(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 December 2007 (21.12.2007)

PCT

(10) International Publication Number
WO 2007/146968 A2

(51) International Patent Classification:
C07K 14/435 (2006.01)

(21) International Application Number:
PCT/US2007/071052

(22) International Filing Date: 12 June 2007 (12.06.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/813,261 12 June 2006 (12.06.2006) US
60/853,287 20 October 2006 (20.10.2006) US

(71) Applicant: **TRUBION PHARMACEUTICALS** [US/US]; 2401 Fourth Avenue, Suite 1050, Seattle, WA 98121 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **THOMPSON, Peter, Armstrong** [US/US]; 14075 30th Place N.E., Bellevue, WA 98007 (US). **LEDBETTER, Jeffrey, A.** [US/US]; 18798 Ridgefield Road NW, Shoreline, WA 98177 (US). **HAYDEN-LEDBETTER, Martha, Susan** [US/US]; 18798 Ridgefield Road NW, Shoreline, WA 98178 (US). **GROSMAIRE, Laura, Sue** [US/US]; 19909 280th Avenue, SE, Hobart, WA 98025 (US). **BADER, Robert** [US/US]; 1411 Boylston Avenue, Apt. #314, Seattle, WA 98122 (US). **BRADY, William** [US/US]; 618 219th Place SW, Bothell, WA 98021 (US). **TCHISTIAKOVA, Lioudmila** [CA/US]; 19 Abbot Bridge Drive, Andover, MA 01810 (US). **FOLLETTIE, Maximillian, T.** [US/US]; 12 Greenbrook Way, Belmont, MA 02478 (US). **CALABRO, Valerie** [FR/US]; 195 Binney Street #1205, Cambridge, MA 02142 (US). **SCHULER, Alwin** [NL/US]; 2 Centennial Street, Hopedale, MA 01747 (US).

(74) Agent: **MERKEL, William, K.**; Marshall, Gerstein & Borun LLP, 233 S. Wacker Drive, Suite 6300, Sears Tower, Chicago, IL 60606-6357 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SINGLE-CHAIN MULTIVALENT BINDING PROTEINS WITH EFFECTOR FUNCTION

(57) Abstract: Multivalent binding peptides, including bi-specific binding peptides, having immunoglobulin effector function are provided, along with encoding nucleic acids, vectors and host cells as well as methods for making such peptides and methods for using such peptides to treat or prevent a variety of diseases, disorders or conditions, as well as to ameliorate at least one symptom associated with such a disease, disorder or condition.

WO 2007/146968 A2

SINGLE-CHAIN MULTIVALENT BINDING PROTEINS WITH EFFECTOR FUNCTION

FIELD OF THE INVENTION

The invention relates generally to the field of multivalent binding molecules and therapeutic applications thereof.

The sequence listing is being submitted as a text file and as a PDF file in compliance with applicable requirements for electronic filing. The sequence listing was created on June 12, 2007. The sequence listing is incorporated herein by reference in its entirety.

BACKGROUND

In a healthy mammal, the immune system protects the body from damage from foreign substances and pathogens. In some instances though, the immune system goes awry, producing traumatic insult and/or disease. For example, B-cells can produce antibodies that recognize self-proteins rather than foreign proteins, leading to the production of the autoantibodies characteristic of autoimmune diseases such as lupus erythematosus, rheumatoid arthritis, and the like. In other instances, the typically beneficial effect of the immune system in combating foreign materials is counterproductive, such as following organ transplantation. The power of the mammalian immune system, and in particular the human immune system, has been recognized and efforts have been made to control the system to avoid or ameliorate the deleterious consequences to health that result either from normal functioning of the immune system in an abnormal environment (e.g., organ transplantation) or from abnormal functioning of the immune system in an otherwise apparently normal environment (e.g., autoimmune disease progression). Additionally, efforts have been made to exploit the immune system to provide a number of target-specific diagnostic and therapeutic methodologies, relying on the capacity of antibodies to specifically recognize and bind antigenic targets with specificity.

One way in which the immune system protects the body is by production of specialized cells called B lymphocytes or B-cells. B-cells produce antibodies that bind to, and in some cases mediate destruction of, a foreign substance or pathogen. In some instances though, the human immune system, and specifically the B lymphocytes of the human immune system, go awry and disease results. There are

numerous cancers that involve uncontrolled proliferation of B-cells. There are also numerous autoimmune diseases that involve B-cell production of antibodies that, instead of binding to foreign substances and pathogens, bind to parts of the body. In addition, there are numerous autoimmune and inflammatory diseases that involve B-cells in their pathology, for example, through inappropriate B-cell antigen presentation to T-cells or through other pathways involving B-cells. For example, autoimmune-prone mice deficient in B-cells do not develop autoimmune kidney disease, vasculitis or autoantibodies. (Shlomchik et al., J Exp. Med. 1994, 180:1295-306). Interestingly, these same autoimmune-prone mice which possess B-cells but are deficient in immunoglobulin production, do develop autoimmune diseases when induced experimentally (Chan et al., J Exp. Med. 1999, 189:1639-48), indicating that B-cells play an integral role in development of autoimmune disease.

B-cells can be identified by molecules on their cell surface. CD20 was the first human B-cell lineage-specific surface molecule identified by a monoclonal antibody. It is a non-glycosylated, hydrophobic 35 kDa B-cell transmembrane phosphoprotein that has both its amino and carboxy ends situated inside the cell. Einfeld et al., EMBO J. 1988, 7:711-17. CD20 is expressed by all normal mature B-cells, but is not expressed by precursor B-cells or plasma cells. Natural ligands for CD20 have not been identified, and the function of CD20 in B-cell biology is still incompletely understood.

Another B-cell lineage-specific cell surface molecule is CD37. CD37 is a heavily glycosylated 40-52 kDa protein that belongs to the tetraspanin transmembrane family of cell surface antigens. It traverses the cell membrane four times forming two extracellular loops and exposing its amino and carboxy ends to the cytoplasm. CD37 is highly expressed on normal antibody-producing (sIg+)B-cells, but is not expressed on pre-B-cells or plasma cells. The expression of CD37 on resting and activated T cells, monocytes and granulocytes is low and there is no detectable CD37 expression on NK cells, platelets or erythrocytes. See, Belov et al., Cancer Res., 61(11):4483-4489 (2001); Schwartz-Albiez et al., J. Immunol., 140(3): 905-914 (1988); and Link et al., J. Immunol., 137(9): 3013-3018 (1988). Besides normal B-cells, almost all malignancies of B-cell origin are positive for CD37 expression, including CLL, NHL, and hairy cell leukemia (Moore, et al. 1987; Merson and Brochier 1988; Faure, et al.

1990). CD37 participates in regulation of B-cell function, since mice lacking CD37 were found to have low levels of serum IgG1 and to be impaired in their humoral response to viral antigens and model antigens. It appears to act as a nonclassical costimulatory molecule or by directly influencing antigen presentation via complex formation with MHC class II molecules. See Knobeloch et al., Mol. Cell. Biol., 20(15):5363-5369 (2000).

Research and drug development has occurred based on the concept that B-cell lineage-specific cell surface molecules such as CD37 and CD20 can themselves be targets for antibodies that would bind to, and mediate destruction of, cancerous and autoimmune disease-causing B-cells that have CD37 and CD20 on their surfaces. Termed "immunotherapy," antibodies made (or based on antibodies made) in a non-human animal that bind to CD37 or CD20 were given to a patient to deplete cancerous or autoimmune disease-causing B-cells.

Monoclonal antibody technology and genetic engineering methods have facilitated development of immunoglobulin molecules for diagnosis and treatment of human diseases. The domain structure of immunoglobulins is amenable to engineering, in that the antigen binding domains and the domains conferring effector functions may be exchanged between immunoglobulin classes and subclasses. Immunoglobulin structure and function are reviewed, for example, in Harlow et al., Eds., Antibodies: A Laboratory Manual, Chapter 14, Cold Spring Harbor Laboratory, Cold Spring Harbor (1988). An extensive introduction as well as detailed information about all aspects of recombinant antibody technology can be found in the textbook "Recombinant Antibodies" (John Wiley & Sons, NY, 1999). A comprehensive collection of detailed antibody engineering lab Protocols can be found in R. Kontermann and S. Dübel (eds.), "The Antibody Engineering Lab Manual" (Springer Verlag, Heidelberg/New York, 2000).

An immunoglobulin molecule (abbreviated Ig), is a multimeric protein, typically composed of two identical light chain polypeptides and two identical heavy chain polypeptides (H_2L_2) that are joined into a macromolecular complex by interchain disulfide bonds, *i.e.*, covalent bonds between the sulfhydryl groups of neighboring cysteine residues. Five human immunoglobulin classes are defined on the basis of their heavy chain composition, and are named IgG, IgM, IgA, IgE, and

IgD. The IgG-class and IgA-class antibodies are further divided into subclasses, namely, IgG1, IgG2, IgG3, and IgG4, and IgA1 and IgA2, respectively. Intrachain disulfide bonds join different areas of the same polypeptide chain, which results in the formation of loops that, along with adjacent amino acids, constitute the immunoglobulin domains. At the amino-terminal portion, each light chain and each heavy chain has a single variable region that shows considerable variation in amino acid composition from one antibody to another. The light chain variable region, V_L, has a single antigen-binding domain and associates with the variable region of a heavy chain, V_H (also containing a single antigen-binding domain), to form the antigen binding site of the immunoglobulin, the Fv.

In addition to variable regions, each of the full-length antibody chains has a constant region containing one or more domains. Light chains have a constant region containing a single domain. Thus, light chains have one variable domain and one constant domain. Heavy chains have a constant region containing several domains. The heavy chains in IgG, IgA, and IgD antibodies have three domains, which are designated C_{H1}, C_{H2}, and C_{H3}; the heavy chains in IgM and IgE antibodies have four domains, C_{H1}, C_{H2}, C_{H3} and C_{H4}. Thus, heavy chains have one variable domain and three or four constant domains. Noteworthy is the invariant organization of these domains in all known species, with the constant regions, containing one or more domains, being located at or near the C-terminus of both the light and heavy chains of immunoglobulin molecules, with the variable domains located towards the N-termini of the light and heavy chains. Immunoglobulin structure and function are reviewed, for example, in Harlow *et al.*, Eds., *Antibodies: A Laboratory Manual*, Chapter 14, Cold Spring Harbor Laboratory, Cold Spring Harbor (1988).

The heavy chains of immunoglobulins can also be divided into three functional regions: the Fd region (a fragment comprising V_H and C_{H1}, *i.e.*, the two N-terminal domains of the heavy chain), the hinge region, and the Fc region (the "fragment crystallizable" region). The Fc region contains the domains that interact with immunoglobulin receptors on cells and with the initial elements of the complement cascade. Thus, the Fc region or fragment is generally considered responsible for the effector functions of an immunoglobulin, such as ADCC (antibody-dependent cell-mediated cytotoxicity), CDC (complement-dependent

cytotoxicity) and complement fixation, binding to Fc receptors, greater half-life *in vivo* relative to a polypeptide lacking an F_C region, protein A binding, and perhaps even placental transfer. Capon et al., Nature, 337: 525-531, (1989). Further, a polypeptide containing an Fc region allows for dimerization/multimerization of the polypeptide. These terms are also used for analogous regions of the other immunoglobulins.

Although all of the human immunoglobulin isotypes contain a recognizable structure in common, each isotype exhibits a distinct pattern of effector function. IgG, by way of nonexhaustive example, neutralizes toxins and viruses, opsonizes, fixes complement (CDC) and participates in ADCC. IgM, in contrast, neutralizes blood-borne pathogens and participates in opsonization. IgA, when associated with its secretory piece, is secreted and provides a primary defense to microbial infection via the mucosa; it also neutralizes toxins and supports opsonization. IgE mediates inflammatory responses, being centrally involved in the recruitment of other cells needed to mount a full response. IgD is known to provide an immunoregulatory function, controlling the activation of B cells. These characterizations of isotype effector functions provide a non-comprehensive illustration of the differences that can be found among human isotypes.

The hinge region, found in IgG, IgA, IgD, and IgE class antibodies, acts as a flexible spacer, allowing the Fab portion to move freely in space. In contrast to the constant regions, the hinge domains are structurally diverse, varying in both sequence and length among immunoglobulin classes and subclasses. For example, the length and flexibility of the hinge region varies among the IgG subclasses. The hinge region of IgG1 encompasses amino acids 216-231 and, because it is freely flexible, the Fab fragments can rotate about their axes of symmetry and move within a sphere centered at the first of two inter-heavy chain disulfide bridges. IgG2 has a shorter hinge than IgG1, with 12 amino acid residues and four disulfide bridges. The hinge region of IgG2 lacks a glycine residue, is relatively short, and contains a rigid poly-proline double helix, stabilized by extra inter-heavy chain disulfide bridges. These properties restrict the flexibility of the IgG2 molecule. IgG3 differs from the other subclasses by its unique extended hinge region (about four times as long as the IgG1 hinge), containing 62 amino acids (including 21 prolines and 11 cysteines), forming an

inflexible poly-proline double helix. In IgG3, the Fab fragments are relatively far away from the Fc fragment, giving the molecule a greater flexibility. The elongated hinge in IgG3 is also responsible for its higher molecular weight compared to the other subclasses. The hinge region of IgG4 is shorter than that of IgG1 and its flexibility is intermediate between that of IgG1 and IgG2. The flexibility of the hinge regions reportedly decreases in the order IgG3>IgG1>IgG4>IgG2. The four IgG subclasses also differ from each other with respect to their effector functions. This difference is related to differences in structure, including differences with respect to the interaction between the variable region, Fab fragments, and the constant Fc fragment.

According to crystallographic studies, the immunoglobulin hinge region can be further subdivided functionally into three regions: the upper hinge region, the core region, and the lower hinge region. Shin *et al.*, 1992 *Immunological Reviews* 130:87. The upper hinge region includes amino acids from the carboxyl end of C_{H1} to the first residue in the hinge that restricts motion, generally the first cysteine residue that forms an interchain disulfide bond between the two heavy chains. The length of the upper hinge region correlates with the segmental flexibility of the antibody. The core hinge region contains the inter-heavy chain disulfide bridges, and the lower hinge region joins the amino terminal end of the C_{H2} domain and includes residues in C_{H2}. *Id.* The core hinge region of human IgG1 contains the sequence Cys-Pro-Pro-Cys which, when dimerized by disulfide bond formation, results in a cyclic octapeptide believed to act as a pivot, thus conferring flexibility. The hinge region may also contain one or more glycosylation sites, which include a number of structurally distinct types of sites for carbohydrate attachment. For example, IgA1 contains five glycosylation sites within a 17-amino-acid segment of the hinge region, conferring resistance of the hinge region polypeptide to intestinal proteases, considered an advantageous property for a secretory immunoglobulin.

Conformational changes permitted by the structure and flexibility of the immunoglobulin hinge region polypeptide sequence may also affect the effector functions of the Fc portion of the antibody. Three general categories of effector functions associated with the Fc region include (1) activation of the classical complement cascade, (2) interaction with effector cells, and (3) compartmentalization

of immunoglobulins. The different human IgG subclasses vary in the relative efficacies with which they fix complement, or activate and amplify the steps of the complement cascade. *See, e.g.*, Kirschfink, 2001 *Immunol. Rev.* 180:177; Chakraborti *et al.*, 2000 *Cell Signal* 12:607; Kohl *et al.*, 1999 *Mol. Immunol.* 36:893; Marsh et al., 1999 *Curr. Opin. Nephrol. Hypertens.* 8:557; Speth *et al.*, 1999 *Wien Klin. Wochenschr.* 111:378.

Exceptions to the H_2L_2 structure of conventional antibodies occur in some isotypes of the immunoglobulins found in camelids (camels, dromedaries and llamas; Hamers-Casterman *et al.*, 1993 *Nature* 363:446; Nguyen *et al.*, 1998 *J. Mol. Biol* 275:413), nurse sharks (Roux *et al.*, 1998 *Proc. Nat. Acad. Sci. USA* 95:11804), and in the spotted ratfish (Nguyen, *et al.*, 2002 *Immunogenetics* 54(1):39-47). These antibodies can apparently form antigen-binding regions using only heavy chain variable region, *i.e.*, these functional antibodies are homodimers of heavy chains only (referred to as "heavy-chain antibodies" or "HCAbs"). Despite the advantages of antibody technology in disease diagnosis and treatment, there are some disadvantageous aspects of developing whole-antibody technologies as diagnostic and/or therapeutic reagents. Whole antibodies are large protein structures exemplified by the heterotetrameric structure of the IgG isotype, containing two light and two heavy chains. Such large molecules are sterically hindered in certain applications. For example, in treatments of solid tumors, whole antibodies do not readily penetrate the interior of the tumor. Moreover, the relatively large size of whole antibodies presents a challenge to ensure that the *in vivo* administration of such molecules does not induce an immune response. Further, generation of active antibody molecules typically involves the culturing of recombinant eukaryotic cells capable of providing appropriate post-translational processing of the nascent antibody molecules, and such cells can be difficult to culture and difficult to induce in a manner that provides commercially useful yields of active antibody.

Recently, smaller immunoglobulin molecules have been constructed to overcome problems associated with whole immunoglobulin methodologies. A single-chain variable antibody fragment (scFv) comprises an antibody heavy chain variable domain joined via a short peptide to an antibody light chain variable domain (Huston et al., Proc. Natl. Acad. Sci. USA, 1988, 85: 5879-83). Because of the small size of

scFv molecules, they exhibit more effective penetration into tissues than whole immunoglobulin. An anti-tumor scFv showed more rapid tumor penetration and more even distribution through the tumor mass than the corresponding chimeric antibody (Yokota et al., Cancer Res. 1992, 52:3402-08).

Despite the advantages that scFv molecules bring to serotherapy, several drawbacks to this therapeutic approach exist. An scFv is rapidly cleared from the circulation, which may reduce toxic effects in normal cells, but such rapid clearance impedes delivery of a minimum effective dose to the target tissue. Manufacturing adequate amounts of scFv for administration to patients has been challenging due to difficulties in expression and isolation of scFv that adversely affect the yield. During expression, scFv molecules lack stability and often aggregate due to pairing of variable regions from different molecules. Furthermore, production levels of scFv molecules in mammalian expression systems are low, limiting the potential for efficient manufacturing of scFv molecules for therapy (Davis et al, J Biol. Chem. 1990, 265:10410-18); Traunecker et al., EMBO J 1991, 10: 3655-59). Strategies for improving production have been explored, including addition of glycosylation sites to the variable regions (Jost, C. R. U.S. Pat. No. 5,888,773, Jost et al, J. Biol. Chem. 1994, 69: 26267-73).

Another disadvantage to using scFv for therapy is the lack of effector function. An scFv without a cytolytic function, such as the antibody-dependent cell-mediated cytotoxicity (ADCC) and complement dependent-cytotoxicity (CDC) associated with the constant region of an immunoglobulin, may be ineffective for treating disease. Even though development of scFv technology began over 12 years ago, currently no scFv products are approved for therapy.

Alternatively, it has been proposed that fusion of an scFv to another molecule, such as a toxin, could take advantage of the specific antigen-binding activity and the small size of an scFv to deliver the toxin to a target tissue. Chaudary et al., Nature 1989, 339:394; Batra et al., Mol. Cell. Biol. 1991, 11:2200. Conjugation or fusion of toxins to scFvs has thus been offered as an alternative strategy to provide potent, antigen-specific molecules, but dosing with such conjugates or chimeras can be limited by excessive and/or non-specific toxicity due to the toxin moiety of such preparations. Toxic effects may include supraphysiological elevation of liver

enzymes and vascular leak syndrome, and other undesired effects. In addition, immunotoxins are themselves highly immunogenic upon administration to a host, and host antibodies generated against the immunotoxin limit potential usefulness for repeated therapeutic treatments of an individual.

Nonsurgical cancer therapy, such as external irradiation and chemotherapy, can suffer from limited efficacy because of toxic effects on normal tissues and cells, due to the lack of specificity these treatments exhibit towards cancer cells. To overcome this limitation, targeted treatment methodologies have been developed to increase the specificity of the treatment for the cells and tissues in need thereof. An example of such a targeted methodology for *in vivo* use is the administration of antibody conjugates, with the antibody designed to specifically recognize a marker associated with a cell or tissue in need of treatment, and the antibody being conjugated to a therapeutic agent, such as a toxin in the case of cancer treatment. Antibodies, as systemic agents, circulate to sensitive and undesirable body compartments, such as the bone marrow. In acute radiation injury, destruction of lymphoid and hematopoietic compartments is a major factor in the development of septicemia and subsequent death. Moreover, antibodies are large, globular proteins that can exhibit poor penetration of tissues in need of treatment.

Human patients and non-human subjects suffering from a variety of end-stage disease processes frequently require organ transplantation. Organ transplantation, however, must contend with the untoward immune response of the recipient and guard against immunological rejection of the transplanted organ by depressing the recipient's cellular immune response to the foreign organ with cytotoxic agents which affect the lymphoid and other parts of the hematopoietic system. Graft acceptance is limited by the tolerance of the recipient to these cytotoxic chemicals, many of which are similar to the anticancer (antiproliferative) agents. Likewise, when using cytotoxic antimicrobial agents, particularly antiviral drugs, or when using cytotoxic drugs for autoimmune disease therapy, e.g., in treatment of systemic lupus erythematosis, a serious limitation is the toxic effects of the therapeutic agents on the bone marrow and the hematopoietic cells of the body.

Use of targeted therapies, such as targeted antibody conjugate therapy, is designed to localize a maximum quantity of the therapeutic agent at the site of desired

action as possible, and the success of such therapies is revealed by the relatively high signal-to-background ratio of therapeutic agent. Examples of targeted antibodies include diagnostic or therapeutic agent conjugates of antibody or antibody fragments, cell-or tissue-specific peptides, and hormones and other receptor-binding molecules. For example, antibodies against different determinants associated with pathological and normal cells, as well as associated with pathogenic microorganisms, have been used for the detection and treatment of a wide variety of pathological conditions or lesions. In these methods, the targeting antibody is directly conjugated to an appropriate detecting or therapeutic agent as described, for example, in Hansen et al., U.S. Pat. No. 3,927,193 and Goldenberg, U.S. Pat. Nos. 4,331,647, 4,348,376, 4,361,544, 4,468,457, 4,444,744, 4,460,459, 4,460,561, 4,624,846 and 4,818,709.

One problem encountered in direct targeting methods, i.e., in methods wherein the diagnostic or therapeutic agent (the "active agent") is conjugated directly to the targeting moiety, is that a relatively small fraction of the conjugate actually binds to the target site, while the majority of conjugate remains in circulation and compromises in one way or another the function of the targeted conjugate. To ensure maximal localization of the active agent, an excess of the targeted conjugate is typically administered, ensuring that some conjugate will remain unbound and contribute to background levels of the active agent. A diagnostic conjugate, e.g., a radioimmunoscintigraphic or magnetic resonance imaging conjugate that does not bind its target can remain in circulation, thereby increasing background and decreasing resolution of the diagnostic technique. In the case of a therapeutic conjugate having a toxin as an active agent (e.g., a radioisotope, drug or toxic compound) attached to a long-circulating targeting moiety such as an antibody, circulating conjugate can result in unacceptable toxicity to the host, such as marrow toxicity or systemic side effects.

U.S. Pat. No. 4,782,840 discloses a method for reducing the effect of elevated background radiation levels during surgery. The method involves injection of a patient with antibodies specific for neoplastic tissue, with the antibodies labeled with radioisotopes having a suitably long half-life, such as Iodine-125. After injection of the radiolabeled antibody, the surgery is delayed at least 7-10 days, preferably 14-21

days, to allow any unbound radiolabeled antibody to be cleared to a low background level.

U.S. Pat. No. 4,932,412 discloses methods for reducing or correcting for non-specific background radiation during intraoperative detection. The methods include the administration to a patient who has received a radiolabeled primary antibody, of a contrast agent, subtraction agent or second antibody which binds the primary antibody.

Apart from producing the antibodies described above, the immune system includes a variety of cell types that have powerful biological effects. During hematopoiesis, bone marrow-derived stem cells differentiate into either mature cells of the immune system ("B" cells) or into precursors of cells that migrate out of the bone marrow to mature in the thymus ("T" cells).

B cells are central to the humoral component of an immune response. B cells are activated by an appropriate presentation of an antigen to become antibody-secreting plasma cells; antigen presentation also results in clonal expansion of the activated B cell. B cells are primarily responsible for the humoral component of an immune response. A plasma cell typically exhibits about 10^5 antibody molecules (IgD and IgM) on its surface.

T lymphocytes can be divided into two categories. The cytotoxic T cells, Tc lymphocytes or CTLs (CD8+ T cells), kill cells bearing foreign surface antigen in association with Class I MHC and can kill cells that are harboring intracellular parasites (either bacteria or viruses) as long as the infected cell is displaying a microbial antigen on its surface. Tc cells kill tumor cells and account for the rejection of transplanted cells. Tc cells recognize antigen-Class I MHC complexes on target cells, contact them, and release the contents of granules directly into the target cell membrane, which lyses the cell.

A second category of T cells is the helper T cell or Th lymphocyte (CD4+ T cells), which produces lymphokines that are "helper" factors in the maturation of B cells into antibody-secreting plasma cells. Th cells also produce certain lymphokines that stimulate the differentiation of effector T lymphocytes and the activity of macrophages. Th1 cells recognize antigen on macrophages in association with Class

II MHC and become activated (by IL-1) to produce lymphokines, including the IFN-γ that activates macrophages and NK cells. These cells mediate various aspects of the cell-mediated immunity response including delayed-type hypersensitivity reactions. Th2 cells recognize antigen in association with Class II MHC on an antigen presenting cell or APC (e.g., migratory macrophages and dendritic cells) and then produce interleukins and other substances that stimulate specific B-cell and T-cell proliferation and activity.

Beyond serving as APCs that initiate T cell interactions, development, and proliferation, macrophages are involved in expression of cell-mediated immunity because they become activated by IFN-γ produced in a cell-mediated immune response. Activated macrophages have increased phagocytic potential and release soluble substances that cause inflammation and destroy many bacteria and other cells. Natural Killer cells are cytotoxic cells that lyse cells bearing new antigen, regardless of their MHC type, and even lyse some cells that bear no MHC proteins. Natural Killer T cells, or NK cells, are defined by their ability to kill cells displaying a foreign antigen (*e.g.*, tumor cells), regardless of MHC type, and regardless of previous sensitization (exposure) to the antigen. NK cells can be activated by IL-2 and IFN-γ, and lyse cells in the same manner as cytotoxic T lymphocytes. Some NK cells have receptors for the Fc domain of the IgG antibody (e.g, CD16 or $F_C\gamma RIII$) and are thus able to bind to the Fc portion of IgG on the surface of a target cell and release cytolytic components that kill the target cell via antibody-dependent cell-mediated cytotoxicity.

Another group of cells is the granulocytes or polymorphonuclear leukocytes (PMNs). Neutrophils, one type of PMN, kill bacterial invaders and phagocytose the remains. Eosinophils are another type of PMN and contain granules that prove cytotoxic when released upon another cell, such as a foreign cell. Basophils, a third type of PMN, are significant mediators of powerful physiological responses (e.g., inflammation) that exert their effects by releasing a variety of biologically active compounds, such as histamine, serotonin, prostaglandins, and leukotrienes. Common to all of these cell types is the capacity to exert a physiological effect within an organism, frequently by killing, and optionally scavenging, deleterious compositions such as foreign cells.

Although a variety of mammalian cells, including cells of the immune system, are capable of directly exerting a physiological effect (e.g., cell killing, typified by Tc, NK, some PMN, macrophage, and the like), other cells indirectly contribute to a physiological effect. For example, initial presentation of an antigen to a naïve T cell of the immune system requires MHC presentation that mandates cell-cell contact. Further, there often needs to be contact between an activated T cell and an antigen-specific B cell to obtain a particular immunogenic response. A third form of cell-cell contact often seen in immune responses is the contact between an activated B cell and follicular dendritic cells. Each of these cell-cell contact requirements complicates the targeting of a biologically active agent to a given target.

Complement-dependent cytotoxicity (CDC) is believed to be a significant mechanism for clearance of specific target cells such as tumor cells. CDC is a series of events that consists of a collection of enzymes that become activated by each other in a cascade fashion. Complement has an important role in clearing antigen, accomplished by its four major functions: (1) local vasodilation; (2) attraction of immune cells, especially phagocytes (chemotaxis); (3) tagging of foreign organisms for phagocytosis (opsonization); and (4) destruction of invading organisms by the membrane attack complex (MAC attack). The central molecule is the C3 protein. It is an enzyme that is split into two fragments by components of either the classical pathway or the alternative pathway. The classical pathway is induced by antibodies, especially IgG and IgM, while the alternative pathway is nonspecifically stimulated by bacterial products like lipopolysaccharide (LPS). Briefly, the products of the C3 split include a small peptide C3a which is chemotactic for phagocytic immune cells and results in local vasodilation by causing the release of C5a fragment from C5. The other part of C3, C3b, coats antigens on the surface of foreign organisms and acts to opsonize the organism for destruction. C3b also reacts with other components of the complement system to form an MAC consisting of C5b, C6, C7, C8 and C9.

There are problems associated with the use of antibodies in human therapy because the response of the immune system to any antigen, even the simplest, is "polyclonal," *i.e.*, the system manufactures antibodies of a great range of structures both in their binding regions as well as in their effector regions.

Two approaches have been used in an attempt to reduce the problem of immunogenic antibodies. The first is the production of chimeric antibodies in which the antigen-binding part (variable regions) of a mouse monoclonal antibody is fused to the effector part (constant region) of a human antibody. In a second approach, antibodies have been altered through a technique known as complementarity determining region (CDR) grafting or "humanization." This process has been further improved to include changes referred to as "reshaping" (Verhoeyen, *et al.*, 1988 *Science* 239:1534-1536; Riechmann, *et al.*, 1988 *Nature* 332:323-337; Tempest, *et al.*, *Bio/Technol* 1991 9:266-271), "hyperchimerization" (Queen, *et al.*, 1989 *Proc Natl Acad Sci USA* 86:10029-10033; Co, *et al.*, 1991 *Proc Natl Acad Sci USA* 88:2869 – 2873; Co, *et al.*, 1992 *J Immunol* 148:1149-1154), and "veneering" (Mark, *et al.*, *In*: Metcalf BW, Dalton BJ, eds. Cellular adhesion: molecular definition to therapeutic potential. New York: Plenum Press, 1994:291-312).

An average of less than one therapeutic antibody per year has been introduced to the market beginning in 1986, eleven years after the publication of monoclonal antibodies. Five murine monoclonal antibodies were introduced into human medicine over a ten year period from 1986-1995, including "muromonab-CD3" (OrthoClone OKT3®) for acute rejection of organ transplants; "edrecolomab" (Panorex®) for colorectal cancer; "odulimomab" (Antilfa®) for transplant rejection; and, "ibritumomab" (Zevalin® yiuxetan) for non-Hodgkin's lymphoma. Additionally, a monoclonal Fab, "abciximab" (ReoPro®) has been marketed for preventing coronary artery reocclusion. Three chimeric monoclonal antibodies were also launched: "rituximab" (Rituxan®) for treating B cell lymphomas; "basiliximab" (Simulect®)for transplant rejection; and "infliximab" (Remicade®) for treatment of rheumatoid arthritis and Crohn's disease. Additionally, "abciximab" (ReoPro®), a 47.6 kD Fab fragment of a chimeric human-murine monoclonal antibody is marketed as an adjunct to percutaneous coronary intervention for the prevention of cardiac ischemic complications in patients undergoing percutaneous coronary intervention. Finally, seven "humanized" monoclonal antibodies have been launched. "Daclizumab" (Zenapax®) is used to prevent acute rejection of transplanted kidneys; "palivizumab" (Synagis®) for RSV; "trastuzumab" (Herceptin®) binds HER-2, a growth factor receptor found on breast cancers cells; "gemtuzumab" (Mylotarg®) for acute myelogenous leukemia (AML); and "alemtuzumab" (MabCampath®) for chronic

lymphocytic leukemia; "adalimumab" (Humira® (D2E7)) for the treatment of rheumatoid arthritis; and, "omalizumab" (Xolair®), for the treatment of persistent asthma.

Thus, a variety of antibody technologies have received attention in the effort to develop and market more effective therapeutics and palliatives. Unfortunately, problems continue to compromise the promise of each of these therapies. For example, the majority of cancer patients treated with rituximab relapse, generally within about 6-12 months, and fatal infusion reactions within 24 hours of rituximab infusion have been reported. Acute renal failure requiring dialysis with instances of fatal outcome has also been reported in treatments with rituximab, as have severe, occasionally fatal, mucocutaneous reactions. Additionally, high doses of rituximab are required for intravenous injection because the molecule is large, approximately 150 kDa, and diffusion into the lymphoid tissues, where many tumor cells may reside is limited.

Trastuzumab administration can result in the development of ventricular dysfunction, congestive heart failure, and severe hypersensitivity reactions (including anaphylaxis), infusion reactions, and pulmonary events. Daclizumab immunosuppressive therapy poses an increased risk for developing lymphoproliferative disorders and opportunistic infections. Death from liver failure, arising from severe hepatotoxicity, and from veno-occlusive disease (VOD), has been reported in patients who received gemtuzumab.

Hepatotoxicity was also reported in patients receiving alemtuzumab. Serious and, in some rare instances fatal, pancytopenia/marrow hypoplasia, autoimmune idiopathic thrombocytopenia, and autoimmune hemolytic anemia have occurred in patients receiving alemtuzumab therapy. Alemtuzumab can also result in serious infusion reactions as well as opportunistic infections. In patients treated with adalimumab, serious infections and sepsis, including fatalities, have been reported, as has the exacerbation of clinical symptoms and/or radiographic evidence of demyelinating disease, and patients treated with adalimumab in clinical trials had a higher incidence of lymphoma than the expected rate in the general population. Omalizumab reportedly induces malignancies and anaphylaxis.

Cancer includes a broad range of diseases, affecting approximately one in four individuals worldwide. Rapid and unregulated proliferation of malignant cells is a hallmark of many types of cancer, including hematological malignancies. Although patients with a hematologic malignant condition have benefited from advances in cancer therapy in the past two decades, Multani *et al.*, 1998 *J. Clin. Oncology* 16:3691-3710, and remission times have increased, most patients still relapse and succumb to their disease. Barriers to cure with cytotoxic drugs include, for example, tumor cell resistance and the high toxicity of chemotherapy, which prevents optimal dosing in many patients.

Treatment of patients with low grade or follicular B cell lymphoma using a chimeric CD20 monoclonal antibody has been reported to induce partial or complete responses in patients. McLaughlin *et al.*, 1996 *Blood* 88:90a (abstract, suppl. 1); Maloney *et al.*, 1997 *Blood* 90:2188-95. However, as noted above, tumor relapse commonly occurs within six months to one year. Further improvements in serotherapy are needed to induce more durable responses, for example, in low grade B cell lymphoma, and to allow effective treatment of high grade lymphoma and other B cell diseases.

Another approach has been to target radioisotopes to B cell lymphomas using monoclonal antibodies specific for CD20. While the effectiveness of therapy is reportedly increased, associated toxicity from the long *in vivo* half-life of the radioactive antibody increases, sometimes requiring that the patient undergo stem cell rescue. Press *et al.*, 1993 *N. Eng. J. Med.* 329:1219-1224; Kaminski *et al.*, 1993 *N. Eng.J. Med.* 329:459-65. Monoclonal antibodies to CD20 have also been cleaved with proteases to yield $F(ab')_2$ or Fab fragments prior to attachment of radioisotope. This has been reported to improve penetration of the radioisotope conjugate into the tumor and to shorten the *in vivo* half-life, thus reducing the toxicity to normal tissues. However, these molecules lack effector functions, including complement fixation and/or ADCC.

Autoimmune diseases include autoimmune thyroid diseases, which include Graves' disease and Hashimoto's thyroiditis. In the United States alone, there are about 20 million people who have some form of autoimmune thyroid disease. Autoimmune thyroid disease results from the production of autoantibodies that either

stimulate the thyroid to cause hyperthyroidism (Graves' disease) or destroy the thyroid to cause hypothyroidism (Hashimoto's thyroiditis). Stimulation of the thyroid is caused by autoantibodies that bind and activate the thyroid stimulating hormone (TSH) receptor. Destruction of the thyroid is caused by autoantibodies that react with other thyroid antigens. Current therapy for Graves' disease includes surgery, radioactive iodine, or antithyroid drug therapy. Radioactive iodine is widely used, since antithyroid medications have significant side effects and disease recurrence is high. Surgery is reserved for patients with large goiters or where there is a need for very rapid normalization of thyroid function. There are no therapies that target the production of autoantibodies responsible for stimulating the TSH receptor. Current therapy for Hashimoto's thyroiditis is levothyroxine sodium, and lifetime therapy is expected because of the low likelihood of remission. Suppressive therapy has been shown to shrink goiters in Hashimoto's thyroiditis, but no therapies that reduce autoantibody production to target the disease mechanism are known.

Rheumatoid arthritis (RA) is a chronic disease characterized by inflammation of the joints, leading to swelling, pain, and loss of function. RA affects an estimated 2.5 million people in the United States. RA is caused by a combination of events including an initial infection or injury, an abnormal immune response, and genetic factors. While autoreactive T cells and B cells are present in RA, the detection of high levels of antibodies that collect in the joints, called rheumatoid factor, is used in the diagnosis of RA. Current therapy for RA includes many medications for managing pain and slowing the progression of the disease. No therapy has been found that can cure the disease. Medications include nonsteroidal anti-inflammatory drugs (NSAIDS), and disease modifying anti-rheumatic drugs (DMARDS). NSAIDS are useful in benign disease, but fail to prevent the progression to joint destruction and debility in severe RA. Both NSAIDS and DMARDS are associated with significant side effects. Only one new DMARD, Leflunomide, has been approved in over 10 years. Leflunomide blocks production of autoantibodies, reduces inflammation, and slows progression of RA. However, this drug also causes severe side effects including nausea, diarrhea, hair loss, rash, and liver injury.

Systemic Lupus Erythematosus (SLE) is an autoimmune disease caused by recurrent injuries to blood vessels in multiple organs, including the kidney, skin, and

joints. SLE is estimated to affect over 500,000 people in the United States. In patients with SLE, a faulty interaction between T cells and B cells results in the production of autoantibodies that attack the cell nucleus. These include anti-double stranded DNA and anti-Sm antibodies. Autoantibodies that bind phospholipids are also found in about half of SLE patients, and are responsible for blood vessel damage and low blood counts. Immune complexes accumulate in the kidneys, blood vessels, and joints of SLE patients, where they cause inflammation and tissue damage. No treatment for SLE has been found to cure the disease. NSAIDS and DMARDS are used for therapy depending upon the severity of the disease. Plasmapheresis with plasma exchange to remove autoantibodies can cause temporary improvement in SLE patients. There is general agreement that autoantibodies are responsible for SLE, so new therapies that deplete the B cell lineage, allowing the immune system to reset as new B cells are generated from precursors, would offer hope for long lasting benefit in SLE patients.

Sjogren's syndrome is an autoimmune disease characterized by destruction of the body's moisture-producing glands. Sjogren's syndrome is one of the most prevalent autoimmune disorders, striking up to an estimated 4 million people in the United States. About half of the people stricken with Sjogren's syndrome also have a connective tissue disease, such as RA, while the other half have primary Sjogren's syndrome with no other concurrent autoimmune disease. Autoantibodies, including anti-nuclear antibodies, rheumatoid factor, anti-fodrin, and anti-muscarinic receptor are often present in patients with Sjogren's syndrome. Conventional therapy includes corticosteroids, and additional more effective therapies would be of benefit.

Immune thrombocytopenic purpura (ITP) is caused by autoantibodies that bind to blood platelets and cause their destruction. Some cases of ITP are caused by drugs, and others are associated with infection, pregnancy, or autoimmune disease such as SLE. About half of all cases are classified as being of idiopathic origin. The treatment of ITP is determined by the severity of the symptoms. In some cases, no therapy is needed although in most cases immunosuppressive drugs, including corticosteroids or intravenous infusions of immune globulin to deplete T cells, are provided. Another treatment that usually results in an increased number of platelets is removal of the spleen, the organ that destroys antibody-coated platelets. More potent

immunosuppressive drugs, including cyclosporine, cyclophosphamide, or azathioprine are used for patients with severe cases. Removal of autoantibodies by passage of patients' plasma over a Protein A column is used as a second line treatment in patients with severe disease. Additional more effective therapies are needed.

Multiple sclerosis (MS) is also an autoimmune disease. It is characterized by inflammation of the central nervous system and destruction of myelin, which insulates nerve cell fibers in the brain, spinal cord, and body. Although the cause of MS is unknown, it is widely believed that autoimmune T cells are primary contributors to the pathogenesis of the disease. However, high levels of antibodies are present in the cerebrospinal fluid of patients with MS, and some predict that the B cell response leading to antibody production is important for mediating the disease. No B cell depletion therapies have been studied in patients with MS, and there is no cure for MS. Current therapy is corticosteroids, which can reduce the duration and severity of attacks, but do not affect the course of MS over time. New biotechnology interferon (IFN) therapies for MS have recently been approved but additional more effective therapies are required.

Myasthenia Gravis (MG) is a chronic autoimmune neuromuscular disorder that is characterized by weakness of the voluntary muscle groups. MG affects about 40,000 people in the United States. MG is caused by autoantibodies that bind to acetylcholine receptors expressed at neuromuscular junctions. The autoantibodies reduce or block acetylcholine receptors, preventing the transmission of signals from nerves to muscles. There is no known cure for mg. Common treatments include immunosuppression with corticosteroids, cyclosporine, cyclophosphamide, or azathioprine. Surgical removal of the thymus is often used to blunt the autoimmune response. Plasmapheresis, used to reduce autoantibody levels in the blood, is effective in mg, but is short-lived because the production of autoantibodies continues. Plasmapheresis is usually reserved for severe muscle weakness prior to surgery. New and effective therapies would be of benefit.

Psoriasis affects approximately five million people, and is characterized by autoimmune inflammation in the skin. Psoriasis is also associated with arthritis in 30% (psoriatic arthritis). Many treatments, including steroids, uv light retinoids,

vitamin D derivatives, cyclosporine, and methotrexate have been used but it is also clear that psoriasis would benefit from new and effective therapies. Scleroderma is a chronic autoimmune disease of the connective tissue that is also known as systemic sclerosis. Scleroderma is characterized by an overproduction of collagen, resulting in a thickening of the skin, and approximately 300,000 people in the United States have scleroderma, which would also benefit from new and effective therapies.

Apparent from the foregoing discussion are needs for improved compositions and methods to treat, ameliorate or prevent a variety of diseases, disorders and conditions, including cancer and autoimmune diseases.

SUMMARY

The invention satisfies at least one of the aforementioned needs in the art by providing proteins containing at least two specific binding domains, wherein those two domains are linked by a constant sub-region derived from an antibody molecule attached at its C-terminus to a linker herein referred to as a scorpion linker, and nucleic acids encoding such proteins, as well as production, diagnostic and therapeutic uses of such proteins and nucleic acids. The constant sub-region comprises a domain derived from an immunoglobulin C_{H2} domain, and preferably a domain derived from an immunoglobulin C_{H3} domain, but does not contain a domain or region derived from, or corresponding to, an immunoglobulin C_{H1} domain. Previously, it had been thought that the placement of a constant region derived from an antibody in the interior of a protein would interfere with antibody function, such as effector function, by analogy to the conventional placement of constant regions of antibodies at the carboxy termini of antibody chains. In addition, placement of a scorpion linker, which may be an immunoglobulin hinge-like peptide, C-terminal to a constant sub-region is an organization that differs from the organization of naturally occurring immunoglobulins. Placement of a constant sub-region (with a scorpion linker attached C-terminal to the constant region) in the interior of a polypeptide or protein chain in accordance with the invention, however, resulted in proteins exhibiting effector function and multivalent (mono- or multi-specific) binding capacities relatively unencumbered by steric hindrances. As will be apparent to one of skill in the art upon consideration of this disclosure, such proteins are modular in design and may be constructed by selecting any of a variety of binding domains for

binding domain 1 or binding domain 2 (or for any additional binding domains found in a particular protein according to the invention), by selecting a constant sub-region having effector function, and by selecting a scorpion linker, hinge-like or non-hinge like (e.g., type II C-lectin receptor stalk region peptides), with the protein exhibiting a general organization of N-binding domain 1-constant sub-region-scorpion linker-binding domain 2-C. Those of skill will further appreciate that proteins of such structure, and the nucleic acids encoding those proteins, will find a wide variety of applications, including medical and veterinary applications.

One aspect of the invention is drawn to a multivalent single-chain binding protein with effector function, or scorpion (the terms are used interchangeably), comprising a first binding domain derived from an immunoglobulin (e.g., an antibody) or an immunoglobulin-like molecule, a constant sub-region providing an effector function, the constant sub-region located C-terminal to the first binding domain; a scorpion linker located C-terminal to the constant sub-region; and a second binding domain derived from an immunoglobulin (such as an antibody) or immunoglobulin-like molecule, located C-terminal to the constant sub-region; thereby localizing the constant sub-region between the first binding domain and the second binding domain. The single-chain binding protein may be multispecific, e.g., bispecific in that it could bind two or more distinct targets, or it may be monospecific, with two binding sites for the same target. Moreover, all of the domains of the protein are found in a single chain, but the protein may form homo-multimers, e.g., by interchain disulfide bond formation. In some embodiments, the first binding domain and/or the second binding domain is/are derived from variable regions of light and heavy immunoglobulin chains from the same, or different, immunoglobulins (e.g., antibodies). The immunoglobulin(s) may be from any vertebrate, such as a mammal, including a human, and may be chimeric, humanized, fragments, variants or derivatives of naturally occurring immunoglobulins.

The invention contemplates proteins in which the first and second binding domains are derived from the same, or different immunoglobulins (e.g., antibodies), and wherein the first and second binding domains recognize the same, or different, molecular targets (e.g., cell surface markers, such as membrane-bound proteins). Further, the first and second binding domains may recognize the same, or different,

epitopes. The first and second molecular targets may be associated with first and second target cells, viruses, carriers and/or objects. In preferred embodiments according to this aspect of the invention, each of the first binding domain, second binding domain, and constant sub-region is derived from a human immunoglobulin, such as an IgG antibody. In yet other embodiments, the multivalent binding protein with effector function has at least one of the first binding domain and the second binding domain that recognizes at least one cell-free molecular target, e.g., a protein unassociated with a cell, such as a deposited protein or a soluble protein. Cell-free molecular targets include, e.g., proteins that were never associated with a cell, e.g., administered compounds such as proteins, as well as proteins that are secreted, cleaved, present in exosomes, or otherwise discharged or separated from a cell.

The target molecules recognized by the first and second binding domains may be found on, or in association with, the same, or different, prokaryotic cells, eukaryotic cells, viruses (including bacteriophage), organic or inorganic target molecule carriers, and foreign objects. Moreover, those target molecules may be on physically distinct cells, viruses, carriers or objects of the same type (e.g., two distinct eukaryotic cells, prokaryotic cells, viruses or carriers) or those target molecules may be on cells, viruses, carriers, or objects that differ in type (e.g., a eukaryotic cell and a virus). Target cells are those cells associated with a target molecule recognized by a binding domain and includes endogenous or autologous cells as well as exogenous or foreign cells (e.g., infectious microbial cells, transplanted mammalian cells including transfused blood cells). The invention comprehends targets for the first and/or second binding domains that are found on the surface of a target cell(s) associated with a disease, disorder or condition of a mammal such as a human. Exemplary target cells include a cancer cell, a cell associated with an autoimmune disease or disorder, and an infectious cell (e.g., an infectious bacterium). A cell of an infectious organism, such as a mammalian parasite, is also contemplated as a target cell. In some embodiments, a protein of the invention is a multivalent (e.g., multispecific) binding protein with effector function wherein at least one of the first binding domain and the second binding domain recognizes a target selected from the group consisting of a tumor antigen, a B-cell target, a TNF receptor superfamily member, a Hedgehog family member, a receptor tyrosine kinase, a proteoglycan-related molecule, a TGF-beta superfamily member, a Wnt-related molecule, a receptor ligand, a T-cell target, a

Dendritic cell target, an NK cell target, a monocyte/macrophage cell target and an angiogenesis target.

In some embodiments of the above-described protein, the tumor antigen is selected from the group consisting of SQUAMOUS CELL CARCINOMA ANTIGEN 1 (SCCA-1), (PROTEIN T4-A), SQUAMOUS CELL CARCINOMA ANTIGEN 2 (SCCA-2), Ovarian carcinoma antigen CA125 (1A1-3B) (KIAA0049), MUCIN 1 (TUMOR-ASSOCIATED MUCIN), (CARCINOMA-ASSOCIATED MUCIN), (POLYMORPHIC EPITHELIAL MUCIN),(PEM),(PEMT),(EPISIALIN), (TUMOR-ASSOCIATED EPITHELIAL MEMBRANE ANTIGEN),(EMA),(H23AG), (PEANUT-REACTIVE URINARY MUCIN), (PUM), (BREAST CARCINOMA- ASSOCIATED ANTIGEN DF3), CTCL tumor antigen se1-1, CTCL tumor antigen se14-3, CTCL tumor antigen se20-4, CTCL tumor antigen se20-9, CTCL tumor antigen se33-1, CTCL tumor antigen se37-2, CTCL tumor antigen se57-1, CTCL tumor antigen se89-1, Prostate-specific membrane antigen, 5T4 oncofetal trophoblast glycoprotein, Orf73 Kaposi's sarcoma-associated herpesvirus, MAGE-C1 (cancer/testis antigen CT7), MAGE-B1 ANTIGEN (MAGE-XP ANTIGEN) (DAM10), MAGE-B2 ANTIGEN (DAM6), MAGE-2 ANTIGEN, MAGE-4a antigen, MAGE-4b antigen, Colon cancer antigen NY-CO-45, Lung cancer antigen NY-LU-12 variant A, Cancer associated surface antigen, Adenocarcinoma antigen ART1, Paraneoplastic associated brain-testis-cancer antigen (onconeuronal antigen MA2; paraneoplastic neuronal antigen), Neuro-oncological ventral antigen 2 (NOVA2), Hepatocellular carcinoma antigen gene 520, TUMOR-ASSOCIATED ANTIGEN CO-029, Tumor-associated antigen MAGE-X2, Synovial sarcoma, X breakpoint 2, Squamous cell carcinoma antigen recognized by T cell, Serologically defined colon cancer antigen 1, Serologically defined breast cancer antigen NY-BR-15, Serologically defined breast cancer antigen NY-BR-16, Chromogranin A; parathyroid secretory protein 1, DUPAN-2, CA 19-9, CA 72-4, CA 195 and L6.

Embodiments of the above-described method comprise a B cell target selected from the group consisting of CD10, CD19, CD20, CD21, CD22, CD23, CD24, CD37, CD38, CD39, CD40, CD72, CD73, CD74, CDw75, CDw76, CD77, CD78, CD79a/b, CD80, CD81, CD82, CD83, CD84, CD85, CD86, CD89, CD98, CD126, CD127, CDw130, CD138 and CDw150.

In other embodiments of the above-described method, the TNF receptor superfamily member is selected from the group consisting of 4-1BB/TNFRSF9, NGF R/TNFRSF16, BAFF R/TNFRSF13C, Osteoprotegerin/TNFRSF11B, BCMA/TNFRSF17, OX40/TNFRSF4, CD27/TNFRSF7, RANK/TNFRSF11A, CD30/TNFRSF8, RELT/TNFRSF19L, CD40/TNFRSF5, TACI/TNFRSF13B, DcR3/TNFRSF6B, TNF RI/TNFRSF1A, DcTRAIL R1/TNFRSF23, TNF RII/TNFRSF1B, DcTRAIL R2/TNFRSF22, TRAIL R1/TNFRSF10A, DR3/TNFRSF25, TRAIL R2/TNFRSF10B, DR6/TNFRSF21, TRAIL R3/TNFRSF10C, EDAR, TRAIL R4/TNFRSF10D, Fas/TNFRSF6, TROY/TNFRSF19, GITR/TNFRSF18, TWEAK R/TNFRSF12, HVEM/TNFRSF14, XEDAR, Lymphotoxin beta R/TNFRSF3, 4-1BB Ligand/TNFSF9, Lymphotoxin, APRIL/TNFSF13, Lymphotoxin beta/TNFSF3, BAFF/TNFSF13C, OX40 Ligand/TNFSF4, CD27 Ligand/TNFSF7, TL1A/TNFSF15, CD30 Ligand/TNFSF8, TNF-alpha/TNFSF1A, CD40 Ligand/TNFSF5, TNF-beta/TNFSF1B, EDA-A2, TRAIL/TNFSF10, Fas Ligand/TNFSF6, TRANCE/TNFSF11, GITR Ligand/TNFSF18, TWEAK/TNFSF12 and LIGHT/TNFSF14.

The above-described method also includes embodiments in which the Hedgehog family member is selected from the group consisting of Patched and Smoothened. In yet other embodiments, the proteoglycan-related molecule is selected from the group consisting of proteoglycans and regulators thereof.

Additional embodiments of the method are drawn to processes in which the receptor tyrosine kinase is selected from the group consisting of Axl, FGF R4, C1q R1/CD93, FGF R5, DDR1, Flt-3, DDR2, HGF R, Dtk, IGF-I R, EGF R, IGF-II R, Eph, INSRR, EphA1, Insulin R/CD220, EphA2, M-CSF R, EphA3, Mer, EphA4, MSP R/Ron, EphA5, MuSK, EphA6, PDGF R alpha, EphA7, PDGF R beta, EphA8, Ret, EphB1, ROR1, EphB2, ROR2, EphB3, SCF R/c-kit, EphB4, Tie-1, EphB6, Tie-2, ErbB2, TrkA, ErbB3, TrkB, ErbB4, TrkC, FGF R1, VEGF R1/Flt-1, FGF R2, VEGF R2/Flk-1, FGF R3 and VEGF R3/Flt-4.

In other embodiments of the method, the Transforming Growth Factor (TGF)-beta superfamily member is selected from the group consisting of Activin RIA/ALK-2, GFR alpha-1, Activin RIB/ALK-4, GFR alpha-2, Activin RIIA, GFR alpha-3, Activin RIIB, GFR alpha-4, ALK-1, MIS RII, ALK-7, Ret, BMPR-IA/ALK-3, TGF-

beta RI/ALK-5, BMPR-IB/ALK-6, TGF-beta RII, BMPR-II, TGF-beta RIIb, Endoglin/CD105 and TGF-beta RIII.

Yet other embodiments of the method comprise a Wnt-related molecule selected from the group consisting of Frizzled-1, Frizzled-8, Frizzled-2, Frizzled-9, Frizzled-3, sFRP-1, Frizzled-4, sFRP-2, Frizzled-5, sFRP-3, Frizzled-6, sFRP-4, Frizzled-7, MFRP, LRP 5, LRP 6, Wnt-1, Wnt-8a, Wnt-3a, Wnt-10b, Wnt-4, Wnt-11, Wnt-5a, Wnt-9a and Wnt-7a.

In other embodiments of the method, the receptor ligand is selected from the group consisting of 4-1BB Ligand/TNFSF9, Lymphotoxin, APRIL/TNFSF13, Lymphotoxin beta/TNFSF3, BAFF/TNFSF13C, OX40 Ligand/TNFSF4, CD27 Ligand/TNFSF7, TL1A/TNFSF15, CD30 Ligand/TNFSF8, TNF-alpha/TNFSF1A, CD40 Ligand/TNFSF5, TNF-beta/TNFSF1B, EDA-A2, TRAIL/TNFSF10, Fas Ligand/TNFSF6, TRANCE/TNFSF11, GITR Ligand/TNFSF18, TWEAK/TNFSF12, LIGHT/TNFSF14, Amphiregulin, NRG1 isoform GGF2, Betacellulin, NRG1 Isoform SMDF, EGF, NRG1-alpha/HRG1-alpha, Epigen, NRG1-beta 1/HRG1-beta 1, Epiregulin, TGF-alpha, HB-EGF, TMEFF1/Tomoregulin-1, Neuregulin-3, TMEFF2, IGF-I, IGF-II, Insulin, Activin A, Activin B, Activin AB, Activin C, BMP-2, BMP-7, BMP-3, BMP-8, BMP-3b/GDF-10, BMP-9, BMP-4, BMP-15, BMP-5, Decapentaplegic, BMP-6, GDF-1, GDF-8, GDF-3, GDF-9, GDF-5, GDF-11, GDF-6, GDF-15, GDF-7, Artemin, Neurturin, GDNF, Persephin, TGF-beta, TGF-beta 2, TGF-beta 1, TGF-beta 3, LAP (TGF-beta 1), TGF-beta 5, Latent TGF-beta 1, Latent TGF-beta bp1, TGF-beta 1.2, Lefty, Nodal, MIS/AMH, FGF acidic, FGF-12, FGF basic, FGF-13, FGF-3, FGF-16, FGF-4, FGF-17, FGF-5, FGF-19, FGF-6, FGF-20, FGF-8, FGF-21, FGF-9, FGF-23, FGF-10, KGF/FGF-7, FGF-11, Neuropilin-1, PlGF, Neuropilin-2, PlGF-2, PDGF, PDGF-A, VEGF, PDGF-B, VEGF-B, PDGF-C, VEGF-C, PDGF-D, VEGF-D and PDGF-AB.

In still other embodiments, the T-cell target is selected from the group consisting of 2B4/SLAMF4, IL-2 R alpha, 4-1BB/TNFRSF9, IL-2 R beta, ALCAM, B7-1/CD80, IL-4 R, B7-H3, BLAME/SLAMF8, BTLA, IL-6 R, CCR3, IL-7 R alpha, CCR4, CXCR1/IL-8 RA, CCR5, CCR6, IL-10 R alpha, CCR7, IL-10 R beta, CCR8, IL-12 R beta 1, CCR9, IL-12 R beta 2, CD2, IL-13 R alpha 1, IL-13, CD3, CD4, ILT2/CD85j, ILT3/CD85k, ILT4/CD85d, ILT5/CD85a, Integrin alpha 4/CD49d,

CD5, Integrin alpha E/CD103, CD6, Integrin alpha M/CD11b, CD8, Integrin alpha X/CD11c, Integrin beta 2/CD18, KIR/CD158, CD27/TNFRSF7, KIR2DL1, CD28, KIR2DL3, CD30/TNFRSF8, KIR2DL4/CD158d, CD31/PECAM-1, KIR2DS4, CD40 Ligand/TNFSF5, LAG-3, CD43, LAIR1, CD45, LAIR2, CD83, Leukotriene B4 R1, CD84/SLAMF5, NCAM-L1, CD94, NKG2A, CD97, NKG2C, CD229/SLAMF3, NKG2D, CD2F-10/SLAMF9, NT-4, CD69, NTB-A/SLAMF6, Common gamma Chain/IL-2 R gamma, Osteopontin, CRACC/SLAMF7, PD-1, CRTAM, PSGL-1, CTLA-4, RANK/TNFRSF11A, CX3CR1, CX3CL1, L-Selectin, CXCR3, SIRP beta 1, CXCR4, SLAM, CXCR6, TCCR/WSX-1, DNAM-1, Thymopoietin, EMMPRIN/CD147, TIM-1, EphB6, TIM-2, Fas/TNFRSF6, TIM-3, Fas Ligand/TNFSF6, TIM-4, Fc gamma RIII/CD16, TIM-6, GITR/TNFRSF18, TNF RI/TNFRSF1A, Granulysin, TNF RII/TNFRSF1B, HVEM/TNFRSF14, TRAIL R1/TNFRSF10A, ICAM-1/CD54, TRAIL R2/TNFRSF10B, ICAM-2/CD102, TRAIL R3/TNFRSF10C, IFN-gamma R1, TRAIL R4/TNFRSF10D, IFN-gamma R2, TSLP, IL-1 RI and TSLP R.

In other embodiments, the NK cell receptor is selected from the group consisting of 2B4/SLAMF4, KIR2DS4, CD155/PVR, KIR3DL1, CD94, LMIR1/CD300A, CD69, LMIR2/CD300c, CRACC/SLAMF7, LMIR3/CD300LF, DNAM-1, LMIR5/CD300LB, Fc epsilon RII, LMIR6/CD300LE, Fc gamma RI/CD64, MICA, Fc gamma RIIB/CD32b, MICB, Fc gamma RIIC/CD32c, MULT-1, Fc gamma RIIA/CD32a, Nectin-2/CD112, Fc gamma RIII/CD16, NKG2A, FcRH1/IRTA5, NKG2C, FcRH2/IRTA4, NKG2D, FcRH4/IRTA1, NKp30, FcRH5/IRTA2, NKp44, Fc Receptor-like 3/CD16-2, NKp46/NCR1, NKp80/KLRF1, NTB-A/SLAMF6, Rae-1, Rae-1 alpha, Rae-1 beta, Rae-1 delta, H60, Rae-1 epsilon, ILT2/CD85j, Rae-1 gamma, ILT3/CD85k, TREM-1, ILT4/CD85d, TREM-2, ILT5/CD85a, TREM-3, KIR/CD158, TREML1/TLT-1, KIR2DL1, ULBP-1, KIR2DL3, ULBP-2, KIR2DL4/CD158d and ULBP-3.

In other embodiments, the monocyte/macrophage cell target is selected from the group consisting of B7-1/CD80, ILT4/CD85d, B7-H1, ILT5/CD85a, Common beta Chain, Integrin alpha 4/CD49d, BLAME/SLAMF8, Integrin alpha X/CD11c, CCL6/C10, Integrin beta 2/CD18, CD155/PVR, Integrin beta 3/CD61, CD31/PECAM-1, Latexin, CD36/SR-B3, Leukotriene B4 R1, CD40/TNFRSF5,

LIMPII/SR-B2, CD43, LMIR1/CD300A, CD45, LMIR2/CD300c, CD68, LMIR3/CD300LF, CD84/SLAMF5, LMIR5/CD300LB, CD97, LMIR6/CD300LE, CD163, LRP-1, CD2F-10/SLAMF9, MARCO, CRACC/SLAMF7, MD-1, ECF-L, MD-2, EMMPRIN/CD147, MGL2, Endoglin/CD105, Osteoactivin/GPNMB, Fc gamma RI/CD64, Osteopontin, Fc gamma RIIB/CD32b, PD-L2, Fc gamma RIIC/CD32c, Siglec-3/CD33, Fc gamma RIIA/CD32a, SIGNR1/CD209, Fc gamma RIII/CD16, SLAM, GM-CSF R alpha, TCCR/WSX-1, ICAM-2/CD102, TLR3, IFN-gamma R1, TLR4, IFN-gamma R2, TREM-1, IL-1 RII, TREM-2, ILT2/CD85j, TREM-3, ILT3/CD85k, TREML1/TLT-1, 2B4/SLAMF4, IL-10 R alpha, ALCAM, IL-10 R beta, Aminopeptidase N/ANPEP, ILT2/CD85j, Common beta Chain, ILT3/CD85k, C1q R1/CD93, ILT4/CD85d, CCR1, ILT5/CD85a, CCR2, Integrin alpha 4/CD49d, CCR5, Integrin alpha M/CD11b, CCR8, Integrin alpha X/CD11c, CD155/PVR, Integrin beta 2/CD18, CD14, Integrin beta 3/CD61, CD36/SR-B3, LAIR1, CD43, LAIR2, CD45, Leukotriene B4 R1, CD68, LIMPII/SR-B2, CD84/SLAMF5, LMIR1/CD300A, CD97, LMIR2/CD300c, CD163, LMIR3/CD300LF, Coagulation Factor III/Tissue Factor, LMIR5/CD300LB, CX3CR1, CX3CL1, LMIR6/CD300LE, CXCR4, LRP-1, CXCR6, M-CSF R, DEP-1/CD148, MD-1, DNAM-1, MD-2, EMMPRIN/CD147, MMR, Endoglin/CD105, NCAM-L1, Fc gamma RI/CD64, PSGL-1, Fc gamma RIII/CD16, RP105, G-CSF R, L-Selectin, GM-CSF R alpha, Siglec-3/CD33, HVEM/TNFRSF14, SLAM, ICAM-1/CD54, TCCR/WSX-1, ICAM-2/CD102, TREM-1, IL-6 R, TREM-2, CXCR1/IL-8 RA, TREM-3 and TREML1/TLT-1.

In yet other embodiments of the method, a Dendritic cell target is selected from the group consisting of CD36/SR-B3, LOX-1/SR-E1, CD68, MARCO, CD163, SR-AI/MSR, CD5L, SREC-I, CL-P1/COLEC12, SREC-II, LIMPII/SR-B2, RP105, TLR4, TLR1, TLR5, TLR2, TLR6, TLR3, TLR9, 4-1BB Ligand/TNFSF9, IL-12/IL-23 p40, 4-Amino-1,8-naphthalimide, ILT2/CD85j, CCL21/6Ckine, ILT3/CD85k, 8-oxo-dG, ILT4/CD85d, 8D6A, ILT5/CD85a, A2B5, Integrin alpha 4/CD49d, Aag, Integrin beta 2/CD18, AMICA, Langerin, B7-2/CD86, Leukotriene B4 R1, B7-H3, LMIR1/CD300A, BLAME/SLAMF8, LMIR2/CD300c, C1q R1/CD93, LMIR3/CD300LF, CCR6, LMIR5/CD300LB, CCR7, LMIR6/CD300LE, CD40/TNFRSF5, MAG/Siglec-4a, CD43, MCAM, CD45, MD-1, CD68, MD-2, CD83, MDL-1/CLEC5A, CD84/SLAMF5, MMR, CD97, NCAM-L1, CD2F-

10/SLAMF9, Osteoactivin/GPNMB, Chem 23, PD-L2, CLEC-1, RP105, CLEC-2, Siglec-2/CD22, CRACC/SLAMF7, Siglec-3/CD33, DC-SIGN, Siglec-5, DC-SIGNR/CD299, Siglec-6, DCAR, Siglec-7, DCIR/CLEC4A, Siglec-9, DEC-205, Siglec-10, Dectin-1/CLEC7A, Siglec-F, Dectin-2/CLEC6A, SIGNR1/CD209, DEP-1/CD148, SIGNR4, DLEC, SLAM, EMMPRIN/CD147, TCCR/WSX-1, Fc gamma RI/CD64, TLR3, Fc gamma RIIB/CD32b, TREM-1, Fc gamma RIIC/CD32c, TREM-2, Fc gamma RIIA/CD32a, TREM-3, Fc gamma RIII/CD16, TREML1/TLT-1, ICAM-2/CD102 and Vanilloid R1.

In still other embodiments of the method, the angiogenesis target is selected from the group consisting of Angiopoietin-1, Angiopoietin-like 2, Angiopoietin-2, Angiopoietin-like 3, Angiopoietin-3, Angiopoietin-like 7/CDT6, Angiopoietin-4, Tie-1, Angiopoietin-like 1, Tie-2, Angiogenin, iNOS, Coagulation Factor III/Tissue Factor, nNOS, CTGF/CCN2, NOV/CCN3, DANCE, OSM, EDG-1, Plfr, EG-VEGF/PK1, Proliferin, Endostatin, ROBO4, Erythropoietin, Thrombospondin-1, Kininostatin, Thrombospondin-2, MFG-E8, Thrombospondin-4, Nitric Oxide, VG5Q, eNOS, EphA1, EphA5, EphA2, EphA6, EphA3, EphA7, EphA4, EphA8, EphB1, EphB4, EphB2, EphB6, EphB3, Ephrin-A1, Ephrin-A4, Ephrin-A2, Ephrin-A5, Ephrin-A3, Ephrin-B1, Ephrin-B3, Ephrin-B2, FGF acidic, FGF-12, FGF basic, FGF-13, FGF-3, FGF-16, FGF-4, FGF-17, FGF-5, FGF-19, FGF-6, FGF-20, FGF-8, FGF-21, FGF-9, FGF-23, FGF-10, KGF/FGF-7, FGF-11, FGF R1, FGF R4, FGF R2, FGF R5, FGF R3, Neuropilin-1, Neuropilin-2, Semaphorin 3A, Semaphorin 6B, Semaphorin 3C, Semaphorin 6C, Semaphorin 3E, Semaphorin 6D, Semaphorin 6A, Semaphorin 7A, MMP, MMP-11, MMP-1, MMP-12, MMP-2, MMP-13, MMP-3, MMP-14, MMP-7, MMP-15, MMP-8, MMP-16/MT3-MMP, MMP-9, MMP-24/MT5-MMP, MMP-10, MMP-25/MT6-MMP, TIMP-1, TIMP-3, TIMP-2, TIMP-4, ACE, IL-13 R alpha 1, IL-13, C1q R1/CD93, Integrin alpha 4/CD49d, VE-Cadherin, Integrin beta 2/CD18, CD31/PECAM-1, KLF4, CD36/SR-B3, LYVE-1, CD151, MCAM, CL-P1/COLEC12, Nectin-2/CD112, Coagulation Factor III/Tissue Factor, E-Selectin, D6, P-Selectin, DC-SIGNR/CD299, SLAM, EMMPRIN/CD147, Tie-2, Endoglin/CD105, TNF RI/TNFRSF1A, EPCR, TNF RII/TNFRSF1B, Erythropoietin R, TRAIL R1/TNFRSF10A, ESAM, TRAIL R2/TNFRSF10B, FABP5, VCAM-1, ICAM-1/CD54, VEGF R2/Flk-1, ICAM-2/CD102, VEGF R3/Flt-4, IL-1 RI and VG5Q.

Other embodiments of the method provide multivalent binding proteins wherein at least one of binding domain 1 and binding domain 2 specifically binds a target selected from the group consisting of Prostate-specific Membrane Antigen (Folate Hydrolase 1), Epidermal Growth Factor Receptor (EGFR), Receptor for Advanced Glycation End products (RAGE, also known as Advanced Glycosylation End product Receptor or AGER), IL-17 A, IL-17 F, P19 (IL23A and IL12B), Dickkopf-1 (Dkk1), NOTCH1, NG2 (Chondroitin Sulfate ProteoGlycan 4 or CSPG4), IgE (IgHE or IgH2), IL-22R (IL22RA1), IL-21, Amyloid β oligomers (Ab oligomers), Amyloid β Precursor Protein (APP), NOGO Receptor (RTN4R), Low Density LipoproteinReceptor-Related Protein 5 (LRP5), IL-4, Myostatin (GDF8), Very Late Antigen 4, an alpha 4, beta 1 integrin (VLA4 or ITGA4), an alpha 4, beta 7 integrin found on leukocytes, and IGF-1R. For example, a VLA4 target may be recognized by a multivalent binding protein in which at least one of binding domain 1 and binding domain 2 is a binding domain derived from Natalizumab (Antegren).

In some embodiments, the cancer cell is a transformed, or cancerous, hematopoietic cell. In certain of these embodiments, at least one of the first binding domain and the second binding domain recognizes a target selected from the group consisting of a B-cell target, a monocyte/macrophage target, a dendritic cell target, an NK-cell target and a T-cell target, each as herein defined. Further, at least one of the first binding domain and the second binding domain can recognize a myeloid targets, including but not limited to, CD5, CD10, CD11b, CD11c, CD13, CD14, CD15, CD18, CD19, CD20, CD21, CD22, CD23, CD25, CD27, CD29, CD30, CD31, CD33, CD34, CD35, CD38, CD43, CD45, CD64, CD66, CD68, CD70, CD80, CD86, CD87, CD88, CD89, CD98, CD100, CD103, CD111, CD112, CD114, CD115, CD116, CD117, CD118, CD119, CD120a, CD120b, CDw123, CDw131, CD141, CD162, CD163, CD177, CD312, IRTA1, IRTA2, IRTA3, IRTA4, IRTA5, B-B2, B-B8 and B-cell antigen receptor.

Other embodiments of the invention are drawn to the multivalent binding protein, as described herein, comprising a sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 103, 105, 107, 109, 332, 333, 334, and 345. Other embodiments are directed to the multivalent binding protein comprising a sequence

selected from the group consisting of SEQ ID NOS:355, 356, 357, 358, 359, 360, 361, 362, 363, 364 and 365.

In other embodiments, the multivalent and multispecific binding protein with effector function has a first binding domain and a second binding domain that recognize a target pair selected from the group consisting of EPHB4-KDR and TIE-TEK. In such embodiments, the protein has a first binding domain recognizing EPHB4 and a second binding domain recognizing KDR or a first binding domain recognizing KDR and a second binding domain recognizing EPHB4. Analogously, the protein may have a first binding domain recognizing TIE and a second binding domain recognizing TEK, or a first binding domain recognizing TEK and a second binding domain recognizing TIE.

In a related aspect, the invention provides a multivalent binding protein with effector function, wherein the constant sub-region recognizes an effector cell F_C receptor (e.g., $F_C\gamma RI$, $F_C\gamma RII$, $F_C\gamma RIII$, $F_C\alpha R$, and $F_C\varepsilon RI$. In particular embodiments, the constant sub-region recognizes an effector cell surface protein selected from the group consisting of CD2, CD3, CD16, CD28, CD32, CD40, CD56, CD64, CD89, $F_C\varepsilon RI$, KIR, thrombospondin R, NKG2D, 2B4/NAIL and 41BB. The constant sub-region may comprise a C_{H2} domain and a C_{H3} domain derived from the same, or different, immunoglobulins, antibody isotypes, or allelic variants. In some embodiments, the C_{H3} domain is truncated and comprises a C-terminal sequence selected from the group consisting of SEQ ID NOS: 366, 367, 368, 369, 370 and 371. Preferably, the C_{H2} domain and the scorpion linker are derived from the same class, or from the same sub-class, of immunoglobulin, when the linker is a hinge-like peptide derived from an immunoglobulin.

Some proteins according to the invention are also contemplated as further comprising a scorpion linker of at least about 5 amino acids attached to the constant sub-region and attached to the second binding domain, thereby localizing the scorpion linker between the constant sub-region and the second binding domain. Typically, the scorpion linker peptide length is between 5-45 amino acids. Scorpion linkers include hinge-like peptides derived from immunoglobulin hinge regions, such as IgG1, IgG2, IgG3, IgG4, IgA, and IgE hinge regions. Preferably, a hinge-like scorpion linker will retain at least one cysteine capable of forming an interchain disulfide bond under

physiological conditions. Scorpion linkers derived from IgG1 may have 1 cysteine or two cysteines, and will preferably retain the cysteine corresponding to an N-terminal hinge cysteine of IgG1. In some embodiments, the scorpion linker is extended relative to a cognate immunoglobulin hinge region and, in exemplary embodiments, comprises a sequence selected from the group consisting of SEQ ID NOS:351, 352, 353 and 354. Non-hinge-like peptides are also contemplated as scorpion linkers, provided that such peptides provide sufficient spacing and flexibility to provide a single-chain protein capable of forming two binding domains, one located towards each protein terminus (N and C) relative to a more centrally located constant sub-region domain. Exemplary non-hinge-like scorpion linkers include peptides from the stalk region of type II C-lectins, such as the stalk regions of CD69, CD72, CD94, NKG2A and NKG2D. In some embodiments, the scorpion linker comprises a sequence selected from the group consisting of SEQ ID NOS:373, 374, 375, 376 and 377.

The proteins may also comprise a linker of at least about 5 amino acids attached to the constant sub-region and attached to the first binding domain, thereby localizing the linker between the constant sub-region and the first binding domain. In some embodiments, linkers are found between the constant sub-region and each of the two binding domains, and those linkers may be of the same or different sequence, and of the same or different lengths.

The constant sub-region of the proteins according to the invention provides at least one effector function. Any effector function known in the art to be associated with an immunoglobulin (e.g., an antibody) is contemplated, such as an effector function selected from the group consisting of antibody-dependent cell-mediated cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC), relatively extended *in vivo* half-life (relative to the same molecule lacking a constant sub-region), FcR binding, protein A binding, and the like. In some embodiments, the extended half-lives of proteins of the invention are at least 28 hours in a human. Of course, proteins intended for administration to non-human subjects will exhibit relatively extended half-lives in those non-human subjects, and not necessarily in humans.

In general, the proteins (including polypeptides and peptides) of the invention exhibit a binding affinity of less than 10^{-9} M, or at least 10^{-6} M, for at least one of the first binding domain and the second binding domain.

Another aspect of the invention is drawn to a pharmaceutical composition comprising a protein as described herein and a pharmaceutically acceptable adjuvant, carrier or excipient. Any adjuvant, carrier, or excipient known in the art is useful in the pharmaceutical compositions of the invention.

Yet another aspect of the invention provides a method of producing a protein as described above comprising introducing a nucleic acid encoding the protein into a host cell and incubating the host cell under conditions suitable for expression of the protein, thereby expressing the protein, preferably at a level of at least 1 mg/liter. In some embodiments, the method further comprises isolating the protein by separating it from at least one protein with which it is associated upon intracellular expression. Suitable host cells for expressing the nucleic acids to produce the proteins of the invention include, but are not limited to, a host cell selected from the group consisting of a VERO cell, a HeLa cell, a CHO cell, a COS cell, a W138 cell, a BHK cell, a HepG2 cell, a 3T3 cell, a RIN cell, an MDCK cell, an A549 cell, a PC12 cell, a K562 cell, a HEK293 cell, an N cell, a *Spodoptera frugiperda* cell, a *Saccharomyces cerevisiae* cell, a *Pichia pastoris* cell, any of a variety of fungal cells and any of a variety of bacterial cells (including, but not limited to, *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, and a Streptomycete).

The invention also provides a method of producing a nucleic acid encoding the protein, as described above, comprising covalently linking the 3' end of a polynucleotide encoding a first binding domain derived from an immunoglobulin variable region to the 5' end of a polynucleotide encoding a constant sub-region , covalently linking the 5' end of a polynucleotide encoding a scorpion linker to the 3' end of the polynucleotide encoding the constant sub-region, and covalently linking the 5' end of a polynucleotide encoding a second binding domain derived from an immunoglobulin variable region to the 3' end of the polynucleotide encoding the scorpion linker, thereby generating a nucleic acid encoding a multivalent binding protein with effector function. Each of these coding regions may be separated by a coding region for a linker or hinge-like peptide as part of a single-chain structure

according to the invention. In some embodiments, the method produces a polynucleotide encoding a first binding domain that comprises a sequence selected from the group consisting of SEQ ID NO: 2 (anti-CD20 variable region, oriented V_L-V_H), SEQ ID NO: 4 (anti-CD28 variable region, oriented V_L-V_H) and SEQ ID NO: 6 (anti-CD28 variable region, oriented V_H-V_L) in single-chain form, rather than requiring assembly from separately encoded polypeptides as must occur for heteromultimeric proteins, including natural antibodies. Exemplary polynucleotide sequences encoding first binding domains are polynucleotides comprising any of SEQ ID NOS: 1, 3 or 5.

This aspect of the invention also provides methods for producing encoding nucleic acids that further comprise a linker polynucleotide inserted between the polynucleotide encoding a first binding domain and the polynucleotide encoding a constant sub-region, the linker polynucleotide encoding a peptide linker of at least 5 amino acids. Additionally, these methods produce nucleic acids that further comprise a linker polynucleotide inserted between the polynucleotide encoding a constant sub-region and the polynucleotide encoding a second binding domain, the linker polynucleotide encoding a peptide linker of at least 5 amino acids. Preferably, the encoded peptide linkers are between 5 and 45 amino acids.

The identity of the linker regions present either between BD1 and EFD or EFD and BD2 may be derived from other sequences identified from homologous –Ig superfamily members. In developing novel linkers derived from existing sequences present in homologous members of the –Ig superfamily, it may be preferable to avoid sequence stretches similar to those located between the end of a C-like domain and the transmembrane domain, since such sequences are often substrates for protease cleavage of surface receptors from the cell to create soluble forms. Sequence comparisons between different members of the –Ig superfamily and subfamilies can be compared for similarities between molecules in the linker sequences that join multiple V-like domains or between the V and C like domains. From this analysis, conserved, naturally occurring sequence patterns may emerge; these sequences when used as the linkers between subdomains of the multivalent fusion proteins should be more protease resistant, might facilitate proper folding between Ig loop regions, and

would not be immunogenic since they occur in the extracellular domains of endogenous cell surface molecules.

The nucleic acids themselves comprise another aspect of the invention. Contemplated are nucleic acids encoding any of the proteins of the invention described herein. As such, the nucleic acids of the invention comprise, in 5' to 3' order, a coding region for a first binding domain, a constant sub-region sequence, and a coding region for a second binding domain. Also contemplated are nucleic acids that encode protein variants wherein the two binding domains and the constant sub-region sequences are collectively at least 80%, and preferably at least 85%, 90%, 95%, or 99% identical in amino acid sequence to the combined sequences of a known immunoglobulin variable region sequence and a known constant sub-region sequence. Alternatively, the protein variants of the invention are encoded by nucleic acids that hybridize to a nucleic acid encoding a non-variant protein of the invention under stringent hybridization conditions of 0.015 M sodium chloride, 0.0015 M sodium citrate at 65-68°C or 0.015 M sodium chloride, 0.0015 M sodium citrate, and 50% formamide at 42°C. Variant nucleic acids of the invention exhibit the capacity to hybridize under the conditions defined immediately above, or exhibit 90%, 95%, 99%, or 99.9% sequence identity to a nucleic acid encoding a non-variant protein according to the invention.

In related aspects, the invention provides a vector comprising a nucleic acid as described above, as well as host cells comprising a vector or a nucleic acid as described herein. Any vector known in the art may be used (e.g., plasmids, phagemids, phasmids, cosmids, viruses, artificial chromosomes, shuttle vectors and the like) and those of skill in the art will recognize which vectors are particularly suited for a given purpose. For example, in methods of producing a protein according to the invention, an expression vector operable in the host cell of choice is selected. In like manner, any host cell capable of being genetically transformed with a nucleic acid or vector of the invention is contemplated. Preferred host cells are higher eukaryotic host cells, although lower eukaryotic (e.g., yeast) and prokaryotic (bacterial) host cells are contemplated.

Another aspect of the invention is drawn to a method of inducing damage to a target cell comprising contacting a target cell with a therapeutically effective amount

of a protein as described herein. In some embodiments, the target cell is contacted *in vivo* by administration of the protein, or an encoding nucleic acid, to an organism in need. Contemplated within this aspect of the invention are methods wherein the multivalent single-chain binding protein induces an additive amount of damage to the target cell, which is that amount of damage expected from the sum of the damage attributable to separate antibodies comprising one or the other of the binding domains. Also contemplated are methods wherein the multivalent single-chain binding protein induces a synergistic amount of damage to the target cell compared to the sum of the damage induced by a first antibody comprising the first binding domain but not the second binding domain and a second antibody comprising the second binding domain but not the first binding domain. In some embodiments, the multivalent single-chain binding protein is multispecific and comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD19/CD20, CD20/CD21, CD20/CD22, CD20/CD40, CD20/CD79a, CD20/CD79b, CD20/CD81, CD21/CD79b, CD37/CD79b, CD79b/CD81, CD19/CL II (i.e., MHC class II), CD20/CL II, CD30/CL II, CD37/CL II, CD72/CL II, and CD79b/CL II.

This aspect of the invention also comprehends methods wherein the multispecific, multivalent single-chain binding protein induces an inhibited amount of damage to the target cell compared to the sum of the damage induced by a first antibody comprising the first binding domain but not the second binding domain and a second antibody comprising the second binding domain but not the first binding domain. Exemplary embodiments include methods wherein the multi-specific, multivalent single-chain binding protein comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20/CL II, CD21/CD79b, CD22/CD79b, CD40/CD79b, CD70/CD79b, CD72/CD79b, CD79a/CD79b, CD79b/CD80, CD79b/CD86, CD21/CL II, CD22/CL II, CD23/CL II, CD40/CL II, CD70/CL II, CD80/CL II, CD86/CL II, CD19/CD22, CD20/CD22, CD21/CD22, CD22/CD23, CD22/CD30, CD22/CD37, CD22/CD40, CD22/CD70, CD22/CD72, CD22/79a, CD22/79b, CD22/CD80, CD22/CD86 and CD22/CL II.

In a related aspect, the invention provides a method of treating a cell proliferation disorder, e.g., cancer, comprising administering a therapeutically effective amount of a protein (as described herein), or an encoding nucleic acid, to an

organism in need. Those of skill in the art, including medical and veterinary professionals, are proficient at identifying organisms in need of treatment. Disorders contemplated by the invention as amenable to treatment include a disorder selected from the group consisting of a cancer, an autoimmune disorder, Rous Sarcoma Virus infection and inflammation. In some embodiments, the protein is administered by *in vivo* expression of a nucleic acid encoding the protein as described herein. The invention also comprehends administering the protein by a route selected from the group consisting of intravenous injection, intraarterial injection, intramuscular injection, subcutaneous injection, intraperitoneal injection and direct tissue injection.

Another aspect of the invention is directed to a method of ameliorating a symptom associated with a cell proliferation disorder comprising administering a therapeutically effective amount of a protein, as described herein, to an organism in need. Those of skill in the art are also proficient at identifying those disorders, or diseases or conditions, exhibiting symptoms amenable to amelioration. In some embodiments, the symptom is selected from the group consisting of pain, heat, swelling and joint stiffness.

Yet another aspect of the invention is drawn to a method of treating an infection associated with an infectious agent comprising administering a therapeutically effective amount of a protein according to the invention to a patient in need, wherein the protein comprises a binding domain that specifically binds a target molecule of the infectious agent. Infectious agents amenable to treatment according to this aspect of the invention include prokaryotic and eukaryotic cells, viruses (including bacteriophage), foreign objects, and infectious organisms such as parasites (e.g., mammalian parasites).

A related aspect of the invention is directed to a method of ameliorating a symptom of an infection associated with an infectious agent comprising administering an effective amount of a protein according to the invention to a patient in need, wherein the protein comprises a binding domain that specifically binds a target molecule of the infectious agent. Those of skill in the medical and veterinary arts will be able to determine an effective amount of a protein on a case-by-case basis, using routine experimentation.

Yet another related aspect of the invention is a method of reducing the risk of infection attributable to an infectious agent comprising administering a prophylactically effective amount of a protein according to the invention to a patient at risk of developing the infection, wherein the protein comprises a binding domain that specifically binds a target molecule of the infectious agent. Those of skill in the relevant arts will be able to determine a prophylactically effective amount of a protein on a case-by-case basis, using routine experimentation.

Another aspect of the invention is drawn to the above-described multivalent single-chain binding protein wherein at least one of the first binding domain and the second binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide.

In certain embodiments, one of the first binding domain and the second binding domain specifically binds CD20, and in some of these embodiments, the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide. For example, in one embodiment, the first binding domain is capable of specifically binding CD20 while the second binding domain is capable of specifically binding, e.g., CD19. In another embodiment, the first binding domain binds CD19 while the second binding domain binds CD20. An embodiment in which both binding domains bind CD20 is also contemplated.

In certain other embodiments according to this aspect of the invention, one of the first binding domain and the second binding domain specifically binds CD79b, and in some of these embodiments, the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide. Exemplary embodiments include distinct multi-specific, multivalent single-chain binding proteins in which a first binding domain:second binding domain specifically binds CD79b:CD19 or CD19:CD79b. A

multivalent binding protein having first and second binding domains recognizing CD79b is also comprehended.

In still other certain embodiments, one of the first binding domain and the second binding domain specifically binds a major histocompatibility complex class II peptide, and in some of these embodiments, the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide. For example, in one embodiment, the first binding domain is capable of specifically binding a major histocompatibility complex class II peptide while the second binding domain is capable of specifically binding, e.g., CD19. In another embodiment, the first binding domain binds CD19 while the second binding domain binds a major histocompatibility complex class II peptide. An embodiment in which both binding domains bind a major histocompatibility complex class II peptide is also contemplated.

In yet other embodiments according to this aspect of the invention, one of the first binding domain and the second binding domain specifically binds CD22, and in some of these embodiments, the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide. Exemplary embodiments include distinct multi-specific, multivalent single-chain binding proteins in which a first binding domain:second binding domain specifically binds CD22:CD19 or CD19:CD22. A multivalent binding protein having first and second binding domains recognizing CD22 is also comprehended.

A related aspect of the invention is directed to the above-described multivalent single-chain binding protein wherein the protein has a synergistic effect on a target cell behavior relative to the sum of the effects of each of the binding domains. In some embodiments, the protein comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20-CD19, CD20-CD21, CD20-CD22, CD20-CD40, CD20-CD79a, CD20-CD79b and CD20-CD81.

The invention further comprehends a multivalent single-chain binding protein as described above wherein the protein has an additive effect on a target cell behavior relative to the sum of the effects of each of the binding domains. Embodiments according to this aspect of the invention include multi-specific proteins comprising a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20-CD23, CD20-CD30, CD20-CD37, CD20-CD70, CD20-CD80, CD20-CD86, CD79b-CD37, CD79b-CD81, major histocompatibility complex class II peptide-CD30, and major histocompatibility complex class II peptide-CD72.

Yet another related aspect of the invention is a multivalent single-chain binding protein as described above wherein the protein has an inhibitory effect on a target cell behavior relative to the sum of the effects of each of the binding domains. In some embodiments, the protein is multispecific and comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20-major histocompatibility complex class II peptide, CD79b-CD19, CD79b-CD20, CD79b-CD21, CD79b-CD22, CD79b-CD23, CD79b-CD30, CD79b-CD40, CD79b-CD70, CD79b-CD72, CD79b-CD79a, CD79b-CD80, CD79b-CD86, CD79b-major histocompatibility complex class II peptide, major histocompatibility complex class II peptide-CD19, major histocompatibility complex class II peptide-CD20, major histocompatibility complex class II peptide-CD21, major histocompatibility complex class II peptide-CD22, major histocompatibility complex class II peptide-CD23, major histocompatibility complex class II peptide-CD37, major histocompatibility complex class II peptide-CD40, major histocompatibility complex class II peptide-CD70, major histocompatibility complex class II peptide-CD79a, major histocompatibility complex class II peptide-CD79b, major histocompatibility complex class II peptide-CD80, major histocompatibility complex class II peptide-CD81, major histocompatibility complex class II peptide-CD86, CD22-CD19, CD22-CD40, CD22-CD79b, CD22-CD86 and CD22-major histocompatibility complex class II peptide.

Another aspect of the invention is a method of identifying at least one of the binding domains of the multivalent binding molecule, such as a multispecific binding molecule, described above comprising: (a) contacting an anti-isotypic antibody with an antibody specifically recognizing a first antigen and an antibody specifically

recognizing a second antigen; (b) further contacting a target comprising at least one of said antigens with the composition of step (a); and (c) measuring an activity of the target, wherein the activity is used to identify at least one of the binding domains of the multivalent binding molecule. In some embodiments, the target is a diseased cell, such as a cancer cell (e.g., a cancerous B-cell) or an auto-antibody-producing B-cell.

In each of the foregoing methods of the invention, it is contemplated that the method may further comprise a plurality of multivalent single-chain binding proteins. In some embodiments, a binding domain of a first multivalent single-chain binding protein and a binding domain of a second multivalent single-chain binding protein induce a synergistic, additive, or inhibitory effect on a target cell, such as a synergistic, additive, or inhibitory amount of damage to the target cell. The synergistic, additive or inhibitory effects of a plurality of multivalent single-chain binding proteins is determined by comparing the effect of such a plurality of proteins to the combined effect of an antibody comprising one of the binding domains and an antibody comprising the other binding domain.

A related aspect of the invention is directed to a composition comprising a plurality of multivalent single-chain binding proteins as described above. In some embodiments, the composition comprises a plurality of multivalent single-chain binding proteins wherein a binding domain of a first multivalent single-chain binding protein and a binding domain of a second multivalent single-chain binding protein are capable of inducing a synergistic, additive, or inhibitory effect on a target cell, such as a synergistic, additive or inhibitory amount of damage to the target cell.

The invention further extends to a pharmaceutical composition comprising the composition described above and a pharmaceutically acceptable carrier, diluent or excipient. In addition, the invention comprehends a kit comprising the composition as described herein and a set of instructions for administering said composition to exert an effect on a target cell, such as to damage the target cell.

Finally, the invention also comprehends a kit comprising the protein as described herein and a set of instructions for administering the protein to treat a cell proliferation disorder or to ameliorate a symptom of the cell proliferation disorder.

Other features and advantages of the present invention will be better understood by reference to the following detailed description, including the examples.

BRIEF DESCRIPTION OF THE DRAWING

Figure 1 shows a schematic representation of the multivalent single-chain molecules envisioned by the invention. Individual subdomains of the fusion protein expression cassette are indicated by separate shapes/blocks on the figure. BD1 refers to binding domain 1, linker 1 refers to any potential linker or hinge like peptide between BD1 and the "effector function domain", indicated as EFD. This subdomain is usually an engineered form of the Fc domain of human IgG1, but may include other subdomains with one or more effector functions as defined herein. Linker 2 refers to the linker sequence, if any, present between the carboxy terminus of the EFD and the binding domain 2, BD2.

Figure 2 shows a Western blot of non-reduced proteins expressed in COS cells. Protein was secreted into the culture medium, and culture supernatants isolated after 48-72 hours from transiently transfected cells by centrifugation. Thirty microliters, 30 μl of crude supernatant were loaded into each well of the gel. Lane identifications: 1- molecular weight markers, with numerals indicating kilodaltons; 2-2H7-IgG-STD1-2E12 LH; 3- 2H7-IgG-STD1-2E12 HL, 4- 2H7-IgG-STD2-2E12 LH; 5- 2H7-IgG-STD2-2E12 HL; 6- 2E12 LH SMIP; 7- 2E12 HL SMIP; 8- 2H7 SMIP. "2H7" refers to a single-chain construct, where BD1 encodes the CD20 specific binding domain (2H7) in the VLVH orientation; "2E12" refers to a binding domain specific for CD28; -IgG-refers to a single-chain construct, with a hinge encoding a sequence where all C are mutated to S (sss), and the CH2 and CH3 domains of IgG1 contain mutations which eliminate both ADCC and CDC effector functions (P238S and P331 S), "STD 1 refers to a 20-amino-acid linker (identified in Figure 7 as "STD1=20aa ") inserted adjacent to the BD2 in the VL-VH orientation, or 2E12 (V_L-V_H). "STD1- HL" refers to a similar construct as just described, but with the BD2 V regions in the VH-VL orientation as follows: 2H7-sssIgG (P238/331S)-20-amino-acid linker-2E12 (V_H-V_L). "STD2- LH" refers to 2H7-sssIgG (P238/331S)-38-amino-acid linker-2E12 (V_L-V_H); "STD2-LH" refers to 2H7-sssIgG (P238/331SS)-38-amino-acid linker-2E12 (V_H-V_L); "SMIP" refers to small modular immunopharmaceutical; and "H" generally refers to V_H, while "L" generally refers to

V_L. Unless otherwise indicated, all protein orientations are N-terminal to C-terminal orientations.

Figure 3 shows two columnar graphs illustrating the binding properties of the 2H7-sssIgG (P238S/P331S)-STD1-2e12 LH and HL derivatives expressed from COS cells. These experiments were performed with crude culture supernatants rather than purified proteins. Serial dilutions from undiluted to 16X of the culture supernatants were incubated with either CD20 expressing cells (WIL-2S) or CD28 expressing cells (CD28 CHO). Binding activity in the supernatants was compared to control samples testing binding of the relevant single specificity SMIP, such as TRU-015, or 2e12 VLVH, or 2e12VHVL SMIPs. Binding in each sample was detected using fluorescein isothyocyanate (FITC) conjugated goat anti-human IgG at a dilution of 1:100.

Figure 4 is a histogram showing the binding pattern of protein A purified versions of the proteins tested in Figure 3 to WIL2-S cells. "TRU015" is a SMIP specific for CD20. Two multispecific binding proteins with effector function were also analyzed: "2H7-2E12 LH" has binding domain 2, specific for CD28, in V_L-V_H orientation; "2H7-2E12 HL" has binding domain 2, specific for CD28, in V_H-V_L orientation. Each of the proteins was tested for binding at 5 μg/ml, and binding detected with FITC goat anti-human IgG at 1:100. See the description for Figure 2 above for more complete descriptions of the molecules tested.

Figure 5 shows two histograms illustrating the binding by protein A purified multispecific binding proteins with effector function to CHO cells expressing CD28. "2H7-2E12 LH" has binding domain 2, specific for CD28, in V_L-V_H orientation; "2H7-2E12 HL" has binding domain 2, specific for CD28, in V_H-V_L orientation. Each of the proteins was tested for binding at 5 μg/ml, and binding was detected with FITC goat anti-human IgG at 1:100. See the descriptions in Figure 2 for a more complete description of the molecules tested.

Figure 6 A) shows a table which identifies the linkers joining the constant sub-region and binding domain 2. The linkers are identified by name, sequence, sequence identifier, sequence length, and the sequence of the fusion with binding domain 2. B) shows a table identifying a variety of constructs identifying elements of

exemplified molecules according to the invention.. In addition to identifying the multivalent binding molecules by name, the elements of those molecules are disclosed in terms of binding domain 1 (BD1), the constant sub-region (hinge and effector domain or EFD), a linker (see Fig. 6A for additional information regarding the linkers), and binding domain 2 (BD2). The sequences of a number of exemplified multivalent binding proteins are provided, and are identified in the figure by a sequence identifier. Other multivalent binding proteins have altered elements, or element orders, with predictable alterations in sequence from the disclosed sequences.

Figure 7 shows a composite columnar graph illustrating the binding of purified proteins at a single, fixed concentration to CD20 expressing WIL-2S cells and to CHO cells expressing CD28. "H1-H6" refers to the 2H7-sss-hIgG-Hx-2e12 molecules with the H1-H6 linkers and the 2e12 V regions in the orientation of V_H-V_L. "L1-L6" refers to the 2H7-sss-hIgG-Lx-2e12 molecules with the L1-L6 linkers and the 2e12 V regions in the orientation of V_L-V_H. All the molecules were tested at a concentration of 0.72 μg/ml, and the binding detected using FITC conjugated goat anti-human IgG at 1:100. The mean fluorescence intensity for each sample was then plotted as paired bar graphs for the two target cell types tested versus each of the multivalent constructs being tested, L1-L6, or H1-H6.

Figure 8 shows photographs of Coomassie stained non-reducing and reducing SDS-PAGE gels. These gels show the effect of the variant linker sequence/length on the 2H7-sss-hIgG-Hx-2e12 HL protein on the amounts of the two predominate protein bands visualized on the gel.

Figure 9 shows Western Blots of the [2H7-sss-hIgG-H6-2e12] fusion proteins and the relevant single specificity SMIPs probed with either (a) CD28mIgG or with (b) a Fab reactive with the 2H7 specificity. The results show that the presence of the H6 linker results in the generation of cleaved forms of the multivalent constructs which are missing the CD28 binding specificity.

Figure 10 shows binding curves of the different linker variants for the [TRU015-sss-IgG-Hx-2e12 HL] H1-H6 linker forms. The first panel shows the binding curves for binding to CD20 expressing WIL-2S cells. The second panel shows the binding curves for binding of the different forms to CD28 CHO cells.

These binding curves were generated with serial dilutions of protein A purified fusion protein, and binding detected using FITC conjugated goat anti-human IgG at 1:100.

Figure 11 shows a table summarizing the results of SEC fractionation of 2H7-sss-IgG-2e12 HL multispecific fusion proteins with variant linkers H1-H7. Each row in the table lists a different linker variant of the [2H7-sss-IgG-Hx-2e12-HL] fusion proteins. The retention time of the peak of interest (POI), and the percentage of the fusion protein present in POI, and the percentage of protein found in other forms is also tabulated. The cleavage of the molecules is also listed, with the degree of cleavage indicated in a qualitative way, with (Yes), Yes, and YES, or No being the four possible choices.

Figure 12 shows two graphs with binding curves for [2H7-sss-hIgG-Hx-2e12] multispecific fusion proteins with variant linkers H3, H6, and H7 linkers to cells expressing CD20 or CD28. Serial dilutions of the protein A purified fusion proteins from 10 μg/ml down to 0.005 μg/ml were incubated with either CD20 expressing WIL-2S cells or CD28 CHO cells. Binding was detected using FITC conjugated goat anti-human IgG at 1:100. Panel A shows the binding to WIL-2S cells, and panel B shows the binding to CD28 CHO cells.

Figure 13 shows the results of an alternative binding assay generated by the molecules used for Figure 12. In this case, the fusion proteins were first bound to WIL-2S CD20 expressing cells, and binding was then detected with CD28mIgG (5 μg/ml)and FITC anti-mouse reagent at 1:100. These results demonstrate the simultaneous binding to both CD20 and CD28 in the same molecule.

Figure 14 shows results obtained using another multispecific fusion construct variant. In this case, modifications were made in the specificity for BD2, so that the V regions for the G28-1 antibody were used to create a CD37 specific binding domain. Shown are two graphs which illustrate the relative ability of CD20 and/or CD37 antibodies to block the binding of the [2H7-sss-IgG-Hx-G28-1] multispecific fusion protein to Ramos or BJAB cells expressing the CD20 and CD37 targets. Each cell type was preincubated with either the CD20 specific antibody (25 μg/ml) or the CD37 specific antibody (10 μg/ml) or both reagents (these are mouse anti-human reagents) prior to incubation with the multispecific fusion protein. Binding of the

multispecific fusion protein was then detected with a FITC goat anti-human IgG reagent at 1:100, (preadsorbed to mouse to eliminate cross-reactivity).

Figure 15 shows the results of an ADCC assay performed with BJAB target cells, PBMC effector cells, and with the CD20-hIgG-CD37 specific fusion protein as the test reagent. For a full description of the procedure see the appropriate example. The graph plots the concentration of fusion protein versus the % specific killing at each dosage tested for the single specificity SMIP reagents, and for the [2H7-sss-hIgG-STD1-G28-1] LH and HL variants. Each data series plots the dose-response effects for one of these single specificity or multispecific single-chain fusion proteins.

Figure 16 shows a table tabulating the results of a co-culture experiment where PBMC were cultured in the presence of TRU 015, G28-1 SMIP, both molecules together, or the [2H7-sss-IgG-H7-G28-1HL] variant. The fusion proteins were used at 20 μg/ml, and incubated for 24 hours or 72 hours. Samples were then stained with CD3 antibodies conjugated to FITC, and either CD19 or CD40 specific antibodies conjugated to PE, then subjected to flow cytometry. The percentage of cells in each gate was then tabulated.

Figure 17 shows two columnar graphs of the effects on B cell line apoptosis after 24 hour incubation with the [2H7-sss-hIgG-H7-G28-1 HL] molecule or control single CD20 and/or CD37 specificity SMIPs alone or in combination. The percentage of annexin V-propidium iodide positive cells is plotted as a function of the type of test reagent used for the coincubation experiments. Panel A shows the results obtained using Ramos cells, and panel B shows those for Daudi cells. Each single CD20 or CD37 directed SMIP is shown at the concentrations indicated; in addition, where combinations of the two reagents were used, the relative amount of each reagent is shown in parentheses. For the multispecific CD20-CD37 fusion protein, concentrations of 5, 10, and 20 μg/ml were tested.

Figure 18 shows two graphs of the [2H7-hIgG-G19-4] molecule variants and their binding to either CD3 expressing cells (Jurkats) or CD20 expressing cells (WIL-2S). The molecules include [2H7-sss-hIgG-STD1-G19-4 HL], LH, and [2H7-csc-hIgG-STD1-G19-4 HL]. Protein A purified fusion proteins were titrated from 20 μg/ml down to 0.05 μg/ml, and the binding detected using FITC goat anti-human IgG

at 1:100. MFI (mean fluorescence intensity) is plotted as a function of protein concentration.

Figure 19 shows the results of ADCC assays performed with the [2H7-hIgG-STD1-G19-4 HL] molecule variants with either an SSS hinge or a CSC hinge, BJAB target cells, and either total human PBMC as effector cells or NK cell depleted PBMC as effector cells. Killing was scored as a function of concentration of the multispecific fusion proteins. The killing observed with these molecules was compared to that seen using G19-4, TRU 015, or a combination of these two reagents. Each data series plots a different test reagent, with the percent specific killing plotted as a function of protein concentration.

Figure 20 shows the percentage of Ramos B-cells that stained positive with Annexin V (Ann) and/or propidium iodide (PI) after overnight incubation with each member of a matrix panel of B-cell antibodies (2 μg/ml) in the presence, or absence, of an anti-CD20 antibody (present at 2 μg/ml where added). Goat-anti-mouse secondary antibody was always present at a two-fold concentration ratio relative to other antibodies (either matrix antibody alone, or matrix antibody and anti-CD20 antibody). Vertically striped bars – matrix antibody (2 μg/ml) denoted on X-axis and goat anti-mouse antibody (4 μg/ml). Horizontally striped bars - matrix antibody (2 μg/ml) denoted on X-axis, anti-CD20 antibody (2 μg/ml), and goat anti-mouse antibody (4 μg/ml). The "2nd step" condition served as a control and involved the addition of goat anti-mouse antibody at 4 μg/ml (vertically striped bar) or 8 μg/ml (horizontally striped bar), without a matrix antibody or anti-CD20 antibody. "CL II" (MHC class II) in the figures refers to a monoclonal antibody cross-reactive to HLA DR, DQ and DP, i.e., to MHC Class II antigens.

Figure 21 shows the percentage of Ramos B-cells that stained positive with Annexin V (Ann) and/or propidium iodide (PI) after overnight incubation with each member of a matrix panel of B-cell antibodies (2 μg/ml) in the presence, or absence, of an anti-CD79b antibody (present at either 0.5 or 1.0 μg/ml where added). See the description of Figure 20 for identification of "CL II" and "2nd step" samples. Vertically striped bars – matrix antibody (2 μg/ml) and goat anti-mouse antibody (4 μg/ml); horizontally striped bars – matrix antibody (2 μg/ml), anti-CD79b antibody

(1.0 μg/ml) and goat anti-mouse antibody (6 μg/ml); stippled bars – matrix antibody (2 μg/ml), anti-CD79b antibody (0.5 μg/ml) and goat anti-mouse antibody (5 μg/ml).

Figure 22 shows the percentage of Ramos B-cells that stained positive with Annexin V (Ann) and/or propidium iodide (PI) after overnight incubation with each member of a matrix panel of B-cell antibodies (2 μg/ml) in the presence, or absence, of an anti-CL II antibody (present at either 0.25 or 0.5 μg/ml where added). See the description of Figure 20 for identification of "CL II" and "2nd step" samples. Vertically striped bars – matrix antibody (2 μg/ml) and goat anti-mouse antibody (4 μg/ml); horizontally striped bars – matrix antibody (2 μg/ml), anti-CL II antibody (0.5 μg/ml) and goat anti-mouse antibody (5 μg/ml); stippled bars – matrix antibody (2 μg/ml), anti-CL II antibody (0.25 μg/ml) and goat anti-mouse antibody (4.5 μg/ml).

Figure 23 shows the percentage of DHL-4 B-cells that stained positive with Annexin V (Ann) and/or propidium iodide (PI) after overnight incubation with each member of a matrix panel of B-cell antibodies (2 μg/ml) in the presence, or absence, of an anti-CD22 antibody (present at 2 μg/ml where added). See the description of Figure 20 for identification of "CL II" and "2nd step" samples. Solid bars – matrix antibody (2 μg/ml) and goat anti-mouse antibody (4 μg/ml); slant-striped bars – matrix antibody (2 μg/ml), anti-CD22 antibody (2 μg/ml) and goat anti-mouse antibody (8 μg/ml).

Figure 24 provides a graph demonstrating direct growth inhibition of lymphoma cell lines Su-DHL6 (triangles) and DoHH2 (squares) by free CD20 SMIP (closed symbols) or monospecific CD20xCD20 scorpion (open symbols).

Figure 25 is a graph showing direct growth inhibition of lymphoma cell lines Su-DHL-6 (triangles) and DoHH2 (squares) by free anti-CD37 SMIP (closed symbols) or monospecific anti-CD37 scorpion (open symbols).

Figure 26 presents a graph showing direct growth inhibition of lymphoma cell lines Su-DHL-6 (triangles) and DoHH2 (squares) by a combination of two different monospecific SMIPs (closed symbols) or by a bispecific CD20-CD37 scorpion (open symbols).

Figure 27 is a graph revealing direct growth inhibition of lymphoma cell lines Su-DHL-6 (triangles) and WSU-NHL (squares) by free CD20 SMIP and CD37 SMIPcombination (closed symbols) or bispecific CD20xCD37 scorpion (open symbols).

Figure 28 provides histograms showing the cell-cycle effects of scorpions. Samples of DoHH2 lymphoma cells were separately left untreated, treated with SMIP 016 or treated with the monospecific CD37 x CD37 scorpion. Open bars: sub-G_1 phase of the cell cyle; black bars: G_0/G_1 phase; shaded: S phase; and striped: G_2/M phase.

Figure 29 presents graphs of data establishing that treatment of lymphoma cells with scorpions resulted in increased signaling capacity compared to free SMIPs, as measured by calcium ion flux.

Figure 30 provides graphs demonstrating scorpion-dependent cellular cytotoxicity

Figure 31 shows graphs of data indicating that scorpions mediate Complement Dependent Cytotoxicity.

Figure 32 provides data in graphical form showing comparative ELISA binding of a SMIP and a scorpion to low- (B) and high-affinity (A) isoforms of FcγRIII (CD16).

Figure 33 presents graphs establishing the binding of a SMIP and a scorpion to low (A)- and high (B)-affinity allelotypes of FcγRIII (CD16) in the presence of target cells.

Figure 34 is a histogram showing the expression level of a CD20 x CD20 scorpion in two experiments (flask 1 and flask 2) under six different culturing conditions. Solid black bars: flask 1; striped bars: flask 2.

Figure 35 provides a histogram showing the production yield of a CD20 x CD37 scorpion.

Figure 36 presents SDS-PAGE gels (under reducing and non-reducing conditions) of a SMIP and a scorpion.

Figure 37 provides a graph showing that scorpions retain the capacity to bind to target cells. Filled squares: CD20 SMIP; filled triangles: CD37 SMIP; filled circles: humanized CD20 (2Lm20-4) SMIP; open diamond: CD37 x CD37 monospecific scorpion; open squares: CD20 x CD37 bi-specific scorpion; and open triangles: humanized CD20 (2Lm20-4) x humanized CD20 (2Lm20-4) scorpion.

Figure 38 contains graphs showing the results of competitive binding assays establishing that both N- and C-terminal scorpion binding domains participate in target cell binding.

Figure 39 presents data in the form of graphs showing that scorpions have lower off-rates than SMIPs.

Figure 40 shows a graph establishing that scorpions are stable in serum *in vivo*, characterized by a reproducible,sustained circulating half-life for the scorpion.

Figure 41 provides a dose-response graph for a CD20 x CD37 bispecific scorpion, demonstrating the *in* vivo efficacy of scorpion administration.

Figure 42 shows target B-cell binding by a monospecific CD20xCD20 scorpion (S0129) and glycovariants.

Figure 43 provides graphs illustrating CD20xCD20 scorpions (parent and glycovariants) inducing ADCC-mediated killing of BJAB B-cells.

Figure 44 shows a gel revealing the effects on scorpion stability arising from changes in the scorpion linker, including changing the sequence of that linker and extending the linker by adding an H7 sequence to the linker, indicated by a "+" in the H7 line under the gel.

Figure 45 shows the binding to WIL2S cells of a CD20xCD20 scorpion (S0129) and scorpion linker variants thereof.

Figure 46 shows the direct cell killing of a variety of B-cells by a CD20xCD20 scorpion and by a CD20 SMIP.

Figure 47 reveals the direct cell killing of additional B-cell lines by a monospecific CD20xCD20 scorpion.

Figure 48 shows the direct cell killing capacities of each of two monospecific scorpions, i.e., CD20xCD20 and CD37xCD37, and a bispecific CD20xCD37 scorpion, the latter exhibiting a different form of kill curve.

Figure 49 graphically depicts the response of Su-DHL-6 B-cells to each of a CD20xCD20 (S0129), a CD37xCD37, and a CD20xCD37 scorpion.

Figue 50 shows the capacity of a bispecific CD19xCD37 scorpion and Rituxan® to directly kill Su-DHL-6 B-cells.

Figure 51 provides histograms showing the direct killing of DHL-4 B-cells by a variety of CD20-binding scorpions and SMIPs, as well as by Rituxan®, as indicated in the figure. Blue bars: live cells; maroon bars on the right of each pair: Annexin+/PI+.

Figure 52 provides a graphic depiction of the direct cell killing of various CD20-binding scorpions and SMIPs, as well as by Rituxan®, as indicated in the figure.

Figure 53 provides graphs of the ADCC activity induced by various CD20-binding scorpions and SMIPs, as indicated in the figure, as well as by Rituxan®.

Figure 54 provides graphs of the CDC activity induced by vawrious CD20-binding scorpions and SMIPs, as indicated in the figure, as well as by Rituxan®.

Figure 55 provides histograms showing the levels of C1q binding to CD20-binding scorpions bound to Ramos B-cells.

Figure 56 provides scatter plots of FACS analyses showing the loss of mitochondrial membrane potential attributable to CD20-binding scorpions (2Lm20-4x2Lm20-4 and 011x2Lm20-4) and Rituxan®, relative to controls (upper panel); histograms of the percentage of cells with disrupted mitochondrial membrane potential (disrupted MMP: black bars) are shown in the lower panel.

Figure 57 provides histograms showing the relative lack of caspase 3 activation by CD20-binding scorpions (2Lm20-4x2Lm20-4 and 011x2Lm20-4), Rituximab, CD95, and controls.

Figure 58 provides a composite of four Western blot analyses of Poly (ADP-ribose) Polymerase and caspases 3, 7, and 9 from B-cells showing little degradation of any of these proteins attributable to CD20-binding scorpions binding to the cells.

Figure 59 is a gel electrophoretogram of B-cell chromosomal DNAs showing the degree of fragmentation attributable to CD20-binding scorpions binding to the cells.

Figure 60 is a gel electrophoretogram of immunoprecipitates obtained with each of an anti-phosphotyrosine antibody and an anti-SYK antibody. The immunoprecipitates were obtained from lysates of B-cells contacted with CD20-binding scorpions, as indicated in the figure.

Figure 61 provides combination index plots of CD20-binding scorpions in combination therapies with each of doxorubicin, vincristine and rapamycin.

DETAILED DESCRIPTION

The present invention provides compositions of relatively small peptides having at least two binding regions or domains, which may provide one or more binding specificities, derived from variable binding domains of immunoglobulins, such as antibodies, disposed terminally relative to an effector domain comprising at least part of an immunoglobulin constant region (i.e., a source from which a constant sub-region, as defined herein, may be derived), as well as nucleic acids, vectors and host cells involved in the recombinant production of such peptides and methods of using the peptide compositions in a variety of diagnostic and therapeutic applications, including the treatment of a disorder as well as the amelioration of at least one symptom of such a disorder. The peptide compositions advantageously arrange a second binding domain C-terminal to the effector domain, an arrangement that unexpectedly provides sterically unhindered or less hindered binding by at least two binding domains of the peptide, while retaining an effector function or functions of the centrally disposed effector domain.

The first and second binding domains of the multivalent peptides according to the invention may be the same (i.e., have identical or substantially identical amino acid sequences and be monospecific) or different (and be multispecific). Although different in terms of primary structure, the first and second binding domains may recognize and bind to the same epitope of a target molecule and would therefore be monospecific. In many instances, however, the binding domains will differ structurally and will bind to different binding sites, resulting in a multivalent, multispecific protein. Those different binding sites may exist on a single target molecule or on different target molecules. In the case of the two binding molecules recognizing different target molecules, those target molecules may exist, e.g., on or in the same structure (e.g., the surface of the same cell), or those target molecules may exist on or in separate structures or locales. For example, a multispecific binding protein according to the invention may have binding domains that specifically bind to target molecule on the surfaces of distinct cell types. Alternatively, one binding domain may specifically bind to a target on a cell surface and the other binding domain may specifically bind to a target not found associated with a cell, such as an extracellular structural (matrix) protein or a free (e.g., soluble or stromal) protein.

The first and second binding domains are derived from one or more regions of the same, or different, immunoglobulin protein structures such as antibody molecules. The first and/or second binding domain may exhibit a sequence identical to the sequence of a region of an immunoglobulin, or may be a modification of such a sequence to provide, e.g., altered binding properties or altered stability. Such modifications are known in the art and include alterations in amino acid sequence that contribute directly to the altered property such as altered binding, for example by leading to an altered secondary or higher order structure for the peptide. Also contemplated are modified amino acid sequences resulting from the incorporation of non-native amino acids, such as non-native conventional amino acids, unconventional amino acids and imino acids. In some embodiments, the altered sequence results in altered post-translational processing, for example leading to an altered glycosylation pattern.

Any of a wide variety of binding domains derived from an immunoglobulin or immunoglobulin-like polypeptide (e.g., receptor) are contemplated for use in

scorpions. Binding domains derived from antibodies comprise the CDR regions of a V_L and a V_H domain, seen, e.g., in the context of using a binding domain from a humanized antibody. Binding domains comprising complete V_L and V_H domains derived from an antibody may be organized in either orientation. A scorpion according to the invention may have any of the binding domains herein described. For scorpions having at least one binding domain recognizing a B-cell, exemplary scorpions have at least one binding domain derived from CD3, CD10, CD19, CD20, CD21, CD22, CD23, CD24, CD37, CD38, CD39, CD40, CD72, CD73, CD74, CDw75, CDw76, CD77, CD78, CD79a/b, CD80, CD81, CD82, CD83, CD84, CD85, CD86, CD89, CD98, CD126, CD127, CDw130, CD138 or CDw150. In some embodiments, the scorpion is a multivalent binding protein comprising at least one binding domain having a sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 103, 105, 107 and 109. In some embodiments, a scorpion comprises a binding domain comprising a sequence selected from the group consisting of any of SEQ ID NOS: 332-345. In some embodiments, a scorpion comprises a binding domain comprising a sequence derived from immunoglobulin V_L and V_H domains, wherein the sequence is selected from the group consisting of any of SEQ ID NOS: 355-365. The invention further contemplates scorpions comprising a binding domain that has the opposite orientation of V_L and V_H having sequences deducible from any of SEQ ID NOS:355-365.

For embodiments in which either, or both, of the binding domains are derived from more than one region of an immunoglobulin (e.g., an Ig V_L region and an Ig V_H region), the plurality of regions may be joined by a linker peptide. Moreover, a linker may be used to join the first binding domain to a constant sub-region. Joinder of the constant sub-region to a second binding domain (i.e., binding domain 2 disposed towards the C-terminus of a scorpion) is accomplished by a scorpion linker. These scorpion linkers are preferably between about 2-45 amino acids, or 2-38 amino acids, or 5-45 amino acids. For example, the H1 linker is 2 amino acids in length and the STD2 linker is 38 amino acids in length. Beyond general length considerations, a scorpion linker region suitable for use in the scorpions according to the invention includes an antibody hinge region selected from the group consisting of IgG, IgA, IgD and IgE hinges and variants thereof. For example, the scorpion linker may be an antibody hinge region selected from the group consisting of human IgG1, human

IgG2, human IgG3, and human IgG4, and variants thereof. In some embodiments, the scorpion linker region has a single cysteine residue for formation of an interchain disulfide bond. In other embodiments, the scorpion linker has two cysteine residues for formation of interchain disulfide bonds. In some embodiments, a scorpion linker region is derived from an immunoglobulin hinge region or a C-lectin stalk region and comprises a sequence selected from the group consisting of SEQ ID NOS:111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 287, 289, 297, 305, 307, 309, 310, 311, 313, 314, 315, 316,317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 346, 351, 352, 353, 354, 373, 374, 375, 376 and 377. More generally, any sequence of amino acids identified in the sequence listing as providing a sequence derived from a hinge region is contemplated for use as a scorpion linker in the scorpion molecules according to the invention. In addition, a scorpion linker derived from an Ig hinge is a hinge-like peptide domain having at least one free cysteine capable of participating in an interchain disulfide bond. Preferably, a scorpion linker derived from an Ig hinge peptide retains a cysteine that corresponds to the hinge cysteine disposed towards the N-terminus of that hinge. Preferably, a scorpion linker derived from an IgG1 hinge has one cysteine or has two cysteines corresponding to hinge cysteines. Additionally, a scorpion linker is a stalk region of a Type II C-lectin molecule. In some embodiments, a scorpion comprises a scorpion linker having a sequence selected from the group consisting of SEQ ID NOS:373-377.

The centrally disposed constant sub-region is derived from a constant region of an immunoglobulin protein. The constant sub-region generally is derived from a C_{H2} portion of a C_H region of an immunoglobulin in the abstract, although it may be derived from a C_{H2}-C_{H3} portion. Optionally, the constant sub-region may be derived from a hinge-C_{H2} or hinge-C_{H2}-C_{H3} portion of an immunoglobulin, placing a peptide corresponding to an Ig hinge region N-terminal to the constant sub-region and disposed between the constant sub-region and binding domain 1. Also, portions of the constant sub-region may be derived from the C_H regions of different immunoglobulins. Further, the peptide corresponding to an Ig CH3 may be truncated, leaving a C-terminal amino acid sequence selected from the group consisting of SEQ ID NOS:366-371. It is preferred, however, that in embodiments in which a scorpion

hinge is a hinge-like peptide derived from an immunoglobulin hinge, that the scorpion linker and the constant sub-region be derived from the same type of immunoglobulin. The constant sub-region provides at least one activity associated with a C_H region of an immunoglobulin, such as antibody-dependent cell-mediated cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC), protein A binding, binding to at least one F_C receptor, reproducibly detectable stability relative to a protein according to the invention except for the absence of a constant sub-region, and perhaps placental transfer where generational transfer of a molecule according to the invention would be advantageous, as recognized by one of skill in the art. As with the above-described binding domains, the constant sub-region is derived from at least one immunoglobulin molecule and exhibits an identical or substantially identical amino acid sequence to a region or regions of at least one immunoglobulin. In some embodiments, the constant sub-region is modified from the sequence or sequences of at least one immunoglobulin (by substitution of one or more non-native conventional or unconventional, e.g., synthetic, amino acids or imino acids), resulting in a primary structure that may yield an altered secondary or higher order structure with altered properties associated therewith, or may lead to alterations in post-translational processing, such as glycosylation.

For those binding domains and constant sub-regions exhibiting an identical or substantially identical amino acid sequence to one or more immunoglobulin polypeptides, the post-translational modifications of the molecule according to the invention may result in a molecule modified relative to the immunoglobulin(s) serving as a basis for modification. For example, using techniques known in the art, a host cell may be modified, e.g. a CHO cell, in a manner that leads to an altered polypeptide glycosylation pattern relative to that polypeptide in an unmodified (e.g., CHO) host cell.

Provided with such molecules, and the methods of recombinantly producing them *in vivo*, new avenues of targeted diagnostics and therapeutics have been opened to allow, e.g., for the targeted recruitment of effector cells of the immune system (e.g., cytotoxic T lymphocytes, natural killer cells, and the like) to cells, tissues, agents and foreign objects to be destroyed or sequestered, such as cancer cells and infectious agents. In addition to localizing therapeutic cells to a site of treatment, the peptides

are useful in localizing therapeutic compounds, such as radiolabeled proteins. Further, the peptides are also useful in scavenging deleterious compositions, for example by associating a deleterious composition, such as a toxin, with a cell capable of destroying or eliminating that toxin (e.g., a macrophage). The molecules of the invention are useful in modulating the activity of binding partner molecules, such as cell surface receptors. This is shown in Figure 17 where apoptotic signaling through CD20 and/or CD37 is markedly enhanced by a molecule of the present invention. The effect of this signaling is the death of the targeted cell. Diseases and conditions where the elimination of defined cell populations is beneficial would include infectious and parasitic diseases, inflammatory and autoimmune conditions, malignancies, and the like. One skilled in the art would recognize that there is no limitation of the approach to the enhancement of apoptotic signaling. Mitotic signaling and signaling leading to differentiation, activation, or inactivation of defined cell populations can be induced by molecules of the present invention through the appropriate selection of binding partner molecules. Further consideration of the disclosure of the invention will be facilitated by a consideration of the following express definitions of terms used herein.

A "single-chain binding protein" is a single contiguous arrangement of covalently linked amino acids, with the chain capable of specifically binding to one or more binding partners sharing sufficient determinants of a binding site to be detectably bound by the single-chain binding protein. Exemplary binding partners include proteins, carbohydrates, lipids and small molecules.

For ease of exposition, "derivatives" and "variants" of proteins, polypeptides, and peptides according to the invention are described in terms of differences from proteins and/or polypeptides and/or peptides according to the invention, meaning that the derivatives and variants, which are proteins/polypeptides/peptides according to the invention, differ from underivatized or non-variant proteins, polypeptides or peptides of the invention in the manner defined. One of skill in the art would understand that the derivatives and variants themselves are proteins, polypeptides and peptides according to the invention.

An "antibody" is given the broadest definition consistent with its meaning in the art, and includes proteins, polypeptides and peptides capable of binding to at least

one binding partner, such as a proteinaceous or non-proteinaceous antigen. An "antibody" as used herein includes members of the immunoglobulin superfamily of proteins, of any species, of single- or multiple-chain composition, and variants, analogs, derivatives and fragments of such molecules. Specifically, an "antibody" includes any form of antibody known in the art, including but not limited to, monoclonal and polyclonal antibodies, chimeric antibodies, CDR-grafted antibodies, humanized antibodies, single-chain variable fragments, bi-specific antibodies, diabodies, antibody fusions, and the like.

A "binding domain" is a peptide region, such as a fragment of a polypeptide derived from an immunoglobulin (e.g., an antibody), that specifically binds one or more specific binding partners. If a plurality of binding partners exists, those partners share binding determinants sufficient to detectably bind to the binding domain. Preferably, the binding domain is a contiguous sequence of amino acids.

An "epitope" is given its ordinary meaning herein of a single antigenic site, i.e., an antigenic determinant, on a substance (e.g., a protein) with which an antibody specifically interacts, for example by binding. Other terms that have acquired well-settled meanings in the immunoglobulin (e.g., antibody) art, such as a "variable light region," variable heavy region," "constant light region," constant heavy region," "antibody hinge region," "complementarity determining region," "framework region," "antibody isotype," "F_C region," "single-chain variable fragment" or "scFv," "diabody," "chimera," "CDR-grafted antibody," "humanized antibody," "shaped antibody," "antibody fusion," and the like, are each given those well-settled meanings known in the art, unless otherwise expressly noted herein.

Terms understood by those in the art as referring to antibody technology are each given the meaning acquired in the art, unless expressly defined herein. Examples of such terms are "V_L" and "V_H", referring to the variable binding region derived from an antibody light and heavy chain, respectively; and C_L and C_H, referring to an "immunoglobulin constant region," i.e., a constant region derived from an antibody light or heavy chain, respectively, with the latter region understood to be further divisible into C_{H1}, C_{H2}, C_{H3} and C_{H4} constant region domains, depending on the antibody isotype (IgA, IgD, IgE, IgG, IgM) from which the region was derived. CDR means "complementarity determining region." A "hinge region" is derived

from the amino acid sequence interposed between, and connecting, the C_{H1} and C_{H2} regions of a single chain of an antibody, which is known in the art as providing flexibility, in the form of a "hinge," to whole antibodies.

A "constant sub-region" is a term defined herein to refer to a peptide, polypeptide, or protein sequence that corresponds to, or is derived from, one or more constant region domains of an antibody. Thus, a constant sub-region may include any or all of the following domains: a C_{H1} domain, a hinge region, a C_{H2} domain, a C_{H3} domain (IgA, IgD, IgG, IgE, and IgM), and a C_{H4} domain (IgE, IgM). A constant sub-region as defined herein, therefore, can refer to a polypeptide region corresponding to an entire constant region of an antibody, or a portion thereof. Typically, a constant sub-region of a polypeptide, or encoding nucleic acid, of the invention has a hinge, C_{H2} domain, and C_{H3} domain.

An "effector function" is a function associated with or provided by a constant region of an antibody. Exemplary effector functions include antibody-dependent cell-mediated cytotoxicity (ADCC), complement activation and complement-dependent cytotoxicity (CDC), F_C receptor binding, and increased plasma half-life, as well as placental transfer. An effector function of a composition according to the invention is detectable; preferably, the specific activity of the composition according to the invention for that function is about the same as the specific activity of a wild-type antibody with respect to that effector function, i.e., the constant sub-region of the multivalent binding molecule preferably has not lost any effector function relative to a wild-type antibody]

A "linker" is a peptide, or polynucleotide, that joins or links other peptides or polynucleotides. Typically, a peptide linker is an oligopeptide of from about 2-50 amino acids, with typical polynucleotide linkers encoding such a peptide linker and, thus, being about 6-150 nucleotides in length. Linkers join the first binding domain to a constant sub-region domain. An exemplary peptide linker is $(Gly_4Ser)_3$. A scorpion linker is used to join the C-terminal end of a constant sub-region to a second binding domain. The scorpion linker may be derived from an immunoglobulin hinge region or from the stalk region of a type II C-lectin, as described in greater detail below.

A "target" is given more than one meaning, with the context of usage defining an unambiguous meaning in each instance. In its narrowest sense, a "target" is a binding site, i.e., the binding domain of a binding partner for a peptide composition according to the invention. In a broader sense, "target" or "molecular target" refers to the entire binding partner (e.g., a protein), which necessarily exhibits the binding site. Specific targets, such as "CD20," "CD37," and the like, are each given the ordinary meaning the term has acquired in the art. A "target cell" is any prokaryotic or eukaryotic cell, whether healthy or diseased, that is associated with a target molecule according to the invention. Of course, target molecules are also found unassociated with any cell (i.e., a cell-free target) or in association with other compositions such as viruses (including bacteriophage), organic or inorganic target molecule carriers, and foreign objects.

Examples of materials with which a target molecule may be associated include autologous cells (e.g., cancer cells or other diseased cells), infectious agents (e.g., infectious cells and infectious viruses), and the like. A target molecule may be associated with an enucleated cell, a cell membrane, a liposome, a sponge, a gel, a capsule, a tablet, and the like, which may be used to deliver, transport or localize a target molecule, regardless of intended use (e.g., for medical treatment, as a result of benign or unintentional provision, or to further a bioterrorist threat). "Cell-free," "virus-free," "carrier-free," "object-free," and the like refer to target molecules that are not associated with the specified composition or material.

"Binding affinity" refers to the strength of non-covalent binding of the peptide compositions of the invention and their binding partners. Preferably, binding affinity refers to a quantitative measure of the attraction between members of a binding pair.

An "adjuvant" is a substance that increases or aids the functional effect of a compound with which it is in association, such as in the form of a pharmaceutical composition comprising an active agent and an adjuvant. An "excipient" is an inert substance used as a diluent in formulating a pharmaceutical composition. A "carrier" is a typically inert substance used to provide a vehicle for delivering a pharmaceutical composition.

"Host cell" refers to any cell, prokaryotic or eukaryotic, in which is found a polynucleotide, protein or peptide according to the invention.

"Introducing" a nucleic acid or polynucleotide into a host cell means providing for entry of the nucleic acid or polynucleotide into that cell by any means known in the art, including but not limited to, *in vitro* salt-mediated precipitations and other forms of transformation of naked nucleic acid/polynucleotide or vector-borne nucleic acid/polynucleotide, virus-mediated infection and optionally transduction, with or without a "helper" molecule, ballistic projectile delivery, conjugation, and the like.

"Incubating" a host cell means maintaining that cell under environmental conditions known in the art to be suitable for a given purpose, such as gene expression. Such conditions, including temperature, ionic strength, oxygen tension, carbon dioxide concentration, nutrient composition, and the like, are well known in the art.

"Isolating" a compound, such as a protein or peptide according to the invention, means separating that compound from at least one distinct compound with which it is found associated in nature, such as in a host cell expressing the compound to be isolated, e.g. by isolating spent culture medium containing the compound from the host cells grown in that medium.

An "organism in need" is any organism at risk of, or suffering from, any disease, disorder or condition that is amenable to treatment or amelioration with a composition according to the invention, including but not limited to any of various forms of cancer, any of a number of autoimmune diseases, radiation poisoning due to radiolabeled proteins, peptides and like compounds, ingested or internally produced toxins, and the like, as will become apparent upon review of the entire disclosure. Preferably, an organism in need is a human patient.

"Ameliorating" a symptom of a disease means detectably reducing the severity of that symptom of disease, as would be known in the art. Exemplary symptoms include pain, heat, swelling and joint stiffness.

Unless clear from context, the terms "protein," "peptide," and "polypeptide" are used interchangeably herein, with each referring to at least one contiguous chain of amino acids. Analogously, the terms "polynucleotide," "nucleic acid," and "nucleic acid molecule" are used interchangeably unless it is clear from context that a particular, and non-interchangeable, meaning is intended.

"Pharmaceutically acceptable salt" refers to salts of the compounds of the present invention derived from the combination of such compounds and an organic or inorganic acid (acid addition salts) or an organic or inorganic base (base addition salts).

Using the terms as defined above, a general description of the various aspects of the invention is provided below. Following the general description, working examples are presented to provide supplementary evidence of the operability and usefulness of the invention disclosed herein.

Proteins and polypeptides

In certain embodiments of the invention, there are provided any of the herein-described multivalent binding proteins with effector function, including binding domain-immunoglobulin fusion proteins, wherein the multivalent binding protein or peptide with effector function comprises two or more binding domain polypeptide sequences. Each of the binding domain polypeptide sequences is capable of binding or specifically binding to a target(s), such as an antigen(s), which target(s) or antigen(s) may be the same or may be different. The binding domain polypeptide sequence may be derived from an antigen variable region or it may be derived from immunoglobulin-like molecules, e.g., receptors that fold in ways that mimic immunoglobulin molecules. The antibodies from which the binding domains are derived may be antibodies that are polyclonal, including monospecific polyclonal, monoclonal (mAbs), recombinant, chimeric, humanized (such as CDR-grafted), human, single-chain, catalytic, and any other form of antibody known in the art, as well as fragments, variants or derivatives thereof. In some embodiments, each of the binding domains of the protein according to the invention is derived from a complete variable region of an immunoglobulin. In preferred embodiments, the binding domains are each based on a human Ig variable region. In other embodiments, the protein is derived from a fragment of an Ig variable region. In such embodiments, it

is preferred that each binding domain polypeptide sequence correspond to the sequences of each of the complementarity determining regions of a given Ig variable region. Also contemplated within the invention are binding domains that correspond to fewer than all CDRs of a given Ig variable region, provided that such binding domains retain the capacity to specifically bind to at least one target.

The multivalent binding protein with effector function also has a constant sub-region sequence derived from an immunoglobulin constant region, preferably an antibody heavy chain constant region, covalently juxtaposed between the two binding domains in the multivalent binding protein with effector function.

The multivalent binding protein with effector function also has a scorpion linker that joins the C-terminal end of the constant sub-region to the N-terminal end of binding domain 2. The scorpion linker is not a helical peptide and may be derived from an antibody hinge region, from a region connecting binding domains of an immunoglobulin, or from the stalk region of type II C-lectins. The scorpion linker may be derived from a wild-type hinge region of an immunoglobulin, such as an IgG1, IgG2, IgG3, IgG4, IgA, IgD or an IgE hinge region. In other embodiments, the invention provides multivalent binding proteins with altered hinges. One category of altered hinge regions suitable for inclusion in the multivalent binding proteins is the category of hinges with an altered number of Cysteine residues, particularly those Cys residues known in the art to be involved in interchain disulfide bond formation in immunoglobulin counterpart molecules having wild-type hinges. Thus, proteins may have an IgG1 hinge in which one of the three Cys residues capable of participating in interchain disulfide bond formations is missing. To indicate the Cysteine sub-structure of altered hinges, the Cys subsequence is presented from N- to C-terminus. Using this identification system, the multivalent binding proteins with altered IgG hinges include hinge structures characterized as cxc, xxc, ccx, xxc, xcx, cxx, and xxx. The Cys residue may be either deleted or substituted by an amino acid that results in a conservative substitution or a non-conservative substitution. In some embodiments, the Cysteine is replaced by a Serine. For proteins with scorpion linkers comprising IgG1 hinges, the number of cysteines corresponding to hinge cysteines is reduced to 1 or 2, preferably with one of those cysteines corresponding to the hinge cysteine disposed closest to the N-terminus of the hinge.

For proteins with scorpion linkers comprising IgG2 hinges, there may be 0, 1, 2, 3, or 4 Cys residues. Forscorpion linkers comprising altered IgG2 hinges containing 1, 2 or 3 Cys residues, all possible subsets of Cys residues are contemplated. Thus, for such linkers having one Cys, the multivalent binding proteins may have the following Cys motif in the hinge region: cxxx, xcxx, xxcx, or xxxc. For scorpion linkers comprising IgG2 hinge variants having 2 or 3 Cys residues, all possible combinations of retained and substituted (or deleted) Cys residues are contemplated. For multivalent binding proteins with scorpion linkers comprising altered IgG3 or altered IgG4 hinge regions, a reduction in Cys residues from 1 to one less than the complete number of Cys residues in the hinge region is contemplated, regardless of whether the loss is through deletion or substitution by conservative or non-conservative amino acids (e.g., Serine). In like manner, multivalent binding proteins having a scorpion linker comprising a wild-type IgA, IgD or IgE hinge are contemplated, as are corresponding altered hinge regions having a reduced number of Cys residues extending from 0 to one less than the total number of Cys residues found in the corresponding wild-type hinge. In some embodiments having an IgG1 hinge, the first, or N-terminal, Cys residue of the hinge is retained. For proteins with either wild-type or altered hinge regions, it is contemplated that the multivalent binding proteins will be single-chain molecules capable of forming homo-multimers, such as dimers, e.g., by disulfide bond formation. Further, proteins with altered hinges may have alterations at the termini of the hinge region, e.g., loss or substitution of one or more amino acid residues at the N-terminus, C-terminus or both termini of a given region or domain, such as a hinge domain, as disclosed herein.

In another exemplary embodiment, the constant sub-region is derived from a constant region that comprises a native, or an engineered, IgD hinge region. The wild-type human IgD hinge has one cysteine that forms a disulfide bond with the light chain in the native IgD structure. In some embodiments, this IgD hinge cysteine is mutated (e.g., deleted) to generate an altered hinge for use as a connecting region between binding domains of, for example, a bispecific molecule. Other amino acid changes or deletions or alterations in an IgD hinge that do not result in undesired hinge inflexibility are within the scope of the invention. Native or engineered IgD hinge regions from other species are also within the scope of the invention, as are humanized native or engineered IgD hinges from non-human species, and (other non

IgD) hinge regions from other human, or non-human, antibody isotypes, (such as the llama IgG2 hinge).

The invention further comprehends constant sub-regions attached to scorpion linkers that may be derived from hinges that correspond to a known hinge region, such as an IgG1 hinge or an IgD hinge, as noted above. The constant sub-region may contain a modified or altered (relative to wild-type) hinge region in which at least one cysteine residue known to participate in inter-chain disulfide bond linkage is replaced by another amino acid in a conservative substitution (e.g., Ser for Cys) or a non-conservative substitution. The constant sub-region does not include a peptide region or domain that corresponds to an immunoglobulin C_{H1} domain.

Alternative hinge and linker sequences that can be used as connecting regions are from portions of cell surface receptors that connect immunoglobulin V-like or immunoglobulin C-like domains. Regions between Ig V-like domains where the cell surface receptor contains multiple Ig V-like domains in tandem, and between Ig C-like domains where the cell surface receptor contains multiple tandem Ig C-like regions are also contemplated as connecting regions. Hinge and linker sequences are typically from 5 to 60 amino acids long, and may be primarily flexible, but may also provide more rigid characteristics. In addition, linkers frequently provide spacing that facilitates minimization of steric hindrance between the binding domains. Preferably, these hinge and linker peptides are primarily α helical in structure, with minimal β sheet structure. The preferred sequences are stable in plasma and serum and are resistant to proteolytic cleavage. The preferred sequences may contain a naturally occurring or added motif such as the CPPC motif that confers a disulfide bond to stabilize dimer formation. The preferred sequences may contain one or more glycosylation sites. Examples of preferred hinge and linker sequences include, but are not limited to, the interdomain regions between the Ig V-like and Ig C-like regions of CD2, CD4, CD22, CD33, CD48, CD58, CD66, CD80, CD86, CD150, CD166, and CD244.

The constant sub-region may be derived from a camelid constant region, such as either a llama or camel IgG2 or IgG3.

Specifically contemplated is a constant sub-region having the C_{H2}-C_{H3} region from any Ig class, or from any IgG subclass, such as IgG1 (e.g., human IgG1). In preferred embodiments, the constant sub-region and the scorpion linker derived from an immunoglobulin hinge are both derived from the same Ig class. In other preferred embodiments, the constant sub-region and the scorpion linker derived from an immunoglobulin hinge are both derived from the same Ig sub-class. The constant sub-region also may be a CH3 domain from any Ig class or subclass, such as IgG1 (e.g., human IgG1), provided that it is associated with at least one immunoglobulin effector function.

The constant sub-region does not correspond to a complete immunoglobulin constant region (i.e., C_{H1}-hinge-C_{H2}-C_{H3}) of the IgG class. The constant sub-region may correspond to a complete immunoglobulin constant region of other classes., IgA constant domains, such as an IgA1 hinge, an IgA2 hinge, an IgA C_{H2} and an IgA C_{H3} domains with a mutated or missing tailpiece are also contemplated as constant sub-regions. Further, any light chain constant domain may function as a constant sub-region, e.g., C_K or any C_L. The constant sub-region may also include JH or JK, with or without a hinge. The constant sub-region may also correspond to engineered antibodies in which, e.g., a loop graft has been constructed by making selected amino acid substitutions using an IgG framework to generate a binding site for a receptor other than a natural F_CR (CD16, CD32, CD64, $F_C\varepsilon R1$), as would be understood in the art. An exemplary constant sub-region of this type is an IgG C_{H2}-C_{H3} region modified to have a CD89 binding site.

This aspect of the invention provides a multivalent binding protein or peptide having effector function, comprising, consisting essentially of, or consisting of (a) an N-terminally disposed binding domain polypeptide sequence derived from an immunoglobulin that is fused or otherwise connected to (b) a constant sub-region polypeptide sequence derived from an immunoglobulin constant region, which preferably includes a hinge region sequence, wherein the hinge region polypeptide may be as described herein, and may comprise, consist essentially of, or consist of, for example, an alternative hinge region polypeptide sequence, in turn fused or otherwise connected to (c) a C-terminally disposed second native or engineered binding domain polypeptide sequence derived from an immunoglobulin.

The centrally disposed constant sub-region polypeptide sequence derived from an immunoglobulin constant region is capable of at least one immunological activity selected from the group consisting of antibody dependent cell-mediated cytotoxicity, CDC, complement fixation, and F_C receptor binding, and the binding domain polypeptides are each capable of binding or specifically binding to a target, such as an antigen, wherein the targets may be the same or different, and may be found in effectively the same physiological environment (e.g., the surface of the same cell) or in different environments (e.g., different cell surfaces, a cell surface and a cell-free location, such as in solution).

This aspect of the invention also comprehends variant proteins or polypeptides exhibiting an effector function that are at least 80%, and preferably 85%, 90%, 95% or 99% identical to a multivalent protein with effector function of specific sequence as disclosed herein.

Polynucleotides

The invention also provides polynucleotides (isolated or purified or pure polynucleotides) encoding the proteins or peptides according to the invention, vectors (including cloning vectors and expression vectors) comprising such polynucleotides, and cells (e.g., host cells) transformed or transfected with a polynucleotide or vector according to the invention. In encoding the proteins or polypeptides of the invention, the polynucleotides encode a first binding domain, a second binding domain and an F_C domain, all derived from immunoglobulins, preferably human immunoglobulins. Each binding domain may contain a sequence corresponding to a full-length variable region sequence (either heavy chain and/or light chain), or to a partial sequence thereof, provided that each such binding domain retains the capacity to specifically bind. The F_C domain may have a sequence that corresponds to a full-length immunoglobulin F_C domain sequence or to a partial sequence thereof, provided that the F_C domain exhibits at least one effector function as defined herein. In addition, each of the binding domains may be joined to the F_C domain via a linker peptide that typically is at least 8, and preferably at least 13, amino acids in length. A preferred linker sequence is a sequence based on the Gly_4Ser motif, such as $(Gly_4Ser)_3$.

Variants of the multivalent binding protein with effector function are also comprehended by the invention. Variant polynucleotides are at least 90%, and

preferably 95%, 99%, or 99.9% identical to one of the polynucleotides of defined sequence as described herein, or that hybridizes to one of those polynucleotides of defined sequence under stringent hybridization conditions of 0.015 M sodium chloride, 0.0015 M sodium citrate at 65-68°C or 0.015 M sodium chloride, 0.0015M sodium citrate, and 50% formamide at 42°C. The polynucleotide variants retain the capacity to encode a multivalent binding protein with effector function.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Hybridization stringency is principally determined by temperature, ionic strength, and the concentration of denaturing agents such as formamide. Examples of stringent conditions for hybridization and washing are 0.015 M sodium chloride, 0.0015 M sodium citrate at 65-68°C or 0.015 M sodium chloride, 0.0015M sodium citrate, and 50% formamide at 42°C. See Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, (Cold Spring Harbor, N.Y. 1989).

More stringent conditions (such as higher temperature, lower ionic strength, higher formamide, or other denaturing agent) may also be used; however, the rate of hybridization will be affected. In instances wherein hybridization of deoxyoligonucleotides is concerned, additional exemplary stringent hybridization conditions include washing in 6x SSC, 0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligonucleotides), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

In a related aspect of the invention, there is provided a method of producing a polypeptide or protein or other construct of the invention, for example, including a multivalent binding protein or peptide having effector function, comprising the steps of (a) culturing a host cell as described or provided for herein under conditions that permit expression of the construct; and (b) isolating the expression product, for example, the multivalent binding protein or peptide with effector function from the host cell or host cell culture.

Constructs

The present invention also relates to vectors, and to constructs prepared from known vectors, that each include a polynucleotide or nucleic acid of the invention,

and in particular to recombinant expression constructs, including any of various known constructs, including delivery constructs, useful for gene therapy, that include any nucleic acids encoding multivalent, for example, multispecific, including bi-specific, binding proteins and polypeptides with effector function, as provided herein; to host cells which are genetically engineered with vectors and/or other constructs of the invention and to methods of administering expression or other constructs comprising nucleic acid sequences encoding multivalent, for example, multispecific, including bi-specific, binding proteins with effector function, or fragments or variants thereof, by recombinant techniques.

Various constructs of the invention including multivalent, for example, multispecific binding proteins with effector function, can be expressed in virtually any host cell, including *in vivo* host cells in the case of use for gene therapy, under the control of appropriate promoters, depending on the nature of the construct (*e.g.,* type of promoter, as described above), and on the nature of the desired host cell (*e.g.,* postmitotic terminally differentiated or actively dividing; *e.g.,* maintenance of an expressible construct as an episome or integrated into the host cell genome).

Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described, for example, in Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, NY, (1989). Exemplary cloning/expression vectors include, but are not limited to, cloning vectors, shuttle vectors, and expression constructs, that may be based on plasmids, phagemids, phasmids, cosmids, viruses, artificial chromosomes, or any nucleic acid vehicle suitable for amplification, transfer, and/or expression of a polynucleotide contained therein that is known in the art. As noted herein, in preferred embodiments of the invention, recombinant expression is conducted in mammalian cells that have been transfected, transformed or transduced with a nucleic acid according to the invention. *See also, for example,* Machida, CA., "Viral Vectors for Gene Therapy: Methods and Protocols"; Wolff, JA, "Gene Therapeutics: Methods and Applications of Direct Gene Transfer" (Birkhauser 1994); Stein, U and Walther, W (eds., "Gene Therapy of Cancer: Methods and Protocols" (Humana Press 2000); Robbins, PD (ed.), "Gene Therapy Protocols" (Humana Press 1997); Morgan, JR (ed.), "Gene Therapy Protocols" (Humana Press 2002); Meager, A (ed.), "Gene Therapy Technologies,

Applications and Regulations: From Laboratory to Clinic" (John Wiley & Sons Inc. 1999); MacHida, CA and Constant, JG, "Viral Vectors for Gene Therapy: Methods and Protocols" (Humana Press 2002);"New Methods Of Gene Therapy For Genetic Metabolic Diseases NIH Guide," Volume 22, Number 35, October 1, 1993. *See also* U.S. Pat. Nos. 6,384,210; 6,384,203; 6,384,202; 6,384,018; 6,383,814; 6,383,811; 6,383,795; 6,383,794; 6,383,785; 6,383,753; 6,383,746; 6,383,743; 6,383,738; 6,383,737; 6,383,733; 6,383,522; 6,383,512; 6,383,481; 6,383,478; 6,383,138; 6,380,382; 6,380,371; 6,380,369; 6,380,362; 6,380,170; 6,380,169; 6,379,967; and 6,379,966.

Typically, expression constructs are derived from plasmid vectors. One preferred construct is a modified pNASS vector (Clontech, Palo Alto, CA), which has nucleic acid sequences encoding an ampicillin resistance gene, a polyadenylation signal and a T7 promoter site. Other suitable mammalian expression vectors are well known (*see, e.g.*, Ausubel *et al.*, 1995; Sambrook *et al.*, *supra*; *see also, e.g.,* catalogues from Invitrogen, San Diego, CA; Novagen, Madison, WI; Pharmacia, Piscataway, NJ). Presently preferred constructs may be prepared that include a dihydrofolate reductase (DHFR)-encoding sequence under suitable regulatory control, for promoting enhanced production levels of the multivalent binding protein with effector function, which levels result from gene amplification following application of an appropriate selection agent (*e.g.*, methotrexate).

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence, as described above. A vector in operable linkage with a polynucleotide according to the invention yields a cloning or expression construct. Exemplary cloning/expression constructs contain at least one expression control element, e.g., a promoter, operably linked to a polynucleotide of the invention. Additional expression control elements, such as enhancers, factor-specific binding sites, terminators, and ribosome binding sites are also contemplated in the vectors and cloning/expression constructs according to the invention. The heterologous structural sequence of the polynucleotide according to the invention is assembled in appropriate phase with translation initiation and termination sequences. Thus, for example, the

multivalent binding protein-encoding nucleic acids as provided herein may be included in any one of a variety of expression vector constructs as a recombinant expression construct for expressing such a protein in a host cell. In certain preferred embodiments the constructs, are included in formulations that are administered *in vivo*. Such vectors and constructs include chromosomal, nonchromosomal and synthetic DNA sequences, *e.g.*, derivatives of SV40; bacterial plasmids; phage DNA; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA, such as vaccinia, adenovirus, fowl pox virus, and pseudorabies, or replication deficient retroviruses as described below. However, any other vector may be used for preparation of a recombinant expression construct, and in preferred embodiments such a vector will be replicable and viable in the host.

The appropriate DNA sequence(s) may be inserted into a vector, for example, by a variety of procedures. In general, a DNA sequence is inserted into an appropriate restriction endonuclease cleavage site(s) by procedures known in the art. Standard techniques for cloning, DNA isolation, amplification and purification, for enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like, and various separation techniques are contemplated. A number of standard techniques are described, for example, in Ausubel *et al.* (1993 *Current Protocols in Molecular Biology*, Greene Publ. Assoc. Inc. & John Wiley & Sons, Inc., Boston, MA); Sambrook *et al.* (1989 *Molecular Cloning*, Second Ed., Cold Spring Harbor Laboratory, Plainview, NY); Maniatis *et al.* (1982 Molecular Cloning, Cold Spring Harbor Laboratory, Plainview, NY); Glover (Ed.) (1985 *DNA Cloning Vol. I and II*, IRL Press, Oxford, UK); Hames and Higgins (Eds.), (1985 *Nucleic Acid Hybridization*, IRL Press, Oxford, UK); and elsewhere.

The DNA sequence in the expression vector is operatively linked to at least one appropriate expression control sequence (*e.g.,* a constitutive promoter or a regulated promoter) to direct mRNA synthesis. Representative examples of such expression control sequences include promoters of eukaryotic cells or their viruses, as described above. Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the

appropriate vector and promoter is well within the level of ordinary skill in the art, and preparation of certain particularly preferred recombinant expression constructs comprising at least one promoter or regulated promoter operably linked to a nucleic acid encoding a protein or polypeptide according to the invention is described herein.

Transcription of the DNA encoding proteins and polypeptides of the invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Gene therapies using the nucleic acids of the invention are also contemplated, comprising strategies to replace defective genes or add new genes to cells and/or tissues, and is being developed for application in the treatment of cancer, the correction of metabolic disorders and in the field of immunotherapy. Gene therapies of the invention include the use of various constructs of the invention, with or without a separate carrier or delivery vehicle or constructs, for treatment of the diseases, disorders, and/or conditions noted herein. Such constructs may also be used as vaccines for treatment or prevention of the diseases, disorders, and/or conditions noted herein. DNA vaccines, for example, make use of polynucleotides encoding immunogenic protein and nucleic acid determinants to stimulate the immune system against pathogens or tumor cells. Such strategies can stimulate either acquired or innate immunity or can involve the modification of immune function through cytokine expression. *In vivo* gene therapy involves the direct injection of genetic material into a patient or animal, typically to treat, prevent or ameliorate a disease or symptoms associated with a disease. Vaccines and immune modulation are systemic therapies. With tissue-specific *in vivo* therapies, such as those that aim to treat cancer, localized gene delivery and/or expression/targeting systems are preferred. Diverse gene therapy vectors that target specific tissues are known in the art, and procedures have been developed to physically target specific tissues, for example, using catheter-based technologies, all of which are contemplated herein.

Ex vivo approaches to gene therapy are also contemplated herein and involve the removal, genetic modification, expansion and re-administration of a subject's, e.g., human patient's, own cells. Examples include bone marrow transplantation for

cancer treatment or the genetic modification of lymphoid progenitor cells. *Ex vivo* gene therapy is preferably applied to the treatment of cells that are easily accessible and can survive in culture during the gene transfer process (such as blood or skin cells).

Useful gene therapy vectors include adenoviral vectors, lentiviral vectors, Adeno-associated virus (AAV) vectors, Herpes Simplex Virus (HSV) vectors, and retroviral vectors. Gene therapies may also be carried out using "naked DNA," liposome-based delivery, lipid-based delivery (including DNA attached to positively charged lipids), electroporation, and ballistic projection.

In certain embodiments, including but not limited to gene therapy embodiments, the vector may be a viral vector such as, for example, a retroviral vector. Miller *et al*., 1989 *BioTechniques* 7:980; Coffin and Varmus, 1996 Retroviruses, *Cold Spring Harbor Laboratory Press*, NY. For example, retroviruses from which the retroviral plasmid vectors may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

Retroviruses are RNA viruses which can replicate and integrate into the genome of a host cell via a DNA intermediate. This DNA intermediate, or provirus, may be stably integrated into the host cell DNA. According to certain embodiments of the present invention, an expression construct may comprise a retrovirus into which a foreign gene that encodes a foreign protein is incorporated in place of normal retroviral RNA. When retroviral RNA enters a host cell coincident with infection, the foreign gene is also introduced into the cell, and may then be integrated into host cell DNA as if it were part of the retroviral genome. Expression of this foreign gene within the host results in expression of the foreign protein.

Most retroviral vector systems that have been developed for gene therapy are based on murine retroviruses. Such retroviruses exist in two forms, as free viral particles referred to as virions, or as proviruses integrated into host cell DNA. The virion form of the virus contains the structural and enzymatic proteins of the

retrovirus (including the enzyme reverse transcriptase), two RNA copies of the viral genome, and portions of the source cell plasma membrane containing viral envelope glycoprotein. The retroviral genome is organized into four main regions: the Long Terminal Repeat (LTR), which contains *cis*-acting elements necessary for the initiation and termination of transcription and is situated both 5' and 3' to the coding genes, and the three genes encoding *gag, pol,* and *env*. These three genes, *gag, pol,* and *env*, encode, respectively, internal viral structures, enzymatic proteins (such as integrase), and the envelope glycoprotein (designated gp70 and p15e) which confers infectivity and host range specificity of the virus, as well as the "R" peptide of undetermined function.

Separate packaging cell lines and vector-producing cell lines have been developed because of safety concerns regarding the uses of retroviruses, including uses in expression constructs. Briefly, this methodology employs the use of two components, a retroviral vector and a packaging cell line (PCL). The retroviral vector contains long terminal repeats (LTRs), the foreign DNA to be transferred and a packaging sequence (y). This retroviral vector will not reproduce by itself because the genes which encode structural and envelope proteins are not included within the vector genome. The PCL contains genes encoding the *gag, pol,* and *env* proteins, but does not contain the packaging signal "y." Thus, a PCL can only form empty virion particles by itself. Within this general method, the retroviral vector is introduced into the PCL, thereby creating a vector-producing cell line (VCL). This VCL manufactures virion particles containing only the foreign genome of the retroviral vector, and therefore has previously been considered to be a safe retrovirus vector for therapeutic use.

A "retroviral vector construct" refers to an assembly which is, within preferred embodiments of the invention, capable of directing the expression of a sequence(s) or gene(s) of interest, such as multivalent binding protein-encoding nucleic acid sequences. Briefly, the retroviral vector construct must include a 5' LTR, a tRNA binding site, a packaging signal, an origin of second strand DNA synthesis and a 3' LTR. A wide variety of heterologous sequences may be included within the vector construct including, for example, sequences which encode a protein (*e.g.,* cytotoxic protein, disease-associated antigen, immune accessory molecule, or replacement

gene), or which are useful as a molecule itself (*e.g.,* as a ribozyme or antisense sequence).

Retroviral vector constructs of the present invention may be readily constructed from a wide variety of retroviruses, including for example, B, C, and D type retroviruses as well as spumaviruses and lentiviruses (see, *e.g.,* RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985). Such retroviruses may be readily obtained from depositories or collections such as the American Type Culture Collection ("ATCC"; Rockville, Maryland), or isolated from known sources using commonly available techniques. Any of the above retroviruses may be readily utilized in order to assemble or construct retroviral vector constructs, packaging cells, or producer cells of the invention, given the disclosure provided herein and standard recombinant techniques (*e.g.*, Sambrook *et al*, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, 1989; Kunkle, 1985 *Proc. Natl. Acad. Sci. (USA) 82*:488).

Suitable promoters for use in viral vectors generally may include, but are not limited to, the retroviral LTR; the SV40 promoter; and the human cytomegalovirus (CMV) promoter described in Miller, *et al*., 1989 *Biotechniques* 7:980-990, or any other promoter (*e.g.*, cellular promoters such as eukaryotic cellular promoters including, but not limited to, the histone, pol III, and β-actin promoters). Other viral promoters that may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein, and may be from among either regulated promoters or promoters as described above.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, ψ-2, ψ-AM, PA12, T19-14X, VT-19-17-H2, ψCRE, ψCRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller, *Human Gene Therapy*, *1*:5-14 (1990). The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and $CaPO_4$ precipitation. In one

alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include the nucleic acid sequence(s) encoding the multivalent binding proteins with effector function. Such retroviral vector particles then may be employed to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the protein or polypeptide. Eukaryotic cells that may be transduced include, but are not limited to, embryonic stem cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, circulating peripheral blood mononuclear and polymorphonuclear cells including myelomonocytic cells, lymphocytes, myoblasts, tissue macrophages, dendritic cells, Kupffer cells, lymphoid and reticuloendothelial cells of the lymph nodes and spleen, keratinocytes, endothelial cells, and bronchial epithelial cells.

Host cells

A further aspect of the invention provides a host cell transformed or transfected with, or otherwise containing, any of the polynucleotides or cloning/expression constructs of the invention. The polynucleotides and cloning/expression constructs are introduced into suitable cells using any method known in the art, including transformation, transfection and transduction. Host cells include the cells of a subject undergoing *ex vivo* cell therapy including, for example, *ex vivo* gene therapy. Eukaryotic host cells contemplated as an aspect of the invention when harboring a polynucleotide, vector, or protein according to the invention include, in addition to a subject's own cells (e.g., a human patient's own cells), VERO cells, HeLa cells, Chinese hamster ovary (CHO) cell lines (including modified CHO cells capable of modifying the glycosylation pattern of expressed multivalent binding molecules, see Published US Patent Application No. 2003/0115614 A1), incorporated herein by reference, COS cells (such as COS-7), W138, BHK, HepG2, 3T3, RIN, MDCK, A549, PC12, K562, HEK293 cells, HepG2 cells, N cells, 3T3 cells, *Spodoptera frugiperda* cells (e.g., Sf9 cells), *Saccharomyces cerevisiae* cells, and any other eukaryotic cell known in the art to be useful in expressing, and optionally isolating, a protein or peptide according to the invention. Also contemplated are prokaryotic cells, including but not limited to, *Escherichia coli*, *Bacillus subtilis*,

Salmonella typhimurium, a Streptomycete, or any prokaryotic cell known in the art to be suitable for expressing, and optionally isolating, a protein or peptide according to the invention. In isolating protein or peptide from prokaryotic cells, in particular, it is contemplated that techniques known in the art for extracting protein from inclusion bodies may be used. The selection of an appropriate host is within the scope of those skilled in the art from the teachings herein.

The engineered host cells can be cultured in a conventional nutrient medium modified as appropriate for activating promoters, selecting transformants, or amplifying particular genes. The culture conditions for particular host cells selected for expression, such as temperature, pH and the like, will be readily apparent to the ordinarily skilled artisan. Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, 1981 *Cell* 23:175, and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and, optionally, enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences, for example as described herein regarding the preparation of multivalent binding protein expression constructs. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Introduction of the construct into the host cell can be effected by a variety of methods with which those skilled in the art will be familiar, including but not limited to, calcium phosphate transfection, DEAE-Dextran-mediated transfection, or electroporation (Davis *et al.*, 1986 *Basic Methods in Molecular Biology*).

In one embodiment, a host cell is transduced by a recombinant viral construct directing the expression of a protein or polypeptide according to the invention. The transduced host cell produces viral particles containing expressed protein or polypeptide derived from portions of a host cell membrane incorporated by the viral particles during viral budding.

Pharmaceutical compositions

In some embodiments, the compositions of the invention, such as a multivalent binding protein or a composition comprising a polynucleotide encoding such a protein as described herein, are suitable to be administered under conditions and for a time sufficient to permit expression of the encoded protein in a host cell *in vivo* or *in vitro*, for gene therapy, and the like. Such compositions may be formulated into pharmaceutical compositions for administration according to well known methodologies. Pharmaceutical compositions generally comprise one or more recombinant expression constructs, and/or expression products of such constructs, in combination with a pharmaceutically acceptable carrier, excipient or diluent. Such carriers will be nontoxic to recipients at the dosages and concentrations employed. For nucleic acid-based formulations, or for formulations comprising expression products according to the invention, about 0.01 μg/kg to about 100 mg/kg body weight will be administered, for example, by the intradermal, subcutaneous, intramuscular or intravenous route, or by any route known in the art to be suitable under a given set of circumstances. A preferred dosage, for example, is about 1 μg/kg to about 1 mg/kg, with about 5 μg/kg to about 200 μg/kg particularly preferred.

It will be evident to those skilled in the art that the number and frequency of administration will be dependent upon the response of the host. Pharmaceutically acceptable carriers for therapeutic use are well known in the pharmaceutical art, and are described, for example, in *Remingtons Pharmaceutical Sciences*, Mack Publishing Co. (A.R. Gennaro edit. 1985). For example, sterile saline and phosphate-buffered saline at physiological pH may be used. Preservatives, stabilizers, dyes and the like may be provided in the pharmaceutical composition. For example, sodium benzoate, sorbic acid and esters of *p*-hydroxybenzoic acid may be added as preservatives. *Id.* at 1449. In addition, antioxidants and suspending agents may be used. *Id.* The compounds of the present invention may be used in either the free base or salt forms, with both forms being considered as being within the scope of the present invention.

The pharmaceutical compositions that contain one or more nucleic acid constructs of the invention, or the proteins corresponding to the products encoded by such nucleic acid constructs, may be in any form which allows for the composition to be administered to a patient. For example, the composition may be in the form of a

solid, liquid or gas (aerosol). Typical routes of administration include, without limitation, oral, topical, parenteral (*e.g.*, sublingually or buccally), sublingual, rectal, vaginal, and intranasal. The term parenteral as used herein includes subcutaneous injections, intravenous, intramuscular, intrasternal, intracavernous, intrathecal, intrameatal, intraurethral injection or infusion techniques. The pharmaceutical composition is formulated so as to allow the active ingredients contained therein to be bioavailable upon administration of the composition to a patient. Compositions that will be administered to a patient take the form of one or more dosage units, where for example, a tablet may be a single dosage unit, and a container of one or more compounds of the invention in aerosol form may hold a plurality of dosage units.

For oral administration, an excipient and/or binder may be present. Examples are sucrose, kaolin, glycerin, starch dextrins, sodium alginate, carboxymethylcellulose and ethyl cellulose. Coloring and/or flavoring agents may be present. A coating shell may be employed.

The composition may be in the form of a liquid, *e.g.*, an elixir, syrup, solution, emulsion or suspension. The liquid may be for oral administration or for delivery by injection, as two examples. When intended for oral administration, preferred compositions contain, in addition to one or more binding domain-immunoglobulin fusion construct or expressed product, one or more of a sweetening agent, preservatives, dye/colorant and flavor enhancer. In a composition intended to be administered by injection, one or more of a surfactant, preservative, wetting agent, dispersing agent, suspending agent, buffer, stabilizer and isotonic agent may be included.

A liquid pharmaceutical composition as used herein, whether in the form of a solution, suspension or other like form, may include one or more of the following adjuvants: sterile diluents such as water for injection, saline solution, preferably physiological saline, Ringer's solution, isotonic sodium chloride, fixed oils such as synthetic mono or digylcerides which may serve as the solvent or suspending medium, polyethylene glycols, glycerin, propylene glycol or other solvents; antibacterial agents such as benzyl alcohol or methyl paraben; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of

tonicity such as sodium chloride or dextrose. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic. Physiological saline is a preferred adjuvant. An injectable pharmaceutical composition is preferably sterile.

It may also be desirable to include other components in the preparation, such as delivery vehicles including, but not limited to, aluminum salts, water-in-oil emulsions, biodegradable oil vehicles, oil-in-water emulsions, biodegradable microcapsules, and liposomes. Examples of immunostimulatory substances (adjuvants) for use in such vehicles include N-acetylmuramyl-L-alanine-D-isoglutamine (MDP), lipopolysaccharides (LPS), glucan, IL-12, GM-CSF, gamma interferon and IL-15.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will vary depending on the mode of administration and whether a sustained release is desired. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.,* polylactic galactide) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109. In this regard, it is preferable that the microsphere be larger than approximately 25 microns.

Pharmaceutical compositions may also contain diluents such as buffers, antioxidants such as ascorbic acid, low molecular weight (less than about 10 residues) polypeptides, proteins, amino acids, carbohydrates (*e.g.*, glucose, sucrose or dextrins), chelating agents (*e.g.*, EDTA), glutathione and other stabilizers and excipients. Neutral buffered saline or saline mixed with nonspecific serum albumin are exemplary appropriate diluents. Preferably, product is formulated as a lyophilizate using appropriate excipient solutions (*e.g.*, sucrose) as diluents.

The pharmaceutical compositions according to the invention also include stabilized proteins and stable liquid pharmaceutical formulations in accordance with technology known in the art, including the technology disclosed in Published US Patent Application No. 2006/0008415 A1, incorporated herein by reference. Such technologies include derivatization of a protein, wherein the protein comprises a thiol group coupled to N-acetyl-L-cysteine, N-ethyl-maleimide, or cysteine.

As described above, the subject invention includes compositions capable of delivering nucleic acid molecules encoding multivalent binding proteins with effector function. Such compositions include recombinant viral vectors, *e.g.,* retroviruses (*see* WO 90/07936, WO 91/02805, WO 93/25234, WO 93/25698, and WO 94/03622), adenovirus (see Berkner, 1988 *Biotechniques* 6:616-627; Li *et al.*, 1993 *Hum. Gene Ther.* 4:403-409; Vincent *et al.*, *Nat. Genet.* 5:130-134; and Kolls *et al.*, 1994 *Proc. Natl. Acad. Sci. USA* 91:215-219), pox virus (*see* U.S. Patent No. 4,769,330; U.S. Patent No. 5,017,487; and WO 89/01973)), recombinant expression construct nucleic acid molecules complexed to a polycationic molecule (*see* WO 93/03709), and nucleic acids associated with liposomes (*see* Wang *et al.*, 1987 *Proc. Natl. Acad. Sci. USA* 84:7851). In certain embodiments, the DNA may be linked to killed or inactivated adenovirus (*see* Curiel *et al.*, 1992 *Hum. Gene Ther.* 3:147-154; Cotton *et al.*, 1992 *Proc. Natl. Acad. Sci. USA 89*:6094). Other suitable compositions include DNA-ligand (*see* Wu *et al.*, 1989 *J. Biol. Chem. 264*:16985-16987) and lipid-DNA combinations (see Felgner *et al.*, 1989 *Proc. Natl. Acad. Sci. USA* 84:7413-7417).

In addition to direct *in vivo* procedures, *ex vivo* procedures may be used in which cells are removed from a host (e.g., a subject, such as a human patient), modified, and placed into the same or another host animal. It will be evident that one can utilize any of the compositions noted above for introduction of constructs of the invention, either the proteins/polypeptides or the nucleic acids encoding them into tissue cells in an *ex vivo* context. Protocols for viral, physical and chemical methods of uptake are well known in the art.

Generation of antibodies

Polyclonal antibodies directed toward an antigen polypeptide generally are produced in animals (e.g., rabbits, hamsters, goats, sheep, horses, pigs, rats, gerbils, guinea pigs, mice, or any other suitable mammal, as well as other non-mammal

species) by means of multiple subcutaneous or intraperitoneal injections of antigen polypeptide or a fragment thereof and an adjuvant. Adjuvants include, but are not limited to, complete or incomplete Freund's adjuvant, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, and dinitrophenol. BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are also potentially useful adjuvants. It may be useful to conjugate an antigen polypeptide to a carrier protein that is immunogenic in the species to be immunized; typical carriers include keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, or soybean trypsin inhibitor. Also, aggregating agents such as alum are used to enhance the immune response. After immunization, the animals are bled and the serum is assayed for anti-antigen polypeptide antibody titer using conventional techniques. Polyclonal antibodies may be utilized in the sera from which they were detected, or may be purified from the sera using, e.g., antigen affinity chromatography.

Monoclonal antibodies directed toward antigen polypeptides are produced using any method which provides for the production of antibody molecules by continuous cell lines in culture. For example, monoclonal antibodies may be made by the hybridoma method as described in Kohler et al., Nature 256:495 [1975]; the human B-cell hybridoma technique (Kosbor et al., Immunol Today 4:72, 1983 ; Cote et al., Proc Natl Acad Sci 80: 2026-2030, 1983) and the EBV-hybridoma technique (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R Liss Inc, New York N.Y., pp 77-96, (1985).

When the hybridoma technique is employed, myeloma cell lines may be used. Cell lines suited for use in hybridoma-producing fusion procedures preferably do not produce endogenous antibody, have high fusion efficiency, and exhibit enzyme deficiencies that render them incapable of growing in certain selective media which support the growth of only the desired fused cells (hybridomas). For example, where the immunized animal is a mouse, one may use P3-X63/Ag8, P3-X63-Ag8.653, NS1/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 and S194/5XX0 Bul; for rats, one may use R210.RCY3, Y3-Ag 1.2.3, IR983F and 4B210; and U-266, GM1500-GRG2, LICR-LON-HMy2 and UC729-6 are all useful in connection with cell fusions.

In an alternative embodiment, human antibodies can be produced from phage-display libraries (Hoogenboom et al., J. Mol. Biol. 227: 381 [1991]; Marks et al., J. Mol. Biol. 222: 581, see also U.S. Patent No. 5,885,793).). These processes mimic immune selection through the display of antibody repertoires on the surface of filamentous bacteriophage, and subsequent selection of phage by their binding to an antigen of choice. One such technique is described in PCT Application No. PCT/US98/17364, filed in the name of Adams et al., which describes the isolation of high affinity and functional agonistic antibodies for MPL- and msk-receptors using such an approach. In this approach, a complete repertoire of human antibody genes can be created by cloning naturally rearranged human V genes from peripheral blood lymphocytes as previously described (Mullinax, et al., Proc. Natl. Acad. Sci.(USA) 87: 8095-8099 [1990]).

Alternatively, an entirely synthetic human heavy chain repertoire can be created from unrearranged V gene segments by assembling each human VH segment with D segments of random nucleotides together with a human J segment (Hoogenboom, et al., J. Mol. Biol. 227:381-388 [1992]). Likewise, a light chain repertoire can be constructed by combining each human V segment with a J segment (Griffiths, et al, EMBO J. 13:3245-3260 [1994]). Nucleotides encoding the complete antibody (i.e., both heavy and light chains) are linked as a single-chain Fv fragment and this polynucleotide is ligated to a nucleotide encoding a filamentous phage minor coat protein. When this fusion protein is expressed on the surface of the phage, a polynucleotide encoding a specific antibody can be identified by selection using an immobilized antigen.

Beyond the classic methods of generating polyclonal and monoclonal antibodies, any method for generating any known antibody form is contemplated. In addition to polyclonals and monoclonals, antibody forms include chimerized antibodies, humanized antibodies, CDR-grafted antibodies, and antibody fragments and variants.

Variants and Derivatives of Specific Binding Agents

In one example, insertion variants are provided wherein one or more amino acid residues supplement a specific binding agent amino acid sequence. Insertions may be located at either or both termini of the protein, or may be positioned within

internal regions of the specific binding agent amino acid sequence. Variant products of the invention also include mature specific binding agent products, i.e., specific binding agent products wherein leader or signal sequences are removed, and the resulting protein having additional amino terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from a specific protein. Polypeptides with an additional methionine residue at position -1 (e.g., Met-1-multivalent binding peptides with effector function) are contemplated, as are polypeptides of the invention with additional methionine and lysine residues at positions -2 and -1 (Met-2-Lys-1-multivalent binding proteins with effector function). Variants of the polypeptides of the invention having additional Met, Met-Lys, or Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

The invention also embraces specific polypeptides of the invention having additional amino acid residues which arise from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide. Variants which result from expression in other vector systems are also contemplated, including those wherein histidine tags are incorporated into the amino acid sequence, generally at the carboxy and/or amino terminus of the sequence.

In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a polypeptide of the invention are removed. Deletions can be effected at one or both termini of the polypeptide, or from removal of one or more residues within the amino acid sequence. Deletion variants necessarily include all fragments of a polypeptide according to the invention.

Antibody fragments refer to polypeptides having a sequence corresponding to at least part of an immunoglobulin variable region sequence. Fragments may be generated, for example, by enzymatic or chemical cleavage of polypeptides corresponding to full-length antibodies. Other binding fragments include those generated by synthetic techniques or by recombinant DNA techniques, such as the

expression of recombinant plasmids containing nucleic acid sequences encoding partial antibody variable regions. Preferred polypeptide fragments display immunological properties unique to, or specific for, a target as described herein. Fragments of the invention having the desired immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

In still another aspect, the invention provides substitution variants of multivalent binding polypeptides having effector function. Substitution variants include those polypeptides wherein one or more amino acid residues in an amino acid sequence are removed and replaced with alternative residues. In some embodiments, the substitutions are conservative in nature; however, the invention embraces substitutions that ore also non-conservative. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table A (*see* WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

Table A

Conservative Substitutions I		
SIDE CHAIN	CHARACTERISTIC	AMINO ACID
Aliphatic	Non-polar	G A P I L V
	Polar – uncharged	S T M N Q
	Polar - charged	D E K R
Aromatic		H F W Y
Other		N Q D E

Alternatively, conservative amino acids can be grouped as described in Lehninger, [Biochemistry, Second Edition; Worth Publishers, Inc. NY:NY (1975), pp.71-77] as set out in Table B, immediately below.

Table B

Conservative Substitutions II		
SIDE CHAIN	CHARACTERISTIC	AMINO ACID
Non-polar (hydrophobic)	A. Aliphatic:	A L I V P
	B. Aromatic	F W
	C. Sulfur-containing	M
	D. Borderline	G
Uncharged-polar	A. Hydroxyl	S T Y
	B. Amides	N Q
	C. Sulfhydryl	C
	D. Borderline	G
Positively Charged (Basic)		K R H
Negatively Charged (Acidic)		D E

Conservative Substitutions II

SIDE CHAIN	CHARACTERISTIC	AMINO ACID
Non-polar (hydrophobic)		
	A. Aliphatic:	A L I V P
	B. Aromatic:	F W

	C. Sulfur-containing:	M
	D. Borderline:	G
Uncharged-polar		
	A. Hydroxyl:	S T Y
	B. Amides:	N Q
	C. Sulfhydryl:	C
	D. Borderline:	G
Positively Charged (Basic)		K R H
Negatively Charged (Acidic)		D E

The invention also provides derivatives of specific binding agent polypeptides. Derivatives include specific binding agent polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. Preferably, the modifications are covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic, and inorganic moieties. Derivatives of the invention may be prepared to increase circulating half-life of a specific binding agent polypeptide, or may be designed to improve targeting capacity for the polypeptide to desired cells, tissues, or organs.

The invention further embraces multivalent binding proteins with effector function that are covalently modified or derivatized to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol, or polypropylene glycol, as described U.S. Patent Nos: 4,640,835, 4,496,689, 4,301,144, 4,670,417, 4,791,192 and 4,179,337. Still other useful polymers known in the art include monomethoxy-polyethylene glycol, dextran, cellulose, and other carbohydrate-based polymers, poly-(N-vinyl pyrrolidone)-polyethylene glycol, propylene glycol homopolymers, a polypropylene oxide/ethylene oxide co-polymer, polyoxyethylated polyols (e.g., glycerol) and polyvinyl alcohol, as well as mixtures of

these polymers. Particularly preferred are polyethylene glycol (PEG) –derivatized proteins. Water-soluble polymers may be bonded at specific positions, for example at the amino terminus of the proteins and polypeptides according to the invention, or randomly attached to one or more side chains of the polypeptide. The use of PEG for improving therapeutic capacities is described in US Pat. No. 6, 133, 426 to Gonzales, et al.

Target Sites for Immunoglobulin Mutagenesis

Certain strategies are available to manipulate inherent properties of an antigen-specific immunoglobulin (e.g., an antibody) that are not available to non-immunoglobulin-based binding molecules. A good example of the strategies favoring, e.g., antibody-based molecules, over these alternatives is the *in vivo* modulation of the affinity of an antibody for its target through affinity maturation, which takes advantage of the somatic hypermutation of immunoglobulin genes to yield antibodies of increasing affinity as an immune response progresses. Additionally, recombinant technologies have been developed to alter the structure of immunoglobulins and immunoglobulin regions and domains. Thus, polypeptides derived from antibodies may be produced that exhibit altered affinity for a given antigen, and a number of purification protocols and monitoring screens are known in the art for identifying and purifying or isolating these polypeptides. Using these known techniques, polypeptides comprising antibody-derived binding domains can be obtained that exhibit decreased or increased affinity for an antigen. Strategies for generating the polypeptide variants exhibiting altered affinity include the use of site-specific or random mutagenesis of the DNA encoding the antibody to change the amino acids present in the protein, followed by a screening step designed to recover antibody variants that exhibit the desired change, e.g., increased or decreased affinity relative to the unmodified parent or referent antibody.

The amino acid residues most commonly targeted in mutagenic strategies to alter affinity are those in the complementarity-determining region (CDR) or hyper-variable region of the light and the heavy chain variable regions of an antibody. These regions contain the residues that physicochemically interact with an antigen, as well as other amino acids that affect the spatial arrangement of these residues. However, amino acids in the framework regions of the variable domains outside the

CDR regions have also been shown to make substantial contributions to the antigen-binding properties of an antibody, and can be targeted to manipulate such properties. *See* Hudson, P.J. Curr. Opin. Biotech., 9: 395-402 (1999) and references therein.

Smaller and more effectively screened libraries of antibody variants can be produced by restricting random or site-directed mutagenesis to sites in the CDRs that correspond to areas prone to "hyper-mutation" during the somatic affinity maturation process. *See* Chowdhury, et al., Nature Biotech., 17: 568-572 (1999) and references therein. The types of DNA elements known to define hyper-mutation sites in this manner include direct and inverted repeats, certain consensus sequences, secondary structures, and palindromes. The consensus DNA sequences include the tetrabase sequence Purine-G-Pyrimidine-A/T (i.e., A or G - G - C or T – A or T) and the serine codon AGY (wherein Y can be C or T).

Thus, another aspect of the invention is a set of mutagenic strategies for modifying the affinity of an antibody for its target. These strategies include mutagenesis of the entire variable region of a heavy and/or light chain, mutagenesis of the CDR regions only, mutagenesis of the consensus hypermutation sites within the CDRs, mutagenesis of framework regions, or any combination of these approaches ("mutagenesis" in this context could be random or site-directed). Definitive delineation of the CDR regions and identification of residues comprising the binding site of an antibody can be accomplished though solving the structure of the antibody in question, and the antibody:ligand complex, through techniques known to those skilled in the art, such as X-ray crystallography. Various methods based on analysis and characterization of such antibody crystal structures are known to those of skill in the art and can be employed to approximate the CDR regions. Examples of such commonly used methods include the Kabat, Chothia, AbM and contact definitions.

The Kabat definition is based on sequence variability and is the most commonly used definition to predict CDR regions. Johnson, et al., Nucleic Acids Research, 28: 214-8 (2000). The Chothia definition is based on the location of the structural loop regions. (Chothia et al., J. Mol. Biol., 196: 901-17 [1986]; Chothia et al., Nature, 342: 877-83 [1989].) The AbM definition is a compromise between the Kabat and Chothia definitions. AbM is an integral suite of programs for antibody structure modeling produced by the Oxford Molecular Group (Martin , et al., Proc.

Natl. Acad. Sci (USA) 86:9268-9272 [1989]; Rees, et al., ABMTM, a computer program for modeling variable regions of antibodies, Oxford, UK; Oxford Molecular, Ltd.). The AbM suite models the tertiary structure of an antibody from primary sequence using a combination of knowledge databases and ab initio methods An additional definition, known as the contact definition, has been recently introduced. *See* MacCallum et al., J. Mol. Biol., 5:732-45 (1996). This definition is based on an analysis of the available complex crystal structures.

By convention, the CDR domains in the heavy chain are typically referred to as H1, H2 and H3, and are numbered sequentially in order moving from the amino terminus to the carboxy terminus. The CDR regions in the light chain are typically referred to as L1, L2 and L3, and are numbered sequentially in order moving from the amino terminus to the carboxy terminus.

The CDR-H1 is approximately 10 to 12 residues in length and typically starts 4 residues after a Cys according to the Chothia and AbM definitions, or typically 5 residues later according to the Kabat definition. The H1 is typically followed by a Trp, typically Trp-Val, but also Trp-Ile, or Trp-Ala. The length of H1 is approximately 10 to 12 residues according to the AbM definition, while the Chothia definition excludes the last 4 residues.

The CDR-H2 typically starts 15 residues after the end of H1 according to the Kabat and AbM definitions. The residues preceding H2 are typically Leu-Glu-Trp-Ile-Gly but there are a number of variations. H2 is typically followed by the amino acid sequence Lys/Arg-Leu/Ile/Val/Phe/Thr/Ala-Thr/Ser/Ile/Ala. According to the Kabat definition, the length of H2 is approximately 16 to 19 residues, where the AbM definition predicts the length to be typically 9 to 12 residues.

The CDR-H3 typically starts 33 residues after the end of H2 and is typically preceded by the amino acid sequence Cys-Ala-Arg. H3 is typically followed by the amino acid Gly. The length of H3 ranges from 3 to 25 residues

The CDR-L1 typically starts at approximately residue 24 and will typically follow a Cys. The residue after the CDR-L1 is always Trp and will typically begin one of the following sequences: Trp-Tyr-Gln, Trp-Leu-Gln, Trp-Phe-Gln, or Trp-Tyr-Leu. The length of CDR-L1 is approximately 10 to 17 residues.

The CDR-L2 starts approximately 16 residues after the end of L1. It will generally follow residues Ile-Tyr, Val-Tyr, Ile-Lys or Ile-Phe. The length of CDR-L2 is approximately 7 residues.

The CDR-L3 typically starts 33 residues after the end of L2 and typically follows a Cys. L3 is typically followed by the amino acid sequence Phe-Gly-XXX-Gly. The length of L3 is approximately 7 to 11 residues.

Various methods for modifying antibodies have been described in the art, including, e.g., methods of producing humanized antibodies wherein the sequence of the humanized immunoglobulin heavy chain variable region framework is 65% to 95% identical to the sequence of the donor immunoglobulin heavy chain variable region framework. Each humanized immunoglobulin chain will usually comprise, in addition to the CDRs, amino acids from the donor immunoglobulin framework that are, e.g., capable of interacting with the CDRs to effect binding affinity, such as one or more amino acids that are immediately adjacent to a CDR in the donor immunoglobulin or those within about 3 angstroms, as predicted by molecular modeling. The heavy and light chains may each be designed by using any one or all of various position criteria. When combined into an intact antibody, humanized immunoglobulins are substantially non-immunogenic in humans and retain substantially the same affinity as the donor immunoglobulin to the antigen, such as a protein or other compound containing an epitope.

In one example, methods for the production of antibodies, and antibody fragments, are described that have binding specificity similar to a parent antibody, but which have increased human characteristics. Humanized antibodies are obtained by chain shuffling using, for example, phage display technology and a polypeptide comprising the heavy or light chain variable region of a non-human antibody specific for an antigen of interest, which is then combined with a repertoire of human complementary (light or heavy) chain variable regions. Hybrid pairings which are specific for the antigen of interest are identified and human chains from the selected pairings are combined with a repertoire of human complementary variable domains (heavy or light). In another embodiment, a component of a CDR from a non-human antibody is combined with a repertoire of component parts of CDRs from human antibodies. From the resulting library of antibody polypeptide dimers, hybrids are

selected and may used in a second humanizing shuffling step; alternatively, this second step is eliminated if the hybrid is already of sufficient human character to be of therapeutic value. Methods of modification to increase human character are known in the art.

Another example is a method for making humanized antibodies by substituting a CDR amino acid sequence for the corresponding human CDR amino acid sequence and/or substituting a FR amino acid sequence for the corresponding human FR amino acid sequences.

Yet another example provides methods for identifying the amino acid residues of an antibody variable domain that may be modified without diminishing the native affinity of the antigen binding domain while reducing its immunogenicity with respect to a heterologous species and methods for preparing these modified antibody variable regions as useful for administration to heterologous species.

Modification of an immunoglobulin such as an antibody by any of the methods known in the art is designed to achieve increased or decreased binding affinity for an antigen and/or to reduce immunogenicity of the antibody in the recipient and/or to modulate effector activity levels. In one approach, humanized antibodies can be modified to eliminate glycosylation sites in order to increase affinity of the antibody for its cognate antigen (Co, et al., Mol. Immunol. 30:1361-1367 [1993]). Techniques such as "reshaping," hyperchimerization," and "veneering/resurfacing" have produced humanized antibodies with greater therapeutic potential. Vaswami, et al., Annals of Allergy, Asthma, & Immunol 81:105 (1998); Roguska, et al., Prot. Engineer. 9:895-904 (1996)]. *See also* US Pat. No. 6,072,035, which describes methods for reshaping antibodies. While these techniques diminish antibody immunogenicity by reducing the number of foreign residues, they do not prevent anti-idiotypic and anti-allotypic responses following repeated administration of the antibodies. Alternatives to these methods for reducing immunogenicity are described in Gilliland et al., J. Immunol. 62(6):3663-71 (1999).

In many instances, humanizing antibodies results in a loss of antigen binding capacity. It is therefore preferable to "back mutate" the humanized antibody to include one or more of the amino acid residues found in the original (most often

rodent) antibody in an attempt to restore binding affinity of the antibody. *See, for example,* Saldanha et al., Mol. Immunol. 36:709-19 (1999).

Glycosylation of immunoglobulins has been shown to affect effector functions, structural stability, and the rate of secretion from antibody-producing cells (see Leatherbarrow et al., Mol. Immunol. 22:407 (1985), incorporated herein by reference). The carbohydrate groups responsible for these properties are generally attached to the constant regions of antibodies. For example, glycosylation of IgG at Asn 297 in the C_{H2} domain facilitates full capacity of the IgG to activate complement-dependent cytolysis (Tao et al., J. Immunol. 143:2595 (1989)). Glycosylation of IgM at Asn 402 in the C_{H3} domain, for example, facilitates proper assembly and cytolytic activity of the antibody (Muraoka et al., J. Immunol. 142:695 (1989)). Removal of glycosylation sites at positions 162 and 419 in the C_{H1} and C_{H3} domains of an IgA antibody led to intracellular degradation and at least 90% inhibition of secretion (Taylor et al., Wall, Mol. Cell. Biol. 8:4197 (1988)). Accordingly, the molecules of the invention include mutationally altered immunoglobulins exhibiting altered glycosylation patterns by mutation of specific residues in, e.g., a constant sub-region to alter effector function. *See* Co et al., Mol. Immunol. 30:1361-1367 (1993), Jacquemon et al., J. Thromb. Haemost. 4:1047-1055 (2006), Schuster et al., Cancer Res. 65:7934-7941 (2005), and Warnock et al., Biotechnol Bioeng. 92:831-842 (2005), each incorporated herein by reference.

The invention also includes multivalent binding molecules having at least one binding domain that is at least 80%, preferably 90% or 95% or 99% identical in sequence to a known immunoglobulin variable region sequence and which has at least one residue that differs from such immunoglobulin variable region, wherein the changed residue adds a glycosylation site, changes the location of one or more glycosylation site(s), or preferably removes a glycosylation site relative to the immunoglobulin variable region. In some embodiments, the change removes an N-linked glycosylation site in a an immunoglobulin variable region framework, or removes an N-linked glycosylation site that occurs in the immunoglobulin heavy chain variable region framework in the region spanning about amino acid residue 65 to about amino acid residue 85, using the numbering convention of Co et al., J. Immunol. 148: 1149, (1992).

Any method known in the art is contemplated for producing the multivalent binding molecules exhibiting altered glycosylation patterns relative to an immunoglobulin referent sequence. For example, any of a variety of genetic techniques may be employed to alter one or more particular residues. Alternatively, the host cells used for production may be engineered to produce the altered glycosylation pattern. One method known in the art, for example, provides altered glycosylation in the form of bisected, non-fucosylated variants that increase ADCC. The variants result from expression in a host cell containing an oligosaccharide-modifying enzyme. Alternatively, the Potelligent technology of BioWa/Kyowa Hakko is contemplated to reduce the fucose content of glycosylated molecules according to the invention. In one known method, a CHO host cell for recombinant immunoglobulin production is provided that modifies the glycosylation pattern of the immunoglobulin F_C region, through production of GDP-fucose. This technology is available to modify the glycosylation pattern of a constant sub-region of a multivalent binding molecule according to the invention.

In addition to modifying the binding properties of binding domains, such as the binding domains of immunoglobulins, and in addition to such modifications as humanization, the invention comprehends the modulation of effector function by changing or mutating residues contributing to effector function, such as the effector function of a constant sub-region. These modifications can be effected using any technique known in the art, such as the approach disclosed in Presta et al., Biochem. Soc. Trans. 30:487-490 (2001), incorporated herein by reference. Exemplary approaches would include the use of the protocol disclosed in Presta et al. to modify specific residues known to affect binding in one or more constant sub-regions corresponding to FCγRI, FCγRII, FCγRIII, FCαR, and FCεR.

In another approach, the Xencor XmAb technology is available to engineer constant sub-regions corresponding to F_C domains to enhance cell killing effector function. See Lazar et al., Proc. Natl. Acad. Sci. (USA) 103(11):4005-4010 (2006), incorporated herein by reference. Using this approach, for example, one can generate constant sub-regions optimized for $F_C\gamma R$ specificity and binding, thereby enhancing cell killing effector function.

Production of multivalent binding proteins with effector function

A variety of expression vector/host systems may be utilized to contain and express the multivalent binding protein (with effector function) of the invention. These systems include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, cosmid, or other expression vectors; yeast transformed with yeast expression or shuttle vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transfected with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with bacterial expression vectors (e.g., Ti or pBR322 plasmid); or animal cell systems. Mammalian cells that are useful in recombinant multivalent binding protein productions include, but are not limited to, VERO cells, HeLa cells, Chinese hamster ovary (CHO) cell lines, COS cells (such as COS-7), W138, BHK, HepG2, 3T3, RIN, MDCK, A549, PC12, K562 and HEK293 cells. Exemplary protocols for the recombinant expression of the multivalent binding protein are described herein below.

An expression vector can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, a promoter, enhancer, or factor-specific binding site, (2) a structural or sequence that encodes the binding agent which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant multivalent binding protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final multivalent binding protein.

For example, the multivalent binding proteins may be recombinantly expressed in yeast using a commercially available expression system, e.g., the Pichia Expression System (Invitrogen, San Diego, CA), following the manufacturer's instructions. This system also relies on the pre-pro-alpha sequence to direct secretion, but transcription of the insert is driven by the alcohol oxidase (AOX1) promoter upon induction by methanol. The secreted multivalent binding peptide may be purified

from the yeast growth medium by, e.g., the methods used to purify the peptide from bacterial and mammalian cell supernatants.

Alternatively, the cDNA encoding the multivalent binding peptide may be cloned into the baculovirus expression vector pVL1393 (PharMingen, San Diego, CA). This vector can be used according to the manufacturer's directions (PharMingen) to infect Spodoptera frugiperda cells in SF9 protein-free medium and to produce recombinant protein. The multivalent binding protein can be purified and concentrated from the medium using a heparin-Sepharose column (Pharmacia, Piscataway, NJ). Insect systems for protein expression, such as the SF9 system, are well known to those of skill in the art. In one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) can be used as a vector to express foreign genes in the *Spodoptera frugiperda* cells or in Trichoplusia larvae. The multivalent binding peptide coding sequence can be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the multivalent binding peptide will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can be used to infect S. frugiperda cells or Trichoplusia larvae in which peptide is expressed (Smith et al., J Virol 46: 584, 1983; Engelhard et al., Proc Nat Acad Sci (USA) 91: 3224-7, 1994).

In another example, the DNA sequence encoding the multivalent binding peptide can be amplified by PCR and cloned into an appropriate vector, for example, pGEX-3X (Pharmacia, Piscataway, NJ). The pGEX vector is designed to produce a fusion protein comprising glutathione-S-transferase (GST), encoded by the vector, and a multivalent binding protein encoded by a DNA fragment inserted into the cloning site of the vector. The primers for the PCR can be generated to include for example, an appropriate cleavage site. Where the multivalent binding protein fusion moiety is used solely to facilitate expression or is otherwise not desirable as an attachment to the peptide of interest, the recombinant multivalent binding protein fusion may then be cleaved from the GST portion of the fusion protein. The pGEX-3X/multivalent binding peptide construct is transformed into *E. coli* XL-1 Blue cells (Stratagene, La Jolla CA), and individual transformants isolated and grown. Plasmid DNA from individual transformants is purified and may be partially sequenced using

an automated sequencer to confirm the presence of the desired multivalent binding protein-encoding nucleic acid insert in the proper orientation.

The fused multivalent binding protein, which may be produced as an insoluble inclusion body in the bacteria, can be purified as follows. Host cells can be harvested by centrifugation; washed in 0.15 M NaCl, 10 mM Tris, pH 8, 1 mM EDTA; and treated with 0.1 mg/ml lysozyme (Sigma Chemical Co.) for 15 minutes at room temperature. The lysate can be cleared by sonication, and cell debris can be pelleted by centrifugation for 10 minutes at 12,000 X g. The multivalent binding protein fusion-containing pellet can be resuspended in 50 mM Tris, pH 8, and 10 mM EDTA, layered over 50% glycerol, and centrifuged for 30 minutes at 6000*g*. The pellet can be resuspended in standard phosphate buffered saline solution (PBS) free of Mg^{++} and Ca^{++}. The multivalent binding protein fusion can be further purified by fractionating the resuspended pellet in a denaturing SDS polyacrylamide gel (Sambrook et al.). The gel is soaked in 0.4 M KCl to visualize the protein, which is excised and electroeluted in gel-running buffer lacking SDS. If the GST/multivalent binding peptide fusion protein is produced in bacteria as a soluble protein, it can be purified using the GST Purification Module (Pharmacia Biotech).

The multivalent binding protein fusion is preferably subjected to digestion to cleave the GST from the multivalent binding peptide of the invention. The digestion reaction (20-40 μg fusion protein, 20-30 units human thrombin (4000 U/mg (Sigma) in 0.5 ml PBS) can be incubated 16-48 hours at room temperature and loaded on a denaturing SDS-PAGE gel to fractionate the reaction products. The gel can be soaked in 0.4 M KCl to visualize the protein bands. The identity of the protein band corresponding to the expected molecular weight of the multivalent binding peptide can be confirmed by amino acid sequence analysis using an automated sequencer (Applied Biosystems Model 473A, Foster City, CA). Alternatively, the identity can be confirmed by performing HPLC and/or mass spectrometry of the peptides.

Alternatively, a DNA sequence encoding the multivalent binding peptide can be cloned into a plasmid containing a desired promoter and, optionally, a leader sequence (see, e.g., Better et al., Science, 240:1041-43, 1988). The sequence of this construct can be confirmed by automated sequencing. The plasmid can then be transformed into a suitable *E. coli* strain, such as strain MC1061, using standard

procedures employing $CaCl_2$ incubation and heat shock treatment of the bacteria (Sambrook et al.). The transformed bacteria can be grown in LB medium supplemented with carbenicillin or another suitable form of selection as would be known in the art, and production of the expressed protein can be induced by growth in a suitable medium. If present, the leader sequence can effect secretion of the multivalent binding peptide and be cleaved during secretion. The secreted recombinant protein can be purified from the bacterial culture medium by the methods described herein below.

Mammalian host systems for the expression of the recombinant protein are well known to those of skill in the art and are preferred systems. Host cell strains can be chosen for a particular ability to process the expressed protein or produce certain post-translation modifications that will be useful in providing protein activity. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Different host cells such as CHO, HeLa, MDCK, 293, WI38, and the like, have specific cellular machinery and characteristic mechanisms for such post-translational activities and can be chosen to ensure the correct modification and processing of the foreign protein.

It is preferable that the transformed cells be used for long-term, high-yield protein production and, as such, stable expression is desirable. Once such cells are transformed with vectors that preferably contain at least one selectable marker along with the desired expression cassette, the cells are grown for 1-2 days in an enriched medium before being switched to selective medium. The selectable marker is designed to confer resistance to selection and its presence allows growth and recovery of cells that successfully express the foreign protein. Resistant clumps of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell.

A number of selection systems can be used to recover the cells that have been transformed for recombinant protein production. Such selection systems include, but are not limited to, HSV thymidine kinase, hypoxanthine-guanine phosphoribosyltransferase and adenine phosphoribosyltransferase genes, in tk-, hgprt- or aprt- cells, respectively. Also, anti-metabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate; gpt, which confers

resistance to mycophenolic acid; neo, which confers resistance to the aminoglycoside G418 and confers resistance to chlorsulfuron; and hygro, which confers resistance to hygromycin. Additional selectable genes that may be useful include trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine. Markers that give a visual indication for identification of transformants include anthocyanins, β-glucuronidase and its substrate, GUS, and luciferase and its substrate, luciferin.

Purification of Proteins

Protein purification techniques are well known to those of skill in the art. These techniques involve, at one level, the crude fractionation of the polypeptide and non-polypeptide fractions. Having separated the multivalent binding polypeptide from at least one other protein, the polypeptide of interest is purified, but further purification using chromatographic and electrophoretic techniques to achieve partial or complete purification (or purification to homogeneity) is frequently desired. Analytical methods particularly suited to the preparation of a pure multivalent binding peptide are ion-exchange chromatography, exclusion chromatography; polyacrylamide gel electrophoresis; and isoelectric focusing. Particularly efficient methods of purifying peptides are fast protein liquid chromatography and HPLC.

Certain aspects of the present invention concern the purification, and in particular embodiments, the substantial purification, of an encoded multivalent binding protein or peptide. The term "purified multivalent binding protein or peptide" as used herein, is intended to refer to a composition, isolatable from other components, wherein the multivalent binding protein or peptide is purified to any degree relative to its naturally obtainable state. A purified multivalent binding protein or peptide therefore also refers to a multivalent binding protein or peptide, free from the environment in which it may naturally occur.

Generally, "purified" will refer to a multivalent binding protein composition that has been subjected to fractionation to remove various other components, and which composition substantially retains its expressed biological activity. Where the term "substantially purified" is used, this designation refers to a multivalent binding protein composition in which the multivalent binding protein or peptide forms the major component of the composition, such as constituting about 50%, about 60%,

about 70%, about 80%, about 90%, about 95%, about 99% or more of the protein, by weight, in the composition.

Various methods for quantifying the degree of purification of the multivalent binding protein will be known to those of skill in the art in light of the present disclosure. These include, for example, determining the specific binding activity of an active fraction, or assessing the amount of multivalent binding polypeptides within a fraction by SDS/PAGE analysis. A preferred method for assessing the purity of a multivalent binding protein fraction is to calculate the binding activity of the fraction, to compare it to the binding activity of the initial extract, and to thus calculate the degree of purification, herein assessed by a "-fold purification number." The actual units used to represent the amount of binding activity will, of course, be dependent upon the particular assay technique chosen to follow the purification and whether or not the expressed multivalent binding protein or peptide exhibits a detectable binding activity.

Various techniques suitable for use in multivalent binding protein purification are well known to those of skill in the art. These include, for example, precipitation with ammonium sulfate, PEG, antibodies and the like, or by heat denaturation, followed by centrifugation; chromatography steps such as ion exchange, gel filtration, reverse phase, hydroxylapatite and affinity chromatography; isoelectric focusing; gel electrophoresis; and combinations of these and other techniques. As is generally known in the art, it is believed that the order of conducting the various purification steps may be changed, or that certain steps may be omitted, and still result in a suitable method for the preparation of a substantially purified multivalent binding protein.

There is no general requirement that the multivalent binding protein always be provided in its most purified state. Indeed, it is contemplated that less substantially multivalent binding proteins will have utility in certain embodiments. Partial purification may be accomplished by using fewer purification steps in combination, or by utilizing different forms of the same general purification scheme. For example, it is appreciated that a cation-exchange column chromatography performed utilizing an HPLC apparatus will generally result in greater purification than the same technique utilizing a low pressure chromatography system. Methods exhibiting a lower degree

of relative purification may have advantages in total recovery of multivalent binding protein product, or in maintaining binding activity of an expressed multivalent binding protein.

It is known that the migration of a polypeptide can vary, sometimes significantly, with different conditions of SDS/PAGE (Capaldi et al., Biochem. Biophys. Res. Comm., 76:425, 1977). It will therefore be appreciated that under differing electrophoresis conditions, the apparent molecular weights of purified or partially purified multivalent binding protein expression products may vary.

Effector cells

Effector cells for inducing, e.g., ADCC, ADCP (antibody-dependent cellular phagocytosis), and the like, against a target cell include human leukocytes, macrophages, monocytes, activated neutrophils, activated natural killer (NK) cells, and eosinophils. Effector cells express $F_C\alpha R$ (CD89), FcγRI, FcγRII, FcγRIII, and/or $F_C\varepsilon R1$ and include, for example, monocytes and activated neutrophils. Expression of FcγRI, e.g., has been found to be up-regulated by interferon gamma (IFN-γ). This enhanced expression increases the cytotoxic activity of monocytes and neutrophils against target cells. Accordingly, effector cells may be activated with (IFN-γ) or other cytokines (e.g., TNF-α or β, colony stimulating factor, IL-2) to increase the presence of FcγRI on the surface of the cells prior to being contacted with a multivalent protein of the invention.

The multivalent proteins of the invention provide an antibody effector function, such as antibody-dependent effector cell-mediated cytotoxicity (ADCC), for use against a target cell. Multivalent proteins with effector function are administered alone, as taught herein, or after being coupled to an effector cell, thereby forming an "activated effector cell." An "activated effector cell" is an effector cell, as defined herein, linked to a multivalent protein with effector function, also as defined herein, such that the effector cell is effectively provided with a targeting function prior to administration.

Activated effector cells are administered *in vivo* as a suspension of cells in a physiologically acceptable solution. The number of cells administered is on the order of 10^8-10^9, but will vary depending on the therapeutic purpose. In general, the

amount will be sufficient to obtain localization of the effector cell at the target cell, and to provide a desired level of effector cell function in that locale, such as cell killing by ADCC and/or phagocytosis. The term physiologically acceptable solution, as used herein, is intended to include any carrier solution which stabilizes the targeted effector cells for administration *in vivo* including, for example, saline and aqueous buffer solutions, solvents, antibacterial and antifungal agents, isotonic agents, and the like.

Accordingly, another aspect of the invention provides a method of inducing a specific antibody effector function, such as ADCC, against a cell in a subject, comprising administering to the subject a multivalent protein (or encoding nucleic acid) or activated effector cell in a physiologically acceptable medium. Routes of administration can vary and suitable administration routes will be determined by those of skill in the art based on a consideration of case-specific variables and routine procedures, as is known in the art.

Cell-free effects

Cell-free effects are also provided by the multivalent molecules of the invention, e.g., by providing a CDC functionality. The complement system is a biochemical cascade of the immune system that helps clear foreign matter such as pathogens from an organism. It is derived from many small plasma proteins that work together in inducing cytolysis of a target cell by disrupting the target cell's plasma membrane. The complement system consists of more than 35 soluble and cell-bound proteins, 12 of which are directly involved in the complement pathways. The proteins are active in three biochemical pathways leading to the activation of the complement system: the classical complement pathway, the alternate complement pathway, and the mannose-binding lectin pathway. Antibodies, in particular the IgG1 class, can also "fix" complement. A detailed understanding of these pathways has been achieved in the art and will not be repeated here, but it is worth noting that complement-dependent cytotoxicity is not dependent on the interaction of a binding molecule with a cell, e.g., a B cell, of the immune system. Also worth noting is that the complement system is regulated by complement regulating proteins. These proteins are present at higher concentrations in the blood plasma than the complement proteins. The complement regulating proteins are found on the surfaces of self-cells,

providing a mechanism to prevent self-cells from being targeted by complement proteins. It is expected that the complement system plays a role in several diseases with an immune component, such as Barraquer-Simons Syndrome, Alzheimer's disease, asthma, lupus erythematosus, various forms of arthritis, autoimmune heart disease, and multiple sclerosis. Deficiencies in the terminal pathway predispose an individual to both autoimmune disease and infections (particularly meningitis).

Diseases, disorders and conditions

The invention provides a multivalent binding proteins with effector function, and variant and derivative thereof, that bind to one or more binding partners and those binding events are useful in the treatment, prevention, or amelioration of a symptom associated with a disease, disorder or pathological condition, preferably one afflicting humans. In preferred embodiments of these methods, the multivalent (and multispecific) binding protein with effector function associates a cell bearing a target, such as a tumor-specific cell-surface marker, with an effector cell, such as a cell of the immune system exhibiting cytotoxic activity. In other embodiments, the multispecific, multivalent binding protein with effector function specifically binds two different disease-, disorder- or condition-specific cell-surface markers to ensure that the correct target is associated with an effector cell, such as a cytotoxic cell of the immune system. Additionally, the multivalent binding protein with effector function can be used to induce or increase antigen activity, or to inhibit antigen activity. The multivalent binding proteins with effector function are also suitable for combination therapies and palliative regimes.

In one aspect, the present invention provides compositions and methods useful for treating or preventing diseases and conditions characterized by aberrant levels of antigen activity associated with a cell. These diseases include cancers and other hyperproliferative conditions, such as hyperplasia, psoriasis, contact dermatitis, immunological disorders, and infertility. A wide variety of cancers, including solid tumors and leukemias are amenable to the compositions and methods disclosed herein. Types of cancer that may be treated include, but are not limited to: adenocarcinoma of the breast, prostate, and colon; all forms of bronchogenic carcinoma of the lung; myeloid; melanoma; hepatoma; neuroblastoma; papilloma; apudoma; choristoma; branchioma; malignant carcinoid syndrome; carcinoid heart

disease; and carcinoma (e.g., Walker, basal cell, basosquamous, Brown-Pearce, ductal, Ehrlich tumor, Krebs 2, merkel cell, mucinous, non-small cell lung, oat cell, papillary, scirrhous, bronchiolar, bronchogenic, squamous cell, and transitional cell). Additional types of cancers that may be treated include: histiocytic disorders; leukemia; histiocytosis malignant; Hodgkin's disease; immunoproliferative small; non-Hodgkin's lymphoma; plasmacytoma; reticuloendotheliosis; melanoma; chondroblastoma; chondroma; chondrosarcoma; fibroma; fibrosarcoma; giant cell tumors; histiocytoma; lipoma; liposarcoma; mesothelioma; myxoma; myxosarcoma; osteoma; osteosarcoma; chordoma; craniopharyngioma; dysgerminoma; hamartoma; mesenchymoma; mesonephroma; myosarcoma; ameloblastoma; cementoma; odontoma; teratoma; thymoma; trophoblastic tumor. Further, the following types of cancers are also contemplated as amenable to treatment: adenoma; cholangioma; cholesteatoma; cyclindroma; cystadenocarcinoma; cystadenoma; granulosa cell tumor; gynandroblastoma; hepatoma; hidradenoma; islet cell tumor; Leydig cell tumor; papilloma; sertoli cell tumor; theca cell tumor; leimyoma; leiomyosarcoma; myoblastoma; myomma; myosarcoma; rhabdomyoma; rhabdomyosarcoma; ependymoma; ganglioneuroma; glioma; medulloblastoma; meningioma; neurilemmoma; neuroblastoma; neuroepithelioma; neurofibroma; neuroma; paraganglioma; paraganglioma nonchromaffin. The types of cancers that may be treated also include, but are not limited to, angiokeratoma; angiolymphoid hyperplasia with eosinophilia; angioma sclerosing; angiomatosis; glomangioma; hemangioendothelioma; hemangioma; hemangiopericytoma; hemangiosarcoma; lymphangioma; lymphangiomyoma; lymphangiosarcoma; pinealoma; carcinosarcoma; chondrosarcoma; cystosarcoma phyllodes; fibrosarcoma; hemangiosarcoma; leiomyosarcoma; leukosarcoma; liposarcoma; lymphangiosarcoma; myosarcoma; myxosarcoma; ovarian carcinoma; rhabdomyosarcoma; sarcoma; neoplasms; nerofibromatosis; and cervical dysplasia. The invention further provides compositions and methods useful in the treatment of other conditions in which cells have become immortalized or hyperproliferative due to abnormally high expression of antigen.

Exemplifying the variety of hyperproliferative disorders amenable to the compositions and methods of the invention are B-cell cancers, including B-cell lymphomas (such as various forms of Hodgkin's disease, non-Hodgkins lymphoma

(NHL) or central nervous system lymphomas), leukemias (such as acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia (CLL), Hairy cell leukemia and chronic myoblastic leukemia) and myelomas (such as multiple myeloma). Additional B cell cancers include small lymphocytic lymphoma, B-cell prolymphocytic leukemia, lymphoplasmacytic lymphoma, splenic marginal zone lymphoma, plasma cell myeloma, solitary plasmacytoma of bone, extraosseous plasmacytoma, extra-nodal marginal zone B-cell lymphoma of mucosa-associated (MALT) lymphoid tissue, nodal marginal zone B-cell lymphoma, follicular lymphoma, mantle cell lymphoma, diffuse large B-cell lymphoma, mediastinal (thymic) large B-cell lymphoma, intravascular large B-cell lymphoma, primary effusion lymphoma, Burkitt's lymphoma/leukemia, B-cell proliferations of uncertain malignant potential, lymphomatoid granulomatosis, and post-transplant lymphoproliferative disorder.

Disorders characterized by autoantibody production are often considered autoimmune diseases. Autoimmune diseases include, but are not limited to: arthritis, rheumatoid arthritis, juvenile rheumatoid arthritis, osteoarthritis, polychondritis, psoriatic arthritis, psoriasis, dermatitis, polymyositis/dermatomyositis, inclusion body myositis, inflammatory myositis, toxic epidermal necrolysis, systemic scleroderma and sclerosis, CREST syndrome, responses associated with inflammatory bowel disease, Crohn's disease, ulcerative colitis, respiratory distress syndrome, adult respiratory distress syndrome (ARDS), meningitis, encephalitis, uveitis, colitis, glomerulonephritis, allergic conditions, eczema, asthma, conditions involving infiltration of T cells and chronic inflammatory responses, atherosclerosis, autoimmune myocarditis, leukocyte adhesion deficiency, systemic lupus erythematosus (SLE), subacute cutaneous lupus erythematosus, discoid lupus, lupus myelitis, lupus cerebritis, juvenile onset diabetes, multiple sclerosis, allergic encephalomyelitis, neuromyelitis optica, rheumatic fever, Sydenham's chorea, immune responses associated with acute and delayed hypersensitivity mediated by cytokines and T-lymphocytes, tuberculosis, sarcoidosis, granulomatosis including Wegener's granulomatosis and Churg-Strauss disease, agranulocytosis, vasculitis (including hypersensitivity vasculitis/angiitis, ANCA and rheumatoid vasculitis), aplastic anemia, Diamond Blackfan anemia, immune hemolytic anemia including autoimmune hemolytic anemia (AIHA), pernicious anemia, pure red cell aplasia

(PRCA), Factor VIII deficiency, hemophilia A, autoimmune neutropenia, pancytopenia, leukopenia, diseases involving leukocyte diapedesis, central nervous system (CNS) inflammatory disorders, multiple organ injury syndrome, myasthenia gravis, antigen-antibody complex mediated diseases, anti-glomerular basement membrane disease, anti-phospholipid antibody syndrome, allergic neuritis, Behcet disease, Castleman's syndrome, Goodpasture's syndrome, Lambert-Eaton Myasthenic Syndrome, Reynaud's syndrome, Sjorgen's syndrome, Stevens-Johnson syndrome, solid organ transplant rejection, graft versus host disease (GVHD), bullous pemphigoid, pemphigus, autoimmune polyendocrinopathies, seronegative spondyloarthropathies, Reiter's disease, stiff-man syndrome, giant cell arteritis, immune complex nephritis, IgA nephropathy, IgM polyneuropathies or IgM mediated neuropathy, idiopathic thrombocytopenic purpura (ITP), thrombotic throbocytopenic purpura (TTP), Henoch-Schonlein purpura, autoimmune thrombocytopenia, autoimmune disease of the testis and ovary including autoimmune orchitis and oophoritis, primary hypothyroidism; autoimmune endocrine diseases including autoimmune thyroiditis, chronic thyroiditis (Hashimoto's Thyroiditis), subacute thyroiditis, idiopathic hypothyroidism, Addison's disease, Grave's disease, autoimmune polyglandular syndromes (or polyglandular endocrinopathy syndromes), Type I diabetes also referred to as insulin-dependent diabetes mellitus (IDDM) and Sheehan's syndrome; autoimmune hepatitis, lymphoid interstitial pneumonitis (HIV), bronchiolitis obliterans (non-transplant) vs NSIP, Guillain-Barré Syndrome, large vessel vasculitis (including polymyalgia rheumatica and giant cell (Takayasu's) arteritis), medium vessel vasculitis (including Kawasaki's disease and polyarteritis nodosa), polyarteritis nodosa (PAN) ankylosing spondylitis, Berger's disease (IgA nephropathy), rapidly progressive glomerulonephritis, primary biliary cirrhosis, Celiac sprue (gluten enteropathy), cryoglobulinemia, cryoglobulinemia associated with hepatitis, amyotrophic lateral sclerosis (ALS), coronary artery disease, familial Mediterranean fever, microscopic polyangiitis, Cogan's syndrome, Whiskott-Aldrich syndrome and thromboangiitis obliterans.

Rheumatoid arthritis (RA) is a chronic disease characterized by inflammation of the joints, leading to swelling, pain, and loss of function. Patients having RA for an extended period usually exhibit progressive joint destruction, deformity, disability and even premature death. Beyond RA, inflammatory diseases, disorders and

conditions in general are amenable to treatment, prevention or amelioration of symptoms (e.g., heat, pain, swelling, redness) associated with the process of inflammation, and the compositions and methods of the invention are beneficial in treating, preventing or ameliorating aberrant or abnormal inflammatory processes, including RA.

Crohn's disease and a related disease, ulcerative colitis, are the two main disease categories that belong to a group of illnesses called inflammatory bowel disease (IBD). Crohn's disease is a chronic disorder that causes inflammation of the digestive or gastrointestinal (GI) tract. Although it can involve any area of the GI tract from the mouth to the anus, it most commonly affects the small intestine and/or colon. In ulcerative colitis, the GI involvement is limited to the colon. Crohn's disease may be characterized by antibodies against neutrophil antigens, i.e., the "perinuclear anti-neutrophil antibody" (pANCA), and *Saccharomyces cervisiae*, i.e. the "anti-Saccharomyces cerevisiae antibody" (ASCA). Many patients with ulcerative colitis have the pANCA antibody in their blood, but not the ASCA antibody, while many Crohn's patients exhibit ASCA antibodies, and not pANCA antibodies. One method of evaluating Crohn's disease is using the Crohn's disease Activity Index (CDAI), based on 18 predictor variables scores collected by physicians. CDAI values of 150 and below are associated with quiescent disease; values above that indicate active disease, and values above 450 are seen with extremely severe disease [Best et al., "Development of a Crohn's disease activity index." Gastroenterology 70:439-444 (1976)]. However, since the original study, some researchers use a 'subjective value' of 200 to 250 as an healthy score.

Systemic Lupus Erythematosus (SLE) is an autoimmune disease caused by recurrent injuries to blood vessels in multiple organs, including the kidney, skin, and joints. In patients with SLE, a faulty interaction between T cells and B-cells results in the production of autoantibodies that attack the cell nucleus. There is general agreement that autoantibodies are responsible for SLE, so new therapies that deplete the B-cell lineage, allowing the immune system to reset as new B-cells are generated from precursors, would offer hope for long lasting benefit in SLE patients.

Multiple sclerosis (MS) is also an autoimmune disease. It is characterized by inflammation of the central nervous system and destruction of myelin, which insulates

nerve cell fibers in the brain, spinal cord, and body. Although the cause of MS is unknown, it is widely believed that autoimmune T cells are primary contributors to the pathogenesis of the disease. However, high levels of antibodies are present in the cerebral spinal fluid of patients with MS, and some theories predict that the B-cell response leading to antibody production is important for mediating the disease.

Autoimmune thyroid disease results from the production of autoantibodies that either stimulate the thyroid to cause hyperthyroidism (Graves' disease) or destroy the thyroid to cause hypothyroidism (Hashimoto's thyroiditis). Stimulation of the thyroid is caused by autoantibodies that bind and activate the thyroid stimulating hormone (TSH) receptor. Destruction of the thyroid is caused by autoantibodies that react with other thyroid antigens.

Additional diseases, disorders, and conditions amenable to the benefits provided by the compositions and methods of the invention include Sjogren's syndrome is an autoimmune disease characterized by destruction of the body's moisture-producing glands. Further, immune thrombocytopenic purpura (ITP) is caused by autoantibodies that bind to blood platelets and cause their destruction, and this condition is suitable for application of the materials and methods of the invention. Myasthenia Gravis (MG), a chronic autoimmune neuromuscular disorder characterized by autoantibodies that bind to acetylcholine receptors expressed at neuromuscular junctions leading to weakness of the voluntary muscle groups, is a disease having symptoms that are treatable using the composition and methods of the invention, and it is expected that the invention will be beneficial in treating and/or preventing MG. Still further, Rous Sarcoma Virus infections are expected to be amenable to treatment, or amelioration of at least one symptom, with the compositions and methods of the invention.

Another aspect of the present invention is using the materials and methods of the invention to prevent and/or treat any hyperproliferative condition of the skin including psoriasis and contact dermatitis or other hyperproliferative disease. Psoriasis, is characterized by autoimmune inflammation in the skin and is also associated with arthritis in 30% of cases, as well as scleroderma, inflammatory bowel disease, including Crohn's disease and ulcerative colitis. It has been demonstrated that patients with psoriasis and contact dermatitis have elevated antigen activity

within these lesions (Ogoshi et al., J. Inv. Dermatol., 110:818-23 [1998]). The multispecific, multivalent binding proteins can deliver a cytotoxic cell of the immune system, for example, directly to cells within the lesions expressing high levels of antigen. The multivalent, e.g., multispecific, binding proteins can be administered subcutaneously in the vicinity of the lesions, or by using any of the various routes of administration described herein and others which are well known to those of skill in the art.

Also contemplated is the treatment of idiopathic inflammatory myopathy (IIM), including dermatomyositis (DM) and polymyositis (PM). Inflammatory myopathies have been categorized using a number of classification schemes. Miller's classification schema (Miller, Rheum Dis Clin North Am. 20:811-826, 1994) identifies 2 idiopathic inflammatory myopathies (IIM), polymyositis (PM) and dermatomyositis (DM).

Polymyositis and dermatomyositis are chronic, debilitating inflammatory diseases that involve muscle and, in the case of DM, skin. These disorders are rare, with a reported annual incidence of approximately 5 to 10 cases per million adults and 0.6 to 3.2 cases per million children per year in the United States (Targoff, Curr Probl Dermatol. 1991, 3:131-180). Idiopathic inflammatory myopathy is associated with significant morbidity and mortality, with up to half of affected adults noted to have suffered significant impairment (Gottdiener et al., Am J Cardiol. 1978, 41:1141-49). Miller (Rheum Dis Clin North Am. 1994, 20:811-826 and Arthritis and Allied Conditions, Ch. 75, Eds. Koopman and Moreland, Lippincott Williams and Wilkins, 2005) sets out five groups of criteria used to diagnose IIM, i.e., Idiopathic Inflammatory Myopathy Criteria (IIMC) assessment, including muscle weakness, muscle biopsy evidence of degeneration, elevation of serum levels of muscle-associated enzymes, electromagnetic triad of myopathy, evidence of rashes in dermatomyositis, and also includes evidence of autoantibodies as a secondary criteria.

IIM associated factors, including muscle-associated enzymes and autoantibodies include, but are not limited to, creatine kinase (CK), lactate dehydrogenase, aldolase, C-reactive protein, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and antinuclear autoantibody (ANA), myositis-specific antibodies (MSA), and antibody to extractable nuclear antigens.

Preferred autoimmune diseases amenable to the methods of the invention include Crohn's disease, Guillain-Barré syndrome (GBS; also known as acute inflammatory demyelinating polyneuropathy, acute idiopathic polyradiculoneuritis, acute idiopathic polyneuritis and Landry's ascending paralysis), lupus erythematosus, multiple sclerosis, myasthenia gravis, optic neuritis, psoriasis, rheumatoid arthritis, hyperthyroidism (e.g., Graves' disease), hypothyroidism (e.g., Hashimoto's disease), Ord's thyroiditis (a thyroiditis similar to Hashimoto's disease), diabetes mellitus (type 1), aplastic anemia, Reiter's syndrome, autoimmune hepatitis, primary biliary cirrhosis, antiphospholipid antibody syndrome (APS), opsoclonus myoclonus syndrome (OMS), temporal arteritis (also known as "giant cell arteritis"), acute disseminated encephalomyelitis (ADEM), Goodpasture's syndrome, Wegener's granulomatosis, coeliac disease, pemphigus, canine polyarthritis, warm autoimmune hemolytic anemia. In addition, the invention contemplates methods for the treatment, or amelioration of a symptom associated with, the following diseases, endometriosis, interstitial cystitis, neuromyotonia, scleroderma, vitiligo, vulvodynia, Chagas' disease leading to Chagasic cardiopathy (cardiomegaly), sarcoidosis, chronic fatigue syndrome, and dysautonomia.

The complement system is believed to play a role in many diseases with an immune component, such as Alzheimer's disease, asthma, lupus erythematosus, various forms of arthritis, autoimmune heart disease and multiple sclerosis, all of which are contemplated as diseases, disorders or conditions amenable to treatment or symptom amelioration using the methods according to the invention.

Certain constant sub-regions are preferred, depending on the particular effector function or functions to be exhibited by a multivalent single-chain binding molecule. For example, IgG (IgG1, 2, or 3) and IgM are preferred for complement activation, IgG of any subtype is preferred for opsonization and toxin neutralization; IgA is preferred for pathogen binding; and IgE for binding of such parasites as worms.

By way of example, F_CRs recognizing the constant region of IgG antibodies have been found on human leukocytes as three distinct types of Fcγ receptors, which are distinguishable by structural and functional properties, as well as by antigenic

structures detected by CD monoclonal antibodies. They are known as FcγRI, FcγRII, and FcγRIII, and are differentially expressed on (overlapping) subsets of leukocytes.

FcgRI (CD64), a high-affinity receptor expressed on monocytes, macrophages, neutrophils, myeloid precursors and dendritic cells, comprised isoforms la and lb. FcgRI has a high affinity for monomeric human IgG1 and IgG3. Its affinity for IgG4 is about 10 times lower, while it does not bind IgG2. FcgRI does not show genetic polymorphism.

FcγRII (CD32), comprised of isoforms lla, llb1, llb2, llb3 and llc, is the most widely distributed human FcγR type, being expressed on most types of blood leukocytes, as well as on Langerhans cells, dendritic cells and platelets. FcγRII is a low-affinity receptor that only binds aggregated IgG. It is the only FcγR class able to bind IgG2. FcγRIIa shows genetics polymorphism, resulting in two distinct allotypes, FcγRlla-H131 and FcγRlla-R131, respectively. This functional polymorphism is attributable to a single amino acid difference: a histidine (H) or an arginine (R) residue at position 131, which is critical for IgG binding. FcγRlla readily binds human IgG and IgG3 and appears not to bind IgG4. The FcγRlla-H131 has a much higher affinity for complexed IgG2 than the FcγRlla-R131 allotype.

FcγRIII (CD16) has two isoforms or allelotypes, both of which are able to bind IgG1 and IgG3. The FcγRIIa, with an intermediate affinity for IgG, is expressed on macrophages, monocytes, natural killer (NK) cells and subsets of T cells. FcγRIIIb is a low-affinity receptor for IgG, selectively expressed on neutrophils. It is a highly mobile receptor with efficient collaboration with other membrane receptors. Studies with myeloma IgG dimers have shown that only IgG1 and IgG3 bind to FcγRIIIb (with low affinity), while no binding of IgG2 and IgG4 has been found. The FcγRIIIb bears a co-dominant, bi-allelic polymorphism, the allotypes being designated NA1 (Neutrophil Antigen) and NA2.

Yet another aspect of the invention is use of the materials and methods of the invention to combat, by treating, preventing or mitigating the effects of, infection, resulting from any of a wide variety of infectious agents. The multivalent, multispecific binding molecules of the invention are designed to efficiently and effectively recruit the host organism's immune system to resist infection arising from

a foreign organism, a foreign cell, a foreign virus or a foreign inanimate object. For example, a multispecific binding molecule may have one binding domain that specifically binds to a target on an infectious agent and another binding domain that specifically binds to a target on an Antigen Presenting Cell, such as CD 40, CD80, CD86, DC-SIGN, DEC-205, CD83, and the like). Alternatively, each binding domain of a multivalent binding molecule may specifically bind to an infectious agent, thereby more effectively neutralizing the agent. In addition, the invention contemplates multispecific, multivalent binding molecules that specifically bind to a target on an infectious agent and to a non-cell-associated binding partner, which may be effective in conjunction with an effector function of the multispecific binding molecule in treating or preventing infection arising from an infectious agent.

Infectious cells contemplated by the invention include any known infectious cell, including but not limited to any of a variety of bacteria (e.g., pathogenic *E. coli, S. typhimurium, P. aeruginosa, B. anthracis, C. botulinum, C. difficile, C. perfringens, H. pylori, V. cholerae*, and the like), mycobacteria, mycoplasma, fungi (including yeast and molds), and parasites (including any known parasitic member of the Protozoa, Trematoda, Cestoda and Nematoda). Infectious viruses include, but are not limited to, eukaryotic viruses (e.g., adenovirus, bunyavirus, herpesvirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retroviruses, and the like) as well as bacteriophage. Foreign objects include objects entering an organism, preferably a human, regardless of mode of entry and regardless of whether harm is intended. In view of the increasing prevalence of multi-drug-resistant infectious agents (e.g., bacteria), particularly as the causative agents of nosocomial infection, the materials and methods of the invention, providing an approach to treatment that avoids the difficulties imposed by increasing antibiotic resistance.

Diseases, conditions or disorders associated with infectious agents and amenable to treatment (prophylactic or therapeutic) with the materials and methods disclosed herein include, but are not limited to, anthrax, aspergillosis, bacterial meningitis, bacterial pneumoniae (e.g., chlamydia pneumoniae), blastomycosis, botulism, brucellosis, candidiasis, cholera, ciccidioidomycosis, cryptococcosis, diahhreagenic, enterohemorrhagic or enterotoxigenic *E. coli*, diphtheria, glanders, histoplasmosis, legionellosis, leprosy, listeriosis, nocardiosis, pertussis, salmonellosis,

scarlet fever, sporotrichosis, strep throat, toxic shock syndrome, traveler's diarrhea, and typhoid fever.

Additional aspects and details of the invention will be apparent from the following examples, which are intended to be illustrative rather than limiting. Example 1 describes recombinant cloning of immunoglobulin heavy and light chain variable regions. Example 2 describes the construction of Small Modular ImmunoPharmaceuticals. Example 3 describes the construction of a prototype cassette for a multivalent binding protein with effector function. Example 4 describes binding and expression studies with this initial prototype molecule. Example 5 describes construction of alternative constructs derived from this initial prototype molecule where the sequence of the linker region between the EFD and BD2 was changed in both length and sequence. In addition, it describes alternative forms where the orientation of V regions in binding domain 2 were also altered. Example 6 describes subsequent binding and functional studies on these alternative constructs with variant linker forms, identifying a cleavage in the linker region in several of these derivative forms, and the new sequence variants developed to address this problem. Example 7 describes the construction of an alternative preferred embodiment of the multispecific, multivalent fusion proteins, where both BD1 and BD2 bind to antigens on the same cell type (CD20 and CD37), or another multispecific fusion protein where the antigen binding specificity for BD2 has been changed to human CD3 instead of CD28. Example 8 describes the binding and functional studies performed with the CD20-hIgG-CD37 multispecific constructs. Example 9 describes the binding and functional studies with the CD20-hIgG-CD3 multivalent fusion protein constructs. Example 10 discloses multivalent binding molecules having linkers based on specific regions of the extracellular domains of members of the immunoglobulin superfamily. Example 11 discloses assays for identifying binding domains expected to be effective in multivalent binding molecules in achieving at least one beneficial effect identified as being associated with such molecules (e.g., disease treatment).

Example 1

Cloning of Immunoglobulin Heavy and Light Chain Variable Regions

Any methods known in the art can be used to elicit antibodies to a given antigenic target. Further, any methods known in the art can be used to clone the immunoglobulin light and/or heavy chain variable regions, as well as the constant sub-region of an antibody or antibodies. The following method provides an exemplary cloning method.

A. Isolation of Total RNA

To clone the immunoglobulin heavy and light chain variable regions, or the constant sub-region, total RNA is isolated from hybridoma cells secreting the appropriate antibody. Cells ($2x10^7$) from the hybridoma cell line are washed with 1x PBS and pelleted via centrifugation in a 12 x 75 mm round bottom polypropylene tube (Falcon no. 2059). TRIzol™ Total RNA Isolation Reagent (Gibco BRL, Life Technologies, Cat no. 15596-018) is added (8 ml) to each tube and the cells are lysed via repeated pipetting. The lysate is incubated for 5 minutes at room temperature prior to the addition of 1.6 ml (0.2 x volume) of chloroform and vigorous shaking for 15 seconds. After standing 3 minutes at room temperature, the lysates are centrifuged at 9,000 rpm for 15 minutes in a 4°C pre-chilled Beckman JA-17 rotor in order to separate the aqueous and organic phases. The top aqueous phase (about 4.8 ml) is transferred into a new tube and mixed gently with 4 ml of isopropanol. After a 10 minute incubation at room temperature, the RNA is precipitated by centrifugation at 9,000 rpm in a 4°C JA-17 rotor for 11 minutes. The RNA pellet is washed with 8 ml of ice-cold 75% ethanol and re-pelleting by centrifugation at 7,000 x rpm for 7 minutes in a JA-17 rotor at 4°C. The ethanol wash is decanted and the RNA pellets are air-dried for 10 minutes. The RNA pellets are resuspended in 150 µl of diethylpyrocarbonate (DEPC)-treated ddH_2O containing 1 µl of RNase Inhibitor (Catalog No. 799017; Boehringer Mannheim/Roche) per 1 ml of DEPC-treated ddH_2O. The pellets are resuspended by gentle pipetting and are incubated for 20 minutes at 55°C. RNA samples are quantitated by measuring the $OD_{260\ nm}$ of diluted aliquots (1.0 $OD_{260\ nm}$ unit = 40 µg/ml RNA).

B. Rapid Amplification of cDNA Ends

5' RACE is carried out to amplify the ends of the heavy and light chain variable regions, or the constant sub-region. The 5' RACE System for Rapid Amplification of cDNA Ends Kit version 2.0 (Life Technologies, cat. no. 18374-058) is used according to the manufacturer's instructions. Degenerate 5' RACE oligonucleotide primers are designed to match, e.g., the constant regions of two common classes of mouse immunoglobulin heavy chains (IgG1 and IgG2b) using the oligonucleotide design program Oligo version 5.1 (Molecular Biology Insights, Cascade CO). Primers are also designed to match the constant region of the mouse IgG kappa light chain. This is the only class of immunoglobulin light chain, so no degeneracy is needed in the primer design. The sequences of the primers are as follows:

Name	Sequence	SEQ ID NO
Heavy Chain GSP1		
	5'AGGTGCTGGAGGGGACAGTCACTGAGCTGC3'	7
Nested Heavy Chain		
	5'GTCACWGTCACTGRCTCAGGGAARTAGC3'	8
	(W = A or T; R = A or G)	
Light Chain GSP1		
	5'GGGTGCTGCTCATGCTGTAGGTGCTGTCTTTGC3'	9
Nested Light Chain	5'CAAGAAGCACACGACTG	
	AGGCACCTCCAGATG3'	10
5' Race Abridged Anchor Primer		
	5'GGCCACGCGTCGACTAGTACGG	

GNNGGGNNGGGNNG3' 11

To amplify the mouse immunoglobulin heavy chain component, the reverse transcriptase reaction is carried in a 0.2 ml thin-walled PCR tube containing 2.5 pmoles of heavy chain GSP1 primer (SEQ ID NO: 7), 4 μg of total RNA isolated from a suitable hybridoma clone (e.g., either clone 4A5 or clone 4B5), and 12 μl of DEPC treated ddH2O. Likewise, for the mouse light chain component, the reverse transcriptase reaction is carried out in a 0.2 ml thin-walled PCR tube containing 2.5 pmoles of a light chain GSP1 primer (SEQ ID NO: 9), 4 μg of total RNA from a suitable hybridoma clone (e.g., either clone 4A5 or clone 4B5), and 12 μl of DEPC treated ddH2O.

The reactions are carried out in a PTC-100 programmable thermal cycler (MJ research Inc., Waltham, MA). The mixture is incubated at 70°C for 10 minutes to denature the RNA and then chilled on wet ice for 1 minute. The tubes are centrifuged briefly in order to collect moisture from the lids of the tubes. Subsequently, the following components are added to the reaction: 2.5 μl of 10x PCR buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl), 2.5 μl of 25 mM $MgCl_2$, 1 μl of 10 mM dNTP mix, and 2.5 μl of 0.1 M DTT. After mixing each tube by gentle pipetting, the tubes are placed in a PTC-100 thermocycler at 42°C for 1 minute to pre-warm the mix. Subsequently, 1 μl (200 units) of SuperScript™ II Reverse Transcriptase (Gibco-BRL; cat no. 18089-011) is added to each tube, gently mixed by pipetting, and incubated for 45 minutes at 42°C. The reactions are cycled to 70°C for 15 minutes to terminate the reaction, and then cycled to 37°C. RNase mix (1 μl) is then added to each reaction tube, gently mixed, and incubated at 37°C for 30 minutes.

The first-strand cDNA generated by the reverse transcriptase reaction is purified with the GlassMAX DNA Isolation Spin Cartridge (Gibco-BRL) according to the manufacturer's instructions. To each first-strand reaction, 120 μl of 6 M NaI binding solution is added. The cDNA/NaI solution is then transferred into a GlassMAX spin cartridge and centrifuged for 20 seconds at 13,000 x g. The cartridge inserts are carefully removed and the flow-through is discarded from the tubes. The spin cartridges are then placed back into the empty tubes and 0.4 ml of cold (4°C) 1x

wash buffer is added to each spin cartridge. The tubes are centrifuged at 13,000 x g for 20 seconds and the flow-through is discarded. This wash step is repeated three additional times. The GlassMAX cartridges are then washed 4 times with 0.4 ml of cold (4°C) 70% ethanol. After the flow-through from the final 70% ethanol wash is discarded, the cartridges are placed back in the tubes and centrifuged at 13,000 x g for an additional 1 minute in order to completely dry the cartridges. The spin cartridge inserts are then transferred to a fresh sample recovery tube where 50 µl of 65°C (pre-heated) DEPC-treated ddH_2O is quickly added to each spin cartridge. The cartridges are centrifuged at 13,000 x g for 30 seconds to elute the cDNA.

C. Terminal Deoxynucleotidyl Transferase (TdT) Tailing

For each first-strand cDNA sample, the following components are added to a 0.2 ml thin-walled PCR tube: 6.5 µl of DEPC-treated ddH_2O, 5.0 µl of 5x tailing buffer, 2.5 µl of 2 mM dCTP, and 10 µl of the appropriate GlassMAX-purified cDNA sample. Each 24 µl reaction is incubated 2-3 minutes in a thermal cycler at 94°C to denature the DNA, and chilled on wet ice for 1 minute. The contents of the tube are collected by brief centrifugation. Subsequently, 1 µl of terminal deoxynucleotidyl transferase (TdT) is added to each tube. The tubes are mixed via gentle pipetting and incubated for 10 minutes at 37°C in a PTC-100 thermal cycler. Following this 10 minute incubation, the TdT is heat inactivated by cycling to 65°C for 10 minutes. The reactions are cooled on ice and the TdT-tailed first-strand cDNA is stored at –20°C.

D. PCR of dC-tailed First-Strand cDNA

Duplicate PCR amplifications (two independent PCR reactions for each dC-tailed first-strand cDNA sample) are performed in a 50 µl volume containing 200 µM dNTPs, 0.4 µM of 5' RACE Abridged Anchor Primer (SEQ ID NO: 11), and 0.4 µM of either Nested Heavy Chain GSP2 (SEQ ID NO: 8) or Nested Light Chain GSP2 (SEQ ID NO: 10), 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl, 5 µl of dC-tailed cDNA, and 5 units of Expand™ Hi-Fi DNA polymerase (Roche/ Boehringer Mannheim GmbH, Germany). The PCR reactions are amplified using a "Touch-down/Touch-up" annealing temperature protocol in a PTC-100 programmable thermal cycler (MJ Research Inc.) with the following conditions: initial denaturation of 95°C for 40 seconds, 5 cycles at 94°C for 20 seconds, 61°C - 2°C/cycle for 20 seconds, 72°C for 40 seconds + 1 second/cycle, followed by 5 cycles

at 94°C for 25 seconds, 53°C + 1°C/cycle for 20 seconds, 72°C for 46 seconds + 1 second/cycle, followed by 20 cycles at 94°C for 25 seconds, 55°C for 20 seconds, 72°C for 51 seconds + 1 second/cycle, and a final incubation of 72°C for 5 minutes.

E. TOPO TA-Cloning

The resulting PCR products are gel-purified from a 1.0% agarose gel using the QIAQuick Gel purification system (QIAGEN Inc., Chatsworth, CA), TA-cloned into pCR2.1 using the TOPO TA Cloning® kit (Invitrogen, San Diego, CA, cat. no. K4550-40), and transformed into *E. coli* TOP10F' cells (Invitrogen), according to manufacturers' instructions. Clones with inserts are identified by blue/white screening according to the manufacturer's instructions, where white clones are considered positive clones. Cultures of 3.5 ml liquid Luria Broth (LB) containing 50 μg/ml ampicillin are inoculated with white colonies and grown at 37°C overnight (about 16 hours) with shaking at 225 rpm.

The QIAGEN Plasmid Miniprep Kit (QIAGEN Inc., cat. no. 12125) is used to purify plasmid DNA from the cultures according to the manufacturer's instructions. The plasmid DNA is suspended in 34 μl of 1x TE buffer (pH 8.0) and then positive clones sequenced as previously described by fluorescent dideoxy nucleotide sequencing and automated detection using ABI Big Dye Terminator 3.1 reagents at 1:4-1:8 dilutions and analyzed using an ABI 3100 DNA sequencer. Sequencing primers used include T7 (5'GTAATACGACTCACTATAGG3'; SEQ ID NO: 12) and M13 Reverse (5'CAGGAAACAGCTATGACC3'; SEQ ID NO: 13) primers. Sequencing results will confirm that the clones correspond to mouse IgG sequences.

F. De novo gene synthesis using overlapping oligonucleotide extension PCR

This method involves the use of overlapping oligonucleotide primers and PCR using either a high fidelity DNA polymerase or a mix of polymerases to synthesize an immunoglobulin V-region or other gene. Starting at the middle of the V-region sequence, 40-50 base primers are designed such that the growing chain is extended by 20-30 bases, in either direction, and contiguous primers overlap by a minimum of 20 bases. Each PCR step requires two primers, one priming on the anti-sense strand (forward or sense primer) and one priming on the sense strand (reverse or anti-sense primer) to create a growing double-stranded PCR product. During primer design,

changes can be made in the nucleotide sequence of the final product to create restriction enzyme sites, destroy existing restriction enzyme sites, add flexible linkers, change, delete or insert bases that alter the amino acid sequence, optimize the overall DNA sequence to enhance primer synthesis and conform to codon usage rules for the organism contemplated for use in expressing the synthetic gene.

Primer pairs are combined and diluted such that the first pair are at 5 μM an each subsequent pair has a 2-fold greater concentration up to 80 μM. One μL from each of these primer mixes is amplified in a 50 μL PCR reaction using Platinum PCR SuperMix-High Fidelity (Invitrogen, San Diego, CA, cat. no. 12532-016). After a 2-minute initial denaturation at 94°C, 30 cycles of PCR are performed using a cycling profile of 94°C for 20 seconds, 60°C for 10 seconds; and 68°C for 15 seconds. PCR products are purified using Qiaquick PCR Purification columns (Qiagen Inc., cat. no. 28704) to remove excess primers and enzyme. This PCR product is then reamplified with the next set of similarly diluted primer pairs using PCR conditions exactly as described above, but increasing the extension time of each cycle to 68°C for 30 seconds. The resultant PCR product is again purified from primers and enzymes as described above and TOPO-TA cloned and sequenced exactly as described in section E above.

Example 2

Construction of Small Modular ImmunoPharmaceuticals (SMIPs)

A multispecific, multivalent binding protein with effector function was constructed that contained a binding domain 1 in the form of a single-chain recombinant (murine/human) scFv designated 2H7 (VL-linker-VH). The scFv 2H7 is a small modular immunopharmacaceutical (SMIP) that specifically recognizes CD20. The binding domain was based on a publicly available human CD20 antibody sequence GenBank Accession Numbers, M17953 for VH, and M17954 for VL. CD20-specific SMIPs are described in co-owned US Patent Publications 2003/133939, 2003/0118592 and 2005/0136049, incorporated herein in their entireties by reference. The peptide linker separating VL and VH was a 15-amino acid linker encoding the sequence: Asp-Gly_3Ser-$(Gly_4Ser)_2$. Binding domain 1 was located at

the N-terminus of the multispecific binding protein, with the C-terminus of that domain linked directly to the N-terminus of a constant sub-region containing a hinge, C_{H2} and C_{H3} domains (in amino-to-carboxy orientation). The constant sub-region was derived from an IgG1 antibody, which was isolated by PCR amplification of human IgG1 from human PBMCs. The hinge region was modified by substituting three Ser residues in place of the three Cys residues present in the wild type version of the human IgG1 hinge domain, encoded by the 15 amino acid sequence: EPKS**C**DKTHT**C**PP**C**P (SEQ ID NO: 14; the three Cys residues replaced by Ser residues are indicated in bold). In alternative embodiments, the hinge region was modified at one or more of the cysteines, so that SSS and CSC type hinges were generated. In addition, the final proline was sometimes substituted with a serine as well as the cysteine substitutions.

The C-terminal end of the C_{H3} domain was covalently attached to a series of alternative linker domains juxtaposed between the constant sub-region C-terminus and the amino terminus of binding domain 2. Preferred multivalent binding proteins with effector function will have one of these linkers to space the constant sub-region from binding domain 2, although the linker is not an essential component of the compositions according to the invention, depending on the folding properties of BD2. For some specific multivalent molecules, the linker might be important for separation of domains, while for others it may be less important. The linker was attached to the N-terminal end of scFv 2E12 ((V_H-linker-V_L), which specifically recognizes CD28. The linker separating the VH and VL domains of the scFv 2E12 part of the multivalent binding molecule was a 20-amino acid linker $(Gly_4Ser)_4$, rather than the standard $(Gy_4Ser)_3$ linker usually inserted between V domains of an scFv. The longer linker was observed to improve the binding properties of the 2e12 scFv in the VH-VL orientation.

The multispecific, multivalent binding molecule as constructed contained a binding domain 1, which comprises the 2E12 leader peptide sequence from amino acids 1-23 of SEQ ID NO: 171; the 2H7 murine anti-human CD20 light chain variable region, which is reflected at position 24 in SEQ ID NO: 171; an Asp-Gly_3-Ser-$(Gly_4Ser)_2$ linker, beginning at residue 130 in SEQ ID NO: 171, the 2H7 murine anti-human CD20 heavy chain variable region with a leucine to serine (VHL11S) amino

acid substitution at residue 11 in the variable domain for VH, and which has a single serine residue at the end of the heavy chain region (i.e., VTVS where a canonical sequence would be VTVSS) (Genbank Acc. No. M17953), and interposed between the two binding domains BD1 (2H7) and BD2 (2E12) is a human IgG1 constant sub-region, including a modified hinge region comprising a "CSC" or an "SSS" sequence, and wild-type C_{H2} and C_{H3} domains. The nucleotide and amino acid sequences of the multivalent binding protein with effector function are set out in SEQ ID NOS: 228 and 229 for the CSC forms, respectively and SEQ ID NOS: 170 and 171, for the SSS forms.

Stably expressing cell lines were created by transfection via electroporation of either uncut or linearized, recombinant expression plasmid into Chinese hamster ovary cells (CHO DG44 cells) followed by selection in methotrexate containing medium. Bulk cultures and master wells producing the highest level of multivalent binding protein were amplified in increasing levels of methotrexate, and adapted cultures were subsequently cloned by limiting dilution. Transfected CHO cells producing the multivalent binding protein were cultured in bioreactors or wave bags using serum-free medium obtained from JRH Biosciences (Excell 302, cat. no. 14324-1000M, supplemented with 4 mM glutamine (Invitrogen, 25030-081), sodium pyruvate (Invitrogen 11360-070, diluted to 1X), non-essential amino acids (Invitrogen, 11140-050, final dilution to 1X), penicillin-streptromycin 100 IU/ml (Invitrogen, 15140-122), and recombulin insulin at 1 μg/ml (Invitrogen, 97-503311). Other serum free CHO basal medias may also be used for production, such as CD-CHO, and the like.

Fusion protein was purified from spent CHO culture supernatants by Protein A affinity chromatography. The multivalent binding protein was purified using a series of chromatography and filtration steps, including a virus reduction filter. Cell culture supernatants were filtered, then subjected to protein A Sepharose affinity chromatography over a GE Healthcare XK 16/40 column. After binding of protein to the column, the column was washed in dPBS, then 1.0 M NaCl, 20 mM sodium phosphate pH 6.0, and then 25 mM NaCl, 25 mN NaOAc, pH 5.0 to remove nonspecific binding proteins. Bound protein was eluted from the column in 100 mM Glycine (Sigma), pH 3.5, and brought to pH 5.0 with 0.5 M 2-(N-Morpholino) ethanesulfonic acid (MES), pH 6.0. Protein samples were concentrated to 25 mg/ml

in preparation for GPC purification. Size exclusion chromatography was performed on a GE Healthcare AKTA Explorer 100 Air apparatus, using a GE healthcare XK column and Superdex 200 preparative grade (GE healthcare).

The material was then concentrated and formulated with 20 mM sodium phosphate and 240 mM sucrose, with a resulting pH of 6.0. The composition was filtered before filling into sterile vials at various concentrations, depending on the amount of material recovered.

Example 3

Construction of Scorpion Expression Cassette

A nucleic acid containing the synthetic 2H7 scFv (anti-CD20; SEQ ID NO: 1) linked to a constant sub-region as described in Example 2 has been designated TRU-015. TRU-015 nucleic acid, as well as synthetic scFv 2E12 (anti-CD28 VL-VH; SEQ ID NO: 3) and synthetic scFv 2E12 (anti-CD28 VH-VL; SEQ ID NO: 5) nucleic acids encoding small modular immunopharmaceuticals, were used as templates for PCR amplification of the various components of the scorpion cassettes The template, or scaffold, for binding domain 1 and the constant sub-region was provided by TRU-015 (the nucleic acid encoding scFv 2H7 (anti-CD20) linked to the constant sub-region) and this template was constructed in the expression vector pD18. The above-noted nucleic acids containing scFv 2E12 in either of two orientations (V_L-V_H and V_H-V_L) provided the coding region for binding domain 2.

TRU 015 SSS hinge $C_{H2}C_{H3}$ for BD2/Linker Insertion

A version of the synthetic 2H7 scFv IgG1 containing the SSS hinge was used to create a scorpion cassette by serving as the template for addition of an EcoRI site to replace the existing stop codon and XbaI site. This molecule was amplified by PCR using primer 9 (SEQ ID NO: 23; see Table 1) and primer 87 (SEQ ID NO: 40; see Table 1) as well as a Platinum PCR High Fidelity mix (Invitrogen). The resultant 1.5 Kbp fragment was purified and cloned into the vector pCR2.1-TOPO (Invitrogen), transformed into *E. coli* strain TOP10 (Invitrogen), and the DNA sequence verified.

Table 1

No.	Name	Sequence 5'-3'	SEQ ID NO.
	PCR Primers		
1	hVK3L-F3H3	GCGATAAAGCTTGCCGCCATGGAA GCACCAGCGCAGCTTCTCTTCC	15
2	hVK3L-F2	ACCAGCGCAGCTTCTCTTCCTCCTG CTACTCTGGCTCCCAGATACCACCG	16
3	hVK3L-F1-2H7VL	GGCTCCCAGATACCACCGGTCAAAT TGTTCTCTCCCAGTCTCCAG	17
4	2H7VH-NheF	GCGATAGCTAGCCAGGCTTATCTAC AGCAGTCTGG	18
5	G4S-NheR	GCGATAGCTAGCCCCACCTCCTCCA GATCCACCACCGCCCGAG	19
6	015VH-XhoR	GCGTACTCGAGGAGACGGTGACCGT GGTCCCTGTG	20
7	G1H-C-XHO	GCAGTCTCGAGCGAGCCCAAATCTTG TGACAAAACTC	21
8	G1H-S-XHO	GCAGTCTCGAGCGAGCCCAAATCTTC TGACAAAACTC	22
9	CH3R-EcoR1	GCGTGAGAATTCTTACCCGGAGACAGG GAGAGGCTC	23
10	G1-XBA-R	GCGACGTCTAGAGTCATTTACCCGGAG ACAGG	24
11	G4SLinkR1-S	AATTATGGTGGCGGTGGCTCGGGCGGT GGTGGATCTGGAGGAGGTGGGAGTGGG	25
12	G4SLinkR1-AS	AATTCCCACTCCCACCTCCTCCAGATCCA CCACCGCCCGAGCCACCGCCACCAT	26
13	2E12VLXbaR	GCGTGTCTAGATTAACGTTTGATTTCCAG CTTGGTG	27
14	2E12VLR1F	GCGATGAATTCTGACATTGTGCTCACCCA ATCTCC	28
15	2E12VHR1F	GCGATGAATTCTCAGGTGCAGCTGAAGGA GTCAG	29
16	2E12VHXbaR	GCGAGTCTAGATTAAGAGGAGACGGTGAC TGAGGTTC	30
17	2e12VHdXbaF1	GGGTCTGGAGTGGCTGGGAATGATATG	31
18	2e12VHdXbaR1	ATTCCCAGCCACTCCAGACCCTTTCCTG	32
19	IgBsrG1F	GAGAACCACAGGTGTACACCCTG	33
20	IgBsrG1R	GCAGGGTGTACACCTGTGGTTCTCG	34
	Sequencing Primers		
82	M13R	CAGGAAACAGCTATGAC	35
83	M13F	GTAAAACGACGGCCAGTG	36
84	T7	GTAATACGACTCACTATAGG	37
85	pD18F-17	AACTAGAGAACCCACTG	38
86	pD18F-20	GCTAACTAGAGAACCCACTG	39
87	pD18F-1	ATACGACTCACTATAGGG	40
88	pD18R-s	GCTCTAGCATTTAGGTGAC	41
89	CH3seqF1	CATGAGGCTCTGCACAAC	

No.	Name	Sequence 5'-3'	SEQ ID NO.
			42
90	CH3seqF2	CCTCTACAGCAAGCTCAC	43
91	CH3seqR1	GGTTCTTGGTCAGCTCATC	44
92	CH3seqR2	GTGAGCTTGCTGTAGAGG	45

Table 1. Oligonucleotide primers used to construct CD20-CD28 scorpion cassette. Primers are separated into 2 groups, PCR and Sequencing. PCR primers were used to construct the cassette and sequencing primers were used to confirm the DNA sequence of all intermediates and final constructs.

n2H7 V_K and human V_{K3} leader sequence fusion

Oligonucleotide-directed PCR mutagenesis was used to introduce an AgeI (ACCGGT) restriction site at the 5' end of the coding region for TRU 015 VK and an Nhe I (GCTAGC) restriction site at the 3' end of the coding region for the (G4S)3 linker using primers 3 and 5 from Table 1. Since primer 3 also encodes the last 6 amino acids of the human VK3 leader (gb:X01668), overlapping PCR was used to sequentially add the N-terminal sequences of the leader including a consensus Kozak box and HinDIII (AAGCTT) restriction site using primers 1, 2 and 5 from Table 1.

n2H7 IgG1 SSS hinge-$C_{H2}C_{H3}$ Construction

Primers 4 and 6 (SEQ ID NOS: 18 and 20, respectively; Table 1) were used to re-amplify the TRU-015 V_H with an NheI site 5' to fuse with the V_K for TRU-015 and an Xho I (5'-CTCGAG-3') site at the 3' end junction with the IgG1 hinge-$C_{H2}C_{H3}$ domains. Likewise, the IgG1 hinge-C_{H2}-C_{H3} region was amplified using primers 8 and 9 from Table 1, introducing a 5' Xho I site and replacing the existing 3' end with an EcoRI (5'-GAATTC-3') site for cloning, and destroying the stop codon to allow translation of Binding Domain 2 attached downstream of the CH3 domain.. This version of the scorpion cassette is distinguished from the previously described cassette by the prefix "n."

In addition to the multivalent binding protein described above, a protein according to the invention may have a binding domain, either binding domain 1 or 2 or both, that corresponds to a single variable region of an immunoglobulin. Exemplary embodiments of this aspect of the invention would include binding

domains corresponding to the V_H domain of a camelid antibody, or a single modified or unmodified V region of another species antibody capable of binding to the target antigen, although any single variable domain is contemplated as useful in the proteins of the invention.

2E12 VL-VH and VH-VL constructions

In order to make the 2E12 scFvs compatible with the cassette, an internal Xba I (5'-TCTAGA-3') site had to be destroyed using overlapping oligonucleotide primers 17 and 18 from Table 1. These two primers in combination with primer pairs 14/16 (VL-VH) or 13/15 (VH-VL) were used to amplify the two oppositely oriented binding domains such that they both carried EcoRI and XbaI sites at their 5' and 3' ends, respectively. Primers 13 and 16 also encode a stop codon (TAA) immediately in front of the Xba I site.

2H7 SSS IgG1 2e12 LH/HL Construction

Effector Domain- Binding Domain 2 Linker addition. (STD linkers - STD1 and STD2)

Complementary primers 11 and 12 from Table 1 were combined, heated to 70°C and slow-cooled to room temperature to allow annealing of the two strands. 5' phosphate groups were added using T4 polynucleotide kinase (Roche) in 1X Ligation buffer with 1mM ATP(Roche) using the manufacturer's protocol. The resulting double-stranded linker was then ligated into the EcoRI site between the coding regions for the IgG1 C_{H3} terminus and the beginning of Binding Domain 2 using T4 DNA ligase (Roche). The resultant DNA constructs were screened for the presence of an EcoRI site at the linker-BD2 junction and the nucleotide sequence GAATTA at the C_{H3}-linker junction. The correct STD 1 linker construct was then re-digested with EcoRI and the linker ligation repeated to produce a molecule that had a linker composed of two (STD 2) identical iterations of the Lx1 sequence. DNA constructs were again screened as above.

Example 4

Expression studies

Expression studies were performed on the nucleic acids described above that encode multivalent binding proteins with effector function. Nucleic acids encoding

multivalent binding proteins were transiently transfected into COS cells and the transfected cells were maintained under well known conditions permissive for heterologous gene expression in these cells. DNA was transiently transfected into COS cells using PEI or DEAE-Dextran as previously described (PEI= Boussif O. et al., PNAS 92: 7297-7301, (1995), incorporated herein by reference; Pollard H. et al., JBC 273: 7507-7511, (1998), incorporated herein by reference). Multiple independent transfections of each new molecule were performed in order to determine the average expression level for each new form. For transfection by PEI, COS cells were plated onto 60 mm tissue culture plates in DMEM/10%FBS medium and incubated overnight so that they would be approximately 90% confluent on the day of transfection. Medium was changed to serum free DMEM containing no antibiotics and incubated for 4 hours. Transfection medium (4ml/plate) contained serum free DMEM with 50 μg PEI and 10-20 ug DNA plasmid of interest. Transfection medium was mixed by vortexing, incubated at room temperature for 15 minutes, and added to plates after aspirating the existing medium. Cultures were incubated for 3-7 days prior to collection of supernatants. Culture supernatants were assayed for protein expression by SDS-PAGE, Western blotting, binding verified by flow cytometry, and function assayed using a variety of assays including ADCC, CDC, and coculture experiments.

SDS-PAGE Analysis and Western Blotting Analysis

Samples were prepared either from crude culture supernatants (usually 30 μl/well) or purified protein aliquots, containing 8 ug protein per well, and 2X Tris-Glycine SDS Buffer (Invitrogen) was added to a 1X final concentration. Ten (10) μl SeeBlue Marker (Invitrogen, Carlsbad, CA) were run to provide MW size standards. The multivalent binding (fusion) protein variants were subjected to SDS-PAGE analysis on 4-20% Novex Tris-glycine gels (Invitrogen, San Diego, CA). Samples were loaded using Novex Tris-glycine SDS sample buffer (2X) under reducing or non-reducing conditions after heating at 95°C for 3 minutes, followed by electrophoresis at 175V for 60 minutes. Electrophoresis was performed using 1X Novex Tris-Glycine SDS Running Buffer (Invitrogen).

After electrophoresis, proteins were transferred to PVDF membranes using a semi-dry electroblotter apparatus (Ellard, Seattle, WA) for 1 hour at 100 mAmp. Western

transfer buffers included the following three buffers present on saturated Whatman filter paper, and stacked in succession: no. 1 contains 36.34 g/liter Tris, pH 10.4, and 20% methanol; no. 2 contains 3.02 g/liter Tris, pH 10.4, and 20% methanol; and no. 3 contains 3.03 g/liter Tris, pH 9.4, 5.25 g/liter ε-amino caproic acid, and 20% methanol. Membranes were blocked in BLOTTO=5% nonfat milk in PBS overnight with agitation. Membranes were incubated with HRP conjugated goat anti-human IgG (Fc specific, Caltag) at 5 ug/ml in BLOTTO for one hour, then washed 3 times for 15 minutes each in PBS-0.5% Tween 20. Wet membranes were incubated with ECL solution for 1 minute, followed by exposure to X-omat film for 20 seconds. Figure 2 shows a Western Blot of proteins expressed in COS cell culture supernatant (30 μl/well) electrophoresed under non-reducing conditions. Lanes are indicated with markers 1-9 and contain the following samples: Lane 1 (cut off= See Blue Markers, kDa are indicated to the side of the blot. Lane 2= 2H7-sssIgG P238S/P331S-STD1-2e12 VLVH; lane 3= 2H7-sssIgG P238S/P331S-STD1-2e12 VHVL, Lane 4=2H7-sssIgG P238S/P331S-STD2-2e12 VLVH; Lane 5=2H7-sssIgG P238S/P331S-STD2-2e12 VHVL; Lane 6=2e12 VLVH SMIP; Lane 7=2e12 VHVL SMIP; Lane 8=2H7 SMIP. 2H7 in these constructs is always in the V_LV_H orientation, sssIgG indicates the identity of the hinge/linker located at linker position 1 as shown in Figure 5, P238S/P331S indicates the version of human IgG1 with mutations from wild type (first aa listed) to mutant (second aa listed) and the amino acid position at which they occur in wild type human IgG1 C_{H2} and CH3 domains, STD1 indicates the 20-amino-acid (18 + restriction site) linker located in linker position 2 as shown in Figure 5, and STD2 indicates the 38 amino acid (36+restriction site) linker located in linker position 2 as shown in Figure 6.

Binding Studies

Binding studies were performed to assess the bispecific binding properties of the CD20/CD28 multispecific, multivalent binding peptides. Initially, WIL2-S cells were added to 96 well plates and centrifuged to pellet cells. To the seeded plates, CD20/CD28 purified protein was added, using two-fold titrations across the plate from 20 μg/ml down to 0.16 μg/ml. A two-fold dilution series of TRU-015 (source of binding domain 1) purified protein was also added to seeded plate wells, the

concentration of TRU-015 extending from 20 μg/ml down to 0.16 μg/ml. One well containing no protein served as a background control.

Seeded plates containing the proteins were incubated on ice for one hour. Subsequently, the wells were washed once with 200 μl 1% FBS in PBS. Goat anti-human antibody labeled with FITC (Fc Sp) at 1:100 was then added to each well, and the plates were again incubated on ice for one hour. The plates were then washed once with 200 μl 1% FBS in PBS and the cells were re-suspended in 200 μl 1% FBS and analyzed by FACS.

To assess the binding properties of the anti-CD28 peptide 2E12 V_HV_L, CD28-expressing CHO cells were plated by seeding in individual wells of a culture plate. The CD20/CD28 purified protein was then added to individual wells using a two-fold dilution scheme, extending from 20 μg/ml down to 0.16 μg/ml. The 2E12IgG-VHVL SMIP purified protein was added to individual seeded wells, again using a two-fold dilution scheme, i.e., from 20 μg/ml down to 0.16 μg/ml. One well received no protein to provide a background control. The plates were then incubated on ice for one hour, washed once with 200 μl 1% FBS in PBS, and goat anti-human antibody labeled with FITC (Fc Sp, CalTag, Burlingame, CA) at 1:100 was added to each well. The plates were again incubated on ice for one hour and subsequently washed once with 200 μl 1% FBS in PBS. Following re-suspension of the cells in 200 μl 1% FBS, FACS analysis was performed. The results showed that multivalent binding proteins with the N-terminal CD20 binding domain 1 bound CD20; those proteins having the C-terminal CD28 binding domain 2 in the N-V_H-V_L-C orientation also bound CD28.

The expressed proteins were shown to bind to CD20 presented on WIL-2S cells (see Figure 3) and to CD28 presented on CHO cells (refer to Figure 3) by flow cytometry (FACS), thereby demonstrating that either BD1 or BD2 could function to bind the specific target antigen. Each data set on the graphs in Figure 3 shows the binding of serial dilutions of the different multivalent binding (fusion) proteins over the titration ranges indicated. The data obtained using these initial constructs indicate that multivalent binding (fusion) proteins with the binding domain 2 version using 2e12 in the VHVL orientation express better and bind better to CD28 than the form in the VLVH orientation at equivalent concentrations.

Figure 4 shows a graphical presentation of the results of binding studies performed with purified proteins from each of these transfections/constructs. The figure shows binding profiles of the proteins to CD20 expressing WIL-2S cells, demonstrating that the multivalent molecule binds to CD20 as well as the single specificity SMIP at the same concentration. The top and bottom panels for Figure 5 show the binding profiles of the BD2 specificity (2e12=CD28) to CD28 CHO cells. For binding of binding domain 2 to CD28, the orientation of the V regions affected binding of the 2e12. 2H7-sss-hIgG-STD1-2e12 multivalent binding proteins with the 2e12 in the VH-VL (HL) orientation showed binding at a level equivalent to the single specificity SMIP, while the 2e12 LH molecule showed less efficient binding at the same concentration.

Example 5

Construction of Various Linker Forms of the Multivalent Fusion Proteins.

This example describes the construction of the different linker forms listed in the table shown in Figure 6.

Construction of C_{H3}-BD2 linkers H1 through H7

To explore the effect of C_{H3}-BD2 linker length and composition on expression and binding of the scorpion molecules, an experiment was designed to compare the existing molecule 2H7sssIgG1-Lx1-2e12HL to a larger set of similar constructs with different linkers. Using 2H7sssIgG1-Lx1-2e12HL as template, a series of PCR reactions were performed using the primers listed in Oligonucleotide Table 2, which created linkers that varied in length form 0 to 16 amino acids. These linkers were constructed as nucleic acid fragments that spanned the coding region for C_{H3} at the BsrGI site to the end of the nucleic acid encoding the linker-BD2 junction at the EcoRI site.

Table 2

No.	Name	Sequence 5'-3'	SEQ ID NO.
	PCR Primers		
1	L1-11R	GCGATAGAATTCCCAGATCCACCACCGCCCGA GCCACCGCCACCATAATTC	46

2	L1-6R	GCGATAGAATTCCCAGAGCCACCGCCACCATA ATTC	47
3	L3R	GCGTATGAATTCCCCGAGCCACCGCCACCCTTA CCCGGAGACAGG	48
4	L4R	GCGTATGAATTCCCAGATCCACCACCGCCCGAG CCACCGCCACCCTTAC	49
5	L5R	GCGTATGAATTCCCGCTGCCTCCTCCCCCAGATC CACCACCGCC	50
6	IgBsrG1F	GAGAACCACAGGTGTACACCCTG	51
7	L-CPPCPR	GCGATAGAATTCGGACAAGGTGGACACCCCTTAC CCGGAGACAGGGAGAG	52

Table 2. Sequences of primers used to generate CH3-BD2 linker variants.

Figure 6 diagrams the schematic structure of a multivalent binding (fusion) protein and shows the orientation of the V regions for each binding domain, the sequence present at linker position 1 (only the Cys residues are listed), and the sequence and identifier for the linker(s) located at linker position 2 of the molecules.

Example 6

Binding and Functional Studies With Variant Linker Forms of the 2H7-IgG-2e12 Prototype Multivalent Fusion Proteins.

This example shows the results of a series of expression and binding studies on the "prototype" 2H7-sssIgG-Hx-2e12 VHVL construct with various linkers (H1-H7) present in the linker position 2. Each of these proteins was expressed by large-scale COS transient transfection and purification of the molecules using protein A affinity chromatography, as described in the previous examples. Purified proteins were then subjected to analyses including SDS-PAGE, Western blotting, binding studies analyzed by flow cytometry, and functional assays for biological activity.

Binding Studies Comparing the Different BD2 Orientations

Binding studies were performed as described in the previous examples, except that protein A-purified material was used, and a constant amount of binding (fusion) protein was used for each variant studied, i.e., 0.72 ug/ml. Figure 7 shows a columnar graph comparing the binding properties of each linker variant and 2e12 orientation variant to both CD20 and CD28 target cells. H1-H6 refer to constructs with the H1-H6 linkers and 2e12 in the VH-VL orientation. L1-L6 refer to constructs with the H1-H6 linkers and 2e12 in the VL-VH orientation. The data demonstrate that the binding

domain 2 specificity for 2e12 binds much more efficiently when present in the HL orientation (samples H1-H6) than when in the LH orientation (samples L1-L6). The effect of linker length is complicated by the discovery, as shown in the next set of figures, that molecules with the longer linkers contain some single-specificity cleaved molecules which are missing the CD28 binding specificity at the carboxy terminus. Other experiments were performed which address the binding of selected linkers, with the results shown in Figures 10, 12, and 13.

SDS-PAGE Analysis of purified H1-H7 Linker Variants

Samples were prepared from purified protein aliquots, containing 8 μg protein per well, and 2X Tris-Glycine SDS Buffer (Invitrogen) was added to a 1X final concentration. For reduced samples/gels, 10X reducing buffer was added to 1X to samples plus Tris-Glycine SDS buffer. Ten (10) μl SeeBlue Marker (Invitrogen, Carlsbad, CA) was run to provide MW size standards. The multivalent binding (fusion) protein variants were subjected to SDS-PAGE analyses on 4-20% Novex Tris-glycine gels (Invitrogen, San Diego, CA). Samples were loaded using Novex Tris-glycine SDS sample buffer (2X) under reducing or non-reducing conditions after heating at 95°C for 3 minutes, followed by electrophoresis at 175V for 60 minutes. Electrophoresis was performed using 1X Novex Tris-Glycine SDS Running Buffer (Invitrogen). Gels were stained after electrophoresis in Coomassie SDS PAGE R-250 stain for 30 minutes with agitation, and destained for at least one hour. Figure 8 shows the nonreduced and reduced Coomassie stained gels of the [2H7-sss-hIgG P238S/P331S-Hx-2e12 VHVL] multivalent binding (fusion) protein variants, along with TRU-015 and 2e12 HLSMIP as control samples. As the linker length is increased, the amount of protein running at approximately SMIP size (or 52 kDa) increases. The increase in the amount of protein in this band corresponds with a decrease in the amount of protein in the upper band running at about 90 kDa. The gel data indicate that the full-length molecule is being cleaved at or near the linker, to generate a molecule which is missing the BD2 region. A smaller BD2 fragment is not present, indicating (1) that the nucleotide sequence within the linker sequence may be creating a cryptic splice site that removes the smaller fragment from the spliced RNA transcript, or (2) that the protein is proteolytically cleaved after translation of the full-size polypeptide, and that the smaller BD2 fragment is unstable, i.e., susceptible to proteolytic processing. Western blotting of some of these molecules indicates that the

proteins all contain the CD20 BD1 sequence, but the smaller band is missing the CD28 BD2 reactivity. No smaller band migrating at "bare" scFv size (25-27 kDa) was observed on any gels or blots, indicating that this smaller peptide fragment is not present in the samples.

Western Blot Binding of BD1 and BD2 by 2H7 specific Fab or CD28mIg

Figure 9 shows the results of Western blotting of the 2H7-sss-hIgG-H6 multivalent binding (fusion) proteins compared to each single-specificity SMIP.

Electrophoresis was performed under non-reducing conditions, and without boiling samples prior to loading. After electrophoresis, proteins were transferred to PVDF membranes using a semi-dry electroblotter apparatus (Ellard, Seattle, WA) for 1 hour at 100 mAmp. Membranes were blocked in BLOTTO (5% nonfat milk in PBS) overnight with agitation. Figure 9A: Membranes were incubated with the AbyD02429.2, a Fab directed to the 2H7 antibody, at 5 μg/ml in BLOTTO for one hour, then washed 3 times for 5 minutes each in PBS-0.5% Tween 20. Membranes were then incubated in 6X His-HRP for one hour at a concentration of 0.5 μg/ml. Blots were washed three times for 15 minutes each in PBST. Wet membranes were incubated with ECL solution for 1 minute, followed by exposure to X-omat film for 20 seconds.

Figure 9B: Membranes were incubated with CD28Ig (Ancell, Bayport, MN) at 10 μg/ml in BLOTTO, then washed three times for 15 minutes each in PBS-0.5% Tween 20. Membranes were then incubated in goat anti-mouse HRP conjugate (CalTag, Burlingame, CA) at 1:3000 in BLOTTO. Membranes were washed three times, for 15 minutes each, then incubated in ECL solution for 1 minute, followed by exposure to X-omat film for 20 seconds. The results from the Western blots indicated that the CD28 binding domain was present in the multivalent "monomer" fraction migrating at approximately 90 kDa, and in higher order forms. No band was detectable migrating at the position expected for a single SMIP or bare scFv size fragment. When the CD20 anti-idiotype Fab was used, a SMIP-sized fragment was detected, indicating the presence of a peptide fragment containing (2H7-sss-hIgG), and missing the CD28 scFv BD2 portion of the fusion protein.

Binding Studies on Selected Linkers

Figure 10 shows the results of binding studies performed on the purified 2H7-sss-hIgG-Hx-2e12 fusion proteins. Binding studies were performed to assess the bispecific binding properties of the CD20/CD28 multispecific binding peptides. Initially, WIL2-S cells were plated using conventional techniques. To the seeded plates, CD20/CD28 purified protein was added, using two-fold titrations across the plate from 20 μg/ml down to 0.16 μg/ml. A two-fold dilution series of TRU-015 (source of binding domain 1) purified protein was also added to seeded plate wells, the concentration of TRU-015 extending from 20 μg/ml down to 0.16 μg/ml. One well containing no protein served as a background control.

Seeded plates containing the proteins were incubated on ice for one hour. Subsequently, the wells were washed once with 200 μl 1% FBS in PBS. Goat anti-human antibody labeled with FITC (Fc Sp) at 1:100 was then added to each well, and the plates were again incubated on ice for one hour. The plates were then washed once with 200 μl 1% FBS in PBS and the cells were re-suspended in 200 μl 1% FBS and analyzed by FACS.

To assess the binding properties of the anti-CD28 peptide 2E12 V_HV_L, CD28-expressing CHO cells were plated by seeding in individual wells of a culture plate. The CD20/CD28 purified protein was then added to individual wells using a two-fold dilution scheme, extending from 20 μg/ml down to 0.16 μg/ml. The 2E12IgGvHvL SMIP purified protein was added to individual seeded wells, again using a two-fold dilution scheme, i.e., from 20 μg/ml down to 0.16 μg/ml. One well received no protein to provide a background control. The plates were then incubated on ice for one hour, washed once with 200 μl 1% FBS in PBS, and goat anti-human antibody labeled with FITC (Fc Sp) at 1:100 was added to each well. The plates were again incubated on ice for one hour and subsequently washed once with 200 μl 1% FBS in PBS. Following re-suspension of the cells in 200 μl 1% FBS, FACS analysis was performed. The expressed proteins were shown to bind to CD20 presented on WIL-2S cells (see Figure 10A) and to CD28 presented on CHO cells (refer to Figure 10B) by flow cytometry (FACS), thereby demonstrating that either BD1 or BD2 could function to bind the specific target antigen. In addition, the linker used (H1-H6) was not found to significantly affect binding avidity to target antigen.

SEC Fractionation of Multivalent Binding (Fusion) Proteins. The binding (fusion) protein was purified from cell culture supernatants by protein A Sepharose affinity chromatography over a GE Healthcare XK 16/40 column. After binding of protein to the column, the column was washed in dPBS, then 1.0 M NaCl, 20 mM sodium phosphate pH 6.0, and then 25 mM NaCl, 25 mN NaOAc, pH 5.0, to remove nonspecific binding proteins. Bound protein was eluted from the column in 100 mM Glycine (Sigma), pH 3.5, and brought to pH 5.0 with 0.5 M 2-(N-Morpholino) ethanesulfonic acid (MES), pH 6.0. Protein samples were concentrated to 25 mg/ml using conventional techniques in preparation for GPC purification. Size exclusion chromatography (SEC) was performed on a GE Healthcare AKTA Explorer 100 Air apparatus, using a GE healthcare XK column and Superdex 200, preparative grade (GE healthcare).

Figure 12 shows a table summarizing the results of SEC fractionation of the different binding (fusion) proteins. With increasing linker length, the complexity of the molecules in solution also increases, making it difficult to isolate peak of interest, or POI from higher order forms by HPLC. The H7 linker seems to resolve much of this complexity into a more homogeneous form in solution, so that the soluble forms migrate mostly as a single POI at approximately 172 kDa.

Additional Binding Studies

A second series of experiments was performed (see Figures 12 and 13) with a smaller subset of multivalent binding (fusion) proteins, this time comparing linkers H3, H6, and H7. The data demonstrate that the binding level drops more significantly for CD28 than for CD20 binding, but both drop slightly as linker length increases. Further, the data showed that the H7 linker exhibited the highest level of binding to both antigens. These data were obtained using protein A-purified multivalent binding (fusion) proteins, but were not further purified by SEC, so multiple forms of the molecules may have been present in solution. The results also indicated that the truncated form may have been less stable than the true multivalent polypeptide, since the binding curves do not appear to fully reflect the significant amount of single specificity form present in solution for linker H6.

Demonstration of Multispecific Binding From a Single Molecule

An alternative binding assay was performed (see Figure 13), where binding to CD20 on the surface of WIL-2S cells was detected with a reagent specific for the CD28 BD2, thereby demonstrating that simultaneous binding may occur to both target antigens, engaging both BD1 and BD2 on the same multispecific binding (fusion) protein (refer to Figure 12) This assay demonstrates the multispecific binding property of the proteins.

Example 7

Construction of Multispecific Binding (Fusion) Proteins With Alternative Specificities in BD2

In addition to the prototype CD20-CD28 multispecific binding molecule, two other forms were made with alternative binding domain 2 regions, including CD37 and CD3 binding domains. The molecules were also made with several of the linker domains described for the [2H7-sss-IgG-Hx/STDx-2e12 HL] multispecific binding (fusion) proteins. The construction of these additional multispecific binding (fusion) molecules are described below.

Anti-CD37 Binding Domain Construction

Table 3

No.	Name	Sequence	SEQ ID NO.
23	G281LH-NheR	ACTGCTGCAGCTGGACCGCGCT AGCTCCGCCGCCACCCGAC	53
24	G281LH-NheF	GGCGGAGCTAGCGCGGTCCAGC TGCAGCAGTCTGGACCTG	54
25	G281-LH-LPinF	GCGATCACCGGTGACATCCAGAT GACTCAGTCTCCAG	55
26	G281-LH-HXhoR	GCGATACTCGAGGAGACGGTGAC TGAGGTTCCTTGAC	56
27	G281-LH-LEcoF	GCGATCGAATTCAGACATCCAGAT GACTCAGTCTCCAG	57
28	G281-LH-HXbaR	GCGATTCTAGATTAGGAAGAGACG GTGACTGAGGTTCCTTGAC	58
29	G281-HL-HF	GCGATAACCGGTGCGGTCCAGCTG CAGCAGTCTGGAC	59
30	G281-HL-HR3	GACCCACCACCGCCCGAGCCACCG CCACCAGAAGAGACGGTGACTGAGG TTC	60
31	G281-HL-HR2	ACTCCCGCCTCCTCCTGATCCGCCG CCACCCGACCCACCACCGCCCGAG	61
32	G281-HL-HNheR	GAGTCATCTGGATGTCGCTAGCACTC CCGCCTCCTCCTGATC	62
33	G281-HL-LNheF	ATCAGGAGGAGGCGGGAGTGCTAGC	

		GACATCCAGATGACTCAGTC	63
34	G281-HL-LXhoR	GCGATACTCGAGCCTTTGATCTCCAG TTCGGTGCCTC	64
35	G281-HL-LXbaR	GCGATATCTAGACTCAACCTTTGATCT CCAGTTCGGTGCCTC	65
36	G281-HL-EcoF	GCGATAGAATTCGCGGTCCAGCTGCA GCAGTCTGGAC	66

Table 3. Oligonucleotide primers used to generate G28-1 anti-CD37 binding domains for both SMIP molecules and scorpions.

The G28-1 scFv (SEQ ID NO:102) was converted to the G28-1 LH SMIP by PCR using the primers in Table X above. Combining primers 23 and 25 with 10 ng G28-1 scFv, the VK was amplified for 30 cycles of 94C, 20 seconds, 58C, 15 seconds, 68C, 15 seconds using Platinum PCR Supermix Hi-Fidelity PCR mix (Invitrogen, Carlsbad, CA) in an ABI 9700 Thermalcycler. The product of this PCR had the restriction sites PinAI (AgeI) at the 5' end of the VK and NheI at the end of the scFv (G4S)3 linker. The VH was similarly altered by combining primers 24 and 26 with 10 ng G28-1 scFv in a PCR run under the identical conditions as with the VK above. This PCR product had the restriction sites NheI at the 5' end of the VH and XhoI at the 3' end. Because significant sequence identity overlap was engineered into primers 23 and 24, the VK and VH were diluted 5-fold, then added at a 1:1 ratio to a PCR using the flanking primers 25 and 26 and a full-length scFv was amplified as above by lengthening the 68C extension time from 15 seconds to 45 seconds. This PCR product represented the entire G28-1 scFv as a PinAI-XhoI fragment and was purified by MinElute column (Qiagen,) purification to remove excess primers, enzymes and salts. The eluate was digested to completion with PinAI (Invitrogen)and XhoI (Roche) in 1X H buffer (Roche,) at 37C for 4 hours in a volume of 50 μL. The digested PCR product was then electrophoresed in a 1% agarose gel, the fragment was removed from the gel and re-purified on a MinElute column using buffer QG and incubating the gel-buffer mix at 50C for 10 minutes with intermittent mixing to dissolve the agarose after which the purification on the column was identical for primer removal post-PCR. 3 μL PinAI-XhoI digested G28-1 LH was combined with 1 μL PinAI-XhoI digested pD18-n2H7sssIgG1 SMIP in a 10 μL reaction with 5 μL 2X LigaFast Ligation Buffer (Promega, Madison, WI) and 1 μL T4 DNA ligase (Roche), mixed well and incubated at room temperature for 10 minutes. 3 μL of this ligation was then transformed into competent TOP 10 (Invitrogen) using the

manufacturer's protocol. These transformants were plated on LB agar plates with 100 μg/ml carbenicillin(Teknova,) and incubated overnight at 37C. After 18 hours of growth, colonies were picked and inoculated into 1 ml T-Broth (Teknova,) containing 100 μg/ml carbenicillin in a deep well 96-well plate and grown overnight in a 37C shaking incubator. After 18-24 hours of growth, DNA was isolated from each overnight culture using the QIAprep 96 Turbo Kit (Qiagen) on the BioRobot8000 (Qiagen). 10 μL from each clone was then digested with both HindIII and XhoI restriction enzymes in 1X B buffer in a 15 μL reaction volume. The digested DNA was electrophoresed on 1% agarose E-gels (Invitrogen, CA) for restriction site analysis. Clones that contained a HindIII-XhoI fragment of the correct size were sequence verified. The G28-1 HL SMIP was constructed in a similar manner by placing a PinAI site on the 5' end and a (G4S)4 linker ending in an Nhe I site of the G28-1 VH using primers 29, 30 31 and 32 from Table X above. The VK was altered by PCR using primers 33 and 34 from Table X such that an NheI site was introduced at the 5' end of the VK and XhoI at the 3' end. These PCRs were then combined as above and amplified with the flanking primers 29 and 34 to yield an intact G28-1 scFv DNA in the VH-VL orientation which was cloned into PinAI-XhoI digested pD18-(n2H7)sssIgG1 SMIP exactly as with the G28-1 LH SMIP.

2H7sssIgG1-STD1-G28-1 LH/HL Construction

Using the G28-1 LH and G28-1 HL SMIPs as templates, the LH and HL anti-CD37 binding domains were altered by PCR such that their flanking restriction sites were compatible with the scorpion cassette. An EcoRI site was introduced at the 5' end of each scFv using either primer 27 (LH) or 36 (HL) and a stop codon/ XbaI site at the 3' end using either primer 28 (LH) or 35 (HL). The resulting DNAs were cloned into EcoRI-XbaI digested pD18-2H7sssIgG-STD1.

2H7sssIgG1-Hx-G28-1 HL Construction

2H7sssIgG1-Hx-2e12 HL DNAs were digested with BsrGI and EcoRI and the 325 bp fragment consisting of the C-terminal end of the IgG1 and linker. These were substituted for the equivalent region in 2H7sssIgG1-STD1-G19-4 HL by removal of the STD1 linker using BsrGI-EcoRI and replacing it with the corresponding linkers from the 2H7sssIgG1-Hx-2e12 HL clones.

Anti-CD3 Binding domain Construction

Table 4

No.	Name	Sequence	SEQ ID NO.
37	194-LH-LF1	GCGTATGAACCGGTGACATCCAGAT GACACAGACTACATC	67
38	194-LF2	ATCCAGATGACACAGACTACATCCTC CCTGTCTGCCTCTCTGGGAGACAG	68
39	194-LF3	GTCTGCCTCTCTGGGAGACAGAGTCA CCATCAGTTGCAGGGCAAGTCAGGAC	69
40	194-LF4	GTTGCAGGGCAAGTCAGGACATTCGC AATTATTTAAACTGGTATCAGCAG	70
41	194-LF5	ATTTAAACTGGTATCAGCAGAAACCAG ATGGAACTGTTAAACTCCTGATC	71
42	194-LF6	GAACTGTTAAACTCCTGATCTACTACA CATCAAGATTACACTCAGGAGTC	72
43	194-LF7	CAAGATTACACTCAGGAGTCCCATCAA GGTTCAGTGGCAGTGGGTCTGGAAC	73
44	194-LR7	CAGGTTGGCAATGGTGAGAGAATAATC TGTTCCAGACCCACTGCCACTGAAC	74
45	194-LR6	GCAAAAGTAAGTGGCAATATCTTCTGGT TGCAGGTTGGCAATGGTGAGAG	75
46	194-LR5	GAACGTCCACGGAAGCGTATTACCC TGTTGGCAAAAGTAAGTGGCAATATC	76
47	194-LR4	CGTTTGGTTACCAGTTTGGTGCCTCCAC CGAACGTCCACGGAAGCGTATTAC	77
48	194-LR3	ACCACCGCCCGAGCCACCGCCACC CCGTTTGGTTACCAGTTTGGTG	78
49	194-LR2	GCTAGCGCTCCCACCTCCTCCAGATCCA CCACCGCCCGAGCCACCGCCAC	79
50	194-LH-LR1	GTTGCAGCTGGACCTCGCTAGCGCT CCCACCTCCTCCAGATC	80
51	194-LH-HF1	GATCTGGAGGAGGTGGGAGCGCTAGC GAGGTCCAGCTGCAACAGTCTGGACCTG	81
52	194-HF2	AGCTGCAACAGTCTGGACCTGAACT GGTGAAGCCTGGAGCTTCAATGAAG	82
53	194-HF3	AGCCTGGAGCTTCAATGAAGATTTCC TGCAAGGCCTCTGGTTACTCATTC	83
54	194-HF4	GCAAGGCCTCTGGTTACTCATTCACT GGCTACATCGTGAACTGGCTGAAGCAG	84
55	194-HF5	ATCGTGAACTGGCTGAAGCAGAGCC ATGGAAAGAACCTTGAGTGGATTGGAC	85
56	194-HF6	GAACCTTGAGTGGATTGGACTTATTA ATCCATACAAAGGTCTTACTACCTAC	86
57	194-HR6	AATGTGGCCTTGCCCTTGAATTTCTG GTTGTAGGTAGTAAGACCTTTGTATG	87
58	194-HR5	CATGTAGGCTGTGCTGGATGACTTGT CTACAGTTAATGTGGCCTTGCCCTTG	88
59	194-HR4	ACTGCAGAGTCTTCAGATGTCAGACTG AGGAGCTCCATGTAGGCTGTGCTGGATG	89
60	194-HR3	ACCATAGTACCCAGATCTTGCACAG TAATAGACTGCAGAGTCTTCAGATGTC	90
61	194-HR2	GCGCCCCAGACATCGAAGTACCAGTC CGAGTCACCATAGTACCCAGATCTTG	91
62	194-LH-HR1	GCGAATACTCGAGGAGACGGTGACCG TGGTCCCTGCGCCCCAGACATCGAAG	92
63	194-HL-HF1	GCGTATGAACCGGTGAGGTCCAGC	

		TGCAACAGTCTGGACCTG	93
64	194-HL-HR1	ACCGCCACCAGAGGAGACGGTGACCGT GGTCCCTGCGCCCCAGACATCGAAGTAC	94
65	194-HL-HR0	ACCTCCTCCAGATCCACCACCGCCCG AGCCACCGCCACCAGAGGAGACGGTG	95
66	194-HL-LF1	GCGGGGGAGGTGGCAGTGCTAGCGA CATCCAGATGACACAGACTACATC	96
67	194-HL-LR3Xho	GCGAATACTCGAGCGTTTGGTTACCA GTTTGGTG	97
68	194-HL-LR3Xba	GCGATATCTAGATTACCGTTTGGTTAC CAGTTTGGTG	98
69	194-HL-HF1R1	GCGTATGAGAATTCAGAGGTCCAGCTG CAACAGTCTGGACCTG	99
70	194-LH-LF1R1	GCGTATGAGAATTCTGACATCCAGA TGACACAGACTACATC	100
71	194-LH-HR1Xba	GCGTATCTAGATTAGGAGACGGTGACC GTGGTCCCTGCGCCCCAGACATCGAAG	101

Table 4. Oligonucleotides used to generate anti-CD3 binding domain from the G19-4 scFv sequence.

The G19-4 binding domain was synthesized by extension of overlapping oligonucleotide primers as described previously. The light chain PCR was done in two steps, beginning by combining primers 43/44, 42/45, 41/46 and 40/47 at concentrations of 5uM, 10 μM, 20 μM and 40 μM ,respectively, in Platinum PCR Supermix Hi-Fidelity for 30 cycles of 94°C, 20 seconds, 60°C, 10 seconds, 68°C, 15 seconds. 1 μL of the resultant PCR product was reamplified using a primer mix of 39/48 (10 μM), 38/49 (20 μM) and 37/50 (40 μM) for the LH or 66/67 (40 μM) for the HL orientation, using the same PCR conditions with the exception of the 68C extension which was increased to 25 seconds. The VK in the LH orientation was bounded by PinAI at the 5' end and NheI at the 3' end, while the HL orientation had NheI at the 5' end and XhoI at the 3' end.

To synthesize the heavy chain, primer mixes with the same concentrations as above were prepared by combining primers 56/57, 55/58, 54/59 and 53/60 for the first PCR step. In the second PCR, primers 52/61 (20 μM) and 51/62 (50 μM) were amplified with 1 μl from the first PCR using the same PCR conditions as with the second PCR of the light chain to make the LH orientation with NheI at the 5' end and XhoI at the 3' end. Primers 52/61(10 μM), 63/64 (20 μM), 63 (20 μM)/65 (40 μM) and 63(20 μM)/5 (80 μM) were combined in a second PCR with 1uL from the previous PCR to create the heavy chain in the HL orientation with PinAI at the 5' end and NheI at the 3' end. As with previous constructs, sufficient overlap was designed

into the primers centered around the NheI site such that the G19-4 LH was synthesized by combining the heavy and light chain PCRs in the LH orientation and reamplifying with the flanking primers, 37 and 62 and the G19-4 HL was synthesized by combining the HL PCRs and re-amplifying with primers 63 and 67.

Full-length G19-4 LH/HL PCR products were separated by agarose gel electrophoresis, excised from the gel and purified with Qiagen MinElute columns as described earlier. These DNAs were then TOPO-cloned into pCR2.1 (Invitrogen), transformed into TOP10 and colonies screened first by EcoRI fragment size, then by DNA sequencing. G19-4 LH/HL were then cloned into pD18-IgG1 via PinAI-XhoI for expression in mammalian cells.

2H7sssIgG1-STD1-G19-4 LH/HL Construction

Using the G19-4 LH and G19-4 HL SMIPs as templates, the LH and HL anti-CD3 binding domains were altered by PCR such that their flanking restriction sites were compatible with the scorpion cassette. An EcoRI site was introduced at the 5' end of each scFv using either primer 27 (LH) or 36 (HL) and a stop codon/ XbaI site at the 3' end using either primer 28 (LH) or 35 (HL). The resulting DNAs were cloned into EcoRI-XbaI digested pD18-2H7sssIgG-STD1.

2H7sssIgG1-Hx-G19-4 HL Construction

2H7sssIgG1-Hx-2e12 HL DNAs were digested with BsrGI and EcoRI and the 325 bp fragment consisting of the C-terminal end of the IgG1 and linker. These were substituted for the equivalent region in 2H7sssIgG1-STD1-G19-4 HL by removal of the STD1 linker using BsrGI-EcoRI and replacing it with the corresponding linkers from the 2H7sssIgG1-Hx-2e12 HL clones.

Apparent from a consideration of the variety of multivalent binding proteins disclosed herein are features of the molecules that are amenable to combination in forming the molecules of the invention. Those features include binding domain 1, a constant sub-region, including a hinge or hinge-like domain, a linker domain, and a binding domain 2. The intrinsic modularity in the design of these novel binding proteins makes it straightforward for one skilled in the art to manipulate the DNA sequence at the N-terminal and/or C-terminal ends of any desirable module such that it can be inserted at almost any position to create a new molecule exhibiting altered or

enhanced functionality compared to the parental molecule(s) from which it was derived. For example, any binding domain derived from a member of the immunoglobulin superfamily is contemplated as either binding domain 1 or binding domain 2 of the molecules according to the invention. The derived binding domains include domains having amino acid sequences, and even encoding polynucleotide sequences, that have a one-to-one correspondence with the sequence of a member of the immunoglobulin superfamily, as well as variants and derivatives that preferably share 80%, 90%, 95%, 99%, or 99.5% sequence identity with a member of the immunoglobulin superfamily. These binding domains (1 and 2) are preferably linked to other modules of the molecules according to the invention through linkers that may vary in sequence and length as described elsewhere herein, provided that the linkers are sufficient to provide any spacing and flexibility necessary for the molecule to achieve a functional tertiary structure. Another module of the multivalent binding proteins is the hinge region, which may correspond to the hinge region of a member of the immunoglobulin superfamily, but may be a variant thereof, such as the "CSC" or "SSS" hinge regions described herein. Also, the constant sub-region comprises a module of the proteins according to the invention that may correspond to a sub-region of a constant region of an immunoglobulin superfamily member, as is typified by the structure of a hinge-C_{H2}-C_{H3} constant sub-region. Variants and derivatives of constant sub-regions are also contemplated, preferably having amino acid sequences that share 80%, 90%, 95%, 99%, or 99.5% sequence identity with a member of the immunoglobulin superfamily.

Exemplary primary structures of the features of such molecules are presented in Table 5, which discloses the polynucleotide and cognate amino acid sequence of illustrative binding domains 1 and 2, as well as the primary structure of a constant sub-region, including a hinge or hinge-like domain, and a linker that may be interposed, e.g., between the C-terminal end of a constant sub-region and the N-terminal end of a binding domain 2 region of a multivalent binding protein. Additional exemplars of the molecules according to the invention include the above-described features wherein, e.g., either or both of binding domains 1 and 2 comprise a domain derived from a V_L or V_L-like domain of a member of the immunoglobulin superfamily and a V_H or V_H-like domain derived from the same or a different member of the immunoglobulin superfamily, with these domains separated by a linker typified

by any of the linkers disclosed herein. Contemplated are molecules in which the orientation of these domains is V_L-V_H or V_H-V_L for BD1 and/or BD2. A more complete presentation of the primary structures of the various features of the multivalent binding molecules according to the invention is found in the table appended at the end of this disclosure. The invention further comprehends polynucleotides encoding such molecules.

Table 5

Binding Domain	Nucleotide Sequence	Amino Acid Sequence	SEQ ID NOS. (amino acid sequence
2H7 LH	atggattttcaagtgcagattttcag cttcctgctaatcagtgcttcagtca taatgtccagaggacaaattgttctc tcccagtctccagcaatcctgtctgc atctccaggggagaaggtcacaatga cttgcagggccagctcaagtgtaagt tacatgcactggtaccagcagaagcc aggatcctcccccaaaccctggattt atgccccatccaacctggcttctgga gtccctgctcgcttcagtggcagtgg gtctgggacctcttactctctcacaa tcagcagagtggaggctgaagatgct gccacttattactgccagcagtggag ttttaacccacccacgttcggtgctg ggaccaagctggagctgaaagatggc ggtggctcgggcggtggtggatctgg aggaggtgggagctctcaggcttatc tacagcagtctggggctgagtcggtg aggcctggggcctcagtgaagatgtc ctgcaaggcttctggctacacattta ccagttacaatatgcactgggtaaag cagacacctagacagggcctggaatg gattggagctatttatccaggaaatg gtgatacttcctacaatcagaagttc aagggcaaggccacactgactgtaga caaatcctccagcacagcctacatgc agctcagcagcctgacatctgaagac tctgcggtctatttctgtgcaagagt ggtgtactatagtaactcttactggt acttcgatgtctggggcacagggacc acggtcaccgtctct	mdfqvqifsfllisasvimsrgqivls qspailsaspgekvtmtcrasssvsym hwyqqkpgsspkpwiyapsnlasgvpa rfsgsgsgtsysltisrveaedaatyy cqqwsfnpptfgagtklelkdgggsgg ggsggggssqaylqqsgaesvrpgasv kmsckasgytftsynmhwvkqtprqgl ewigaiypgngdtsynqkfkgkatltv dkssstaymqlssltsedsavyfcarv vyysnsywyfdvwgtgttvtvs	1 (2)

Binding Domain	Nucleotide Sequence	Amino Acid Sequence	SEQ ID NOS. (amino acid sequence
2e12 LH	atggattttcaagtgcagattttcag cttcctgctaatcagtgcttcagtca taatgtccagaggagtcgacattgtg ctcacccaatctccagcttctttggc tgtgtctctaggtcagagagccacca tctcctgcagagccagtgaaagtgtt gaatattatgtcacaagtttaatgca gtggtaccaacagaaaccaggacagc cacccaaactcctcatctctgctgct agcaacgtagaatctggggtccctgc caggtttagtggcagtgggtctggga cagactttagcctcaacatccatcct gtggaggaggatgatattgcaatgta tttctgtcagcaaagtaggaaggttc catggacgttcggtggaggcaccaag ctggaaatcaaacggggtggcggtgg atccggcggaggtgggtcgggtggcg gcggatctcaggtgcagctgaaggag tcaggacctggcctggtggcgccctc acagagcctgtccatcacatgcaccg tctcagggttctcattaaccggctat ggtgtaaactgggttcgccagcctcc aggaaagggtctggagtggctgggaa tgatatggggtgatggaagcacagac tataattcagctctcaaatccagact atcgatcaccaaggacaactccaaga gccaagttttcttaaaaatgaacagt ctgcaaactgatgacacagccagata ctactgtgcccgagatggttatagta actttcattactatgttatggactac tggggtcaaggaacctcagtcaccgt ctcctct	MDFQVQIFSFLLISASVIMSRGVDIVL TQSPASLAVSLGQRATISCRASESVEY YVTSLMQWYQQKPGQPPKLLISAASNV ESGVPARFSGSGSGTDFSLNIHPVEED DIAMYFCQQSRKVPWTFGGGTKLEIKR GGGGSGGGGSGGGGSQVQLKESGPGLV APSQSLSITCTVSGFSLTGYGVNWVRQ PPGKGLEWLGMIWGDGSTDYNSALKSR LSITKDNSKSQVFLKMNSLQTDDTARY YCARDGYSNFHYYVMDYWGQGTSVTVS S	3 (4)
2e12 HL	atggattttcaagtgcagattttcag cttcctgctaatcagtgcttcagtca taatgtccagaggagtccaggtgcag ctgaaggagtcaggacctggcctggt ggcgccctcacagagcctgtccatca catgcaccgtctcagggttctcatta accggctatggtgtaaactgggttcg ccagcctccaggaaagggtctggagt ggctgggaatgatatggggtgatgga agcacagactataattcagctctcaa atccagactatcgatcaccaaggaca actccaagagccaagttttcttaaaa atgaacagtctgcaaactgatgacac agccagatactactgtgcccgagatg gttatagtaactttcattactatgtt atggactactggggtcaaggaacctc agtcaccgtctcctctgggggtggag gctctggtggcggtggatccggcgga ggtgggtcgggtggcggcggatctga cattgtgctcacccaatctccagctt ctttggctgtgtctctaggtcagaga gccaccatctcctgcagagccagtga aagtgttgaatattatgtcacaagtt taatgcagtggtaccaacagaaacca	MDFQVQIFSFLLISASVIMSRGVQVQL KESGPGLVAPSQSLSITCTVSGFSLTG YGVNWVRQPPGKGLEWLGMIWGDGSTD YNSALKSRLSITKDNSKSQVFLKMNSL QTDDTARYYCARDGYSNFHYYVMDYWG QGTSVTVSSGGGGSGGGGSGGGGSGGG GSDIVLTQSPASLAVSLGQRATISCRA SESVEYYVTSLMQWYQQKPGQPPKLLI SAASNVESGVPARFSGSGSGTDFSLNI HPVEEDDIAMYFCQQSRKVPWTFGGGT KLEIKR	5 (6)

Binding Domain	Nucleotide Sequence	Amino Acid Sequence	SEQ ID NOS. (amino acid sequence
	ggacagccacccaaactcctcatctc tgctgctagcaacgtagaatctgggg tccctgccaggtttagtggcagtggg tctgggacagactttagcctcaacat ccatcctgtggaggaggatgatattg caatgtatttctgtcagcaaagtagg aaggttccatggacgttcggtggagg caccaagctggaaatcaaacgt		
G28-1 LH	accggtgacatccagatgactcagtc tccagcctccctatctgcatctgtgg gagagactgtcaccatcacatgtcga acaagtgaaaatgtttacagttattt ggcttggtatcagcagaaacagggaa aatctcctcagctcctggtctctttt gcaaaaaccttagcagaaggtgtgcc atcaaggttcagtggcagtggatcag gcacacagttttctctgaagatcagc agcctgcagcctgaagattctggaag ttatttctgtcaacatcattccgata atccgtggacgttcggtggaggcacc gaactggagatcaaaggtggcggtgg ctcgggcggtggtgggtcgggtggcg gcggatctgctagcgcagtccagctg cagcagtctggacctgagctggaaaa gcctggcgcttcagtgaagatttcct gcaaggcttctggttactcattcact ggctacaatatgaactgggtgaagca gaataatggaaagagccttgagtgga ttggaaatattgatccttattatggt ggtactacctacaaccggaagttcaa gggcaaggccacattgactgtagaca aatcctccagcacagcctacatgcag ctcaagagtctgacatctgaggactc tgcagtctattactgtgcaagatcgg tcggccctatggactactggggtcaa ggaacctcagtcaccgtctcgag	DIQMTQSPASLSASVGETVTITCRTSE NVYSYLAWYQQKQGKSPQLLVSFAKTL AEGVPSRFSGSGSGTQFSLKISSLQPE DSGSYFCQHHSDNPWTFGGGTELEIKG GGGSGGGGSGGGGSASAVQLQQSGPEL EKPGASVKISCKASGYSFTGYNMNWVK QNNGKSLEWIGNIDPYYGGTTYNRKFK GKATLTVDKSSSTAYMQLKSLTSEDSA VYYCARSVGPMDYWGQGTSVTVS	102 (103)
G28-1 HL	accggtgaggtccagctgcaacagtc tggacctgaactggtgaagcctggag cttcaatgaagatttcctgcaaggcc tctggttactcattcactggctacat cgtgaactggctgaagcagagccatg gaaagaaccttgagtggattggactt attaatccatacaaaggtcttactac ctacaaccagaaattcaagggcaagg ccacattaactgtagacaagtcatcc agcacagcctacatggagctcctcag tctgacatctgaagactctgcagtct attactgtgcaagatctgggtactat ggtgactcggactggtacttcgatgt ctggggcgcagggaccacggtcaccg tctcctctggtggcggtggctcgggc ggtggtggatctggaggaggtgggag cgggggaggtggcagtgctagcgaca tccagatgacacagactacatcctcc ctgtctgcctctctgggagacagagt caccatcagttgcagggcaagtcagg	EVQLQQSGPELVKPGASMKISCKASGY SFTGYIVNWLKQSHGKNLEWIGLINPY KGLTTYNQKFKGKATLTVDKSSSTAYM ELLSLTSEDSAVYYCARSGYYGDSDWY FDVWGAGTTVTVSSGGGGSGGGGSGGG GSGGGGSASDIQMTQTTSSLSASLGDR VTISCRASQDIRNYLNWYQQKPDGTVK LLIYYTSRLHSGVPSRFSGSGSGTDYS LTIANLQPEDIATYFCQQGNTLPWTFG GGTKLVTKRS	104 (105)

Binding Domain	Nucleotide Sequence	Amino Acid Sequence	SEQ ID NOS. (amino acid sequence
	acattcgcaattatttaaactggtat cagcagaaaccagatggaactgttaa actcctgatctactacacatcaagat tacactcaggagtcccatcaaggttc agtggcagtgggtctggaacagatta ttctctcaccattgccaacctgcaac cagaagatattgccacttacttttgc caacagggtaatacgcttccgtggac gttcggtggaggcaccaaactggtaa ccaaacgctcgag		
G19-4 LH	accggtgacatccagatgacacagac tacatcctccctgtctgcctctctgg gagacagagtcaccatcagttgcagg gcaagtcaggacattcgcaattattt aaactggtatcagcagaaaccagatg gaactgttaaactcctgatctactac acatcaagattacactcaggagtccc atcaaggttcagtggcagtgggtctg gaacagattattctctcaccattgcc aacctgcaaccagaagatattgccac ttacttttgccaacagggtaatacgc ttccgtggacgttcggtggaggcacc aaactggtaaccaaacggggtggcgg tggctcgggcggtggtggatctggag gaggtgggagcgctagcgaggtccag ctgcaacagtctggacctgaactggt gaagcctggagcttcaatgaagattt cctgcaaggcctctggttactcattc actggctacatcgtgaactggctgaa gcagagccatggaaagaaccttgagt ggattggacttattaatccatacaaa ggtcttactacctacaaccagaaatt caagggcaaggccacattaactgtag acaagtcatccagcacagcctacatg gagctcctcagtctgacatctgaaga ctctgcagtctattactgtgcaagat ctgggtactatggtgactcggactgg tacttcgatgtctggggcgcagggac cacggtcaccgtctcctcgag	DIQMTQTTSSLSASLGDRVTISCRASQ DIRNYLNWYQQKPDGTVKLLIYYTSRL HSGVPSRFSGSGSGTDYSLTIANLQPE DIATYFCQQGNTLPWTFGGGTKLVTKR GGGGSGGGGSGGGGSASEVQLQQSGPE LVKPGASMKISCKASGYSFTGYIVNWL KQSHGKNLEWIGLINPYKGLTTYNQKF KGKATLTVDKSSSTAYMELLSLTSEDS AVYYCARSGYYGDSDWYFDVWGAGTTV TVSS	106 (107)
G19-4 HL	accggtgaggtccagctgcaacagtc tggacctgaactggtgaagcctggag cttcaatgaagatttcctgcaaggcc tctggttactcattcactggctacat cgtgaactggctgaagcagagccatg gaaagaaccttgagtggattggactt attaatccatacaaaggtcttactac ctacaaccagaaattcaagggcaagg ccacattaactgtagacaagtcatcc agcacagcctacatggagctcctcag tctgacatctgaagactctgcagtct attactgtgcaagatctgggtactat ggtgactcggactggtacttcgatgt ctggggcgcagggaccacggtcaccg tctcctctggtggcggtggctcgggc ggtggtggatctggaggaggtgggag cgctagcgacatccagatgacacaga	EVQLQQSGPELVKPGASMKISCKASGY SFTGYIVNWLKQSHGKNLEWIGLINPY KGLTTYNQKFKGKATLTVDKSSSTAYM ELLSLTSEDSAVYYCARSGYYGDSDWY FDVWGAGTTVTVSSGGGGSGGGGSGGG GSASDIQMTQTTSSLSASLGDRVTISC RASQDIRNYLNWYQQKPDGTVKLLIYY TSRLHSGVPSRFSGSGSGTDYSLTIAN LQPEDIATYFCQQGNTLPWTFGGGTKL VTKRS	108 (109)

Binding Domain	Nucleotide Sequence	Amino Acid Sequence	SEQ ID NOS. (amino acid sequence
	ctacatcctccctgtctgcctctctg ggagacagagtcaccatcagttgcag ggcaagtcaggacattcgcaattatt taaactggtatcagcagaaaccagat ggaactgttaaactcctgatctacta cacatcaagattacactcaggagtcc catcaaggttcagtggcagtgggtct ggaacagattattctctcaccattgc caacctgcaaccagaagatattgcca cttacttttgccaacagggtaatacg cttccgtggacgttcggtggaggcac caaactggtaaccaaacgctcgag		

Hinge Region	Nucleotide Sequence	Amino Acid Sequence	SEQ ID NO. (amino acid sequence)
sss(s)-hIgG1	gagcccaaatcttctgacaaaact cacacatctccaccgagctca	EPKSSDKTHTSPPSS	230 (231)
csc(s)-hIgG1	gagcccaaatcttgtgacaaaact cacacatctccaccgtgctca	EPKSCDKTHTSPPCS	232 (233)
ssc(s)-hIgG1	gagcccaaatcttctgacaaaact cacacatctccaccgtgctca	EPKSSDKTHTSPPCS	110 (111)
scc(s)-hIgG1	gagcccaaatcttctgacaaaact cacacatgtccaccgtgctca	EPKSSDKTHTCPPCS	112 (113)
css(s)-hIgG1	gagcccaaatcttgtgacaaaact cacacatctccaccgagctca	EPKSCDKTHTSPPSS	114 (115)
scs(s)-hIgG1	gagcccaaatcttgtgacaaaact cacacatgtccaccgagctca	EPKSSDKTHTCPPSS	116 (117)
ccc(s)-hIgG1	gagcccaaatcttgtgacaaaact cacacatgtccaccgtgctca	EPKSCDKTHTSPPCS	118 (119)
ccc(p)-hIgG1	gagcccaaatcttgtgacaaaact cacacatgtccaccgtgccca	EPKSCDKTHTSPPCP	120 (121)
sss(p)-hIgG1	gagcccaaatcttctgacaaaact cacacatctccaccgagccca	EPKSSDKTHTSPPSP	122 (123)
csc(p)-hIgG1	gagcccaaatcttgtgacaaaact cacacatctccaccgtgccca	EPKSCDKTHTSPPCP	124 (125)
ssc(p)-hIgG1	gagcccaaatcttctgacaaaact cacacatctccaccgtgccca	EPKSSDKTHTSPPCP	126 (127)
scc(p)-hIgG1	gagcccaaatcttctgacaaaact cacacatgtccaccgtgccca	EPKSSDKTHTCPPCP	128 (129)
css(p)-hIgG1	gagcccaaatcttgtgacaaaact cacacatctccaccgagccca	EPKSCDKTHTSPPSP	130 (131)
scs(p)-hIgG1	gagcccaaatcttgtgacaaaact cacacatgtccaccgagccca	EPKSSDKTHTCPPSP	132 (133)
scppcp	agttgtccaccgtgccca	SCPPCP	134 (135)
EFD	**Nucleotide Sequence**	**Amino acid Sequence**	**Sequence Identifier (amino acid sequence)**
hIgG1 (P238S) $C_{H2}C_{H3}$	gcacctgaactcctgggtggatcg tcagtcttcctcttcccccaaaa cccaaggacaccctcatgatctcc cggacccctgaggtcacatgcgtg gtggtggacgtgagccacgaagac	APELLGGSSVFLFPPKPKDTLMIS RTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYNSTYRV VSVLTVLHQDWLNGKEYKCKVSNK ALPAPIEKTISKAKGQPREPQVYT	142 (143)

	cctgaggtcaagttcaactggtac gtggacggcgtggaggtgcataat gccaagacaaagccgcgggaggag cagtacaacagcacgtaccgtgtg gtcagcgtcctcaccgtcctgcac caggactggctgaatggcaaggag tacaagtgcaaggtctccaacaaa gccctcccagcccccatcgagaaa acaatctccaaagccaaagggcag ccccgagaaccacaggtgtacacc ctgcccccatcccgggatgagctg accaagaaccaggtcagcctgacc tgcctggtcaaaggcttctatccc agcgacatcgccgtggagtgggag agcaatgggcagccggagaacaac tacaagaccacgcctcccgtgctg gactccgacggctccttcttcctc tacagcaagctcaccgtggacaag agcaggtggcagcaggggaacgtc ttctcatgctccgtgatgcatgag gctctgcacaaccactacacgcag aagagcctctccc tgtctccgggtaaatga	LPPSRDELTKNQVSLTCLVKGFYP SDIAVEWESNGQPENNYKTTPPVL DSDGSFFLYSKLTVDKSRWQQGNV FSCSVMHEALHNHYTQKSLSLSPG K	
hIgG1 (P331S) $C_{H2}C_{H3}$	gcacctgaactcctgggtggaccg tcagtcttcctcttcccccaaaa cccaaggacaccctcatgatctcc cggacccctgaggtcacatgcgtg gtggtggacgtgagccacgaagac cctgaggtcaagttcaactggtac gtggacggcgtggaggtgcataat gccaagacaaagccgcgggaggag cagtacaacagcacgtaccgtgtg gtcagcgtcctcaccgtcctgcac caggactggctgaatggcaaggag tacaagtgcaaggtctccaacaaa gccctcccagcctccatcgagaaa acaatctccaaagccaaagggcag ccccgagaaccacaggtgtacacc ctgcccccatcccgggatgagctg accaagaaccaggtcagcctgacc tgcctggtcaaaggcttctatccc agcgacatcgccgtggagtgggag agcaatgggcagccggagaacaac tacaagaccacgcctcccgtgctg gactccgacggctccttcttcctc tacagcaagctcaccgtggacaag agcaggtggcagcaggggaacgtc ttctcatgctccgtgatgcatgag gctctgcacaaccactacacgcag aagagcctctccc tgtctccgggtaaatga	APELLGGPSVFLFPPKPKDTLMIS RTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYNSTYRV VSVLTVLHQDWLNGKEYKCKVSNK ALPASIEKTISKAKGQPREPQVYT LPPSRDELTKNQVSLTCLVKGFYP SDIAVEWESNGQPENNYKTTPPVL DSDGSFFLYSKLTVDKSRWQQGNV FSCSVMHEALHNHYTQKSLSLSPG K	144 (145)
hIgG1 (P238S/ P331S) $C_{H2}C_{H3}$	gcacctgaactcctgggtggatcg tcagtcttcctcttcccccaaaa cccaaggacaccctcatgatctcc cggacccctgaggtcacatgcgtg gtggtggacgtgagccacgaagac cctgaggtcaagttcaactggtac gtggacggcgtggaggtgcataat gccaagacaaagccgcgggaggag cagtacaacagcacgtaccgtgtg gtcagcgtcctcaccgtcctgcac	APELLGGSSVFLFPPKPKDTLMIS RTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYNSTYRV VSVLTVLHQDWLNGKEYKCKVSNK ALPASIEKTISKAKGQPREPQVYT LPPSRDELTKNQVSLTCLVKGFYP SDIAVEWESNGQPENNYKTTPPVL DSDGSFFLYSKLTVDKSRWQQGNV FSCSVMHEALHNHYTQKSLSLSPG K	146 (147)

	caggactggctgaatggcaaggag tacaagtgcaaggtctccaacaaa gccctcccagcctccatcgagaaa acaatctccaaagccaaagggcag ccccgagaaccacaggtgtacacc ctgcccccatcccgggatgagctg accaagaaccaggtcagcctgacc tgcctggtcaaaggcttctatccc agcgacatcgccgtggagtgggag agcaatgggcagccggagaacaac tacaagaccacgcctcccgtgctg gactccgacggctccttcttcctc tacagcaagctcaccgtggacaag agcaggtggcagcaggggaacgtc ttctcatgctccgtgatgcatgag gctctgcacaaccactacacgcag aagagcctctccctgtctccgggt aaatga		
Linker	**Nucleotide Sequence**	**Amino Acid Sequence**	**Sequence Identifier**
STD1	aattatggtggcggtggctcgggc ggtggtggatctggaggaggtggg agtgggaattct	NYGGGGSGGGGSGGGGSGNS	148 (149)
STD2	aattatggtggcggtggctcgggc ggtggtggatctggaggaggtggg agtgggaattatggtggcggtggc tcgggcggtggtggatctggagga ggtgggagtgggaattct	NYGGGGSGGGGSGGGGSGNYGGGG SGGGGSGGGGSGNS	150 (151)
H1	aattct	NS	152 (153)
H2	ggtggcggtggctcggggaattct	GGGGSGNS	154 (155)
H3	aattatggtggcggtggctctggg aattct	NYGGGGSGNS	156 (157)
H4	ggtggcggtggctcgggcggtggt ggatctgggaattct	GGGGSGGGGSGNS	158 (159)
H5	aattatggtggcggtggctcgggc ggtggtggatctgggaattct	NYGGGGSGGGGSGNS	160 (161)
H6	ggtggcggtggctcgggcggtggt ggatctgggggaggaggcagcggg aattct	GGGGSGGGGSGGGGSGNS	162 (163)
H7	gggtgtccaccttgtccgaattct	GCPPCPNS	164 (165)
(G4S)3	ggtggcggtggatccggcggaggt gggtcgggtggcggcggatct	GGGSGGGSGGGS	166 (167)
(G4S)4	ggtggcggtggctcgggcggtggt ggatctggaggaggtgggagcggg ggaggtggcagt	GGGSGGGSGGGSGGGGS	168 (169)

Table 5. Primary structures (polynucoleotide and cognate amino acid sequences) of exemplary features of multivalent binding molecules.

Example 8

Binding and Functional Studies with Alternative Multispecific Fusion Proteins

Experiments that parallel the experiments described above for the prototypical CD20-IgG-CD28 multispecific binding (fusion) molecule were conducted for each of

the additional multivalent binding molecules described above. In general, the data obtained for these additional molecules parallel the results observed for the prototype molecule. Some of the salient results of these experiments are disclosed below. Figure 14 shows results of blocking studies performed on one of the new molecules where both BD1 and BD2 bind to target antigens on the same cell or cell type, in this case, CD20 and CD37. This multispecific, multivalent binding (fusion) protein was designed with binding domain 1 binding CD20 (2H7; VLVH orientation), and binding domain 2 binding CD37, G28-1 VL-VH (LH) or VH-VL (HL). The experiment was performed in order to demonstrate the multispecific properties of the protein.

Blocking Studies:Ramos or BJAB B lymphoblastoid cells ($2.5x10^5$) were pre-incubated in 96-well V-bottom plates in staining medium (PBS with 2% mouse sera) with murine anti-CD20 (25 μg/ml) antibody, or murine anti-CD37 (10 μg/ml) antibody, both together or staining medium alone for 45 minutes on ice in the dark. Blocking antibodies were pre-incubated with cells for 10 minutes at room temperature prior to addition of the multispecific binding (fusion) protein at the concentration ranges indicated, usually from 0.02 μg/ml to 10 μg/ml, and incubated for a further 45 minutes on ice in the dark. Cells were washed 2 times in staining medium, and incubated for one hour on ice with Caltag (Burlingame, CA) FITC goat anti-human IgG (1:100) in staining medium, to detect binding of the multispecific binding (fusion) proteins to the cells. The cells were then washed 2 times with PBS and fixed with 1% paraformaldehyde (cat. no. 19943, USB, Cleveland, Ohio). The cells were analyzed by flow cytometry using a FACsCalibur instrument and CellQuest software (BD Biosciences, San Jose, CA). Each data series plots the binding of the 2H7-sss-hIgG-STD1-G28-1 HL fusion protein in the presence of CD20, CD37, or both CD20 and CD37 blocking antibodies. Even though this experiment used one of the cleaved linkers, only the presence of both blocking antibodies completely eliminates binding by the multispecific binding (fusion) protein, demonstrating that the bulk of the molecules possess binding function for both CD20 and CD37. The data were similar for two cell lines tested in panels A and B, Ramos and BJAB, where the CD20 blocking antibody was more effective than the CD37 blocking antibody at reducing the level of binding observed by the multispecific binding (fusion) protein.

ADCC Assays

Figure 15 shows the results of ADCC assays performed on the CD20-CD37 multispecific binding (fusion) proteins. ADCC assays were performed using BJAB lymphoblastoid B cells as targets and human PBMC as effector cells. BJAB cells were labeled with 500 μCi/ml ^{51}Cr sodium chromate (250 μCi/μg) for 2 hours at 37°C in IMDM/10%FBS. The labeled cells were washed three times in RPMI.10% FBS and resuspended at $4x10^5$ cells/ml in RPMI. Heparinized, human whole blood was obtained from anonymous, in-house donors and PBMC isolated by fractionation over Lymphocyte Separation Media (LSM, ICN Biomedical) gradients. Buffy coats were harvested and washed twice in RPMI/10% FBS prior to resuspension in RPMI/10% FBS at a final concentration of $5x10^6$ cells/ml. Cells were counted by trypan blue exclusion using a hemacytometer prior to use in subsequent assays. Reagent samples were added to RPMI medium with 10% FBS at 4 times the final concentration and three 10 fold serial dilutions for each reagent were prepared. These reagents were then added to 96-well U-bottom plates at 50 μl/well for the indicated final concentrations. The ^{51}Cr-labeled BJAB cells were added to the plates at 50 μl/well ($2x10^4$ cells/well). The PBMCs were then added to the plates at 100 μl/well ($5x10^5$ cells/well) for a final ratio of 25:1 effector (PBMC):target (BJAB). Effectors and targets were added to medium alone to measure background killing. The ^{51}Cr-labeled cells were added to medium alone to measure spontaneous release of ^{51}Cr and to medium with 5% NP40 (cat. no.28324, Pierce, Rockford, IL) to measure maximal release of ^{51}Cr. Reactions were set up in triplicate wells of a 96-well plate. Multispecific binding (fusion) proteins were added to wells at a final concentration ranging from 0.01 μg/ml to 10μg/ml, as indicated on the graphs. Each data series plots a different multispecific binding (fusion) protein or the corresponding single specificity SMIPs at the titration ranges described. Reactions were allowed to proceed for 6 hours at 37°C in 5% CO_2 prior to harvesting and counting. Twenty-five μl of the supernatant from each well were then transferred to a Luma Plate 96 (cat. no. 6006633, Perkin Elmer, Boston, Mass) and dried overnight at room temperature. CPM released was measured on a Packard TopCounNXT. Percent specific killing was calculated by subtracting (cpm {mean of triplicate samples} of sample – cpm spontaneous release)/(cpm maximal release-cpm spontaneous release) x100. Data are plotted as % specific killing versus protein concentration. The data demonstrate that

the multispecific binding (fusion) protein is able to mediate ADCC activity against cells expressing the target antigen(s) as well as the single specificity SMIPs for CD20 and/or CD37, but does not show augmentation in the level of this effector function.

Co-culture Experiments

Figure 16 shows the results of experiments designed to look at other properties of this type of multispecific binding (fusion) protein, where having two binding domains against targets expressed on the same cell or cell type might result in synergistic effects by signaling/binding through the two surface receptors bound. The co-culture experiments were performed using PBMC isolated as described for the ADCC assays above.. These PBMC were resuspended in culture medium at $2x10^6$ cells/ml in a final volume of 500 μl/well, and cultured alone or incubated with single specificity SMIPs for CD20, CD37, CD20+CD37, or the multispecific binding (fusion) protein using the H7 linker, [2H7-sss-IgG-H7-G28-1 HL]. Each of the test reagents was added at a final concentration of 20 μg/ml. After 24 hours of culture, no real differences were seen in the % of B cells in culture; however, when the cells were subjected to flow cytometry, cell clumping was visible in the FWD X 90 staining pattern for the cultures containing the multispecific binding (fusion) protein, indicating that the B cells expressing the two target antigens were engaged in homotypic adhesion. After 72 hours in culture, the multispecific binding (fusion) protein resulted in the death of almost all the B cells present. The combination of the two single-specificity SMIPs also drastically decreased the percentage of B cells, but not to the level seen with the multispecific binding molecule. These data suggest that engaging both binding domains for CD20 and CD37 on the same multispecific molecule, results in homotypic adhesion between B cells and may also result in binding of both CD20 and CD37 antigens on the same cell. Without wishing to be bound by theory, the synergistic effect in eliminating target cells may be due (1) to the binding through binding domains 1 and 2 on the same cell types, and/or (2) to interactions of the effector function domain (constant sub-region) of the multivalent binding molecules with monocytes or other cell types in the PBMC culture that result in delayed killing. The kinetics of this killing effect are not rapid, taking more than 24 hours to be achieved, indicating that it is may be a secondary effect, requiring production of cytokines or other molecules prior to the effects being observed.

Apoptosis Assays

Figure 17 shows the results of experiments designed to explore the induction of apoptosis after treatment of B cell lines with either the [2H7-sss-hIgG-H7-G28-1 HL] multispecific, multivalent binding (fusion) proteins or the single specificity CD20 and/or CD37 SMIPS, alone and in combination with one another. Ramos cells (panel A; ATCC No. CRL-1596), and Daudi cells (panel B; ATCC No. CCL-213) were incubated overnight (24 hours) at 37°C in 5% CO_2 in Iscoves (Gibco) complete medium with 10% FBS at 3 X 10^5 cells/ml and 5, 10, or 20 μg/ml fusion proteins. For combination experiments with the single specificity SMIPs, the proteins were used at the following concentrations: TRU-015 (CD20 directed SMIP)=10 μg/ml, with 5 μg/ml G28-1 LH (CD37 directed SMIP). Alternatively, TRU-015=20 μg/ml was combined with G28-1 LH at 10 μg/ml. Cells were then stained with Annexin V-FITC and propidium iodide using the BD Pharmingen Apoptosis Detection Kit I cat. no. 556547), and processed according to kit instructions. The cells were gently vortexed, incubated in the dark at room temperature for 15 minutes, and diluted in 400 μl binding buffer prior to analysis. Samples were analyzed by flow cytometry on a FACsCalibur (Becton Dickinson) instrument using Cell Quest software (Becton Dickinson). The data are presented as columnar graphs plotting the percentage of Annexin V/propidium iodide positive cells versus type of treatment. Clearly, the multispecific binding (fusion) protein is able to induce a significantly higher level of apoptotic death in both cell lines than the single specificity reagents, even when used together. This increased functional activity reflects an interaction of the coordinate binding of BD1 and BD2 (specific for CD20 and CD37) receptors on the target cells.

Example 9

Binding and Functional Properties of 2H7-hIgG-G19-4 Multispecific Binding (Fusion) Proteins

This example describes the binding and functional properties of the 2H7-hIgG-G19-4 multispecific fusion proteins. The construction of these molecules is described in Example 7. Expression and purification are as described in previous Examples.

Binding experiments were performed as described for previous molecules, except that the target cells used to measure CD3 binding were Jurkat cells expressing CD3 on their surface. Refer to Figure 18, where the top graph shows binding curves obtained for binding of the CD20-CD3 multispecific molecules to Jurkat cells using purified proteins serially diluted from 20 to 0.05 μg/ml. The HL orientation of the G19-4 specificity seems to bind better to the CD3 antigen than does the LH orientation. The lower panel shows the binding curves obtained for the BD1, the binding domain recognizing CD20. All of the molecules bind well, and at a level nearly equivalent to a single specificity SMIP for CD20.

ADCC Assays

For the data presented in Figure 19, ADCC assays were performed as described in the previous Example. In this case, the fusion proteins were all 2H7-hIgG-G19-4 variants or combinations of the single-specificity SMIPs (2H7, specific for CD20) or antibodies (G19-4, specific for CD3). In addition, for the data presented in the lower panel of Figure 19, NK cells were depleted from PBMC prior to use, by magnetic bead depletion using a MACS (Miltenyi Biotec, Auburn, CA) column separation apparatus and NK cell-specific CD16 magnetic microbeads (cat no.: 130-045-701). The data presented in the two panels demonstrate that all of the CD20-hIgG-CD3 multispecific molecules mediate ADCC, regardless of whether NK cells are depleted or total PBMC are used in the assay. For the TRU 015 or combinations of G19-4 and TRU015, only cultures containing NK cells could mediate ADCC. G19-4 did not work well in either assay against BJAB targets, which do not express CD3, although G19-4 may have bound to CD3 expressing NK T cells and activated these cells in the first assay shown. The killing observed in the lower panel for the multispecific binding (fusion) proteins is probably mediated through activation of cytotoxicity in the T cell population by binding CD3, against the BJAB targets expressing the CD20 antigen. This killing activity appears to be relatively insensitive to the dosage of the molecules over the concentration ranges used, and is still significantly different from the other molecules tested, even at a concentration of 0.01 ug/ml.

Example 10

Multivalent Binding Molecules

Other embodiments include linker domains derived from immunoglobulins. More specifically, the source sequences for these linkers are sequences obtained by comparing regions present between the V-like domains or the V- and C-like domains of other members of the immunoglobulin superfamily. Because these sequences are usually expressed as part of the extracellular domain of cell surface receptors, they are expected to be more stable to proteolytic cleavage, and should also not be immunogenic. One type of sequence that is not expected to be as useful in the role of a linker for the multivalent binding (fusion) proteins is the type of sequence expressed on surface-expressed members of the –Ig superfamily, but that occur in the intervening region between the C-like domain and the transmembrane domain. Many of these molecules have been observed in soluble form, and are cleaved in these intervening regions close to the cell membrane, indicating that the sequences are more susceptible to cleavage than the rest of the molecule.

The linkers described above are inserted into either a single specificity SMIP, between the binding domain and the effector function domain, or are inserted into one of the two possible linker positions in a multivalent binding (fusion) protein, as described herein.

A complete listing of the sequences disclosed in this application is appended, and is incorporated herein by reference in its entirety. The color coding indicating the sequence of various regions or domains of the particular polynucleotides and polypeptides are useful in identifying a corresponding region or domain in the sequence of any of the molecules disclosed herein.

Example 11

Screening matrix for scorpion candidates targeting B-cells

Introduction

As a means of identifying combinations of paired monoclonal antibody binding domains that would most likely yield useful and potent multivalent binding molecules, or scorpions, against a target population, a series of monoclonal antibodies against B cell antigens was tested in a combination matrix against B cell lines

representing various non Hodgkin's lymphomas. To ensure that all possible pairwise comparisons of antibodies known or expected to bind to the cell of interest are assayed, a two-dimensional matrix of antibodies may be used to guide the design of studies using a given cell type. Monoclonal antibodies against numerous B cell antigens known by their cluster designations (CDs) are recorded in the left column. Some of these antibodies (designated by the antigen(s) to which they specifically bind), i.e., CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and CL II (MHC Class II), were incubated, alone or in combination with other members of this monoclonal antibody set, with antigen-positive target cells. The variable domains of these antibodies are contemplated as binding domains in exemplary embodiments of the multivalent binding molecules. Using the knowledge in the art and routine procedures, those of skill in the art are able to identify suitable antibody sequences (nucleic acid encoding sequences as well as amino acid sequences), for example in publicly available databases, to generate a suitable antibody or fragment thereof (e.g., by hybridization-based cloning, PCR, peptide synthesis, and the like), and to construct multivalent binding molecules using such compounds. Sources of exemplary antibodies from which binding domains were obtained as described herein are provided in Table 6. Typically, a cloning or synthesis strategy that realizes the CDR regions of an antibody chain will be used, although any antibody, fragment thereof, or derivative thereof that retains the capacity to specifically bind to a target antigen is contemplated.

Stated in more detail, the cloning of heavy and/or light chain variable regions of antibodies from hybridomas is standard in the art. There is no requirement that the sequence of the variable region of interest be known in order to obtain that region using conventional cloning techniques. See, e.g., Gilliland et al., Tissue Antigens 47(1):1-20 (1996). To prepare single-chain polypeptides comprising a variable region recognizing a murine or human leukocyte antigen, a method was devised for rapid cloning and expression that yielded functional protein within two to three weeks of RNA isolation from hybridoma cells. Variable regions were cloned by poly-G tailing the first-strand cDNA followed by anchor PCT with a forward poly-C anchor primer and a reverse primer specific for the constant region sequence. Both primers contain flanking restriction endonuclease sites for insertion into pUC19. Sets of PCR primers for isolation of murine, hamster and rat V_L and V_H genes were generated. Following

determination of consensus sequences for a specific V_L and V_H pair, the V_L and V_H genes were linked by DNA encoding an intervening peptide linker (typically encoding $(Gly_4Ser)_3$) and the V_L-linker-V_H gene cassettes were transferred into the pCDM8 mammalian expression vector. The constructs were transfected into COS cells and sFvs were recovered from conditioned culture medium supernatant. This method has been successfully used to generate functional sFv to human CD2, CD3, CD4, CD8, CD28, CD40, CD45 and to murine CD3 and gp39, from hybridomas producing murine, rat, or hamster antibodies. Initially, the sFvs were expressed as fusion proteins with the hinge-C_{H2}-C_{H3} domains of human IgG1 to facilitate rapid characterization and purification using goat anti-human IgG reagents or protein A. Active sFv could also be expressed with a small peptide, e.g., a tag, or in a tailless form. Expression of CD3 (G19-4) sFv tailless forms demonstrated increased cellular signaling activity and revealed that sFvs have potential for activating receptors.

Alternatively, identification of the primary amino acid sequence of the variable domains of monoclonal antibodies can be achieved directly, e.g., by limited proteolysis of the antibody followed by N-terminal peptide sequencing using, e.g., the Edman degradation method or by fragmentation mass spectroscopy. N-terminal sequencing methods are well known in the art. Following determination of the primary amino acid sequence, the variable domains, a cDNA encoding this sequence is assembled by synthetic nucleic acid synthesis methods (e.g., PCR) followed by scFv generation. The necessary or preferred nucleic acid manipulation methods are standard in the art.

Fragments, derivatives and analogs of antibodies, as described above, are also contemplated as suitable binding domains. Further, any of the constant sub-regions described above are contemplated, including constant sub-regions comprising any of the above-described hinge regions. Additionally, the multivalent single-chain binding molecules described in this example may include any or all of the linkers described herein.

Monoclonal antibodies were initially exposed to cells and then cross-linked using a goat anti-mouse second-step antibody (2^{nd} step). Optionally, one could cross-link the antibodies prior to contacting cells with the antibodies, e.g., by cross-linking the antibodies in solution. As another alternative, monoclonal antibodies could be

cross-linked in a solid phase by adsorbing onto the plastic bottom of tissue culture wells or "trapped" on this plastic by means of goat anti-mouse antibody adsorbed to the plastic, followed by plate-based assays to evaluate, e.g., growth arrest or cell viability.

Inversion of phosphatidylserine from the cytosolic side of the cell membrane to the exterior cell surface of that plasma membrane is an accepted indicator of pro-apoptotic events. Progression to apoptosis leads to loss of cell membrane integrity, which can be detected by entry of a cell-impermeant intercalating dye, e.g., propidium iodide (PI). Following cell exposure to monoclonal antibodies alone or in combination, a dual, pro-apoptotic assay was performed and treated cell populations were scored for cell surface-positive annexin V (ANN) and/or PI inclusion.

Annexin V binding /Propidium iodide internalization analysis

Cells and cell culture conditions. Experiments were performed to examine the effect of cross-linking two different monoclonal antibodies against targets expressed on four human B-cell lines. Effects on cell lines were measured by determining levels of ANN and/or PI staining following exposure. The human B cell lines BJAB, Ramos (ATCC#CRL-1596), Daudi (ATCC#CCL-213), and DHL-4 (DSMZ#ACC495) were incubated for 24 hours at 37°C in 5% CO_2 in Iscoves (Gibco) complete medium with 10% FBS. Cells were maintained at a density between 2-8 x 10^5 cells/ml and a viability typically >95% prior to study.

Experiments were conducted at a cell density of 2 x 10^5 cells/ml and 2 μg/ml of each comparative monoclonal antibody from a matrix against B-cell antigens. Each comparator monoclonal antibody was added at 2 μg/ml alone or individually when combined with each matrix monoclonal antibody, also at 2 μg/ml. Table 6 lists the catalog number and sources of monoclonal antibodies used in these experiments. For cross-linking these monoclonal antibodies in solution, goat anti-mouse IgG (Jackson Labs catalog no.115-001-008) was added to each well at a concentration ratio of 2:1 (goat anti-mouse: each monoclonal antibody), e.g., a well with only one monoclonal antibody at 2 μg/ml would have goat anti-mouse added to a final concentration of 4 μg/ml, while wells with both comparator monoclonal antibody (2 μg/ml) and a monoclonal antibody from the matrix (2 μg/ml) would have 8 μg/ml of goat anti-mouse antibody added to the well.

After 24 hours of incubation at 37°C in 5% CO_2, cells were stained with Annexin V-FITC and propidium iodide using the BD Pharmingen Annexin V-FITC Apoptosis Detection Kit I (#556547). Briefly, cells were washed twice with cold PBS and resuspended in "binding buffer" at 1 x 10^6 cells/ml. One hundred microliters of the cells in binding buffer were then stained with 5 µl of Annexin V-FITC and 5 µl of propidium iodide. The cells were gently mixed and incubated in the dark at room temperature for 15 minutes. Four hundred microliters of binding buffer were then added to each sample. The samples were then read on a FACsCalibur (Becton Dickinson) and analyzed using Cell Quest software (Becton Dickinson).

Table 6

Name	Catalog number	Commercial supplier
Anti-CD19	#C2269-74	US Biological (Swampscott, MA)
Anti-CD20	#169-820	Ancell Corp (Bayport, MN)
Anti-CD21	#170-820	Ancell Corp (Bayport, MN)
Anti-CD22	#171-820	Ancell Corp (Bayport, MN)
Anti-CD23	#172-820	Ancell Corp (Bayport, MN)
Anti-CD30	#179-820	Ancell Corp (Bayport, MN)
Anti-CD37	#186-820	Ancell Corp (Bayport, MN)
Anti-CD40	#300-820	Ancell Corp (Bayport, MN)
Anti-CD70	#222-820	Ancell Corp (Bayport, MN)
Anti-CD72	#C2428-41B1	US Biological (Swampscott, MA)
Anti-CD79a	#235-820	Ancell Corp (Bayport, MN)
Anti-CD79b	#301-820	Ancell Corp (Bayport, MN)
Anti-CD80	#110-820	Ancell Corp (Bayport, MN)
Anti-CD81	#302-820	Ancell Corp (Bayport, MN)
Anti-CD86	#307-820	Ancell Corp (Bayport, MN)
Anti-CL II DR, DQ, DP	#131-820	Ancell Corp (Bayport, MN)

Table 6. Antibodies against B cell antigens used in this study and their sources.

Addition of the cross-linking antibody (e.g., goat anti-mouse antibody) to monoclonal antibody A alone resulted in increased cell sensitivity, suggesting that a multivalent binding molecule, or scorpion, constructed with two binding domains recognizing the same antigen would be effective at increasing cell sensitivity. Without wishing to be bound by theory, this increased sensitivity could be due to antigen clustering and altered signaling. TNF receptor family members, for example, require homo-multimerization for signal transduction and scorpions with equivalent binding domains on each end of the molecule could facilitate this interaction. The

clustering and subsequent signaling by CD40 is an example of this phenomenon in the B cell system.

As shown in Figures 20, 21 and 22, the addition of monoclonal antibody A and monoclonal antibody B against different antigens will produce additive or in some combinations greater than additive (i.e., synergistic) pro-apoptotic effects on treated cells. In Figure 20, for example, the combination of anti-CD20 with monoclonal antibodies against other B cell antigens all resulted, to varying extents, in increased cell sensitivity. Some combinations, such as anti-CD20 combined with anti-CD19 or anti-CD20 combined with anti-CD21, however, produced greater than additive pro-apoptotic effects, indicating that multivalent binding molecules or scorpions composed of these binding domains should be particularly effective at eliminating transformed B cells. Referring to Figure 20, the percentage of cells exhibiting pro-apoptotic activities when exposed to anti-CD 20 antibody alone is about 33% (vertically striped bar corresponding to "20," i.e., the anti-CD20 antibody); the percentage of pro-apoptotic cells upon exposure to anti-CD19 antibody is about 12% (vertically striped bar in Fig. 20 corresponding to "19," i.e., the anti-CD19 antibody); and the percentage of pro-apoptotic cells upon exposure to both anti-CD20 and anti-CD19 antibodies is about 73% (horizontally striped bar in Fig. 20 corresponding to "19"). The 73% of pro-apoptotic cells following exposure to both antibodies is significantly greater than the 45% (33% + 12%) sum of the effects attributable to each individual antibody, indicating a synergistic effect attributable to the anti-CD19 and anti-CD 20 antibody pair. Useful multivalent binding molecules include molecules in which the two binding domains lead to an additive effect on B-cell behavior as well as multivalent binding molecules in which the two binding domains lead to synergistic effects on B-cell behavior. In some embodiments, one binding domain will have no detectable effect on the measured parameter of cell behavior, with each of the paired binding domains contributing to distinct aspects of the activities of the multivalent binding molecule, such as a multispecific, multivalent binding molecule (e.g., binding domain A binds to a target cell and promotes apoptosis while binding domain B binds to a soluble therapeutic such as a cytotoxin). Depending on the design of a multivalent binding molecule, the issue of the type of combined effect (additive, synergistic, or inhibitory) of the two binding domains on a target cell may not be relevant because one of the binding domains is specific for a

non-cellular (e.g., soluble) binding partner or is specific for a cell-associated binding partner, but on a different cell type.

Exemplary binding domain pairings producing additive, synergistic or inhibitory effects, as shown in Figures 20-23, are apparent from Tables 7 and 8. Table 7 provides quantitative data extracted from each of Figures 20-23 in terms of the percentage of cells staining positive for ANN and/or PI. Table 8 provides calculations using the data of Table 7 that provided a basis for determining whether the interaction of a given pair of antibodies yielded an additive, synergistic, or inhibitory effect, again as assessed by the percentage of cells staining positive for ANN and/or PI.

Table 7

Name	**Anti-CD20**	**Anti-CD79b**	**Anti-CL II**	**Anti-CD22**
Anti-CD19	13/73*	18/76/66	14/47/46	12/11
Anti-CD20	33/NA	42/94/92	33/71/76	28/33
Anti-CD21	14/75	22/50/76	18/24/40	11/11
Anti-CD22	8/55	12/39/33	12/19/17	10/12
Anti-CD23	8/41	12/63/55	14/22/17	10/12
Anti-CD30	8/38	14/72/61	12/56/61	10/11
Anti-CD37	15/45	19/92/86	20/60/62	19/20
Anti-CD40	10/48	12/44/30	13/21/28	14/13
Anti-CD70	9/40	12/56/39	15/21/15	10/10
Anti-CD72	NA	16/60/64	30/78/63	17/17
Anti-CD79a	21/66	43/42/50	28/55/51	14/14
Anti-CD79b	46/88	70/70/68	45/80/76	26/16

Name	Anti-CD20	Anti-CD79b	Anti-CL II	Anti-CD22
Anti-CD80	7/41	14/35/30	15/19/17	11/11
Anti-CD81	14/65	25/86/83	25/54/43	19/20
Anti-CD86	7/38	16/58/42	15/24/18	14/11
Anti- CL II	53/77	52/96/98	47/52/43	72/57

*In columns 2-4 of Table 7, the numerical values reflect the heights of histogram bars in Figures 20-22, respectively, with the first number in each cell denoting the height of a vertically striped bar, the second number denoting the height of a horizontally striped bar and, where present, the third number reflecting the height of a stippled bar. In column 5, the first number reflects the height of a solid bars and the second number reflects the height of a slant-striped bar in Figure 23.

Table 8

Name	Anti-CD20	Anti-CD79b	Anti-CL II	Anti-CD22
Anti-CD19	S: 13+33=46*	A: 18+56=74 A: 18+43=61	S: 14+26=40 S: 14+18=32	I: 12+10=22
Anti-CD20	NA	A: 42+56=98 A: 42+43=85	S: 33+26=59 S: 33+18=51	A/I:28+10=38
Anti-CD21	S: 14+33=47	I: 22+56=78 S: 22+43=65	I: 18+26=44 A: 18+18=36	I: 11+10=21
Anti-CD22	S: 8+33=41	I: 12+56=68 I: 12+43=55	I: 12+26=38 I: 12+18=30	NA
Anti-CD23	A: 8+33=41	A: 12+56=68 A: 12+43=55	I: 14+26=40 I: 14+18=32	I: 10+10=20

Name	Anti-CD20	Anti-CD79b	Anti-CL II	Anti-CD22
Anti-CD30	A: 8+33=41	A: 14+56=70 A: 14+43=57	S: 12+26=38 S: 12+18=30	I: 10+10=20
Anti-CD37	A: 15+33=48	S: 19+56=75 S: 19+43=62	S: 20+26=46 S: 20+18=38	I: 19+10=29
Anti-CD40	A/S:10+33=43	I: 12+56=68 I: 12+43=55	I: 13+26=39 A: 13+18=31	I: 14+10=24
Anti-CD70	A: 9+33=42	I: 12+56=68 I: 12+43=55	I: 15+26=41 I: 15+18=33	I: 10+10=20
Anti-CD72	NA	I: 16+56=72 A: 16+43=59	S: 30+26=56 S: 30+18=48	I: 17+10=27
Anti-CD79a	S: 21+33=54	I: 43+56=99 I: 43+43=86	A: 28+26=54 A: 28+18=46	I: 14+10=24
Anti-CD79b	S: 46+33=79	NA	S: 45+26=71 S: 45+18=63	I: 26+10=36
Anti-CD80	A: 7+33=40	I: 14+56=70 I: 14+43=57	I: 15+26=41 I: 15+18=33	I: 11+10=21
Anti-CD81	S: 14+33=47	A: 25+56=81 S: 25+43=68	A: 25+26=51 A: 25+18=43	I: 19+10=29
Anti-CD86	A: 7+33=40	I: 16+56=72 I: 16+43=59	I: 15+26=41 I: 15+18=33	I: 14+11=25

Name	Anti-CD20	Anti-CD79b	Anti-CL II	Anti-CD22
Anti- CL II	I: 53+33=86	A: 52+56=108 A: 52+43=95	NA	I: 72+10=82

"A" means an "additive" effect was observed
"S" means a "synergistic" effect was observed
"I" means an "inhibitory" effect was observed
*Equation schematic: A+B=C, where "A" is the percent ANN and/or PI positive cells due to matrix antibody alone, "B" is the percent ANN and/or PI positive cells due to the common antibody (anti-CD20 for Fig. 20, anti-CD79b for Fig. 21, anti-CLII for Fig. 22, and anti-CD22 for Fig. 23), and "C" is the expected additive effect. (See Table 7, above, for the quantitative data corresponding to Figures 20-23.) Where two equations are present in a cell, the upper equation reflects results use of the higher indicated concentration of common antibody; the lower equation reflects use of the lower indicated concentration of common antibody.

In some embodiments, the two binding domains interact in an inhibitory, additive or synergistic manner in sensitizing (or de-sensitizing) a target cell such as a B cell. Figure 23 shows the protective, or inhibitory, effects resulting from combining anti-CD22 antibody with strongly pro-apoptotic monoclonal antibodies such as the anti-CD79b antibody or anti-MHC class II (i.e., anti-CL II) antibody. For example, Figure 23 and Table 7 show that anti-CD22 antibody alone induces no more than about 10% of cells to exhibit pro-apoptotic behavior (solid bar corresponding to "22" in Fig. 23) and anti-CD79b induces about 26% pro-apoptotic cells (solid bar corresponding to "CD79b" in Fig. 23). In combination, however, anti-CD22 and anti-CD79b induce only about 16% pro-apoptotic cells (slant-striped bar corresponding to "79b" in Fig. 23). Thus, the combined antibodies induce 16% pro-apoptotic cells, which is less than the 38% sum of the individual effects attributable to anti-CD22 (12%) and anti-CD79b (26%). Using this approach, an inspection of Figure 23 and/or Tables 7-8 reveals that anti-CD22 antibody, and by extension a multispecific, multivalent binding molecule comprising an anti-CD22 binding domain, when used in separate combination with each of the following antibodies (or corresponding binding domains): anti-CD19, anti-CD20, anti-CD21, anti-CD23, anti-CD30, anti-CD37, anti-CD40, anti-CD70, anti-CD72, anti-CD79a, anti-CD79b, anti-CD80, anti-CD81, anti-CD86 and anti-MHC class II antibodies/binding domains, will result in an inhibited overall effect.

Without wishing to be bound by theory, the data can be interpreted as indicating that anti-CD22 antibody, or a multispecific, multivalent binding molecule comprising an anti-CD22 binding domain, will protect against, or mitigate an effect of, any of the antibodies listed immediately above. More generally, a multispecific, multivalent binding molecule comprising an anti-CD22 binding domain will inhibit the effect arising from interaction with any of CD19, CD20, CD21, CD23, CD30, CD37, CD40, CD70, CD72, CD 79a, CD79b, CD80, CD81, CD86, and MHC class II molecules. It can be seen in Figure 23 and Table 8 that anti-CD22 antibody, and by extension a binding domain comprising an anti-CD22 binding domain, will function as an inhibitor or mitigator of the activity of any antibody/binding domain recognizing a B-cell surface marker such as a CD antigen. Multivalent binding molecules, including multispecific, multivalent binding molecules, are expected to be useful in refining treatment regimens for a variety of diseases wherein the activity of a binding domain needs to be attenuated or controlled.

In addition to the inhibitory, additive or synergistic combined effect of two binding domains interacting with a target cell, typically through the binding of cell-surface ligands, the experimental results disclosed herein establish that a given pair of binding domains may provide a different type of combined effect depending on the relative concentrations of the two binding domains, thereby increasing the versatility of the invention. For example, Table 8 discloses that anti-CD21 and anti-CD79b interact in an inhibitory manner at the higher tested concentration of anti-CD79b, but these two antibodies interact in a synergistic manner at the lower tested concentration of anti-CD79b. Although some embodiments will use a single type of multivalent binding molecule, i.e., a monospecific, multivalent binding molecule, comprising, e.g., a single CD21 binding domain and a single CD79b binding domain, the invention comprehends mixtures of multivalent binding molecules that will allow adjustments of relative binding domain concentrations to achieve a desired effect, such as an inhibitory, additive or synergistic effect. Moreover, the methods of the invention encompass use of a single multivalent binding molecule in combination with another binding molecule, such as a conventional antibody molecule, to adjust or optimize the relative concentrations of binding domains. Those of skill in the art will be able to determine useful relative concentrations of binding domains using standard

techniques (e.g., by designing experimental matrices of two dilution series, one for each binding domain).

Without wishing to be bound by theory, it is recognized that the binding of one ligand may induce or modulate the surface appearance of a second ligand on the same cell type, or it may alter the surface context of the second ligand so as to alter its sensitivity to binding by a specific binding molecule such as an antibody or a multivalent binding molecule.

Although exemplified herein using B cell lines and antigens, these methods to determine optimally effective multivalent binding molecules (i.e., scorpions) are applicable to other disease settings and target cell populations, including other normal cells, their aberrant cell counterparts including chronically stimulated hematopoietic cells, carcinoma cells and infected cells.

Other signaling phenotypes such as Ca^{2+} mobilization; tyrosine phosphoregulation; caspase activation; NF-κB activation; cytokine, growth factor or chemokine elaboration; or gene expression (e.g., in reporter systems) are also amenable to use in methods of screening for the direct effects of monoclonal antibody combinations.

As an alternative to using a secondary antibody to cross-link the primary antibodies and mimic the multivalent binding molecule or scorpion structure, other molecules that bind the Fc portion of antibodies, including soluble Fc receptors, protein A, complement components including C1q, mannose binding lectin, beads or matrices containing reactive or cross-linking agents, bifunctional chemical cross-linking agents, and adsorption to plastic, could be used to cross-link multiple monoclonal antibodies against the same or different antigens.

Example 12

Multivalent Binding Protein with Effector Function, or Scorpion, Structures

The general schematic structure of a scorpion polypeptide is H2N-binding domain 1-scorpion linker-constant sub-region-binding domain 2. scorpions may also have a hinge-like region, typically a peptide region derived from an antibody hinge, disposed N-terminal to binding domain 1. In some scorpion embodiments, binding

domain 1 and binding domain 2 are each derived from an immunoglobulin binding domain, e.g., derived from a V_L and a V_H. The V_L and a V_H are typically joined by a linker. Experiments have been conducted to demonstrate that scorpion polypeptides may have binding domains that differ from an immunoglobulin binding domain, including an Ig binding domain from which the scorpion binding domain was derived, by amino acid sequence differences that result in a sequence divergence of typically less than 5%, and preferably less than 1%, relative to the source Ig binding domain.

Frequently, the sequence differences result in single amino acid changes, such as substitutions. A preferred location for such amino acid changes is in one or more regions of a scorpion binding domain that correspond, or exhibit at least 80% and preferably 85% or 90%, sequence identity to an Ig complementarity determining region (CDR) of an Ig binding domain from which the scorpion binding domain was derived. Further guidance is provided by comparing models of peptides binding the same target, such as CD20. With respect to CD20, epitope mapping has revealed that the 2H7 antibody, which binds CD20, recognizes the Ala-Asn-Pro-Ser (ANPS) motif of CD20 and it is expected that CD20-binding scorpions will also recognize this motif. Amino acid sequence changes that result in the ANPS motif being deeply embedded in a pocket formed of scorpion binding domain regions corresponding to Ig CDRs are expected to be functional binders of CD20. Modeling studies have also revealed that scorpion regions corresponding to CDR3 (V_L), CDR1-3 (V_H) contact CD20 and changes that maintain or facilitate these contacts are expected to yield scorpions that bind CD20.

In addition to facilitating interaction of a scorpion with its target, changes to the sequences of scorpion binding domains (relative to cognate Ig binding domain sequences) that promote interaction between scorpion binding domain regions that correspond to Ig V_L and V_H domains are contemplated. For example, in a CD20-binding scorpion region corresponding to V_L, the sequence SYIV may be changed by substituting an amino acid for Val (V33), such as His, resulting in the sequence SYIH. This change is expected to improve interaction between scorpion regions corresponding to V_L and V_H domains. Further, it is expected that the addition of a residue at the N-terminus of a scorpion region corresponding to V_H-CDR3 will alter the orientation of that scorpion region, likely affecting its binding characteristics,

because the N-terminal Ser of V_H-CDR3 makes contact with CD20. Routine assays will reveal those orientations that produce desirable changes in binding characteristics. It is also contemplated that mutations in scorpion regions corresponding to V_H-CDR2 and/or V_H-CDR3 will create potential new contacts with a target, such as CD20. For example, based on modeling studies, it is expected that substitutions of either Y105 and W106 (found in the sequence NSYW) in a region corresponding to V_H-CDR3 will alter the binding characteristics of a scorpion in a manner amenable to routine assay for identifying scorpions with modified binding characteristics. By way of additional example, it is expected that an alteration in the sequence of a scorpion binding domain corresponding to an Ig VL-CDR3, such as the Trp (W) in the sequence CQQW, will affect binding. Typically, alterations in a scorpion region corresponding to an Ig CDR will be screened for those scorpions exhibiting an increase in affinity for the target.

Based on the model structure of the humanized CD20 scFv binding domain 20-4, on the published information relating to the CD20 extracellular loop structure (Du, et al., J Biol. Chem. 282(20):15073-80 (2007)), and on the CD20 binding epitope recognized by the mouse 2H7 antibody (which was the source of CDRs for the humanized 20-4 scFv binding domain), mutations were engineered in the CDR regions of the 2Lm20-4x2Lm20-4 scorpion with the aim of improving the affinity of its binding to CD20. First, the mutations were design to influence the 20-4 CDR conformation and to promote more efficient binding to the CD20 extracellular loop. Second, the introduced changes were designed to provide new intermolecular interactions between the 2Lm20-4x2Lm20-4 scorpion and its target. These mutations include: VL CDR1 V33H (i.e., a substitution of His for Val at position 33 of CDR1 in the VL region), VL CDR3 W90Y, VH CDR2 D57E, VH CDR3 insertion of V after residue S99, VH CDR3 Y101K, VH CDR3 N103G, VH CDR3 N104G, and VH CDR3 Y105D. Due to expected synergistic effects of combining some of theses mutations, 11 mutants were designed, combining different mutations as shown in Table 9 (residues introduced by mutation are bolded and underscored).

Table 9

V_L CDR1	V_L CDR3	V_H CDR2	V_H CDR3

RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	**SV**YYSNYWYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	**SV**YY**GG**YWYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SYYSNS**D**WYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SYYS**GGD**WYFDL
RASSSVSYIV	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SY**K**SNSYWYFDL
RASSSVSYIV	QQWSFNPPT	AIYPGNG**E**TSYNQKFKG	SYYSNSYWYFDL
RASSSVSYIV	QQ**Y**SFNPPT	AIYPGNGDTSYNQKFKG	SYYSNSYWYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SY**K**SNS**D**WYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNG**E**TSYNQKFKG	SYYSNS**D**WYFDL
RASSSVSYI**H**	QQ**Y**SFNPPT	AIYPGNGDTSYNQKFKG	SYYSNS**D**WYFDL
RASSSVSYI**H**	QQ**Y**SFNPPT	AIYPGNG**E**TSYNQKFKG	SY**K**S**GGD**WYFDL

Mutations were introduced into binding domains of the CD20xCD20 scorpion (2Lm20-4x2Lm20-4) by PCR mutagenesis using primers encoding the altered sequence region. After sequence confirmation, DNA fragments encoding the 2Lm20-4 scFv fragments with corresponding mutations were cloned into a conventional expression vector containing a coding region for the constant sub-region of a scorpion, resulting in a polynucleotide containing the complete DNA sequence of new versions of the 2Lm20-4x2Lm20-4 scorpion. The variants of the 2Lm20-4x2Lm20-4 scorpion with CDR mutations were produced by expression in a transient COS cell system and purified through Protein A and size-exclusion (SEC) chromatography. The binding properties of 2Lm20-4x2Lm20-4 scorpion variants were evaluated by FACS analysis using primary B-cells and the WIL2-S B-lymphoma cell line.

Other mutants have also been generated using a similar approach to optimize CD20 binding domains. The CD20 SMIP designated TRU015 served as a substrate for generating mutants and, unless noted to the contrary, all domains were human domains. The following mutants were found to contain useful and functional CD20

binding domains. The 018008 molecule contained a substitution of Q (single-letter amino acid code) for S at position 27 of CDR1 in VL, a substitution of S for T at position 28 in CDR1 of VH and a substitution of L for V at position 102 in CDR3 of VH. The following partial scorpion linker sequences, corresponding to the CCCP sequence in an IgG1 hinge, were separately combined with the mutated VL and VH: CSCS, SCCS and SCCP, consistent with the modular design of scorpions. The 018009 molecule contained a substitution of Q for S at position 27 of CDR1 of VL, a substitution of S for T at position 28 of CDR1 of VH and substitutions of S for V at position 96, L for V at position 102 and deletion of the V at position 95, all in CDR3 of VH. The same scorpion linkers sub-sequences described above as being found in the scorpion linkers used in 018008 were used in 018009. The 018010 molecule contained substitutions of a Q for S at position 27, an I for M at position 33 and a V for H at position 34, all in CDR1 of VL, along with an S for T substitution at position 28 of CDR1 of VH and an L for V substitution at position 102 in CDR3 of VH. Scorpion linkers defined by the CSCS and SCCS sub-sequences were used with 018010. 018011 contained the same mutations in CDR1 of VL and in CDR1 of VH as described for 018010, along with deletion of V at position 95, substitution of S for V at position 96 and substitution of L for V at postion 102, all in CDR3 of VH. Scorpion linkers defined by the CSCS, SCCS and SCCP sub-sequences were used in 018011 molecules. The 018014 VL was an unmutated mouse VL, with a human VH containing the S for T change at 28 in CDR1 and the L for V change at 102 in CDR3. 018015 also contained an unmutated mouse VL along with a human VH containing an S for T change at 28 of CDR1 and, in CDR3, a deletion of V at 95, substitution of S for V at 96, and substitution of L for V at 102. The 2Lm5 molecule had a Q for S at 27in CDR1 of VL, an F for Y at 27 and an S for T at 30, both in CDR1 of VH, as well as deletion of the V at 95, S for V at 96 and L for V at 102, all in CDR3 of VH. Scorpion linkers defined by the CSCS, SCCS and SCCP were separately used in each of 018014 and 018015. 2Lm5-1 was the same as 2Lm5 except 2Lm5-1 had no mutations in CDR1 of VH, and only a scorpion linker defined by the CSSS sub-sequence was used. 2Lm6-1 had the mutations of 2Lm5 and a substitution of T for S at 92 and S for F at 93 in CDR3 of VL, and only the scorpion linker defined by the CSSS sub-sequence was used. The only mutations in 2Lm16 were the mutations in CDR3 of VH listed above for 2Lm5-1. Scorpion linkers defined by the sub-sequences CSCS, SCCS, and SCCP were separately used in 2Lm16. 2Lm16-1 substituted Q for

S at 27 in CDR1 of VL and substituted T for S at 92, and S for F at 93, both in CDR3 of VL, and, in CDR3 of VH, deleted V at 95, substituted S for V at 96 and substituted L for V at 102; only the scorpion linker defined by the CSSS sub-sequence was used. 2Lm19-3 substituted Q for S at 27, I for M at 33, and V for H at 34, all in CDR1 of VL, along with the mutations in CDR3 of VH listed for 2Lm16-1. Scorpion linkers defined by the sub-sequences CSCS, SCCS, and SCCP were separately used in 2Lm19-3. The 2Lm20-4 molecule contained an I for M at 33 and a V for H at 34, both in CDR1 of VL, along with the mutations in CDR3 of VH listed for 2Lm16-1. For 2Lm5-1, 2Lm6-1, 2Lm16, 2Lm16-1, 2Lm19-3, and 2Lm20-4, there also was an S for L substitution at position 11 in the framework region of VH. Scorpion linkers defined by the CSCS, SCCS and SCCP sub-sequences were separately used in 2Lm20-4. Finally, the substitution of S for P at position 331 was present in the following mutants: 018008 with the scorpion linker defined by CSCS, 018009 with each of scorpion linkers defined by CSCS and SCCP, 018010 with the scorpion linker defined by CSCS, 018011 with the scorpion linker defined by SCCP, 018014 with the scorpion linker defined by CSCS, 018015 with the scorpion linker defined by CSCS, 2Lm16 with scorpion linkers defined by any of CSCS, SCCS, and SCCP, 2Lm19-3 with a scorpion linker defined by CSCS or SCCP, and 2Lm20-4 with a scorpion linker defined by CSCS or SCCP.

In addition, changes in the length of a linker joining two regions of a binding domain, such as regions of a scorpion binding domain that correspond to an Ig V_L and V_H, are contemplated. For example, removal of a C-terminal Asp in interdomain linkers where it is found is expected to affect the binding characteristics of a scorpion, as is a substitution of Gly for Asp.

Also contemplated are scorpions that have a scorpion linker (interposed C-terminal to the constant sub-region and N-terminal to binding domain 2) that is lengthened relative to a hinge region of an Ig, with amino acid residues being added C-terminal to any cysteine in the scorpion that corresponds to an Ig hinge cysteine, with the scorpion cysteine being capable of forming an interchain disulfide bond. Scorpions containing these features have been constructed and are characterized below.

Efforts were undertaken to improve the expression, stability and therapeutic potency of scorpions through the optimization of the scorpion linker covalently joining the constant sub-region and the C-terminally disposed binding domain 2. The prototypical scorpion used for optimization studies contained an anti-CD20 scFV (binding domain 1) fused N-terminal to the constant sub-region derived from IgG1 C_{H2} and C_{H3}, with a second anti-CD20 scFv fused C-terminal to that constant sub-region. This scorpion, like immunoglobulin molecules, is expected to associate through the constant region (or sub-region) to form a homodimeric complex with peptide chains linked by disulfide bonds. To obtain high level of expression of a stable, tetravalent molecule with high affinity for its CD20 target, the scorpion linker between the constant sub-region and the second binding domain must accommodate the following considerations. First, steric hindrance between the homologous binding domains carried by the two scFv fragments (one scFv fragment on each of two scorpion monomers) should be minimized to facilitate maintenance of the native conformations of each binding domain. Second, the configurations and orientations of binding domains should allow productive association of domains and high-affinity binding of each binding domain to its target. Third, the scorpion linker itself should be relatively protease-resistant and non-immunogenic.

In the exemplary CD20xCD20 scorpion construct S0129, the C-terminus of C_{H3} and the second anti-CD20 scFV domain were linked by the 2H7 scorpion linker, a peptide derived from, and corresponding to, a fragment of a natural human hinge sequence of IgG1. The 2H7 scorpion linker served as a base for design efforts using computer-assisted modeling that were aimed at improving the expression of scorpions and improving the binding characteristics of the expressed molecules.

To analyze the 2H7 scorpion linker, the 3-dimensional structure of a dimeric form of the human IgG1 hinge was modeled using Insight II software. The crystal structure of anti-CD20 scFV in the V_H-V_L orientation was chosen as a reference structure for the 20-4 binding domains (RCSB Protein Data Bank entry code: 1A14). In intact IgG1, the hinge connects the C-terminus of theC_{H1} domain to the N-terminus of the C_{H2} domain, with the configuration of each domain being such that hinge cysteine residues can pair to form a homodimer. In the exemplary scorpion molecule, the hinge-derived 2H7 linker connected the C-terminal end of the scorpion domain

derived from the IgG1 C_{H3} domain to the N-terminal end of that portion of scorpion binding domain 2 derived from an IgG1 V_{H2} domain. Using a 3-D modeled structure of the V_H-V_L scFV, expectations of the optimal distance between the C-terminal ends of the 2H7 linkers was influenced by three considerations. First, hinge stability must be maintained, and stability is aided by dimerization, e.g., homodimerization, which means that the hinge cysteines must be able to pair in the presence of the two folded binding domains. Second, two binding domains, e.g., scFVs, must accommodate the 2H7 linker C-termini without steric interference in order to allow for proper protein folding. Third, the CDRs of each binding domain should be able to face the same direction, as in a native antibody, because each binding domain of the prototypical scorpion can bind adjacent receptors (CD20) on the same cell surface. Given these considerations, the distance between the two N-terminal ends of scFvs is expected to be approximately 28Å. The distance between the C-terminal ends of the theoretically designed 2H7 linkers in dimeric scorpion forms is expected to be about 16Å. To accommodate the distances expected to be needed for optimizing the performance of a scorpion, the C-terminus of the 2H7 linker was extended by at least 3 amino acids. Such an extension is expected to allow for the formation of disulfide bonds between 2H7 linker cysteine residues, to allow for proper folding of the C-terminal binding domain 2, and to facilitate a correct orientation of the CDRs. In addition, in intact IgG1, due to the presence of the C_{H1} and V_{L1} domains between the hinge and binding domains, the distance between the binding domains carried by the two chains is further increased and is expected to further favor the cross-linking of adjacent receptors on the same cell surface. In view of the considerations described above, a set of linkers with different lengths was designed (Table 10). To minimize immunogenicity, natural residues present at the N-terminal end of the C_{H2} domain (Ala-Pro-Glu-Leu or APEL) were used to lengthen the 2H7 scorpion linker by sequence addition to the C-terminus of the scorpion linker. The longer constructs contained one or multiple (Gly4Ser) linker units known to be protease-resistant and flexible.

The CD20xCD20 scorpion constructs containing extended scorpion linkers between the C_{H3} domain of the constant sub-region and the C-terminal scFv binding domain were constructed using PCR mutagenesis and subcloned into a conventional mammalian expression vector. The effect of linker length on CD20xCD20 scorpion

expression could be analyzed be comparing the yield of secreted protein in transient expression experiments using COS or HEK293 cells, or by analysis of protein synthesis and accumulation in the cells by Western blot analyses or pulse-chase studies with [35]S-labeled methionine/cysteine.

Table 10

Construct Number	Scorpion linker core (2H7) sequence	Extension sequence	Extended scorpion linker sequence
1	GCPPCPNS	APEL	GCPPCPNSAPEL
2	GCPPCPNS	APELGGGGS	GCPPCPNSAPELGGGGS
3	GCPPCPNS	APELGGGGSGGGGS	GCPPCPNS APELGGGGSGGGGS
4	GCPPCPNS	APELGGGGSGGGGSGGGGS	GCPPCPNS APELGGGGSGGGGSGGGGS

Glycosylated scorpions are also contemplated and, in this context, it is contemplated that host cells expressing a scorpion may be cultured in the presence of a carbohydrate modifier, which is defined herein as a small organic compound, preferably of molecular weight less than 1000 daltons, that inhibits the activity of an enzyme involved in the addition, removal, or modification of sugars that are part of a carbohydrate attached to a polypeptide, such as occurs during N-linked carbohydrate maturation of a protein. Glycosylation is a complex process that takes place in the endoplasmic reticulum ("core glycosylation") and in the Golgi bodies ("terminal glycosylation"). A variety of glycosidase and/or mannosidase inhibitors provide one or more of desired effects of increasing ADCC activity, increasing Fc receptor binding, and altering glycosylation pattern. Exemplary inhibitors include, but are not

limited to, castanospermine and kifunensine. The effects of expressing scorpions in the presence of at least one such inhibitor are disclosed in the following example.

Example 13

Scorpion protein expression levels and characterization

Scorpion protein expression levels were determined and the expressed proteins were characterized to demonstrate that the protein design led to products having practical benefits. A monospecific CD20xCD20 scorpion and a bispecific CD20xCD37 scorpion were expressed in CHO DG44 cells in culture using conventional techniques.

Basal level, stable expression of the CD20xCD20 scorpion S0129 (2lm20-4x2lm20-4) in CHO DG44 cells cultured in the presence of various feed supplements was observed as shown in Fig. 34. All culture media contained 50 nM methotrexate, a concentration that maintained copy number of the scorpion-encoding polynucleotide. The polynucleotide contained a coding region for the scorpion protein that was not codon-optimized for expression in CHO DG44 cells. The polynucleotide was introduced into cells using the pD18 vector Apparent from Fig. 34, expression levels of about 7-46 μg/ml were obtained.

Expression levels following amplification of the polynucleotide encoding a bispecific CD20xCD37 scorpion were also determined. The pD18 vector was used to clone the CD20xCD37 scorpion coding region and the plasmid was introduced into CHO DG44 cells. Amplification of the encoding polynucleotide was achieved using the *dhfr*-methotrexate technique known in the art, where increasing concentrations of MTX are used to select for increased copy number of the Dihydrofolate Reductase gene (*dhfr*), which leads to co-amplification of the tightly linked polynucleotide of interest. Fig. 35 shows that stable expression levels of about 22-118 μg/ml of the bispecific CD20xCD37 scorpion were typically observed. Variability in yield was seen under different conditions, including methotrexate concentration used for amplification, but these variables are amenable to optimization by those of skill in the art. A variety of other scorpion molecules described herein were also subjected to expression analyses in CHO and/or COS cells, with the results provided in Table 11, below. These results demonstrate that significant yields of scorpion proteins can be

obtained using conventional techniques and routine optimization of the amplification technique.

Expressed proteins were also characterized by SDS-PAGE analysis to assess the degrees of homogeneity and integrity of the expressed proteins and to confirm molecular weight of monomeric peptides. The denaturing polyacrylamide gels (4-20% Tris Glycine) were run under reducing and non-reducing conditions. The results presented in Fig. 36 reveal single protein bands for each of a 2Lm20-4 SCC SMIP and S1000 (CD20(2lm20-4)xCD20(2lm20-4) monospecific scorpion. S0126) of the expected monomeric molecular weights under reducing conditions. These data establish that SMIPs and scorpions are amenable to purification in an intact form. Under non-reducing conditions, a trace amount of a peptide consistent with the expected size of a monomeric SMIP was seen, with the vast majority of the protein appearing in a single well-defined band consistent with a dimeric structure. Under these non-reducing conditions, the monospecific scorpion protein showed a single well-defined band of a molecular weight consistent with a dimeric structure. The dimeric structures for both the SMIP and the scorpion are consistent with their monomeric structures, each of which contains a hinge-like scorpion linker containing at least one Cysteine capable of participating in disulfide bond formation.

The effect of scorpion linkers on the expression and integrity of scorpions was also assessed, and results are shown in Table 12. This table lists scorpion linker variants of the monospecific CD20xCD20 (2Lm20-4x2Lm20-4) S0129 scorpion and the CD20xCD28 S0033 scorpion (2H7sccpIgG1-H7-2e12), their integrity as single chain molecules, and their transient expression levels in COS cells relative to the parent scorpion S0129 or S0033, as appropriate, with an H7 linker (set as 100%). Table 13 provides data resulting from an evaluation of scorpion linker variants incorporated into the CD20xCD20 scorpion, along with analogous data for the CD20xCD28 scorpion. Table 13 provides data resulting from an evaluation of S0129 variants containing scorpion linkers that are not hinge-like linkers containing at least one Cysteine capable of disulfide bond formation; rather, the scorpion linkers in these molecules are derived from Type II C-lectin stalks. Apparent from the data presented in Table 13 is that hinge-like scorpion linkers may be associated with scorpions expressed at higher or lower levels than an unmodified parent scorpion linker in

transient expression assays. Further, some of the linker variants exhibit greater resistance to proteolytic cleavage than the unmodified parent linker, a concern for all or almost all expressed proteins. The data of Table 13 show that non-hinge-like linkers such as linkers derived from the stalk region of Type II C-lectins are found in scorpions that exhibit binding characteristics that vary slightly from scorpions containing hinge-like scorpion linkers. Additionally, the scorpion containing a non-hinge-like scorpion linker exhibits effector function (ADCC) that either equals or exceeds the ADCC associated with scorpions having hinge-like scorpion linkers.

Table 11

Linker Name	Upstream (CH3) Sequence	S0129 (2Lm20-4 x 2Lm20-4) linker variants - aa seq[1]	based on	#AAs	Expression COS[2]	Cleavage[3]	Expression CHO[2]
H7	QKSLSLSPGK	GCPPCPNS	H7	18	100	-	100
H16	QKSLSLSPGK	LSVKADFLTPSIGNS	CD80	25	174	+	
H18	QKSLSLSPGK	LSVLANFSQPEIGNS	CD86	25	165	++	
H19	QKSLSLSPGK	LSVLANFSQPEISCPPCPNS	CD86 + H7	30	161	+	109
H28	QKSLSLSPGK	RIHQMNSELSVLANS	CD86	25	170	++	
H32	QKSLSLSPGK	RIHLNVSERPFPPNS	CD22	25	184	++	
H47	QKSLSLSPG	LSVKADFLTPSIGNS	H16	24	141	-	206
H48	QKSLSLSPG	KADFLTPSIGNS	H16	21	137	-	
H50	Q	LSVLANFSQPEIGNS	H18	36	21	-	
H51	QKS	LSVLANFSQPEIGNS	H18	18	116	-	
H52	QKSLSLSPG	SQPEMPISNS	H18	20	95	-	
H53	QKSLSL	SQPEMPISCPPCPNS	H19	26	93	-	
H54	Q	SVLANFSQPEISCPPCPNS	H19	31	72	+/-	
H55	QKSLSLSPG	RIHQMNSELSVLANS	H28	24	118	+	
H56	QKSLSLSPG	QMNSELSVLANS	H28	21	130	-	183
H57	QKSLSLSPG	VSERPFPPNS	H32	19	118	-	
H58	QKSLSLSPG	KPFFTCGSADTCPNS	CD72	24	103	-	
H59	QKSLS	KPFFTCGSADTCPNS	CD72	20	94	-	

[1]NPS is a glycosylation consensus motif

[2]Transient expression in COS (6W plates), or CHO (single flask) relative to S0129-H7 (%)

[3]Cleavage products observed by SDS-PAGE/silver stain: -=none, +=faint band, ++=major band(s), +++>50% cleaved

Table 12

Linker Name	S0129 (2Lm20-4 x 2Lm20-4) linker variants - aa seq	Changes in CH3?[1]	Linker seq. based on	20x20 Expression[2]	20x20 Cleavage?[3]
H7	GCPPCPNS	N	H7	100	-
H8	GSPPSPNS	N	H7	107	+
H9	GSPPSPNS	Y	H7	142	-
H10	EPKSTDKTHTCPPCPNS	N	IgG1 hinge	98	-
H11	EPKSTDKTHTSPPSPNS	N	IgG1 hinge	126	+
H16	LSVKADFLTPSIGNS	N	CD80	174	+
H17	LSVKADFLTPSISCPPCPNS	N	CD80 + H7	113	+
H18	LSVLANFSQPEIGNS	N	CD86	165	++
H19	LSVLANFSQPEISCPPCPNS	N	CD86 + H7	161	+

H20	LKIQERVSKPKISNS	N	CD2	115	+++
H21	LKIQERVSKPKISCPPCPNS	N	CD2 + H7	90	+++
H22	LNVSERPFPPHIQNS	N	CD22	149	++
H23	LDVSERPFPPHIQSCPPCPNS	N	CD22 + H7	121	++
H24	REQLAEVTLSLKANS	N	CD80	145	++
H25	REQLAEVTLSLKACPPCPNS	N	CD80 + H7	98	+
H26	RIHQMNSELSVLANS	N	CD86	170	++
H27	RIHQMNSELSVLACPPCPNS	N	CD86 + H7	154	++
H28	DTKGKNVLEKIFSNS	N	CD2	153	+
H30	LPPETQESQEVTLNS	N	CD22	78	+
H32	RIHLNVSERPFPPNS	N	CD22	184	++
H33	RIHLNVSERPFPPCPPCPNS	N	CD22 + H7	74	+
H36	GCPPCPGGGGSNS	N	H7	110	+
H40	GCPPCPANS	Y	H7	110	+
H41	GCPPCPANS	Y	H7	102	-
H42	GCPPCPNS	Y	H7	99	-
H44	GGGASCPPCPGNS	Y	H7	108	+
H45	GGGASCPPCAGNS	Y	H7	107	-
H46	GGGASCPPCANS	Y	H7	98	-
H47	LSVKADFLTPSIGNS	Y	CD80	141	-
H48	ADFLTPSIGNS	N	CD80	137	-
H50	LSVLANFSQPEIGNS	Y	CD86	21	-
H51	LSVLANFSQPEIGNS	Y	CD86	110	-
H52	SQPEIVPISNS	Y	CD86	95	-
H53	SQPEIVPISCPPCPNS	Y	CD86 + H7	95	-
H54	SVLANFSQPEISCPPCPNS	Y	CD86 + H7	72	+/-
H55	RIHQMNSELSVLANS	Y	CD86	118	+
H56	QMNSELSVLANS	Y	CD86	130	-
H57	VSERPFPPNS	Y	CD22	118	-
H58	KPFFTCGSADTCPNS	Y	CD72	103	-
H59	KPFFTCGSADTCPNS	Y	CD72	94	-
H60	QYNCPGQYTFSMNS	Y	CD69	>100[5]	-
H61	EPAFTPGPNIELQKDSDCNS	Y	CD94	>100	-
H62	QRHNNSSLNTRTQKARHCNS	Y	NKG2A	>100	-
H63	NSLFNQEVQIPLTESYCNS	Y	NKG2D	>100	-

[1]Additional changes to the end of CH3 such as 1-9 aa deletion and/or codon optimization

[2]Transient expression in COS (6W plates), relative to S0129-H7 parent (%)

[3]Cleavage product(s) observed by SDS-PAGE/silver stain:

	-=none, +=faint band, ++=major band(s), +++>50% cleaved	
[5]H60-H63 variants compared by estimation of recovery of protein purified from COS spent media.		

Table 13

Proteins	**Description**	**Production Yield (ug protein purified/ml sup)**	**% POI (M.wt in Kd by MALS)**	**Improvement over S0129wt POI**	Binding to Ramos	ADCC assay	Sequence of scorpion linker
S0129wt	H7 linker	1.6	67 (167)	-	-	-	GCPPC
S0129-CD69	CD69 stalk	2.9	66 (167)	1.8	Weaker than S0129 wt	*Slightly better than S0129wt POI	QYNCPGQYTFSM
S0129-CD72	CD72 truncated stalk	2.0	69 (165)	1.2	Similar to S0129wt	*Slightly better than S0129wt POI	PFFTCGSADTC
S0129-CD94	CD94 stalk	2.9	67 (171)	1.8	Similar to S0129wt	*Slightly better than S0129wt POI	EPAFTPGPNIELQKDSDC
S0129-NKG2A	NKG2A stalk	2.5	**93 (170)**	2.2	Slightly better than S0129wt	Similar to S0129wt POI	QRHNNSSLNTRTQKARHC
S0129-NKG2D	NKG2D stalk	1.9	70 (166)	1.2	Similar to S0129wt	*Slightly better than S0129wt POI	NSLFNQEVQIPLTESYC

As noted in the preceding example, production by expression of scorpions in cultures containing a carbohydrate modifier is contemplated. In exemplary embodiments, castanospermine (MW 189.21) is added to the culture medium to a final concentration of about 200 μM (corresponding to about 37.8 μg/mL), or concentration ranges greater than about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 μM, and up to about 300, 275, 250, 225, 200, 175, 150, 125, 100, 75, 60, or 50 μg/mL. For example, ranges of 10-50, or 50-200, or 50-300, or 100-300, or 150-250 μM are contemplated. In other exemplary embodiments, DMJ, for example DMJ-HCl (MW 199.6) is added to the culture medium to a final concentration of about 200 μM (corresponding to about 32.6 μg DMJ/mL), or

concentration ranges greater than about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 μM, and up to about 300, 275, 250, 225, 200, 175, 150, 125, 100, 75, 60, or 50 μg/mL. For example, ranges of 10-50, or 50-200, or 50-300, or 100-300, or 150-250 μM are contemplated. In other exemplary embodiments, kifunensine (MW 232.2) is added to the culture medium to a final concentration of about 10 μM (corresponding to about 2.3 μg/mL), or concentration ranges greater than about 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 μM, and up to about 50, 45, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, or 11 μM. For example, ranges of 1-10, or 1-25, or 1-50, or 5-10, or 5-25, or 5-15 μM are contemplated.

In one experiment, a monospecific CD20xCD20 scorpion (S0129) was expressed in cells cultured in 200 μM castanospermine (S0129 CS200) or 10 μM (excess) kifunensine (S0129 KF 10) and the binding, or staining, of WIL2S cells by the expressed scorpion was measured, as shown in Fig. 42. In comparative binding studies, moreover, a glycosylated S0129 scorpion bound CD16 (FCγRIII) approximately three times better than the unglycosylated S0129 scorpion.

In another study, the ADCC-mediated killing of BJAB B-cells by humanized CD20xCD20 scorpion (S0129) was explored. The results shown in Fig. 43 establish that the scorpion, when expressed in cells being cultured in the presence of either castanospermine or kifunensine, led to significantly more potent ADCC-mediatd BJAB B-cell death for a given concentration of scorpion exposure.

Example 14

Scorpion binding

a. Domain spacing

Bispecific scorpions are capable of binding at least two targets simultaneously, utilizing the pairs of binding domains at the N- and C-terminus of the molecule. In so doing, for cell-surface targets, the composition can cross-link or cause the physical co-approximation of the targets. It will be appreciated by those skilled in the art that many receptor systems are activated upon such cross-linking, resulting in signal induction causing changes in cellular phenotype. The design of the compositions disclosed herein was intended, in part, to maximize such signaling and to control the resultant phenotype.

Approximate dimensions of domains of the scorpion compositions, as well as expectations of interdomain flexibility in terms of ranges of interdomain angles, are known and were considered in designing the scorpion architecture. For scorpions using scFv binding domains for binding domains 1 and 2 (BD1 and BD2), an IgG1 N-terminal hinge (H1), and the H7 PIMS linker described herein, the binding domain at the N-terminus and the binding domain at the C-terminus may be maximally about 150-180Å apart and minimally about 20-30Å apart. Binding domains at the N-terminus may be maximally about 90-100Å apart and minimally about 10-20Å apart (Deisenhofer, et al., 1976, Hoppe-Seyler's Z. Physiol. Chem. Bd. 357, S. 435-445; Gregory, et al., 1987, Mol. Immunol. 24(8):821-9.; Poljak, et al., 1973, Proc. Natl. Acad. Sci., 1973, 70: 3305-3310; Bongini, et al., 2004, Proc. Natl. Acad. Sci. 101: 6466-6471; Kienberger, et al., 2004, EMBO Reports, 5: 579 – 583, each incorporated herein by reference). The choice of these dimensions was done in part to allow for receptor-receptor distances of less than about 50Å in receptor complexes bound by the scorpion as distances less than this may be optimal for maximal signaling of certain receptor oligomers (Paar, et al., 2002, J. Immunol., 169: 856-864, incorporated herein by reference) while allowing for the incorporation of F_C structures required for effector function.

The binding domains at the N- and C-terminus of scorpions were designed to be flexible structures to facilitate target binding and to allow for a range of geometries of the bound targets. It will also be appreciated by those skilled in the art that flexibility between the N- or C-terminal binding domains (BD1 and BD2, respectively) and between the binding domains and the F_C domain of the molecule, as well as the maximal and minimal distances between receptors bound by BD1 and/or BD2, can be modified, for example by choice of N-terminal hinge domain (H1) and, by structural analogy, the more C-terminally located scorpion linker domain (H2). For example hinge domains from IgG1, IgG2, IgG3, IgG4, IgE, IgA2, synthetic hinges and the hinge-like C_{H2} domain of IgM show different degrees of flexibility, as well as different lengths. Those skilled in the art will understand that the optimal choice of H1 and scorpion linker (H2) will depend upon the receptor system(s) the scorpion is designed to interact with as well as the desired signaling phenotype induced by scorpion binding.

In some embodiments, scorpions have a scorpion linker (H2) that is a hinge-like linker corresponding to an Ig hinge, such as an IgG1 hinge. These embodiments include scorpions having an amino acid sequence of the scorpion hinge that is an N-terminally extended sequence relative to, e.g., the H7 sequence or the wild-type IgG1 hinge sequence. Exemplary scorpion linkers of this type would have the sequence of the H7 hinge N-termanally extended by H_2N-APEL$(x)_y$-CO_2H, where x is a unit of the Gly_4Ser linker and y is a number between 0 and 3. Exemplifying the influence of the scorpion linker on scorpion stability is a study done using two scorpions, a bispecific CD20xCD28 scorpion and a monospecific CD20xCD20 scorpion. For each of these two scorpion designs, a variety of scorpion linkers were inserted. In particular, scorpion linkers H16 and H17, which primarily differ in that H17 has the sequence of H16 with the sequence of H7 appended at the C-terminus, and scorpion linkers H18 and 19, in which analogously the sequence of H7 is appended at the C-terminus of H18 in generating H19. For each of the two scorpion backbones (20x28 and 20x20), each of the four above-described scorpion linkers were inserted at the appropriate location. Transient expression of these constructs was obtained in COS cells and the scorpion proteins found in the culture supernatants were purified on protein A/G-coated wells (Pierce SEIZE IP kit). Purified proteins were fractionated on SDS-PAGE gels and visualized by silver stain. Inspection of Fig. 44 reveals that the additional H7 sequence in the scorpion linker adds to the stability of each type of scorpion linker and each type of scorpion protein. In other words, appending H7 to the C-terminus of either H16 or H18 added to the stability of the scorpion molecule, and this observation held regardless of whether the scorpion was CD20xCD28 or CD20xCD20. In terms of target binding, the scorpion proteins having the CD20xCD20 architecture exhibited similar binding properties to the parent monospecific humanized CD20xCD20 scorpion S0129, as shown in Fig. 45.

Beyond the preceding embodiments, however, it may be desirable to prevent bound receptors from approaching within about 50Å of each other to intentionally create submaximal signals (Paar, et al., J. Immunol., 169: 856-864). In such a case, choices of H1 and Scorpion linker (H2) that are shorter and less flexible than those described above would be expected to be appropriate.

The same spacing considerations apply to scorpion linkers that are not hinge-like. These scorpion linkers are exemplified by the class of peptides having the amino acid sequence of a stalk region of a C-lectin. Exemplary scorpion hinges comprising a C-lectin stalk region are scorpion hinges derived from the CD72 stalk region, the CD94 stalk region, and the NKG2A stalk region. Scorpions containing such scorpion hinges were constructed and characterized in terms of expression, susceptibility to cleavage, and amenability to purification. The data are presented in Table 14.

Table 14

Linker Name	G_4S Codon optimization[1]	End of CH3	S0129 Scorpion Linker variants amino acid seq	Linker seq. based on	Expression (%S0129)[2]	Cleavage[3]	Bench-top purification %POI
H7	N	K	GCPPCPNS	H7	100	-	70
H60	Y(17)	K	GCPPCPNS	H7	114	-	ND
H61	Y(15)	K	GCPPCPNS	H7	90	-	66
H62	N	G	QRHNNSSLNTRTQKARHCPNS	NKG2A stalk	129	-	89
H63	Y(17)	G	QRHNNSSLNTRTQKARHCPNS	NKG2A stalk	100	-	85
H64	Y(15)	G	QRHNNSSLNTRTQKARHCPNS	NKG2A stalk	81	-	83
H65	N	G	EPAFTPGPNIELQKDSDCPNS	CD94 stalk	133	-	66
H66	Y(17)	G	EPAFTPGPNIELQKDSDCPNS	CD94 stalk	200	-	64
H67	Y(15)	G	EPAFTPGPNIELQKDSDCPNS	CD94 stalk	129	-	65
H68	N	G	RTRYLQVSQQLQQTNRVLEVTNSSLRQQLR LKITQLGQSAEDLQGSRRELAQSQEALQVEQ RAHQAAEGQLQACQADRQKTKETLQSEEQ QRRALEQKLSNMENRLKPFFTCGSADTC	CD72 full stalk	110	-	75

[1]Codon optimization of Gly_4Ser linker, with (17) or without (15) restriction site

[2]Estimate of expression in COS based on recovery of protein in benchtop purification

[3]Cleavage product(s) observed by SDS-PAGE/Coomassie Blue stain of purified protein

b. Binding of N- and C-terminal binding domains

Both N- and C-terminal domains participate in target cell binding

The target cell binding abilities of a CD20 SMIP (TRU015), a CD37 SMIP (SMIP016), a combination of CD20 and CD37 SMIPS (TRU015+SMIP016), and the CD20xCD37 bispecific scorpion (015x016), were assessed by measuring the capacity of each of these molecules to block the binding of an antibody specifically competing for binding to the relevant target, either CD37 or CD20. The competing antibodies were FITC-labeled monoclonal anti-CD37 antibody or PE-labeled monoclonal anti-CD20 antibody, as appropriate. Ramos B-cells provided the targets.

Ramos B-cells at $1.2x10^7$/ml in PBS with 5% mouse sera (#100-113, Gemini Bio-Products, West Sacramento, CA) (staining media) were added to 96-well V-bottom plates (25 µl/well). The various SMIPs and scorpions were diluted to 75

μg/ml in staining media and 4-fold dilutions were performed to theconcentrations indicated in Fig. 38. The diluted compounds were added to the plated cells in addition to media alone for control wells. The cells were incubated for 10 minutes with the compounds and then FITC anti-CD37 antibody (#186-040, Ancell, Bayport, MN) at 5 μg/ml and PE anti-CD20 antibody (#555623, BD Pharmingen, San Jose, CA) at 3 μg/ml (neet) were added together to the wells in 25 μl staining media. The cells were incubated on ice in the dark for 45 minutes and then washed 2.5 times with PBS. Cells were fixed with 1% paraformaldehyde (#19943 1 LT, USB Corp, Cleveland, OH) and then run on a FACs Calibur (BD Biosciences, San Jose, CA). The data were analyzed with Cell Quest software (BD Biosciences, San Jose, CA). The results shown in Fig. 38 establish that all SMIPs, SMIP combinations and scorpions containing a CD20 binding site successfully competed with PE-labeled anti-CD 20 antibody for binding to Ramos B-cells (upper panel); all SMIPs, SMIP combinations and scorpions containing a CD37 binding site successfully competed with FITC-labeled anti-CD 37 antibody for binding to Ramos B-cells (lower panel). The bispecific CD20xCD37 scorpion, therefore, was shown to have operable N- and C-terminal binding sites for targets on B-cells.

c. Cell-surface persistence

An investigation of the cell-surface persistence of bound SMIPs and scorpions (monospecific and bispecific) on the surface of B-cells revealed that scorpions exhibited greater cell-surface persistence than SMIPs. Ramos B-cells at $6x10^6$ /ml ($3x10^5$/well) in staining media (2.5% goat sera, 2.5% mouse sera in PBS) were added to 96-well V-bottom plates. Test reagents were prepared at two-fold the final concentration in staining media by making a 5-fold serial dilution of a 500 nM initial stock and then were added 1:1 to the Ramos B-cells. In addition, media controls were also plated. The cells were incubated in the dark, on ice, for 45 minutes. The plates were then washed 3.5 times with cold PBS. The secondary reagent, FITC goat anti-human IgG (#H10501, Caltag/Invitrogen, Carlsbad,CA) was then added at a 1:100 dilution in staining media. The cells were incubated for 30 minutes in the dark, on ice. Cells were then washed 2.5 times by centrifugation with cold PBS, fixed with a 1% paraformaldehyde solution (#19943 1 LT ,USB Corp, Cleveland, OH) and then run on a FACs Calibur (BD Biosciences, San Jose, CA). The data were analyzed with

CellQuest software (BD Biosciences, San Jose, CA). Results of the data analysis are presented in Fig. 37, which shows the binding of several SMIPs, a monospecific CD20xCD20 scorpion and a bispecific CD20xCD37 scorpion to their targets on Ramos B cells.

Two tubes of Ramos B-cells ($7x10^5$/ml) were incubated for 30 minutes on ice with each of the two compounds being investigated, i.e., a humanized CD20 (2Lm20-4) SMIP and a humanized CD20xCD20 (2Lm20-4x2Lm20-4) scorpion, each at 25 μg/ml in Iscoves media with 10% FBS. At the end of the incubation period, both tubes were washed 3 times by centrifugation. One tube of cells was then plated into 96-well flat-bottom plates at $2x10^5$ cells/well in 150 μl of Iscoves media with one plate then going into the 37°C incubator and the other plate incubated on ice. The second tube of each set was resuspended in cold PBS with 2% mouse serum and 1% sodium azide (staining media) and plated into a 96-well V-bottom plate at $2x10^5$ cells/well for immediate staining with the secondary antibody, i.e., FITC goat anti-human IgG (#H10501, Caltag/Invitrogen, Carlsbad, CA). The secondary antibody was added at a 1:100 final dilution in staining media and the cells were stained on ice, in the dark, for 30 minutes. Cells were then washed 2.5 times with cold PBS, and fixed with 1% paraformaldehyde (#19943 1 LT, USB Corp, Cleveland, OH).

At the time points designated in Fig. 39, samples were harvested from the 96-well flat-bottom plates, incubated at either 37°C or on ice, and placed into 96-well V-bottom plates ($2x10^5$ cells/well). The cells were washed once with cold staining media, resuspended, and the secondary antibody was added at a final dilution of 1:100 in staining media. These cells were incubated on ice, in the dark, for 30 minutes. The cells were then washed 2.5 times by centrifugation in cold PBS, and subsequently fixed with 1% paraformaldehyde. The samples were run on a FACS Calibur (BD Biosciences, San Jose, CA) and the data was analyzed with CellQuest software (BD Biosciences, San Jose, CA). Results presented in Fig. 39 demonstrate that the binding of a SMIP and a scorpion to the surface of B-cells persists for at least six hours, with the monospecific hu CD20xCD20 (2Lm20-4x2Lm20-4) scorpion persisting to a greater extent than the hu CD20 (2Lm20-4) SMIP.

Example 15

Direct cell killing by monospecific and bispecific scorpions

Experiments were conducted to assess the capacity of monospecific and bispecific scorpion molecules to directly kill lymphoma cells, i.e., to kill these cells without involvement of ADCC or CDC. In particular, the Su-DHL-6 and DoHH2 lymphoma cell lines were separately subjected to a monospecific scorpion, i.e., a CD20xCD20 scorpion or a CD37xCD37 scorpion, or to a bispecific CD20xCD37 scorpion.

Cultures of Su-DHL-6, DoHH2, Rec-1, and WSU-NHL lymphoma cells were established using conventional techniques and some of these cultures were then individually exposed to a monospecific CD20 SMIP, a monospecific scorpion (CD20xCD20 or CD37xCD37), or a bispecific scorpion (CD20xCD37 or CD19xCD37). The exposure of cells to SMIPs or scorpions was conducted under conditions that did not result in cross-linking. The cells remained in contact with the molecules for 96 hours, after which growth was measured by detection of ATP, as would be known in the art. The cell killing attributable to the CD20 SMIP and the CD20xCD20 monospecific scorpion are apparent in Fig. 24 and Table 15. The cell killing capacity of the CD37xCD37 monospecific scorpion is apparent from Fig. 25 and Table 15, the ability of the CD20xCD37 bispecific scorpion to kill lymphoma cells is apparent from Fig. 26 and Table 15, and the capacity of the CD19xCD37 bispecific scorpion to kill lymphoma cells is evident from Fig. 27 and Table 15. Data were pooled from three independent experiments and points represent the mean ± SEM. IC_{50} values in Table 15 were determined from the curves in Figs. 24, 25, and 26, as noted in the legend to Table 15, and are defined as the concentration resulting in 50% inhibition compared to untreated cultures. The data in the figures and table demonstrate that scorpions are greater than 10-fold more potent in killing these cell lines than the free SMIP using the same binding domains.

Table 15

	Cell Line		
IC_{50} (nM)	SU-DHL-6	DoHH2	WSU-NHL

	Cell Line		
CD20 SMIP*	>100	60	NA
CD20xCD20 scorpion*	0.3	4.0	NA
CD37 SMIP**	>100	>100	NA
CD37xCD37 scorpion**	10	1.2	NA
CD20 SMIP and CD37 SMIP***	6	2	NA
CD20xCD37 scorpion***	0.05	0.05	NA
CD19 SMIP and CD37 SMIP****	0.16	NA	0.40
CD19xCD37 scorpion****	0.005	NA	0.04

* Data derived from Fig. 24.

** Data derived from Fig. 25.

*** Data derived from Fig. 26.

**** Data derived from Fig. 27.

Additional experiments with the humanized CD20xCD20 scorpion S0129 were conducted in Su-DHL-4, Su-DHL-6, DoHH2, Rec-1, and WSU-NHL cells. The results are presented in Fig. 46 and Fig. 47. The data provided in these figures extends the findings discussed above in showing that scorpions have the capacity to directly kill a variety of cell lines.

The above findings were extended to other monospecific and bispecific scorpions, with each scorpion demonstrating capacity to directly kill B cells. DoHH2 B-cells were exposed *in vitro* to the monospecific CD20xCD20 scorpion, a

monospecific CD37xCD37 scorpion, or a bispecific CD20xCD37 scorpion. The results presented in Fig. 48 demonstrate that bispecific scorpions have kill curves that are different in form from monospecific scorpions.

Culturing Su-DHL-6 cells in the presence of 70 nM CD20xCD20 scorpion (S0129), CD20xCD37 scorpion, or CD37xCD37 scorpion also led to direct B-cell killing in an *in vitro* environment (Fig. 49). Consistently, Su-DHL-6 cells exposed to either a bispecific CD19xCD37 scorpion or to Rituxan® led to direct cell killing, with the bispecific scorpion exhibiting lethality at lower doses, as revealed in Fig. 50.

Another demonstration of direct cell killing was provided by exposing DHL-4 cells to four independent monospecific scorpions recognizing CD20. Two versions of CD20xCD20 scorpion were designed to incorporate two 20-4 binding domains (20-4x20-4 and S0129) and the second two incorporate a hybrid of the 011 and 20-4 binding domains. All four of the independently constructed and purified versions of the two CD20xCD20 scorpion designs, (20-4x20-4 and S0129) and hybrid (011x20-4 and 011x20-4ΔAsp), efficiently killed the DHL-4 cells in a direct manner. For this study, DHL-4 cells were treated *in vitro* with 1 μg/ml of the indicated proteins for 24 hours. Cells were then stained with Annexin V and Propidium Iodide, early and late markers of cell death, respectively, and cell populations were quantified by FACS. The results presented in Fig. 51 establish the direct killing capacity of each of the CD20xCD20 constructs as evidenced by increased staining shown in black bars. In addition, the results demonstrate that the hybrid 011x20-4 proteins exhibited a slight increase in direct cell killing relative to 20-4x20-4-based scorpions, despite the fact that each of these scorpions monospecifically recognized CD20. In a separate set of experiments, the dose-response of the four independent scorpion constructs was determined by FACS analysis of Annexin V- and Propidium Iodide-stained cell populations. The results, shown in Fig. 52, demonstrate dose-responsive increases in cell death resulting from treatment of the DHL-4 cells with each of the independent scorpion constructs.

Example 16

Accessory functions mediated by scorpions (ADCC & CDC)

a. Scorpion-dependent cellular cytotoxicity

Experiments were conducted to determine whether scorpions would mediate the killing of BJAB B lymphoma cells. BJAB B lymphoma cells were observed to be killed with CD20 and/or CD37 scorpions.

Initially, $1x10^7$/ml BJAB B-cells were labeled with 500 μCi/ml ^{51}Cr sodium chromate (#CJS1, Amersham Biosciences, Piscataway, NJ) for 2 hours at 37°C in Iscoves media with 10% FBS. The ^{51}Cr-loaded BJAB B cells were then washed 3 times in RPMI media with 10% FBS and resuspended at $4x10^5$/ml in RPMI. Peripheral blood mononuclear cells (PBMC) from in-house donors were isolated from heparinized whole blood via centrifugation over Lymphocyte Separation Medium (#50494, MP Biomedicals, Aurora, Oh), washed 2 times with RPMI media and resuspended at $5x10^6$/ml in RPMI with 10% FBS. Reagent samples were added to RPMI media with 10% FBS at 4 times the final concentration and three 10-fold serial dilutions for each reagent were prepared. These reagents were then added to 96-well U- bottom plates at 50 μl/well to the indicated final concentrations. The ^{51}Cr-labeled BJAB were then added to the plates at 50 μl/well ($2x10^4$/well). The PBMC were then added to the plates at 100 μl/well ($5x10^5$/well) for a final ratio of 25:1 effectors (PBMC):target (BJAB). Effectors and targets were added to media alone to measure background killing. The ^{51}Cr-labeled BJAB were added to media alone to measure spontaneous release of ^{51}Cr and to media with 5% NP40 (#28324, Pierce, Rockford, Ill) to measure maximal release of ^{51}Cr. The plates were incubated for 6 hours at 37°C in 5%CO_2. Fifty μl (25 μl would also be suitable) of the supernatant from each well were then transferred to a LumaPlate-96 (#6006633, Perkin Elmer, Boston, Mass) and dried overnight at room temperature.

After drying, radioactive emissions were quantitated as cpm on a Packard TopCount-NXT. Sample values were the mean of triplicate samples. Percent specific killing was calculated using the following equation: % Kill = ((sample – spontaneous release)/(maximal release – spontaneous release)) x 100. The plots in Fig. 30 show that BJAB B cells were killed by monospecific scorpions CD20xCD20 and CD37xCD37. The combination of CD20 SMIP and CD37 SMIP also killed BJAB B

cells. These results demonstrate that scorpions exhibit scorpion-dependent cellular cytotoxicity and it is expected that this functionality is provided by the constant sub-region of the scorpion, providing ADCC activity.

b. Scorpion role in complement-dependent cytotoxicity

Experiments also demonstrated that scorpions have Complement-Dependent Cytotoxicity (CDC) activity. The experiment involved exposure of Ramos B-cells to CD19 and/or CD37 SMIPs and scorpions, as described below and as shown in Fig. 31.

The experiment was initiated by adding from 5 to 2.5 x 10^5 Ramos B-cells to wells of 96-well V-bottomed plates in 50 µl of Iscoves media (no FBS). The test compounds in Iscoves, (or Iscoves alone) were added to the wells in 50 µl at twice the indicated final concentration. The cells and reagents were incubated for 45 minutes at 37°C. The cells were washed 2.5 times in Iscoves with no FBS and resuspended in Iscoves with human serum (# A113, Quidel, San Diego, CA) in 96-well plates at the indicated concentrations. The cells were then incubated for 90 minutes at 37°C. The cells were washed by centrifugation and resuspended in 125 µl cold PBS. Cells were then transferred to FACs cluster tubes (#4410, CoStar, Corning, NY) and 125 µl PBS with propidium iodide (# P-16063, Molecular Probes, Eugene, OR) at 5 µg/ml was added. The cells were incubated with the propidium iodide for 15 minutes at room temperature in the dark and then placed on ice, quantitated, and analyzed on a FACsCalibur with CellQuest software (Becton Dickinson). The results presented in Fig. 31 establish that the CD19 SMIP, but not the CD37 SMIP, exhibits CDC activity, with a combination of the two SMIPs exhibiting approximately the same level of CDC activity as CD19 SMIP alone. The CD19xCD37 scorpion, however, exhibited significantly greater CDC activity than either SMIP alone or in combination, establishing that the scorpion architecture provides a greater level of Complement-dependent Cytotoxicity than other molecular designs.

c. ADCC/CDC activity of CD20xCD20 monospecific scorpions

Three distinct CD20xCD20 monospecific scorpions were examined for ADCC and CDC functionality, along with appropriate controls. ADCC was assayed using conventional techniques, and the results are presented in Fig. 53. Apparent from the

Figure is the appreciable, but not identical, ADCC activity associated with each of the tested CD20xCD20 monospecific scorpions.

To assess CDC, Ramos B-cell samples ($4x10^5$) were incubated with each of the CD20xCD20 scorpions (0, 0.5, 5, 50 and 500 nM) and serum (10%) for 3.5 hour at 37°C. Cell death was assessed by 7-AAD staining and FACS analysis. The results are presented in Fig. 54, which reveals that the scorpions exhibit some CDC activity. In a similar experiment, Ramos B-cell samples ($4x10^5$) were incubated with CD20xCD20 scorpion protein (5, 50, 100 nM) and serum (10%) for 2 hour at 37°C. Cells were washed 2X and incubated with anti-human C1q FITC antibody. Bound C1q was assessed by FACS analysis and the results are presented in Fig. 55. These results are consistent with the results presented in Fig. 54 that each of the CD20xCD20 monospecific scorpions was associated with some CDC activity, although less activity than was associated with a CD20 SMIP.

d. Interactions of scorpions with $F_C\gamma$RIII

ELISA studies showed that scorpions bound to FcγRIII (CD16) low (a low affinity isoform or allelotype) at increased levels in the absence of target cells. ELISA plates were initially coated with either low- or high-affinity CD16mIgG using conventional techniques. The ability of this immobilized fusion protein to capture either a CD20 SMIP or a CD20xCD20 monospecific scorpion was assessed. Bound SMIPs and scorpions were detected with goat anti-human IgG (HRP) secondary antibody and mean fluorescence intensity (MFI) was determined. PBS alone (negative control) is shown as a single point. The results are presented in Fig. 32A (capture by CD16 high affinity isoform fusion) and 32B (capture by CD16 low affinity isoform fusion). Apparent from a consideration of Figs. 32A and 32B is that both CD20 SMIP and CD20xCD20 monospecific scorpion showed increased binding to both the high- and low-affinity CD16 isoform fusions, with the CD20xCD20 scorpion showing a dramatic increase in binding to the low affinity isoform fusion with increasing protein concentration.

The binding of scorpions to the FcγRIII isoforms in the presence of target cells was also assessed. The data show the increased binding of scorpions to both FcγRIII (CD16) low- and high-affinity isoforms or allelotypes in the presence of target cells with increasing protein concentration.

In conducting the experiment, CD20-positive target cells were exposed to CD20 SMIPs or CD20xCD20 monospecific scorpions under conditions that allowed the binding of the SMIP or scorpion to the CD20-positive target cell. Subsequently, the SMIP- or scorpion-bearing target cell was exposed to either CD16 high- or low-affinity isoform tagged with mouse IgFc. A labeled goat anti-mouse Fc was then added as a secondary antibody to label the immobilized CD16 tagged with the mouse IgFc. Cells were then detected using flow cytometry on a FACs Calibur (BD Biosciences, San Jose, CA) and analyzed with Cell Quest software (BD Biosciences, San Jose, CA). As shown in Fig. 33, increased concentrations of each of the CD20 SMIP and the CD20xCD20 monospecific scorpion led to increased binding to the CD16 isoforms in the presence of target cells, with the increase in binding of the CD20xCD20 scorpion being more significant than the increased binding seen with the CD20 SMIP.

Example 17

Cell-cycle effects of scorpions on target lymphoma cells

The cell-cycle effects of scorpions were assessed by exposing lymphoma cells to SMIPs, monospecific scorpions and bispecific scorpions. More particularly, DoHH2 lymphoma cells (0.5×10^6) were treated for 24 hours with 0.4 nM rituximab, CD20xCD37 scorpion, TRU-015 (CD20 SMIP) + SMIP-016 combination (0.2 nM each), 100 nM SMIP-016 or 100 nM CD37xCD37 scorpion. These concentrations respresent about 10-fold more than the IC50 value of the scorpion in a 96-hour growth inhibition assay (see Figs. 24-27). Cultures were labeled for 20 minutes at 37°C with 10 μM BrdU (bromodeoxyuridine). Following fixation, cells were stained with anti-BrdU-FITC antibody and counterstained with propidium iodide. Values in Fig. 28 are the mean +/- SD of 4 replicate cultures from 2-3 independent experiments. All sample data were analyzed at the same time and pooled for presentation using both the BrdU and PI incorporation dot plots. Plots demonstrate that a major effect of scorpion treatment is a depletion of cells in S-phase, as well as an increase in the G_0/G_1 compartment.

Example 18

Physiological effects of scorpions

a. Mitochondrial potential

CD20xCD20 scorpions induced loss of mitochondrial membrane potential in DHL4 B-cells, as revealed in a JC-1 assay. JC-1 is a cationic carbocyanine dye that exhibits potential-dependent accumulation in the mitochondria (Mitoprobe® JC-1 Assay Kit for Flow Cytometry from Molecular Probes). JC-1 is more specific to the mitochondrial membrane than the plasma membrane and is used to determine changes in mitochondrial membrane potential. Accumulation in mitochondria is indicated by a fluorescence shift from green (529nm) to red (590nm).

In conducting the experiment, DHL-4 B-cells ($5x10^5$ cells/ml) were initially cultured in 24-well plates and treated for 24 hours with 1 μg/ml CD20xCD20 scorpion, Rituximab, IgG control antibody, or 5 μM staurosporine at 37°C, 5%CO_2, in a standard tissue-culture incubator. JC-1 dye (10 μl/ml, 2 μM final concentration) was added and cells were incubated for another 30 minutes at 37°C. Cells were harvested by centrifugation (5minutes at 1200 rpm), washed with 1ml PBS, and resuspended in 500 μl PBS. Cells were analyzed by flow cytometry (FACSCalibur, BD) with 488 nM excitation and 530 nM and 585 nM emission filters. For the representative scatter plots shown in Fig. 56, red fluorescence was measured on the Y-axis and green fluorescence was measured on the X-axis. Depolarization of the mitochondrial membrane was measured as a decrease in red fluorescence, as seen in the positive control CCCP (carbonyl cyanide 3-chlorophenylhydrazone), a known mitochondrial membrane potential disrupter. To confirm that JC-1 was responsive to changes in membrane potential, DHL-4 B-cells were treated with two concentrations of CCCP (50μM and 250μM) for 5 minutes at 37°C, 5%CO_2. An additional positive control was cells treated with the broad-spectrum kinase inhibitor staurosporine to induce apoptosis. The results shown in Fig. 56 are dot-plot graphs of 10,000 counts, with red fluorescence plotted on the Y-axis and green fluorescence plotted on the X-axis. A summary histogram of the percentage of cells with disrupted mitochondrial membrane potential (disrupted MMP: black bars) is shown in Fig 56. These results demonstrate that treatment with either the 20-4x20-4 scorpion or the 011x20-4

scorpion generated a decrease in the mitochondrial membrane potential associated with cell death.

b. Calcium flux

Scorpion molecules were analyzed for influences on cell signaling pathways, using Ca^{++} mobilization, a common feature of cell signaling, as a measure therefor. SU-DHL-6 lymphoma cells were labeled with Calcium 4 dye and treated with the test molecules identified below. Cells were read for 20 seconds to determine background fluorescence, and then SMIPs/scorpions were added (first dashed line in Fig. 28) and fluorescence was measured out to 600 seconds. At 600 seconds, an 8-fold excess of cross-linked goat-anti-human F(ab)'2 was added and fluorescence was measured for a further 300 seconds. Panel (A) of Fig. 28 shows the results obtained with a combination of CD20 SMIP and CD37 SMIP (red line); or obtained with a CD20xCD37 bispecific scorpion (black line), compared with unstimulated cells (blue line). In panel B of Fig. 28, the results of treating cells withCD20 SMIP alone (red line) resulted in Ca^{++} mobilization, but this was not as robust as the signal generated by the monospecific CD20xCD20 scorpion (black line). The Ca^{++} mobilization plots of Fig. 28 represent the fluorescence from triplicate wells treated with equimolar amounts of scorpion and SMIP/SMIP combinations.

c. Caspases 3, 7 and 9

The ability of CD20-binding scorpions to directly kill B-cells as evidenced by increased Annexin V and Propidium Iodide staining and the loss of mitochondrial membrane potential led to an further investigation of additional apoptosis-related effects of CD20-binding scorpions in B-cells. The approach taken was to perform Apo1 assays on DHL-4 B-cells exposed to CD20xCD20 scorpions or appropriate controls. The Apo1 assay is based on a synthetic peptide substrate for caspase 3 and 7. The assay components are available from Promega (Apo-ONE® Homogeneuous Caspase-3/7 Assay). Caspase-mediated cleavage of the labeled peptide Z-DEVD-Rhodamine 110 releases the fluorescent rhodamine 110 label, which is measured using 485 nm excitation and 530nm detection.

In the experiment, 100 µl DHL-4 B-cells ($1x10^6$ cells/ml) were plated in black 96-well flat-bottom tissue culture plates and treated for 24 or 48 hours with 1 µg/ml

CD20xCD20 scorpion, Rituximab, an IgG control antibody, or 5 μM staurosporine at 37°C, 5%CO_2 in a standard tissue-culture incubator. (Staurosporine is a small-molecule, broad-spectrum protein kinase inhibitor that is known in the art as a potent inducer of classical apoptosis in a wide variety of cell types.) After 24 or 48 hours, 100 μl of the 100-fold diluted substrate was added to each well, gently mixed for one minute on a plate shaker (300 rpm) and incubated at room temperature for two hours. Fluorescence was measured using 485 nM excitation and 527 nM emission filter (Fluoroskan Ascent FL, Thermo Labsystems). Graphs showing average fluorescent intensity of triplicate treatments plus/minus standard deviation after 24 hours and 48 hours (24 hours only for staurosporine) are presented in Fig. 57. These results establish that CD20-binding scorpions do not directly kill B-cells by an apoptotic pathway involving activation of caspase 3/7.

The results obtained in the Apo-1 assay were confirmed by Western blot analyses designed to detect pro-caspase cleavage resulting in activated caspase or to detect cleavage of PARP (Poly (ADP-Ribose) Polymerase), a protein known to be cleaved by activated caspase 3. DHL-4 B-cells were exposed to a CD20 binding scorpion or a control for 4, 24, or 48 hours and cell lysates were fractionated on SDS-PAGE and blotted for Western analyses using conventional techniques. Fig. 58 presents the results in the form of a collection of Western blots. The bottom three Westerns utilized anti-caspase antibodies to detect shifts in molecular weight of the caspase enzyme, reflecting proteolytic activation. For caspases 3, 7, and 9, there was no evidence of caspase activation by any of the CD20-binding molecules. Staurosporine served as a positive control for the assay, and induced pro-caspase cleavage to active caspase for each of caspases 3, 7 and 9. The fourth Western blot shown in Fig. 58 reveals that PARP, a known substrate of activated caspase 3, was not cleaved, consistent with a failure of CD20-binding scorpions to activate caspase 3. The results of all of these experiments are consistent in showing that caspase 3 activation is not a significant feature of the direct cell killing of DHL-4 B-cells induced by CD20 binding scorpions.

In addition, a time series study was conducted to determine the effect of CD20 binding proteins, including a CD20xCD20 scorpion, on Caspase 3. DoHH2 or Su-DHL-6 B-cells were incubated with 10nM CD20 binding protein (S0129 scorpion,

2Lm20-4 SMIP, or Rituxan®) +/- soluble CD16 Ig (40nM), soluble CD16 Ig alone, or media. The cells were cultured in complete RPMI with 10% FBS at $3x10^5$/well/300 µl and harvested at 4 hours, 24 hours or 72 hours. The 72-hour time-point samples were plated in 500 µl of the test agent. Cells were washed with PBS and then stained for intracellular active caspase-3 using the BD Pharmingen Caspase 3, Active Form, mAB Apoptosis Kit:FITC (cat no.55048, BD Pharmingen, San Jose, CA). Briefly, after 2 additional washes in cold PBS, the cells were suspended in cold cytofix/cytoperm solution and incubated on ice for 20 minutes. Cells were then washed by centrifugation, aspirated, and washed two times with Perm/Wash buffer at room temperature. The samples were then stained with 20 µl FITC-anti-caspase 3 in 100 µl of Perm-Wash buffer at room temperature in the dark for thirty minutes. The samples were then washed two times with Perm-Wash buffer, and resuspended in 500 µl of Perm-Wash buffer. Washed cells were then transferred to FACs tubes and run on a FACs Calibur (BD Biosciences, San Jose, CA) and analyzed with Cell Quest software (BD Biosciences, San Jose, CA). The results are shown in Table 16.

Table 16

Molecule (10 nM)	Percentage Caspase-3 positive cells			Percentage in live gate		
	4 hours	24 hours	48 hours	4 hours	24 hours	48 hours
RTXN	7	25	7	75	53	56
and CD16 hi (4X)	27	47	21	79	60	43
CD20 SMIP (2Lm20-4)	5	5	10	89	85	81
and CD16 hi	28	54	21	61	60	41
Humanized CD20xCD20 scorpion	7	13	14	69	68	61

(S00129)						
and CD16 hi	30	31	15	67	75	72
Media	7	5	9	89	82	80
and CD16 hi	6	5	9	91	83	80

The results of all of these experiments are consistent in showing that there is limited activation of caspase 3 in the absence of CD16, which does not implicate caspase 3 activation as a significant feature of the direct cell killing induced by CD20 binding scorpions.

d. DNA fragmentation

Induction of classical apoptotic signaling pathways ultimately results in condensation and fragmented degradation of chromosomal DNA. To determine whether CD20-binding scorpions directly killed B-cells through a classical apoptotic mechanism, the state of B-cell chromosomal DNA was examined following exposure of the cells to CD20-binding scorpions, or controls. Initially, DHL-4 B-cells were treated *in vitro* for 4, 24 or 48 hours with a CD20-binding molecule, i.e., the monospecific CD20xCD20 (2Lm20-4x2Lm20-4) scorpion, the CD20xCD20 (011x2Lm20-4) scorpion, or Rituximab, or with a control. Subsequently, cells were lysed and chromosomal DNA was purified using conventional techniques. The chromosomal DNA was then size-fractionated by gel electrophoresis. The gel electrophoretogram shown in Fig. 59 reveals a lack of DNA fragmentation that demonstrated that the cell death generated by CD20-binding scorpions was not mediated by a classical apoptotic pathway. Staurosporine-treated cells were used as positive control in these assays.

e. SYK phosphorylation

SYK is a phospho-regulated protein with several phosphorylation sites that functions as a transcriptional repressor. SYK is localized to the cell nucleus, but is capable of rapid relocation to the membrane upon activation. For activation, SYK must retain its nuclear localization sequence. Activated SYK has a role in suppressing breast cancer tumors and SYK is activated by pro-apoptotic signals such

as ionizing radiation, BCR ligation and MHC class II cross-linking. Further, SYK has been shown to affect the PLC-γ and Ca^{++} pathways. Given these observations, the capacity of CD20-binding scorpions to affect SYK was investigated.

DHL-6 B-cells were exposed to a bispecific CD20xCD37 scorpion for 0, 5, 7 or 15 hours and the cells were lysed. Lysates were immunoprecipitated with either an anti-phosphotyrosine antibody or with an anti-SYK antibody. Immunoprecipitates were fractionated by gel electrophoresis and the results are shown in Fig. 60. Apparent from an inspection of Fig. 60 is the failure of the bispecific CD20xCD37 scorpion to induce phosphorylation of SYK, thereby activating it. Consistent with the above-described studies on caspase activation and chromosomal DNA fragmentation, it does not appear that CD20-binding scorpions directly kill B-cells using a classic apoptotic pathway, such as the caspase-dependent pathway. While not wishing to be bound by theory, it is expected that the CD20-binding scorpions directly kill B-cells through a caspase-, and SYK-, independent pathway that does not prominently feature chromosomal DNA fragmentation, at least not on the same time frame as fragmentation occurs during caspase-dependent apoptosis.

Example 19

Scorpion applications

a. *In vivo* activity of scorpions

The activity of scorpions was also assessed using a mouse model. Measurements of scorpion activity *in vivo* involved administration of 10-300 μg scorpion and subsequent time-series determinations of serum concentrations of that scorpion. Results of these studies, presented as serum concentration curves for each of two bispecific scorpions (i.e., S0033, a CD20xCD27 scorpion and a CD20xCD37 scorpion) from three-week pharmacokinetic studies in mice are presented in Fig. 40. The data in Fig. 40 show that it took at least 500 hours after administration before the serum levels of each of the two bispecific scorpions fell back to baseline levels. Thus, scorpions show serum stability and reproducible, sustained circulating half-lives *in vivo*.

The *in vivo* efficacy of scorpions was also assessed. An aggressive Ramos xenograft model was used in parallel experiments with SMIPs versus historical

immunoglobulin controls. The survival curves provided in Fig. 41 reveal that administration of 10 μg bispecific scorpion had negligible influence on survival, but administration of 100-300 μg had significant positive effect on the survival of mice bearing Ramos xenografts.

b. Combination therapies

it is contemplated that scorpions will find application in the prevention, treatment or amelioration of a symptom of, a wide variety of conditions affecting man, other mammals and other organisms. For example, CD20-binding scorpions are expected to be useful in treating or preventing a variety of diseases associated with excessive or aberrant B-cells. In fact, any disease amenable to a treatment involving the depletion of B-cells would be amenable to treatment with a CD20-binding scorpion. In addition, scorpions, e.g., CD20-binding scorpions, may be used in combination therapies with other therapeutics. To illustrate the feasibility of a wide variety of combination therapies, the monospecific CD20xCD20 scorpion (S0129) was administered to Su-DHL-6 B-cells in combination with doxorubicin, vincristine or rapamycin. Doxorubicin is a topoisomerase II poison that interferes with DNA biochemistry and belongs to a class of drugs contemplated for anti-cancer treatment. Rapamycin (Sirolimus) is a macrolide antibiotic that inhibits the initiation of protein synthesis and suppresses the immune system, finding application in organ transplantation and as an anti-proliferative used with coronary stents to inhibit or prevent restenosis. Vincristine is a vinca alkaloid that inhibits tubule formation and has been used to treat cancer.

The experimental results shown in Fig. 61 are presented as Combination Index values for each combination over a range of effect levels. The interactions of the monospecific CD20xCD20 scorpion S0129 are different for each drug class, while with Rituxan® (RTXN) the plots forms are similar. The effect seen with doxorubicin at high concentrations may reflect a shift towards monovalent binding. The data establish that CD20-binding scorpions may be used in combination with a variety of other therapeutics and such combinations would be apparent to one of skill in the art in view of the present disclosure.

Variations on the structural themes for multivalent binding molecules with effector function, or scorpions, will be apparent to those of skill in the art upon review

of the present disclosure, and such variant structures are within the scope of the invention.

CLAIMS

What is claimed is:

1. A multivalent single-chain binding protein with effector function comprising

a. a first binding domain derived from an immunoglobulin or immunoglobulin-like molecule;

b. a constant sub-region providing an effector function, the immunoglobulin constant sub-region located C-terminal to the first binding domain;

c. a scorpion linker located C-terminal to the constant sub-region; and

d. a second binding domain derived from an immunoglobulin or immunoglobulin-like molecule located C-terminal to the constant sub-region;

thereby localizing the constant sub-region between the first binding domain and the second binding domain.

2. The protein according to claim 1 wherein the first binding domain comprises a variable light region derived from a first immunoglobulin and a variable heavy region derived from a second immunoglobulin.

3. The protein according to claim 2 wherein the first immunoglobulin and the second immunoglobulin are the same immunoglobulin.

4. The protein according to claim 1 wherein the second binding domain comprises a variable light region derived from a first immunoglobulin and a variable heavy region derived from a second immunoglobulin.

5. The protein according to claim 4 wherein the first immunoglobulin and the second immunoglobulin are the same immunoglobulin.

6. The protein according to claim 1 wherein the first binding domain and the second binding domain recognize the same molecular target.

7. The protein according to claim 1 wherein the first binding domain and the second binding domain recognize the same epitope.

8. The protein according to claim 1 wherein the first binding domain and the second binding domain recognize different molecular targets on the same eukaryotic cell, prokaryotic cell, virus, carrier, or object.

9. The protein according to claim 1 wherein the first binding domain and the second binding domain recognize molecular targets associated with physically distinct eukaryotic cells, prokaryotic cells, viruses, carriers or objects.

10. The protein according to claim 1 wherein at least one of the first binding domain and the second binding domain recognize at least one molecular target not associated with a eukaryotic cell, prokaryotic cell, virus, carrier or object.

11. The protein according to claim 1 wherein the first binding domain, the second binding domain and the constant sub-region are each derived from a human immunoglobulin.

12. The protein according to claim 1 wherein at least one of the first binding domain and the second binding domain recognizes a target on a cancer cell.

13. The protein according to claim 11 wherein at least one of the first binding domain and the second binding domain recognizes a target selected from the group consisting of a tumor antigen, a B-cell target, a TNF receptor superfamily member, a Hedgehog family member, a receptor tyrosine kinase, a proteoglycan-related molecule, a TGF-beta superfamily member, a Wnt-related molecule, a receptor ligand, a T-cell target, a Dendritic cell target, an NK cell target, a monocyte/macrophage cell target and an angiogenesis target.

14. The protein according to claim 13 wherein the tumor antigen is selected from the group consisting of SQUAMOUS CELL CARCINOMA ANTIGEN 1, SQUAMOUS CELL CARCINOMA ANTIGEN 2, Ovarian carcinoma antigen CA125, MUCIN 1, CTCL tumor antigen se1-1, CTCL tumor antigen se14-3, CTCL tumor antigen se20-4, CTCL tumor antigen se20-9, CTCL tumor antigen se33-1, CTCL tumor antigen se37-2, CTCL tumor antigen se57-1, CTCL tumor antigen se89-1, Prostate-specific membrane antigen, 5T4 oncofetal trophoblast glycoprotein, Orf73 Kaposi's sarcoma-associated herpesvirus, MAGE-C1, MAGE-B1 ANTIGEN, MAGE-B2 ANTIGEN, MAGE-2 ANTIGEN, MAGE-4a antigen, MAGE-4b antigen,

Colon cancer antigen NY-CO-45, Lung cancer antigen NY-LU-12 variant A, Cancer associated surface antigen, Adenocarcinoma antigen ART1, Paraneoplastic associated brain-testis-cancer antigen, Neuro-oncological ventral antigen 2, Hepatocellular carcinoma antigen gene 520, TUMOR-ASSOCIATED ANTIGEN CO-029, Tumor-associated antigen MAGE-X2, Synovial sarcoma X breakpoint 2, Squamous cell carcinoma antigen recognized by T cell, Serologically defined colon cancer antigen 1, Serologically defined breast cancer antigen NY-BR-15, Serologically defined breast cancer antigen NY-BR-16, Chromogranin A; parathyroid secretory protein 1, DUPAN-2, CA 19-9, CA 72-4, CA 195 and L6.

15. The protein according to claim 13 wherein the B cell target is selected from the group consisting of CD10, CD19, CD20, CD21, CD22, CD23, CD24, CD37, CD38, CD39, CD40, CD72, CD73, CD74, CDw75, CDw76, CD77, CD78, CD79a/b, CD80, CD81, CD82, CD83, CD84, CD85, CD86, CD89, CD98, CD126, CD127, CDw130, CD138 and CDw150.

16. The protein according to claim 13 wherein the TNF receptor superfamily member is selected from the group consisting of 4-1BB/TNFRSF9, NGF R/TNFRSF16, BAFF R/TNFRSF13C, Osteoprotegerin/TNFRSF11B, BCMA/TNFRSF17, OX40/TNFRSF4, CD27/TNFRSF7, RANK/TNFRSF11A, CD30/TNFRSF8, RELT/TNFRSF19L, CD40/TNFRSF5, TACI/TNFRSF13B, DcR3/TNFRSF6B, TNF RI/TNFRSF1A, DcTRAIL R1/TNFRSF23, TNF RII/TNFRSF1B, DcTRAIL R2/TNFRSF22, TRAIL R1/TNFRSF10A, DR3/TNFRSF25, TRAIL R2/TNFRSF10B, DR6/TNFRSF21, TRAIL R3/TNFRSF10C, EDAR, TRAIL R4/TNFRSF10D, Fas/TNFRSF6, TROY/TNFRSF19, GITR/TNFRSF18, TWEAK R/TNFRSF12, HVEM/TNFRSF14, XEDAR, Lymphotoxin beta R/TNFRSF3, 4-1BB Ligand/TNFSF9, Lymphotoxin, APRIL/TNFSF13, Lymphotoxin beta/TNFSF3, BAFF/TNFSF13C, OX40 Ligand/TNFSF4, CD27 Ligand/TNFSF7, TL1A/TNFSF15, CD30 Ligand/TNFSF8, TNF-alpha/TNFSF1A, CD40 Ligand/TNFSF5, TNF-beta/TNFSF1B, EDA-A2, TRAIL/TNFSF10, Fas Ligand/TNFSF6, TRANCE/TNFSF11, GITR Ligand/TNFSF18, TWEAK/TNFSF12 and LIGHT/TNFSF14.

17. The protein according to claim 13 wherein the is Hedgehog family member is selected from the group consisting of Patched and Smoothened.

18. The protein according to claim 13 wherein the receptor tyrosine kinase is selected from the group consisting of Axl, FGF R4, C1q R1/CD93, FGF R5, DDR1, Flt-3, DDR2, HGF R, Dtk, IGF-I R, EGF R, IGF-II R, Eph, INSRR, EphA1, Insulin R/CD220, EphA2, M-CSF R, EphA3, Mer, EphA4, MSP R/Ron, EphA5, MuSK, EphA6, PDGF R alpha, EphA7, PDGF R beta, EphA8, Ret, EphB1, ROR1, EphB2, ROR2, EphB3, SCF R/c-kit, EphB4, Tie-1, EphB6, Tie-2, ErbB2, TrkA, ErbB3, TrkB, ErbB4, TrkC, FGF R1, VEGF R1/Flt-1, FGF R2, VEGF R2/Flk-1, FGF R3 and VEGF R3/Flt-4.

19. The protein according to claim 13 wherein the proteoglycan-related molecule is selected from the group consisting of proteoglycans and regulators thereof.

20. The protein according to claim 13 wherein the Transforming Growth Factor (TGF)-beta superfamily member is selected from the group consisting of Activin RIA/ALK-2, GFR alpha-1, Activin RIB/ALK-4, GFR alpha-2, Activin RIIA, GFR alpha-3, Activin RIIB, GFR alpha-4, ALK-1, MIS RII, ALK-7, Ret, BMPR-IA/ALK-3, TGF-beta RI/ALK-5, BMPR-IB/ALK-6, TGF-beta RII, BMPR-II, TGF-beta RIIb, Endoglin/CD105 and TGF-beta RIII.

21. The protein according to claim 13 wherein the Wnt-related molecule is selected from the group consisting of Frizzled-1, Frizzled-8, Frizzled-2, Frizzled-9, Frizzled-3, sFRP-1, Frizzled-4, sFRP-2, Frizzled-5, sFRP-3, Frizzled-6, sFRP-4, Frizzled-7, MFRP, LRP 5, LRP 6, Wnt-1, Wnt-8a, Wnt-3a, Wnt-10b, Wnt-4, Wnt-11, Wnt-5a, Wnt-9a and Wnt-7a.

22. The protein according to claim 13 wherein the receptor ligand is selected from the group consisting of 4-1BB Ligand/TNFSF9, Lymphotoxin, APRIL/TNFSF13, Lymphotoxin beta/TNFSF3, BAFF/TNFSF13C, OX40 Ligand/TNFSF4, CD27 Ligand/TNFSF7, TL1A/TNFSF15, CD30 Ligand/TNFSF8, TNF-alpha/TNFSF1A, CD40 Ligand/TNFSF5, TNF-beta/TNFSF1B, EDA-A2, TRAIL/TNFSF10, Fas Ligand/TNFSF6, TRANCE/TNFSF11, GITR Ligand/TNFSF18, TWEAK/TNFSF12, LIGHT/TNFSF14, Amphiregulin, NRG1 isoform GGF2, Betacellulin, NRG1 Isoform SMDF, EGF, NRG1-alpha/HRG1-alpha, Epigen, NRG1-beta 1/HRG1-beta 1, Epiregulin, TGF-alpha, HB-EGF,

TMEFF1/Tomoregulin-1, Neuregulin-3, TMEFF2, IGF-I, IGF-II, Insulin, Activin A, Activin B, Activin AB, Activin C, BMP-2, BMP-7, BMP-3, BMP-8, BMP-3b/GDF-10, BMP-9, BMP-4, BMP-15, BMP-5, Decapentaplegic, BMP-6, GDF-1, GDF-8, GDF-3, GDF-9, GDF-5, GDF-11, GDF-6, GDF-15, GDF-7, Artemin, Neurturin, GDNF, Persephin, TGF-beta, TGF-beta 2, TGF-beta 1, TGF-beta 3, LAP (TGF-beta 1), TGF-beta 5, Latent TGF-beta 1, Latent TGF-beta bp1, TGF-beta 1.2, Lefty, Nodal, MIS/AMH, FGF acidic, FGF-12, FGF basic, FGF-13, FGF-3, FGF-16, FGF-4, FGF-17, FGF-5, FGF-19, FGF-6, FGF-20, FGF-8, FGF-21, FGF-9, FGF-23, FGF-10, KGF/FGF-7, FGF-11, Neuropilin-1, PlGF, Neuropilin-2, PlGF-2, PDGF, PDGF-A, VEGF, PDGF-B, VEGF-B, PDGF-C, VEGF-C, PDGF-D, VEGF-D and PDGF-AB.

23. The protein according to claim 13 wherein the T-cell receptor is selected from the group consisting of 2B4/SLAMF4, IL-2 R alpha, 4-1BB/TNFRSF9, IL-2 R beta, ALCAM, B7-1/CD80, IL-4 R, B7-H3, BLAME/SLAMF8, BTLA, IL-6 R, CCR3, IL-7 R alpha, CCR4, CXCR1/IL-8 RA, CCR5, CCR6, IL-10 R alpha, CCR7, IL-10 R beta, CCR8, IL-12 R beta 1, CCR9, IL-12 R beta 2, CD2, IL-13 R alpha 1, IL-13, CD3, CD4, ILT2/CD85j, ILT3/CD85k, ILT4/CD85d, ILT5/CD85a, Integrin alpha 4/CD49d, CD5, Integrin alpha E/CD103, CD6, Integrin alpha M/CD11b, CD8, Integrin alpha X/CD11c, Integrin beta 2/CD18, KIR/CD158, CD27/TNFRSF7, KIR2DL1, CD28, KIR2DL3, CD30/TNFRSF8, KIR2DL4/CD158d, CD31/PECAM-1, KIR2DS4, CD40 Ligand/TNFSF5, LAG-3, CD43, LAIR1, CD45, LAIR2, CD83, Leukotriene B4 R1, CD84/SLAMF5, NCAM-L1, CD94, NKG2A, CD97, NKG2C, CD229/SLAMF3, NKG2D, CD2F-10/SLAMF9, NT-4, CD69, NTB-A/SLAMF6, Common gamma Chain/IL-2 R gamma, Osteopontin, CRACC/SLAMF7, PD-1, CRTAM, PSGL-1, CTLA-4, RANK/TNFRSF11A, CX3CR1, CX3CL1, L-Selectin, CXCR3, SIRP beta 1, CXCR4, SLAM, CXCR6, TCCR/WSX-1, DNAM-1, Thymopoietin, EMMPRIN/CD147, TIM-1, EphB6, TIM-2, Fas/TNFRSF6, TIM-3, Fas Ligand/TNFSF6, TIM-4, Fc gamma RIII/CD16, TIM-6, GITR/TNFRSF18, TNF RI/TNFRSF1A, Granulysin, TNF RII/TNFRSF1B, HVEM/TNFRSF14, TRAIL R1/TNFRSF10A, ICAM-1/CD54, TRAIL R2/TNFRSF10B, ICAM-2/CD102, TRAIL R3/TNFRSF10C, IFN-gamma R1, TRAIL R4/TNFRSF10D, IFN-gamma R2, TSLP, IL-1 RI and TSLP R.

24. The protein according to claim 13 wherein the NK cell receptor is selected from the group consisting of 2B4/SLAMF4, KIR2DS4, CD155/PVR, KIR3DL1, CD94, LMIR1/CD300A, CD69, LMIR2/CD300c, CRACC/SLAMF7, LMIR3/CD300LF, DNAM-1, LMIR5/CD300LB, Fc epsilon RII, LMIR6/CD300LE, Fc gamma RI/CD64, MICA, Fc gamma RIIB/CD32b, MICB, Fc gamma RIIC/CD32c, MULT-1, Fc gamma RIIA/CD32a, Nectin-2/CD112, Fc gamma RIII/CD16, NKG2A, FcRH1/IRTA5, NKG2C, FcRH2/IRTA4, NKG2D, FcRH4/IRTA1, NKp30, FcRH5/IRTA2, NKp44, Fc Receptor-like 3/CD16-2, NKp46/NCR1, NKp80/KLRF1, NTB-A/SLAMF6, Rae-1, Rae-1 alpha, Rae-1 beta, Rae-1 delta, H60, Rae-1 epsilon, ILT2/CD85j, Rae-1 gamma, ILT3/CD85k, TREM-1, ILT4/CD85d, TREM-2, ILT5/CD85a, TREM-3, KIR/CD158, TREML1/TLT-1, KIR2DL1, ULBP-1, KIR2DL3, ULBP-2, KIR2DL4/CD158d and ULBP-3.

25. The protein according to claim 13 wherein the monocyte/macrophage cell receptor is selected from the group consisting of B7-1/CD80, ILT4/CD85d, B7-H1, ILT5/CD85a, Common beta Chain, Integrin alpha 4/CD49d, BLAME/SLAMF8, Integrin alpha X/CD11c, CCL6/C10, Integrin beta 2/CD18, CD155/PVR, Integrin beta 3/CD61, CD31/PECAM-1, Latexin, CD36/SR-B3, Leukotriene B4 R1, CD40/TNFRSF5, LIMPII/SR-B2, CD43, LMIR1/CD300A, CD45, LMIR2/CD300c, CD68, LMIR3/CD300LF, CD84/SLAMF5, LMIR5/CD300LB, CD97, LMIR6/CD300LE, CD163, LRP-1, CD2F-10/SLAMF9, MARCO, CRACC/SLAMF7, MD-1, ECF-L, MD-2, EMMPRIN/CD147, MGL2, Endoglin/CD105, Osteoactivin/GPNMB, Fc gamma RI/CD64, Osteopontin, Fc gamma RIIB/CD32b, PD-L2, Fc gamma RIIC/CD32c, Siglec-3/CD33, Fc gamma RIIA/CD32a, SIGNR1/CD209, Fc gamma RIII/CD16, SLAM, GM-CSF R alpha, TCCR/WSX-1, ICAM-2/CD102, TLR3, IFN-gamma R1, TLR4, IFN-gamma R2, TREM-1, IL-1 RII, TREM-2, ILT2/CD85j, TREM-3, ILT3/CD85k, TREML1/TLT-1, 2B4/SLAMF4, IL-10 R alpha, ALCAM, IL-10 R beta, Aminopeptidase N/ANPEP, ILT2/CD85j, Common beta Chain, ILT3/CD85k, C1q R1/CD93, ILT4/CD85d, CCR1, ILT5/CD85a, CCR2, Integrin alpha 4/CD49d, CCR5, Integrin alpha M/CD11b, CCR8, Integrin alpha X/CD11c, CD155/PVR, Integrin beta 2/CD18, CD14, Integrin beta 3/CD61, CD36/SR-B3, LAIR1, CD43, LAIR2, CD45, Leukotriene B4 R1, CD68, LIMPII/SR-B2, CD84/SLAMF5, LMIR1/CD300A, CD97, LMIR2/CD300c, CD163, LMIR3/CD300LF, Coagulation Factor III/Tissue

Factor, LMIR5/CD300LB, CX3CR1, CX3CL1, LMIR6/CD300LE, CXCR4, LRP-1, CXCR6, M-CSF R, DEP-1/CD148, MD-1, DNAM-1, MD-2, EMMPRIN/CD147, MMR, Endoglin/CD105, NCAM-L1, Fc gamma RI/CD64, PSGL-1, Fc gamma RIII/CD16, RP105, G-CSF R, L-Selectin, GM-CSF R alpha, Siglec-3/CD33, HVEM/TNFRSF14, SLAM, ICAM-1/CD54, TCCR/WSX-1, ICAM-2/CD102, TREM-1, IL-6 R, TREM-2, CXCR1/IL-8 RA, TREM-3 and TREML1/TLT-1.

26. The protein according to claim 13 wherein the Dendritic cell receptor is selected from the group consisting of CD36/SR-B3, LOX-1/SR-E1, CD68, MARCO, CD163, SR-AI/MSR, CD5L, SREC-I, CL-P1/COLEC12, SREC-II, LIMPII/SR-B2, RP105, TLR4, TLR1, TLR5, TLR2, TLR6, TLR3, TLR9, 4-1BB Ligand/TNFSF9, IL-12/IL-23 p40, 4-Amino-1,8-naphthalimide, ILT2/CD85j, CCL21/6Ckine, ILT3/CD85k, 8-oxo-dG, ILT4/CD85d, 8D6A, ILT5/CD85a, A2B5, Integrin alpha 4/CD49d, Aag, Integrin beta 2/CD18, AMICA, Langerin, B7-2/CD86, Leukotriene B4 R1, B7-H3, LMIR1/CD300A, BLAME/SLAMF8, LMIR2/CD300c, C1q R1/CD93, LMIR3/CD300LF, CCR6, LMIR5/CD300LB, CCR7, LMIR6/CD300LE, CD40/TNFRSF5, MAG/Siglec-4a, CD43, MCAM, CD45, MD-1, CD68, MD-2, CD83, MDL-1/CLEC5A, CD84/SLAMF5, MMR, CD97, NCAM-L1, CD2F-10/SLAMF9, Osteoactivin/GPNMB, Chem 23, PD-L2, CLEC-1, RP105, CLEC-2, Siglec-2/CD22, CRACC/SLAMF7, Siglec-3/CD33, DC-SIGN, Siglec-5, DC-SIGNR/CD299, Siglec-6, DCAR, Siglec-7, DCIR/CLEC4A, Siglec-9, DEC-205, Siglec-10, Dectin-1/CLEC7A, Siglec-F, Dectin-2/CLEC6A, SIGNR1/CD209, DEP-1/CD148, SIGNR4, DLEC, SLAM, EMMPRIN/CD147, TCCR/WSX-1, Fc gamma RI/CD64, TLR3, Fc gamma RIIB/CD32b, TREM-1, Fc gamma RIIC/CD32c, TREM-2, Fc gamma RIIA/CD32a, TREM-3, Fc gamma RIII/CD16, TREML1/TLT-1, ICAM-2/CD102 and Vanilloid R1.

27. The protein according to claim 13 wherein the angiogenesis target is selected from the group consisting of Angiopoietin-1, Angiopoietin-like 2, Angiopoietin-2, Angiopoietin-like 3, Angiopoietin-3, Angiopoietin-like 7/CDT6, Angiopoietin-4, Tie-1, Angiopoietin-like 1, Tie-2, Angiogenin, iNOS, Coagulation Factor III/Tissue Factor, nNOS, CTGF/CCN2, NOV/CCN3, DANCE, OSM, EDG-1, Plfr, EG-VEGF/PK1, Proliferin, Endostatin, ROBO4, Erythropoietin, Thrombospondin-1, Kininostatin, Thrombospondin-2, MFG-E8, Thrombospondin-4,

Nitric Oxide, VG5Q, eNOS, EphA1, EphA5, EphA2, EphA6, EphA3, EphA7, EphA4, EphA8, EphB1, EphB4, EphB2, EphB6, EphB3, Ephrin-A1, Ephrin-A4, Ephrin-A2, Ephrin-A5, Ephrin-A3, Ephrin-B1, Ephrin-B3, Ephrin-B2, FGF acidic, FGF-12, FGF basic, FGF-13, FGF-3, FGF-16, FGF-4, FGF-17, FGF-5, FGF-19, FGF-6, FGF-20, FGF-8, FGF-21, FGF-9, FGF-23, FGF-10, KGF/FGF-7, FGF-11, FGF R1, FGF R4, FGF R2, FGF R5, FGF R3, Neuropilin-1, Neuropilin-2, Semaphorin 3A, Semaphorin 6B, Semaphorin 3C, Semaphorin 6C, Semaphorin 3E, Semaphorin 6D, Semaphorin 6A, Semaphorin 7A, MMP, MMP-11, MMP-1, MMP-12, MMP-2, MMP-13, MMP-3, MMP-14, MMP-7, MMP-15, MMP-8, MMP-16/MT3-MMP, MMP-9, MMP-24/MT5-MMP, MMP-10, MMP-25/MT6-MMP, TIMP-1, TIMP-3, TIMP-2, TIMP-4, ACE, IL-13 R alpha 1, IL-13, C1q R1/CD93, Integrin alpha 4/CD49d, VE-Cadherin, Integrin beta 2/CD18, CD31/PECAM-1, KLF4, CD36/SR-B3, LYVE-1, CD151, MCAM, CL-P1/COLEC12, Nectin-2/CD112, Coagulation Factor III/Tissue Factor, E-Selectin, D6, P-Selectin, DC-SIGNR/CD299, SLAM, EMMPRIN/CD147, Tie-2, Endoglin/CD105, TNF RI/TNFRSF1A, EPCR, TNF RII/TNFRSF1B, Erythropoietin R, TRAIL R1/TNFRSF10A, ESAM, TRAIL R2/TNFRSF10B, FABP5, VCAM-1, ICAM-1/CD54, VEGF R2/Flk-1, ICAM-2/CD102, VEGF R3/Flt-4, IL-1 RI and VG5Q.

28. The protein according to claim 1 wherein at least one of the first binding domain and the second binding domain specifically binds a target selected from the group consisting of Prostate-specific Membrane Antigen (Folate Hydrolase 1), Epidermal Growth Factor Receptor, Receptor for Advanced Glycation End products, IL-17 A, IL-17 F, P19, Dickkopf-1, NOTCH1, NG2, IgHE, IL-22R, IL-21, an Amyloid β oligomer, an Amyloid β Precursor Protein, NOGO Receptor (RTN4R), Low Density LipoproteinReceptor-Related Protein 5, IL-4, Myostatin, Very Late Antigen 4 and IGF-1R.

29. The protein according to claim 12 wherein the cancer cell is a transformed hematopoietic cell.

30. The protein according to claim 29 wherein at least one of the first binding domain and the second binding domain recognizes a target selected from the group consisting of a B-cell target, a T-cell target, a dendritic cell target, or an NK-cell target.

31. The protein according to claim 29 wherein at least one of the first binding domain and the second binding domain recognizes a target selected from the group consisting of CD5, CD10, CD11c, CD15, CD19, CD20, CD21, CD22, CD23, CD25, CD27, CD30, CD38, CD45, CD70, CD80, CD86, CD103, IRTA1, IRTA2, IRTA3, IRTA4, IRTA5, B-B2, B-B8, and B-cell antigen receptor.

32, The protein according to claim 1 comprising a sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 103, 105, 107, 109, 332, 333, 334, and 345.

33. The proteinaccording to claim 1 comprising a sequence selected from the group consisting of SEQ ID NOS:355, 356, 357, 358, 359, 360, 361, 362, 363, 364 and 365.

34. The protein according to claim 1 wherein the constant sub-region recognizes an effector cell F_C receptor.

35. The protein according to claim 1 wherein the constant sub-region recognizes an effector cell surface protein selected from the group consisting of CD16, CD32a, CD32b, CD64, CD89, FcεR1, FcRn, and pIgR.

36. The protein according to claim 1 wherein the constant sub-region comprises a C_{H2} domain and a C_{H3} domain.

37. The protein according to claim 36, wherein the C_{H3} domain is truncated and comprises a C-terminal sequence selected from the group consisting of SEQ ID NOS: 366, 367, 368, 369, 370 and 371.

38. The protein according to claim 1, wherein the scorpion linker is derived from an immunoglobulin hinge.

39. The protein according to claim 38, wherein the scorpion linker has a reduced, but non-zero, number of cysteines relative to the hinge from which the scorpion linker is derived, and wherein one cysteine corresponds to the N-terminal hinge cysteine of the immunoglobulin hinge.

40. The protein according to claim 39 wherein the scorpion linker is derived from an IgG1 hinge region and wherein the linker comprises 1 or 2 cysteines, further wherein the linker retains the cysteine corresponding to an N-terminal hinge cysteine of the IgG1 hinge region.

41. The protein according to claim 1, wherein the scorpion linker is derived from the stalk region of a C-lectin.

42. The protein according to claim 1, wherein the scorpion linker comprises a sequence selected from the group consisting of SEQ ID NOS:373, 374, 375, 376 and 377.

43. The protein according to claim 1, wherein the scorpion linker comprises a sequence selected from the group consisting of SEQ ID NOS:351, 352, 353 and 354.

44. The protein according to claim 1 further comprising a linker of at least about 5 amino acids attached to the constant sub-region and attached to the first binding domain, thereby localizing the linker between the constant sub-region and the first binding domain.

45. The protein according to claim 1 in which the binding affinity of at least one of the first binding domain and the second binding domain is less than 10^{-9} M.

46. The protein according to claim 1 in which the binding affinity of at least one of the first binding domain and the second binding domain is at least 10^{-6} M.

47. The protein according to claim 1, wherein the effector function is selected from the group consisting of antibody-dependent cell-mediated cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC).

48. The protein according to claim 1, wherein the *in vivo* half-life of the protein is at least 28 hours in a human.

49. A pharmaceutical composition comprising the protein according to claim 1 and a pharmaceutically acceptable adjuvant, carrier or excipient.

50. A nucleic acid encoding a protein according to claim 1.

51. A vector comprising a nucleic acid according to claim 50.

52. A host cell comprising a vector according to claim 51.

53. A method of producing the protein according to claim 1 comprising:

a. introducing a nucleic acid encoding the protein according to claim 1 into a host cell; and

b. incubating the host cell under conditions suitable for expression of the protein, thereby expressing the protein at a level of at least 1 mg/ml.

54. The method according to claim 53 further comprising isolating the protein.

55. The method according to claim 53 wherein the host cell is selected from the group consisting of a VERO cell, a HeLa cell, a CHO cell, a COS cell, a W138 cell, a BHK cell, a HepG2 cell, a 3T3 cell, a RIN cell, an MDCK cell, an A549 cell, a PC12 cell, a K562 cell, a HEK293 cell, an N cell, a *Spodoptera frugiperda* cell, a *Saccharomyces cerevisiae* cell, a *Pichia pastoris* cell, a fungal cell and a bacterial cell.

56. A method of producing a nucleic acid encoding the protein according to claim 1 comprising:

a. covalently linking the 3' end of a polynucleotide encoding a first binding domain derived from an immunoglobulin variable region to the 5' end of a polynucleotide encoding a constant sub-region;

b. covalently linking the 5' end of a polynucleotide encoding a scorpion linker to the 3' end of the polynucleotide encoding the constant sub-region; and

c. covalently linking the 5' end of a polynucleotide encoding a second binding domain derived from an immunoglobulin variable region to the 3' end of the polynucleotide encoding the scorpion linker,

thereby generating a nucleic acid encoding a multivalent binding protein with effector function.

57. The method according to claim 56 wherein the polynucleotide encoding a first binding domain comprises a sequence selected from the group consisting of SEQ ID NO:1 (anti-CD20 variable region), SEQ ID NO:3 (anti-CD28 variable region, oriented V_L-V_H) and SEQ ID NO:5 (anti-CD28 variable region, oriented V_H-V_L).

58. The method according to claim 56, further comprising a linker polynucleotide inserted between the polynucleotide encoding a first binding domain and the polynucleotide encoding a constant sub-region, the linker polynucleotide encoding a peptide linker of at least 5 amino acids.

59. The method according to claim 56, further comprising a linker polynucleotide inserted between the polynucleotide encoding a constant sub-region and the polynucleotide encoding a second binding domain, the linker polynucleotide encoding a peptide linker of at least 5 amino acids.

60. A method of inducing damage to a target cell comprising contacting a target cell with a therapeutically effective amount of a protein according to claim 1.

61. The method according to claim 60 wherein the target cell is contacted *in vivo* by administration of the protein to an organism in need.

62. A method of treating a cell proliferation disorder comprising administering a therapeutically effective amount of a protein according to claim 1 to an organism in need.

63. The method according to claim 62 wherein the disorder is selected from the group consisting of a cancer, an autoimmune disorder, Rous Sarcoma Virus infection and inflammation.

64. The method according to claim 62 wherein the protein is administered by *in vivo* expression of a nucleic acid encoding the protein according to claim 1.

65. The method according to claim 62 wherein the protein is administered by a route selected from the group consisting of intravenous injection, intraarterial injection, intramuscular injection, subcutaneous injection, intraperitoneal injection and direct tissue injection.

66. A method of ameliorating a symptom associated with a cell proliferation disorder comprising administering a therapeutically effective amount of a protein according to claim 1 to an organism in need.

67. The method according to claim 66, wherein the symptom is selected from the group consisting of pain, heat, swelling and joint stiffness.

68. A method of treating an infection associated with an infectious agent comprising administering a therapeutically effective amount of the protein according to claim 1 to a patient in need, wherein the protein comprises a binding domain that specifically binds a target molecule of the infectious agent.

69. A method of ameliorating a symptom of an infection associated with an infectious agent comprising administering an effective amount of the protein according to claim 1 to a patient in need, wherein the protein comprises a binding domain that specifically binds a target molecule of the infectious agent.

70. A method of reducing the risk of infection attributable to an infectious agent comprising administering a prophylactically effective amount of the protein according to claim 1 to a patient at risk of developing the infection, wherein the protein comprises a binding domain that specifically binds a target molecule of the infectious agent.

71. A kit comprising the protein according to claim 1 and a set of instructions for administering the protein to treat a cell proliferation disorder or to ameliorate a symptom of the cell proliferation disorder.

72. The multivalent single-chain binding protein according to claim 1 wherein at least one of the first binding domain and the second binding domain specifically binds an antigen selected from the group consisting of CD19, CD20,

CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide.

73. The protein according to claim 72 wherein one of the first binding domain and the second binding domain specifically binds CD20.

74. The protein according to claim 73 wherein the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide.

75. The protein according to claim 72 wherein one of the first binding domain and the second binding domain specifically binds CD79b.

76. The protein according to claim 75 wherein the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide.

77. The protein according to claim 72 wherein one of the first binding domain and the second binding domain specifically binds a major histocompatibility complex class II peptide.

78. The protein according to claim 77 wherein the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86, and a major histocompatibility complex class II peptide.

79. The protein according to claim 72 wherein one of the first binding domain and the second binding domain specifically binds CD22.

80. The protein according to claim 79 wherein the other binding domain specifically binds an antigen selected from the group consisting of CD19, CD20, CD21, CD22, CD23, CD30, CD37, CD40, CD70, CD72, CD79a, CD79b, CD80, CD81, CD86 and a major histocompatibility complex class II peptide.

81. The protein according to claim 72 wherein the protein has a synergistic effect on a target cell behavior relative to the sum of the effects of each of the binding domains, and wherein the protein comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20-CD19, CD20-CD21, CD20-CD22, CD20-CD40, CD20-CD79a, CD20-CD79b and CD20-CD81.

82. The protein according to claim 72 wherein the protein has an additive effect on a target cell behavior relative to the sum of the effects of each of the binding domains, and wherein the protein comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20-CD23, CD20-CD30, CD20-CD37, CD20-CD70, CD20-CD80, CD20-CD86, CD79b-CD37, CD79b-CD81, major histocompatibility complex class II peptide-CD30, and major histocompatibility complex class II peptide-CD72.

83. The protein according to claim 72 wherein the protein has an inhibitory effect on a target cell behavior relative to the sum of the effects of each of the binding domains, and wherein the protein comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20-major histocompatibility complex class II peptide, CD79b-CD19, CD79b-CD20, CD79b-CD21, CD79b-CD22, CD79b-CD23, CD79b-CD30, CD79b-CD40, CD79b-CD70, CD79b-CD72, CD79b-CD79a, CD79b-CD80, CD79b-CD86, CD79b-major histocompatibility complex class II peptide, major histocompatibility complex class II peptide-CD19, major histocompatibility complex class II peptide-CD20, major histocompatibility complex class II peptide-CD21, major histocompatibility complex class II peptide-CD22, major histocompatibility complex class II peptide-CD23, major histocompatibility complex class II peptide-CD37, major histocompatibility complex class II peptide-CD40, major histocompatibility complex class II peptide-CD70, major histocompatibility complex class II peptide-CD79a, major histocompatibility complex class II peptide-CD79b, major histocompatibility complex class II peptide-CD80, major histocompatibility complex class II peptide-CD81, major histocompatibility complex class II peptide-CD86, CD22-CD19, CD22-CD40, CD22-CD79b, CD22-CD86 and CD22-major histocompatibility complex class II peptide.

84. A method of identifying at least one of the binding domains of the multivalent binding molecule according to claim 1 comprising:

(a) contacting an anti-isotypic antibody with an antibody specifically recognizing a first antigen and an antibody specifically recognizing a second antigen;

(b) further contacting a target comprising at least one of the antigens with the composition of step (a); and

(c) measuring an activity of the target, wherein the activity is used to identify at least one of the binding domains of the multivalent binding molecule.

85. The method according to claim 84 wherein the target is a diseased cell.

86. The method according to claim 60 wherein the multivalent single-chain binding protein induces a synergistic amount of damage to the target cell compared to the sum of the damage induced by a first antibody comprising the first binding domain but not the second binding domain and a second antibody comprising the second binding domain but not the first binding domain.

87. The method according to claim 86 wherein the multivalent single-chain binding protein comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD19/CD20, CD20/CD21, CD20/CD22, CD20/CD40, CD20/CD79a, CD20/CD79b, CD20/CD81, CD21/CD79b, CD37/CD79b, CD79b/CD81, CD19/CL II, CD20/CL II, CD30/CL II, CD37/CL II, CD72/CL II, and CD79b/CL II.

88. The method according to claim 60 wherein the multivalent single-chain binding protein induces an inhibited amount of damage to the target cell compared to the sum of the damage induced by a first antibody comprising the first binding domain but not the second binding domain and a second antibody comprising the second binding domain but not the first binding domain.

89. The method according to claim 88 wherein the multivalent single-chain binding protein comprises a binding domain pair specifically recognizing a pair of antigens selected from the group consisting of CD20/CL II, CD21/CD79b, CD22/CD79b, CD40/CD79b, CD70/CD79b, CD72/CD79b, CD79a/CD79b,

CD79b/CD80, CD79b/CD86, CD21/CL II, CD22/CL II, CD23/CL II, CD40/CL II, CD70/CL II, CD80/CL II, CD86/CL II, CD19/CD22, CD20/CD22, CD21/CD22, CD22/CD23, CD22/CD30, CD22/CD37, CD22/CD40, CD22/CD70, CD22/CD72, CD22/79a, CD22/79b, CD22/CD80, CD22/CD86 and CD22/CL II.

90. The method according to claim 60 further comprising a plurality of multivalent single-chain binding proteins.

91. The method according to claim 90 wherein a binding domain of a first multivalent single-chain binding protein and a binding domain of a second multivalent single-chain binding protein induce a synergistic amount of damage to the target cell.

92. The method according to claim 90 wherein a binding domain of a first multivalent single-chain binding protein and a binding domain of a second multivalent single-chain binding protein induce an inhibited amount of damage to the target cell.

93. A composition comprising a plurality of multivalent single-chain binding proteins according to claim 1.

94. The composition according to claim 93 wherein a binding domain of a first multivalent single-chain binding protein and a binding domain of a second multivalent single-chain binding protein are capable of inducing a synergistic amount of damage to a target cell.

95. The composition according to claim 93 wherein a binding domain of a first multivalent single-chain binding protein and a binding domain of a second multivalent single-chain binding protein are capable of inducing an additive amount of damage to a target cell.

96. The composition according to claim 93 wherein a binding domain of a first multivalent single-chain binding protein and a binding domain of a second multivalent single-chain binding protein are capable of inducing an inhibited amount of damage to a target cell.

97. A pharmaceutical composition comprising the composition according to claim 93 and a pharmaceutically acceptable carrier, diluent or excipient.

98. A kit comprising the composition according to claim 93 and a set of instructions for administering the composition to damage a target cell.

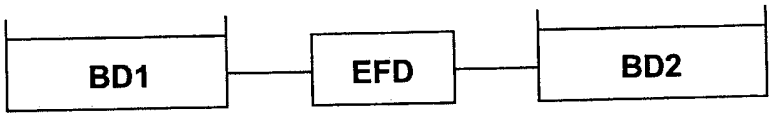


FIG. 1

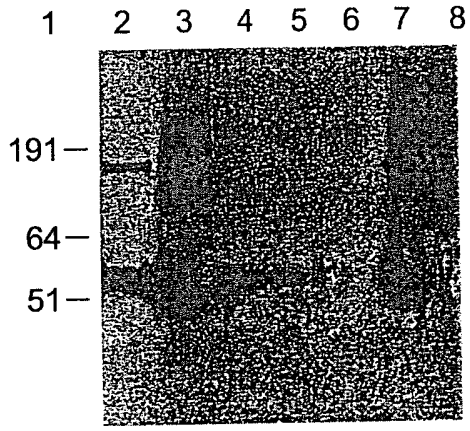
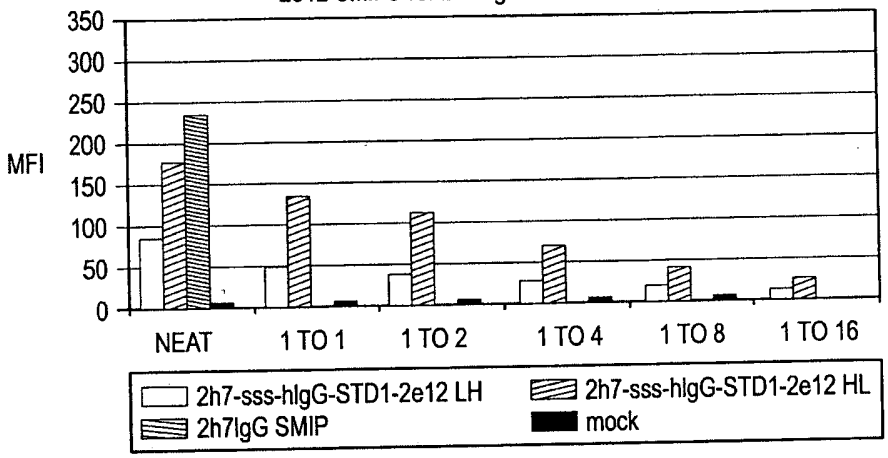


FIG. 2

FIG. 3

BINDING OF PROTEIN EXPRESSED IN COS SUPERNATANTS TO CELLS EXPRESSING TARGET ANTIGENS

A. BINDING OF COS SUPERNATANTS TO WIL-2S CELLS
2e12 SMIPs vs. 2H7-IgG-2e12 SCORPIONS



B. BINDING OF COS SUPERNATANTS TO CD28 OHO CELLS
2e12 SMIPs vs. 2H7-2e12 SCORPIONS

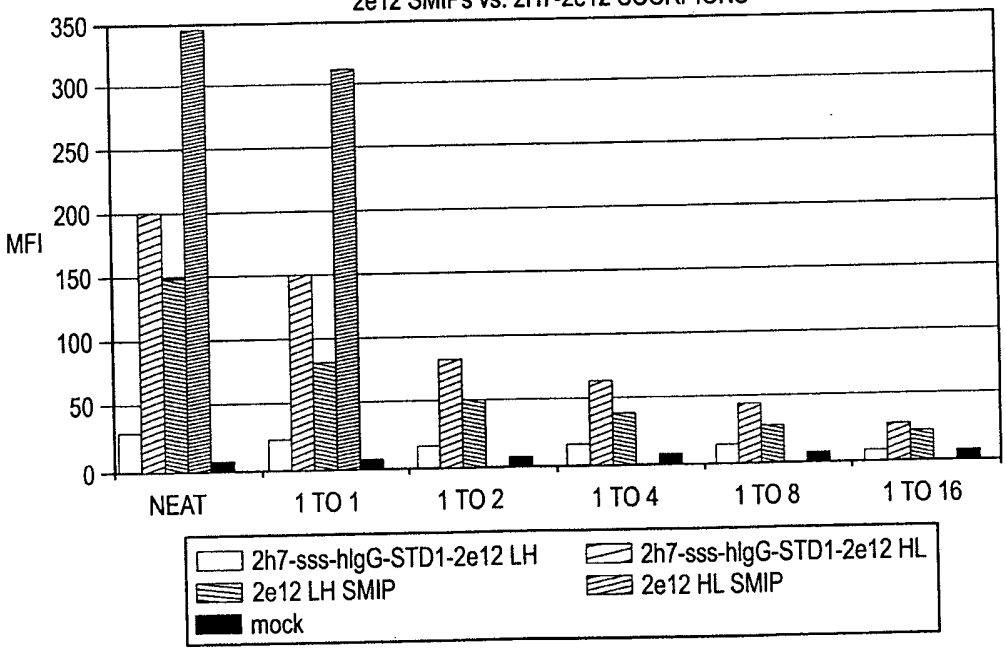


FIG. 4

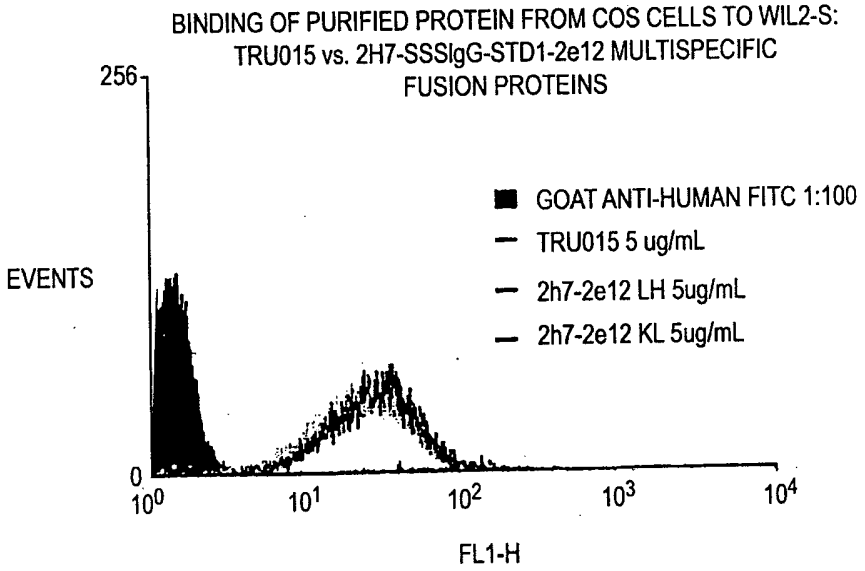

BINDING OF PURIFIED PROTEIN FROM COS CELLS TO WIL2-S:
TRU015 vs. 2H7-SSSIgG-STD1-2e12 MULTISPECIFIC
FUSION PROTEINS
256
EVENTS
0
GOAT ANTI-HUMAN FITC 1:100
TRU015 5 ug/mL
2h7-2e12 LH 5ug/mL
2h7-2e12 KL 5ug/mL
10^0
10^1
10^2
10^3
10^4
FL1-H

FIG. 5

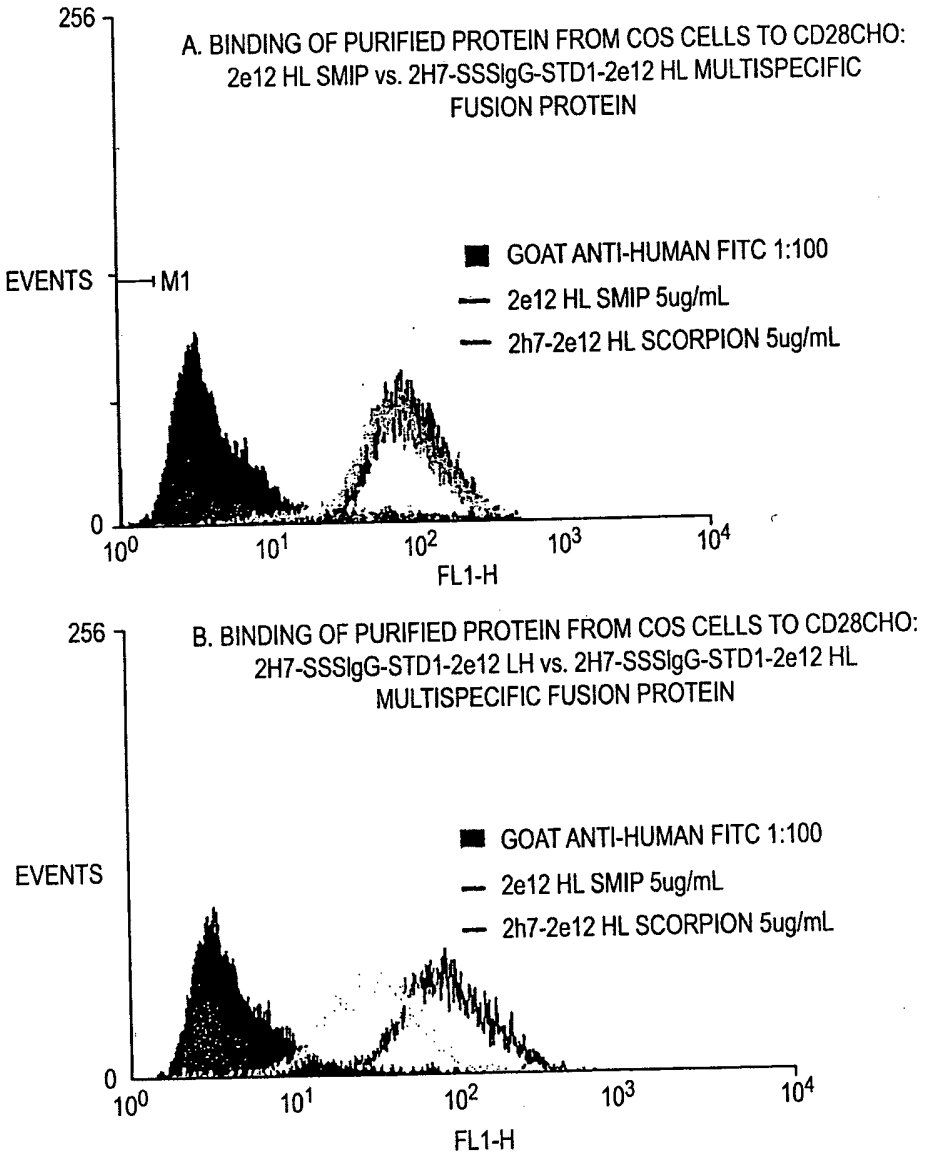

256
A. BINDING OF PURIFIED PROTEIN FROM COS CELLS TO CD28CHO: 2e12 HL SMIP vs. 2H7-SSSIgG-STD1-2e12 HL MULTISPECIFIC FUSION PROTEIN
EVENTS
M1
GOAT ANTI-HUMAN FITC 1:100
2e12 HL SMIP 5ug/mL
2h7-2e12 HL SCORPION 5ug/mL
0
10^0
10^1
10^2
10^3
10^4
FL1-H
256
B. BINDING OF PURIFIED PROTEIN FROM COS CELLS TO CD28CHO: 2H7-SSSIgG-STD1-2e12 LH vs. 2H7-SSSIgG-STD1-2e12 HL MULTISPECIFIC FUSION PROTEIN
EVENTS
GOAT ANTI-HUMAN FITC 1:100
2e12 HL SMIP 5ug/mL
2h7-2e12 HL SCORPION 5ug/mL
0
10^0
10^1
10^2
10^3
10^4
FL1-H

FIG. 6A

TABLE IDENTIFYING FUNCTIONAL ELEMENTS OF MULTISPECIFIC FUSION PROTEINS



BD1 Binding Domain 1	Linker 1	EFD Effector Function Domain	Linker 2 ID	Linker Length (aa) Linker Plus R. Site	CH3 COOH	Linker 2 Sequence (- restriction site)	Fusion Junction with BD2 (L2 RS--QVQ 2e12)	BD2 Binding Domain 2
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	STD	20	PGK	NYGGGGSGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	STD2	38	PGK	NYGGGGSGGGGSGGGGSG NYGGGGSGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	H1 L1=2e12 V_L-V_H	2 (RS)	PGK	--	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	H2 L2=2e12 V_L-V_H	8	PGK	GGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	H3 L3=2e12 V_L-V_H	10	PGK	NYGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	H4 L4=2e12 V_L-V_H	13	PGK	GGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	H5 L5=2e12 V_L-V_H	15	PGK	NYGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	H6 L6=2e12 V_L-V_H	18	PGK	GGGGSGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	Modified hIgGl hinge ccc->sss	hIgGl-CH2-CH3	H7	8	PGK	GCPPCP	NS QVQ	2e12 V_H-V_L

FIG. 6B (1 of 5)

Constructs

Name	BD1	Hinge	EFD	Linker	BD2
2H7-sss-hIgG-STD1-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD1	2e12 (VL-VH)
2H7-sss-hIgG-STD1-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD1	2e12 (VH-VL)
2H7-sss-hIgG-STD2-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD2	2e12 (VL-VH)
2H7-sss-hIgG-STD2-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD2	2e12 (VH-VL)
2H7-csc-hIgG-STD1-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD1	2e12 (VL-VH)
2H7-csc-hIgG-STD2-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD2	2e12 (VL-VH)
2H7-csc-hIgG-STD1-2e12 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD1	2e12 (VL-VH)
2H7-csc-hIgG-STD2-2e12 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD2	2e12 (VH-VL)
2H7-sss-hIgG-H1-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H1	2e12 (VL-VH)
2H7-csc-hIgG-H1-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H1	2e12 (VL-VH)
2H7-sss-hIgG-H1-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H1	2e12 (VH-VL)
2H7-csc-hIgG-H1-2e12 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H1	2e12 (VH-VL)
2H7-sss-hIgG-H2-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H2	2e12 (VH-VL)
2H7-csc-hIgG-H2-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H2	2e12 (VL-VH)
2H7-sss-hIgG-H2-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H2	2e12 (VH-VL)
2H7-csc-hIgG-H2-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H2	2e12 (VH-VL)
2H7-sss-hIgG-H3-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H3	2e12 (VL-VH)
2H7-csc-hIgG-H3-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H3	2e12 (VL-VH)
2H7-sss-hIgG-H3-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H3	2e12 (VH-VL)
2H7-csc-hIgG-H3-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H3	2e12 (VH-VL)
2H7-sss-hIgG-H4-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H4	2e12 (VL-VH)
2H7-csc-hIgG-H4-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H4	2e12 (VL-VH)
2H7-sss-hIgG-H4-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H4	2e12 (VH-VL)
2H7-csc-hIgG-H4-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H4	2e12 (VH-VL)

FIG. 6B (2 of 5)

Name	BD1	Hinge	EFD	Linker	BD2
2H7-sss-hIgG-H5-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H5	2e12 (VL-VH)
2H7-csc-hIgG-H5-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H5	2e12 (VL-VH)
2H7-sss-hIgG-H5-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H5	2e12 (VH-VL)
2H7-csc-hIgG-H5-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H5	2e12 (VH-VL)
2H7-sss-hIgG-H6-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H6	2e12 (VL-VH)
2H7-csc-hIgG-H6-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H6	2e12 (VL-VH)
2H7-sss-hIgG-H6-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H6	2e12 (VH-VL)
2H7-csc-hIgG-H6-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H6	2e12 (VH-VL)
2H7-sss-hIgG-H7-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H7	2e12 (VL-VH)
2H7-csc-hIgG-H7-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H7	2e12 (VL-VH)
2H7-sss-hIgG-H7-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H7	2e12 (VH-VL)
2H7-csc-hIgG-H7-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H7	2e12 (VH-VL)
2H7-sss-hIgG-STD1-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-sss-hIgG-STD1-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-sss-hIgG-STD2-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD2	G28-1 (VL-VH)
2H7-sss-hIgG-STD2-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD2	G28-1 (VH-VL)
2H7-csc-hIgG-STD1-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-csc-hIgG-STD2-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD2	G28-1 (VH-VL)
2H7-csc-hIgG-STD1-G28-1 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-csc-hIgG-STD2-G28-1 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD2	G28-1 (VH-VL)
2H7-sss-hIgG-H1-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H1	G28-1 (VL-VH)
2H7-csc-hIgG-H1-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H1	G28-1 (VH-VL)
2H7-sss-hIgG-H1-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H1	G28-1 (VL-VH)
2H7-csc-hIgG-H1-G28-1 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H1	G28-1 (VH-VL)
2H7-sss-hIgG-H2-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H2	G28-1 (VL-VH)

FIG. 6B (3 of 5)

Name	BD1	Hinge	EFD	Linker	BD2
2H7-csc-hIgG-H2-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H2	G28-1 (VH-VL)
2H7-sss-hIgG-H2-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H2	G28-1 (VL-VH)
2H7-csc-hIgG-H2-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H2	G28-1 (VH-VL)
2H7-sss-hIgG-H3-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H3	G28-1 (VL-VH)
2H7-csc-hIgG-H3-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H3	G28-1 (VH-VL)
2H7-sss-hIgG-H3-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H3	G28-1 (VL-VH)
2H7-csc-hIgG-H3-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H3	G28-1 (VH-VL)
2H7-sss-hIgG-H4-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H4	G28-1 (VL-VH)
2H7-csc-hIgG-H4-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H4	G28-1 (VH-VL)
2H7-sss-hIgG-H4-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H4	G28-1 (VL-VH)
2H7-csc-hIgG-H4-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H4	G28-1 (VH-VL)
2H7-sss-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-sss-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-sss-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-csc-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-sss-hIgG-H6-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H6	G28-1 (VL-VH)
2H7-csc-hIgG-H6-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H6	G28-1 (VL-VH)
2H7-sss-hIgG-H6-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H6	G28-1 (VH-VL)
2H7-csc-hIgG-H6-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H6	G28-1 (VH-VL)
2H7-sss-hIgG-H7-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H7	G28-1 (VL-VH)
2H7-csc-hIgG-H7-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H7	G28-1 (VL-VH)
2H7-sss-hIgG-H7-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H7	G28-1 (VH-VL)
2H7-csc-hIgG-H7-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H7	G28-1 (VH-VL)
2H7-sss-hIgG-STD1-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD1	G19-4 (VL-VH)
2H7-sss-hIgG-STD1-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD1	G19-4 (VH-VL)

FIG. 6B (4 of 5)

Name	BD1	Hinge	EFD	Linker	BD2
2H7-sss-hIgG-STD2-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD2	G19-4 (VL-VH)
2H7-sss-hIgG-STD2-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	STD2	G19-4 (VH-VL)
2H7-csc-hIgG-STD1-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD1	G19-4 (VL-VH)
2H7-csc-hIgG-STD2-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD2	G19-4 (VH-VL)
2H7-csc-hIgG-STD1-G19-4 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD1	G19-4 (VL-VH)
2H7-csc-hIgG-STD2-G19-4 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	STD2	G19-4 (VH-VL)
2H7-sss-hIgG-H1-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H1	G19-4 (VL-VH)
2H7-csc-hIgG-H1-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H1	G19-4 (VH-VL)
2H7-sss-hIgG-H1-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H1	G19-4 (VL-VH)
2H7-csc-hIgG-H1-G19-4 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H1	G19-4 (VH-VL)
2H7-sss-hIgG-H2-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-csc-hIgG-H2-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-sss-hIgG-H2-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-csc-hIgG-H2-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-sss-hIgG-H3-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H3	G19-4 (VL-VH)
2H7-csc-hIgG-H3-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H3	G19-4 (VH-VL)
2H7-sss-hIgG-H3-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H3	G19-4 (VL-VH)
2H7-csc-hIgG-H3-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H3	G19-4 (VH-VL)
2H7-sss-hIgG-H4-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H4	G19-4 (VL-VH)
2H7-csc-hIgG-H4-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H4	G19-4 (VH-VL)
2H7-sss-hIgG-H4-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H4	G19-4 (VL-VH)
2H7-csc-hIgG-H4-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H4	G19-4 (VH-VL)
2H7-sss-hIgG-H5-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H5	G19-4 (VL-VH)
2H7-csc-hIgG-H5-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H5	G19-4 (VH-VL)
2H7-sss-hIgG-H5-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H5	G19-4 (VL-VH)

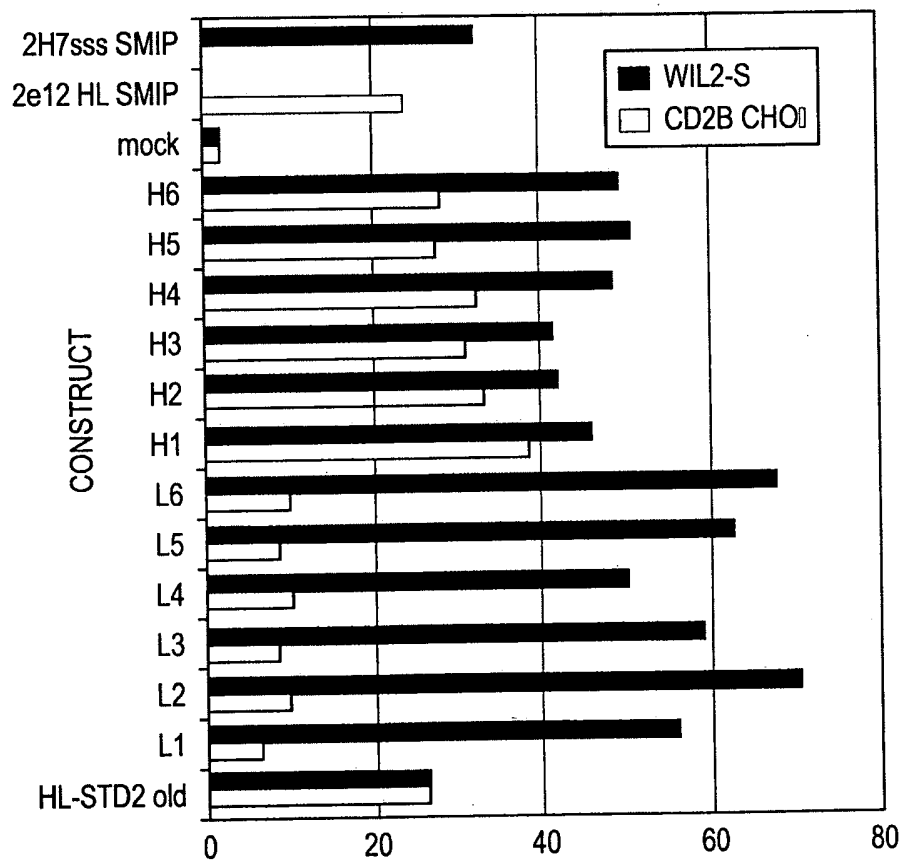
FIG. 6B (5 of 5)

Name	BD1	Hinge	EFD	Linker	BD2
2H7-csc-hIgG-H5-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H5	G19-4 (VH-VL)
2H7-sss-hIgG-H6-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H6	G19-4 (VL-VH)
2H7-csc-hIgG-H6-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H6	G19-4 (VL-VH)
2H7-sss-hIgG-H6-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H6	G19-4 (VH-VL)
2H7-csc-hIgG-H6-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H6	G19-4 (VH-VL)
2H7-sss-hIgG-H7-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H7	G19-4 (VL-VH)
2H7-csc-hIgG-H7-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H7	G19-4 (VL-VH)
2H7-sss-hIgG-H7-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG or hIgG (P238S/P331S)	H7	G19-4 (VH-VL)
2H7-csc-hIgG-H7-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG or hIgG (P238S/P331S)	H7	G19-4 (VH-VL)

FIG. 7

ALTERING THE LINKER AFTER THE EFD AND THE ORIENTATION OF V REGIONS IN BD2 AFFECTS BINDING EFFICIENCY OF MULTISPECIFIC FUSION PROTEINS

BINDING OF 2H7-SSShIgG-L/Hx-2e12 MULTISPECIFIC FUSION PROTEINS WITH VARIOUS LINKERS TO WIL2-S AND CD28 CHO



**ALL FUSION PROTEINS USED AT 0.72 ug/mL, PROTEIN A PURIFIED, GENERATED FROM TRANSIENT COS SUPERNATANTS.

FIG. 8

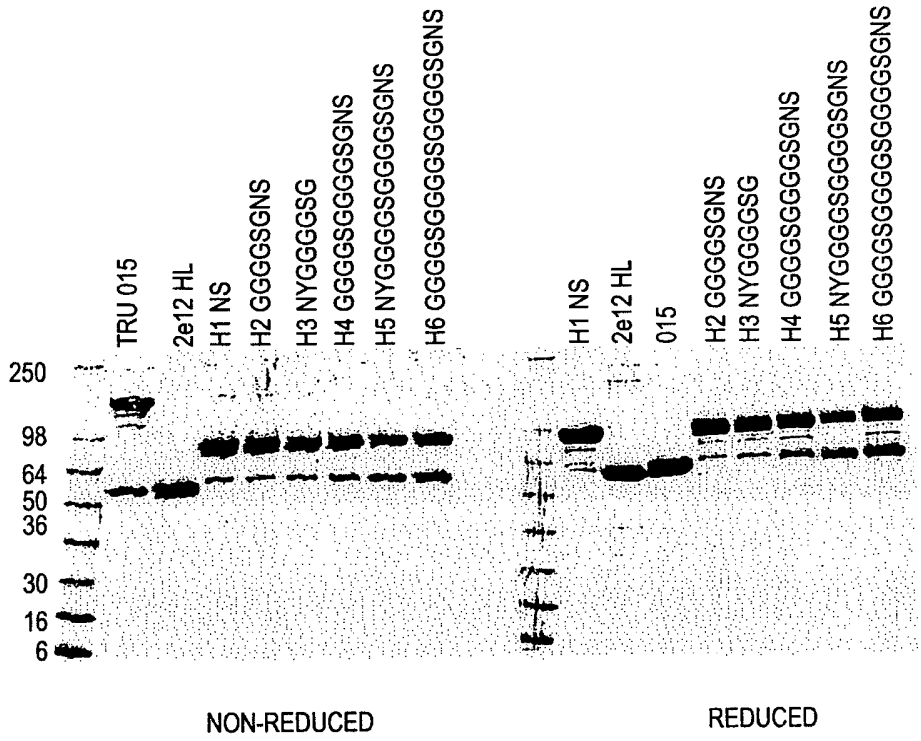

TRU 015
2e12 HL
H1 NS
H2 GGGGSGNS
H3 NYGGGGSG
H4 GGGGSGGGGSGNS
H5 NYGGGGSGGGGSGNS
H6 GGGGSGGGGSGGGGSGNS
250
98
64
50
36
30
16
6
H1 NS
2e12 HL
015
H2 GGGGSGNS
H3 NYGGGGSG
H4 GGGGSGGGGSGNS
H5 NYGGGGSGGGGSGNS
H6 GGGGSGGGGSGGGGSGNS
NON-REDUCED
REDUCED

FIG. 9

Western Blots of Multispecific Fusion Proteins With H6 Linker
A. Absence of SMIP or smaller CD28 detectable forms
B. Presence of a SMIP sized form using CD20 anti-id Fab

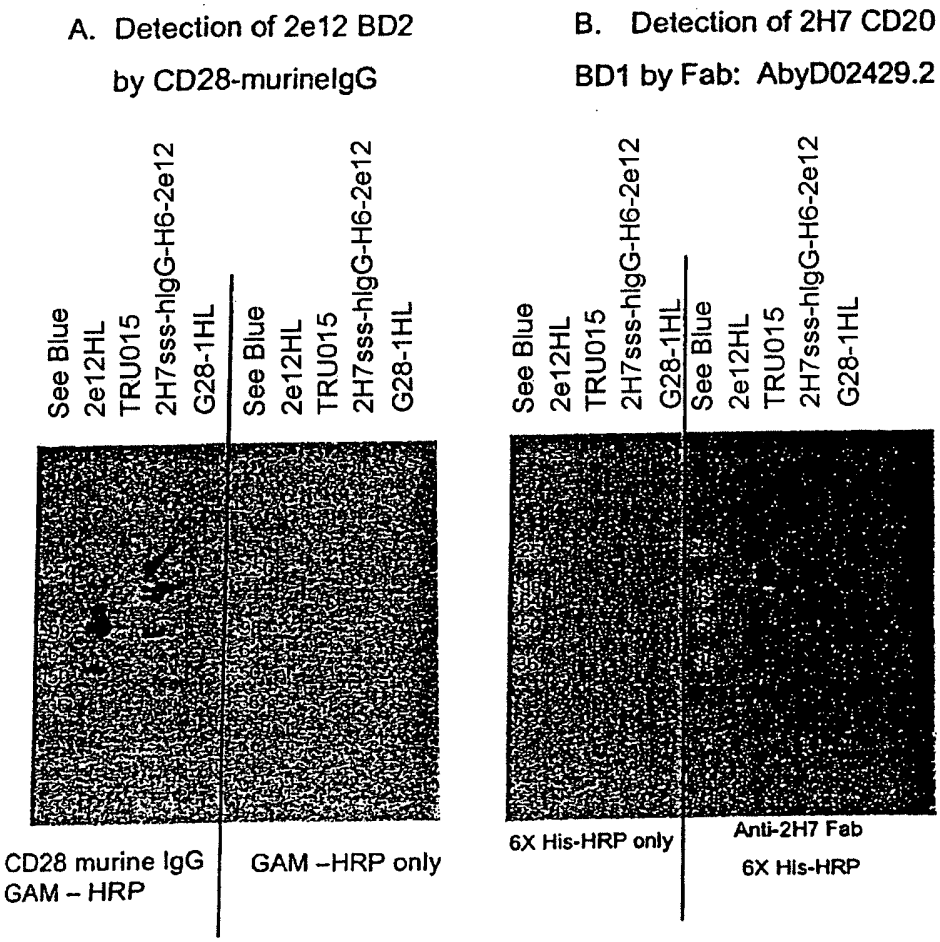
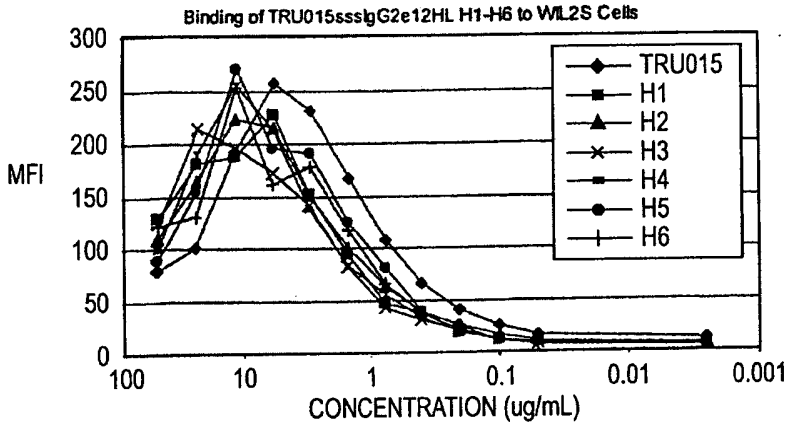


FIG. 10

A. Binding of Multispecific Fusion Proteins With Variant Linkers to WIL-2S Cells



B. Binding of Multispecific Fusion Protein With Variant Linkers to CD28 CHO Cells

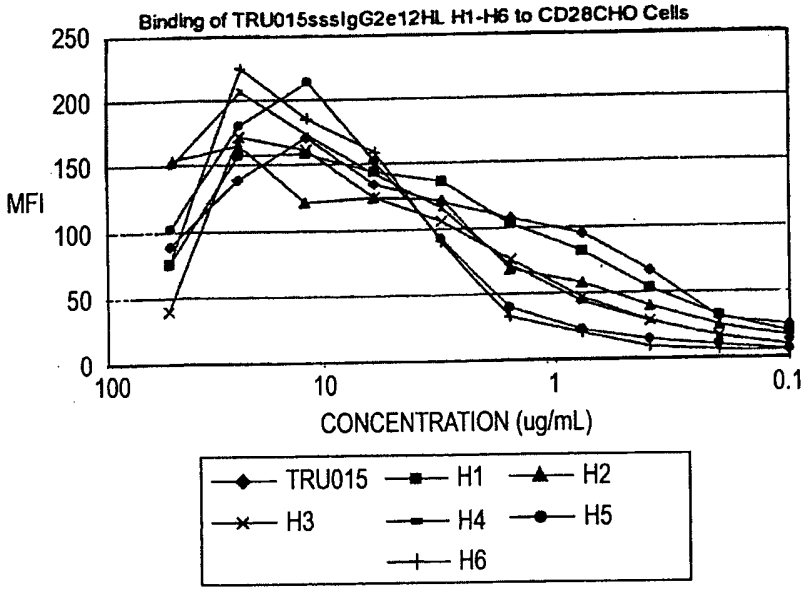


FIG. 11

Summary of SEC Fractionation of 2H7-sssIgG-2e12 HL Multispecific Fusion Proteins with Variant Linkers

Linker ID	Retention Time for POI	% POI	% Other Forms	SDS-PAGE Analysis of Cleavage at EFD/BD2 junction
H1	7.620	70.9	29.0	(Yes)
H2	7.589	67.5	32.5	(Yes)
H3	7.605	68.7	31.4	Yes
H4	7.622 (shoulder)	68.0	32.0	Yes
H5	7.933/7.680 (doublet with shoulder)	32.42/33.94	33.63	YES
H6	7.901/7.69 (doublet with shoulder)	34.5/29.3	36.1	YES
H7	7.788	84.2	15.8 (HMW)	NO

FIG. 12

Binding of [2H7-sss-hIgG-H_X-2e12 HL] Fusion Proteins with Different Linkers to Cells Expressing Target Antigen for BD1 or BD2

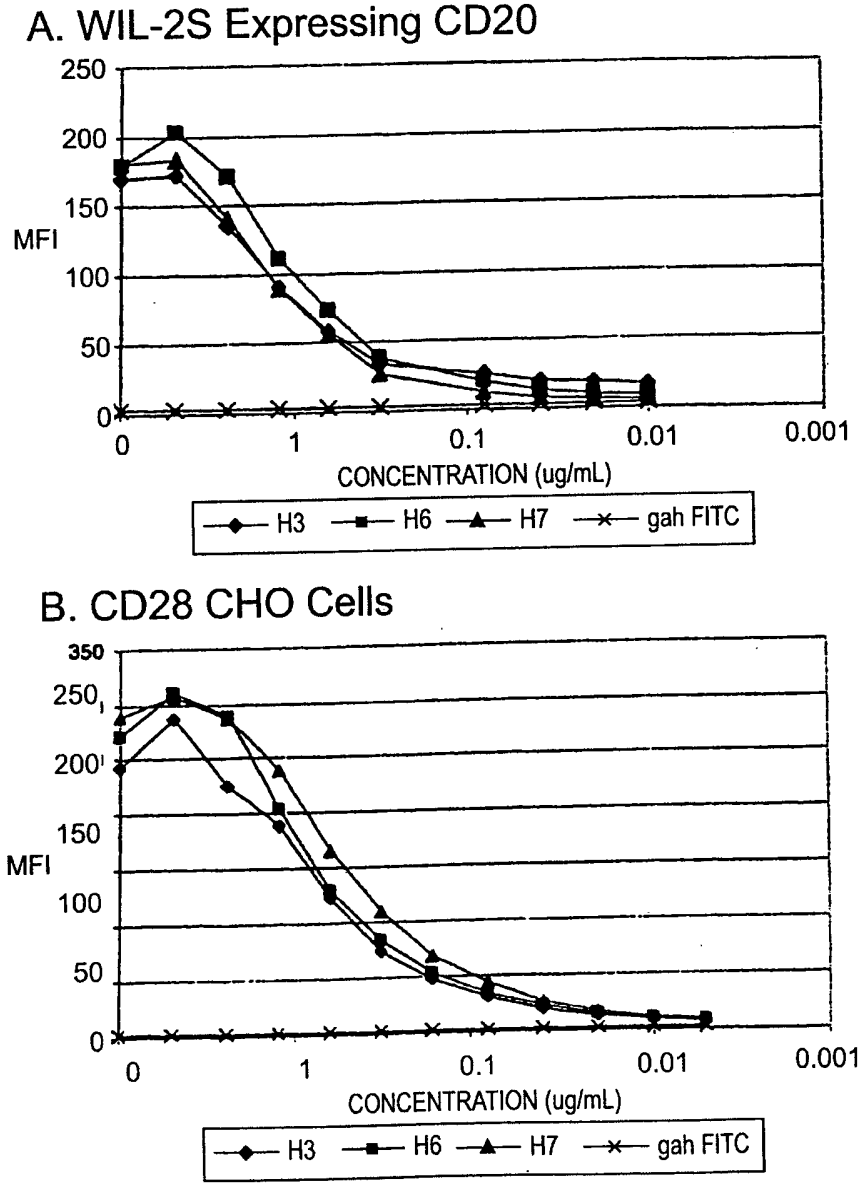


FIG. 13

Simultaneous Binding of BD1 and BD2:

Binding of [2H7-sss-hIgG-H_x-2e12 HL] Fusion Proteins with H3, H6, and H7 Linkers to WIL-2S Cells can be detected with CD28mIgG + FITC anti-mouse

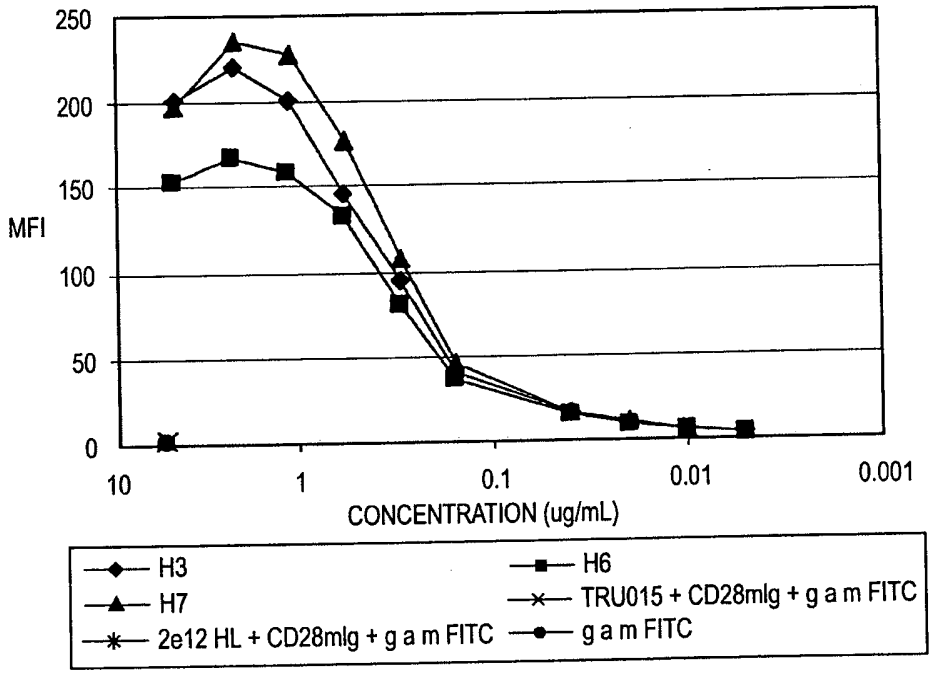
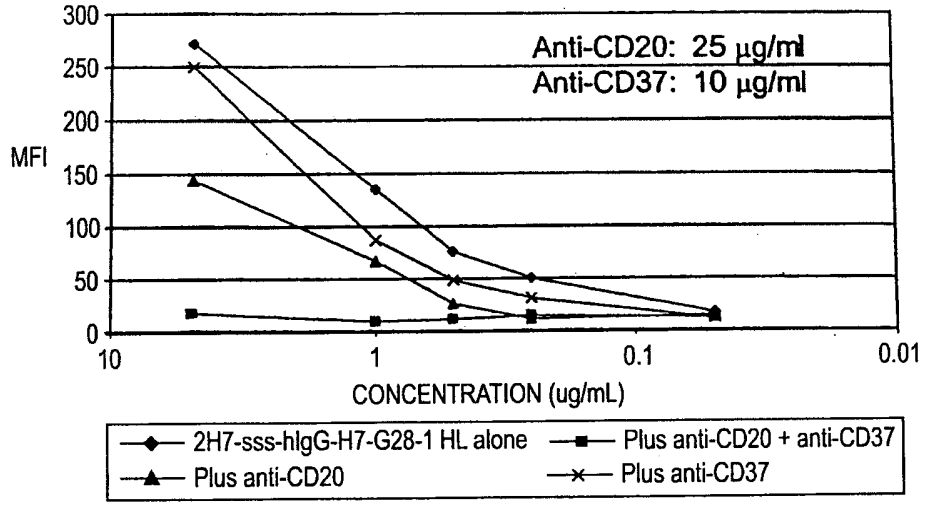


FIG. 14

A. Blocking of Binding of 2H7-sss-hIgG-H7-G28-1 HL Protein to Ramos Cells by CD20 and/or CD37 Targeted Antibodies



B. Blocking of Binding of 2H7-sss-hIgG-H7-G28-1 HL to BJAB Cells by CD20 and/or CD37 Targeted Antibodies

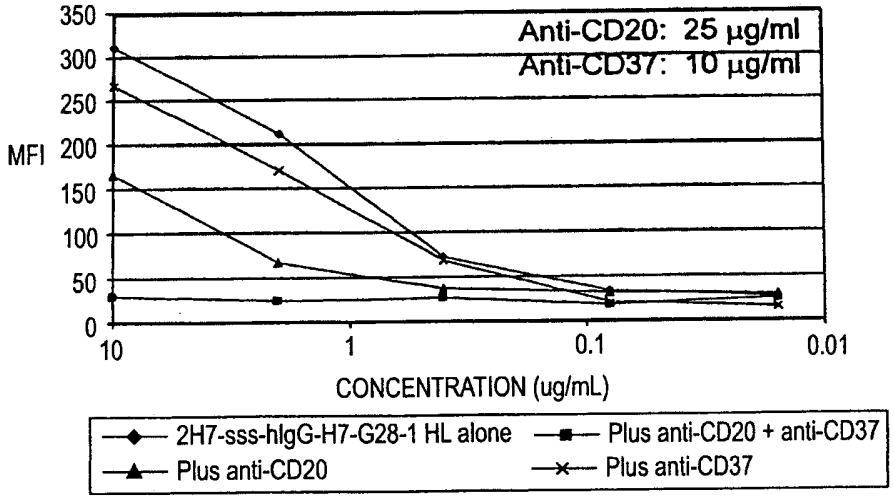


FIG. 15

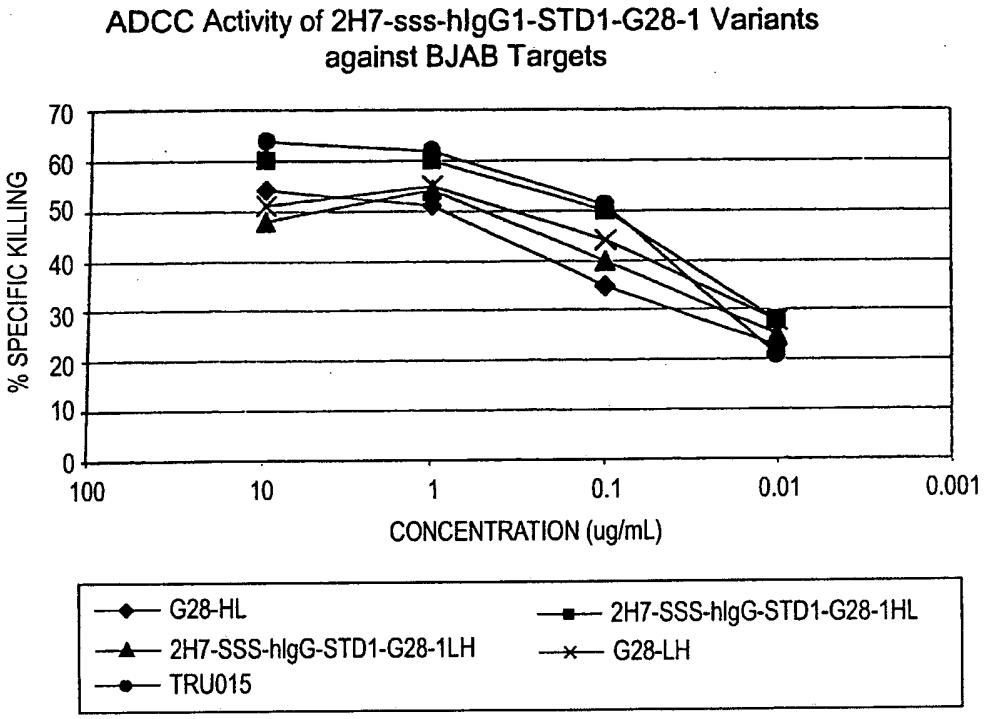

ADCC Activity of 2H7-sss-hIgG1-STD1-G28-1 Variants against BJAB Targets
% SPECIFIC KILLING
70
60
50
40
30
20
10
0
100
10
1
0.1
0.01
0.001
CONCENTRATION (ug/mL)
G28-HL
2H7-SSS-hIgG-STD1-G28-1HL
2H7-SSS-hIgG-STD1-G28-1LH
G28-LH
TRU015

FIG. 16

The Percent CD3, CD19, and CD40 Positive Lymphocytes Present in Culture after Incubation with TRU-015, G28-1 SMIP, TRU-015+G28-1 SMIP, or 2H7-sss-hIgG1-H7-G28-1 HL

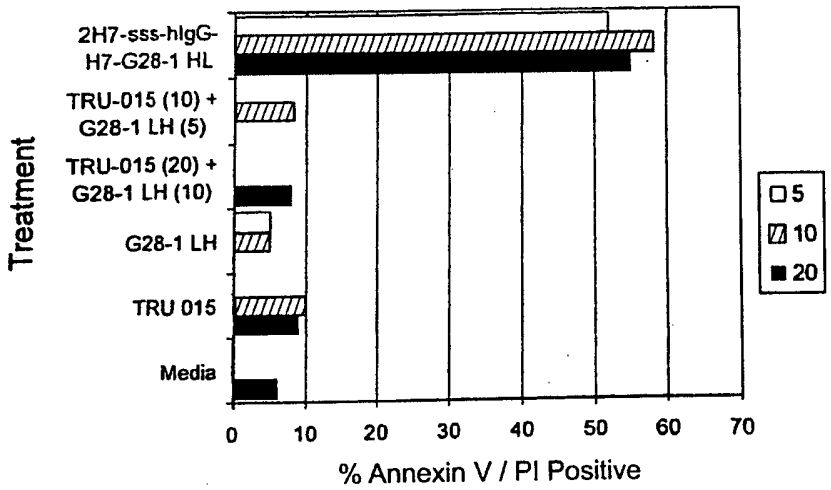
24 Hours	$CD3^+$	$CD19^+$	$CD40^+$
Media	77%	8%	7%
TRU015	76%	8.5%	7.9%
G28-1 SMIP	78%	8. 9%	7.8%
TRU015+G28-1 SMIP	76.3%	8.1%	7.5%
2H7-ssshIgG1-G28-1HL (H7)	78.4%	8.9%	7.9%

72 Hours	$CD3^+$	$CD19^+$	$CD40^+$
Media	78.5%	7.9%	7.7%
TRU015	85.2%	2.3%	2.7%
G28-1 SMIP	86.7%	2.2%	2.3%
TRU015+G28-1 SMIP	87.1%	1.1%	1.1%
2H7-ssshIgG1-G28-1 HL (H7)	96.7%	.08%	.77%

FIG. 17

The Percent Annexin and/or PI Positive B Cells in Culture after 24 Hour Incubation with Single or Multispecific Fusion Proteins.

A. Ramos Cells:



B. Daudi Cells:

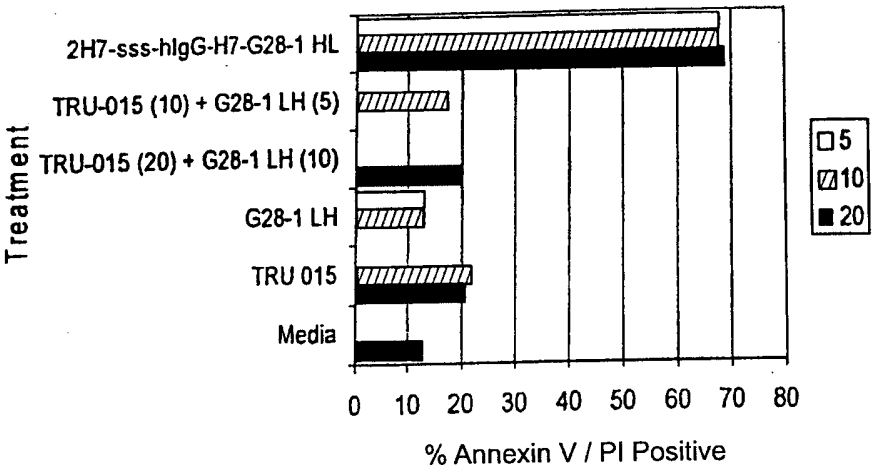


FIG. 18

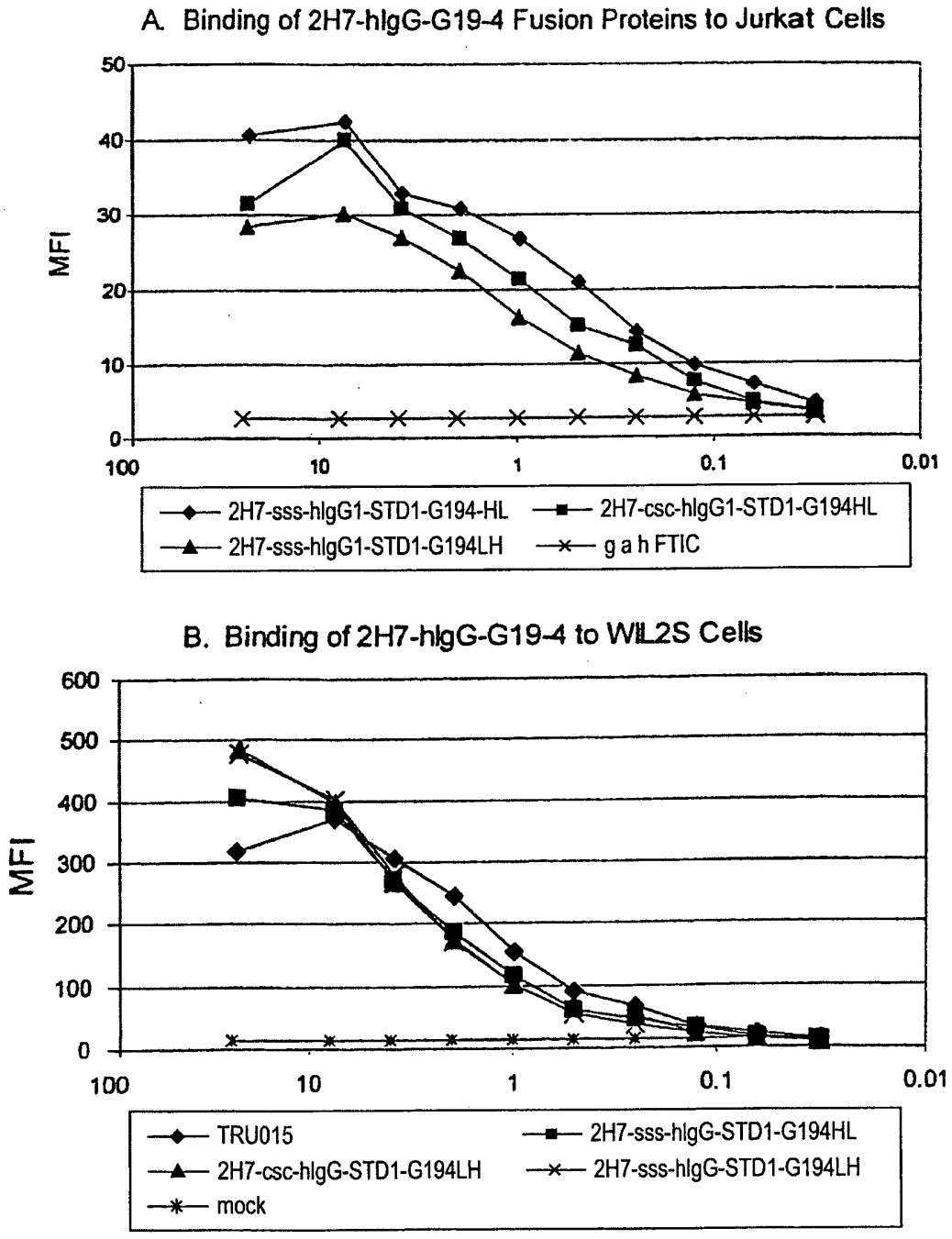
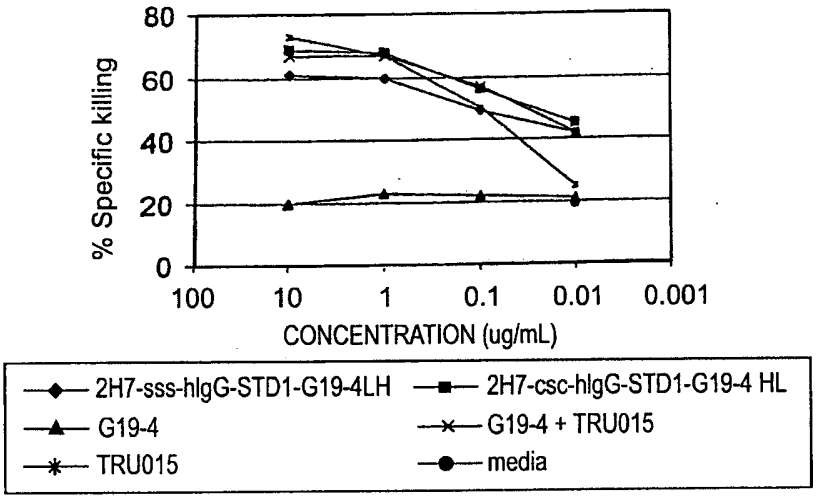

A. Binding of 2H7-hIgG-G19-4 Fusion Proteins to Jurkat Cells
MFI
50
40
30
20
10
0
100
10
1
0.1
0.01
2H7-sss-hIgG1-STD1-G194-HL
2H7-csc-hIgG1-STD1-G194HL
2H7-sss-hIgG1-STD1-G194LH
g a h FTIC
B. Binding of 2H7-hIgG-G19-4 to WIL2S Cells
MFI
600
500
400
300
200
100
0
100
10
1
0.1
0.01
TRU015
2H7-sss-hIgG-STD1-G194HL
2H7-csc-hIgG-STD1-G194LH
2H7-sss-hIgG-STD1-G194LH
mock

FIG. 19

A. ADCC Activity of 2H7-G19-4 Multispecific Fusion Proteins Against BJAB Targets Using Human PBMC Effectors at 25:1



B. ADCC Activity in NK Cell Depiated PBMC Effector Cultures Using BJAB Targets (30:1, E:T) and 2H7-G19-4 Fusion Proteins

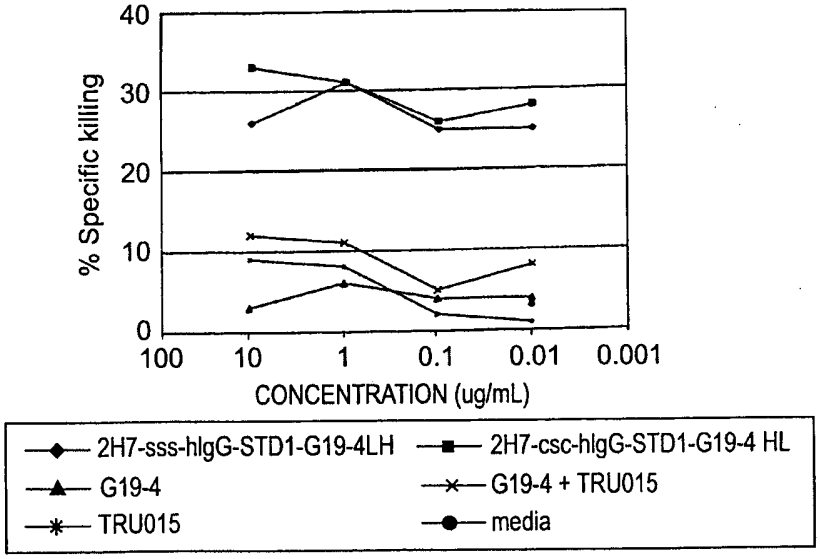


FIG. 20

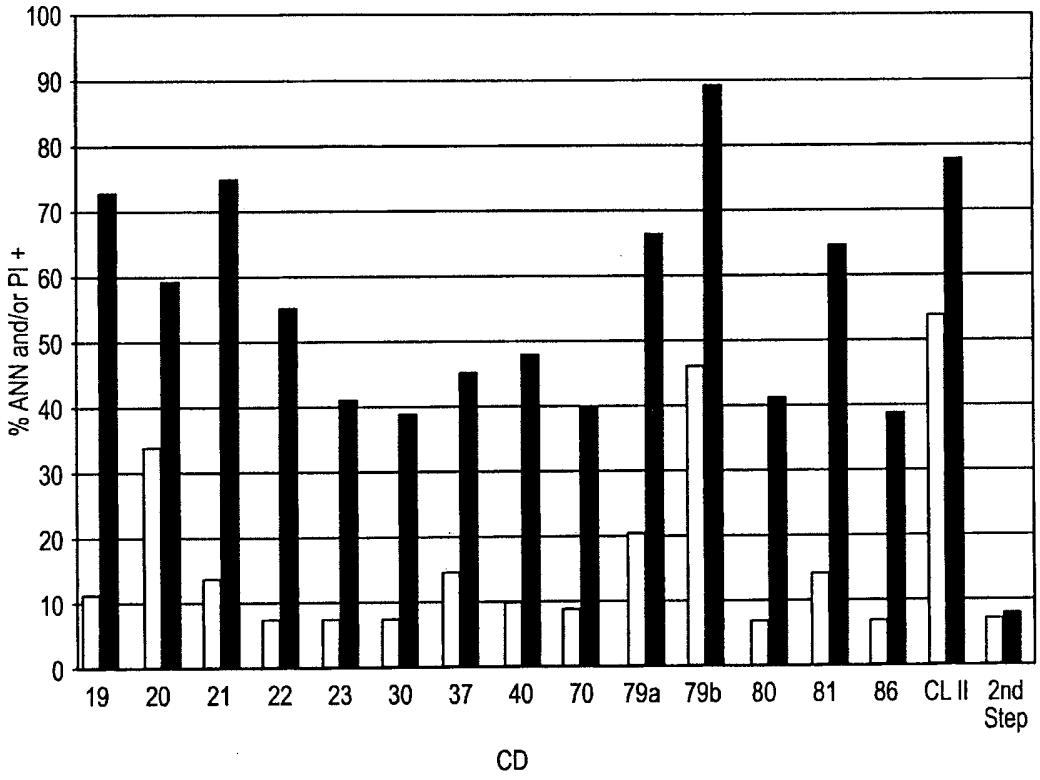
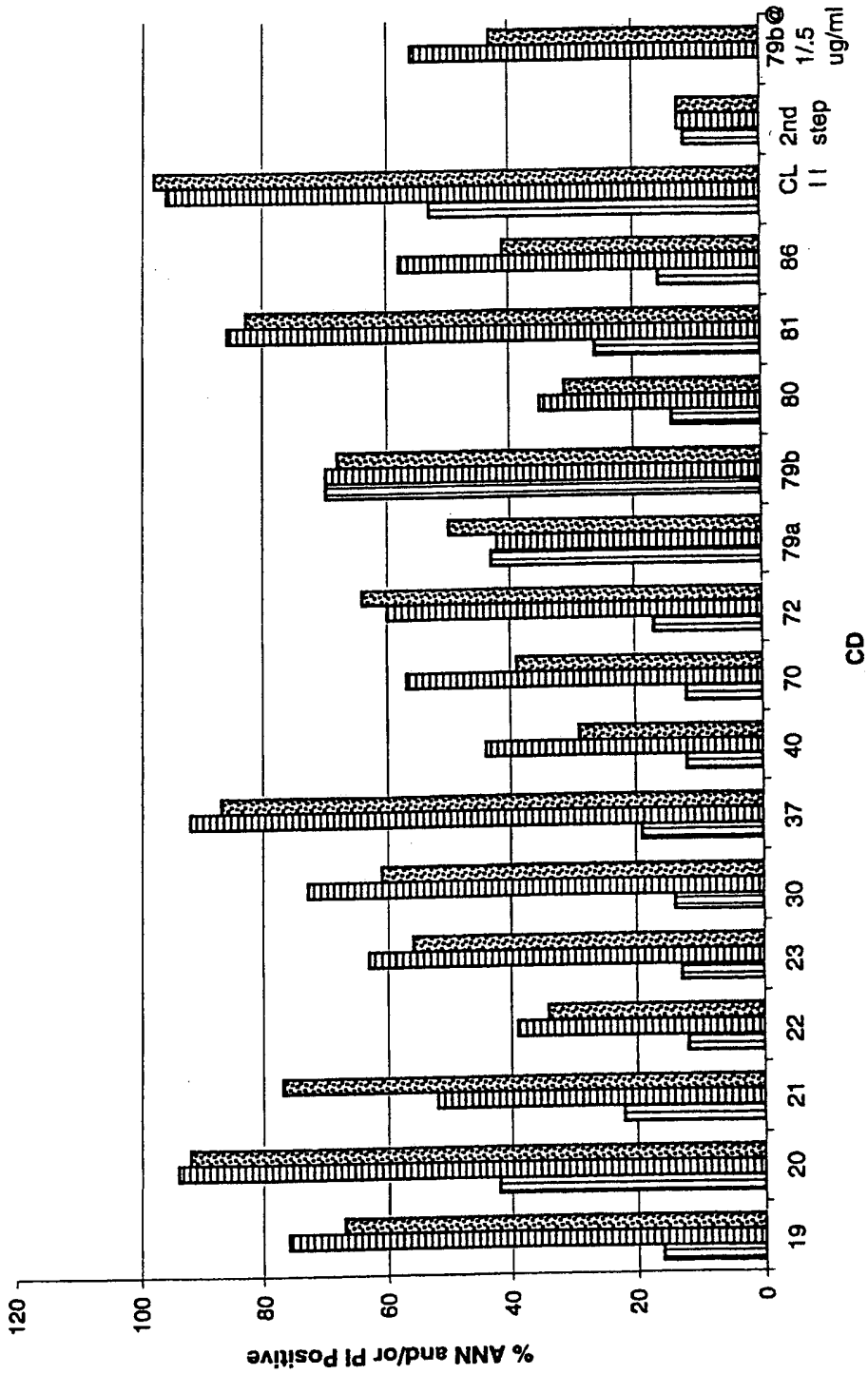

100
90
80
70
60
50
40
30
20
10
0
% ANN and/or PI +
19
20
21
22
23
30
37
40
70
79a
79b
80
81
86
CL II
2nd Step
CD

FIG. 21


% ANN and/or PI Positive
120
100
80
60
40
20
0
CD
19
20
21
22
23
30
37
40
70
72
79a
79b
80
81
86
CL 2nd step 1/.5 ug/ml
79b@

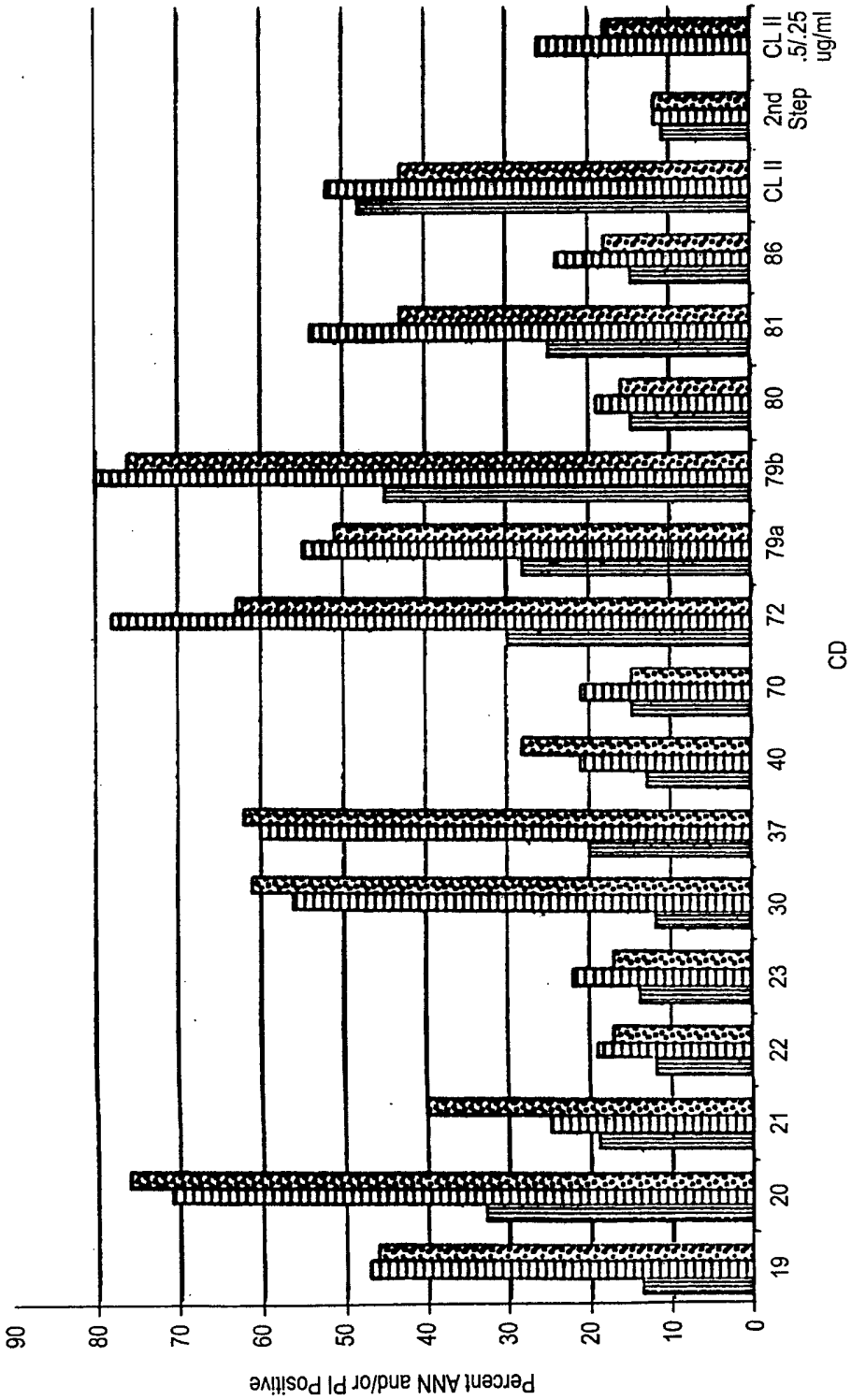

Percent ANN and/or PI Positive
0
10
20
30
40
50
60
70
80
90
CD
19
20
21
22
23
30
37
40
70
72
79a
79b
80
81
86
CL II
2nd Step
CL II .5/.25 ug/ml

FIG. 22

FIG. 23

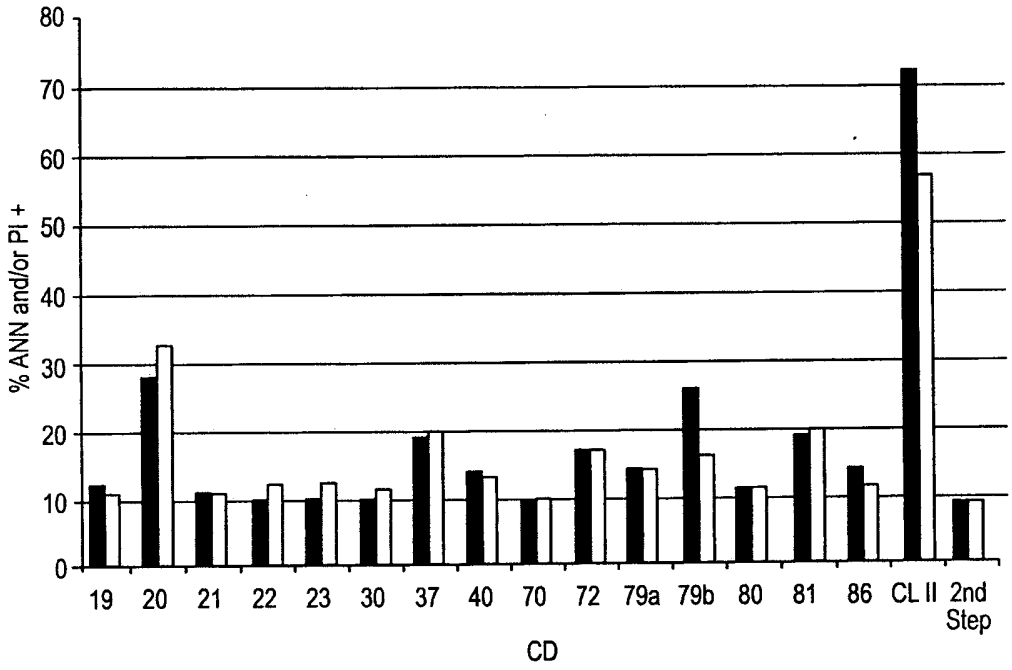

80
70
60
50
40
30
20
10
0
% ANN and/or PI +
19
20
21
22
23
30
37
40
70
72
79a
79b
80
81
86
CL II
2nd Step
CD

FIG. 24

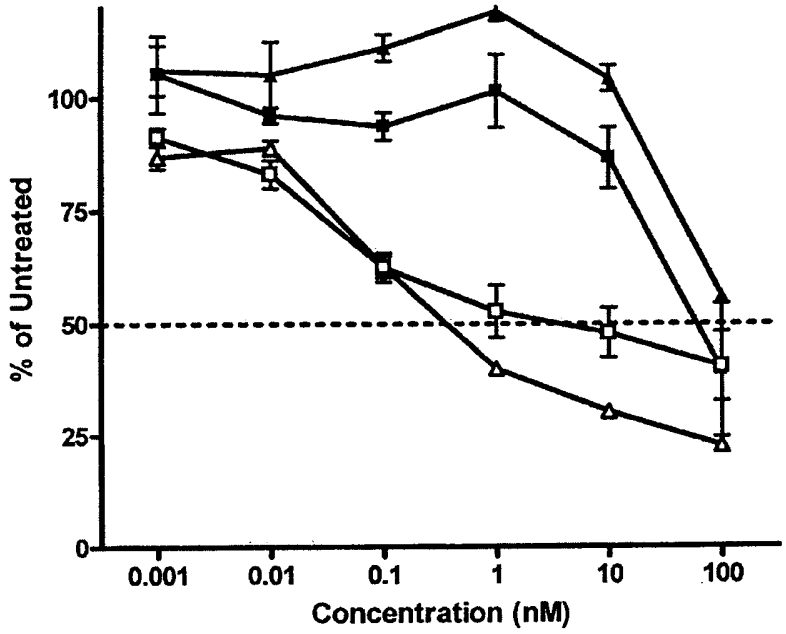

% of Untreated
100
75
50
25
0
0.001
0.01
0.1
1
10
100
Concentration (nM)

FIG. 25

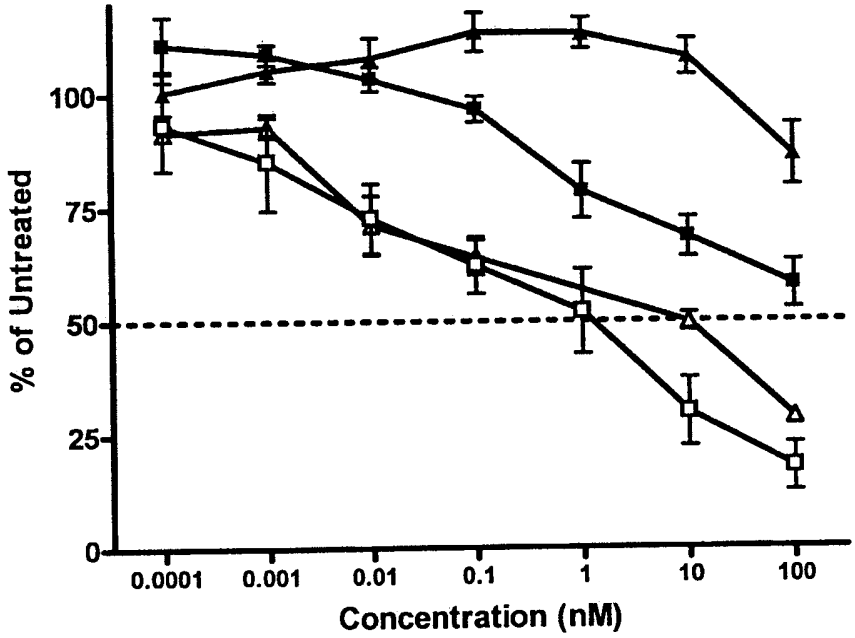

% of Untreated
100
75
50
25
0
0.0001
0.001
0.01
0.1
1
10
100
Concentration (nM)

FIG. 26

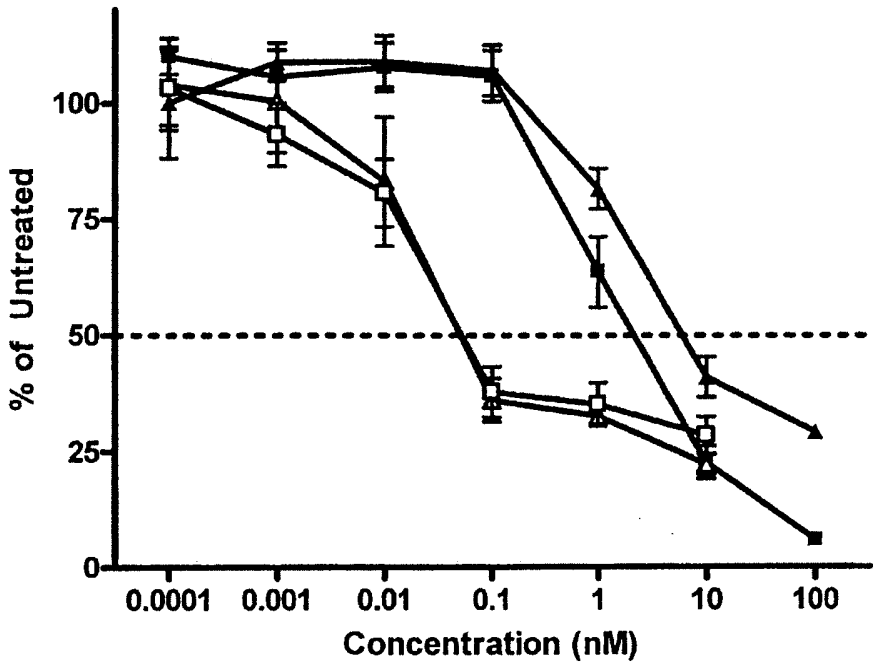

% of Untreated
100
75
50
25
0
0.0001
0.001
0.01
0.1
1
10
100
Concentration (nM)

FIG. 27

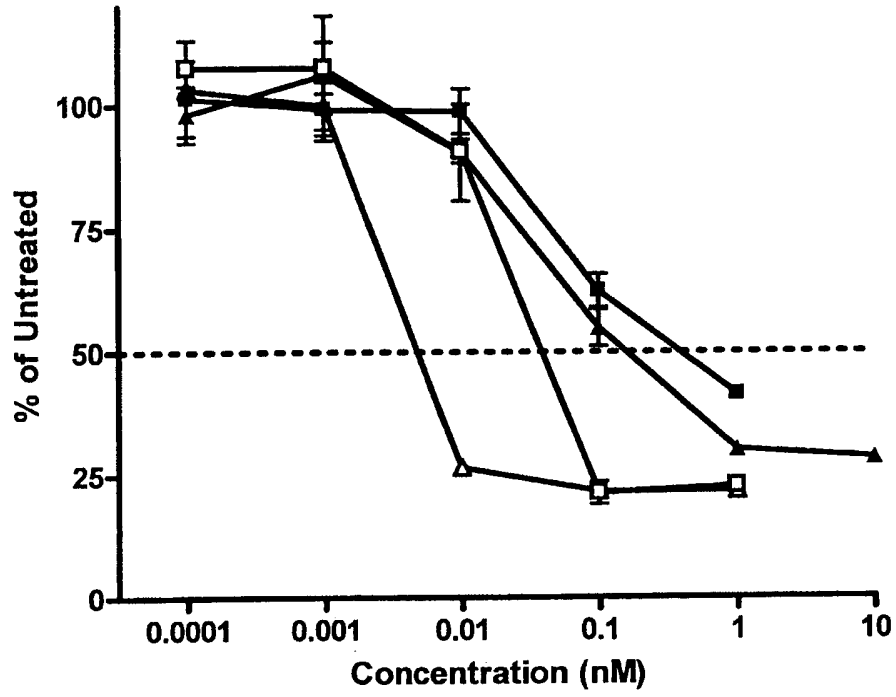

% of Untreated
100
75
50
25
0
0.0001
0.001
0.01
0.1
1
10
Concentration (nM)

FIG. 28

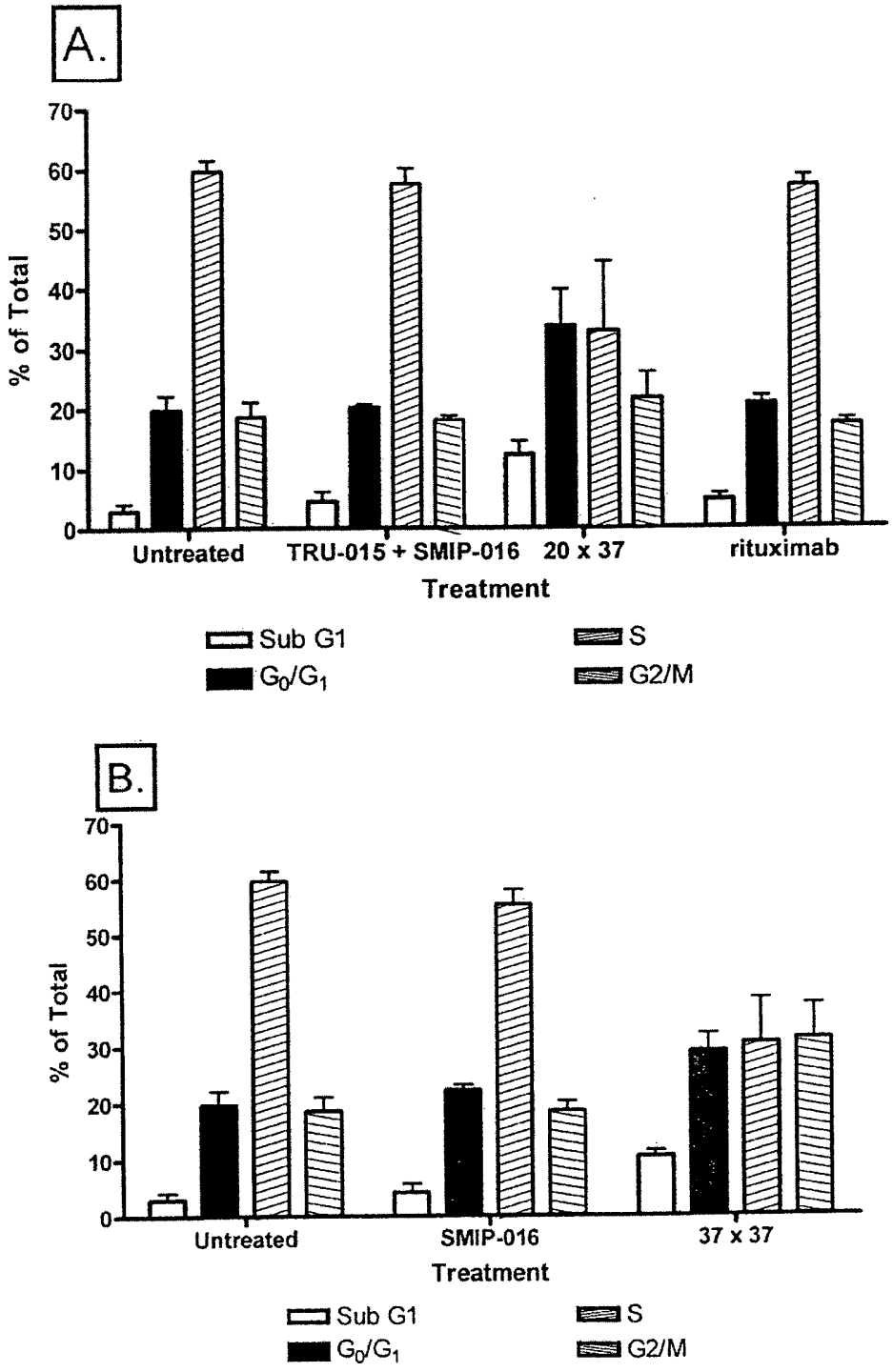

A.
70
60
50
40
30
20
10
0
% of Total
Untreated
TRU-015 + SMIP-016
20 x 37
rituximab
Treatment
Sub G1
G_0/G_1
S
G2/M
B.
70
60
50
40
30
20
10
0
% of Total
Untreated
SMIP-016
37 x 37
Treatment
Sub G1
G_0/G_1
S
G2/M

FIG. 29

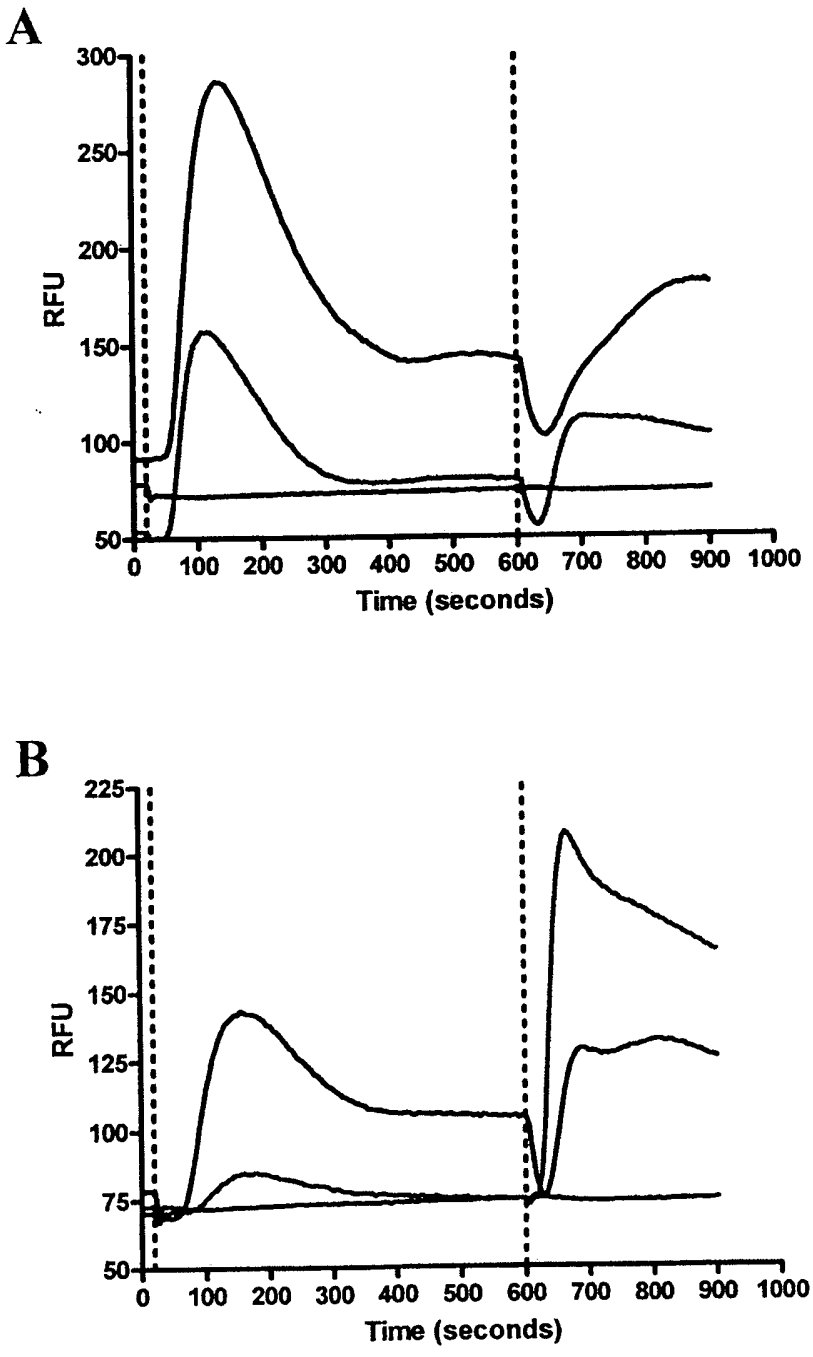

A
300
250
200
150
100
50
RFU
0 100 200 300 400 500 600 700 800 900 1000
Time (seconds)
B
225
200
175
150
125
100
75
50
RFU
0 100 200 300 400 500 600 700 800 900 1000
Time (seconds)

FIG. 30

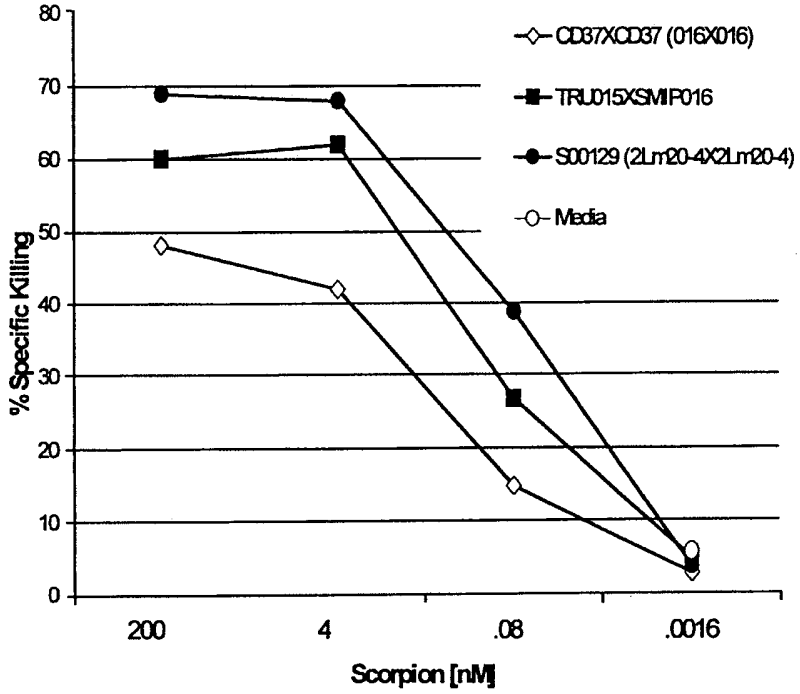

80
70
60
50
40
30
20
10
0
% Specific Killing
200
4
.08
.0016
Scorpion [nM]
CD37XCD37 (016X016)
TRU015XSMIP016
S00129 (2Lm20-4X2Lm20-4)
Media

FIG. 31

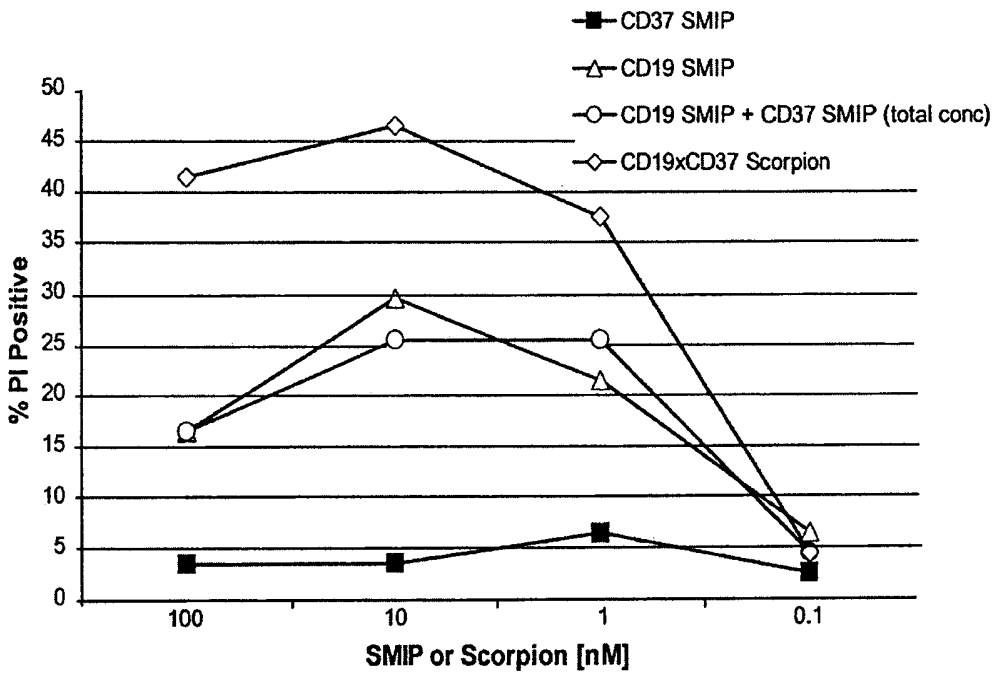

CD37 SMIP
CD19 SMIP
CD19 SMIP + CD37 SMIP (total conc)
CD19xCD37 Scorpion
% PI Positive
50
45
40
35
30
25
20
15
10
5
0
100
10
1
0.1
SMIP or Scorpion [nM]

FIG. 32A

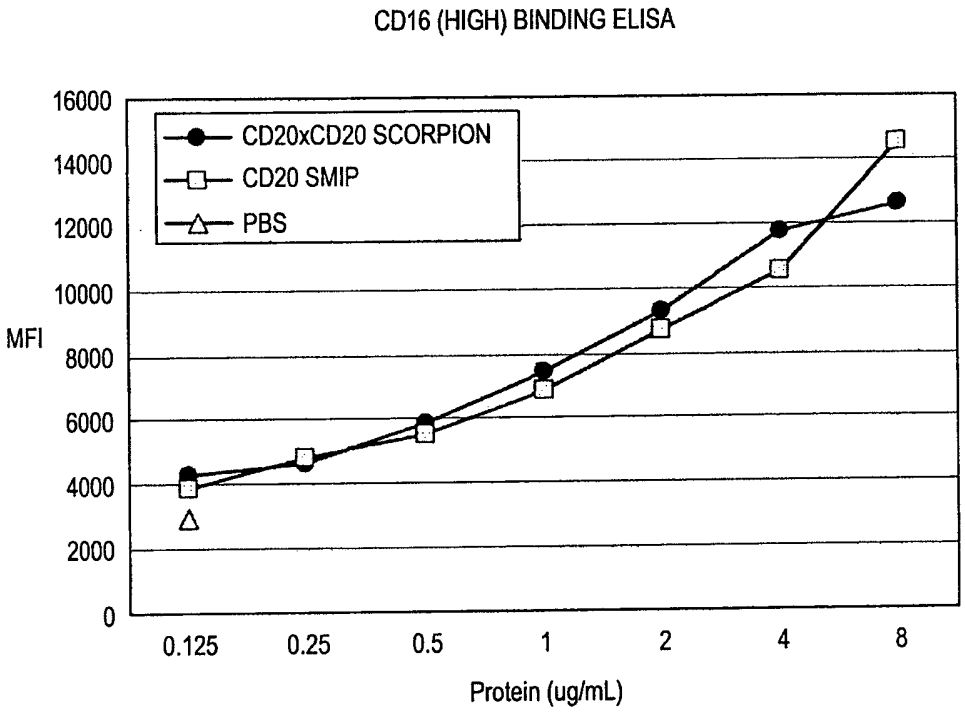

CD16 (HIGH) BINDING ELISA
16000
14000
12000
10000
8000
6000
4000
2000
0
MFI
CD20xCD20 SCORPION
CD20 SMIP
PBS
0.125
0.25
0.5
1
2
4
8
Protein (ug/mL)

FIG. 32B

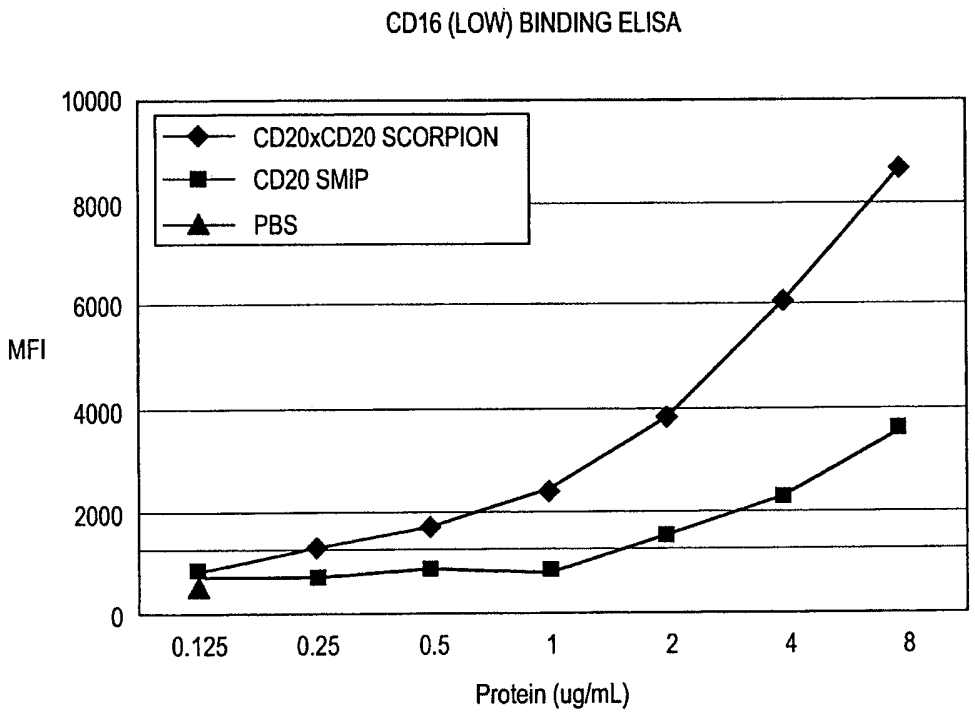

CD16 (LOW) BINDING ELISA
10000
8000
6000
4000
2000
0
MFI
CD20xCD20 SCORPION
CD20 SMIP
PBS
0.125
0.25
0.5
1
2
4
8
Protein (ug/mL)

FIG. 33

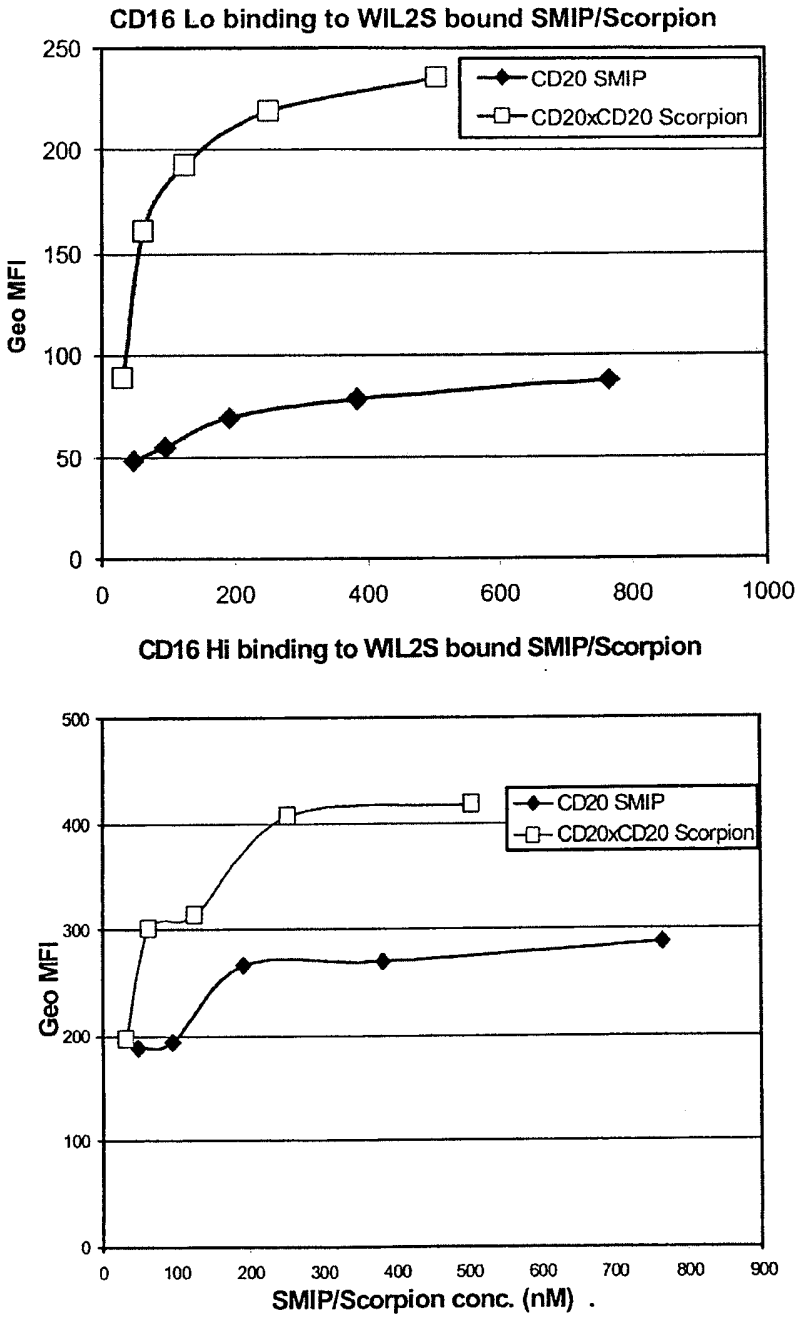

CD16 Lo binding to WIL2S bound SMIP/Scorpion
Geo MFI
250
200
150
100
50
0
0
200
400
600
800
1000
CD20 SMIP
CD20xCD20 Scorpion
CD16 Hi binding to WIL2S bound SMIP/Scorpion
Geo MFI
500
400
300
200
100
0
0
100
200
300
400
500
600
700
800
900
SMIP/Scorpion conc. (nM)
CD20 SMIP
CD20xCD20 Scorpion

FIG. 34

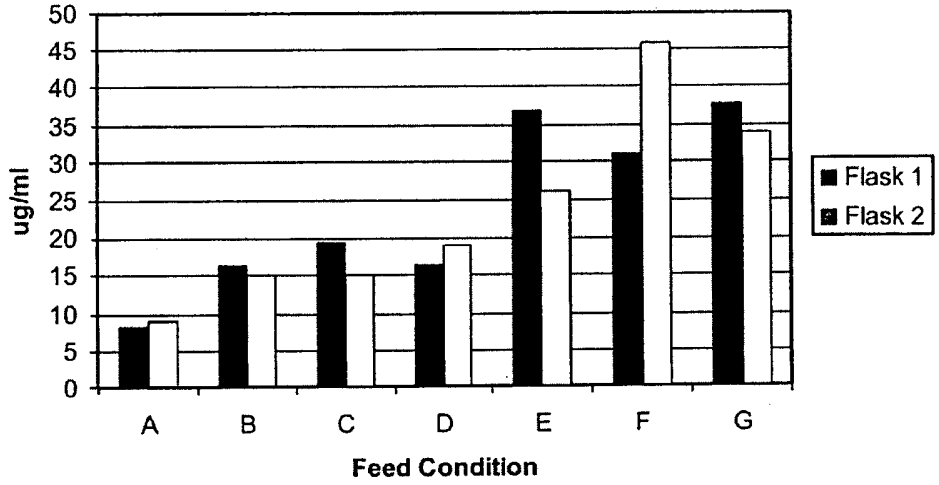

50
45
40
35
30
25
20
15
10
5
0
ug/ml
A
B
C
D
E
F
G
Feed Condition
Flask 1
Flask 2

FIG. 35

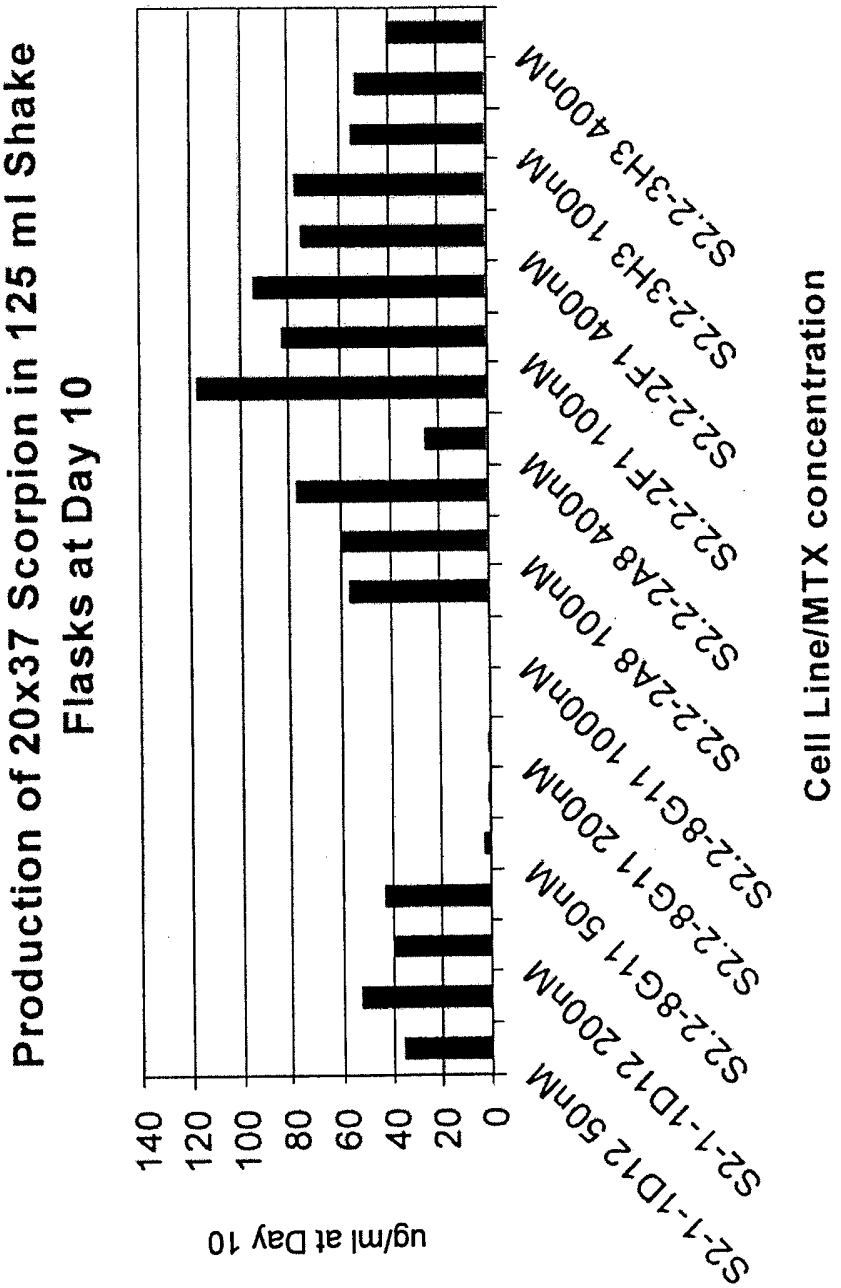

Production of 20x37 Scorpion in 125 ml Shake Flasks at Day 10
ug/ml at Day 10
0
20
40
60
80
100
120
140
S2-1-1D12 50nM
S2-1-1D12 200nM
S2.2-8G11 50nM
S2.2-8G11 200nM
S2.2-8G11 1000nM
S2.2-2A8 100nM
S2.2-2A8 400nM
S2.2-2F1 100nM
S2.2-2F1 400nM
S2.2-3H3 100nM
S2.2-3H3 400nM
Cell Line/MTX concentration

FIG. 36

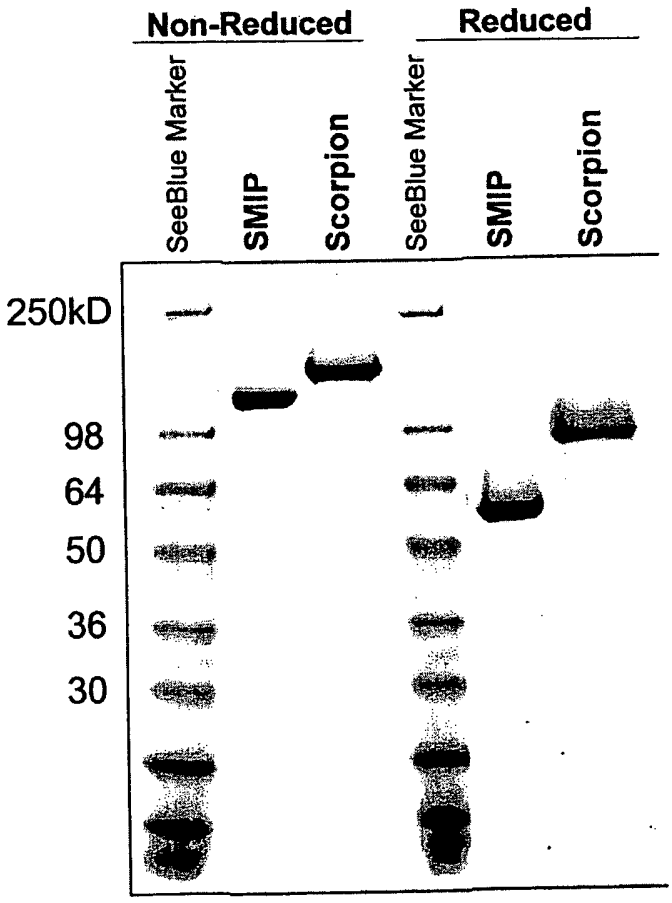

Non-Reduced
Reduced
SeeBlue Marker
SMIP
Scorpion
SeeBlue Marker
SMIP
Scorpion
250kD
98
64
50
36
30

FIG. 37

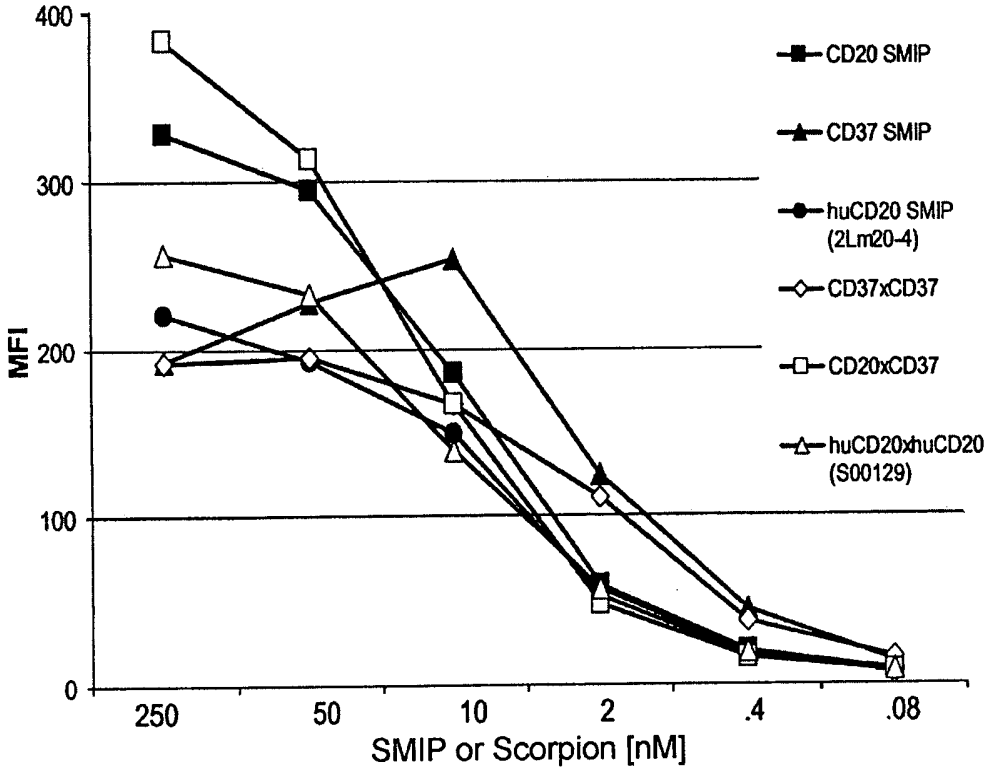

400
300
200
100
0
MFI
250
50
10
2
.4
.08
SMIP or Scorpion [nM]
CD20 SMIP
CD37 SMIP
huCD20 SMIP (2Lm20-4)
CD37xCD37
CD20xCD37
huCD20xhuCD20 (S00129)

FIG. 38

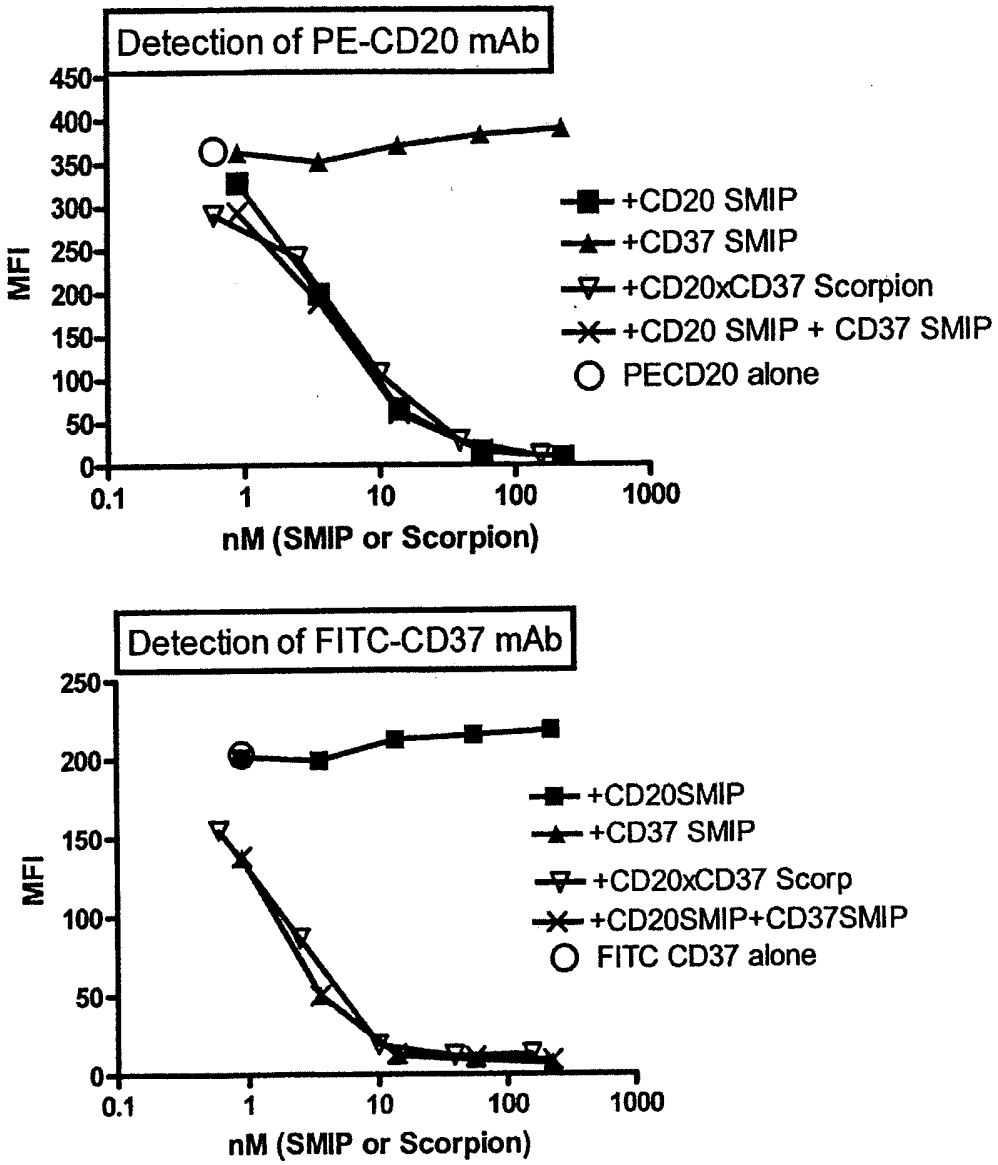

Detection of PE-CD20 mAb
MFI
450
400
350
300
250
200
150
100
50
0
0.1
1
10
100
1000
nM (SMIP or Scorpion)
+CD20 SMIP
+CD37 SMIP
+CD20xCD37 Scorpion
+CD20 SMIP + CD37 SMIP
PECD20 alone
Detection of FITC-CD37 mAb
MFI
250
200
150
100
50
0
0.1
1
10
100
1000
nM (SMIP or Scorpion)
+CD20SMIP
+CD37 SMIP
+CD20xCD37 Scorp
+CD20SMIP+CD37SMIP
FITC CD37 alone

FIG. 39

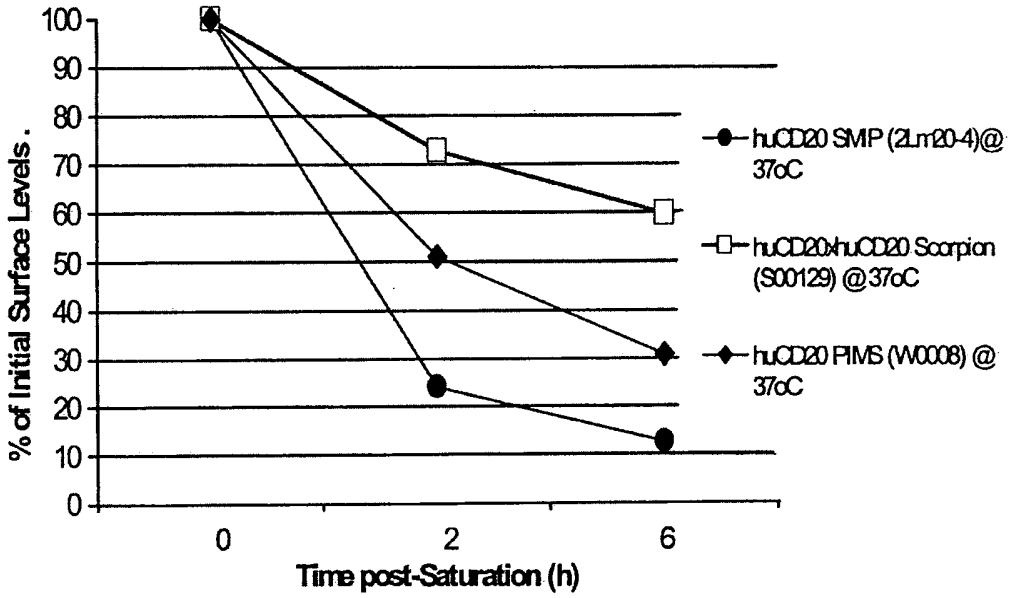

100
90
80
70
60
50
40
30
20
10
0
% of Initial Surface Levels .
0
2
6
Time post-Saturation (h)
huCD20 SMP (2Lm20-4)@ 37oC
huCD20xhuCD20 Scorpion (S00129) @37oC
huCD20 PIMS (W0008) @ 37oC

FIG. 40

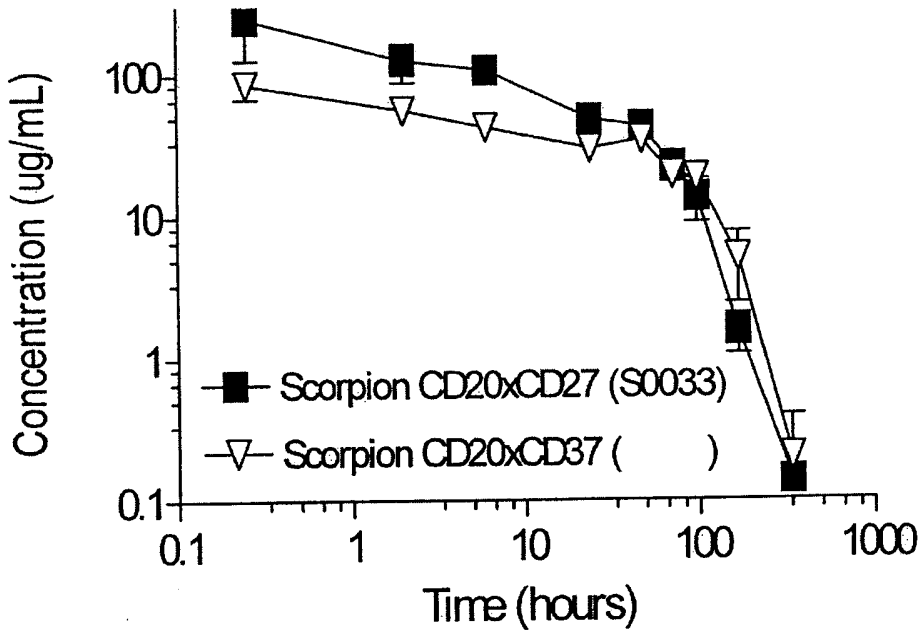

Concentration (ug/mL)
100
10
1
0.1
0.1
1
10
100
1000
Time (hours)
Scorpion CD20xCD27 (S0033)
Scorpion CD20xCD37 ( )

FIG. 41

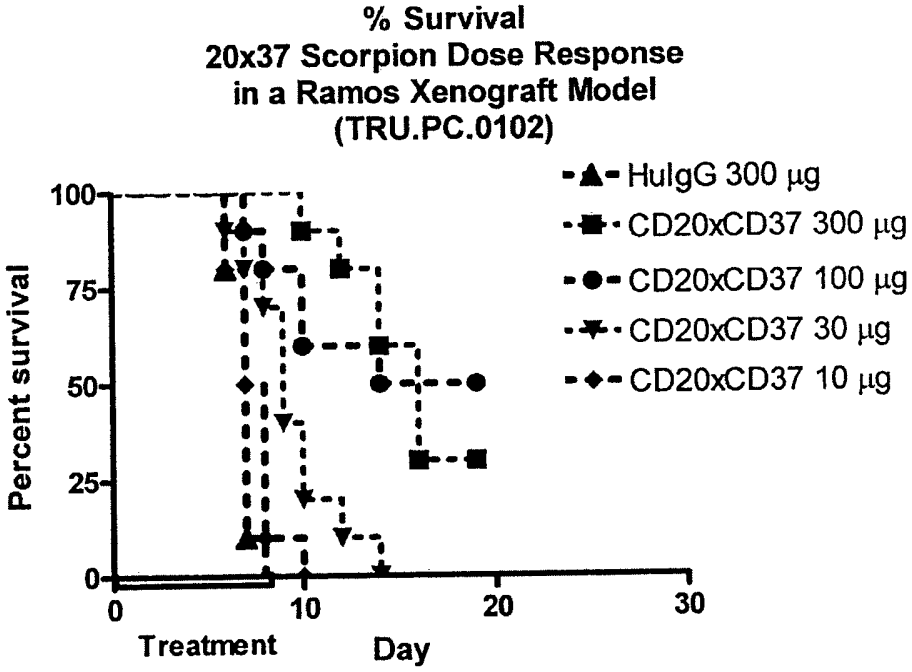
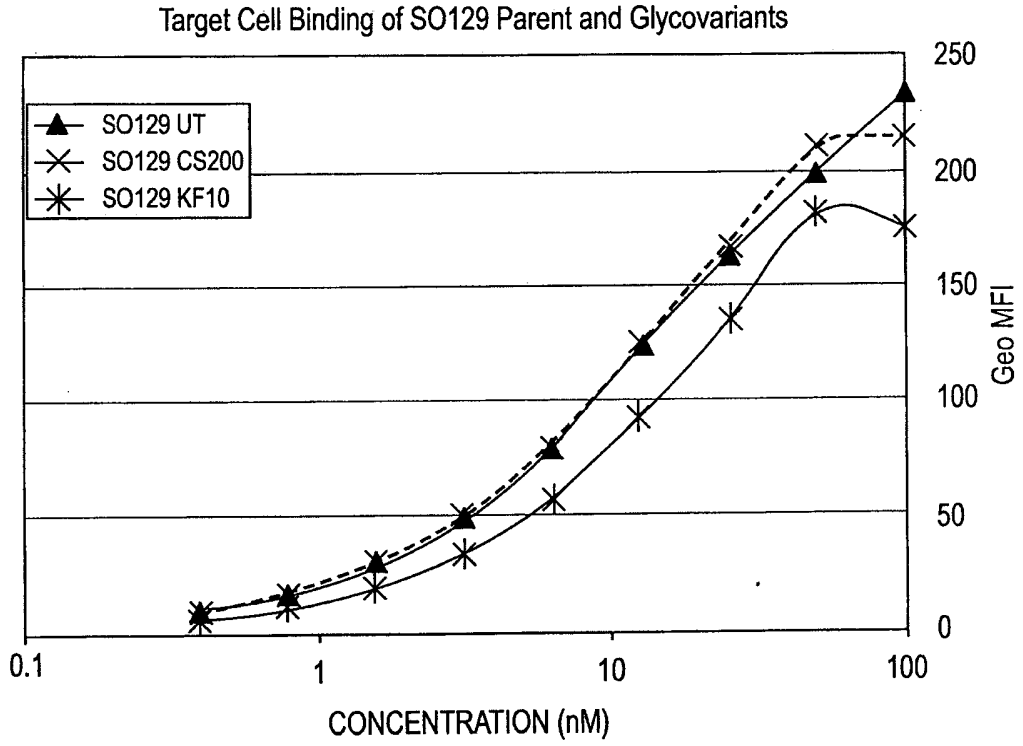

% Survival
20x37 Scorpion Dose Response
in a Ramos Xenograft Model
(TRU.PC.0102)
HulgG 300 µg
CD20xCD37 300 µg
CD20xCD37 100 µg
CD20xCD37 30 µg
CD20xCD37 10 µg
Percent survival
100
75
50
25
0
0
10
20
30
Treatment
Day

FIG. 42



WIL2S staining with 2Im20-4, and SO129 +/- carbohydrate modification
Kifunensine@10 uM (excess), Castanospermine@200 uM

FIG. 43

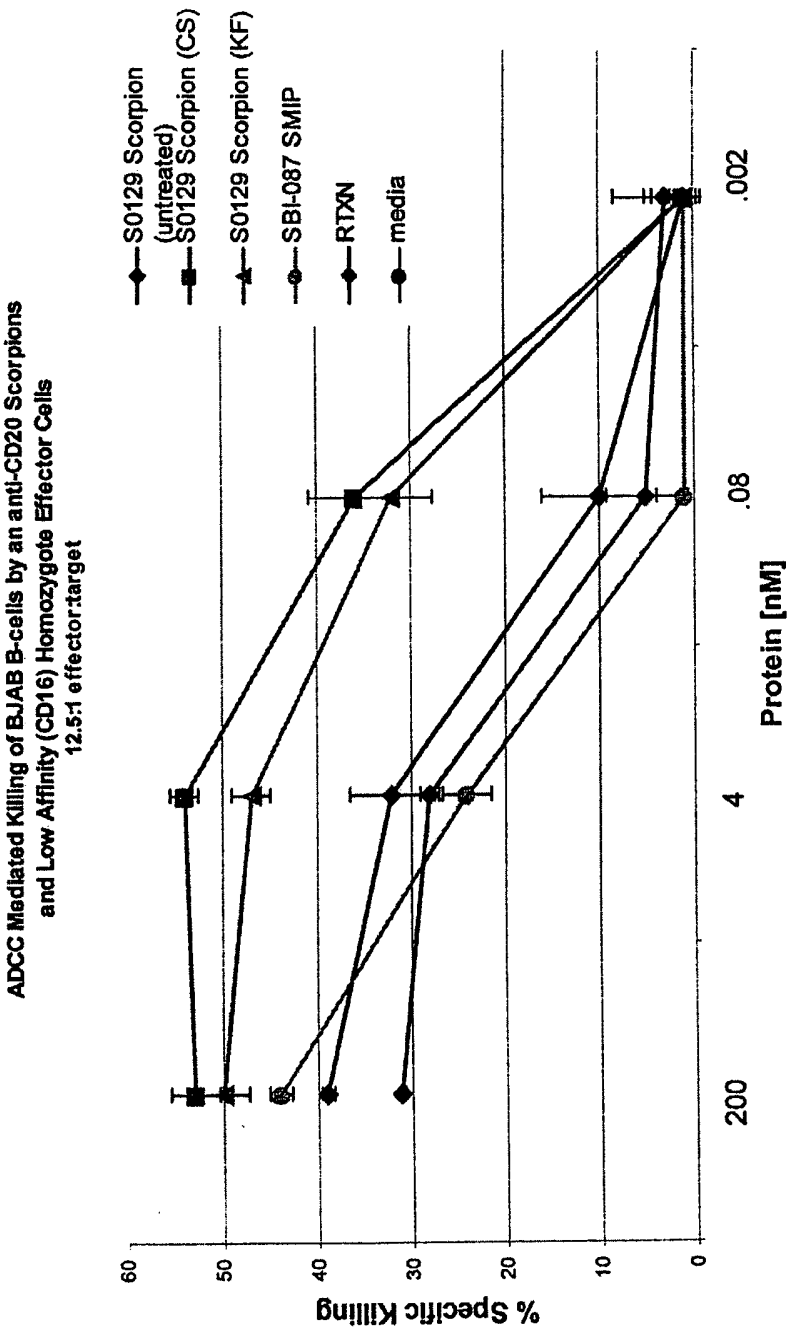
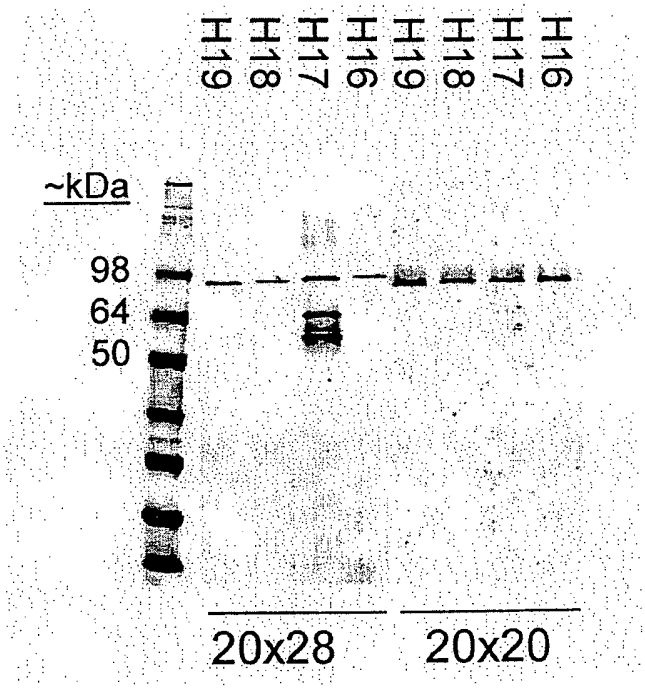

ADCC Mediated Killing of BJAB B-cells by an anti-CD20 Scorpions and Low Affinity (CD16) Homozygote Effector Cells
12.5:1 effector:target
S0129 Scorpion (untreated)
S0129 Scorpion (CS)
S0129 Scorpion (KF)
SBI-087 SMIP
RTXN
media
% Specific Killing
0
10
20
30
40
50
60
200
4
.08
.002
Protein [nM]

FIG. 44



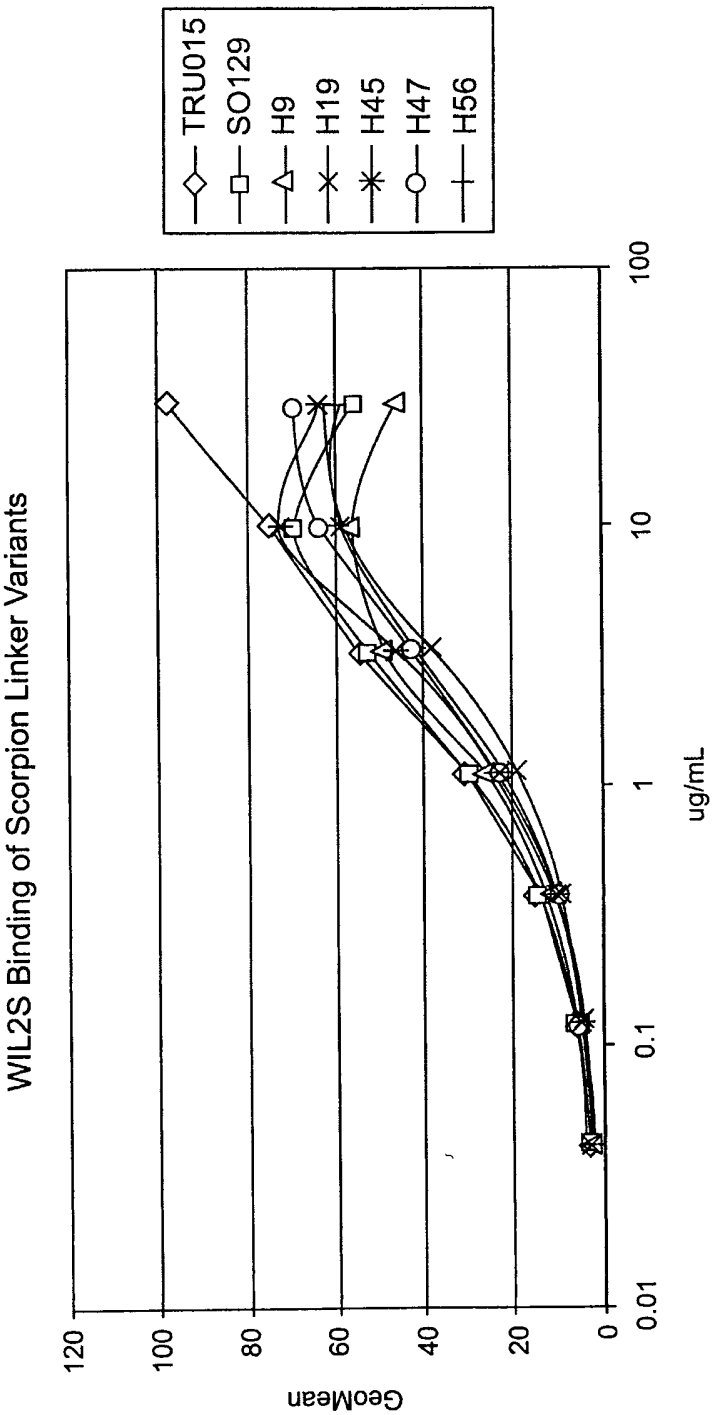

TRU015
SO129
H9
H19
H45
H47
H56
WIL2S Binding of Scorpion Linker Variants
GeoMean
120
100
80
60
40
20
0
0.01
0.1
1
10
100
ug/mL

FIG. 45

FIG. 46

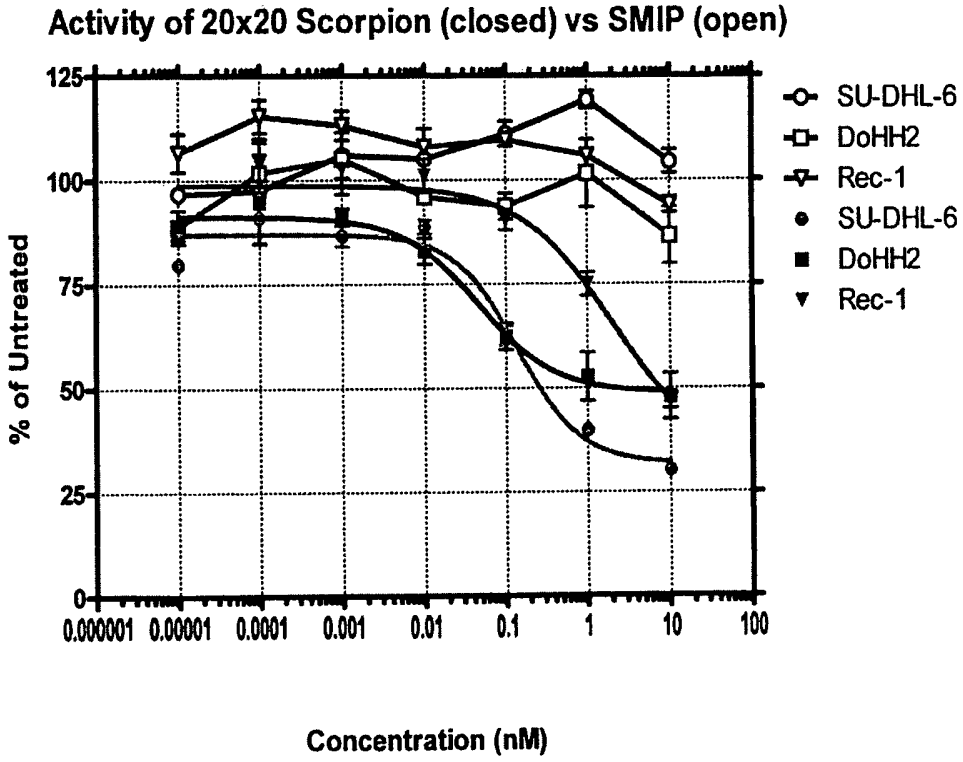

Activity of 20x20 Scorpion (closed) vs SMIP (open)
% of Untreated
125
100
75
50
25
0
0.000001
0.00001
0.0001
0.001
0.01
0.1
1
10
100
Concentration (nM)
SU-DHL-6
DoHH2
Rec-1
SU-DHL-6
DoHH2
Rec-1

FIG. 47

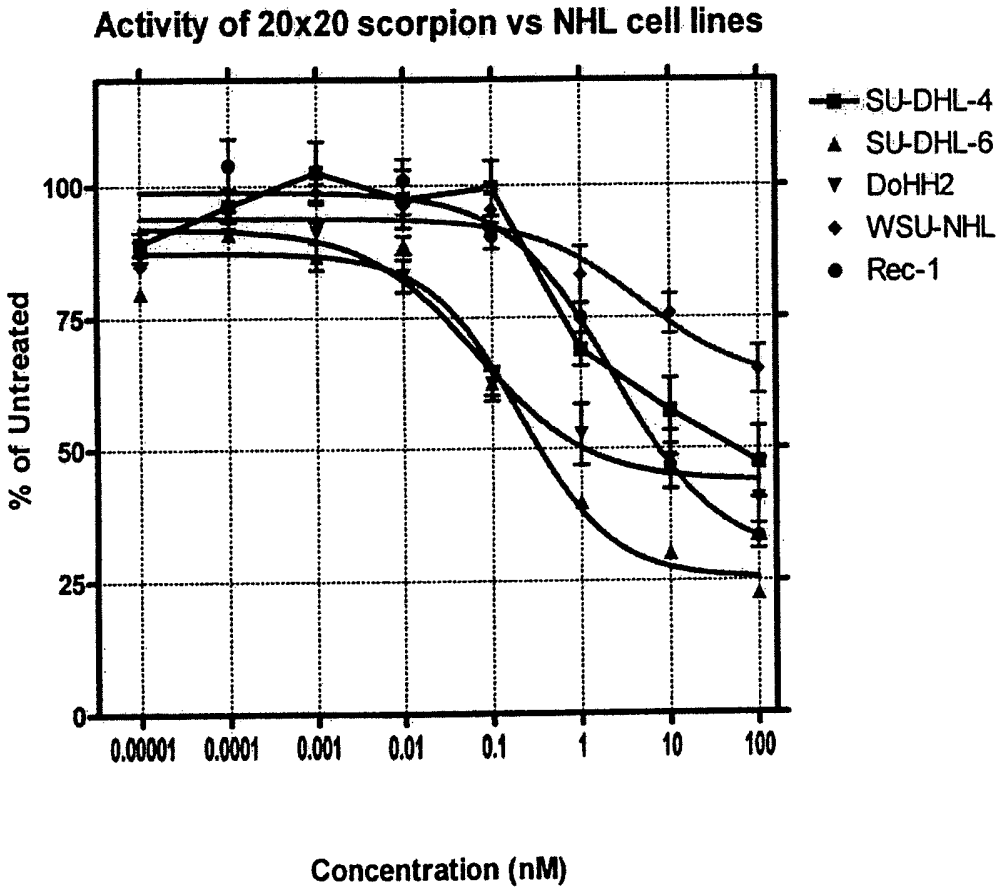

Activity of 20x20 scorpion vs NHL cell lines
% of Untreated
100
75
50
25
0
0.00001
0.0001
0.001
0.01
0.1
1
10
100
Concentration (nM)
SU-DHL-4
SU-DHL-6
DoHH2
WSU-NHL
Rec-1

FIG. 48

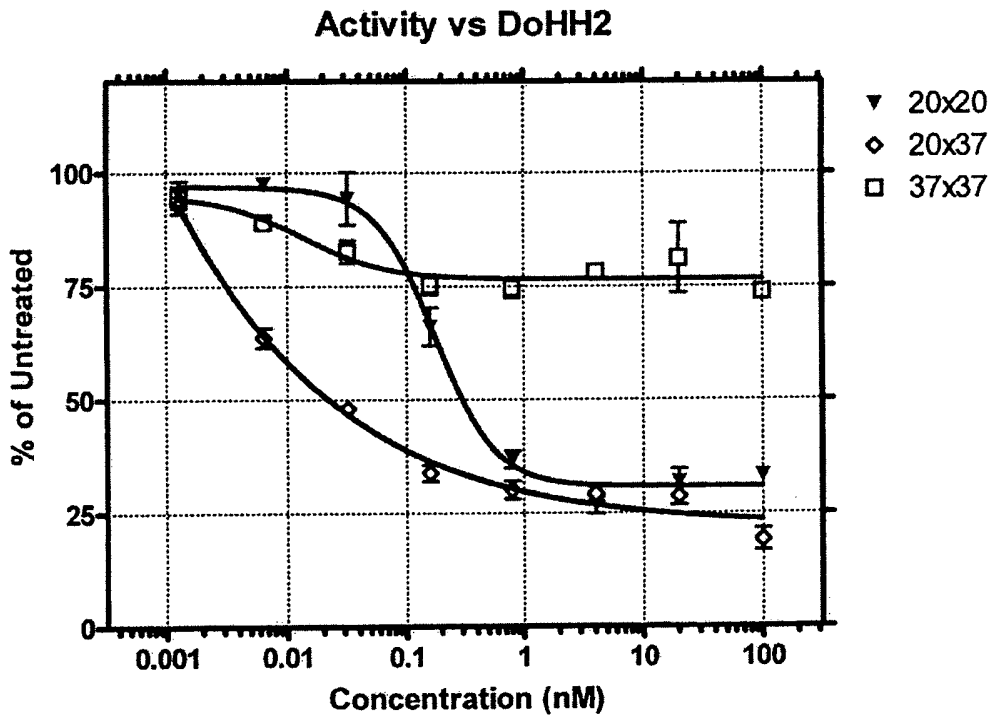

Activity vs DoHH2
% of Untreated
100
75
50
25
0
0.001
0.01
0.1
1
10
100
Concentration (nM)
20x20
20x37
37x37

FIG. 49

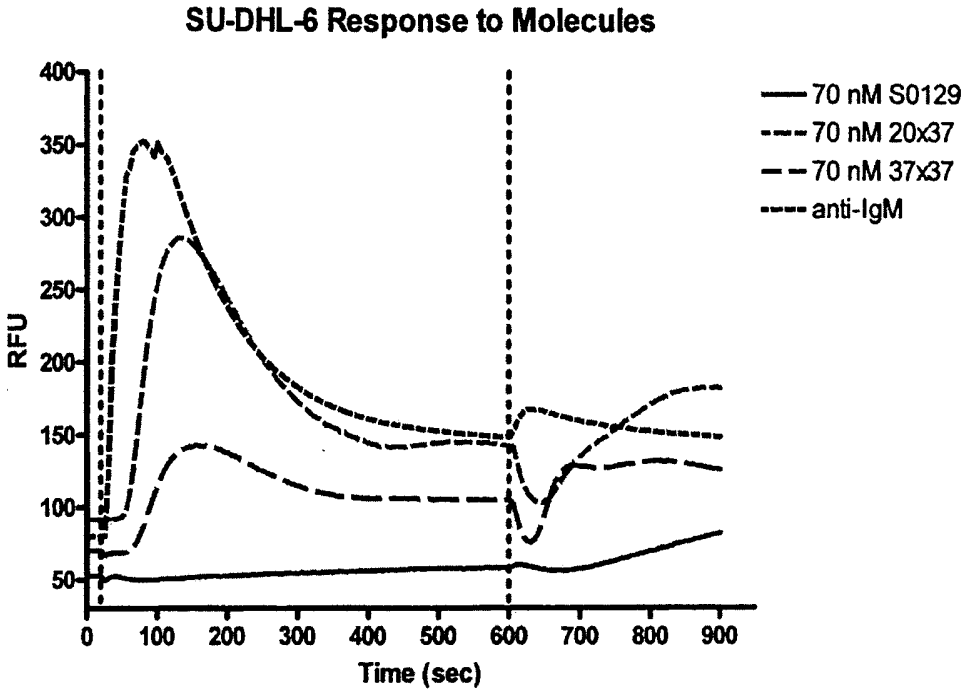

SU-DHL-6 Response to Molecules
RFU
400
350
300
250
200
150
100
50
0
100
200
300
400
500
600
700
800
900
Time (sec)
70 nM S0129
70 nM 20x37
70 nM 37x37
anti-IgM

FIG. 50

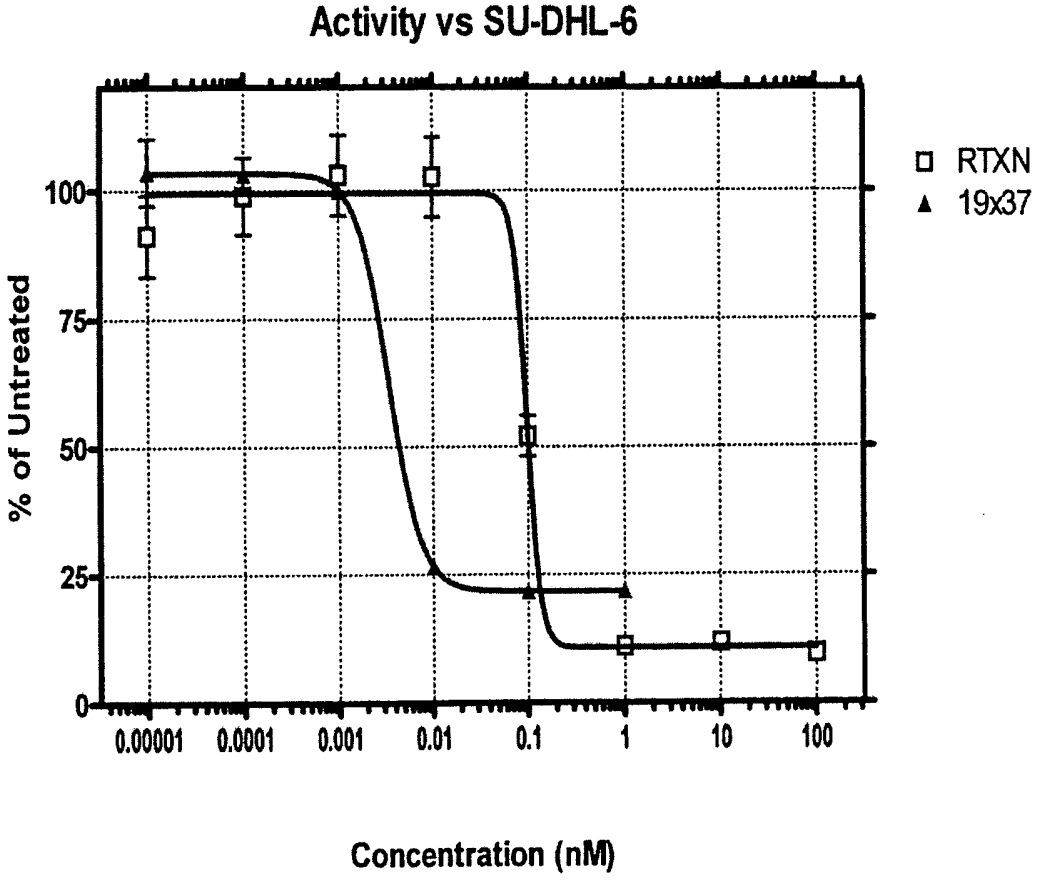

Activity vs SU-DHL-6
% of Untreated
100
75
50
25
0
0.00001
0.0001
0.001
0.01
0.1
1
10
100
Concentration (nM)
RTXN
19x37

FIG. 51

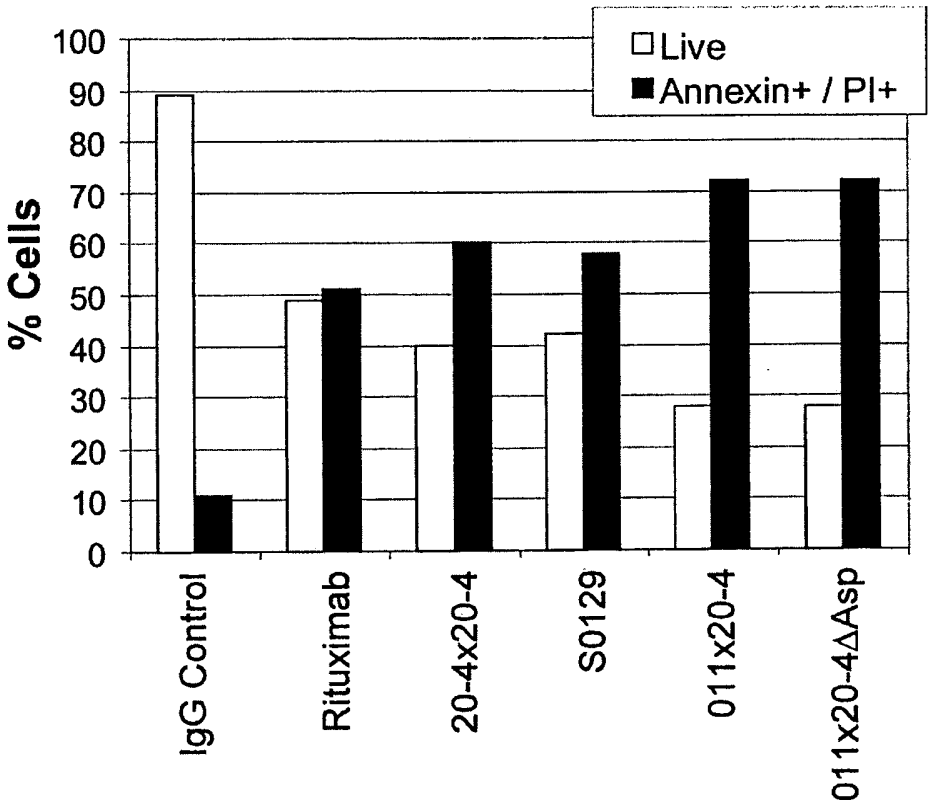

Live
Annexin+ / PI+
% Cells
100
90
80
70
60
50
40
30
20
10
0
IgG Control
Rituximab
20-4x20-4
S0129
011x20-4
011x20-4ΔAsp

FIG. 52

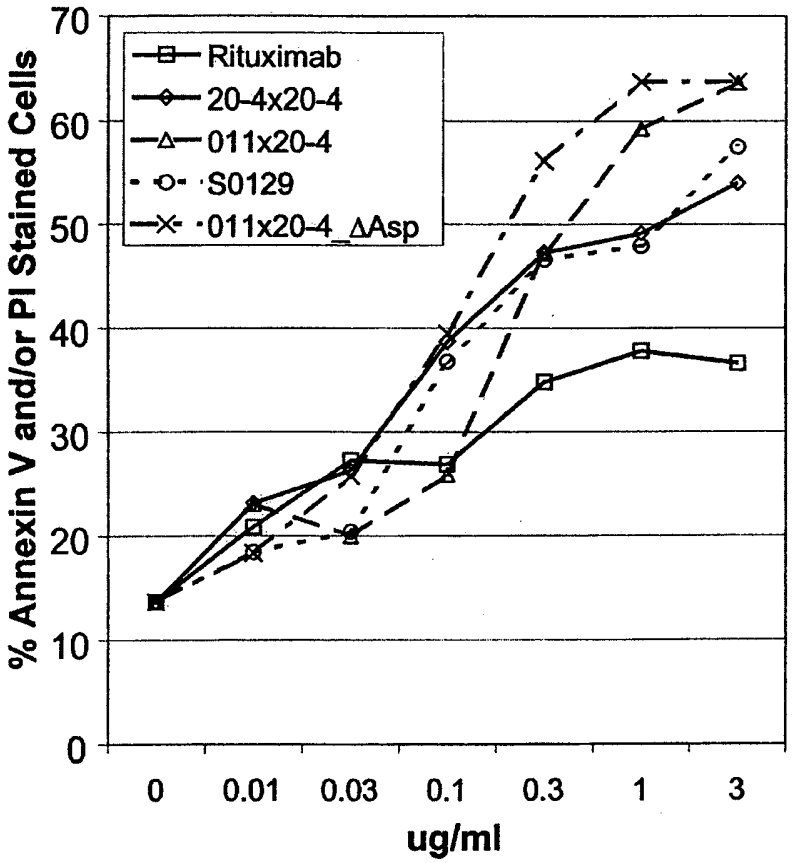

% Annexin V and/or PI Stained Cells
70
60
50
40
30
20
10
0
Rituximab
20-4x20-4
011x20-4
S0129
011x20-4_ΔAsp
0
0.01
0.03
0.1
0.3
1
3
ug/ml

FIG. 53

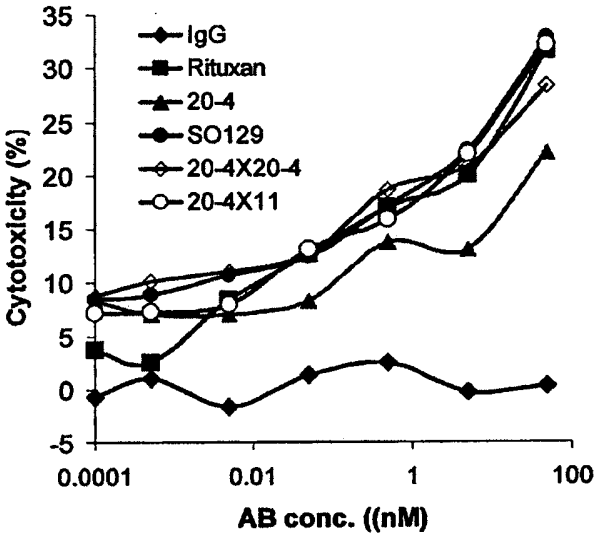

35
30
25
20
15
10
5
0
-5
Cytotoxicity (%)
IgG
Rituxan
20-4
SO129
20-4X20-4
20-4X11
0.0001
0.01
1
100
AB conc. ((nM)

FIG. 54

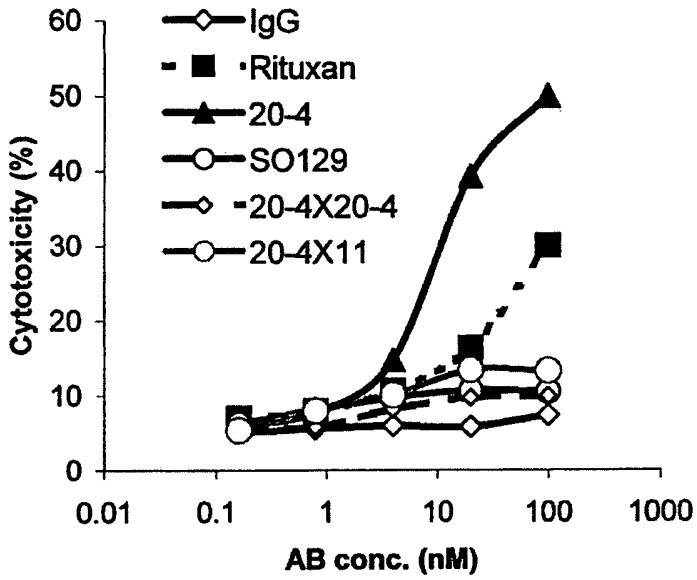

IgG
Rituxan
20-4
SO129
20-4X20-4
20-4X11
Cytotoxicity (%)
60
50
40
30
20
10
0
0.01
0.1
1
10
100
1000
AB conc. (nM)

FIG. 55

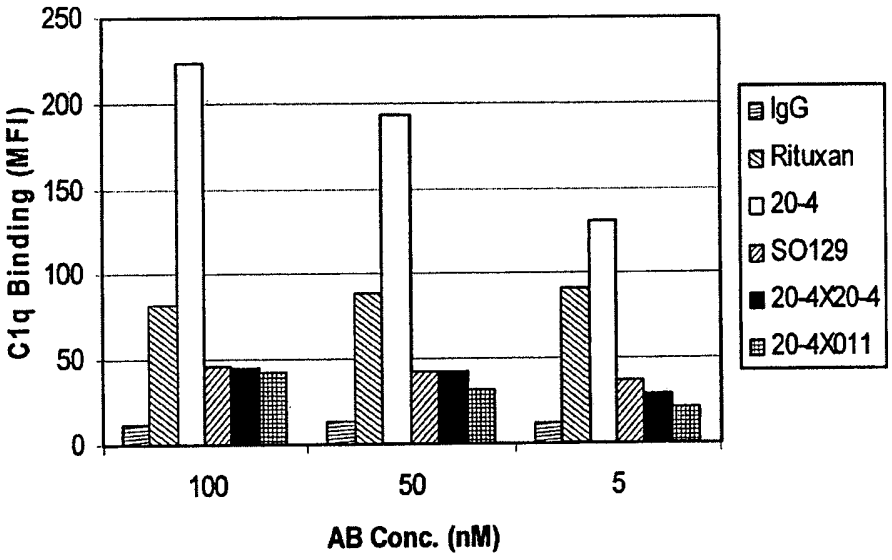

250
200
150
100
50
0
C1q Binding (MFI)
100
50
5
AB Conc. (nM)
IgG
Rituxan
20-4
SO129
20-4X20-4
20-4X011

FIG. 56

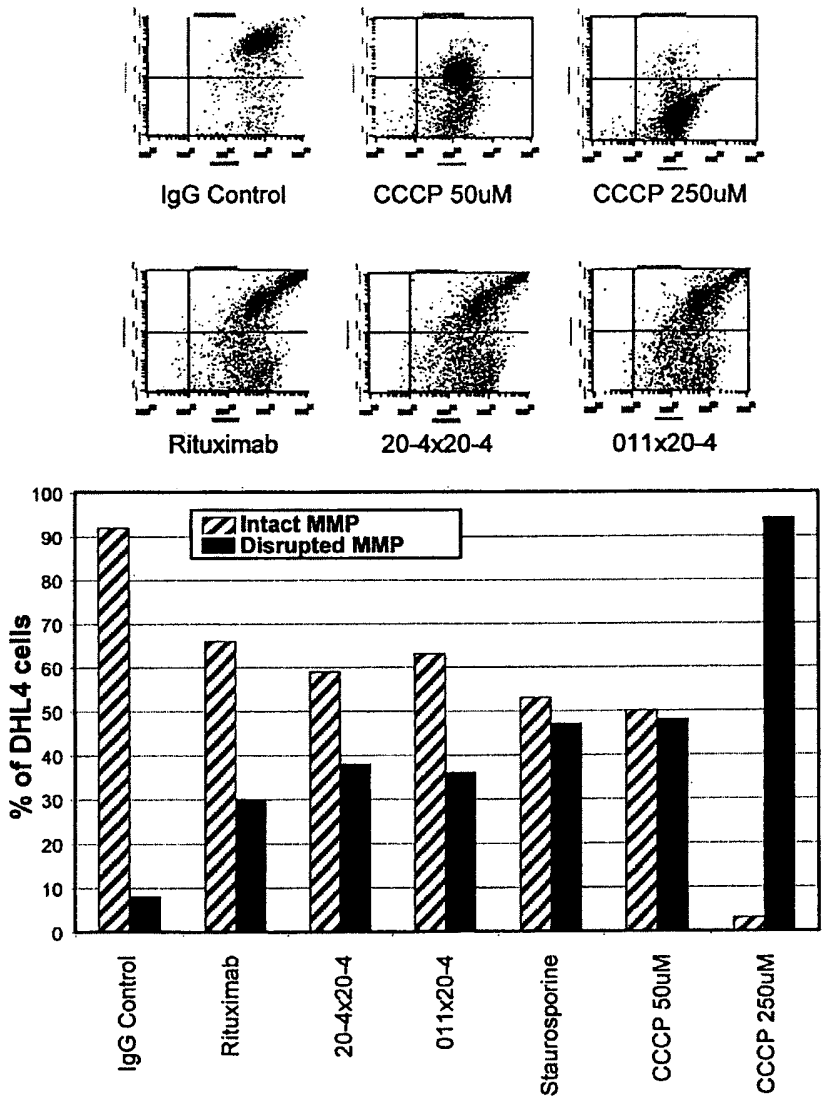

IgG Control
CCCP 50uM
CCCP 250uM
Rituximab
20-4x20-4
011x20-4
Intact MMP
Disrupted MMP
% of DHL4 cells
100
90
80
70
60
50
40
30
20
10
0
IgG Control
Rituximab
20-4x20-4
011x20-4
Staurosporine
CCCP 50uM
CCCP 250uM

FIG. 57

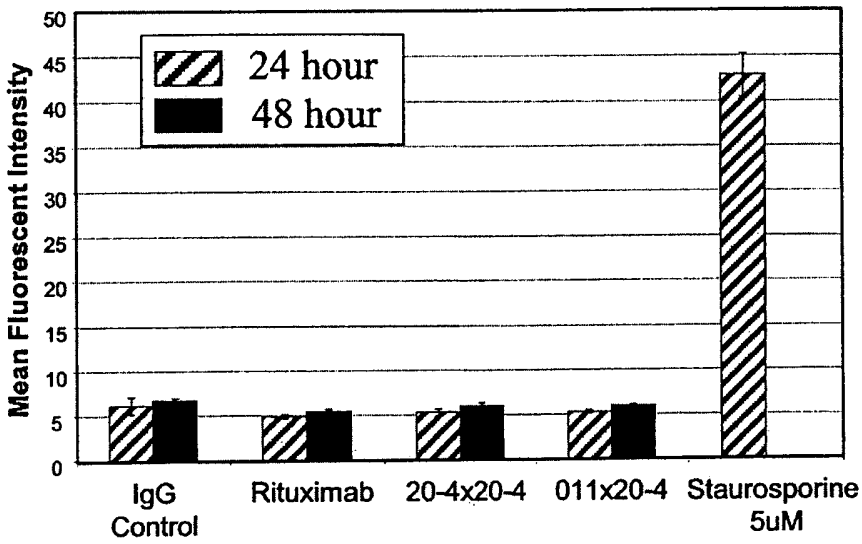
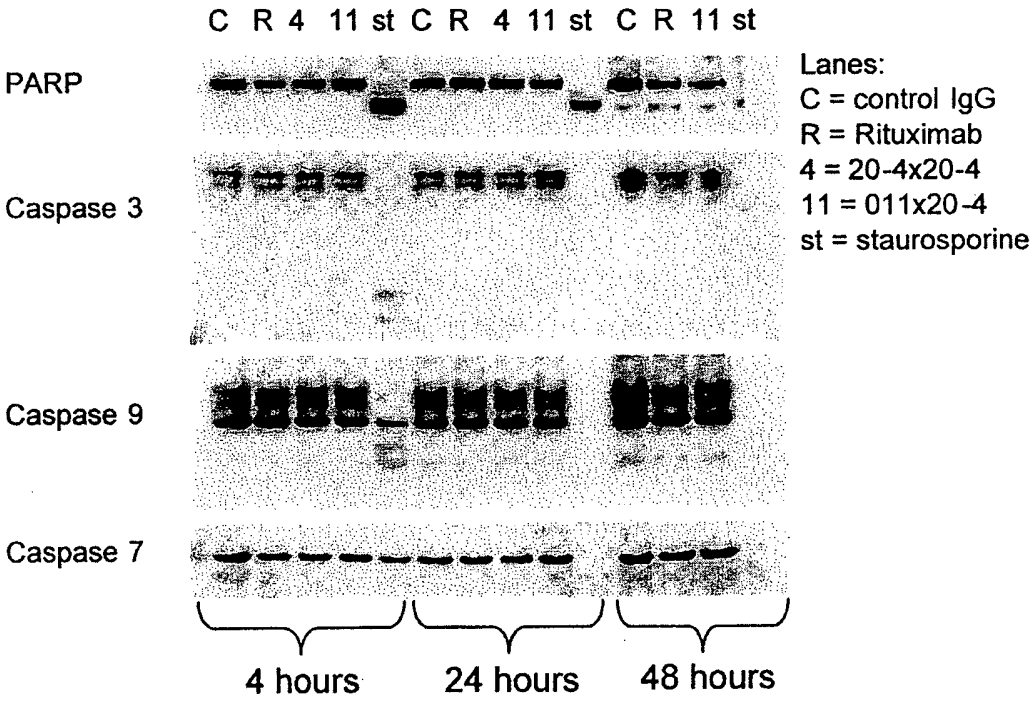

Mean Fluorescent Intensity
50
45
40
35
30
25
20
15
10
5
0
24 hour
48 hour
IgG Control
Rituximab
20-4x20-4
011x20-4
Staurosporine 5uM

FIG. 58


C R 4 11 st C R 4 11 st C R 11 st
PARP
Caspase 3
Caspase 9
Caspase 7
Lanes:
C = control IgG
R = Rituximab
4 = 20-4x20-4
11 = 011x20-4
st = staurosporine
4 hours
24 hours
48 hours

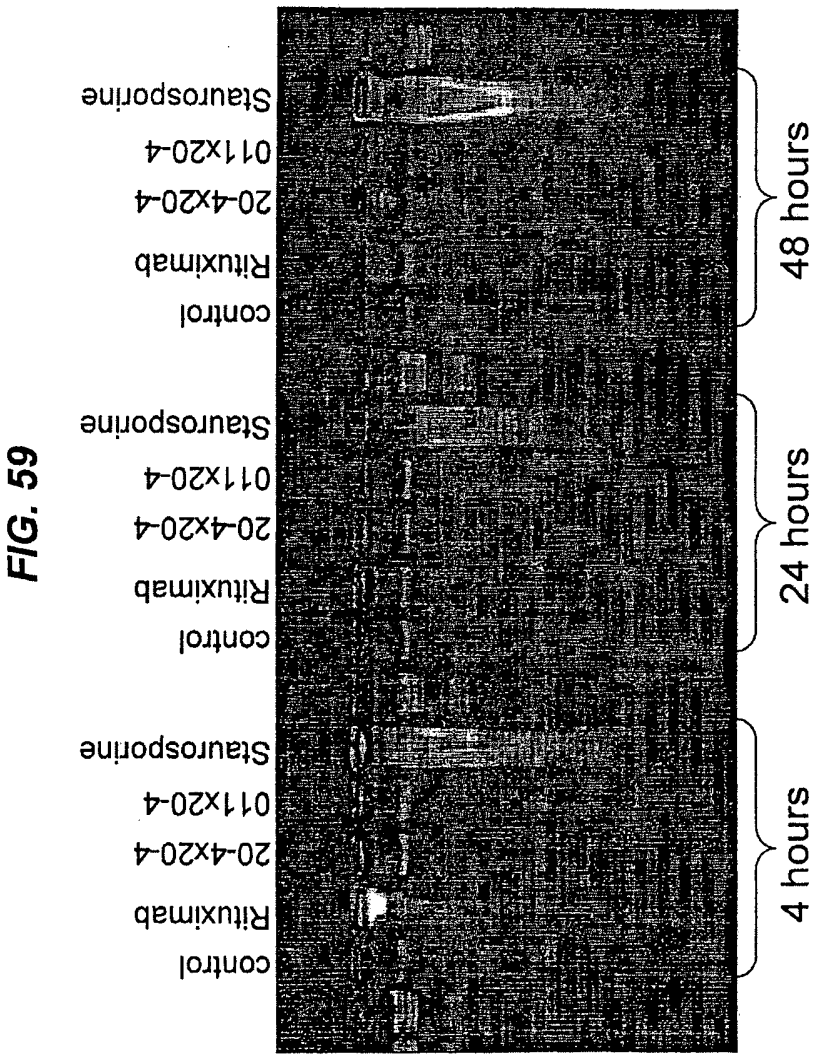
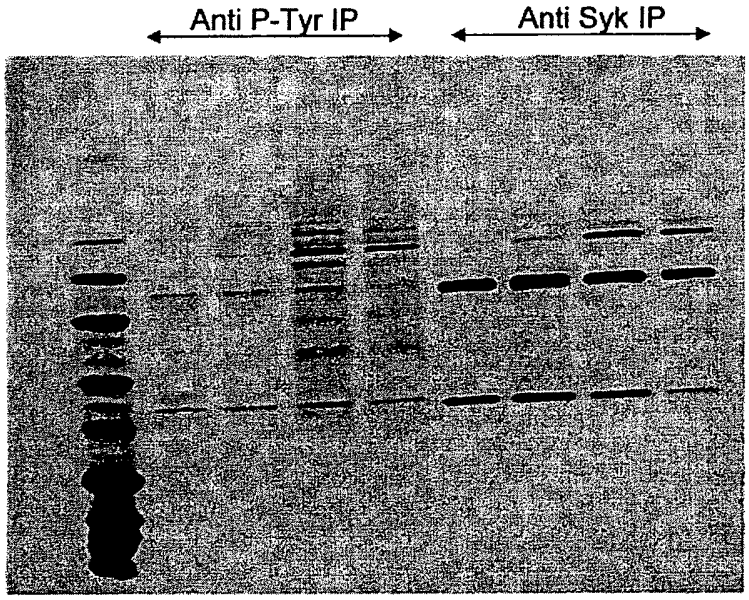

4 hours
48 hours
24 hours
control
Rituximab
20-4x20-4
011x20-4
Staurosporine
control
Rituximab
20-4x20-4
011x20-4
Staurosporine
control
Rituximab
20-4x20-4
011x20-4
Staurosporine

FIG. 59

FIG. 60

DHL6 Cells Treated with 20x37 Scorpion



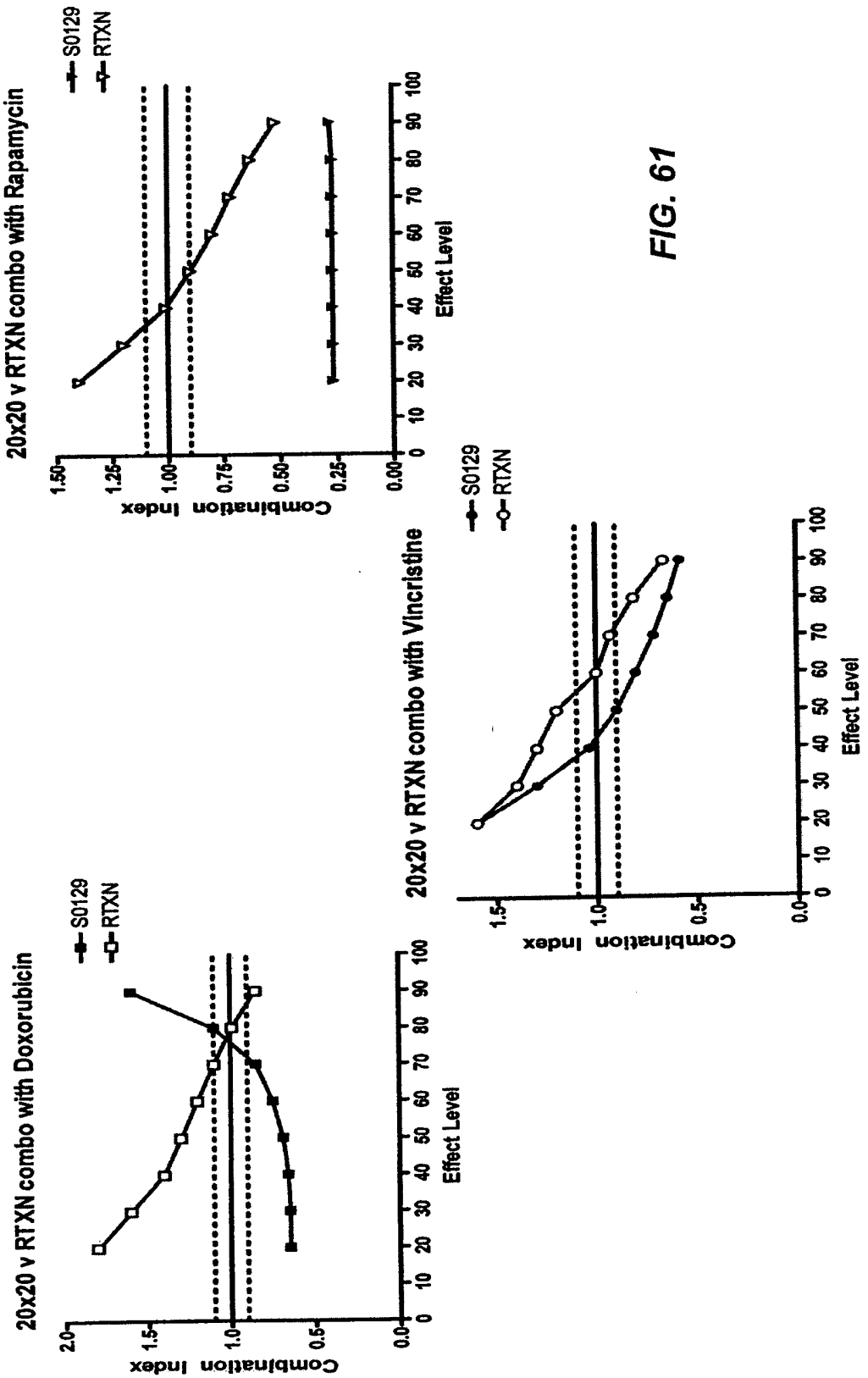

20x20 v RTXN combo with Rapamycin
S0129
RTXN
Combination Index
Effect Level
20x20 v RTXN combo with Vincristine
S0129
RTXN
Combination Index
Effect Level
20x20 v RTXN combo with Doxorubicin
S0129
RTXN
Combination Index
Effect Level

FIG. 61

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 December 2007 (21.12.2007)

PCT

(10) International Publication Number
WO 2007/146968 A3

(51) International Patent Classification:
C07K 16/28 (2006.01) *A61P 35/00* (2006.01)
A61K 39/395 (2006.01) *A61P 31/00* (2006.01)

(21) International Application Number:
PCT/US2007/071052

(22) International Filing Date: 12 June 2007 (12.06.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/813,261 12 June 2006 (12.06.2006) US
60/853,287 20 October 2006 (20.10.2006) US

(71) Applicants *(for all designated States except US)*: **TRUBION PHARMACEUTICALS, INC.** [US/US]; 2401 Fourth Avenue, Suite 1050, Seattle, WA 98121 (US). **Wyeth** [US/US]; 5 Giralda Farms, Madison, New Jersey 07940 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **THOMPSON, Peter, Armstrong** [US/US]; 14075 30th Place N.E., Bellevue, WA 98007 (US). **LEDBETTER, Jeffrey, A.** [US/US]; 18798 Ridgefield Road NW, Shoreline, WA 98177 (US). **HAYDEN-LEDBETTER, Martha, Susan** [US/US]; 18798 Ridgefield Road NW, Shoreline, WA 98178 (US). **GROSMAIRE, Laura, Sue** [US/US]; 19909 280th Avenue, SE, Hobart, WA 98025 (US). **BADER, Robert** [US/US]; 1411 Boylston Avenue, Apt. #314, Seattle, WA 98122 (US). **BRADY, William** [US/US]; 618 219th Place SW, Bothell, WA 98021 (US). **TCHISTIAKOVA, Lioudmila** [CA/US]; 19 Abbot Bridge Drive, Andover, MA 01810 (US). **FOLLETTIE, Maximillian, T.** [US/US]; 12 Greenbrook Way, Belmont, MA 02478 (US). **CALABRO, Valerie** [FR/US]; 195 Binney Street #1205, Cambridge, MA 02142 (US). **SCHULER, Alwin** [NL/US]; 2 Centennial Street, Hopedale, MA 01747 (US).

(74) Agent: **MERKEL, William, K.**; Marshall, Gerstein & Borun LLP, 233 S. Wacker Drive, Suite 6300, Sears Tower, Chicago, IL 60606-6357 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
19 June 2008

(54) Title: SINGLE-CHAIN MULTIVALENT BINDING PROTEINS WITH EFFECTOR FUNCTION



(57) Abstract: Multivalent binding peptides, including bi-specific binding peptides, having immunoglobulin effector function are provided, along with encoding nucleic acids, vectors and host cells as well as methods for making such peptides and methods for using such peptides to treat or prevent a variety of diseases, disorders or conditions, as well as to ameliorate at least one symptom associated with such a disease, disorder or condition.

WO 2007/146968 A3

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/071052

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K16/28 A61K39/395 A61P35/00 A61P31/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K A61K A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2004/076489 A (MEDINNOVA AS [NO]; BOGEN BJARNE [NO]; FREDRIKSEN AGNETE BRUNSVIK [NO];) 10 September 2004 (2004-09-10) page 17 - page 29 claims 78-81 page 13, line 3 - line 14	1-3, 49-52, 56,59-71
Y		88

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 January 2008	17/04/2008

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Bumb, Peter

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/071052

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	COLOMA M J ET AL: "DESIGN AND PRODUCTION OF NOVEL TETRAVALENT BISPECIFIC ANTIBODIES" NATURE BIOTECHNOLOGY, NATURE PUBLISHING GROUP, NEW YORK, NY, US, vol. 15, no. 2, 1 February 1997 (1997-02-01), pages 159-163, XP000647731 ISSN: 1087-0156 the whole document	1-3, 49-52, 56, 59-67,71
Y		88
X	US 2006/104971 A1 (GARBER ELLEN [US] ET AL) 18 May 2006 (2006-05-18) examples 1-6,9 figures 9,10 paragraphs [0076] - [0078]	1-3, 49-52, 56, 59-67,71
Y		88
X	US 2005/158829 A1 (FANDL JAMES P [US] ET AL) 21 July 2005 (2005-07-21) examples 3,4	1-3, 49-52, 56, 59-67,71
Y		88
X	WO 03/074569 A (IMMUNOMEDICS INC [US]; MCCALL JOHN DOUGLAS [GB]) 12 September 2003 (2003-09-12) page 40 - page 46	1-3, 49-52, 56, 59-67,71
Y		88
X	WO 2005/000899 A (BIOGEN IDEC INC [US]; GLASER SCOTT [US]; REFF MITCHELL [US]; YANG TZUN) 6 January 2005 (2005-01-06) figures 4,7 examples 8-10 page 47, lines 5-7	1-3, 49-52, 56, 59-67,71
Y		88
A	WO 2005/017148 A (TRUBION PHARMACEUTICALS INC [US]; LEDBETTER JEFFREY A [US]; HAYDEN-LED) 24 February 2005 (2005-02-24) the whole document	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2007/071052

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 60-70, 86, 88 (of invention 1)
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers allsearchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search reportcovers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1, 49 - 52, 56, 59 - 71, 86, 88 (in part), 2 - 3 (complete)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

[ ] The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.1

Although claims 60-70, 86, 88 (of invention 1) are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box II.1

Claims Nos.: 60-70, 86, 88 (of invention 1)

Although claims 60-70, 86, 88 (of invention 1) are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 2-3 (complete)

   AB1-CHX-L2-AB2, wherein AB1 comprises VH+VL of an immunoglobulin; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

   ---

2. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 4-5 (complete)

   AB1-CHX-L2-AB2, wherein AB2 comprises VH+VL of an immunoglobulin; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

   ---

3. claims: 1, 9, 49-52, 56, 59-71, 86, 88 (in part), 6-7 (complete)

   AB1-CHX-L2-AB2, wherein AB1 and AB2 recognize the same molecular target or epitope, i.e. on the same cell or on two distinct cells; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

   ---

4. claims: 1, 9, 49-52, 56, 59-71, 86, 88 (in part), 8, 10 (complete)

   AB1-CHX-L2-AB2, wherein AB1 and AB2 recognize two different molecular targets on the same or different cells/virii/carriers/objects; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

   ---

5. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 11 (complete)

   AB1-CHX-L2-AB2, wherein AB1, CHX and AB2 are derived from a human Ig; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

   ---

6. claims: 1, 12-13, 28-30, 49-52, 56, 59-71, 86, 88 (in part), 14 (complete)

   AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a tumour antigen, i.e. those mentioned in claim 14, or i.e. Prostate Specific Membrane Antigen mentioned in claim 28; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

   ---

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

7. claims: 1, 12-13, 29-30, 32, 49-52, 56-57, 59-71, 86, 88 (in part), 15, 31, 33, 72-83, 87, 89
(complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a B-cell target, i.e. those mentioned in claims 15, 31, 72, or i.e. wherein AB1 or AB2 comprise anti-CD20 or anti-CD37 amino acid sequences SEQ ID Nos: 2, 103, 105, 332-345, 355-365; related nucleotides, vectors, host cells and methods of production, i.e. comprising SEQ ID NO: 1; methods of treatment making use thereof

---

8. claims: 1, 12-13, 49-52, 56, 59-71, 86, 88 (in part), 16
(complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a TNF-R superfamily member, i.e. those mentioned in claim 16; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

9. claims: 1, 12-13, 49-52, 56, 59-71, 86, 88 (in part), 17
(complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a Hedgehog family member, i.e. those mentioned in claim 17; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

10. claims: 1, 12-13, 28, 49-52, 56, 59-71, 86, 88 (in part), 18
(complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a receptor tyrosine kinase member, i.e. those mentioned in claim 18; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

11. claims: 1, 12-13, 28, 49-52, 56, 59-71, 86, 88 (in part), 19
(complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a proteoglycan related molecule, i.e. those mentioned in claim 19 or Receptor for Advanced Glycation End products mentioned in claim 28; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

12. claims: 1, 12-13, 49-52, 56, 59-71, 86, 88 (in part), 20
(complete)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a TGF-superfamily member, i.e. those mentioned in claim 20; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

13. claims: 1, 12-13, 49-52, 56, 59-71, 86, 88 (in part), 21 (complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a Wnt-related molecule, i.e. those mentioned in claim 21; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

14. claims: 1, 12-13, 49-52, 56, 59-71, 86, 88 (in part), 22 (complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a receptor ligand, i.e. those mentioned in claim 22 or IL-17A, IL-17F, P19, Dickkopf-1, IL-21, IL-4, Myostatin mentioned in claim 28; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

15. claims: 1, 12-13, 29-30, 32, 49-52, 57, 59-71, 86, 88 (in part), 23 (complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize a T cell target, i.e. the TCR mentioned in claim 23 or the anti-CD3 or anti-CD28 amino acid sequences SEQ ID ID Nos: 4, 6, 107, 109; related nucleotides, vectors, host cells and methods of production, i.e. comprising SEQ ID NO: 3 or 5; methods of treatment making use thereof

---

16. claims: 1, 12-13, 29-30, 49-52, 56, 59-71, 86, 88 (in part), 24 (complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize an NK cell target, i.e.those mentioned in claim 24; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

17. claims: 1, 12-13, 49-52, 56, 59-71, 86, 88 (in part), 25 (complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize an monocyte/macrophage cell receptor, i.e. those mentioned in claim 25; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

18. claims: 1, 12-13, 29-30, 49-52, 56, 59-71, 86, 88 (in part), 26 (complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize an Dendritic cell receptor, i.e. those mentioned in claim 26; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

19. claims: 1, 12-13, 49-52, 56, 59-71, 86, 88 (in part), 27 (complete)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize an angiogenesis target, i.e. those mentioned in claim 27; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

20. claims: 1, 28, 49-52, 56, 59-71, 86, 88 (in part)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize NOTCH1; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

21. claims: 1, 28, 49-52, 56, 59-71, 86, 88 (in part)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize IgHE; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

22. claims: 1, 28, 49-52, 56, 59-71, 86, 88 (in part)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize IL-22R; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

23. claims: 1, 28, 49-52, 56, 59-71, 86, 88 (in part)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize Amyloid oligomer or Amyloid Precursor Protein; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

24. claims: 1, 28, 49-52, 56, 59-71, 86, 88 (in part)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize NOGO receptor; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

25. claims: 1, 28, 49-52, 56, 59-71, 86, 88 (in part)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize Low density Lipoprotein Receptor - Related Protein 5; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

26. claims: 1, 28, 49-52, 56, 59-71, 86, 88 (in part)

AB1-CHX-L2-AB2, wherein AB1 or AB2 recognize Very Late Antigen 4; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

27. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 34-36 (complete)

AB1-CHX-L2-AB2, wherein CHX recognizes an Fc receptor, CD16, CD32a, CD64, CD89, Fc R1, FcRn or pIgR or has a CH2 + CH3 domain; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

28. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 37 (complete)

AB1-CHX-L2-AB2, comprising CH2 and truncated CH3, wherein CH3 ends in SEQ ID Nos: 366-371; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

29. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 38-40 (complete)

AB1-CHX-L2-AB2, wherein L2 is derived from an immunoglobulin hinge, i.e. comprising SEQ ID Nos: 351-354 (see Table 10); related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

30. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 41 (complete)

AB1-CHX-L2-AB2, wherein L2 is derived from a C-lectin stalk, i.e. comprising SEQ ID Nos: 373-377 (see Table 13); related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

31. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 44, 58 (complete)

AB1-CHX-L2-AB2, wherein a linker L1 >5aa is inserted between AB1 and CHX; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

---

32. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 45, 46 (complete)

AB1-CHX-L2-AB2, wherein AB1 and AB2 have a binding affinity less than 10 -9 M or greater than 10 -6 M; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

33. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 47 (complete)

AB1-CHX-L2-AB2, wherein Chx has CDC- or ADCC-function; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

34. claims: 1, 49-52, 56, 59-71, 86, 88 (in part), 48 (complete)

AB1-CHX-L2-AB2, which has an in vivo half-life of >28 hours in a human; related nucleotides, vectors, host cells and methods of production; methods of treatment making use thereof

---

35. claims: Claims 53-55 (complete)

A method for producing a protein AB1-CHX-L2-AB2, by expressing it from a host-cell at a level of > 1 mg/ml

---

36. claims: Claims 84-85 (complete)

A method for identifying AB1 or AB2 comprising, contacting at least one antigen with two antibodies having two different specificities, contacting said two antibodies with an anti-isotye secondary antibody and measuring activity of the target-antigen

---

37. claims: Claims 90-98 (complete)

Compositions comprising and methods using a plurality of multivalent single-chain binding proteins.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2007/071052

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2004076489 A	10-09-2004	AU 2004215489 A1	10-09-2004
		CA 2517054 A1	10-09-2004
		EP 1599504 A1	30-11-2005
		US 2004253238 A1	16-12-2004
US 2006104971 A1	18-05-2006	NONE	
US 2005158829 A1	21-07-2005	NONE	
WO 03074569 A	12-09-2003	AU 2003209446 A1	16-09-2003
		CA 2478011 A1	12-09-2003
		EP 1487879 A2	22-12-2004
		JP 2006502091 T	19-01-2006
WO 2005000899 A	06-01-2005	AU 2004252170 A1	06-01-2005
		AU 2004252171 A1	06-01-2005
		CA 2529945 A1	06-01-2005
		CA 2530388 A1	06-01-2005
		EP 1641826 A2	05-04-2006
		EP 1641827 A2	05-04-2006
		JP 2007535296 T	06-12-2007
		JP 2007535297 T	06-12-2007
		WO 2005000898 A2	06-01-2005
WO 2005017148 A	24-02-2005	AU 2003300092 A1	07-03-2005
		BR PI0318417 A	25-07-2006
		CA 2533921 A1	24-02-2005
		CN 1852976 A	25-10-2006
		EP 1654358 A1	10-05-2006
		HR 20060074 A2	31-08-2006
		IS 8305 A	16-02-2006
		JP 2007528194 T	11-10-2007
		KR 20060070530 A	23-06-2006
		US 2007237779 A1	11-10-2007

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
C07K 16/28 (2006.01)
A61K 39/395 (2006.01)
A61P 35/00 (2006.01)
A61P 31/00 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200780025783.7

[43] 公开日 2009 年 7 月 22 日

[11] 公开号 CN 101490085A

[22] 申请日 2007.6.12
[21] 申请号 200780025783.7
[30] 优先权
[32] 2006. 6.12 [33] US [31] 60/813,261
[32] 2006.10.20 [33] US [31] 60/853,287
[86] 国际申请 PCT/US2007/071052 2007.6.12
[87] 国际公布 WO2007/146968 英 2007.12.21
[85] 进入国家阶段日期 2009.1.7
[71] 申请人 特鲁比昂药品公司
地址 美国华盛顿州
共同申请人 惠氏公司
[72] 发明人 彼得·阿姆斯壮·汤普森
杰弗里·A·莱德佰特
马萨·苏珊·海登 - 莱德佰特
劳拉·休·格劳斯麦瑞
罗伯特·贝德尔 威廉·布莱蒂
里奥德米拉·奇斯蒂阿阔娃
马克西米立安·T·弗拉蒂耶
瓦莱丽·卡拉布罗 阿尔文·舒勒
[74] 专利代理机构 北京英赛嘉华知识产权代理有限责任公司
代理人 王达佐 崔建丽
权利要求书 19 页 说明书 428 页 附图 67 页

[54] 发明名称

具有效应功能的单链多价结合蛋白

[57] 摘要

本发明提供具有免疫球蛋白效应功能的多价结合肽(包括双特异性结合肽)，以及编码核酸、载体和宿主细胞，以及制造所述肽的方法，和使用所述肽治疗或预防各种疾病、病症或状态，以及改善至少一种与该疾病、病症或状态相关的症状的方法。



1. 具有效应功能的多价单链结合蛋白，其包含:

a. 第一结合结构域，其源自免疫球蛋白或免疫球蛋白样分子;

b. 提供效应功能的恒定亚区，该免疫球蛋白恒定亚区位于所述第一结合结构域的 C-端;

c. 蝎形连接子，其位于所述恒定亚区的 C-端；及

d. 第二结合结构域，其源自免疫球蛋白或免疫球蛋白样分子，并位于所述恒定亚区的 C-端;

因此，所述恒定亚区定位于所述第一结合结构域与所述第二结合结构域之间。

2. 如权利要求 1 所述的蛋白，其中所述第一结合结构域包含源自第一免疫球蛋白的轻链可变区和源自第二免疫球蛋白的重链可变区。

3. 如权利要求 2 所述的蛋白，其中所述第一免疫球蛋白与所述第二免疫球蛋白为相同的免疫球蛋白。

4. 如权利要求 1 所述的蛋白，其中所述第二结合结构域包含源自第一免疫球蛋白的轻链可变区和源自第二免疫球蛋白的重链可变区。

5. 如权利要求 4 所述的蛋白，其中所述第一免疫球蛋白与所述第二免疫球蛋白为相同的免疫球蛋白。

6. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域识别相同的分子靶标。

7. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域识别相同的表位。

8. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域识别相同的真核细胞、原核细胞、病毒、载体或物体上的不同分子靶标。

9. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域识别与实体不同的真核细胞、原核细胞、病毒、载体或物体相关的分子靶标。

10. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中的至少一者识别至少一种分子靶标，所述分子靶标与真核细胞、原核细胞、病毒、载体或物体不相关。

11. 如权利要求 1 所述的蛋白，其中所述第一结合结构域、所述第二结合结构域和所述恒定亚区中的每个源自人类免疫球蛋白。

12. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中的至少一者识别癌细胞上的靶标。

13. 如权利要求 11 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中的至少一者识别选自由以下组成的组的靶标：肿瘤抗原、B-细胞靶标、TNF 受体超家族成员、Hedgehog 家族成员、受体酪氨酸激酶、蛋白聚糖相关分子、TGF-β 超家族成员、Wnt 相关分子、受体配体、T-细胞靶标、树突状细胞靶标、NK 细胞靶标、单核细胞/巨噬细胞靶标和血管生成靶标。

14. 如权利要求 13 所述的蛋白，其中所述肿瘤抗原是选自由以下组成的组：鳞状细胞癌抗原 1、鳞状细胞癌抗原 2、卵巢癌抗原 CA125、粘蛋白 1、CTCL 肿瘤抗原 se1-1、CTCL 肿瘤抗原 se14-3、CTCL 肿瘤抗原 se20-4、CTCL 肿瘤抗原 se20- 9、CTCL 肿瘤抗原 se33-1、CTCL 肿瘤抗原 se37-2、CTCL 肿瘤抗原 se57-1、CTCL 肿瘤抗原 se89-1、前

列腺特异性膜抗原、5T4 癌胚滋养层糖蛋白、Orf73 卡波西氏肉瘤相关疱疹病毒、MAGE-C1、MAGE-B1 抗原、MAGE-B2 抗原、MAGE-2 抗原、MAGE-4a 抗原、MAGE-4b 抗原、结肠癌抗原 NY-CO-45、肺癌抗原 NY-LU-12 变异体 A、癌症相关表面抗原、腺癌抗原 ART1、副肿瘤相关脑-睾丸癌抗原、神经肿瘤腹侧抗原 2、肝细胞癌抗原基因 520、肿瘤相关抗原 CO-029、肿瘤相关抗原 MAGE-X2、滑膜肉瘤 X 断裂点 2、T 细胞所识别的鳞状细胞癌抗原、血清学定义的结肠癌抗原 1、血清学定义的乳癌抗原 NY-BR-15、血清学定义的乳癌抗原 NY-BR-16、嗜铬粒蛋白 A; 副甲状腺分泌蛋白 1、DUPAN-2、CA 19-9、CA 72-4、CA 195 和 L6。

15. 如权利要求 13 所述的蛋白，其中所述 B 细胞靶标是选自由以下组成的组：CD10、CD19、CD20、CD21、CD22、CD23、CD24、CD37、CD38、CD39、CD40、CD72、CD73、CD74、CDw75、CDw76、CD77、CD78、CD79a/b、CD80、CD81、CD82、CD83、CD84、CD85、CD86、CD89、CD98、CD126、CD127、CDw130、CD138 和 CDw150。

16. 如权利要求 13 所述的蛋白，其中所述 TNF 受体超家族成员是选自由以下组成的组：4-1BB/TNFRSF9、NGF R/TNFRSF16、BAFF R/TNFRSF13C、护骨素/TNFRSF11B、BCMA/TNFR SF17、OX40/TNFRSF4、CD27/TNFRSF7、RANK/TNFRSF 11A、CD30/TNFRSF8、RELT/TNFRSF19L、CD40/TNFRSF 5、TACI/TNFRSF13B、DcR3/TNFRSF6B、TNF RI/TNFRSF 1A、DcTRAIL R1/TNFRSF23、TNF RII/TNFRSF1B、Dc TRAIL R2/TNFRSF22、TRAIL R1/TNFRSF10A、DR3/ TNFRSF25、TRAIL R2/TNFRSF10B、DR6/TNFRSF21、TRAIL R3/TNFRSF10C、EDAR、TRAIL R4/TNFRSF10D、Fas/TNFRSF6、TROY/TNFRSF19、GITR/TNFRSF18、TWEAK R/TNFRSF12、HVEM/TNFRSF14、XEDAR、淋巴毒素 β R/TNFRSF3、4-1BB 配体/TNFSF9、淋巴毒素、APRIL/TNFSF13、淋巴毒素 β/TNFSF3、BAFF/TNFSF13C、OX40 配体/TNFSF4、CD27 配

体/TNFSF7、TL1A/ TNFSF15、CD30 配体/TNFSF8、TNF-α/TNFSF1A、CD40 配体/TNFSF5、TNF-β/TNFSF1B、EDA-A2、TRAIL/TNFSF10、Fas 配体/TNFSF6、TRANCE/TNFSF11、GITR 配体/TNFSF18、TWEAK/TNFSF12 和 LIGHT/ TNFSF14。

17. 如权利要求 13 所述的蛋白，其中所述 Hedgehog 家族成员是选自由 Patched 和 Smoothened 组成的组。

18. 如权利要求 13 所述的蛋白，其中所述受体酪氨酸激酶是选自由以下组成的组：Axl、FGF R4、C1q R1/CD93、FGF R5、DDR1、Flt-3、DDR2、HGF R、Dtk、IGF-I R、EGF R、IGF-II R、Eph、INSRR、EphA1、胰岛素 R/CD220、EphA2、M-CSF R、EphA3、Mer、EphA4、MSP R/Ron、EphA5、MuSK、EphA6、PDGF R α、EphA7、PDGF R β、EphA8、Ret、EphB1、ROR1、EphB2、ROR2、EphB3、SCF R/c-kit、EphB4、Tie-1、EphB6、Tie-2、ErbB2、TrkA、ErbB3、TrkB、ErbB4、TrkC、FGF R1、VEGF R1/Flt-1、FGF R2、VEGF R2/Flk-1、FGF R3 和 VEGF R3/Flt-4。

19. 如权利要求 13 所述的蛋白，其中所述蛋白聚糖相关分子是选自由蛋白聚糖及其调节剂组成的组。

20. 如权利要求 13 所述的蛋白，其中所述转化生长因子(TGF)-β 超家族成员是选自由以下组成的组：活化素 RIA/ ALK-2、GFR α-1、活化素 RIB/ALK-4、GFR α-2、活化素 RIIA、GFR α-3、活化素 RIIB、GFR α-4、ALK-1、MIS RII、ALK-7、Ret、BMPR-IA/ALK-3、TGF-β RI/ALK-5、BMPR-IB/ALK-6、TGF-β RII、BMPR-II、TGF-β RIIb、内皮因子/CD105 和 TGF-β RIII。

21. 如权利要求 13 所述的蛋白，其中所述 Wnt 相关分子是选自由以下组成的组：Frizzled-1、Frizzled-8、Frizzled-2、Frizzled-9、Frizzled-3、sFRP-1、Frizzled-4、sFRP-2、Frizzled-5、sFRP-3、Frizzled-6、sFRP-4、

Frizzled-7、MFRP、LRP 5、LRP 6、Wnt-1、Wnt-8a、Wnt-3a、Wnt-10b、Wnt-4、Wnt-11、Wnt-5a、Wnt-9a 和 Wnt-7a。

22. 如权利要求 13 所述的蛋白，其中所述受体配体是选自由以下组成的组：4-1BB 配体/TNFSF9、淋巴毒素、APRIL/ TNFSF 13、淋巴毒素 β/TNFSF3、BAFF/TNFSF13C、OX40 配体/TNFSF4、CD27 配体/TNFSF7、TL1A/TNFSF15、CD30 配体/TNFSF8、TNF-α/TNFSF1A、CD40 配体/TNFSF 5、TNF-β/TNFSF1B、EDA-A2、TRAIL/TNFSF10、Fas 配体/TNFSF6、TRANCE/TNFSF11、GITR 配体/TNFSF 18、TWEAK/TNFSF12、LIGHT/TNFSF14、双调蛋白、NRG1 同种型 GGF2、β 细胞生长因子、NRG1 同种型 SMDF、EGF、NRG1-α/ HRG1-α、Epigen、NRG1-β1/HRG1-β1、表皮调节素、TGF-α、HB-EGF、TMEFF1/土莫调节素-1、神经调节蛋白-3、TMEFF2、IGF-I、IGF-II、胰岛素、活化素 A、活化素 B、活化素 AB、活化素 C、BMP-2、BMP-7、BMP-3、BMP-8、BMP-3b/ GDF-10、BMP-9、BMP-4、BMP-15、BMP-5、Dpp 蛋白(Decapentaplegic)、BMP-6、GDF-1、GDF-8、GDF-3、GDF-9、GDF-5、GDF-11、GDF-6、GDF-15、GDF-7、阿特明(Artemin)、纽特元(Neurturin)、GDNF、普塞芬(Persephin)、TGF-β、TGF-β 2、TGF-β 1、TGF-β 3、LAP(TGF-β 1)、TGF-β 5、潜在性 TGF-β1、潜在性 TGF-β bp1、TGF-β 1.2、Lefty、Nodal、MIS/AMH、酸性 FGF、FGF-12、碱性 FGF、FGF-13、FGF-3、FGF-16、FGF-4、FGF-17、FGF-5、FGF-19、FGF-6、FGF-20、FGF-8、FGF-21、FGF-9、FGF-23、FGF-10、KGF/FGF-7、FGF-11、神经菌毛素(Neuropilin)-1、PlGF、神经菌毛素-2、PlGF-2、PDGF、PDGF-A、VEGF、PDGF-B、VEGF-B、PDGF-C、VEGF-C、PDGF-D、VEGF-D 和 PDGF-AB。

23. 如权利要求 13 所述的蛋白，其中所述 T-细胞靶标是选自由以下组成的组：2B4/SLAMF4、IL-2 R α、4-1BB/TNFRSF9、IL-2 R β、ALCAM、B7-1/CD80、IL-4 R、B7-H3、BLAME/ SLAMF8、BTLA、IL-6 R、CCR3、IL-7 R α、CCR4、CXCR1/ IL-8 RA、CCR5、CCR6、

IL-10 R α、CCR7、IL-10 R β、CCR8、IL-12 R β 1、CCR9、IL-12 R β 2、CD2、IL-13 R α 1、IL-13、CD3、CD4、ILT2/CD85j、ILT3/CD85k、ILT4/CD85d、ILT5/CD85a、整合素 α 4/CD49d、CD5、整合素 α E/CD103、CD6、整合素 α M/CD11b、CD8、整合素 α X/ CD11c、整合素 β 2/CD18、KIR/CD158 、 CD27/TNFRSF7 、 KIR2DL1 、 CD28 、 KIR2DL3 、 CD30/TNFRSF8、KIR2DL4/ CD158d、CD31/PECAM-1、KIR2DS4、CD40 配体/ TNFSF5、LAG-3、CD43、LAIR1、CD45、LAIR2、CD83、白三烯 B4 R1、CD84/SLAMF5、NCAM-L1、CD94、NKG2A、CD97、NKG2C、CD229/SLAMF3、NKG2D、CD2F- 10/SLAMF9、NT-4、CD69、NTB-A/SLAMF6、共同 γ 链/IL-2 R γ、骨桥蛋白、CRACC/SLAMF 7、PD-1、CRTAM、PSGL-1、CTLA-4、RANK/TNFRSF11A、CX3CR1、CX3CL1、L-选择素、CXCR3、SIRP β 1、CXCR4、SLAM、CXCR6、TCCR/WSX-1、DNAM-1、胸腺生成素、EMMPRIN/CD147、TIM-1、EphB6、TIM-2、Fas/TNFRSF6、TIM-3、Fas 配体/TNFSF6、TIM-4、Fc γ RIII/CD16、TIM-6、GITR/TNFRSF18、TNF RI/TNFRSF1A、颗粒溶素 、 TNF RII/ TNFRSF 1B 、 HVEM/TNFRSF14 、 TRAIL R1/TNFRSF10A、ICAM-1/CD54、TRAIL R2/TNFRSF10B、ICAM-2/CD102、TRAIL R3/TNFRSF10C、IFN-γ R1、TRAIL R4/TNFRSF10D、IFN-γ R2、TSLP、IL-1 RI 和 TSLP R。

24. 如权利要求 13 所述的蛋白，其中所述 NK 细胞靶标是选自由以下组成的组：2B4/SLAMF4、KIR2DS4、CD155/PVR、KIR3DL1、CD94、LMIR1/CD300A、CD69、LMIR2/CD 300c、CRACC/SLAMF7、LMIR3/CD300LF 、 DNAM-1 、 LMIR5/CD300LB 、 Fcε RII 、LMIR6/CD300LE、Fcγ RI/CD 64、MICA、Fcγ RIIB/CD32b、MICB、Fcγ RIIC/CD32c、MULT-1、Fcγ RIIA/CD32a、柄蛋白-2/CD112、Fcγ RIII/CD16、NKG2A、FcRH1/IRTA5、NKG2C、FcRH2/ IRTA4、NKG2D、FcRH4/IRTA1、NKp30、FcRH5/IRTA2、NKp44、类 Fc 受体 3/CD16-2、NKp46/ NCR1、NKp80/KLRF1、NTB-A/SLAMF6、Rae-1、Rae-1 α、Rae-1 β、Rae-1 δ、H60、Rae-1ε、ILT2/CD85j、Rae-1γ、ILT3/CD85k、

TREM-1、ILT4/CD85d、TREM-2、ILT5/CD 85a、TREM-3、KIR/CD158、TREML1/TLT-1 、 KIR2DL1 、 ULBP-1 、 KIR2DL3 、 ULBP-2 、 KIR2DL4/CD158d 和 ULBP- 3。

25. 如权利要求 13 所述的蛋白，其中所述单核细胞/巨噬细胞靶标是选自由以下组成的组：B7-1/CD80、ILT4/CD85d、B7-H1、ILT5/CD85a、共同 β 链、整合素 α 4/CD49d、BLAME/ SLAMF8、整合素 α X/CD11c、CCL6/C10、整合素 β 2/ CD18、CD155/PVR、整合素 β 3/CD61、CD31/PECAM-1、胶乳素(Latexin)、CD36/SR-B3、白三烯 B4 R1、CD40/TNFRSF5 、 LIMPII/SR-B2 、 CD43 、 LMIR1/CD300A 、 CD45 、 LMIR2/CD300c 、 CD68 、 LMIR3/CD300LF 、 CD84/SLAMF 5 、 LMIR5/CD300LB、CD97、LMIR6/CD300LE、CD163、LRP-1、CD2F-10/SLAMF9 、 MARCO 、 CRACC/SLAMF7 、 MD-1 、 ECF-L 、 MD-2 、 EMMPRIN/CD147、MGL2、内皮因子/CD105、骨活素/GPNMB、Fc γ RI/CD64、骨桥蛋白、Fc γ RIIB/CD32b、PD-L2、Fc γ RIIC/CD32c、Siglec- 3/CD33、Fc γ RIIA/CD32a、SIGNR1/CD209、Fc γ RIII/ CD16、SLAM、GM-CSF R α、TCCR/WSX-1、ICAM-2/ CD102、TLR3、IFN-γ R1、TLR4、IFN-γ R2、TREM-1、IL-1 RII、TREM-2、ILT2/CD85j、TREM-3、ILT3/CD85k、TREML1/TLT-1、2B4/SLAMF4、IL-10 R α、ALCAM、IL-10 R β、氨基肽酶 N/ANPEP、ILT2/CD85j、共同 β 链、ILT3/ CD85k、C1q R1/CD93、ILT4/CD85d、CCR1、ILT5/CD85a、CCR2、整合素 α 4/CD49d、CCR5、整合素 α M/CD11b、CCR8、整合素 α X/CD11c、CD155/PVR、整合素 β 2/CD 18、CD14、整合素 β 3/CD61、CD36/SR-B3、LAIR1、CD43、LAIR2、CD45、白三烯 B4 R1、CD68、LIMPII/SR-B2 、 CD84/SLAMF5 、 LMIR1/CD300A 、 CD97 、 LMIR2/CD300c、CD163、LMIR3/CD300LF、凝血因子 III/组织因子、LMIR5/CD300LB、CX3CR1、CX3CL1、LMIR6/CD300LE、CXCR4、LRP-1、CXCR6 、 M-CSF R 、 DEP-1/CD148 、 MD-1 、 DNAM-1 、 MD-2 、 EMMPRIN/CD147、MMR、内皮因子/ CD105、NCAM-L1、Fcγ RI/CD64、PSGL-1、Fcγ RIII/ CD16、RP105、G-CSF R、L-选择素、GM-CSF R α、

Siglec-3/ CD33、HVEM/TNFRSF14、SLAM、ICAM-1/CD54、TCCR/WSX-1、ICAM-2/CD102、TREM-1、IL-6 R、TREM-2、CXCR1/IL-8 RA、TREM-3 及 TREML1/TLT-1。

26. 如权利要求 13 所述的蛋白，其中所述树突状细胞靶标是选自由以下组成的组: CD36/SR-B3、LOX-1/SR-E1、CD68、MARCO、CD163、SR-AI/MSR 、 CD5L 、 SREC-I 、 CL-P1/COLEC12 、 SREC-II 、 LIMPII/SR-B2、RP105、TLR4、TLR1、TLR5、TLR2、TLR6、TLR3、TLR9、4-1BB 配体/TNFSF9、IL-12/IL-23 p40、4-氨基-1,8-萘酰亚胺、ILT2/CD85j 、 CCL21/6Ckine 、 ILT3/CD85k 、 8- 氧 基 (oxo)-dG 、 ILT4/CD85d、8D6A、ILT5/CD85a、A2B5、整合素 α 4/CD49d、Aag、整合素 β 2/CD18、AMICA、郎罕细胞特异蛋白、B7-2/CD86、白三烯 B4 R1、B7-H3、LMIR1/ CD300A、BLAME/SLAMF8、LMIR2/CD300c、C1q R1/ CD93、LMIR3/CD300LF、CCR6、LMIR5/CD300LB、CCR7、LMIR6/CD300LE、CD40/TNFRSF5、MAG/Siglec-4a、CD43、MCAM、CD45、MD-1、CD68、MD-2、CD83、MDL-1/CLEC5A、CD84/SLAMF5、MMR、CD97、NCAM- L1、CD2F-10/SLAMF9、骨活素/GPNMB、Chem 23 、 PD-L2 、 CLEC-1 、 RP105 、 CLEC-2 、 Siglec-2/CD22 、 CRACC/SLAMF7 、 Siglec-3/CD33 、 DC-SIGN 、 Siglec-5 、 DC-SIGNR/CD299、Siglec-6、DCAR、Siglec-7、DCIR/CLEC4A、Siglec-9、DEC-205、Siglec-10、凝集素(Dectin)-1 /CLEC7A、Siglec-F、凝集素-2/CLEC6A、SIGNR1/CD209、DEP-1/CD148、SIGNR4、DLEC、SLAM、EMMPRIN/CD147 、 TCCR/WSX-1 、 Fcγ RI/CD64 、 TLR3 、 Fcγ RIIB/CD32b、TREM-1、Fcγ RIIC/CD32c、TREM-2、Fcγ RIIA/CD32a、TREM-3、Fcγ RIII/CD16、TREML1/TLT-1、ICAM-2/CD102 和辣椒素 R1(Vanilloid R1)。

27. 如权利要求 13 所述的蛋白，其中所述血管生成靶标是选自由以下组成的组: 血管生成素(Angiopoietin)-1、类血管生成素 2、血管生成素-2、类血管生成素 3、血管生成素-3、类血管生成素 7/CDT6、血

管生成素-4、Tie-1、类血管生成素 1、Tie-2、血管生长素(Angiogenin)、iNOS、凝血因子 III/组织因子、nNOS、CTGF/CCN2、NOV/CCN3、DANCE、OSM、EDG-1、Plfr、EG-VEGF/PK1、增殖素(Proliferin)、内皮抑制素(Endostatin)、ROBO4、红血球生成素(Erythropoietin)、凝血栓蛋白-1、激肽抑制素(Kininostatin)、凝血栓蛋白-2、MFG-E8、凝血栓蛋白-4、一氧化氮(Nitric Oxide)、VG5Q、eNOS、EphA1、EphA5、EphA2、EphA6、EphA3、EphA7、EphA4、EphA8、EphB1、EphB4、EphB2、EphB6、EphB3、Ephrin-A1、Ephrin-A4、Ephrin-A2、Ephrin-A5、Ephrin-A3、Ephrin- B1、Ephrin-B3、Ephrin-B2、酸性 FGF、FGF-12、碱性 FGF、FGF-13、FGF-3、FGF-16、FGF-4、FGF-17、FGF-5、FGF-19、FGF-6、FGF-20、FGF-8、FGF-21、FGF-9、FGF-23、FGF-10、KGF/FGF-7、FGF-11、FGF R1、FGF R4、FGF R2、FGF R5、FGF R3、神经菌毛素-1、神经菌毛素-2、臂板蛋白(Semaphorin) 3A、臂板蛋白 6B、臂板蛋白 3C、臂板蛋白 6C、臂板蛋白 3E、臂板蛋白 6D、臂板蛋白 6A、臂板蛋白 7A、MMP、MMP-11、MMP-1、MMP-12、MMP-2、MMP-13、MMP-3、MMP-14、MMP-7、MMP-15、MMP-8、MMP-16/MT3-MMP、MMP-9、MMP-24/MT5- MMP、MMP-10、MMP-25/MT6-MMP、TIMP-1、TIMP-3、TIMP-2、TIMP-4、ACE、IL-13 Rα1、IL-13、C1q R1/CD 93、整合素 α 4/CD49d、VE-钙粘素(Cadherin)、整合素 β 2/CD18、CD31/PECAM-1、KLF4、CD36/SR-B3、LYVE- 1、CD151、MCAM、CL-P1/COLEC12、粘连蛋白-2/CD112、凝血因子 III/组织因子、E-选择素、D6、P-选择素、DC- SIGNR/CD299、SLAM、EMMPRIN/CD147、Tie-2、内皮因子/CD105、TNF RI/TNFRSF1A、EPCR、TNF RII/TNFRSF1B、红血球生成素 R、TRAIL R1/TNFRSF10A、ESAM、TRAIL R2/TNFRSF10B、FABP5、VCAM-1、ICAM-1/CD54、VEGF R2/Flk-1、ICAM-2/CD102、VEGF R3/Flt-4、IL-1 RI 和 VG5Q。

28. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中至少一者特异性地结合选自由以下组成的组的靶标：前列腺特异性膜抗原(叶酸水解酶 1)、表皮生长因子受体、高级糖

基化最终产物的受体、IL-17 A、IL-17 F、P19、Dickkopf-1、NOTCH1、NG2、IgHE、IL-22R、IL-21、类淀粉β寡聚物、类淀粉β前驱蛋白、NOGO 受体(RTN4R)、低密度脂蛋白受体相关蛋白 5、IL-4、肌肉抑制素、极晚期抗原 4 和 IGF-1R。

29. 如权利要求 12 所述的蛋白，其中所述癌细胞是转化的造血细胞。

30. 如权利要求 29 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中至少一者识别选自由以下组成的组的靶标：B-细胞靶标、T-细胞靶标、树突状细胞靶标或 NK-细胞靶标。

31. 如权利要求 29 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中至少一者识别选自由以下组成的组的靶标：CD5、CD10、CD11c、CD15、CD19、CD20、CD21、CD22、CD23、CD25、CD27、CD30、CD38、CD45、CD70、CD80、CD86、CD103、IRTA1、IRTA2、IRTA3、IRTA4、IRTA5、B-B2、B-B8 和 B-细胞抗原受体。

32. 如权利要求 1 所述的蛋白，其包含选自由以下组成的组的序列：SEQ ID NO: 2、4、6、103、105、107、109、332、333、334 和 345。

33. 如权利要求 1 所述的蛋白，其包含选自由以下组成的组的序列：SEQ ID NO: 355、356、357、358、359、360、361、362、363、364 和 365。

34. 如权利要求 1 所述的蛋白，其中所述恒定亚区识别效应细胞 F_C 受体。

35. 如权利要求 1 所述的蛋白，其中所述恒定亚区识别选自由以

下组成的组的效应细胞表面蛋白：CD16、CD32a、CD32b、CD64、CD89、FcεR1、FcRn 和 pIgR。

36. 如权利要求 1 所述的蛋白，其中所述恒定亚区包含 C_{H2} 结构域和 C_{H3} 结构域。

37. 如权利要求 36 所述的蛋白，其中所述 C_{H3} 结构域为截短型，且包含选自由以下组成的组的 C-端序列：SEQ ID NO: 366、367、368、369、370 及 371。

38. 如权利要求 1 所述的蛋白，其中所述蝎形连接子源自免疫球蛋白铰链。

39. 如权利要求 38 所述的蛋白，其中所述蝎形连接子相对于衍生所述蝎形连接子的铰链具有数目减少但非零的半胱氨酸，且其中一个半胱氨酸对应于所述免疫球蛋白铰链的 N-端铰链半胱氨酸。

40. 如权利要求 39 所述的蛋白，其中所述蝎形连接子源自 IgG1 铰链区，且其中所述连接子包含 1 个或 2 个半胱氨酸，另外其中所述连接子保留对应于所述 IgG1 铰链区的 N-端铰链半胱氨酸的半胱氨酸。

41. 如权利要求 1 所述的蛋白，其中所述蝎形连接子源自 C-凝集素的茎区。

42. 如权利要求 1 所述的蛋白，其中所述蝎形连接子包含选自由 SEQ ID NO: 373、374、375、376 和 377 组成的组的序列。

43. 如权利要求 1 所述的蛋白，其中所述蝎形连接子包含选自由 SEQ ID NO: 351、352、353 和 354 组成的组的序列。

44. 如权利要求 1 所述的蛋白，其进一步包含具有至少约 5 个氨基酸的连接子，所述连接子与所述恒定亚区连接，且与所述第一结合结构域连接，从而使所述连接子定位于所述恒定亚区与所述第一结合结构域之间。

45. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中至少一者的结合亲和性小于 10^{-9} M。

46. 如权利要求 1 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中至少一者的结合亲和性为至少 10^{-6} M。

47. 如权利要求 1 所述的蛋白，其中所述效应功能是选自由抗体依赖性细胞介导的细胞毒性(ADCC)和补体依赖性细胞毒性(CDC)组成的组。

48. 如权利要求 1 所述的蛋白，其中所述蛋白在人类的体内半衰期为至少 28 小时。

49. 药学上的组合物，其包含如权利要求 1 所述的蛋白和药学上可接受的佐剂、载体或赋形剂。

50. 编码如权利要求 1 所述的蛋白的核酸。

51. 包含如权利要求 50 所述的核酸的载体。

52. 包含如权利要求 51 所述的载体的宿主细胞。

53. 制备如权利要求 1 所述的蛋白的方法，所述方法包括:

a. 将编码如权利要求 1 所述的蛋白的核酸引入宿主细胞中；和

b. 将所述宿主细胞在适于表达所述蛋白的条件下孵育，从而以

至少 1 mg/ml 的水平表达所述蛋白。

54. 如权利要求 53 所述的方法，其进一步包含分离所述蛋白。

55. 如权利要求 53 所述的方法，其中所述宿主细胞是选自由以下组成的组：VERO 细胞、海拉细胞、CHO 细胞、COS 细胞、W138 细胞、BHK 细胞、HepG2 细胞、3T3 细胞、RIN 细胞、MDCK 细胞、A549 细胞、PC12 细胞、K562 细胞、HEK293 细胞、N 细胞、草地贪夜蛾(*Spodoptera frugiperda*)细胞、酿酒酵母(*Saccharomyces cerevisiae*)细胞、巴斯德毕赤氏酵母(*Pichia pastoris*)细胞、真菌细胞及细菌细胞。

56. 制备编码如权利要求 1 所述的蛋白的核酸的方法，所述方法包括：

a. 将编码源自免疫球蛋白可变区的第一结合结构域的聚核苷酸的 3'端与编码恒定亚区的聚核苷酸的 5'端共价连接；

b. 将编码蝎形连接子的聚核苷酸的 5'端与编码所述恒定亚区的聚核苷酸的 3'端共价连接；及

c. 将编码源自免疫球蛋白可变区的第二结合结构域的聚核苷酸的 5'端与编码所述蝎形连接子的聚核苷酸的 3'端共价连接；

从而形成编码具有效应功能的多价结合蛋白的核酸。

57. 如权利要求 56 所述的方法，其中所述编码第一结合结构域的聚核苷酸包含选自由以下组成的组的序列：SEQ ID NO: 1 (抗-CD20 可变区)、SEQ ID NO: 3 (抗-CD28 可变区，V_L-V_H取向)和 SEQ ID NO: 5 (抗-CD28 可变区，V_H-V_L取向)。

58. 如权利要求 56 所述的方法，其进一步包含将连接子聚核苷酸插在所述编码第一结合结构域的聚核苷酸与所述编码恒定亚区的聚核苷酸之间，所述连接子聚核苷酸编码具有至少 5 个氨基酸的肽连接子。

59. 如权利要求 56 所述的方法，其进一步包含将连接子聚核苷酸插在所述编码恒定亚区的聚核苷酸与所述编码第二结合结构域的聚核苷酸之间，所述连接子聚核苷酸编码具有至少 5 个氨基酸的肽连接子。

60. 诱导靶标细胞损坏的方法，所述方法包括使靶标细胞与治疗有效量的如权利要求 1 所述的蛋白接触。

61. 如权利要求 60 所述的方法，其中所述靶标细胞被体内接触是通过给予有需要的有机体所述蛋白。

62. 治疗细胞增殖病症的方法，所述方法包括给予有需要的有机体治疗有效量的如权利要求 1 所述的蛋白。

63. 如权利要求 62 所述的方法，其中所述病症是选自由癌症、自体免疫病症、劳氏肉瘤病毒感染及炎症组成的组。

64. 如权利要求 62 所述的方法，其中所述蛋白被给药是通过编码如权利要求 1 所述的蛋白的核酸的体内表达。

65. 如权利要求 62 所述的方法，其中所述蛋白的给药通过选自由以下组成的组的途径：静脉内注射、动脉内注射、肌肉内注射、皮下注射、腹膜内注射及直接组织注射。

66. 改善与细胞增殖病症相关的症状的方法，所述方法包括给予有需要的有机体治疗有效量的如权利要求 1 所述的蛋白。

67. 如权利要求 66 所述的方法，其中所述症状是选自由疼痛、发热、肿胀及关节僵硬组成的组。

68. 治疗与感染物相关的感染的方法，所述方法包括给予有需要

的患者治疗有效量的如权利要求 1 所述的蛋白，其中所述蛋白包含特异性地结合所述感染物的靶标分子的结合结构域。

69. 改善与感染物相关的感染症状的方法，所述方法包括给予有需要的患者治疗有效量的如权利要求 1 所述的蛋白，其中所述蛋白包含特异性地结合所述感染物的靶标分子的结合结构域。

70. 降低归因于感染物的感染风险的方法，所述方法包括给予有感染风险的患者预防有效量的如权利要求 1 所述的蛋白，其中所述蛋白包含特异性地结合所述感染物的靶标分子的结合结构域。

71. 试剂盒，其包含如权利要求 1 所述的蛋白和一套关于给予所述蛋白以治疗细胞增殖病症或改善所述细胞增殖病症的症状的说明书。

72. 如权利要求 1 的多价单链结合蛋白，其中所述第一结合结构域与所述第二结合结构域中至少一者特异性地结合一种选自由以下组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86 及一种主要组织相容性复合物 II 类肽。

73. 如权利要求 72 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中的一者特异性地结合 CD20。

74. 如权利要求 73 所述的蛋白，其中所述另一结合结构域特异性地结合一种选自由以下组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86 及一种主要组织相容性复合物 II 类肽。

75. 如权利要求 72 所述的蛋白，其中所述第一结合结构域与所述

第二结合结构域中的一者特异性地结合 CD79b。

76. 如权利要求 75 所述的蛋白，其中所述另一结合结构域特异性地结合一种选自由以下组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86 及一种主要组织相容性复合物 II 类肽。

77. 如权利要求 72 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中的一者特异性地结合一种主要组织相容性复合物 II 类肽。

78. 如权利要求 77 所述的蛋白，其中所述另一结合结构域特异性地结合一种选自由以下组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86 及一种主要组织相容性复合物 II 类肽。

79. 如权利要求 72 所述的蛋白，其中所述第一结合结构域与所述第二结合结构域中的一者特异性地结合 CD22。

80. 如权利要求 79 所述的蛋白，其中所述另一结合结构域特异性地结合一种选自由以下组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86 及一种主要组织相容性复合物 II 类肽。

81. 如权利要求 72 所述的蛋白，其中所述蛋白对于靶标细胞行为具有一个相对于各结合结构域效应总和的协同效应，且其中所述蛋白包含一对结合结构域特异性地识别一对选自由以下组成的组的抗原：CD20-CD19、CD20-CD21、CD20-CD22、CD20-CD40、CD20-CD79a、CD20-CD79b 及 CD20-CD81。

82. 如权利要求 72 所述的蛋白，其中所述蛋白对于靶标细胞行为具有一个相对于各结合结构域效应总和的加成效应，且其中所述蛋白包含一对结合结构域特异性地识别一对选自由以下组成的组的抗原：CD20-CD23、CD20-CD30、CD20-CD37、CD20-CD70、CD20-CD80、CD20-CD86、CD79b-CD37、CD79b-CD81、主要组织相容性复合物 II 类肽-CD30 及主要组织相容性复合物 II 类肽-CD72。

83. 如权利要求 72 所述的蛋白，其中所述蛋白对于靶标细胞行为具有一个相对于各结合结构域效应总和的抑制效应，且其中所述蛋白包含一对结合结构域特异性地识别一对选自由以下组成的组的抗原：CD20-主要组织相容性复合物 II 类肽、CD79b-CD19、CD79b-CD20、CD79b-CD21 、 CD79b-CD22 、 CD79b-CD23 、 CD79b-CD30 、 CD79b-CD40 、 CD79b-CD70 、 CD79b-CD72 、 CD79b-CD79a 、 CD79b-CD80、CD79b-CD86、CD79b-主要组织相容性复合物 II 类肽、主要组织相容性复合物 II 类肽-CD19、主要组织相容性复合物 II 类肽-CD20、主要组织相容性复合物 II 类肽-CD21、主要组织相容性复合物 II 类肽-CD22、主要组织相容性复合物 II 类肽-CD23、主要组织相容性复合物 II 类肽-CD37、主要组织相容性复合物 II 类肽-CD40、主要组织相容性复合物 II 类肽-CD70、主要组织相容性复合物 II 类肽-CD79a、主要组织相容性复合物 II 类肽-CD79b、主要组织相容性复合物 II 类肽-CD80、主要组织相容性复合物 II 类肽-CD81、主要组织相容性复合物 II 类肽-CD86、CD22-CD19、CD22- CD40、CD22-CD79b、CD22-CD86 及 CD22-主要组织相容性复合物 II 类肽。

84. 鉴别如权利要求 1 所述的多价结合分子的结合结构域中的至少一者的方法，所述方法包括：

(a) 使抗同型抗体与特异性地识别第一抗原的抗体及特异性地识别第二抗原的抗体接触；

(b) 进一步使包含所述抗原中的至少一者的靶标与步骤(a)的所述组合物接触；及

(c) 测量所述靶标的活性，其中所述活性是用于鉴别所述多价结合分子的结合结构域中的至少一者。

85. 如权利要求 84 所述的方法，其中所述靶标为病变细胞。

86. 如权利要求 60 所述的方法，其中与包含所述第一结合结构域但不包含所述第二结合结构域的第一抗体和包含所述第二结合结构域但不包含所述第一结合结构域的第二抗体所诱发的损害总和相比，所述多价单链结合蛋白对所述靶标细胞诱发协同量的损害。

87. 如权利要求 86 所述的方法，其中所述多价单链结合蛋白包含一对结合结构域特异性地识别一对选自由以下组成的组的抗原：CD19/CD20、CD20/CD21、CD20/CD22、CD20/CD40、CD20/CD79a、CD20/CD79b、CD20/CD81、CD21/ CD79b、CD37/CD79b、CD79b/CD81、CD19/CL II、CD20/CL II、CD30/CL II、CD37/CL II、CD72/CL II 及 CD79b/CL II。

88. 如权利要求 60 所述的方法，其中与包含所述第一结合结构域但不包含所述第二结合结构域的第一抗体和包含所述第二结合结构域但不包含所述第一结合结构域的第二抗体所诱发的损害总和相比，所述多价单链结合蛋白对所述靶标细胞诱发抑制量的损害。

89. 如权利要求 88 所述的方法，其中所述多价单链结合蛋白包含一对结合结构域特异性地识别一对选自由以下组成的组的抗原：CD20/CL II、CD21/CD79b、CD22/CD79b、CD40/ CD79b、CD70/CD79b、CD72/CD79b、CD79a/CD79b、CD79b/ CD80、CD79b/CD86、CD21/CL II、CD22/CL II、CD23/CL II、CD40/CL II、CD70/CL II、CD80/CL II、CD86/CL II、CD19/CD22、CD20/CD22、CD21/CD22、CD22/CD23、CD22/ CD30、CD22/CD37、CD22/CD40、CD22/CD70、CD22/ CD72、CD22/79a、CD22/79b、CD22/CD80、CD22/CD86 及 CD22/CL II。

90. 如权利要求 60 所述的方法，其进一步包含多种多价单链结合蛋白。

91. 如权利要求 90 所述的方法，其中第一多价单链结合蛋白的结合结构域和第二多价单链结合蛋白的结合结构域对所述靶标细胞诱发协同量的损害。

92. 如权利要求 90 所述的方法，其中第一多价单链结合蛋白的结合结构域和第二多价单链结合蛋白的结合结构域对所述靶标细胞诱发抑制量的损害。

93. 组合物，其包含多种如权利要求 1 的多价单链结合蛋白。

94. 如权利要求 93 所述的组合物，其中第一多价单链结合蛋白的结合结构域和第二多价单链结合蛋白的结合结构域能够对靶标细胞诱发协同量的损害。

95. 如权利要求 93 所述的组合物，其中第一多价单链结合蛋白的结合结构域和第二多价单链结合蛋白的结合结构域能够对靶标细胞诱发加成量的损害。

96. 如权利要求 93 所述的组合物，其中第一多价单链结合蛋白的结合结构域和第二多价单链结合蛋白的结合结构域能够对靶标细胞诱发抑制量的损害。

97. 药学上的组合物，其包含如权利要求 93 所述的组合物和药学上可接受的载体、稀释剂或赋形剂。

98. 试剂盒，其包含如权利要求 93 所述的组合物和一套关于给予所述组合物以损害靶标细胞的说明书。

具有效应功能的单链多价结合蛋白

技术领域

本发明通常涉及多价结合分子及其治疗应用的领域。

序列表以文本形式且以符合电子归档的应用要求的 PDF 文件递交。序列表创建于 2007 年 6 月 12 日。序列表以引用的方式全文并入本文中。

背景技术

在健康哺乳动物中，免疫系统保护动物身体免受外来物体及病原体的损害。但在某些情况下，免疫系统会出差错，导致创伤性损伤和/或疾病。例如，B-细胞可产生识别自身蛋白而非外来蛋白的抗体，导致产生以自身抗体为特征的自体免疫疾病，诸如红斑性狼疮症、类风湿性关节炎及其类似疾病。在其他情况下，免疫系统抗击外来物质的典型有利效应起相反作用，诸如伴随器官移植。已确认哺乳动物免疫系统(且尤其人类免疫系统)的功效，且已努力控制该系统以避免或改善因免疫系统在异常环境(例如器官移植)中正常行使功能或免疫系统在其他外在正常环境(例如自体免疫疾病进展)中异常行使功能对健康所造成的有害结果。此外，已努力利用免疫系统来提供多种靶标特异性诊断及治疗方法，这视抗体特异性地识别及结合具特异性的抗原性靶标的能力而定。

免疫系统保护动物身体的一种方式是产生特化细胞，称作 B 淋巴细胞或 B-细胞。B-细胞产生结合至外来物体或病原体且在某些情况下介导外来物体或病原体的破坏的抗体。但在某些情况下，人类免疫系统(且特定而言，人类免疫系统的 B 淋巴细胞)会出差错且导致疾病。存在多种癌症涉及 B-细胞的失控增殖。亦存在多种自体免疫疾病涉及结合至动物身体的部分而非结合外来物体及病原体的抗体的 B-细胞产生。此外，存在多种自体免疫疾病及炎性疾病的病理学涉及 B-细胞，

例如，经由 B-细胞抗原向 T-细胞的不当呈递或经由涉及 B-细胞的其他途径。例如，缺乏 B-细胞的易自体免疫小鼠不发展自体免疫肾病、脉管炎或自身抗体。(Shlomchik 等人，J Exp. Med. 1994, 180:1295-306)。有趣的是，所述具有 B-细胞但缺乏免疫球蛋白产生的易自体免疫小鼠在经实验方法诱导时不发展自体免疫疾病(Chan 等人，J Exp. Med. 1999, 189:1639-48)，这表明 B-细胞在自体免疫疾病的发展中具有不可或缺的作用。

B-细胞可由其细胞表面上的分子来识别。CD20 为第一种由单克隆抗体识别的人类 B-细胞谱是特异性表面分子。其为未经糖基化的疏水性 35 kDa B-细胞跨膜磷蛋白，该磷蛋白的氨基端与羧基端均位于细胞内。Einfeld 等人，EMBO J. 1988, 7:711-17。CD20 由所有正常成熟 B-细胞表达，而非由前体 B-细胞或浆细胞表达。CD20 的天然配体还未被鉴定，且对 CD20 在 B-细胞生物学中的功能仍不完全了解。

另一种 B-细胞谱是特异性细胞表面分子为 CD37。CD37 为经严重糖基化的 40-52 kDa 蛋白，其属于细胞表面抗原的四次跨膜蛋白家族。其四次横跨细胞膜，形成两个细胞外环且使其氨基端与羧基端暴露于细胞质。CD37 于产生正常抗体(sIg+)的 B-细胞上得以高度表达，但在前 B-细胞或浆细胞上未经表达。CD37 在静息 T 细胞且活化 T 细胞、单核细胞及粒细胞上的表达较低，且检测不到 CD37 于 NK 细胞、血小板或红血球上的表达。参见 Belov 等人，Cancer Res., 61(11):4483-4489 (2001); Schwartz-Albiez 等人，J. Immunol., 140(3): 905-914 (1988); 及 Link 等人，J. Immunol., 137(9): 3013-3018 (1988)。除正常 B-细胞外，B-细胞起源的几乎所有恶性肿瘤(包括 CLL、NHL 及毛细胞白血病)对于 CD37 表达均呈阳性(Moore 等人，1987; Merson 及 Brochier 1988; Faure 等人，1990)。CD37 参与 B-细胞功能的调节，因为发现缺乏 CD37 的小鼠具有低含量的血清 IgG1 且发现其对病毒抗原及模型抗原的体液反应减弱。其看起来充当非经典型协同刺激分子或经由与 II 类 MHC 分子形成复合物来直接影响抗原呈递。参见 Knobeloch 等人，Mol. Cell. Biol., 20(15):5363-5369 (2000)。

基于以下概念进行研究及药物研发：B-细胞谱是特异性细胞表面

分子(诸如 CD37 和 CD20)本身可靶向抗体，所述抗体可结合至在其表面上具有 CD37 和 CD20、导致癌性疾病及自体免疫疾病的 B-细胞且介导所述 B-细胞的破坏。向患者提供在非人类动物中制备的结合至 CD37 或 CD20 的所谓“免疫治疗”抗体(或基于所制备的抗体)以耗尽导致癌性疾病或自体免疫疾病的 B-细胞。

单克隆抗体技术及遗传工程方法有利于研发用于诊断及治疗人类疾病的免疫球蛋白分子。免疫球蛋白的结构域结构与工程设计的相符之处在于，抗原结合结构域与赋予效应功能的结构域可在免疫球蛋白类与子类之间互换。免疫球蛋白结构及功能是综述于例如 Harlow 等人编的 Antibodies: A Laboratory Manual，第 14 章，Cold Spring Harbor Laboratory, Cold Spring Harbor (1988)中。关于重组抗体技术所有方面的广泛介绍以及详细信息可见于教科书“Recombinant Antibodies (重组抗体)” (John Wiley & Sons, NY, 1999)中。详细抗体工程设计实验室方案的综合性汇编可见于 R. Kontermann 及 S. Dübel(编)，“The Antibody Engineering Lab Manual (抗体工程设计实验室手册)” (Springer Verlag, Heidelberg/New York, 2000)中。

免疫球蛋白分子(缩写为 Ig)为一种多聚蛋白，其通常由两个相同轻链多肽及两个相同重链多肽(H_2L_2)组成，所述多肽由链间二硫键(亦即，相邻半胱氨酸残基的氢硫基之间的共价键)连接成巨分子复合物。基于重链组成定义五类人类免疫球蛋白，且命名为 IgG、IgM、IgA、IgE 及 IgD。IgG 类及 IgA 类抗体进一步分成子类，亦即分别为 IgG1、IgG2、IgG3 及 IgG4，以及 IgA1 及 IgA2。链内二硫键连接同一多肽链中的不同区域，导致形成环，所述环与相邻氨基酸一起构成免疫球蛋白结构域。各轻链及各重链在氨基-末端部分处具有单一可变区，不同抗体的单一可变区在氨基酸组成上显示相当大的不同。轻链可变区 V_L 具有单一抗原结合结构域且与重链可变区 V_H (也含有单一抗原结合结构域)结合以形成免疫球蛋白的抗原结合位点 Fv。

除可变区外，全长抗体链各自具有含有一个或多个结构域的恒定区。轻链具有含有单一结构域的恒定区。因此，轻链具有一个可变结构域及一个恒定结构域。重链具有含有若干个结构域的恒定区。IgG、

IgA 及 IgD 抗体中的重链具有三个指定为 C_{H1}、C_{H2} 及 C_{H3} 的结构域；IgM 及 IgE 抗体中的重链具有四个结构域：C_{H1}、C_{H2}、C_{H3} 及 C_{H4}。因此，重链具有一个可变结构域及三或四个恒定结构域。在所有已知物种中，值得注意的是所述结构域的不变构成，其中含有一或多个结构域的恒定区位于或靠近免疫球蛋白分子的轻链与重链的 C-末端，可变结构域朝向轻链与重链的 N-末端定位。免疫球蛋白结构及功能是综述于例如 Harlow 等人编的 *Antibodies: A Laboratory Manual* (抗体：实验室手册)，第 14 章，Cold Spring Harbor Laboratory, Cold Spring Harbor (1988)中。

免疫球蛋白的重链亦可分成三个功能区：Fd 区(包含 V_H 及 C_{H1}(亦即重链的两个 N-末端结构域)的片段)、铰链区及 Fc 区("可结晶片段"区)。Fc 区含有与细胞上的免疫球蛋白受体相互作用且与补体级联的起始元件相互作用的结构域。因此，一般认为 Fc 区或片段负责免疫球蛋白的效应功能，诸如 ADCC(抗体依赖性细胞介导的细胞毒性)、CDC(补体依赖性细胞毒性)及补体结合反应、结合至 Fc 受体、相对于缺乏 F_C 区的多肽而言更长的体内半衰期、蛋白 A 结合及或许甚至胎盘转移。Capon 等人，Nature, 337: 525-531, (1989)。此外，含有 Fc 区的多肽容许多肽的二聚化/多聚化。所述术语亦可用于其他免疫球蛋白的类似区。

尽管所有人类免疫球蛋白同型均含有共同的可识别结构，但各同型显示不同型式的效应功能。IgG(作为非穷举性实例)中和毒素及病毒，调理、结合补体(CDC)且参与 ADCC。相反，IgM 中和血源性病原体且参与调理作用。IgA 在与其分泌片段结合时被分泌且经由粘膜提供针对微生物感染的主要防御；其亦中和毒素且支持调理作用。IgE 介导炎性反应，主要涉及于确立完全反应所需的其他细胞的募集中。已知 IgD 用于提供免疫调节功能，控制 B 细胞的活化。同型效应功能的所述特征为可存在于人类同型之间的差异提供非全面性说明。

存在于 IgG、IgA、IgD 及 IgE 类抗体中的铰链区充当柔性间隔子，使得 Fab 部分在空间内可自由移动。在免疫球蛋白类及子类中，铰链结构域与恒定区相比在结构上不同，序列与长度均各不相同。例如，

铰链区的长度及柔性在 IgG 子类中各不相同。IgG1 的铰链区包含氨基酸 216-231，且由于其可自由挠曲，因此 Fab 片段可绕其对称轴旋转且可在以两个重链间二硫键的第一键为中心的球体内移动。IgG2 具有比 IgG1 更短的铰链，IgG2 的铰链具有 12 个氨基酸残基及四个二硫键。IgG2 的铰链区缺少甘氨酸残基，相对较短，且含有通过额外的重链间二硫键稳固的刚性聚脯氨酸双螺旋。所述特性限制 IgG2 分子的柔性。IgG3 与其他子类的不同之处在于其独特的延长铰链区(约 IgG1 铰链的四倍长)，该延长铰链区含有 62 个氨基酸(包括 21 个脯氨酸及 11 个半胱氨酸)，形成非柔性聚脯氨酸双螺旋。在 IgG3 中，Fab 片段距 Fc 片段相对较远，使得该分子具有更大柔性。IgG3 中的伸长铰链亦造成其分子量比其他子类更高。IgG4 的铰链区比 IgG1 的铰链区更短且其柔性居于 IgG1 柔性与 IgG2 柔性中间。据报导，铰链区的柔性依如下次序递减：IgG3>IgG1>IgG4>IgG2。该四种 IgG 子类就其效应功能而言亦彼此不同。此差异与结构差异(包括可变区之间相互作用、Fab 片段及恒定 Fc 片段方面的差异)有关。

根据结晶学研究，免疫球蛋白铰链区在功能上可进一步再分为三个区：上铰链区、核心区及下铰链区。Shin 等人，1992 *Immunological Reviews* 130:87。上铰链区包括 C_{H1} 的羧基端至铰链中限制移动的第一残基(通常为在两条重链之间形成链间二硫键的第一半胱氨酸残基)的氨基酸。上铰链区的长度与抗体的区段柔性相关。核心铰链区含有重链内二硫键，且下铰链区连接 C_{H2} 域的氨基末端且包括 C_{H2} 中的残基。同前。人类 IgG1 的核心铰链区含有序列 Cys-Pro-Pro-Cys，该序列在通过形成二硫键而二聚合时形成环状八肽，导致该环状八肽可充当枢轴，因此赋予柔性。铰链区亦可含有一个或多个糖基化位点，其包括多个具有不同结构类型的用于连接碳水化合物的位点。例如，IgA1 在铰链区的 17 氨基酸区段内含有五个糖基化位点，从而使铰链区多肽对肠蛋白酶具有耐受性，被认为是分泌免疫球蛋白的有利特性。

免疫球蛋白铰链区多肽序列的结构及柔性所容许的构形变化亦可影响抗体的 Fc 部分的效应功能。与 Fc 区相关的三类一般效应功能包括(1)经典补体级联的活化；(2)与效应细胞相互作用；及(3)免疫球蛋白

的区室化。不同人类 IgG 子类就其结合补体或活化补体级联及扩增补体级联的步骤的相对效力而言彼此不同。参见例如 Kirschfink, 2001 *Immunol. Rev.* 180:177; Chakraborti 等人，2000 *Cell Signal* 12:607; Kohl 等人，1999 *Mol. Immunol.* 36:893; Marsh 等人，1999 *Curr. Opin. Nephrol. Hypertens.* 8:557; Speth 等人， 1999 *Wien Klin. Wochenschr.* 111:378。

存在于骆驼科动物(骆驼、单峰骆驼及美洲驼； Hamers-Casterman 等人，1993 *Nature* 363:446； Nguyen 等人，1998 *J. Mol. Biol* 275:413)、护士鲨(Roux 等人，1998 *Proc. Nat. Acad. Sci. USA* 95:11804)及斑点兔银鲛(Nguyen 等人，2002 *Immunogenetics* 54 (1):39-47)中的免疫球蛋白的某些同型中存在常规抗体的 H_2L_2 结构的例外情况。显然，所述抗体仅使用重链可变区即可形成抗原结合区，亦即所述功能性抗体为仅重链的同型二聚体(称为“重链抗体”或“HCAb”)。虽然抗体技术在疾病诊断及治疗中具有优势，但在研发全抗体技术作为诊断和/或治疗试剂中仍存在某些不利方面。全抗体为大蛋白结构，例如含有两条轻链及两条重链的 IgG 同型的异型四聚体结构。所述大分子在某些应用中受空间位阻。例如，在实体肿瘤的治疗中，全抗体不易于渗入肿瘤内部。此外，相对大尺寸的全抗体可提供激发以确保所述分子的体内给药不诱发免疫反应。此外，活性抗体分子的产生通常涉及培养能够对初生抗体分子提供适当翻译后加工的重组真核细胞，而所述细胞难以培养且难以用提供活性抗体的商业使用产量的方式诱发。

最近，已建构较小免疫球蛋白分子以克服与全免疫球蛋白方法相关的问题。单链可变抗体片段(scFv)包含经由短肽与抗体轻链可变结构域连接的抗体重链可变结构域(Huston 等人，Proc. Natl. Acad. Sci. USA, 1988, 85: 5879-83)。由于 scFv 分子的尺寸小，因此其与全免疫球蛋白相比显示更有效地渗入组织。抗肿瘤 scFv 与相应嵌合抗体相比显示更快速的肿瘤渗透及在肿瘤块内的更均匀分布(Yokota 等人，Cancer Res. 1992, 52:3402-08)。

虽然 scFv 分子为血清治疗带来优势，但该治疗方法仍存在若干缺点。scFv 可从循环系统中快速清除，从而可减少对正常细胞的毒性作用，但此快速清除阻碍最低有效剂量递送至靶组织。由于表达及分离

scFv 困难而对产量造成不利影响，因此制造足量用于给予患者的 scFv 已成为挑战。在表达期间，scFv 分子缺乏稳定性且经常由于不同分子的可变区配对而聚集。此外，在哺乳动物表达系统中，scFv 分子的产量低，此限制有效制造治疗用 scFv 分子的潜能(Davis 等人，J Biol. Chem. 1990, 265:10410-18; Traunecker 等人，EMBO J 1991, 10: 3655-59)。已探究改良生产的策略，包括向可变区中添加糖基化位点(Jost, C. R. 美国专利第 5,888,773 号，Jost 等人，J. Biol. Chem. 1994, 69: 26267-73)。

使用 scFv 治疗的另一缺点为缺乏效应功能。不具有细胞溶解功能(诸如抗体依赖性细胞介导的细胞毒性(ADCC)及补体依赖性细胞毒性(CDC))、与免疫球蛋白的恒定区结合的 scFv 对治疗疾病而言无效。虽然已在 12 年前开始研发 scFv 技术，但目前尚无 scFv 产品获准用于治疗。

或者，已提议可利用特异性抗原结合活性及小尺寸的 scFv 使 scFv 与诸如毒素的另一分子融合以将毒素递送至靶组织。Chaudary 等人，Nature 1989, 339:394; Batra 等人，Mol. Cell. Biol. 1991, 11:2200。因此，可将毒素与 scFv 的结合或融合作为提供有效抗原特异性分子的替代性策略，但以所述结合体或嵌合体给药却受到过度和/或非特异毒性(因所述制剂中的毒素部分所致)的限制。毒性作用可包括肝酶的超生理性升高及血管渗漏综合症，及其他不良作用。此外，免疫毒素在给予宿主后本身具有高免疫原性，且针对免疫毒素所产生的宿主抗体限制重复治疗性治疗个体的潜在有效性。

非外科癌症治疗(诸如外部照射及化疗)由于所述治疗对癌细胞显示特异性缺乏、因对正常组织及细胞的毒性作用而功效受限。为克服此局限性，已研发靶向治疗方法以增强治疗对需要其的细胞及组织的特异性。该靶向方法用于体内用途的实例为给予抗体结合体，其中抗体被设计为可特异性地识别与需要治疗的细胞或组织相关的标志物且使该抗体与治疗剂(诸如毒素(在癌症治疗的情况下))结合。抗体作为全身性药剂可循环至敏感性及不良身体区室，诸如骨髓。在急性辐射损伤中，淋巴及造血区室的破坏为发展败血病且随后死亡的主要因素。

此外，抗体为球状大蛋白，其对需要治疗的组织可显示不良的渗透。

罹患多种末期疾病过程的人类患者及非人类个体通常需要器官移植。然而，器官移植必须应对受者的不当免疫反应且通过用影响造血系统的淋巴部分及其他部分的细胞毒性剂抑制受者对外来器官的细胞性免疫反应来避免植入器官的免疫排斥反应。移植体接受性受到受者对所述细胞毒性化学品(其中多种与抗癌(抗增殖性)药剂类似)的耐受性的限制。同样，当使用细胞毒性抗微生物剂(尤其抗病毒药物)或当使用细胞毒性药物治疗自体免疫疾病时，例如在治疗全身性红斑性狼疮症中，严重限制为治疗剂对身体的骨髓及造血细胞的毒性作用。

靶向疗法(诸如靶向抗体结合疗法)的用途经设计为尽可能使最大量的治疗剂定位于需要作用的位点，且所述疗法是否成功通过治疗剂的相对高的信号与本底值的比率来体现。靶向抗体的实例包括抗体或抗体片段、细胞特异性肽或组织特异性肽以及激素及其他受体结合分子的诊断剂结合体或治疗剂结合体。例如，已使用针对与病理性细胞及正常细胞相关以及与病原性微生物相关的不同决定子的抗体来检测及治疗多种病理性状态或病灶。在所述方法中，靶向抗体直接偶联至如以下文献中所述的适当检测剂或治疗剂，例如 Hansen 等人，美国专利第 3,927,193 号及 Goldenberg，美国专利第 4,331,647 号、第 4,348,376 号、第 4,361,544 号、第 4,468,457 号、第 4,444,744 号、第 4,460,459 号、第 4,460,561 号、第 4,624,846 号及第 4,818,709 号。

直接靶向方法(亦即其中诊断剂或治疗剂(“活性剂”)直接偶联至靶向部分的方法)中所遭遇的一个问题为，偶联物的相对较小部分实际上结合至靶标位点，而大部分偶联物保持循环且想方设法减损靶向偶联物的功能。为确保活性剂的最大量定位，通常给予过量靶向偶联物，从而确保某些偶联物保持未结合状态且促进活性剂的本底含量。诊断性偶联物(例如不结合其靶标的放射免疫显像偶联物或磁共振成像偶联物)可保持循环，从而增加本底且降低诊断技术的解析度。在治疗性偶联物的情况下(该治疗性偶联物具有与长时循环靶向部分(诸如抗体)连接的作为活性剂的毒素(例如放射性同位素、药物或毒性化合物))，循环偶联物可产生宿主不可接受的毒性，诸如骨髓中毒或全身性副作

用。

美国专利第 4,782,840 号揭示一种在手术期间降低高本底辐射量的作用的方法。该方法包括用对赘生性组织具特异性的抗体对患者进行注射，其中经放射性同位素(诸如碘-125)标记的抗体具有适当长的半衰期。在注射经放射性标记的抗体后，将手术延迟至少 7 至 10 天、优选 14 至 21 天，以使得可将任何未结合的放射性标记抗体清除直至低本底量。

美国专利第 4,932,412 号揭示在手术进行中检测期间用于降低或修正非特异性本底辐射的方法。所述方法包括向已接受经放射性标记的初级抗体的患者给予造影剂、减影剂或结合初级抗体的二级抗体。

除产生上述抗体外，免疫系统包括多种具有强生物效应的细胞类型。在造血期间，源于骨髓的干细胞分化成免疫系统的成熟细胞("B"细胞)或分化成自骨髓中迁出在胸腺中成熟的细胞前体("T"细胞)。

B 细胞对于免疫反应的体液组成至关重要。B 细胞经抗原的适当呈递活化而变成抗体分泌浆细胞；抗原呈递亦导致活化 B 细胞的克隆扩增。B 细胞主要负责免疫反应的体液组成。浆细胞在其表面上通常显示约 10^5 个抗体分子(IgD 及 IgM)。

T 淋巴细胞可分成两类。细胞毒性 T 细胞、Tc 淋巴细胞或 CTL(CD8+T 细胞)杀死携带有与 I 类 MHC 相关的外来表面抗原的细胞且可杀死具有细胞内寄生物(细菌或病毒)的细胞，只要所感染细胞在其表面上显示微生物抗原即可。Tc 细胞杀死肿瘤细胞，且说明对植入细胞的排斥反应。Tc 细胞识别靶标细胞上的抗原-I 类 MHC 复合物，与其接触，且将颗粒内含物直接释放至靶标细胞膜中，从而溶解细胞。

第二类 T 细胞为辅助 T 细胞或 Th 淋巴细胞(CD4+T 细胞)，其产生在 B 细胞成熟为抗体分泌浆细胞中作为"辅助"因子的淋巴因子。Th 细胞亦产生某些刺激效应 T 淋巴细胞的分化及巨噬细胞的活性的淋巴因子。Th1 细胞识别巨噬细胞上与 II 类 MHC 相关的抗原且(通过 IL-1)经活化以产生淋巴因子，包括活化巨噬细胞及 NK 细胞的 IFN-γ。所述细胞介导细胞介导的免疫反应(包括延迟型过敏性反应)的各种方面。Th2 细胞识别抗原呈递细胞或 APC(例如迁移性巨噬细胞及树突状细

胞)上与II类MHC相关的抗原且接着产生介白素及刺激特异性B-细胞及T-细胞增殖及活性的其他物质。

除充当引发T细胞相互作用、发育及增殖的APC之外，巨噬细胞亦涉及于细胞介导免疫的表达中，因为其通过细胞介导免疫反应中所产生的IFN-γ活化。活化巨噬细胞具有增强的吞噬细胞潜能且释放导致炎症且破坏多种细菌及其他细胞的可溶性物质。天然杀伤细胞为可溶解带有新抗原的细胞(不管其是否为MHC型)，且甚至可溶解一些不带有MHC蛋白的细胞的细胞毒性细胞。天然杀伤T细胞或NK细胞依据其杀死显示外来抗原的细胞(例如肿瘤细胞)的能力来定义，无论是否为MHC型且无论是否预先敏化(暴露)于抗原。NK细胞可由IL-2及IFN-γ活化，且以与细胞毒性T淋巴细胞相同的方式溶解细胞。某些NK细胞具有IgG抗体的Fc结构域的受体(例如CD16或$F_C\gamma RIII$)且因此能够结合至靶标细胞表面上的IgG的Fc部分且释放经由抗体依赖性细胞介导的细胞毒性杀死靶标细胞的溶细胞组份。

另一组细胞为粒细胞或多形核白细胞(PMN)。嗜中性白细胞(PMN的一种类型)杀死细菌侵袭者且吞噬残留物。嗜伊红细胞为PMN的另一种类型且含有在对于另一种细胞(诸如外来细胞)释放时显示细胞毒性的颗粒。嗜碱细胞(PMN 的第三种类型)为强生理性反应(例如炎症)的重要介质粒，其通过释放多种生物学活性化合物(诸如组织氨、血清素、前列腺素及白三烯)来发挥其作用。所述所有细胞类型的共同点为在有机体内通常通过杀死且任选地清除有害组合物(诸如外来细胞)来发挥生理作用的能力。

尽管多种哺乳动物细胞(包括免疫系统的细胞)能够直接发挥生理作用(例如，以Tc、NK、某些PMN、巨噬细胞及其类似细胞为代表的细胞杀死作用)，但其他细胞间接促成生理作用。例如，抗原初始呈递至免疫系统的原始T细胞需要命令细胞-细胞接触的MHC呈现。此外，活化T细胞与抗原特异性B细胞之间通常需要接触以获得特定免疫原性反应。免疫反应中常见的细胞-细胞接触的第三种形式为活化B细胞与滤泡性树突状细胞之间的接触。所述细胞-细胞接触要求各自使生物活性剂靶向指定靶标变得复杂化。

补体依赖性细胞毒性(CDC)被认为是清除诸如肿瘤细胞的特定靶标细胞的重要机制。CDC 是由大量酶以级联方式彼此间活化组成的一系列事件。补体在清除抗原中具有重要作用，此由其四种主要功能实现：(1)局部血管扩张；(2)吸引免疫细胞，尤其巨噬细胞(趋化性)；(3)标记外来有机体用于吞噬作用(调理作用)；及(4)通过膜攻击复合物(MAC 攻击)破坏侵袭有机体。主要分子为 C3 蛋白。其为由经典途径或替代途径的组份裂解成两个片段的酶。经典途径通过抗体(尤其 IgG 及 IgM)诱发，而替代途径通过细菌产物(如脂多糖(LPS))非特异性地刺激。简而言之，C3 裂解的产物包括小肽 C3a，其对于吞噬型免疫细胞具有趋化性且通过促使 C5a 片段自 C5 中释放而导致局部血管扩张。C3 的另一部分 C3b 包覆外来有机体表面上的抗原且对有机体发挥调理作用以将其破坏。C3b 亦与补体系统的其他组份反应以形成由 C5b、C6、C7、C8 及 C9 组成的 MAC。

在人类治疗中存在与使用抗体相关的问题，此是因为免疫系统对任何抗原的反应，即使最简单者，亦为“多克隆”反应，亦即系统制造在其结合区以及其效应区中均具有大范围结构的抗体。

已使用两种方法以试图减少免疫原性抗体的问题。第一种方法为制备嵌合抗体，其中使小鼠单克隆抗体的抗原结合部分(可变区)与人类抗体的效应部分(恒定区)融合。在第二种方法中，已经由称为互补决定区(CDR)移植或“人源化”的技术来改变抗体。该方法经进一步改良以包括称为以下各者的变化：“重构”(Verhoeyen 等人，1988 *Science* 239:1534-1536; Riechmann 等人，1988 *Nature* 332:323-337; Tempest 等人，*Bio/Technol* 1991 9:266-271)、“超嵌合” (Queen 等人，1989 *Proc Natl Acad Sci USA* 86:10029-10033; Co 等人，1991 *Proc Natl Acad Sci USA* 88:2869-2873; Co 等人，1992 *J Immunol* 148:1149-1154)，及“面饰” (Mark 等人，于：Metcalf BW, Dalton BJ 编 Cellular adhesion: molecular definition to therapeutic potential. New York: Plenum Press, 1994:291-312 中)。

自 1986 年，即公布单克隆抗体后十一年时起，平均每年有不到一种治疗性抗体引入市场。自 1986 年至 1995 年的十年期间，有五种鼠

单克隆抗体引入人类药物中，包括用于器官移植体的急性排斥反应的“莫罗单抗-CD3(muromonab-CD3)” (OrthoClone OKT3®)；用于结肠直肠癌的“依决洛单抗(edrecolomab)” (Panorex®)；用于移植体排斥反应的“奥度莫单抗(odulimomab)” (Antilfa®)；及用于非霍奇金氏淋巴瘤(non-Hodgkin's lymphoma)的“替伊莫单抗(ibritumomab)” (Zevalin® yiuxetan)。此外，已市售单克隆Fab“阿昔单抗(abciximab)” (ReoPro®)用于预防冠状动脉再阻塞。亦已推出三种嵌合单克隆抗体：用于治疗B细胞淋巴瘤的“利妥昔单抗(rituximab)” (Rituxan®)；用于移植体排斥反应的“巴利昔单抗(basiliximab)”(Simulect®)；及用于治疗类风湿性关节炎及克罗恩氏病(Crohn's disease)的“英利昔单抗(infliximab)” (Remicade®)。此外，“阿昔单抗” (ReoPro®)(嵌合型人类-鼠单克隆抗体的47.6 kD Fab片段)是作为经皮冠状动脉介入术的佐剂市售以用于预防经历经皮冠状动脉介入术的患者的心脏缺血性并发症。最后，已推出七种“人源化”单克隆抗体。“达利珠单抗(Daclizumab)” (Zenapax®)是用于预防移植肾脏的急性排斥反应；“帕利珠单抗(palivizumab)” (Synagis®)是用于RSV；“曲妥珠单抗(trastuzumab)” (Herceptin®)结合HER-2(一种存在于乳癌细胞上的生长因子)；“吉妥珠单抗(gemtuzumab)” (Mylotarg®)是用于急性骨髓性白血病(AML)；且“阿仑单抗(alemtuzumab)” (MabCampath®)是用于慢性淋巴细胞性白血病；“阿达木单抗(adalimumab)” (Humira®(D2E7))是用于治疗类风湿性关节炎；且“奥马佐单抗(omalizumab)” (Xolair®)是用于治疗持续性哮喘。

因此，多种抗体技术已在研发及销售更有效治疗剂及缓和剂的工作中受到关注。遗憾的是，仍存在问题而破坏所述各种疗法的前景。例如，大部分经利妥昔单抗治疗的癌症患者通常在约6至12个月内复发，且已报导在利妥昔单抗输注24小时内发生致命性输注反应。亦已报导，在存在致命性结果的情况下需要透析的急性肾衰竭在以利妥昔单抗治疗中具有严重(有时致命)的粘膜皮肤反应。此外，静脉内注射需要高剂量的利妥昔单抗，此是因为该分子大(约150 kDa)且扩散进入存在有大量肿瘤细胞的淋巴组织受到限制。

曲妥珠单抗给药可导致发生心室机能不良、充血性心脏衰竭及严

重过敏性反应(包括重度过敏)、输注反应及肺部事件。达利珠单抗免疫抑制疗法使得发生淋巴组织增殖症及机会性感染的风险增加。已报导在接受吉妥珠单抗的患者中发生因肝衰竭所致的死亡、因严重肝脏毒性所致的死亡及因静脉闭塞疾病(VOD)所致的死亡。

亦报导接受阿仑单抗的患者体内的肝脏毒性。接受阿仑单抗疗法的患者已发生严重的(在某些罕见情况下为致命性的)全血细胞减少症/骨髓发育不全、自体免疫特发性血小板减少症及自体免疫性溶血性贫血。阿仑单抗亦可导致严重输注反应以及机会性感染。据报导，在经阿达木单抗治疗的患者中，严重感染及败血症(包括致命性严重感染及败血症)出现临床症状的恶化和/或脱髓鞘疾病的放射照相迹象，且在临床试验中经阿达木单抗治疗的患者患淋巴瘤的发病率比一般群体所预期的发病率更高。奥马佐单抗据报导可诱发恶性肿瘤及重度过敏。

癌症包括广泛范围的疾病，波及全世界范围内约四分之一的个体。恶性细胞的快速且不受调节的增殖为多种癌症类型(包括恶性血液病)的标志。尽管患有恶性血液病的患者在过去二十年中因癌症治疗的进步而受益(Multani 等人，1998 *J. Clin. Oncology* 16:3691-3710)，且症状缓解已成倍增加，但大多数患者的疾病仍会复发且死亡。对于细胞毒性药物医治的障碍包括例如肿瘤细胞耐受性及化疗的高毒性，所述障碍妨碍对诸多患者的最佳给药。

据报导，使用嵌合型 CD20 单克隆抗体治疗患有低度恶性或滤泡性 B 细胞淋巴瘤的患者可诱发患者的部分反应或全部反应。McLaughlin 等人，1996 *Blood* 88:90a(增刊 1，摘要)；Maloney 等人，1997 *Blood* 90:2188-95。然而，如上所述，肿瘤复发一般在六个月至一年内发生。需要进一步改良血清疗法以诱发对例如低度恶性 B 细胞淋巴瘤的更持久反应且使得可有效治疗高度恶性淋巴瘤及其他 B 细胞疾病。

另一种方法为使用对 CD20 具特异性的单克隆抗体来使放射性同位素靶向 B 细胞淋巴瘤。尽管据报导治疗有效性得以增加，但与放射性抗体的长体内半衰期相关的毒性亦增强，时常需要患者经历干细胞挽救。Press 等人，1993 *N. Eng. J. Med.* 329:1219-1224; Kaminski 等人，

1993 *N. Eng.J. Med.* 329:459-65。在连接放射性同位素之前，亦已用蛋白酶将 CD20 的单克隆抗体裂解以产生 $F(ab')_2$ 或 Fab 片段。据报导，此可改良放射性同位素结合体渗透至肿瘤中且可缩短体内半衰期，从而降低对正常组织的毒性。然而，所述分子缺乏效应功能，包括补体结合反应和/或 ADCC。

自体免疫疾病包括自体免疫甲状腺疾病(包括格雷氏病(Graves' disease)及桥本氏甲状腺炎(Hashimoto's thyroiditis))。仅在美国，约有两千万人患有某种形式的自体免疫性甲状腺病。自体免疫性甲状腺病是因产生刺激甲状腺而导致甲状腺功能亢进(格雷氏病)或破坏甲状腺而导致甲状腺功能减退(桥本氏甲状腺炎)的自体抗体所致。对甲状腺的刺激是因自身抗体结合且活化促甲状腺激素(TSH)受体所致。对甲状腺的破坏是因自身抗体与其他甲状腺抗原反应所致。格雷氏病的当前疗法包括手术、放射性碘或抗甲状腺药物疗法。放射性碘受到广泛使用，此是因为抗甲状腺药物具有显著副作用且疾病复发率高。手术专供患有大甲状腺肿的患者或需要极快速正常化甲状腺功能时所用。尚不存在靶向负责刺激 TSH 受体的自身抗体的产生的疗法。桥本氏甲状腺炎的当前疗法为左旋甲状腺素钠(levothyroxine sodium)，且由于症状缓解的可能性低而预期为终生疗法。抑制性疗法已证明可缩减桥本氏甲状腺炎的甲状腺肿，但尚未知减少自身抗体产生以靶向疾病机制的疗法。

类风湿性关节炎(RA)为以关节发炎为特征的慢性疾病，其导致肿胀、疼痛及功能丧失。在美国估计有两百五十万人患有 RA。RA 是因包括初始感染或损伤、异常免疫反应及遗传因素在内的事件组合所致。尽管 RA 中存在自反应性 T 细胞及 B 细胞，但可利用检测聚集于关节中的高含量抗体(称为类风湿性因子)来诊断 RA。RA 的当前疗法包括用于控制疼痛且减缓疾病进展的多种药物。尚未发现可治愈该疾病的疗法。药物包括非类固醇消炎药(NSAIDS)及改变病情抗类风湿药物(DMARDS)。NSAIDS 适用于良性疾病，但无法阻止发展为严重 RA 的关节破坏及无力。NSAIDS 与 DMARDS 均与显著副作用相关。10 年来仅一种新 DMARD 来氟米特(Leflunomide)获准。来氟米特阻止产

生自身抗体，减少炎症且减缓 RA 进展。然而，该药物亦导致严重副作用，包括恶心、腹泻、落发、皮疹及肝损伤。

全身性红斑性狼疮症(SLE)为自体免疫疾病，其是因包括肾、皮肤及关节的多种器官中的血管发生反复性损伤所致。在美国估计逾500,000 人患有 SLE。在患有 SLE 的患者中，T 细胞与 B 细胞之间的不当相互作用导致产生攻击细胞核的自身抗体。所述抗体包括抗双链DNA及抗-Sm抗体。约一半SLE患者体内亦存在结合磷脂的自身抗体，且所述自身抗体造成血管损害及低血球计数。免疫复合物积聚于 SLE 患者的肾、血管及关节中，导致炎症及组织损害。尚未发现可治愈该疾病的 SLE 疗法。视疾病的严重程度而定，可使用 NSAIDS 及 DMARDS 治疗。利用血浆交换移除自身抗体的血浆清除术可使 SLE 患者得到暂时性改善。通常一致认为自身抗体导致 SLE，因此耗尽 B 细胞谱是的新颖疗法(当由前体产生新 B 细胞时可使免疫系统恢复)为 SLE 患者持久受益提供希望。

修格兰氏综合征(Sjogren's syndrome)为以身体水分产生腺体的破坏为特征的自体免疫疾病。修格兰氏综合征为最普遍的自体免疫病症之一，在美国估计多达 4 百万人患有此病。修格兰氏综合征患者中约一半亦患有结缔组织疾病，诸如 RA，而另一半患有原发性修格兰氏综合征、但未患其他并发性自体免疫疾病。包括抗细胞核抗体、类风湿性因子、抗胞衬蛋白及抗蕈毒碱受体在内的自身抗体通常存在于患有修格兰氏综合征的患者体内。常规疗法包括皮质类固醇，且其他更有效疗法亦有益。

免疫性血小板减少性紫癜(ITP)是因自身抗体结合至血小板且导致其破坏所致。ITP 的某些情况是因药物所致，而其他情况是与感染、怀孕或诸如 SLE 的自体免疫性疾病相关。所有情况中约一半可归类为特发性起源。ITP 的治疗是依据症状的严重程度来决定。尽管在大多数情况下提供免疫抑制性药物(包括皮质类固醇或静脉内输注免疫球蛋白以耗尽 T 细胞)，但在某些情况下无需治疗。常导致血小板数量增加的另一种治疗为移除脾脏，该器官破坏抗体包覆的血小板。对于严重情况的患者，使用更有效的免疫抑制性药物，包括环孢素

(cyclosporine)、环磷酰胺(cyclophosphamide)或硫唑嘌呤(azathioprine)。通过使患者血浆通过蛋白A管柱移除自身抗体可用作患有严重疾病的患者的第二线治疗。需要其他更有效的疗法。

多发性硬化症(MS)亦为自体免疫疾病。其特征为中枢神经系统的炎症及髓鞘的破坏，使得脑、脊髓及身体中的神经细胞纤维孤立。尽管MS的病因未知，但广泛相信自体免疫T细胞为该疾病发病机制的主要促成因素。然而，MS患者的脑脊髓液中存在高含量的抗体，且某些人预言导致抗体产生的B细胞反应对于介导该疾病而言具有重要作用。尚未对MS患者研究B细胞耗尽疗法，且尚无MS治愈。当前疗法为皮质类固醇，其可减少攻击的持续时间及严重程度，但不影响MS随时间的进程。最近MS的新颖生物技术干扰素(IFN)疗法已获批准，但仍需要其他更有效的疗法。

重症肌无力(MG)为以任意肌肉群虚弱为特征的慢性自体免疫神经肌肉失调症。在美国约40,000人患有MG。MG是因自身抗体结合至在神经肌肉接点处经表达的乙酰胆碱受体所致。自身抗体减少或阻断乙酰胆碱受体，阻止信号由神经传递至肌肉。尚未知MG的治愈方法。普通治疗包括用皮质类固醇、环孢素、环磷酰胺或硫唑嘌呤进行免疫抑制。通常利用手术移除胸腺来减弱自体免疫反应。用于减少血液中的自身抗体含量的血浆清除术可有效用于MG，但因自身抗体持续产生而为短暂性的。血浆清除术一般供严重肌肉虚弱者在手术之前使用。新颖有效的疗法亦为有益的。

约五百万人患有牛皮癣，且其特征为皮肤的自体免疫性炎症。在30%患者中，牛皮癣亦与关节炎相关(牛皮癣性关节炎)。已使用多种疗法，包括类固醇、紫外光类视色素、维生素D衍生物、环孢素及甲氨喋呤(methotrexate)，但显然牛皮癣亦将受益于新颖且有效的疗法。硬皮症为结缔组织的慢性自体免疫疾病，亦称为全身性硬化。硬皮症特征为胶原蛋白的超量产生，导致皮肤变厚，且在美国约300,000人患有硬皮症，硬皮症亦将受益于新颖且有效的疗法。

由以上论述显而易见，需要经改良的组合物及方法以治疗、改善或预防多种疾病、病症及状态，包括癌症及自体免疫疾病。

概述

本发明通过提供以下各者来满足本领域中的上述需要中的至少一者：含有至少两个特异性结合结构域的蛋白，其中该两个结构域是通过源自抗体分子的恒定亚区来连接，该抗体分子的 C-末端连接至本文中称为蝎形连接子的连接子；及编码所述蛋白的核酸；以及所述蛋白及核酸的制备、诊断性及治疗性用途。该恒定亚区包含源自免疫球蛋白 C_{H2} 结构域的结构域及优选源自免疫球蛋白 C_{H3} 结构域的结构域，但不含有源自或对应于免疫球蛋白 C_{H1} 结构域的结构域或区。先前认为将源自抗体的恒定区置于蛋白内部会干扰抗体功能，诸如效应功能，以此类推，通常将抗体的恒定区置于抗体链的羧基末端亦会干扰抗体功能。此外，于恒定亚区的 C-末端布置蝎形连接子(其可为免疫球蛋白铰链样肽)为不同于天然存在的免疫球蛋白构成的构成。然而，根据本发明将恒定亚区(其中蝎形连接子连接恒定区的 C-末端)置于多肽链或蛋白链内部使得显示效应功能及多价(单特异性或多特异性)结合能力的蛋白相对不受空间位阻的阻碍。正如本领域技术人员在考虑本揭示内容后所显而易见，所述蛋白设计为模块且可如下被构建：通过在多种结合结构域中选择任何结构域作为结合结构域 1 或结合结构域 2(或作为存在于本发明的特定蛋白中的任何其他结合结构域)；通过选择具有效应功能的恒定亚区；且通过选择铰链样或非铰链样的蝎形连接子(例如 II 型 C-凝集素受体茎区肽)，其中蛋白显示 N-结合结构域 1-恒定亚区-蝎形连接子-结合结构域 2-C 的一般构成。本领域技术人员将进一步了解具有该结构的蛋白及编码此蛋白的核酸将存在多种应用，包括医学及兽医学应用。

本发明的一方面是关于具有效应功能或蝎形连接子(所述术语可互换使用)的多价单链结合蛋白，该结合蛋白包含源自免疫球蛋白(例如抗体)或免疫球蛋白样分子的第一结合结构域；提供效应功能的恒定亚区，该恒定亚区位于第一结合结构域的 C-末端；位于恒定亚区的 C-末端的蝎形连接子；及源自免疫球蛋白(诸如抗体)或免疫球蛋白样分子的第二结合结构域，该第二结合结构域位于恒定亚区的 C-末端；

从而使恒定亚区定位于第一结合结构域与第二结合结构域之间。单链结合蛋白的多特异性(例如双特异性)在于其可结合两个或两个以上不同靶标，或其可具有单特异性，具有两个针对同一靶标的结合位点。此外，该蛋白的所有结构域均存在于单链中，但该蛋白可形成同型多聚体，例如通过形成链间二硫键形成同型多聚体。在某些实施方案中，第一结合结构域与/或第二结合结构域是源自相同或不同免疫球蛋白(例如抗体)的轻免疫球蛋白链及重免疫球蛋白链的可变区。免疫球蛋白可来源于任何脊椎动物，诸如哺乳动物，包括人类，且可为天然存在的免疫球蛋白的嵌合型、人源化片段、变异体或衍生物。

本发明涵盖其中第一及第二结合结构域源自相同或不同免疫球蛋白(例如抗体)且其中第一及第二结合结构域识别相同或不同分子靶标(例如细胞表面标志物，诸如膜结合蛋白)的蛋白。此外，第一及第二结合结构域可识别相同或不同表位。第一及第二分子靶标可与第一及第二靶标细胞、病毒、载体和/或物体相关。在根据本发明的该方面的优选实施方案中，第一结合结构域、第二结合结构域及恒定亚区各自是源自人类免疫球蛋白，诸如 IgG 抗体。在其他实施方案中，具有效应功能的多价结合蛋白具有可识别至少一种无细胞分子靶标(例如与细胞不相关的蛋白，诸如沉积蛋白或可溶性蛋白)的第一结合结构域及第二结合结构域中的至少一者。例如，无细胞分子靶标包括与细胞从不相关的蛋白，例如所给予的化合物，诸如蛋白；以及经分泌、裂解、存在于外来体中或自细胞排出或分离的蛋白。

由第一及第二结合结构域所识别的靶标分子可存在于相同或不同的原核细胞、真核细胞、病毒(包括细菌性噬菌体)、有机或无机靶标分子载体及外来物体上或与所述物体相关。此外，该靶标分子可位于类型相同、实体不同的细胞、病毒、载体或物体上(例如两个不同的真核细胞、原核细胞、病毒或载体)，或该靶标分子可位于不同类型的细胞、病毒、载体或物体上(例如真核细胞及病毒)。靶标细胞为与由结合结构域所识别的靶标分子相关的那些细胞且包括内源细胞或自体细胞以及外源细胞或外来细胞(感染性微生物细胞、移植哺乳动物细胞(包括输血细胞))。本发明包括用于存在于与哺乳动物(诸如人类)的疾

病、病症或状态相关的靶标细胞表面上的第一和/或第二结合结构域的靶标。示例性靶标细胞包括癌细胞、与自体免疫疾病或病症相关的细胞及感染性细胞(例如感染性细菌)。亦涵盖感染性有机体(诸如哺乳动物寄生虫)的细胞作为靶标细胞。在某些实施方案中，本发明的蛋白为具有效应功能的多价(例如多特异性)结合蛋白，其中第一结合结构域及第二结合结构域中的至少一者识别选自由以下各者组成的组的靶标：肿瘤抗原、B-细胞靶标、TNF 受体超家族成员、Hedgehog 家族成员、受体酪氨酸激酶、蛋白聚糖相关分子、TGF-β 超家族成员、Wnt 相关分子、受体配体、T-细胞靶标、树突状细胞靶标、NK 细胞靶标、单核细胞/巨噬细胞靶标及血管生成靶标。

在上述蛋白的某些实施方案中，肿瘤抗原是选自由以下各者组成的组：鳞状细胞癌抗原 1(SCCA-1)、(蛋白 T4-A)、鳞状细胞癌抗原 2(SCCA-2)、卵巢癌抗原 CA125(1A1-3B) (KIAA0049)、粘蛋白 1(肿瘤相关粘蛋白)、(癌相关粘蛋白)、(多形性上皮粘蛋白)、(PEM)、(PEMT)、(上皮粘蛋白)、(肿瘤相关上皮细胞膜抗原)、(EMA)、(H23AG)、(花生反应性尿粘蛋白)、(PUM)、(乳癌相关抗原 DF3)、CTCL 肿瘤抗原 se1-1、CTCL 肿瘤抗原 se14-3、CTCL 肿瘤抗原 se20-4、CTCL 肿瘤抗原 se20-9、CTCL 肿瘤抗原 se33-1、CTCL 肿瘤抗原 se37-2、CTCL 肿瘤抗原 se57-1、CTCL 肿瘤抗原 se89-1、前列腺特异性细胞膜抗原、5T4 癌胚滋养层糖蛋白、Orf73 卡波西氏肉瘤相关疱疹病毒(Orf73 Kaposi's sarcoma-associated herpesvirus)、MAGE-C1(癌症/睾丸抗原 CT7)、MAGE-B1 抗原(MAGE-XP 抗原) (DAM10)、MAGE-B2 抗原(DAM6)、MAGE-2 抗原、MAGE-4a 抗原、MAGE-4b 抗原、结肠癌抗原 NY-CO-45、肺癌抗原 NY-LU-12 变异体 A、癌症相关表面抗原、腺癌抗原 ART1、副肿瘤相关脑-睾丸-癌症抗原(神经癌抗原 MA2；副肿瘤神经元抗原)、神经肿瘤腹侧抗原 2(NOVA2)、肝细胞癌抗原基因 520、肿瘤相关抗原 CO-029、肿瘤相关抗原 MAGE-X2、滑液肉瘤、X 断点 2、由 T 细胞识别的鳞状细胞癌抗原、血清学定义的结肠癌抗原 1、血清学定义的乳癌抗原 NY-BR-15、血清学定义的乳癌抗原 NY-BR-16、嗜铬粒蛋白 A；副甲状腺分泌蛋白 1、DUPAN-2、CA 19-9、CA 72-4、CA 195 及

L6。

上述方法的实施方案包含选自由以下各者组成的组的 B 细胞靶标：CD10、CD19、CD20、CD21、CD22、CD23、CD24、CD37、CD38、CD39、CD40、CD72、CD73、CD74、CDw75、CDw76、CD77、CD78、CD79a/b、CD80、CD81、CD82、CD83、CD84、CD85、CD86、CD89、CD98、CD126、CD127、CDw130、CD138 及 CDw150。

在上述方法的其他实施方案中，TNF 受体超家族成员是选自由以下各者组成的组：4-1BB/TNFRSF9、NGFR/TNFRSF16、BAFF R/TNFRSF13C、护骨素 / TNFRSF11B、BCMA/TNFRSF17、OX40/TNFRSF4、CD27/TNFRSF7、RANK/TNFRSF11A、CD30/TNFRSF8、RELT/TNFRSF19L、CD40/TNFRSF5、TACI/TNFRSF13B、DcR3/TNFRSF6B、TNF RI/TNFRSF1A、DcTRAIL R1/TNFRSF23、TNF RII/TNFRSF1B、DcTRAIL R2/TNFRSF22、TRAIL R1/TNFRSF10A、DR3/TNFRSF25、TRAIL R2/TNFRSF10B、DR6/TNFRSF21、TRAIL R3/TNFRSF10C、EDAR、TRAIL R4/TNFRSF10D、Fas/TNFRSF6、TROY/TNFRSF19、GITR/TNFRSF18、TWEAK R/TNFRSF12、HVEM/TNFRSF14、XEDAR、淋巴毒素 β R/TNFRSF3、4-1BB 配体/TNFSF9、淋巴毒素、APRIL/TNFSF13、淋巴毒素 β/TNFSF3、BAFF/TNFSF13C、OX40 配体/TNFSF4、CD27 配体/TNFSF7、TL1A/TNFSF15、CD30 配体/TNFSF8、TNF-α/ TNFSF1A、CD40 配体/TNFSF5、TNF-β/TNFSF1B、EDA-A2、TRAIL/TNFSF10、Fas 配体/TNFSF6、TRANCE/TNFSF11、GITR 配体/TNFSF18、TWEAK/TNFSF12 及 LIGHT/TNFSF14。

上述方法亦包括其中 Hedgehog 家族成员是选自由 Patched 及 Smoothened 组成的组的实施方案。在其他实施方案中，蛋白聚糖相关分子是选自由蛋白聚糖及其调节剂组成的组。

该方法的其他实施方案是关于其中受体酪氨酸激酶是选自由以下各者组成的组的方法：Axl、FGF R4、C1q R1/CD93、FGF R5、DDR1、Flt-3、DDR2、HGF R、Dtk、IGF-I R、EGF R、IGF-II R、Eph、INSRR、EphA1、胰岛素 R/CD220、EphA2、M-CSF R、EphA3、Mer、EphA4、

MSP R/Ron、EphA5、MuSK、EphA6、PDGF R α、EphA7、PDGF R β、EphA8、Ret、EphB1、ROR1、EphB2、ROR2、EphB3、SCF R/c-kit、EphB4、Tie-1、EphB6、Tie-2、ErbB2、TrkA、ErbB3、TrkB、ErbB4、TrkC、FGF R1、VEGF R1/Flt-1、FGF R2、VEGF R2/Flk-1、FGF R3及 VEGF R3/Flt-4。

在该方法的其他实施方案中，转化生长因子(TGF)-β 超家族成员是选自由以下各者组成的组：活化素 RIA/ALK-2、GFR α-1、活化素 RIB/ALK-4、GFR α-2、活化素 RIIA、GFR α-3、活化素 RIIB、GFR α-4、ALK-1、MIS RII、ALK-7、Ret、BMPR-IA/ALK-3、TGF-β RI/ALK-5、BMPR-IB/ALK-6、TGF-β RII、BMPR-II、TGF-β RIIb、内皮因子/CD105及 TGF-β RIII。

该方法的其他实施方案包含选自由以下各者组成的组的 Wnt 相关分子：Frizzled-1、Frizzled-8、Frizzled-2、Frizzled-9、Frizzled-3、sFRP-1、Frizzled-4、sFRP-2、Frizzled-5、sFRP-3、Frizzled-6、sFRP-4、Frizzled-7、MFRP、LRP 5、LRP 6、Wnt-1、Wnt-8a、Wnt-3a、Wnt-10b、Wnt-4、Wnt-11、Wnt-5a、Wnt-9a 及 Wnt-7a。

在该方法的其他实施方案中，该受体配体是选自由以下各者组成的组：4-1BB 配体/TNFSF9、淋巴毒素、APRIL/TNFSF13、淋巴毒素β/TNFSF3、BAFF/TNFSF13C、OX40 配体/TNFSF4、CD27 配体/TNFSF7、TL1A/TNFSF15、CD30 配体/TNFSF8、TNF-α/TNFSF1A、CD40 配体/TNFSF5、TNF-β/TNFSF1B、EDA-A2、TRAIL/TNFSF10、Fas 配体/TNFSF6、TRANCE/TNFSF11、GITR 配体/TNFSF18、TWEAK/TNFSF12、LIGHT/TNFSF14、双调蛋白(Amphiregulin)、NRG1同种型 GGF2、β 细胞生长因子(Betacellulin)、NRG1 同种型 SMDF、EGF、NRG1-α/HRG1-α、Epigen、NRG1-β1/HRG1-β1、表皮调节素(Epiregulin)、TGF-α、HB-EGF、TMEFF1/土莫调节素(Tomoregulin)-1、神经调节蛋白-3(Neuregulin- 3)、TMEFF2、IGF-I、IGF-II、胰岛素、活化素 A、活化素 B、活化素 AB、活化素 C、BMP-2、BMP-7、BMP-3、BMP-8、BMP-3b/GDF-10、BMP-9、BMP-4、BMP-15、BMP-5、Dpp 蛋白(Decapentaplegic)、BMP-6、GDF-1、GDF-8、GDF-3、GDF-9、

GDF-5、GDF-11、GDF-6、GDF-15、GDF-7、阿特明(Artemin)、纽特元(Neurturin)、GDNF、普塞芬(Persephin)、TGF-β、TGF-β2、TGF-β1、TGF-β3、LAP(TGF-β1)、TGF-β5、潜在性 TGF-β1、潜在性 TGF-βbp1、TGF-β1.2、Lefty、Nodal、MIS/AMH、酸性 FGF、FGF-12、碱性 FGF、FGF-13、FGF-3、FGF-16、FGF-4、FGF-17、FGF-5、FGF-19、FGF-6、FGF-20、FGF-8、FGF-21、FGF-9、FGF-23、FGF-10、KGF/FGF-7、FGF-11、神经菌毛素-1(Neuropilin-1)、PlGF、神经菌毛素-2、PlGF-2、PDGF、PDGF-A、VEGF、PDGF-B、VEGF-B、PDGF-C、VEGF-C、PDGF-D、VEGF-D 及 PDGF-AB。

在其他实施方案中，T-细胞靶标是选自由以下各者组成的组：2B4/SLAMF4、IL-2Rα、4-1BB/TNFRSF9、IL-2Rβ、ALCAM、B7-1/CD80、IL-4 R、B7-H3、BLAME/ SLAMF8、BTLA、IL-6R、CCR3、IL-7Rα、CCR4、CXCR1/IL-8RA、CCR5、CCR6、IL-10Rα、CCR7、IL-10Rβ、CCR8、IL-12Rβ1、CCR9、IL-12Rβ2、CD2、IL-13Rα1、IL-13、CD3、CD4、ILT2/CD85j、ILT3/CD85k、ILT4/CD85d、ILT5/CD85a、整合素 α4/CD49d、CD5、整合素 αE/CD103、CD6、整合素 αM/CD11b、CD8、整合素 αX/CD11c、整合素 β2/CD18、KIR/CD158、CD27/TNFRSF7、KIR2DL1、CD28、KIR2DL3、CD30/TNFRSF8、KIR2DL4/CD158d、CD31/PECAM-1、KIR2DS4、CD40 配体/TNFSF5、LAG-3、CD43、LAIR1、CD45、LAIR2、CD83、白三烯 B4 R1、CD84/SLAMF5、NCAM-L1、CD94、NKG2A、CD97、NKG2C、CD229/SLAMF3、NKG2D、CD2F-10/SLAMF9、NT-4、CD69、NTB-A/SLAMF6、共同 γ 链/IL-2Rγ、骨桥蛋白(Osteopontin)、CRACC/SLAMF7、PD-1、CRTAM、PSGL-1、CTLA-4、RANK/TNFRSF11A、CX3CR1、CX3CL1、L-选择素(L-选择素)、CXCR3、SIRPβ1、CXCR4、SLAM、CXCR6、TCCR/WSX-1、DNAM-1、胸腺生成素(Thymopoietin)、EMMPRIN/CD147、TIM-1、EphB6、TIM-2、Fas/TNFRSF6、TIM-3、Fas 配体/TNFSF6、TIM-4、FcγRIII/CD16、TIM-6、GITR/TNFRSF18、TNF RI/TNFRSF1A、颗粒溶素(Granulysin)、TNF RII/TNFRSF1B、HVEM/TNFRSF14、TRAIL R1/TNFRSF10A 、 ICAM-1/CD54 、 TRAIL R2/TNFRSF10B 、

ICAM-2/CD102 、 TRAIL R3/TNFRSF10C 、 IFN-γR1 、 TRAIL R4/TNFRSF10D、IFN-γR2、TSLP、IL-1 RI 及 TSLP R。

在其他实施方案中，NK 细胞受体是选自由以下各者组成的组：2B4/SLAMF4 、 KIR2DS4 、 CD155/PVR 、 KIR3DL1 、 CD94 、 LMIR1/CD300A 、 CD69 、 LMIR2/CD300c 、 CRACC/SLAMF7 、 LMIR3/CD300LF 、 DNAM-1 、 LMIR5/CD300LB 、 FcεRII 、 LMIR6/CD300LE、FcγRI/CD64、MICA、FcγRIIB/CD32b、MICB、FcγRIIC/CD32c 、 MULT-1 、 FcγRIIA/CD32a 、 柄蛋白 -2/CD112 、 FcγRIII/CD16、NKG2A、FcRH1/IRTA5、NKG2C、FcRH2/IRTA4、NKG2D、FcRH4/IRTA1、NKp30、FcRH5/IRTA2、NKp44、类 Fc 受体 3/CD16-2、NKp46/NCR1、NKp80/KLRF1、NTB-A/ SLAMF6、Rae-1、Rae-1α、Rae-1β、Rae-1δ、H60、Rae-1ε、ILT2/CD85j、Rae-1γ、ILT3/CD85k、TREM-1、ILT4/CD85d、TREM-2、ILT5/CD85a、TREM-3、KIR/CD158、TREML1/TLT-1 、 KIR2DL1 、 ULBP-1 、 KIR2DL3 、 ULBP-2 、 KIR2DL4/CD158d 及 ULBP-3。

在其他实施方案中，单核细胞/巨噬细胞靶标是选自由以下各者组成的组：B7-1/CD80、ILT4/CD85d、B7-H1、ILT5/CD85a、共同 β 链、整合素 α4/CD49d、BLAME/ SLAMF8、整合素 αX/CD11c、CCL6/C10、整合素 β2/CD18、CD155/PVR、整合素 β3/CD61、CD31/PECAM-1、胶乳素(Latexin)、CD36/SR-B3、白三烯 B4 R1、CD40/TNFRSF5、LIMPII/SR-B2、CD43、LMIR1/CD300A、CD45、LMIR2/CD300c、CD68、LMIR3/CD300LF 、 CD84/SLAMF5 、 LMIR5/CD300LB 、 CD97 、 LMIR6/CD300LE、CD163、LRP-1、CD2F-10/SLAMF9、MARCO、CRACC/SLAMF7、MD-1、ECF-L、MD-2、EMMPRIN/CD147、MGL2、内皮因子/CD105、骨活素(Osteoactivin)/GPNMB、FcγRI/CD64、骨桥蛋白、FcγRIIB/CD32b、PD-L2、FcγRIIC/CD32c、Siglec-3/CD33、FcγRIIA/CD32a、SIGNR1/CD209、FcγRIII/CD16、SLAM、GM-CSF Rα、TCCR/WSX-1、ICAM-2/CD102、TLR3、IFN-γR1、TLR4、IFN-γR2、TREM-1、IL-1 RII、TREM-2、ILT2/CD85j、TREM-3、ILT3/CD85k、TREML1/TLT-1、2B4/SLAMF4、IL-10 Rα、ALCAM、IL-10 Rβ、氨基

肽酶 N/ANPEP、ILT2/CD85j、共同 β 链、ILT3/CD85k、C1q R1/CD93、ILT4/CD85d、CCR1、ILT5/CD85a、CCR2、整合素 α4/CD49d、CCR5、整合素 αM/CD11b、CCR8、整合素 αX/CD11c、CD155/PVR、整合素 β2/CD18、CD14、整合素 β3/CD61、CD36/SR-B3、LAIR1、CD43、LAIR2、CD45、白三烯 B4 R1、CD68、LIMPII/SR-B2、CD84/SLAMF5、LMIR1/CD300A、CD97、LMIR2/CD300c、CD163、LMIR3/CD300LF、凝血因子 III/组织因子、LMIR5/CD300LB、CX3CR1、CX3CL1、LMIR6/CD300LE、CXCR4、LRP-1、CXCR6、M-CSF R、DEP-1/CD148、MD-1、DNAM-1、MD-2、EMMPRIN/CD147、MMR、内皮因子/CD105、NCAM-L1、FcγRI/CD64、PSGL-1、FcγRIII/CD16、RP105、G-CSF R、L-选择素、GM-CSF Rα、Siglec-3/CD33、HVEM/ TNFRSF14、SLAM、ICAM-1/CD54、TCCR/WSX-1、ICAM-2/CD102、TREM-1、IL-6R、TREM-2、CXCR1/IL-8 RA、TREM-3 及 TREML1/TLT-1。

在该方法的其他实施方案中，树突状细胞靶标是选自由以下各者组成的组：CD36/SR-B3、LOX-1/SR-E1、CD68、MARCO、CD163、SR-AI/MSR、CD5L、SREC-I、CL-P1/COLEC12、SREC-II、LIMPII/SR-B2、RP105、TLR4、TLR1、TLR5、TLR2、TLR6、TLR3、TLR9、4-1BB 配体/TNFSF9、IL-12/IL-23 p40、4-氨基-1,8-萘酰亚胺、ILT2/CD85j、CCL21/6Ckine、ILT3/CD85k、8-氧基-dG、ILT4/CD85d、8D6A、ILT5/CD85a、A2B5、整合素 α4/CD49d、Aag、整合素 β2/CD18、AMICA、郎罕细胞特异蛋白(Langerin)、B7-2/CD86、白三烯 B4 R1、B7-H3、LMIR1/CD300A、BLAME/SLAMF8、LMIR2/CD300c、C1qR1/CD93、LMIR3/CD300LF、CCR6、LMIR5/CD300LB、CCR7、LMIR6/CD300LE、CD40/TNFRSF5、MAG/Siglec-4a、CD43、MCAM、CD45、MD-1、CD68、MD-2、CD83、MDL-1/CLEC5A、CD84/SLAMF5、MMR、CD97、NCAM-L1、CD2F-10/SLAMF9、骨活素/GPNMB、Chem 23、PD-L2、CLEC-1、RP105、CLEC-2、Siglec-2/CD22、CRACC/SLAMF7、Siglec-3/CD33、DC-SIGN、Siglec-5、DC-SIGNR/CD299、Siglec-6、DCAR、Siglec-7、DCIR/CLEC4A、Siglec-9、DEC-205、Siglec-10、凝集素-1/CLEC7A、Siglec-F、凝集素-2/CLEC6A、

SIGNR1/CD209 、 DEP-1/CD148 、 SIGNR4 、 DLEC 、 SLAM 、 EMMPRIN/CD147 、 TCCR/WSX-1 、 FcγRI/CD64 、 TLR3 、 FcγRIIB/CD32b、TREM-1、FcγRIIC/CD32c、TREM-2、FcγRIIA/CD32a、TREM-3、FcγRIII/CD16、TREML1/TLT-1、ICAM-2/CD102 及辣椒素 R1(Vanilloid R1)。

在该方法的其他实施方案中，血管生成靶标是选自由以下各者组成的组: 血管生成素(Angiopoietin)-1、类血管生成素 2、血管生成素-2、类血管生成素 3、血管生成素-3、类血管生成素 7/CDT6、血管生成素-4、Tie-1、类血管生成素 1、Tie-2、血管生长素(Angiogenin)、iNOS、凝血因子 III/组织因子、nNOS、CTGF/CCN2、NOV/CCN3、DANCE、OSM、EDG-1、Plfr、EG-VEGF/PK1、增殖素(Proliferin)、内皮抑制素(Endostatin)、ROBO4、红血球生成素(Erythropoietin)、凝血栓蛋白-1、激肽抑制素(Kininostatin)、凝血栓蛋白-2、MFG-E8、凝血栓蛋白-4、一氧化氮(Nitric Oxide)、VG5Q、eNOS、EphA1、EphA5、EphA2、EphA6、EphA3、EphA7、EphA4、EphA8、EphB1、EphB4、EphB2、EphB6、EphB3、Ephrin-A1、Ephrin-A4、Ephrin-A2、Ephrin-A5、Ephrin-A3、Ephrin-B1、Ephrin-B3、Ephrin-B2、酸性 FGF、FGF-12、碱性 FGF、FGF-13、FGF-3、FGF-16、FGF-4、FGF-17、FGF-5、FGF-19、FGF-6、FGF-20、FGF-8、FGF-21、FGF-9、FGF-23、FGF-10、KGF/FGF-7、FGF-11、FGF R1、FGF R4、FGF R2、FGF R5、FGF R3、神经菌毛素-1、神经菌毛素-2、臂板蛋白 3A、臂板蛋白 6B、臂板蛋白 3C、臂板蛋白 6C、臂板蛋白 3E、臂板蛋白 6D、臂板蛋白 6A、臂板蛋白 7A、MMP、MMP-11、MMP-1、MMP-12、MMP-2、MMP-13、MMP-3、MMP-14、MMP-7、MMP-15、MMP-8、MMP-16/MT3-MMP、MMP-9、MMP-24/MT5-MMP、MMP-10、MMP-25/MT6-MMP、TIMP-1、TIMP-3、TIMP-2、TIMP-4、ACE、IL-13 Rα1、IL-13、C1q R1/CD93、整合素 α4/CD49d、VE-钙粘素、整合素 β2/CD18、CD31/PECAM-1、KLF4、CD36/SR-B3、LYVE-1、CD151、MCAM、CL-P1/COLEC12、粘连蛋白-2/CD112、凝血因子 III/组织因子、E-选择素、D6、P-选择素、DC-SIGNR/CD299 、 SLAM 、 EMMPRIN/CD147 、 Tie-2 、 内皮因子

/CD105、TNF RI/TNFRSF1A、EPCR、TNF RII/TNFRSF1B、红血球生成素 R、TRAIL R1/TNFRSF10A、ESAM、TRAIL R2/TNFRSF10B、FABP5、VCAM-1、ICAM-1/CD54、VEGF R2/Flk-1、ICAM-2/CD102、VEGF R3/Flt-4、IL-1 RI 及 VG5Q。

该方法的其他实施方案提供多价结合蛋白，其中结合结构域 1 及结合结构域 2 中的至少一者特异性地结合选自由以下各者组成的组的靶标：前列腺特异性细胞膜抗原(叶酸水解酶 1)、表皮生长因子受体(EGFR)、用于高级糖基化终产物的受体(RAGE；亦称为高级糖基化终产物受体或 AGER)、IL-17A、IL-17F、P19(IL23A 及 IL12B)、Dickkopf-1(Dkk1)、NOTCH1、NG2(硫酸软骨素蛋白聚糖 4 或 CSPG4)、IgE(IgHE 或 IgH2)、IL-22R(IL22RA1)、IL-21、类淀粉 β 寡聚物(Ab 寡聚物)、类淀粉 β 前驱蛋白(APP)、NOGO 受体(RTN4R)、低密度脂蛋白受体相关蛋白 5(LRP5)、IL-4、肌肉抑制素(GDF8)、极晚期抗原 4、α4、β1 整合素(VLA4 或 ITGA4)、存在于白细胞上的 α4、β7 整合素，及 IGF-1R。例如，VLA4 靶标可由其中结合结构域 1 及结合结构域 2 中的至少一者为源自那他珠单抗(Natalizumab)(Antegren)的结合结构域的多价结合蛋白来识别。

在某些实施方案中，癌细胞为转型造血细胞或癌性造血细胞。在某些所述实施方案中，第一结合结构域及第二结合结构域中的至少一者识别选自由以下各者组成的组的靶标：B-细胞靶标、单核细胞/巨噬细胞靶标、树突状细胞靶标、NK-细胞靶标及 T-细胞靶标(其各自如本文中所定义)。此外，第一结合结构域及第二结合结构域中的至少一者可识别骨髓靶标，包括但不限于：CD5、CD10、CD11b、CD11c、CD13、CD14、CD15、CD18、CD19、CD20、CD21、CD22、CD23、CD25、CD27、CD29、CD30、CD31、CD33、CD34、CD35、CD38、CD43、CD45、CD64、CD66、CD68、CD70、CD80、CD86、CD87、CD88、CD89、CD98、CD100、CD103、CD111、CD112、CD114、CD115、CD116、CD117、CD118、CD119、CD120a、CD120b、CDw123、CDw131、CD141、CD162、CD163、CD177、CD312、IRTA1、IRTA2、IRTA3、IRTA4、IRTA5、B-B2、B-B8 及 B-细胞抗原受体。

本发明的其他实施方案是关于如本文所述的多价结合蛋白，其包含选自由以下各者组成的组的序列：SEQ ID NO: 2、4、6、103、105、107、109、332、333、334 及 345。其他实施方案是关于包含选自由以下各者组成的组的序列的多价结合蛋白：SEQ ID NO: 355、356、357、358、359、360、361、362、363、364 及 365。

在其他实施方案中，具有效应功能的多价及多特异性结合蛋白具有识别选自由 EPHB4-KDR 与 TIE-TEK 组成的组的靶标对的第一结合结构域与第二结合结构域。在所述实施方案中，该蛋白具有识别 EPHB4 的第一结合结构域与识别 KDR 的第二结合结构域，或识别 KDR 的第一结合结构域与识别 EPHB4 的第二结合结构域。类似地，该蛋白可具有识别 TIE 的第一结合结构域与识别 TEK 的第二结合结构域，或识别 TEK 的第一结合结构域与识别 TIE 的第二结合结构域。

在一相关方面中，本发明提供具有效应功能的多价结合蛋白，其中恒定亚区识别效应细胞 F_C 受体(例如 $F_C\gamma RI$、$F_C\gamma RII$、$F_C\gamma RIII$、$F_C\alpha R$ 及 $F_C\varepsilon RI$)。在特定实施方案中，恒定亚区识别选自由以下各者组成的组的效应细胞表面蛋白：CD2、CD3、CD16、CD28、CD32、CD40、CD56、CD64、CD89、$F_C\varepsilon RI$、KIR、凝血栓蛋白 R、NKG2D、2B4/NAIL 及 41BB。恒定亚区可包含源自相同或不同免疫球蛋白、抗体同型或等位基因变异体的 C_{H2} 结构域及 C_{H3} 结构域。在某些实施方案中，C_{H3} 结构域为截短型结构域且包含选自由以下各者组成的组的 C-末端序列：SEQ ID NO: 366、367、368、369、370 及 371。当连接子为源自免疫球蛋白的铰链样肽时，C_{H2} 结构域及蝎形连接子优选是源自相同类别或源自相同子类的免疫球蛋白。

本发明亦涵盖如下某些蛋白，所述蛋白进一步包含具有至少约 5 个氨基酸、与恒定亚区连接且与第二结合结构域连接的蝎形连接子，从而使蝎形连接子定位于恒定亚区与第二结合结构域之间。通常，蝎形连接子肽长度介于 5 至 45 个氨基酸之间。蝎形连接子包括源自免疫球蛋白铰链区(诸如 IgG1、IgG2、IgG3、IgG4、IgA 及 IgE 铰链区)的铰链样肽。优选地，铰链样蝎形连接子将保留至少一个在生理条件下能够形成链间二硫键的半胱氨酸。源自 IgG1 的蝎形连接子可具有 1

个半胱氨酸或 2 个半胱氨酸，且优选将保留对应于 IgG1 的 N-末端铰链半胱氨酸的半胱氨酸。在某些实方案中，蝎形连接子相对于同源免疫球蛋白铰链区而言延长，且在示例性实施方案中，包含选自由 SEQ ID NO: 351、352、353 及 354 组成的组的序列。亦涵盖非铰链样肽作为蝎形连接子，其限制条件为所述肽提供足够空间及柔性以得到能够形成两个结合结构域的单链蛋白，该单链蛋白是相对于更靠近中心定位的恒定亚区结构域而朝向各蛋白末端(N 及 C)定位。示例性非铰链样蝎形连接子包括源自 II 型 C-凝集素的茎区(诸如 CD69、CD72、CD94、NKG2A 及 NKG2D 的茎区)的肽。在某些实施方案中，蝎形连接子包含选自由 SEQ ID NO: 373、374、375、376 及 377 组成的组的序列。

所述蛋白亦可包含具有至少约 5 个氨基酸、与恒定亚区连接且与第一结合结构域连接的连接子，从而使该连接子定位于恒定亚区与第一结合结构域之间。在某些实施方案中，连接子是存在于恒定亚区与两个结合结构域中之一者之间，且该连接子可具有相同或不同的序列且可具有相同或不同的长度。

本发明的蛋白的恒定亚区提供至少一种效应功能。涵盖本领域中已知的与免疫球蛋白(例如抗体)相关的任何效应功能，诸如选自由以下各者组成的组的效应功能：抗体依赖性细胞介导的细胞毒性(ADCC)、补体依赖性细胞毒性(CDC)、体内相对延长的半衰期(相对于缺乏恒定亚区的相同分子)、FcR 结合性、蛋白 A 结合性及其类似效应功能。在某些实施方案中，本发明的蛋白在人类中的延长半衰期为至少 28 小时。当然，欲给予至非人类个体的蛋白在此非人类个体中将显示相对延长的半衰期，而在人类中则未必如此。

一般而言，本发明的蛋白(包括多肽和肽)对于第一结合结构域及第二结合结构域中的至少一者显示小于 10^{-9} M 或至少 10^{-6} M 的结合亲和性。

本发明的另一方面是关于包含如本文中所述的蛋白和药学上可接受的佐剂、载体或赋形剂的药学上的组合物。本领域中已知的任何佐剂、载体或赋形剂均适用于本发明的药学上的组合物中。

本发明的又一方面提供一种制备如上所述的蛋白的方法，该方法

包含：将编码蛋白的核酸引入宿主细胞且将该宿主细胞优选以至少 1 毫克/升的量在适于表达蛋白的条件下孵育，从而表达该蛋白。在某些实施方案中，该方法进一步包含如下分离该蛋白：将其与其在细胞内表达后所结合的至少一种蛋白分离。适于表达核酸以制备本发明的蛋白的宿主细胞包括(但不限于)选自由以下各者组成的组的宿主细胞：VERO 细胞、海拉细胞(HeLa cell)、CHO 细胞、COS 细胞、W138 细胞、BHK 细胞、HepG2 细胞、3T3 细胞、RIN 细胞、MDCK 细胞、A549 细胞、PC12 细胞、K562 细胞、HEK293 细胞、N 细胞、草地贪夜蛾(*Spodoptera frugiperda*)细胞、酿酒酵母(*Saccharomyces cerevisiae*)细胞、巴斯德毕赤氏酵母(*Pichia pastoris*)细胞、多种真菌细胞中的任何者及多种细菌细胞中的任何者(包括(但不限于)大肠埃希氏杆菌(*Escherichia coli*)、枯草杆菌(*Bacillus subtilis*)、鼠伤寒沙门氏杆菌(*Salmonella typhimurium*)及链霉菌(*Streptomycete*))。

本发明亦提供一种制备如上所述的编码该蛋白的核酸的方法，该方法包含：使编码源自免疫球蛋白可变区的第一结合结构域的聚核苷酸的 3'端与编码恒定亚区的聚核苷酸的 5'端共价连接，使编码蝎形连接子的聚核苷酸的 5'端与编码恒定亚区的聚核苷酸的 3'端共价连接，且使编码源自免疫球蛋白可变区的第二结合结构域的聚核苷酸的 5'端与编码蝎形连接子的聚核苷酸的 3'端共价连接，从而形成编码具有效应功能的多价结合蛋白的核酸。所述编码区各自可作为本发明的单链结构的部分而由连接子或铰链样肽的编码区分离。在某些实施方案中，该方法制备编码第一结合结构域的单链形式的聚核苷酸，该第一结合结构域包含选自由以下各者组成的组的序列：SEQ ID NO: 2(抗-CD20 可变区、V_L-V_H取向)、SEQ ID NO: 4(抗-CD28 可变区、V_L-V_H取向)及 SEQ ID NO: 6(抗-CD28 可变区、V_H-V_L取向)，然而要求异型多聚蛋白(包括天然抗体)必需由经分别编码的多肽组装而成。编码第一结合结构域的示例性聚核苷酸序列为包含 SEQ ID NO: 1、3 或 5 的任意者的聚核苷酸。

本发明的该方面亦提供制备编码核酸的方法，所述编码核酸进一步包含插于编码第一结合结构域的聚核苷酸与编码恒定亚区的聚核苷

酸之间的连接子聚核苷酸，该连接子聚核苷酸编码具有至少 5 个氨基酸的肽连接子。此外，所述方法制备进一步包含插于编码恒定亚区的聚核苷酸与编码第二结合结构域的聚核苷酸之间的连接子聚核苷酸的核酸，该连接子聚核苷酸编码具有至少 5 个氨基酸的肽连接子。优选地，所编码的肽连接子具有 5 个与 45 个之间的氨基酸。

存在于 BD1 与 EFD 之间或存在于 EFD 与 BD2 之间的连接子区的一致性源自以同源-Ig 超家族成员所识别的其他序列。在研发源自同源-Ig 超家族成员中所存在的现有序列的新颖连接子中，由于所述序列通常为用于使细胞表面受体经蛋白酶裂解以形成可溶性形式的底物，因此优选避免出现类似于位于 C-样结构域末端与跨膜结构域末端之间的那些序列段的序列段。在-Ig 超家族与亚家族的不同成员之间进行序列比对，可比较分子之间在连接多个 V-样结构域或连接 V 样结构域与 C 样结构域的连接子序列方面的类似性。由此分析可知，可出现保守型天然存在的序列模式；所述序列在用作多价融合蛋白的子结构域之间的连接子时应具有更强的蛋白酶耐受性，可促成 Ig 环区之间的适当折叠，且因其发生于内源性细胞表面分子的细胞外结构域中而不具有免疫原性。

所述核酸本身构成本发明的另一方面。涵盖编码本文所述的本发明任何蛋白的核酸。因此，本发明的核酸依 5'至 3'的次序包含第一结合结构域的编码区、恒定亚区序列及第二结合结构域的编码区。亦涵盖编码蛋白变异体的核酸，其中该两个结合结构域及恒定亚区序列在氨基酸序列方面与已知免疫球蛋白可变区序列及已知恒定亚区序列的组合序列总共具有至少 80%且优选至少 85%、90%、95%或 99%的一致性。或者，本发明的蛋白变异体通过在 65-68°C、0.015 M 氯化钠、0.0015 M 柠檬酸钠的严格杂交条件下或在 42°C、0.015 M 氯化钠、0.0015 M 柠檬酸钠及 50%甲酰胺的严格杂交条件下与编码本发明的非变异体蛋白的核酸杂交的核酸来编码。本发明的变异体核酸在上文刚刚定义的条件下显示杂交能力，或与编码本发明的非变异体蛋白的核酸显示 90%、95%、99%或 99.9%的序列一致性。

在相关方面中，本发明提供包含如上所述的核酸的载体以及包含

如本文中所述的载体或核酸的宿主细胞。可使用本领域中已知的任何载体(例如质粒、噬菌粒(phagemid)、噬菌粒(phasmid)、粘粒、病毒、人工染色体、穿梭载体及其类似物)，且本领域技术人员将了解何种载体特别适于指定目的。例如，在制备本发明的蛋白的方法中，选择可于所选宿主细胞中操作的表达载体。同样，涵盖能够经本发明的核酸或载体进行遗传转化的任何宿主细胞。优选的宿主细胞为高等真核生物宿主细胞，但亦涵盖低等真核生物(例如酵母)及原核生物(细菌)宿主细胞。

本发明的另一方面是关于一种诱导靶标细胞损伤的方法，该方法包含使靶标细胞与治疗有效量的如本文中所述的蛋白接触。在某些实施方案中，通过将蛋白或编码核酸给予有此需要的有机体来体内接触靶标细胞。本发明的该方面中涵盖其中多价单链结合蛋白可对靶标细胞诱导加成量的损伤的方法，此量为由因包含所述结合结构域中的一者或另一者的分离抗体所致的损伤总量所预计的损伤量。亦涵盖其中多价单链结合蛋白可对靶标细胞诱导协同量(与由包含第一结合结构域、但不包含第二结合结构域的第一抗体及包含第二结合结构域、但不包含第一结合结构域的第二抗体所诱导的损伤总量相比)的损伤的方法。在某些实施方案 中，多价单链结合蛋白具有多特异性且包含特异性地识别选自由以下各者组成的组的抗原对的结合结构域对：CD19/CD20、CD20/CD21、CD20/CD22、CD20/CD40、CD20/CD79a、CD20/CD79b、CD20/CD81、CD21/CD79b、CD37/CD79b、CD79b/CD81、CD19/CL II(亦即 II 类 MHC)、CD20/CL II、CD30/CL II、CD37/CL II、CD72/CL II 及 CD79b/CL II。

本发明的该方面亦包括其中多特异性多价单链结合蛋白可对靶标细胞诱导抑制量的损伤(与包含第一结合结构域、但不包含第二结合结构域的第一抗体及包含第二结合结构域、但不包含第一结合结构域的第二抗体所诱导的损伤总量相比)的方法。示例性实施方案包括其中多特异性多价单链结合蛋白包含特异性地识别选自由以下各者组成的组的抗原对的结合结构域对的方法：CD20/CL II、CD21/CD79b、CD22/CD79b 、 CD40/CD79b 、 CD70/CD79b 、 CD72/CD79b 、

CD79a/CD79b、CD79b/CD80、CD79b/CD86、CD21/CL II、CD22/CL II、CD23/CL II、CD40/CL II、CD70/CL II、CD80/CL II、CD86/CL II、CD19/CD22、CD20/CD22、CD21/CD22、CD22/CD23、CD22/CD30、CD22/CD37、CD22/CD40、CD22/CD70、CD22/CD72、CD22/79a、CD22/79b、CD22/CD80、CD22/CD86 及 CD22/CL II。

在一相关方面中，本发明提供一种治疗细胞增殖病症(例如癌症)的方法，该方法包含将治疗有效量的蛋白(如本文中所述)或编码核酸给予有此需要的有机体。本领域技术人员(包括医学专业人员及兽医学专业人员)熟知对于需要治疗的有机体的鉴别。本发明所涵盖的易经受治疗的病症包括选自由癌症、自体免疫病症、劳氏(Rous)肉瘤病毒感染及炎症组成的组的病症。在某些实施方案中，该蛋白是通过体内表达如本文中所述的编码蛋白的核酸来给药。本发明亦包括通过选自由以下各者组成的组的途径来给予蛋白：静脉内注射、动脉内注射、肌肉内注射、皮下注射、腹膜内注射及直接组织注射。

本发明的另一方面是关于一种改善与细胞增殖病症相关的症状的方法，该方法包含将治疗有效量的如本文中所述的蛋白给予有此需要的有机体。本领域技术人员亦熟知对于显示易经受改善的症状的那些病症或疾病或状态的鉴别。在某些实施方案中，该症状是选自由疼痛、发热、肿胀及关节僵硬组成的组。

本发明的又一方面是关于一种治疗与感染物相关的感染的方法，该方法包含将治疗有效量的本发明的蛋白给予有此需要的患者，其中该蛋白包含特异性地结合感染物的靶标分子的结合结构域。根据本发明的该方面，易经受治疗的感染物包括原核细胞及真核细胞、病毒(包括噬菌体)、外来物及感染性有机体，诸如寄生虫(例如哺乳动物寄生虫)。

本发明的一相关方面是关于一种改善与感染物相关的感染症状的方法，该方法包含将有效量的本发明的蛋白给予有此需要的患者，其中该蛋白包含特异性地结合感染物的靶标分子的结合结构域。熟习医学及兽医学技术者将能够利用常规实验基于个例来判定蛋白的有效量。

本发明的又一相关方面为一种降低因感染物所致的感染风险的方法，该方法包含将预防有效量的本发明的蛋白给予具有患该感染的风险的患者，其中该蛋白包含特异性地结合感染物的靶标分子的结合结构域。熟习相关技术者将能够利用常规实验基于个例来判定蛋白的预防有效量。

本发明的另一方面是关于上述多价单链结合蛋白，其中第一结合结构域及第二结合结构域中的至少一者特异性地结合选自由以下各者组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86及主要组织相容性复合物II类肽。

在某些实施方案中，第一结合结构域及第二结合结构域中的一者特异性地结合 CD20，而在某些所述实施方案中，另一结合结构域特异性地结合选自由以下各者组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86及主要组织相容性复合物II类肽。例如，在一个实施方案中，第一结合结构域能够特异性地结合 CD20，而第二结合结构域能够特异性地结合例如 CD19。在另一实施方案中，第一结合结构域结合CD19，而第二结合结构域结合CD20。亦涵盖其中两个结合结构域均结合CD20的实施方案。

在根据本发明的该方面的某些其他实施方案中，第一结合结构域及第二结合结构域中的一者特异性地结合CD79b，而在某些所述实施方案中，另一结合结构域特异性地结合选自由以下各者组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86及主要组织相容性复合物II类肽。示例性实施方案包括不同的多特异性多价单链结合蛋白，其中第一结合结构域:第二结合结构域特异性地结合 CD79b:CD19 或CD19:CD79b。亦包括具有可识别 CD79b 的第一及第二结合结构域的多价结合蛋白。

在其他某些实施方案中，第一结合结构域及第二结合结构域中的一者特异性地结合主要组织相容性复合物II类肽，而在某些所述实施

方案中，另一结合结构域特异性地结合选自由以下各者组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86及主要组织相容性复合物II类肽。例如，在一个实施方案中，第一结合结构域能够特异性地结合主要组织相容性复合物II类肽，而第二结合结构域能够特异性地结合例如CD19。在另一实施方案中，第一结合结构域结合CD19，而第二结合结构域结合主要组织相容性复合物II类肽。亦涵盖其中两个结合结构域均结合主要组织相容性复合物II类肽的实施方案。

在根据本发明的该方面的其他实施方案中，第一结合结构域及第二结合结构域中的一者特异性地结合 CD22，而在某些所述实施方案中，另一结合结构域特异性地结合选自由以下各者组成的组的抗原：CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86及主要组织相容性复合物II类肽。例示性实施方案包括不同的多特异性多价单链结合蛋白，其中第一结合结构域:第二结合结构域特异性地结合 CD22:CD19 或CD19:CD22。亦包括具有可识别 CD22 的第一及第二结合结构域的多价结合蛋白。

本发明的一相关方面是关于上述多价单链结合蛋白，其中该蛋白对靶标细胞行为具有协同效应(相对于各结合结构域的效应的总和)。在某些实施方案中，该蛋白包含特异性地识别选自由以下各者组成的组的抗原对的结合结构域对：CD20-CD19、CD20-CD21、CD20-CD22、CD20-CD40、CD20-CD79a、CD20-CD79b及CD20-CD81。

本发明进一步包括如上所述的多价单链结合蛋白，其中该蛋白对靶标细胞行为具有加成效应(相对于各结合结构域的效应的总和)。根据本发明的该方面的实施方案包括包含特异性地识别选自由以下各者组成的组的抗原对的结合结构域对的多特异性蛋白：CD20-CD23、CD20-CD30、CD20-CD37、CD20-CD70、CD20-CD80、CD20-CD86、CD79b-CD37、CD79b-CD81、主要组织相容性复合物II类肽-CD30及主要组织相容性复合物II类肽-CD72。

本发明的又一相关方面为如上所述的多价单链结合蛋白，其中该

蛋白对靶标细胞行为具有抑制效应(相对于各结合结构域的效应的总和)。在某些实施方案中，该蛋白具有多特异性且包含特异性地识别选自由以下各者组成的组的抗原对的结合结构域：CD20-主要组织相容性复合物 II 类肽、CD79b-CD19、CD79b-CD20、CD79b-CD21、CD79b-CD22、CD79b-CD23、CD79b-CD30、CD79b-CD40、CD79b-CD70、CD79b-CD72、CD79b-CD79a、CD79b-CD80、CD79b-CD86、CD79b-主要组织相容性复合物 II 类肽、主要组织相容性复合物 II 类肽-CD19、主要组织相容性复合物 II 类肽-CD20、主要组织相容性复合物 II 类肽-CD21、主要组织相容性复合物 II 类肽-CD22、主要组织相容性复合物 II 类肽-CD23、主要组织相容性复合物 II 类肽-CD37、主要组织相容性复合物 II 类肽-CD40、主要组织相容性复合物 II 类肽-CD70、主要组织相容性复合物 II 类肽-CD79a、主要组织相容性复合物 II 类肽-CD79b、主要组织相容性复合物 II 类肽-CD80、主要组织相容性复合物 II 类肽-CD81、主要组织相容性复合物 II 类肽-CD86、CD22-CD19、CD22-CD40、CD22-CD79b、CD22-CD86及 CD22-主要组织相容性复合物 II 类肽。

本发明的另一方面为一种鉴别上述多价结合分子(诸如多特异性结合分子)的结合结构域中的至少一者的方法，该方法包含：(a)使抗同型抗体与特异性地识别第一抗原的抗体及特异性地识别第二抗原的抗体接触；(b)进一步使包含所述抗原中的至少一者的靶标与步骤(a)的组合物接触；及(c)测量靶标的活性，其中此活性是用于鉴别多价结合分子的结合结构域中的至少一者。在某些实施方案中，该靶标为病态细胞，诸如癌细胞(例如癌性 B-细胞)或产生自身抗体的 B-细胞。

在本发明的上述各方法中，预期该方法可进一步包含多种多价单链结合蛋白。在某些实施方案中，第一多价单链结合蛋白的结合结构域及第二多价单链结合蛋白的结合结构域可对靶标细胞诱导协同效应、加成效应或抑制效应，诸如对靶标细胞诱导协同量、加成量或抑制量的损伤。多种多价单链结合蛋白的协同效应、加成效应或抑制效应是通过将所述多种蛋白的效应与包含结合结构域中的一者的抗体及包含另一结合结构域的抗体的组合效应相比较来判定。

本发明的一相关方面是关于一种包含多种如上所述的多价单链结合蛋白的组合物。在某些实施方案中，该组合物包含多种多价单链结合蛋白，其中第一种多价单链结合蛋白的结合结构域及第二种多价单链结合蛋白的结合结构域能够对靶标细胞诱导协同效应、加成效应或抑制效应，诸如对靶标细胞诱导协同量、加成量或抑制量的损伤。

本发明进一步是关于一种包含上述组合物及药学上可接受的载体、稀释剂或赋形剂的药学上的组合物。此外，本发明包括一种试剂盒，该试剂盒包含如本文中所述的组合物及一组关于给予该组合物以对靶标细胞产生效应(诸如损伤靶标细胞)的说明书。

最后，本发明亦包括一种试剂盒，该试剂盒包含如本文中所述的蛋白及一组关于给予该蛋白以治疗细胞增殖病症或改善细胞增殖病症的症状的说明书。

通过参考以下详细说明(包括实施例)可更清楚地了解本发明的其他特征和优势。

附图简述

图 1 显示本发明所设想的多价单链分子的示意图。融合蛋白表达盒的个别子结构域通过图上的独立形状/方块指示。BD1 是指结合结构域 1，连接子 1 是指介于 BD1 与“效应功能结构域”(以 EFD 指示)之间的任何潜在连接子或铰链样肽。该子结构域通常为人类 IgG1 的 Fc 结构域的工程化形式，但可包括如本文中所定义的具有一或多种效应功能的其他子结构域。连接子 2 是指(若存在，则)存在于 EFD 的羧基末端与结合结构域 2 (BD2)之间的连接子序列。

图 2 显示 COS 细胞中所表达的非还原蛋白的蛋白质印迹。使蛋白分泌于培养基内，且在 48-72 小时后通过离心从经瞬间转染的细胞分离培养基上清液。将三十微升(30 μl)粗上清液载入各凝胶孔内。泳道标识：1-分子量标记，数字指示千道尔顿(kilodalton)；2-2H7-IgG-STD1-2E12 LH；3-2H7-IgG-STD1-2E12 HL，4-2H7-IgG-STD2-2E12 LH；5-2H7-IgG-STD2-2E12 HL；6-2E12 LH SMIP；7-2E12 HL SMIP；8-2H7 SMIP。“2H7”是指单链构建体，其中 BD1 编码 VLVH 取向的 CD20 特

异性结合结构域(2H7); “2E12”是指对 CD28 具特异性的结合结构域; -IgG-是指具有编码其中所有 C 均突变为 S 的序列(sss)的铰链的单链构建体，且 IgG1 的 CH2 结构域和 CH3 结构域含有排除 ADCC 效应功能与 CDC 效应功能的突变(P238S 与 P331S); “STD 1”是指插在 VL-VH 取向的 BD2 相邻处或 2E12 (V_L-V_H)相邻处的 20 氨基酸连接子(在图 7 中标识为“STD1=20aa”)。“STD1-HL”是指与刚才所述类似、但 BD2 V 区为 VH-VL 取向的如下构建体: 2H7-sssIgG(P238/331S)-20 氨基酸连接子-2E12 (V_H-V_L)。“STD2-LH”是指 2H7-sssIgG (P238/331S)-38 氨基酸连接子-2E12 (V_L-V_H); “STD2-LH”是指 2H7-sssIgG (P238/331SS)-38 氨基酸连接子-2E12 (V_H-V_L); “SMIP”是指小模组免疫药物; 且“H”通常是指 V_H，而“L”通常是指 V_L。除非另有说明，否则所有蛋白取向均为 N-末端至 C-末端取向。

图 3 显示两幅柱状图，其说明由 COS 细胞所表达的 2H7-sssIgG (P238S/P331S)-STD1-2e12 LH 衍生物及 2H7-sssIgG (P238S/P331S)-STD1-2e12 HL 衍生物的结合特性。所述实验是用粗培养物上清液、而非经纯化的蛋白来进行。将未经 16 倍稀释的培养物上清液的连续稀释液与表达 CD20 的细胞(WIL-2S)或表达 CD28 的细胞(CD28 CHO)一起孵育。将上清液的结合活性与测试相关单一特异性 SMIP 的结合性的对照样本(诸如 TRU-015 或 2e12 VLVH 或 2e12VHVL SMIP)相比较。各样本的结合是使用异硫氰酸萤光素(FITC)与山羊抗人类 IgG 的结合体、以 1:100 的稀释度来检测。

图 4 为显示图 3 中所测试的蛋白的蛋白 A 纯化型与 WIL2-S 细胞的结合模式的频度分布图。“TRU015”为对 CD20 具特异性的 SMIP。亦对两种具有效应功能的多特异性结合蛋白进行分析: “2H7-2E12 LH”具有 V_L-V_H 取向的对 CD28 具特异性的结合结构域 2; “2H7-2E12 HL”具有 V_H-V_L 取向的对 CD28 具特异性的结合结构域 2。测试所述蛋白各自在 5 μg/mL 下的结合性，且用 FITC 山羊抗人类 IgG、以 1:100 检测结合性。关于所测试分子的更完整说明，参见以上图 2 的说明。

图 5 显示两幅频度分布图，其说明具有效应功能的经蛋白 A 纯化的多特异性结合蛋白与表达 CD28 的 CHO 细胞的结合性。“2H7-2E12

LH”具有 V_L-V_H取向的对 CD28 具特异性的结合结构域 2；“2H7-2E12 HL”具有 V_H-V_L取向的对 CD28 具特异性的结合结构域 2。测试所述蛋白各自在 5 μg/mL 下的结合性，且用 FITC 山羊抗人类 IgG、以 1:100 检测结合性。关于所测试分子的更完整说明，参见图 2 的说明。

图 6A)显示标识连接恒定亚区与结合结构域 2 的连接子的表格。连接子是以名称、序列、序列识别号、序列长度及与结合结构域 2 融合的序列来标识。B)显示标识多种构建体的表格，所述构建体标识本发明的例示分子中的元件。除以名称标识多价结合分子外，亦就结合结构域 1(BD1)、恒定亚区(铰链及效应结构域或 EFD)、连接子(关于连接子的其他信息，参见图 6A)及结合结构域 2(BD2)而言来揭示那些分子中的元件。提供多种例示性多价结合蛋白的序列，且在图中以序列识别号来标识。其他多价结合蛋白具有经改变的元件或元件次序，可由所揭示的序列来预测序列的改变。

图 7 显示组合式柱状图，其说明经纯化蛋白以单一结合浓度结合至表达 CD20 的 WIL-2S 细胞及表达 CD28 的 CHO 细胞。“H1-H6”是指具有 H1-H6 连接子及 V_H-V_L 取向的 2e12 V 区的 2H7-sss-hIgG-Hx-2e12 分子。“L1-L6”是指具有 L1-L6 连接子及 V_L-V_H 取向的 2e12 V 区的 2H7-sss-hIgG-Lx-2e12 分子。所有分子均在 0.72 μg/mL 的浓度下进行测试，且使用 FITC 与山羊抗人类 IgG 的结合体、以 1:100 检测结合性。接着对各样本的平均萤光强度作图，其为所测试的两种靶标细胞类型相对于所测试的各多价构建体(L1-L6 或 H1-H6)的成对柱状图。

图 8 显示库马斯(Coomassie)染色非还原性及还原性 SDS-PAGE 凝胶的像片。所述凝胶显示 2H7-sss-hIgG-Hx-2e12 HL 蛋白上的变异体连接子序列/长度对显现于凝胶上的两种主要蛋白带的量的影响。

图 9 显示用与 2H7 特异性具有反应性的(a) CD28mIgG 或(b) Fab 所探测的[2H7-sss-hIgG-H6-2e12]融合蛋白及相关单一特异性 SMIP 的蛋白质印迹。结果显示 H6 连接子的存在导致产生裂解形式的多价构建体，其缺乏 CD28 结合特异性。

图 10 显示[TRU015-sss-IgG-Hx-2e12 HL] H1-H6 连接子形式的不

同连接子变异体的结合曲线。第一分图显示结合至表达 CD20 的 WIL-2S 细胞的结合曲线。第二分图显示不同形式结合至 CD28 CHO 细胞的结合曲线。所述结合曲线如下产生：连续稀释经蛋白 A 纯化的融合蛋白且使用 FITC 与山羊抗人类 IgG 的结合体以 1:100 来检测结合性。

图 11 显示概述具有变异体连接子 H1-H7 的 2H7-sss-IgG-2e12 HL 多特异性融合蛋白的 SEC 分离结果的表格。表中各列列举[2H7-sss-IgG-Hx-2e12-HL]融合蛋白的不同连接子变异体。所关注峰(POI)的滞留时间及以 POI 存在的融合蛋白的百分率及以其他形式存在的蛋白的百分率亦列入表中。亦列举分子是否裂解，裂解程度以定性方式指示，其中(是)、是、**是**、或否为四种可能的选择。

图 12 显示具有变异体连接子 H3、H6 及 H7 连接子的[2H7-sss-hIgG-Hx-2e12]多特异性融合蛋白与表达 CD20 或 CD28 的细胞的两条结合曲线图。将经蛋白 A 纯化的融合蛋白自 10 μg/ml 至 0.005 μg/ml 的连续稀释液与表达 CD20 的 WIL-2S 细胞或 CD28 CHO 细胞一起孵育。使用 FITC 与山羊抗人类 IgG 的结合体以 1:100 来检测结合性。分图 A 显示与 WIL-2S 细胞的结合，且分图 B 显示与 CD28 CHO 细胞的结合。

图 13 显示由用于图 12 的分子所产生的替代性结合测定的结果。在此情况下，首先使融合蛋白结合至表达 CD20 的 WIL-2S 细胞，且接着用 CD28mIgG (5 μg/ml)及 FITC 抗小鼠试剂以 1:100 来检测结合性。所述结果表明在同一分子中同时结合至 CD20 与 CD28。

图 14 显示使用另一种多特异性融合构建体变异体所获得的结果。在此情况下，可改变 BD2 的特异性，以便使用 G28-1 抗体的 V 区形成 CD37 特异性结合结构域。所显示的两幅图说明 CD20 和/或 CD37 抗体阻断[2H7-sss-IgG-Hx-G28-1]多特异性融合蛋白与表达 CD20 及 CD37 靶标的 Ramos 或 BJAB 细胞结合的相对能力。在与多特异性融合蛋白一起孵育之前，将各细胞类型与 CD20 特异性抗体(25 μg/ml)或 CD37 特异性抗体(10 μg/ml)或两种试剂(所述试剂为小鼠抗人类试剂)一起预孵育。接着用 FITC 山羊抗人类 IgG 试剂(预吸附于小鼠以排除

交叉反应性)以 1:100 检测多特异性融合蛋白的结合性。

图 15 显示用 BJAB 靶标细胞、PBMC 效应细胞且以 CD20-hIgG-CD37 特异性融合蛋白作为测试试剂所进行的 ADCC 测定的结果。该程序的完整说明参见适当实施例。该图为融合蛋白浓度相对于在各种剂量下对单一特异性 SMIP 试剂和 [2H7-sss-hIgG-STD1-G28-1] LH 与 HL 变异体所测试的%特异性杀死率的曲线图。对于所述单一特异性或多特异性单链融合蛋白中之一者的剂量-反应效应，对各数据系列作图。

图 16 显示列举共培养实验结果的表格，其中在 TRU 015、G28-1 SMIP、两种分子一起或[2H7-sss-IgG-H7-G28-1HL]变异体存在下培养 PBMC。使用 20 μg/ml 的融合蛋白，且孵育 24 小时或 72 小时。接着将样本用结合 FITC 的 CD3 抗体及结合 PE 的 CD19 或 CD40 特异性抗体染色，接着经受流式细胞术。接着将各闸中细胞的百分率列入表中。

图 17 显示在与[2H7-sss-hIgG-H7-G28-1 HL]分子或对照物单一 CD20 和/或 CD37 特异性 SMIP 单独或组合一起孵育 24 小时后对 B 细胞系细胞凋亡的影响的两幅柱状图。将膜联蛋白 V (annexin V)-碘化丙啶阳性细胞的百分率作为用于共孵育实验的测试试剂类型的函数来作图。分图 A 显示使用 Ramos 细胞所获得的结果，且分图 B 显示使用 Daudi 细胞所获得的结果。针对各单一 CD20 或 CD37 的 SMIP 是以指定浓度显示；此外，若使用两种试剂的组合，则以括号显示各种试剂的相对量。对于多特异性 CD20-CD37 融合蛋白以 5 μg/ml、10 μg/ml 和 20 μg/ml 的浓度进行测试。

图 18 显示[2H7-hIgG-G19-4]分子变异体和其与 CD3 表达细胞(Jurkats)或 CD20 表达细胞(WIL-2S)结合的两幅图。所述分子包括 [2H7-sss-hIgG-STD1-G19-4 HL]、LH 和 [2H7-csc-hIgG-STD1-G19-4 HL]。经蛋白 A 纯化的融合蛋白通过 20 μg/ml 滴定至 0.05 μg/ml，且使用 FITC 山羊抗人类 IgG 以 1:100 来检测结合性。将 MFI(平均萤光强度)作为蛋白浓度的函数来作图。

图 19 显示用具有 SSS 铰链或 CSC 铰链的[2H7-hIgG-STD1-G19-4 HL]分子变异体、BJAB 靶标细胞及作为效应细胞的全部人类 PBMC

或作为效应细胞的 NK 细胞耗尽型 PBMC 所进行的 ADCC 测定的结果。将杀死率作为多特异性融合蛋白浓度的函数来计分。将以所述分子所观测的杀死率与使用 G19-4、TRU 015 或该两种试剂的组合所见的杀死率相比较。相对于不同测试试剂对各数据系列作图，其中将%特异性杀死率作为蛋白浓度的函数来作图。

图 20 显示在抗-CD20 抗体存在或不存在下(若添加，则以 2 μg/ml 存在)与 B-细胞抗体(2 μg/ml)的基质组合中的各成员一起孵育隔夜后，经膜联蛋白 V(Ann)和/或碘化丙啶(PI)阳性染色的 Ramos B-细胞的百分率。山羊-抗小鼠二级抗体总是以相对于其他抗体(单独基质抗体，或基质抗体与抗-CD20 抗体)而言两倍的浓度比存在。竖直条纹柱—X 轴上表示的基质抗体(2 μg/ml)与山羊抗小鼠抗体(4 μg/ml)。水平条纹柱—X 轴上表示的基质抗体(2 μg/ml)、抗-CD20 抗体(2 μg/ml)及山羊抗小鼠抗体(4 μg/ml)。“第 2 步”条件充当对照且包括添加 4 μg/ml 的山羊抗小鼠抗体(竖直条纹柱)或 8 μg/ml(水平条纹柱)，而不添加基质抗体或抗-CD20 抗体。图中的“CL II”(II 类 MHC)是指与 HLA DR、DQ 及 DP(亦即，与 II 类 MHC 抗原)具有交叉反应性的单克隆抗体。

图 21 显示在抗-CD79b 抗体存在或不存在下(若添加，则以 0.5 或 1.0 μg/ml 存在)与 B-细胞抗体(2 μg/ml)的基质组合中的各成员一起孵育隔夜后，经膜联蛋白 V(Ann)和/或碘化丙啶(PI)阳性染色的 Ramos B-细胞的百分率。“CL II”及 “第 2 步”样本的鉴别参见图 20 的说明。竖直条纹柱—基质抗体(2 μg/ml)及山羊抗小鼠抗体(4 μg/ml)；水平条纹柱—基质抗体(2 μg/ml)、抗-CD79b 抗体(1.0 μg/ml)及山羊抗小鼠抗体(6 μg/ml)；点画柱—基质抗体(2 μg/ml)、抗-CD79b 抗体(0.5 μg/ml)及山羊抗小鼠抗体(5 μg/ml)。

图 22 显示在抗-CL II 抗体存在或不存在下(若添加，则以 0.25 或 0.5 μg/ml 存在)与 B-细胞抗体(2 μg/ml)的基质组合中的各成员一起孵育隔夜后，经膜联蛋白 V(Ann)和/或碘化丙啶(PI)阳性染色的 Ramos B-细胞的百分率。“CL II”及“第 2 步”样本的鉴别参见图 20 的说明。竖直条纹柱—基质抗体(2 μg/ml)及山羊抗小鼠抗体(4 μg/ml)；水平条纹柱—基质抗体(2 μg/ml)、抗-CL II 抗体(0.5 μg/ml)及山羊抗小鼠抗体(5

μg/ml)；点画柱—基质抗体(2 μg/ml)、抗-CL II 抗体(0.25 μg/ml)及山羊抗小鼠抗体(4.5 μg/ml)。

图 23 显示在抗-CD22 抗体存在或不存在下(若添加，则以 2 μg/ml 存在)与 B-细胞抗体(2 μg/ml)的基质组合中的各成员一起孵育隔夜后，经膜联蛋白 V(Ann)和/或碘化丙啶(PI)阳性染色的 DHL-4 B-细胞的百分率。“CL II”及“第 2 步”样本的鉴别参见图 20 的说明。实心柱—基质抗体(2 μg/ml)及山羊抗小鼠抗体(4 μg/ml)；斜条纹柱—基质抗体(2 μg/ml)、抗-CD22 抗体(2 μg/ml)及山羊抗小鼠抗体(8 μg/ml)。

图 24 提供的图表说明游离 CD20 SMIP(实心)或单特异性 CD20×CD20 蝎形分子(空心)对淋巴瘤细胞系 Su-DHL6(三角形)及 DoHH2(方形)的直接生长抑制作用。

图 25 的图表显示游离抗-CD37 SMIP(实心)或单特异性抗-CD37 蝎形分子(空心)对淋巴瘤细胞系 Su-DHL-6(三角形)及 DoHH2(方形)的直接生长抑制作用。

图 26 提供的图表显示两种不同的单特异性 SMIP(实心)的组合或双特异性 CD20-CD37 蝎形分子(空心)对淋巴瘤细胞系 Su-DHL-6(三角形)及 DoHH2(方形)的直接生长抑制作用。

图 27 的图表揭示游离 CD20 SMIP 与 CD37 SMIP 的组合(实心)或双特异性 CD20xCD37 蝎形分子(空心)对淋巴瘤细胞系 Su-DHL-6(三角形)及 WSU-NHL(方形)的直接生长抑制作用。

图 28 提供的直方图显示蝎形分子的细胞周期效应。DoHH2 淋巴瘤细胞的样本分为：未经处理、经 SMIP 016 处理或经单特异性 CD37×CD37 蝎形分子处理。空白柱：细胞周期的亚-G_1 期；黑色柱：G_0/G_1 期；阴影柱：S 期；及条纹柱：G_2/M 期。

图 29 提供的数据图表证明以蝎形分子处理淋巴瘤细胞使得信号传导能力与游离 SMIP 相比增强(如通过钙离子流所测量)。

图 30 提供的图表说明蝎形分子依赖性细胞细胞毒性。

图 31 显示的数据图表说明蝎形分子介导补体依赖性细胞毒性。

图 32 以图表形式提供的数据显示 SMIP 和蝎形分子与 FcγRIII(CD16)的低亲和性同种型(B)及高亲和性同种型(A)的比较性

ELISA 结合。

图 33 提供的图表证明在靶标细胞存在下，SMIP 和蝎形分子与 FcγRIII(CD16)的低亲和性等位基因型(A)及高亲和性等位基因型(B)的结合。

图 34 的直方图显示在两个实验(烧瓶 1 及烧瓶 2)中，在六种不同培养条件下，CD20×CD20 蝎形分子的表达程度。实心黑色柱：烧瓶 1；条纹柱：烧瓶 2。

图 35 提供的直方图显示 CD20×CD37 蝎形分子的产量。

图 36 提供 SMIP 和蝎形分子的 SDS-PAGE 凝胶影像(在还原及非还原条件下)。

图 37 提供的图表显示蝎形分子保持结合至靶标细胞的能力。填充方形：CD20 SMIP；填充三角形：CD37 SMIP；填充圆：人源化 CD20 (2Lm20-4)SMIP；空白菱形：CD37×CD37 单特异性蝎形分子；空心方形：CD20×CD37 双特异性蝎形分子；及空心三角形：人源化 CD20 (2Lm20-4)×人源化 CD20 (2Lm20-4)蝎形分子。

图 38 所含的图表显示竞争性结合测定的结果证明 N-末端与 C-末端蝎形分子结合结构域参与靶标细胞结合。

图 39 以图表形式提供的数据显示蝎形分子具有比 SMIP 更低的解离速率。

图 40 显示的图表证明蝎形分子在体内、在血清中具有稳定性，特征为蝎形分子可再生、循环半衰期持久。

图 41 提供 CD20 x CD37 双特异性蝎形分子的剂量-反应图表，其说明蝎形分子给药的体内功效。

图 42 显示单特异性 CD20xCD20 蝎形分子(S0129)和糖变异体与靶标 B-细胞的结合。

图 43 提供的图表说明 CD20xCD20 蝎形分子(亲本抗体及糖变异体)诱导 ADCC 介导的杀死 BJAB B-细胞。

图 44 显示的凝胶影像揭示因改变蝎形连接子(包括改变此连接子的序列及通过将 H7 序列添加至连接子来延长连接子)所导致的对蝎形分子稳定性的影响，此由凝胶下方 H7 线中的“+”来指示。

图 45 显示 CD20xCD20 蝎形分子(S0129)和其蝎形连接子变异体与 WIL2S 细胞的结合。

图 46 显示 CD20xCD20 蝎形分子和 CD20 SMIP 对多种 B-细胞的直接细胞杀死作用。

图 47 揭示单特异性 CD20xCD20 蝎形分子对其他 B-细胞系的直接细胞杀死作用。

图 48 显示两种单特异性蝎形分子(亦即 CD20×CD20 及 CD37×CD37)和双特异性 CD20xCD37 蝎形分子中每一者的直接细胞杀死能力，后者显示不同形式的杀死率曲线。

图 49 以图表形式描绘 Su-DHL-6 B-细胞对 CD20xCD20 (S0129)、CD37xCD37 及 CD20xCD37 蝎形分子中每一者的反应。

图 50 显示双特异性 CD19xCD37 蝎形分子和 Rituxan®直接杀死 Su-DHL-6 B-细胞的能力。

图 51 提供的直方图显示多种结合 CD20 的蝎形分子和 SMIP 以及 Rituxan®直接杀死 DHL-4 B-细胞(如图中所示)。蓝色柱：活细胞；每一对右侧的粟色柱：膜联蛋白+/PI+。

图 52 提供的图表描述各种结合 CD20 的蝎形分子和 SMIP 以及 Rituxan®的直接细胞杀死作用(如图中所示)。

图 53 提供由各种结合 CD20 的蝎形分子和 SMIP(如图中所示)以及 Rituxan®所诱导的 ADCC 活性的图表。

图 54 提供由各种结合 CD20 的蝎形分子和 SMIP(如图中所示)以及 Rituxan®所诱导的 CDC 活性的图表。

图 55 提供的直方图显示 C1q 与结合至 Ramos B-细胞的 CD20 结合性蝎形分子结合的程度。

图 56 提供的 FACS 分析的散布图显示， CD20 结合性蝎形分子(2Lm20-4x2Lm20-4 和 011x2Lm20-4)和 Rituxan®导致线粒体膜电位相对于对照物下降(上图)；具有断裂性线粒体膜电位的细胞的百分率的直方图(断裂性 MMP：黑色柱)显示于下图中。

图 57 提供的直方图显示 CD20 结合性蝎形分子(2Lm20-4x2Lm20-4 及 011x2Lm20-4)、利妥昔单抗(Rituximab)、CD95

及对照物对半胱天冬酶 3(caspase 3)活化的相对不足。

图 58 提供以蛋白质印迹法对获自 B-细胞的聚(ADP-核糖)聚合酶及半胱天冬酶 3、7 及 9 进行四次分析的组合，显示任何所述蛋白因 CD20 结合性蝎形分子结合至细胞而小量降解。

图 59 为 B-细胞染色体 DNA 的凝胶电泳图，其显示因 CD20 结合性蝎形分子结合至细胞所致的破碎程度。

图 60 为用抗磷酸化酪氨酸抗体和抗-SYK 抗体中每一者所获得的免疫沉淀物的凝胶电泳图。免疫沉淀物是获自 B-细胞与 CD20 结合性蝎形分子接触的溶胞物，如图中所示。

图 61 提供 CD20 结合性蝎形分子与羟道诺红霉素(doxorubicin)、长春新碱(vincristine)及雷帕霉素(rapamycin)中每一者的组合疗法的组合指数曲线图。

发明详述

本发明提供相对小肽的组合物，所述相对小肽具有至少两个结合区或结合结构域，所述结合区或结合结构域可提供一或多种结合特异性、源自免疫球蛋白(诸如抗体)的可变结合结构域、相对于包含至少部分免疫球蛋白恒定区(亦即如本文中所定义的恒定亚区的来源)的效应结构域在末端安置；以及重组制备所述肽中所涉及的核酸、载体及宿主细胞；及在多种诊断性及治疗性应用中使用所述肽组合物的方法，所述应用包括治疗病症以及改善该病症的至少一种症状。所述肽组合物有利地将第二结合结构域 C-末端排列至效应结构域，此排列出乎意料地提供该肽中至少两个结合结构域在无空间位阻或较弱位阻下的结合，同时保持安置于中心的效应结构域的效应功能。

本发明的多价肽的第一结合结构域与第二结合结构域可相同(亦即，具有一致性或大体上一致性的氨基酸序列且具有单特异性)或不同(且具有多特异性)。尽管第一结合结构域及第二结合结构域在初级结构方面不同，但其可识别且结合靶标分子的相同表位且因此将具有单特异性。然而，在多种情况下，所述结合结构域将具有不同结构且将结合至不同结合位点，产生多价多特异性蛋白。此不同结合位点可存

在于单一靶标分子上或不同靶标分子上。在两个结合分子识别不同靶标分子的情况下，该靶标分子可存在于(例如)相同结构上或相同结构中(例如相同细胞的表面)，或该靶标分子可存在于独立结构或位置上或存在于独立结构或位置中。例如，本发明的多特异性结合蛋白可具有特异性地结合至不同细胞类型表面上的靶标分子的结合结构域。或者，一个结合结构域可特异性地结合至细胞表面上的靶标，且另一结合结构域可特异性地结合至发现与细胞无关的靶标，诸如细胞外结构性(基质)蛋白或游离(例如可溶性或基质)蛋白。

第一结合结构域及第二结合结构域是源自相同或不同免疫球蛋白的蛋白结构(诸如抗体分子)的一或多个区。第一结合结构域和/或第二结合结构域可显示与免疫球蛋白区域的序列一致的序列，或可为该序列的修饰形式以提供(例如)经改变的结合特性或经改变的稳定性。所述修饰于本领域中已知且包括直接导致特性改变(诸如结合性改变)的氨基酸序列改变，例如导致肽的二级结构改变或更高级结构改变的氨基酸序列改变。亦涵盖由并入非天然氨基酸(诸如非天然的常规氨基酸、非常规氨基酸及亚氨基酸)所获得的经修饰氨基酸序列。在某些实施方案中，改变序列导致翻译后加工改变，例如导致糖基化模式改变。

涵盖用于蝎形分子的源自免疫球蛋白或免疫球蛋白样多肽(例如受体)的多种结合结构域中的任意者。源自抗体的结合结构域包含 V_L 结构域及 V_H 结构域的 CDR 区(参见例如使用源自人源化抗体的结合结构域的上下文)。包含源自抗体的完整 V_L 及 V_H 结构域的结合结构域可以任一取向来组织。本发明的蝎形分子可具有本文中所述的任何结合结构域。对于具有至少一个识别 B-细胞的结合结构域的蝎形分子而言，例示性蝎形分子具有至少一个源自以下各者的结合结构域：CD3、CD10、CD19、CD20、CD21、CD22、CD23、CD24、CD37、CD38、CD39、CD40、CD72、CD73、CD74、CDw75、CDw76、CD77、CD78、CD79a/b、CD80、CD81、CD82、CD83、CD84、CD85、CD86、CD89、CD98、CD126、CD127、CDw130、CD138 或 CDw150。在某些实施方案中，蝎形分子为包含至少一个结合结构域的多价结合蛋白，该结合结构域具有选自由以下各者组成的组的序列：SEQ ID NO: 2、4、6、

103、105、107及109。在某些实施方案中，蝎形分子包含含有选自由任何以下各者组成的组的序列的结合结构域：SEQ ID NO: 332-345。在某些实施方案中，蝎形分子包含含有源自免疫球蛋白 V_L 及 V_H 结构域的序列的结合结构域，其中该序列是选自由 SEQ ID NO: 355-365 中任何者组成的组。本发明进一步涵盖包含结合结构域的蝎形分子，该结合结构域具有可由 SEQ ID NO: 355-365 中的任何者推断的序列、具有 V_L 与 V_H 的反向取向。

对于其中结合结构域中的任一者或两者源自免疫球蛋白的一个以上区(例如 Ig V_L 区及 Ig V_H 区)的实施方案而言，多个区可通过连接子肽连接。此外，可使用连接子将第一结合结构域与恒定亚区相连接。恒定亚区与第二结合结构域的连接(亦即结合结构域2朝向蝎形分子的C-末端安置)可通过蝎形连接子来达成。所述蝎形连接子优选具有约2至45个氨基酸或2至38个氨基酸或5至45个氨基酸。例如，H1连接子长度为2个氨基酸且STD2连接子长度为38个氨基酸。除一般考虑长度外，适用于本发明的蝎形分子中的蝎形连接子区包括选自由IgG、IgA、IgD及IgE铰链及其变异体组成的组的抗体铰链区。例如，蝎形连接子可为选自由人类IgG1、人类IgG2、人类IgG3及人类IgG4及其变异体组成的组的抗体铰链区。在某些实施方案中，蝎形连接子区具有用于形成链间二硫键的单一半胱氨酸残基。在其他实施方案中，蝎形连接子具有两个用于形成链间二硫键的半胱氨酸残基。在某些实施方案中，蝎形连接子区是源自免疫球蛋白铰链区或C-凝集素茎区且包含选自由以下各者组成的组的序列：SEQ ID NO: 111、113、115、117、119、121、123、125、127、129、131、133、135、149、151、153、155、157、159、161、163、165、167、169、231、233、235、237、239、241、243、245、247、249、251、253、255、257、259、261、263、265、267、269、271、273、275、277、279、281、287、289、297、305、307、309、310、311、313、314、315、316、317、318、319、320、321、322、323、324、325、326、327、328、329、330、331、346、351、352、353、354、373、374、375、376及377。更一般而言，提供源自铰链区的序列时，涵盖序列表中所标识的任何

氨基酸序列用作本发明的蝎形分子中的蝎形连接子。此外，源自 Ig 铰链的蝎形连接子为铰链样肽结构域，该肽结构域具有至少一个能够参与链间二硫键的游离半胱氨酸。优选地，源自 Ig 铰链肽的蝎形连接子保留有半胱氨酸，所述半胱氨酸对应于朝向该铰链的 N-末端安置的铰链半胱氨酸。优选地，源自 IgG1 铰链的蝎形连接子具有一个或具有两个对应于铰链半胱氨酸的半胱氨酸。此外，蝎形连接子为 II 型 C-凝集素分子的茎区。在某些实施方案中，蝎形分子包含具有选自由 SEQ ID NO: 373-377 组成的组的序列的蝎形连接子。

安置于中心的恒定亚区是源自免疫球蛋白的恒定区。恒定亚区通常是源自发明摘要中的免疫球蛋白的 C_H 区的 C_{H2} 部分，尽管其亦可源自 C_{H2}-C_{H3} 部分。恒定亚区任选地可源自免疫球蛋白的铰链-C_{H2} 或铰链-C_{H2}-C_{H3} 部分，将对应于 Ig 铰链区的肽置于恒定亚区的 N-末端且安置于恒定亚区与结合结构域 1 之间。此外，恒定亚区的部分可源自不同免疫球蛋白的 C_H 区。此外，可将对应于 Ig CH3 的肽截短，剩下选自由 SEQ ID NOS: 366-371 组成的组的 C-末端氨基酸序列。然而优选地，在其中蝎铰链为源自免疫球蛋白铰链的铰链样肽的实施方案中，蝎形连接子及恒定亚区是源自相同类型的免疫球蛋白。恒定亚区提供至少一种与免疫球蛋白的 C_H 区相关的活性，诸如抗体依赖性细胞介导的细胞毒性(ADCC)、补体依赖性细胞毒性(CDC)、蛋白 A 结合性、与至少一种 F_C 受体的结合性、可再生性检测的稳定性(相对于除不存在恒定亚区以外的本发明的蛋白而言)，及或许胎盘转移性，正如本领域技术人员所知，其中本发明的分子的继代转移为有利的。如同上述结合结构域，恒定亚区是源自至少一种免疫球蛋白分子且与至少一种免疫球蛋白的一或多个区域显示一致或大体上一致的氨基酸序列。在某些实施方案中，经由至少一种免疫球蛋白的该或所述序列对恒定亚区进行修饰(通过取代一或多个常规或非常规的非天然(例如合成)氨基酸或亚氨基酸)，从而产生可使二级结构或更高级结构改变、特性亦随其改变的初级结构，或可导致翻译后加工改变，诸如糖基化。

若所述结合结构域和恒定亚区与一或多种免疫球蛋白多肽显示一致或大体上一致的氨基酸序列，则对本发明的分子的翻译后修饰可产

生经修饰(相对于作为修饰基础的免疫球蛋白)的分子。例如，可利用本领域中已知的技术，以使得多肽糖基化模式相对于未经修饰(例如CHO)宿主细胞中的此多肽发生改变的方式修饰宿主细胞，例如 CHO 细胞。

在具备所述分子和体内重组制备其的方法的情况下，已开发出新颖靶向诊断及治疗的途径，例如以容许免疫系统的效应细胞(例如细胞毒性 T 淋巴细胞、天然杀伤细胞及其类似者)靶向募集至欲破坏或隔断的细胞、组织、药剂及外来物，诸如癌细胞及感染物。除使治疗性细胞定位于治疗部位外，所述肽适用于定位治疗性化合物，诸如经放射性标记的蛋白。此外，所述肽亦适用于清除有害组合物，例如通过使有害组合物(诸如毒素)与能够破坏或排除该毒素的细胞(例如巨噬细胞)结合。本发明的分子适用于调节结合伙伴分子(诸如细胞表面受体)的活性。此显示于图 17 中，其中本发明的分子使经由 CD20 和/或 CD37 的细胞凋亡信号传导明显增强。此信号传导的结果为靶标细胞死亡。其中排除限定细胞群为有益的疾病及状态包括感染性及寄生性疾病、炎性及自体免疫状态、恶性肿瘤及其类似疾病。本领域技术人员应了解细胞凋亡信号传导的增强方法不受限制。有丝分裂信号传导及导致限定细胞群的分化、活化或灭活的信号传导可通过本发明的分子经由适当选择结合伙伴分子来诱导。通过考虑本文中所用术语的以下表述定义将有利于对本发明揭示内容的进一步理解。

“单链结合蛋白”为共价连接的氨基酸的单一邻接排列，该链能够特异性地结合至一或多个结合伙伴，所述伙伴共用足以与单链结合蛋白可检测性地结合的结合位点的决定子。例示性结合伙伴包括蛋白、碳水化合物、脂质及小分子。

为便于说明，本发明的蛋白、多肽及肽的“衍生物”及“变异体”是就与本发明的蛋白和/或多肽和/或肽的差异来描述，此意谓所述衍生物及变异体(其是本发明的蛋白/多肽/肽)以特定方式不同于本发明的未衍生或非变异蛋白、多肽或肽。本领域技术人员应了解所述衍生物及变异体本身为本发明的蛋白、多肽及肽。

“抗体”被赋予与其在本领域中的含义一致的最广泛定义，且包括

能够结合至少一个结合伙伴(诸如蛋白性或非蛋白性抗原)的蛋白、多肽及肽。如本文中所使用的“抗体”包括任何物种的蛋白的免疫球蛋白超家族的成员、单链或多链组成的蛋白的免疫球蛋白超家族的成员，及所述分子的变异体、类似物、衍生物及片段。特定而言，“抗体”包括本领域中已知的抗体的任何形式，包括(但不限于)单克隆及多克隆抗体、嵌合型抗体、CDR 移植抗体、人源化抗体、单链可变片段、双特异性抗体、双功能抗体、抗体融合体及其类似抗体。

“结合结构域”为特异性地结合一或多个特异性结合伙伴的肽区域，诸如源自免疫球蛋白(例如抗体)的多肽片段。若存在多个结合伙伴，则此伙伴共用足以可检测性地结合该结合结构域的结合决定子。该结合结构域优选为邻接氨基酸序列。

“表位”在本文中所给定的一般含义为在与抗体特异性相互作用(例如通过结合)的物质(例如蛋白)上具有单一抗原位点，亦即抗原决定子。在免疫球蛋白(例如抗体)技术中已获得确定含义的其他术语，诸如“轻链可变区”、“重链可变区”、“轻链恒定区”、“重链恒定区”、“抗体铰链区”、“互补决定区”、“框架区”、“抗体同型”、“F_C 区”、 “单链可变片段”或“scFv”、“双功能抗体”、“嵌合体”、“CDR 移植抗体”、“人源化抗体”、“成形抗体”、“抗体融合体”及其类似术语，各自具有本领域中已知的确定含义，除非本文中另外特别说明。

除非本文中明确定义，否则本领域技术人员参考抗体技术所了解的术语具有本领域中所获得的含义。所述术语的实例为“V_L”及“V_H”，是指分别源自抗体轻链及重链的可变结合区；以及 C_L 和 C_H，是指“免疫球蛋白恒定区”，亦即分别源自抗体轻链或重链的恒定区，其中应了解后者恒定区可进一步分为 C_{H1}、C_{H2}、C_{H3} 及 C_{H4} 恒定区结构域，此视该区所来源的抗体同型(IgA、IgD、IgE、IgG、IgM)而定。CDR 意谓“互补决定区”。“铰链区”是源自插于抗体单链的 C_{H1} 区与 C_{H2} 区之间且连接该两个区的氨基酸序列，在本领域中已知“铰链区”以“铰链”形式向全抗体提供柔性。

“恒定亚区”为本文中定义的术语，其是指对应于或源自抗体的一或多个恒定区结构域的肽、多肽或蛋白序列。因此，恒定亚区可包括

以下任何结构域或所有结构域：C_{H1}结构域、铰链区、C_{H2}结构域、C_{H3}结构域(IgA、IgD、IgG、IgE及IgM)及C_{H4}结构域(IgE、IgM)。因此，如本文中所定义的恒定亚区可是指对应于抗体的完整恒定区或其部分的多肽区。通常，本发明的多肽的恒定亚区或编码核酸具有铰链、C_{H2}结构域及C_{H3}结构域。

“效应功能”为与抗体恒定区相关的功能或由抗体恒定区所提供的功能。例示性效应功能包括抗体依赖性细胞介导的细胞毒性(ADCC)、补体活化及补体依赖性细胞毒性(CDC)、F_C受体结合性及血浆半衰期增加以及胎盘转移。本发明组合物的效应功能具有可检测性；具有此功能的本发明组合物的特异性活性就此效应功能而言优选与野生型抗体的特异性活性大致相同，亦即相对于野生型抗体而言，多价结合分子的恒定亚区优选不损失任何效应功能。

“连接子”为肽或聚核苷酸，其联接或连接其他肽或聚核苷酸。通常，肽连接子为具有约2至50个氨基酸的寡肽，其中典型聚核苷酸连接子编码该肽连接子且因此长度为约6至150个核苷酸。连接子使第一结合结构域与恒定亚区结构域相连接。例示性肽连接子为$(Gly_4Ser)_3$。蝎形连接子是用于将恒定亚区的C-末端与第二结合结构域相连接。蝎形连接子可源自免疫球蛋白铰链区或源自II型C-凝集素的茎区，如下文中更详细描述。

“靶标”被赋予一种以上含义，使用环境定义各种情况下的明确含义。就其最狭窄意义而言，“靶标”为结合位点，亦即本发明的肽组合物的结合伙伴的结合结构域。就更广泛含义而言，“靶标”或“分子靶标”是指必定呈现结合位点的完整结合伙伴(例如蛋白)。特异性靶标(诸如“CD20”、“CD37”及其类似靶标)各自被赋予该术语在本领域中已获得的一般含义。“靶标细胞”为与本发明的靶标分子相关的任何原核细胞或真核细胞(无论健康或患病与否)。当然，亦发现与任何细胞均不相关的靶标分子(亦即无细胞靶标)，或与其他组成相关的靶标分子，诸如病毒(包括噬菌体)、有机或无机靶标分子载体及外来物。

与靶标分子相关的物质的实例包括自体性细胞(例如癌细胞或其他病态细胞)、感染物(例如感染性细胞及感染性病毒)及其类似物。靶

标分子可与去核细胞、细胞膜、脂质体、海绵体、凝胶、胶囊、锭剂及其类似物结合，无论何种预定用途(例如作为善意或无意识措施的结果而用于药物治疗，或进一步为生物恐怖威胁)，其均可用于递送、输送或定位靶标分子。“无细胞”、“无病毒”、“无载体”、“无物体”及其类似者是指未与特定组合物或物质结合的靶标分子。

“结合亲和性”是指本发明的肽组合物与其结合伙伴的非共价结合的强度。结合亲和性优选是指对于结合对成员之间吸引力的定量量测。

“佐剂”为增强或促进与其结合(诸如以包含活性剂及佐剂的药学上的组合物形式结合)的化合物的功能效应的物质。“赋形剂”为在药学上的组合物配制中用作稀释剂的惰性物质。“载体”通常为惰性物质，其是用于提供用以递送药学上的组合物的媒介物。

“宿主细胞”是指其中存在本发明的聚核苷酸、蛋白或肽的任何原核细胞或真核细胞。

将核酸或聚核苷酸“引入”宿主细胞意谓通过本领域中已知的任何方式使核酸或聚核苷酸进入该细胞，所述方式包括(但不限于)裸核酸/聚核苷酸或载体携带的核酸/聚核苷酸的体外盐介导的沉淀及其他转化形式、病毒介导的感染及任选地具有或不具有“辅助”分子的转导、弹道抛射体递送、偶联及其类似方式。

“孵育”宿主细胞意谓将该细胞维持在本领域中已知的适于指定用途(诸如基因表达)的环境条件下。所述条件(包括温度、离子强度、氧张力、二氧化碳浓度、营养物组成及其类似条件)已于本领域中所熟知。

“分离”化合物(诸如本发明的蛋白或肽)意谓将该化合物从与其天然结合的存在于诸如表达欲经分离的化合物的宿主细胞中的至少一种不同化合物分离，例如将含有该化合物的消耗培养基从生长于该培养基中的宿主细胞分离。

“有此需要的有机体”为具有罹患易经受治疗或易用本发明的组合物改善的任何疾病、病症或状态的风险或罹患任何该疾病、病症或状态的任何有机体，所述疾病、病症或状态包括(但不限于)各种癌症形式的任意者、多种自体免疫疾病的任意者、因放射性标记蛋白、肽及类似化合物所致的辐射中毒、摄入性或内部产生性毒素及其类似形式，

此在回顾整个揭示内容后将变得显而易见。有此需要的有机体优选为人类患者。

如本领域中所知，疾病症状的“改善”意谓疾病的该症状的严重程度可检测性地降低。例示性症状包括疼痛、发热、肿胀及关节僵硬。

除非上下文中明确说明，否则术语“蛋白”、“肽”及“多肽”在本文中可互换使用，其各自是指氨基酸的至少一个邻接链。类似地，除非上下文中明确说明特定且不可互换的含义，否则术语“聚核苷酸”、“核酸”及“核酸分子”可互换使用。

“药学上可接受的盐”是指本发明化合物的盐，其是衍生自所述化合物与有机酸或无机酸的组合(酸加成盐)或所述化合物与有机碱或无机碱的组合(碱加成盐)。

下文利用如上文所定义的术语对本发明的各种方面提供一般性说明。在一般性说明之后，描述工作实施例以为本文中所揭示的本发明的可操作性及适用性提供补充证明。

蛋白和多肽

在本发明的某些实施方案中，提供本文中所述的具有效应功能的多价结合蛋白(包括结合结构域-免疫球蛋白融合蛋白)中的任意者，其中具有效应功能的多价结合蛋白或肽包含两个或两个以上结合结构域多肽序列。结合结构域多肽序列中的每一者均能够结合或特异性地结合一个或多个靶标，诸如一个或多个抗原，其中一个或多个靶标或抗原可相同或可不相同。结合结构域多肽序列可源自抗原可变区，或其可源自免疫球蛋白样分子，例如以模仿免疫球蛋白分子的方式折叠的受体。结合结构域所来源的抗体可为多克隆抗体(包括单特异性多克隆抗体)、单克隆抗体(mAb)、重组抗体、嵌合抗体、人源化抗体(诸如CDR-移植抗体)、人类抗体、单链抗体、催化抗体及本领域中已知的抗体的任何其他形式，以及其片段、变异体或衍生物。在某些实施方案中，本发明的蛋白的结合结构域中的每一者均是源自免疫球蛋白的完整可变区。在优选实施方案中，结合结构域各自是基于人类 Ig 可变区。在其他实施方案中，蛋白是源自 Ig 可变区的片段。在所述实施方案中，优选地，各结合结构域多肽序列对应于指定 Ig 可变区的互补决

定区中每一者的序列。本发明亦涵盖对应于比指定 Ig 可变区的全部 CDR 更少的 CDR 的结合结构域，其限制条件为所述结合结构域保持特异性地结合至少一种靶标的能力。

具有效应功能的多价结合蛋白亦具有源自免疫球蛋白恒定区(优选抗体重链恒定区)、共价并置于具有效应功能的多价结合蛋白中的两个结合结构域之间的恒定亚区序列。

具有效应功能的多价结合蛋白亦具有使恒定亚区的 C-末端与结合结构域 2 的 N-末端连接的蝎形连接子。蝎形连接子不为螺旋肽且可源自抗体铰链区、源自连接免疫球蛋白结合结构域的区域或源自 II 型 C-凝集素的茎区。蝎形连接子可源自免疫球蛋白的野生型铰链区，诸如 IgG1、IgG2、IgG3、IgG4、IgA、IgD 或 IgE 铰链区。在其他实施方案中，本发明提供具有经改变的铰链的多价结合蛋白。适于包含于多价结合蛋白中的经改变铰链区的一种类别为涉及于具有野生型铰链的免疫球蛋白对应物分子中的链间二硫键形成的半胱氨酸残基(尤其本领域中已知的那些 Cys 残基)的数量改变的铰链类别。因此，蛋白可具有 IgG1 铰链，其中缺失三个能够参与链间二硫键形成的 Cys 残基中的一者。为指示经改变铰链的半胱氨酸亚结构，自 N-末端至 C-末端显示 Cys 亚序列。利用此识别系统，具有经改变 IgG 铰链的多价结合蛋白包括表征为 cxc、xxc、ccx、xxc、xcx、cxx 及 xxx 的铰链结构。Cys 残基可缺失，或可经氨基酸取代，从而形成保守性取代或非保守性取代。在某些实施方案中，半胱氨酸经丝氨酸置换。对于具有包含 IgG1 铰链的蝎形连接子的蛋白而言，对应于铰链半胱氨酸的半胱氨酸的数目减少为 1 或 2，优选地，对应于铰链半胱氨酸的那些半胱氨酸中的一者最靠近该铰链的 N-末端安置。

对于具有包含 IgG2 铰链的蝎形连接子的蛋白而言，可存在 0、1、2、3 或 4 个 Cys 残基。对于包含含有 1、2 或 3 个 Cys 残基的经改变 IgG2 铰链的蝎形连接子而言，涵盖 Cys 残基的所有可能子集。因此，对于具有一个 Cys 的所述连接子而言，多价结合蛋白可在铰链区中具有以下 Cys 基序：cxxx、xcxx、xxcx 或 xxxc。对于包含具有 2 或 3 个 Cys 残基的 IgG2 铰链变异体的蝎形连接子而言，涵盖经保留的 Cys 残

基与经取代(或缺失)的 Cys 残基的所有可能组合。对于具有包含经改变 IgG3 或经改变 IgG4 铰链区的蝎形连接子的多价结合蛋白而言，涵盖铰链区中的 Cys 残基由 1 减少至小于 Cys 残基的完整数目，无论该减少是经由缺失或经由保守性氨基酸或非保守性氨基酸(例如丝氨酸)取代。同样涵盖具有包含野生型 IgA、IgD 或 IgE 铰链的蝎形连接子的多价结合蛋白，亦涵盖具有数目减少的 Cys 残基(其数目自 0 延续至小于对应野生型铰链中所存在的 Cys 残基的总数的数目)的对应经改变铰链区。在具有 IgG1 铰链的某些实施方案中，保留该铰链的第一或 N-末端 Cys 残基。对于具有野生型铰链区或经改变铰链区的蛋白而言，预期所述多价结合蛋白为能够例如通过二硫键形成而形成同型多聚体(诸如二聚体)的单链分子。此外，如本文中所揭示，具有改变铰链的蛋白的铰链区末端可经修改，例如在指定区或结构域(诸如铰链结构域)的 N-末端、C-末端或两末端缺失或取代一或多个氨基酸残基。

在另一例示性实施方案中，恒定亚区是源自包含天然 IgD 铰链区或工程化 IgD 铰链区的恒定区。野生型人类 IgD 铰链具有一个可与天然 IgD 结构中的轻链形成二硫键的半胱氨酸。在某些实施方案中，此 IgD 铰链半胱氨酸经突变(例如缺失)以产生用作例如双特异性分子的结合结构域之间的连接区的经改变铰链。IgD 铰链中不产生不良铰链刚性的其他氨基酸变化或缺失或修改是属于本发明的范畴。其他物种的天然或工程化 IgD 铰链区亦属于本发明的范畴，来自非人类物种的人源化天然或工程化 IgD 铰链以及来自其他人类或非人类抗体同型的(其他非 IgD)铰链区(诸如美洲驼 IgG2 铰链)亦属于本发明的范畴。

本发明进一步包括与可源自对应于如上所述的已知铰链区(诸如 IgG1 铰链或 IgD 铰链)的铰链的蝎形连接子相连接的恒定亚区。该恒定亚区可含有(相对于野生型)经修饰或经修改的铰链区，其中已知参与链间二硫键连接的至少一个半胱氨酸残基以保守性取代(例如 Ser 取代 Cys)或非保守性取代方式置换为另一种氨基酸。该恒定亚区不包括对应于免疫球蛋白 C_{H1} 结构域的肽区或肽结构域。

可用作连接区的替代铰链及连接子序列是源自连接免疫球蛋白 V-样结构域或免疫球蛋白 C-样结构域的细胞表面受体的部分。亦涵盖 Ig

V-样结构域之间的区(其中细胞表面受体含有多个串联式 Ig V-样结构域)及 Ig C-样结构域之间的区(其中细胞表面受体含有多个串联式 Ig C-样区)作为连接区。铰链及连接子序列通常长为 5 至 60 个氨基酸，且不仅主要为具柔性，而且亦可提供更多刚性特征。此外，连接子通常提供便于结合结构域之间的空间位阻最小化的空间。优选地，所述铰链及连接子肽的结构主要为 α 螺旋，β 片状结构最少。优选序列在血浆及血清中具有稳定性且耐受蛋白水解切割。优选序列可含有天然存在的基序或经添加的基序，诸如赋予二硫键以稳定二聚体形成的 CPPC 基序。优选序列可含有一或多个糖基化位点。优选铰链及连接子序列的实例包括(但不限于)CD2、CD4、CD22、CD33、CD48、CD58、CD66、CD80、CD86、CD150、CD166 及 CD244 的 Ig V-样区与 Ig C-样区之间的结构域间区。

恒定亚区可源自骆驼科动物恒定区，诸如美洲驼或骆驼 IgG2 或 IgG3。

特定而言，涵盖具有源自任何 Ig 类别或源自任何 IgG 亚类，诸如 IgG1(例如人类 IgG1)的 C_{H2}-C_{H3} 区的恒定亚区。在优选实施方案中，恒定亚区与源自免疫球蛋白铰链的蝎形连接子均源自相同 Ig 类别。在其他优选实施方案中，恒定亚区与源自免疫球蛋白铰链的蝎形连接子均源自相同 Ig 亚类。恒定亚区亦可为源自任何 Ig 类别或亚类(诸如 IgG1，例如人类 IgG1)的 CH3 结构域，其限制条件为其与至少一种免疫球蛋白效应功能相关。

恒定亚区不对应于IgG类的完整免疫球蛋白恒定区(亦即 C_{H1}-铰链-C_{H2}-C_{H3})。恒定亚区可对应于其他类别的完整免疫球蛋白恒定区，亦涵盖具有突变尾段或缺失尾段的IgA恒定结构域(诸如IgA1铰链、IgA2铰链、IgA C_{H2} 及 IgA C_{H3} 结构域)作为恒定亚区。此外，任何轻链恒定结构域均可用作恒定亚区，例如 C_K 或任何 C_L。恒定亚区亦可包括有或无铰链的 JH 或 JK。恒定亚区亦可对应于工程化抗体，其中如本领域中所了解，例如一种环移植体已通过使用 IgG 框架进行选定的氨基酸取代来构建以产生一个对于天然 F_CR 以外的受体(CD16、CD32、CD64、$F_C\varepsilon R1$)的结合位点。此类型的一个例示性恒定亚区为一个经修

饰而具有 CD89 结合位点的 IgG C_{H2}-C_{H3} 区。

本发明的该方面提供一种具有效应功能的多价结合蛋白或肽，包含以下、基本上由以下组成或由以下组成：(a)一个安置于 N-端的源自免疫球蛋白的结合结构域多肽序列，融合或连接于(b)一个源自免疫球蛋白恒定区的恒定亚区多肽序列，其优选包括一个铰链区序列，其中该铰链区多肽可如本文中所述，且可包含以下、基本上由以下组成或由以下组成：例如一个替代铰链区多肽序列，该恒定亚区多肽序列又融合或连接于(c)一个安置于 C-端的源自免疫球蛋白的第二天然或工程化结合结构域多肽序列。

该安置于中心的源自免疫球蛋白恒定区的恒定亚区多肽序列能够具有至少一种选自由以下组成的组的免疫活性：抗体依赖性细胞介导的细胞毒性、CDC、补体固定、及 F_C 受体结合，且所述结合结构域多肽各自能够结合或特异性地结合靶标，诸如抗原，其中靶标可相同或不同，且可有效地存在于相同生理环境(例如相同细胞的表面)中或不同环境(例如不同细胞表面，细胞表面及无细胞的位置，诸如于溶液中)中。

本发明的该方面亦包括与如本文中所揭示的特定序列的具有效应功能的多价蛋白具有至少 80%且优选 85%、90%、95%或 99%一致性的显示效应功能的变异体蛋白或多肽。

聚核苷酸

本发明亦提供编码本发明的蛋白或肽的聚核苷酸(经分离或经纯化的聚核苷酸或纯聚核苷酸)、包含所述聚核苷酸的载体(包括克隆载体和表达载体)，及经本发明的聚核苷酸或载体转化或转染的细胞(例如宿主细胞)。在编码本发明的蛋白或多肽方面，聚核苷酸编码第一结合结构域、第二结合结构域及 F_C 结构域，所述结构域均源自免疫球蛋白，优选源自人类免疫球蛋白。各结合结构域可含有对应于全长可变区序列(重链和/或轻链)或对应于其部分序列的序列，其限制条件为各该结合结构域保留特异性结合的能力。F_C 结构域可具有对应于全长免疫球蛋白 F_C 结构域序列或对应于其部分序列的序列，其限制条件为 F_C 结构域显示至少一种如本文中所定义的效应功能。此外，所述结合

结构域中的每一者均可经由通常长度为至少 8 个氨基酸且优选至少 13 个氨基酸的连接子肽而连接至 F_C 结构域。优选连接子序列为基于 Gly_4Ser 基序的序列，诸如$(Gly_4Ser)_3$。

本发明亦包括具有效应功能的多价结合蛋白的变异体。变异体聚核苷酸与具有如本文中所述的定义序列的聚核苷酸中的一者具有至少 90%且优选 95%、99%或 99.9%的一致性，或其可与具有定义序列的那些聚核苷酸中的一者在 65-68°C、0.015 M 氯化钠、0.0015 M 柠檬酸钠或在 42°C、0.015 M 氯化钠、0.0015 M 柠檬酸钠及 50%甲酰胺的严格杂交条件下杂交。聚核苷酸变异体保留编码具有效应功能的多价结合蛋白的能力。

术语“严格”用于是指本领域中通常所理解为严格的条件。杂交严格性主要依据温度、离子强度及变性剂(诸如甲酰胺)的浓度来判定。用于杂交及洗涤的严格条件的实例为在 65-68°C 下、0.015 M 氯化钠、0.0015 M 柠檬酸钠，或在 42°C 下、0.015 M 氯化钠、0.0015 M 柠檬酸钠及 50%甲酰胺。参见 Sambrook 等人，Molecular Cloning: A Laboratory Manual，第 2 版，Cold Spring Harbor Laboratory, (Cold Spring Harbor, N.Y. 1989)。

亦可使用更严格的条件(诸如更高温度、更低离子强度、更高甲酰胺或其他变性剂)；然而，杂交速率将受到影响。在其中涉及去氧寡核苷酸的杂交的情况下，其他例示性严格杂交条件包括在 37°C(对于 14 碱基寡核苷酸而言)、48°C(对于 17 碱基寡核苷酸而言)、55°C(对于 20 碱基寡核苷酸而言)及 60°C(对于 23 碱基寡核苷酸而言)下于 6 倍 SSC、0.05%焦磷酸钠中洗涤。

在本发明的相关方面中，提供一种制备本发明的多肽或蛋白或其他构建体(例如，包括具有效应功能的多价结合蛋白或肽)的方法，该方法包含以下步骤：(a)将如本文中所述或所提供的宿主细胞在容许表达该构建体的条件下培养；及(b)将表达产物(例如具有效应功能的多价结合蛋白或肽)与宿主细胞或宿主细胞培养物分离。

构建体

本发明亦是关于各自包括本发明的聚核苷酸或核酸的载体及由已

知载体所制备的构建体，特别是关于适用于基因疗法的重组表达构建体，包括各种已知构建体中的任意者，包括递送构建体，其包括编码如本文中所提供的具有效应功能的多价(例如多特异性，包括双特异性)结合蛋白及多肽的任何核酸；是关于在遗传上经本发明的载体和/或其他构建体工程化的宿主细胞；且是关于通过重组技术给予包含编码具有效应功能的多价(例如多特异性，包括双特异性)结合蛋白或其片段或变异体的核酸序列的表达构建体或其他构建体的方法。

在用于基因疗法的情况下，在适当启动子的控制下，视构建体的性质(例如如上所述的启动子的类型)而定及视所要宿主细胞的性质(例如有丝分裂后末端分化或活跃分裂；例如可表达构建体维持为附加体或整合于宿主细胞基因组内)而定，本发明的各种构建体(包括具有效应功能的多价(例如多特异性)结合蛋白)实际上可表达于任何宿主细胞中，包括体内宿主细胞。

用于原核及真核宿主的适当克隆载体和表达载体描述于例如Sambrook 等人，*Molecular Cloning: A Laboratory Manual*，第 2 版，Cold Spring Harbor, NY, (1989)中。例示性克隆/表达载体包括(但不限于)克隆载体、穿梭载体及表达构建体，其可基于质粒、噬菌粒(phagemid)、噬菌粒(phasmid)、粘粒、病毒、人工染色体或适于扩增、转移和/或表达其中所含的本领域中已知的聚核苷酸的任何核酸媒介物。如本文中所述，在本发明的优选实施方案中，在已经本发明的核酸转染、转化或转导的哺乳动物细胞中进行重组表达。亦参见例如 Machida, CA., "Viral Vectors for Gene Therapy: Methods and Protocols (基因治疗的病毒载体: 方法和实验方案)"; Wolff, JA, "Gene Therapeutics: Methods and Applications of Direct Gene Transfer (基因疗法: 直接基因转移的方法和应用)" (Birkhauser 1994); Stein, U 及 Walther, W (编), "Gene Therapy of Cancer: Methods and Protocols (癌症的基因治疗: 方法和实验操作)" (Humana Press 2000); Robbins, PD (编), "Gene Therapy Protocols (基因治疗方案)" Humana Press 1997); Morgan, JR (编), "Gene Therapy Protocols (基因治疗方案)" (Humana Press 2002); Meager, A (编), "Gene Therapy Technologies, Applications and Regulations: From Laboratory to

Clinic (基因治疗技术、应用和调节：实验室到临床)" (John Wiley & Sons Inc. 1999); MacHida, CA 和 Constant, JG, "Viral Vectors for Gene Therapy: Methods and Protocols (基因治疗的病毒载体：方法和实验操作)" (Humana Press 2002); "New Methods Of Gene Therapy For Genetic Metabolic Diseases NIH Guide (用于遗传代谢病的基因治疗新方法 NIH 指南)"第 22 卷，第 35 期，1993 年 10 月 1 日。亦参见美国专利第 6,384,210 号、第 6,384,203 号、第 6,384,202 号、第 6,384,018 号、第 6,383,814 号、第 6,383,811 号、第 6,383,795 号、第 6,383,794 号、第 6,383,785 号、第 6,383,753 号、第 6,383,746 号、第 6,383,743 号、第 6,383,738 号、第 6,383,737 号、第 6,383,733 号、第 6,383,522 号、第 6,383,512 号、第 6,383,481 号、第 6,383,478 号、第 6,383,138 号、第 6,380,382 号、第 6,380,371 号、第 6,380,369 号、第 6,380,362 号、第 6,380,170 号、第 6,380,169 号、第 6,379,967 号及第 6,379,966 号。

表达构建体通常是源自质粒载体。一种优选构建体为经修饰的 pNASS 载体(Clontech, Palo Alto, CA)，其具有编码氨苄西林(ampicillin)抗性基因的核酸序列、多聚腺苷酸化信号及 T7 启动子位点。其他合适哺乳动物表达载体已熟知(参见例如 Ausubel 等人，1995; Sambrook 等人，同上；亦参见例如 Invitrogen, San Diego, CA 的目录；Novagen, Madison, WI; Pharmacia, Piscataway, NJ)。为促进具有效应功能的多价结合蛋白的产量增加(所述产量是在应用适当选择剂(例如甲氨喋呤)后因基因扩增所致)，目前可在合适调节控制下制备包括二氢叶酸还原酶(DHFR)编码序列的优选构建体。

通常，重组表达载体将如上所述包括复制起点及容许宿主细胞转化的可选择标志物，及源自经高表达的基因以直接转录下游结构性序列的启动子。载体与本发明的聚核苷酸的可操作性连接产生克隆构建体或表达构建体。例示性克隆/表达构建体含有至少一个与本发明的聚核苷酸操作性连接的表达控制元件，例如启动子。本发明的载体及克隆/表达构建体中亦涵盖其他表达控制元件，诸如增强子、因子特异性结合位点、终止子及核糖体结合位点。本发明的聚核苷酸的异源结构性序列是在适当时期与翻译起始序列及终止序列一起组装。因此，例

如，如本文中所提供的多价结合蛋白编码核酸可作为用于在宿主细胞中表达该蛋白的重组表达构建体而包括于各种表达载体构建体的任一者中。在某些优选实施方案中，所述构建体是包括于体内给药的制剂中。所述载体及构建体包括染色体 DNA 序列、非染色体 DNA 序列及合成 DNA 序列，例如 SV40 的衍生物；细菌质粒；噬菌体 DNA；酵母质粒；源自质粒与噬菌体 DNA、病毒 DNA(诸如牛痘、腺病毒、禽痘病毒及假狂犬病或如下所述的复制缺乏性逆转录病毒)的组合的载体。然而，可使用任何其他载体来制备重组表达构建体，且在优选实施方案中，该载体可复制且可存活于宿主中。

可将适当 DNA 序列例如通过各种程序插入载体中。一般而言，将 DNA 序列通过本领域中已知的程序插入适当限制性核酸内切酶切割位点中。涵盖用于克隆、DNA 分离、扩增及纯化的标准技术，用于涉及 DNA 连接酶、DNA 聚合酶、限制性核酸内切酶及其类似酶的酶促反应的标准技术及各种分离技术。多种标准技术描述于例如 Ausubel 等人(1993 *Current Protocols in Molecular Biology* (分子生物学现行实验方案), Greene Publ. Assoc. Inc. & John Wiley & Sons, Inc., Boston, MA); Sambrook 等人(1989 *Molecular Cloning* (分子克隆)，第 2 版，Cold Spring Harbor Laboratory, Plainview, NY); Maniatis 等人(1982 Molecular Cloning (分子克隆), Cold Spring Harbor Laboratory, Plainview, NY); Glover (编) (1985 *DNA Cloning* (DNA 克隆) *第 I 及 II 卷*，IRL Press, Oxford, UK); Hames 和 Higgins (编), (1985 *Nucleic Acid Hybridization* (核酸杂交), IRL Press, Oxford, UK); 及其他文献中。

使表达载体中的 DNA 序列与至少一个适当表达控制序列(例如组成性启动子或调节性启动子)操作性地连接以引导 mRNA 合成。所述表达控制序列的代表性实例包括如上所述的真核细胞或其病毒的启动子。启动子区可选自使用 CAT(氯霉素转移酶)载体或具有可选择标志物的其他载体的任何所要基因。真核启动子包括 CMV 快速早期启动子、HSV 胸苷激酶、早期及晚期 SV40、源自逆转录病毒的 LTR，及小鼠金属硫蛋白-I。选择适当载体及启动子是属于本领域普通技术人员的水平范围内，且本文中描述某些特别优选的重组表达构建体的制

备，所述重组表达构建体包含至少一个与编码本发明的蛋白或多肽的核酸操作性地连接的启动子或调节性启动子。

通过高等真核细胞转录编码本发明的蛋白及多肽的 DNA 可通过将增强子序列插入载体中来增强。实例包括位于复制起点近侧的 SV40 增强子(bp 100 至 270)、巨细胞病毒早期启动子增强子、复制起点近侧的多形瘤增强子及腺病毒增强子。

亦涵盖使用本发明的核酸的基因疗法(包含置换缺陷型基因或将新颖基因添加至细胞和/或组织中的策略)，且其正在研发以供用于治疗癌症、调整代谢失调及用于免疫疗法领域。本发明的基因疗法包括使用具有或不具有独立载体或递送媒介物或构建体的本发明的各种构建体以供治疗本文中所述的疾病、病症和/或状态。所述构建体亦可用作用于治疗或预防本文中所述的疾病、病症和/或状态的疫苗。DNA 疫苗例如利用编码免疫原性蛋白的聚核苷酸及核酸决定子以刺激免疫系统从而防御病原体或肿瘤细胞。所述策略可刺激后天性或先天性免疫，或可经由细胞因子表达而涉及免疫功能的修饰。体内基因疗法通常涉及将遗传物质直接注入患者或动物体内以治疗、预防或改善疾病或与疾病相关的症状。疫苗及免疫调节为全身性疗法。对于组织特异性体内疗法，诸如旨在治疗癌症的那些疗法，优选为定位基因递送和/或表达/靶向系统。靶向特异性组织的多种基因疗法载体于本领域中已知，且已研发出完全靶向特异性组织的程序，例如使用基于导管的技术，所有所述内容均涵盖于本文中。

本文中亦涵盖离体基因疗法，且所述方法涉及移除、遗传性修饰、扩增及再给予个体(例如人类患者)的自身细胞。实例包括用于癌症治疗的骨髓移植或淋巴祖细胞的遗传性修饰。离体基因疗法优选适用于处理易得到且在基因转移过程期间可存活于培养物中的细胞(诸如血液或皮肤细胞)。

适用的基因疗法载体包括腺病毒载体、慢病毒载体、腺相关病毒(AAV)载体、单纯性疱疹病毒(HSV)载体及逆转录病毒载体。基因疗法亦可使用“裸 DNA”、基于脂质体的递送法、基于脂质的递送法(包括与带正电荷的脂质连接的 DNA)、电穿孔法及弹道投射法来进行。

在某些实施方案(包括但不限于基因疗法实施方案)中，载体可为病毒载体，诸如逆转录病毒载体。Miller 等人，1989 *BioTechniques* 7:980；Coffin 和 Varmus, 1996 Retroviruses, *Cold Spring Harbor Laboratory Press*, NY。例如，逆转录病毒质粒载体所来源的逆转录病毒包括但不限于莫洛尼(Moloney)鼠白血病病毒、脾坏死病毒、逆转录病毒(诸如劳氏(Rous)肉瘤病毒)、哈维(Harvey)肉瘤病毒、禽类白血病病毒、长臂猿白血病病毒、人类免疫缺乏病毒、腺病毒、骨髓增殖性肉瘤病毒及乳腺肿瘤病毒。

逆转录病毒为可复制且可经由 DNA 中间体整合于宿主细胞的基因组中的 RNA 病毒。此 DNA 中间体或前病毒可稳定地整合于宿主细胞 DNA 中。根据本发明的某些实施方案，表达构建体可包含其中并入有编码外来蛋白的外来基因而非正常逆转录病毒 RNA 的逆转录病毒。当逆转录病毒 RNA 进入宿主细胞内、同时发生感染时，外来基因也被引入细胞中，且接着可整合于宿主细胞 DNA 中，仿佛其为逆转录病毒基因组的部分。在宿主内表达此外来基因导致表达外来蛋白。

已研发出用于基因疗法的大部分逆转录病毒载体系统是基于鼠逆转录病毒。所述逆转录病毒以两种形式存在：称为病毒粒子的游离病毒颗粒或整合于宿主细胞 DNA 中的原病毒。病毒粒子形式的病毒含有：逆转录病毒的结构性和酶促性蛋白(包括逆转录酶)、病毒基因组的两个 RNA 拷贝和含有病毒包膜糖蛋白的源细胞质膜的部分。逆转录病毒基因组被组织为四个主要区：长末端重复(LTR)，其含有为转录起始和终止所必需的顺式作用元件且位于编码基因的 5'与 3'；和三个编码 *gag*、*pol* 及 *env* 的基因。该三个基因 *gag*、*pol* 及 *env* 分别编码内部病毒结构、酶促性蛋白(诸如整合酶)及包膜糖蛋白(命名为 gp70 及 p15e)，其赋予病毒以感染性及宿主范围特异性以及“R”肽的未定功能。

出于与使用逆转录病毒(包括用于表达构建体中)相关的安全考虑，已研发出单独包装的细胞系及载体制备性细胞系。简而言之，此方法是利用两种组份(逆转录病毒载体和包装细胞系(PCL))的用途。逆转录病毒载体含有长末端重复(LTR)、待转移的外来 DNA 及包装序列(y)。此逆转录病毒载体因编码结构性和包膜蛋白的基因不包括于载体

基因组中而无法由自身再生。PCL 含有编码 *gag*、*pol* 及 *env* 蛋白的基因，但不含有包装信号“y”。因此，PCL 仅可自身形成空病毒粒子。在该一般方法中，逆转录病毒载体是被引入 PCL 中，从而形成载体制备性细胞系(VCL)。该 VCL 制造仅含有逆转录病毒载体的外来基因组的病毒粒子，且因此在先前被视为安全逆转录病毒载体以供治疗性使用。

“逆转录病毒载体构建体”是指在本发明的优选实施方案中能够引导所关注的一个或多个序列或基因(诸如编码多价结合蛋白的核酸序列)的表达的组合体。简而言之，逆转录病毒载体构建体必须包括 5'LTR、tRNA 结合位点、包装信号、第二链 DNA 合成的起点及 3'LTR。载体构建体中可包括多种异源序列，包括例如编码蛋白(例如细胞毒性蛋白、疾病相关抗原、免疫辅助分子或置换基因)的序列或本身用作分子的序列(例如用作核糖核酸酶或反义序列)。

本发明的逆转录病毒载体构建体可易于由多种逆转录病毒来建构，所述逆转录病毒包括例如 B、C 及 D 型逆转录病毒以及泡沫病毒和慢病毒(参见例如 RNA Tumor Viruses (RNA 肿瘤病毒)，第 2 版，Cold Spring Harbor Laboratory, 1985)。所述逆转录病毒易获自保藏处或保藏中心，诸如美国典型培养物保藏中心(“ATCC”; Rockville, Maryland)，或使用通常可利用的技术由已知来源分离而得。已知本文中所提供的揭示内容和标准重组技术，可易于利用任何上述逆转录病毒以组装或构建本发明的逆转录病毒载体构建体、包装细胞或生产者细胞(例如 Sambrook 等人，*Molecular Cloning: A Laboratory Manual* (分子克隆：实验室指南)，第 2 版，Cold Spring Harbor Laboratory Press, 1989; Kunkle, 1985 *Proc. Natl. Acad. Sci. (USA) 82*:488)。

适用于病毒载体中的启动子通常可包括但不限于：逆转录病毒 LTR; SV40 启动子; 及人类巨细胞病毒(CMV)启动子(描述于 Miller 等人，1989 *Biotechniques* 7: 980-990 中)或其他任何启动子(例如细胞性启动子，诸如真核细胞启动子，包括但不限于组蛋白、pol III 及 β-肌动蛋白启动子)。可使用的其他病毒启动子包括但不限于腺病毒启动子、胸苷激酶(TK)启动子及 B19 小病毒启动子。对于本领域技术人员，合适启动子的选择将经由本文中所含的教导显而易见，且可选自调节

启动子或如上所述的启动子。

使用逆转录病毒质粒载体可转导包装细胞系以形成生产者细胞系。可经转染的包装细胞的实例包括但不限于如 Miller, *Human Gene Therapy* (人基因治疗), 1:5-14 (1990)中所述的 PE501、PA317、ψ-2、ψ-AM、PA12、T19-14X、VT-19-17-H2、ψCRE、ψCRIP、GP+E-86、GP+envAm12 及 DAN 细胞系。载体可经由本领域中已知的任何方式来转导包装细胞。所述方式包括但不限于电穿孔法、使用脂质体及 $CaPO_4$ 沉淀。在一个选择中，可将逆转录病毒质粒载体囊封于脂质体中，或与脂质偶联，且接着给予宿主。

生产者细胞系产生感染性逆转录病毒载体颗粒，所述颗粒包括编码具有效应功能的多价结合蛋白的核酸序列。所述逆转录病毒载体颗粒接着可用于体外或体内转导真核细胞。经转导的真核细胞将表达编码该蛋白或多肽的核酸序列。可经转导的真核细胞包括但不限于胚胎干细胞以及造血干细胞、肝细胞、纤维母细胞、循环外周血单核细胞及多形核细胞(包括骨髓单核细胞)、淋巴细胞、肌母细胞、组织巨噬细胞、树突状细胞、库普弗细胞(Kupffer cell)、淋巴结及脾的淋巴细胞及网状内皮细胞、角化细胞、内皮细胞及支气管上皮细胞。

宿主细胞

本发明的另一方面提供一种经本发明的任何聚核苷酸或克隆/表达构建体转化或转染的宿主细胞，或含有本发明的任何聚核苷酸或克隆/表达构建体的宿主细胞。聚核苷酸及克隆/表达构建体可利用本领域中已知的任何方法(包括转化、转染及转导)而被引入合适细胞中。宿主细胞包括经历离体细胞疗法(包括例如离体基因疗法)的个体的细胞。本发明的一方面所涵盖的真核宿主细胞在具有本发明的聚核苷酸、载体或蛋白时，除包括个体自身的细胞(例如人类患者自身的细胞)以外，亦包括 VERO 细胞、海拉细胞(HeLa cell)、中国仓鼠卵巢(CHO)细胞系(包括能够修饰所表达的多价结合分子的糖基化模式的经修饰 CHO 细胞，参见已公开的美国专利申请第 2003/0115614 A1 号，其以引用方式并入本文中)、COS 细胞(诸如 COS-7)、W138、BHK、HepG2、3T3、RIN、MDCK、A549、PC12、K562、HEK293 细胞、HepG2 细

胞、N细胞、3T3细胞、草地贪夜蛾细胞(例如Sf9细胞)、酿酒酵母细胞及本领域中已知的适用于表达及任选地分离本发明的蛋白或肽的任何其他真核细胞。亦涵盖原核细胞，包括但不限于大肠杆菌(*Escherichia coli*)、枯草杆菌(*Bacillus subtilis*)、鼠伤寒沙门氏杆菌(*Salmonella typhimurium*)、链霉菌(Streptomycete)或本领域中已知的适于表达及任选地分离本发明的蛋白或肽的任何原核细胞。在自原核细胞分离蛋白或肽中，尤其涵盖可使用本领域中已知的用于从包涵体中萃取蛋白的技术。由本文中的教导可知，适当宿主的选择是在本领域技术人员的技术范畴内。

工程化宿主细胞可培养于常规培养基中，该培养基经修饰以适合活化启动子、选择转化体或扩增特定基因。经选择用于表达的特定宿主细胞的培养条件，诸如温度、pH值及其类似条件，对于一般本领域技术人员而言是显而易见的。亦可使用各种哺乳动物细胞培养系统来表达重组蛋白。哺乳动物表达系统的实例包括猴肾纤维母细胞的COS-7系(由Gluzman, 1981 *Cell* 23:175所描述)及能够表达相容性载体的其他细胞系，例如C127、3T3、CHO、海拉细胞系及BHK细胞系。哺乳动物表达载体将包含复制起点、合适启动子及任选的增强子，以及任何必需的核糖体结合位点、多聚腺苷酸化位点、剪接供体及受体位点、转录终止序列及5'侧翼非转录序列，例如如本文中关于多价结合蛋白表达构建体的制备所述。源自SV40剪接的DNA序列及多聚腺苷酸化位点可用于提供所需的非转录遗传元件。构建体导入宿主细胞内可通过本领域技术人员熟知的多种方法来实现，所述方法包括但不限于磷酸钙转染法、DEAE-葡聚糖介导转染法或电穿孔法(Davis等人，1986 *Basic Methods in Molecular Biology* (分子生物学基本方法))。

在一实施方案中，宿主细胞通过引导本发明的蛋白或多肽的表达的重组病毒构建体来转导。经转导的宿主细胞产生含有经表达的蛋白或多肽的病毒颗粒，该经表达的蛋白或多肽是源自在病毒萌芽期间经病毒颗粒合并的宿主细胞膜的部分。

药学上的组合物

在某些实施方案中，本发明的组合物(诸如包含编码如本文中所述

的该蛋白的聚核苷酸的多价结合蛋白或组合物)适于在足以容许所编码的蛋白在宿主细胞中经体内或体外表达的条件及时间下给药以用于基因疗法及其类似疗法。所述组合物可被配制为根据熟知方法给药的药学上的组合物。药学上的组合物通常包含一或多种重组表达构建体和/或所述构建体的表达产物以及药学上可接受的载体、赋形剂或稀释剂。所述载体在所用剂量及浓度下对受者应为无毒的。对于基于核酸的制剂或对于包含本发明的表达产物的制剂而言，可通过例如皮内、皮下、肌肉内或静脉内途径或通过本领域中已知的在给定的一组环境下合适的任何途径，每公斤体重给予约 0.01 μg 至约 100 mg。优选剂量例如为约 1 μg/kg 至约 1 mg/kg，特别优选为约 5 μg/kg 至约 200 μg/kg。

对于本领域技术人员显而易见，给药次数和频率将视宿主的反应而定。适于治疗性用途的药学上可接受的载体已熟知于药物技术中，且描述于例如 *Remingtons Pharmaceutical Sciences* (雷明顿药物科学), Mack Publishing Co.(A.R. Gennaro 编，1985)中。例如，可使用生理学 pH 值下的无菌盐水及磷酸盐缓冲盐水。药学上的组合物中可提供防腐剂、稳定剂、染料及其类似物。例如，可添加苯甲酸钠、山梨酸及对羟基苯甲酸的酯作为防腐剂。同上，在第 1449 页。此外，可使用抗氧化剂和悬浮剂。同上。本发明的化合物可以游离碱或盐形式使用，两种形式均视为属于本发明的范畴。

含有本发明的一或多种核酸构建体或对应于由所述核酸构建体所编码的产物的蛋白的药学上的组合物可为容许组合物给予患者的任何形式。例如，组合物可为固体、液体或气体(气雾剂)的形式。典型的给药途径包括但不限于经口、局部、肠胃外(例如经舌下或经颊)、舌下、直肠、阴道及鼻内。如本文中所使用的术语肠胃外包括皮下注射、静脉内、肌肉内、胸骨内、海绵体内、鞘内、耳道内、尿道内注射或输注技术。药学上的组合物可被配制以使得其中所含的活性成份在将该组合物给予患者后可被生物利用。将给予患者的组合物可采取一或多种剂量单位的形式，其中例如锭剂可为单一剂量单位，且气雾剂形式的本发明的一或多种化合物的容器可具有多数个剂量单位。

对于经口给药而言，可存在赋形剂和/或粘合剂。实例为蔗糖、高岭土、甘油、淀粉糊精、褐藻酸钠、羧甲基纤维素及乙基纤维素。可存在着色剂和/或调味剂。可使用包衣外壳。

组合物可为液体形式，例如酏剂、糖浆、溶液、乳液或悬浮液。举两个实例，液体可供经口给药或通过注射递送。当意欲经口给药时，优选组合物除含有一或多个结合结构域-免疫球蛋白融合构建体或表达产物外，亦含有甜味剂、防腐剂、染料/着色剂及增香剂中的一或多者。在意欲通过注射给药的组合物中，可包括界面活性剂、防腐剂、湿润剂、分散剂、悬浮剂、缓冲剂、稳定剂及等渗剂中的一或多者。

如本文中所使用的液体药学上的组合物(无论溶液、悬浮液形式或其他类似形式)可包括以下佐剂中的一或多者：无菌稀释剂，诸如注射用水、盐水溶液(优选为生理盐水)、林格氏溶液(Ringer's solution)、等渗性氯化钠、不挥发性油类(诸如可用作溶剂或悬浮介质的合成单甘油酯或二甘油酯、聚乙二醇、甘油、丙二醇或其他溶剂)；抗菌剂，诸如苄醇或对羟基苯甲酸甲酯；抗氧化剂，诸如抗坏血酸或亚硫酸氢钠；螯合剂，诸如乙二胺四乙酸；缓冲剂，诸如乙酸盐、柠檬酸盐或磷酸盐，及用于调整张力的药剂，诸如氯化钠或葡萄糖。肠胃外制剂可封装于由玻璃或塑料制成的安瓿、抛弃式针筒或多剂量小瓶中。生理盐水为优选佐剂。可注射药学上的组合物优选为无菌的。

制剂中亦需要包括其他组份，诸如递送媒介物，包括但不限于铝盐、油包水乳液、生物可降解性油媒介物、水包油乳液、生物可降解性微胶囊及脂质体。用于所述媒介物中的免疫刺激性物质(佐剂)的实例包括N-乙酰基胞壁酰基-L-丙氨酸-D-异谷酰胺(MDP)、脂多醣(LPS)、葡聚糖、IL-12、GM-CSF、γ干扰素及IL-15。

尽管本发明的药学上的组合物中可使用本领域普通技术人员已知的任何合适载体，但载体类型将视给药模式及是否需要持续释放而变化。对于肠胃外给药(诸如皮下注射)而言，载体优选包含水、盐水、醇、脂肪、蜡或缓冲剂。对于经口给药而言，可使用上述任何载体或固体载体，诸如甘露糖醇、乳糖、淀粉、硬脂酸镁、糖精钠、滑石粉、纤维素、葡萄糖、蔗糖及碳酸镁。生物可降解性微球体(例如聚乳酸半

乳糖苷)亦可用作本发明的药学上的组合物的载体。合适的生物可降解性微球体是揭示于例如美国专利第 4,897,268 号和第 5,075,109 号中。就此而言，微球体优选大于约 25 微米。

药学上的组合物亦可含有稀释剂(诸如缓冲剂)、抗氧化剂(诸如抗坏血酸)、低分子量(少于约 10 个残基)多肽、蛋白、氨基酸、碳水化合物(例如葡萄糖、蔗糖或糊精)、螯合剂(例如 EDTA)、谷胱甘肽及其他稳定剂及赋形剂。中性缓冲盐水或与非特异性血清白蛋白混合的盐水为例示性的适当稀释剂。优选地，使用适当赋形剂溶液(例如蔗糖)作为稀释剂将产物配制为冻干物。

根据本领域中已知的技术(包括揭示于已公开的美国专利申请第 2006/0008415 A1 号中的技术，其以引用方式并入本文中)，本发明的药学上的组合物亦包括稳定蛋白及稳定液体药物制剂。所述技术包括蛋白的衍生化，其中该蛋白包含与 N-乙酰基-L-半胱氨酸、N-乙基-顺丁烯二酰亚胺或半胱氨酸偶合的硫醇基。

如上所述，本发明包括能够递送编码具有效应功能的多价结合蛋白的核酸分子的组合物。所述组合物包括重组病毒载体，例如逆转录病毒(参见 WO 90/07936、WO 91/02805、WO 93/25234、WO 93/25698 及 WO 94/03622)、腺病毒(参见 Berkner, 1988 *Biotechniques* 6:616-627; Li 等人，1993 *Hum. Gene Ther.* 4:403-409; Vincent 等人，*Nat. Genet.* 5:130-134 及 Kolls 等人，1994 *Proc. Natl. Acad. Sci. USA* 91:215-219)、痘病毒(参见美国专利第 4,769,330 号、美国专利第 5,017,487 号及 WO 89/01973))、与聚阳离子型分子复合的重组表达构建体核酸分子(参见 WO 93/03709)及与脂质体结合的核酸(参见 Wang 等人，1987 *Proc. Natl. Acad. Sci. USA* 84:7851)。在某些实施方案中，可将 DNA 与灭活或失活的腺病毒相连接(参见 Curiel 等人，1992 *Hum. Gene Ther.* 3:147-154; Cotton 等人，1992 *Proc. Natl. Acad. Sci. USA 89*:6094)。其他合适的组合物包括 DNA-配体(参见 Wu 等人，1989 *J. Biol. Chem. 264*:16985-16987)及脂质-DNA 组合(参见 Felgner 等人，1989 *Proc. Natl. Acad. Sci. USA* 84:7413-7417)。

除引导体内程序外，亦可使用离体程序，其中将细胞从宿主(例如

个体，诸如人类患者)中移除、修饰且置于同一或另一宿主动物中。显然，在离体环境中，可使用上述任何组合物将本发明的构建体(蛋白/多肽或编码其的核酸)引入组织细胞中。病毒、物理及化学摄取方法的实验方案已熟知于本领域中。

抗体的形成

借助于多次皮下或腹膜内注射抗原多肽或其片段及佐剂，通常可在动物(例如兔、仓鼠、山羊、绵羊、马、猪、大鼠、沙鼠、豚鼠、小鼠或任何其他合适哺乳动物，以及其他非哺乳动物种类)中产生针对抗原多肽的多克隆抗体。佐剂包括但不限于弗氏完全或不完全佐剂(complete or incomplete Freund's adjuvant)、矿物质凝胶(诸如氢氧化铝)及表面活性物质，诸如溶血卵磷脂、普流尼克(pluronic)多元醇、聚阴离子、肽、油乳液及二硝基苯酚。BCG(卡介苗(bacilli Calmette-Guerin))及短小棒状杆菌亦为潜在适用的佐剂。其可适用于使抗原多肽与在待免疫的物种中具有免疫原性的载体蛋白结合；典型载体包括匙孔螺血蓝蛋白、血清白蛋白、牛甲状腺球蛋白或大豆胰蛋白酶抑制剂。此外，可使用诸如明矾的聚集剂以增强免疫反应。免疫之后，将动物放血且利用常规技术测定血清的抗-抗原多肽抗体效价。可将多克隆抗体用于血清中，自血清中检测多克隆抗体，或可利用例如抗原亲和层析法将多克隆抗体自血清中纯化。

针对抗原多肽的单克隆抗体可利用通过培养物中的连续细胞系产生抗体分子的任何方法来制备。例如，单克隆抗体可通过如 Kohler 等人，Nature 256:495 [1975]中所述的杂交瘤方法；人类 B-细胞杂交瘤技术(Kosbor 等人，Immunol Today 4:72, 1983；Cote 等人，Proc Natl Acad Sci 80: 2026-2030, 1983)及 EBV-杂交瘤技术(Cole 等人，Monoclonal Antibodies and Cancer Therapy, Alan R Liss Inc, New York N.Y.，第 77-96 页，(1985))来制备。

当使用杂交瘤技术时，可使用骨髓瘤细胞系。优选地，适用于产生杂交瘤的融合程序中的细胞系不产生内源抗体，具有高融合效率，且显示酶缺乏，从而使其不能在仅支持所要融合细胞(杂交瘤)生长的某些选择性培养基中生长。例如，若经免疫的动物为小鼠，则可使用

P3-X63/Ag8、P3-X63-Ag8.653、NS1/1.Ag 4 1、Sp210-Ag14、FO、NSO/U、MPC-11、MPC11-X45-GTG 1.7 及 S194/5XX0 Bul；若为大鼠，则可使用 R210.RCY3、Y3-Ag 1.2.3、IR983F 及 4B210；且 U-266、GM1500-GRG2、LICR-LON-HMy2 及 UC729-6 均可结合细胞融合来使用。

在一替代实施方案中，人类抗体可经由噬菌体显示文库来制备(Hoogenboom 等人，J. Mol. Biol. 227: 381 [1991]；Marks 等人，J. Mol. Biol. 222: 581；亦参见美国专利第 5,885,793 号)。所述方法经由以下步骤来模拟免疫选择：在丝状噬菌体表面上显示全部抗体，且随后依据其与所选抗原的结合来选择噬菌体。一种该技术是描述于以 Adams 等人的名义申请的 PCT 申请第 PCT/US98/17364 号中，其描述利用该方法分离 MPL-受体及 msk-受体的高亲和性及功能性促效抗体。在该方法中，人类抗体基因的完整清单可通过自如前所述的外周血淋巴细胞中克隆天然重排的人类 V 基因而形成(Mullinax 等人，Proc. Natl. Acad. Sci. (USA) 87: 8095-8099 [1990])。

或者，完整的全部合成人类重链可通过将各人类 VH 区段与随机核苷酸的 D 区段连同人类 J 区段一起组合而由未经重排的 V 基因区段来形成(Hoogenboom 等人，J. Mol. Biol. 227:381-388 [1992])。同样，全部轻链可通过将各人类 V 区段与 J 区段合并来构建(Griffiths 等人，EMBO J. 13:3245-3260 [1994])。将编码完整抗体(亦即重链与轻链)的核苷酸连接为单链 Fv 片段，且将该聚核苷酸与编码丝状噬菌体次要外壳蛋白的核苷酸接合。当该融合蛋白在噬菌体表面上表达时，编码特异性抗体的聚核苷酸可通过使用固定抗原选择来鉴别。

除形成多克隆及单克隆抗体的典型方法外，亦涵盖用于形成任何已知抗体形式的任何方法。除多克隆及单克隆抗体外，抗体形式亦包括嵌合抗体、人源化抗体、CDR-移植抗体及抗体片段及变异体。

特异性结合剂的变异体及衍生物

在一实施例中，提供插入变异体，其中一或多个氨基酸残基增补特异性结合剂氨基酸序列。插入可位于蛋白的任一末端或两末端，或可位于特异性结合剂氨基酸序列的内部区中。本发明的变异体产物亦

包括成熟特异性结合剂产物，亦即其中前导序列或信号序列已移除而所得蛋白具有其他氨基末端残基的特异性结合剂产物。其他氨基末端残基可源自另一种蛋白或可包括一或多个因源自特异性蛋白而无法鉴别的残基。涵盖在位置-1 具有其他甲硫氨酸残基的多肽(例如具有效应功能的 Met-1-多价结合肽)，亦包括在位置-2 及位置-1 具有其他甲硫氨酸及赖氨酸残基的本发明的多肽(具有效应功能的 Met-2-Lys-1-多价结合蛋白)。具有其他 Met、Met-Lys 或 Lys 残基(或一般而言为一或多个碱性残基)的本发明多肽的变异体尤其适用于增强在细菌宿主细胞中产生重组蛋白。

本发明亦包括因使用特异性表达系统所产生的具有其他氨基酸残基的本发明的特异性多肽。例如，使用表达作为谷胱甘肽-S-转移酶(GST)融合产物的部分的所要多肽的市购载体可在 GST 组份与所要多肽切割之后提供在位置-1 具有其他甘氨酸残基的所要多肽。亦涵盖在其他载体系统中经由表达而产生的变异体，包括其中组氨酸标志物通常在序列的羧基末端和/或氨基末端并入氨基酸序列中的那些变异体。

在另一方面中，本发明提供缺失变异体，其中本发明的多肽中的一或多个氨基酸残基被移除。缺失可在多肽的一个末端或两个末端实现，或经由氨基酸序列中的一或多个残基的移除来实现。缺失变异体必然包括本发明的多肽的所有片段。

抗体片段是指具有对应于免疫球蛋白可变区序列中的至少部分的序列的多肽。片段例如可通过酶促裂解或化学裂解对应于全长抗体的多肽而形成。其他结合片段包括通过合成技术或通过重组 DNA 技术(诸如表达含有编码部分抗体可变区的核酸序列的重组质粒)而形成的那些片段。优选的多肽片段呈现如本文中所述的靶标所特有或所特异性的免疫特性。具有所要免疫特性的本发明的片段可通过本领域中熟知且常规实践的任何方法来制备。

在又一方面中，本发明提供具有效应功能的多价结合多肽的取代变异体。取代变异体包括其中氨基酸序列中的一或多个氨基酸残基被移除且经替代性残基置换的那些多肽。在某些实施方案中，所述取代在性质上为保守的；然而本发明包括亦为非保守性的取代。氨基酸可

根据物理特性和对二级与三级蛋白结构的作用来进行分类。在本领域中，保守性取代可理解为一种氨基酸取代具有类似特性的另一氨基酸。表 A 中列举例示性保守性取代(参见 WO 97/09433，第 10 页，1997 年 3 月 13 日公开(PCT/GB96/02197，于 96 年 6 月 9 日提交))，紧接如下。

表 A

保守性取代 I		
侧链	特征	氨基酸
脂族	非极性	G A P I L V
	极性-不带电荷	S T M N Q
	极性-带电荷	D E K R
芳族		H F W Y
其他		N Q D E

或者，如在表 B 中所列举，保守性氨基酸可如 Lehninger，[Biochemistry，第二版；Worth Publishers, Inc. NY:NY (1975)，第 71-77 页]中所述来分组，紧接如下。

表 B:

保守性取代 II		
侧链	特征	氨基酸
非极性(疏水性)	A.脂族:	A L I V P
	B.芳族	F W
	C.含硫	M
	D.边界	G
不带电荷-极性	A.羟基	S T Y
	B.酰胺	N Q
	C.硫氢基	C
	D.边界	G
带正电荷(碱性)		K R H
带负电荷(酸性)		D E

保守性取代 II

侧链	特征	氨基酸
非极性(疏水性)		
	A.脂族:	A L I V P
	B.芳族:	F W
	C.含硫:	M
	D.边界:	G
不带电荷-极性		
	A.羟基:	S T Y
	B.酰胺:	N Q
	C.硫氢基:	C
	D.边界:	G
带正电荷(碱性)		K R H
带负电荷(酸性)		D E

本发明亦提供特异性结合剂多肽的衍生物。衍生物包括具有除氨基酸残基的插入、缺失或取代以外的修饰的特异性结合剂多肽。优选地，所述修饰具有共价性质，且包括例如与聚合物、脂质、其他有机部分及无机部分的化学键合。本发明的衍生物可被制备以增加特异性结合剂多肽的循环半衰期，或可被设计以改良多肽靶向所要细胞、组织或器官的能力。

本发明进一步包括如美国专利第 4,640,835 号、第 4,496,689 号、第 4,301,144 号、第 4,670,417 号、第 4,791,192 号及第 4,179,337 号中所述、经共价修饰或衍生化以包括一或多种水溶性聚合物连接物(诸如聚乙二醇、聚氧乙二醇或聚丙二醇)的具有效应功能的多价结合蛋白。本领域中已知的其他适用聚合物包括单甲氧基聚乙二醇、葡聚糖、纤维素及其他基于碳水化合物的聚合物、聚(N-乙烯基吡咯啶酮)-聚乙二醇、丙二醇均聚物、聚氧化丙烯/氧化乙烯共聚物、聚氧乙烯化多元醇(例如甘油)及聚乙烯醇，以及所述聚合物的混合物。尤其优选为聚乙二醇(PEG)衍生化蛋白。水溶性聚合物可键合于特定位置，例如键合于

本发明的蛋白及多肽的氨基末端，或与多肽的一或多个侧链随机连接。PEG用于改良治疗能力的用途是描述于颁予Gonzales等人的美国专利第6,133,426号中。

免疫球蛋白诱变的靶点

某些策略可供用于操控抗原特异性免疫球蛋白(例如抗体)的固有特性，所述策略对基于非免疫球蛋白的结合分子无效。所述策略中支持例如基于抗体的分子优于所述替代物的良好实例为，经由亲和性成熟在体内调节抗体对其靶标的亲和性，其是利用免疫球蛋白基因的体细胞超突变以产生随免疫反应进展而亲和性增加的抗体。此外，已研发出改变免疫球蛋白的结构及免疫球蛋白区及结构域的重组技术。因此，可制备对指定抗原显示经改变的亲和性的源自抗体的多肽，且本领域中已知多种用于鉴别及纯化或分离所述多肽的纯化方案及监控屏幕。使用所述已知技术，可获得对抗原显示减少或增加的亲和性、包含源自抗体的结合结构域的多肽。用于产生显示改变的亲和性的多肽变异体的策略包括：对编码抗体的DNA使用定点诱变或随机诱变以改变蛋白中所存在的氨基酸，继而为筛检步骤，该步骤经设计以回收显示所要变化(例如，相对于未经修饰的亲本抗体或参照抗体增加或减少的亲和性)的抗体变异体。

诱变策略中为改变亲和性所最常靶向的氨基酸残基为抗体的轻链及重链可变区的互补性决定区(CDR)或超可变区中的那些残基。所述区含有与抗原进行物化相互作用的残基以及影响所述残基的空间排列的其他氨基酸。然而，亦已证明CDR区以外的可变结构域框架区中的氨基酸对抗体的抗原结合特性具有实质性作用，且可经靶向以操控所述特性。参见Hudson, P.J. Curr. Opin. Biotech., 9: 395-402 (1999)及其中的参考文献。

抗体变异体的更小且更有效筛检文库可通过对CDR中对应于在体细胞亲和性成熟过程期间倾向于“超突变”的区域的位点限制随机诱变或定点诱变来制备。参见Chowdhury等人，Nature Biotech., 17: 568-572 (1999)及其中的参考文献。已知以此方式限定超突变位点的DNA元件的类型包括同向重复和反向重复、某些共有序列、二级结构

及回文序列。共有 DNA 序列包括四碱基序列嘌呤-G-嘧啶-A/T(亦即 A 或 G-G-C 或 T-A 或 T)及丝氨酸密码子 AGY(其中 Y 可为 C 或 T)。

因此，本发明的另一方面为一组用于修饰抗体对其靶标的亲和性的诱变策略。所述策略包括重链和/或轻链的整个可变区的诱变、仅 CDR 区的诱变、CDR 内的共有超突变位点的诱变、框架区的诱变或所述方法的任何组合("诱变"在本上下文中可为随机诱变或定点诱变)。经由本领域技术人员已知的技术(诸如 X 射线晶体分析法)解决所述抗体和抗体:配体复合物的结构从而可完成对 CDR 区的明确描述及对包含抗体的结合位点的残基的鉴别。基于对所述抗体晶体结构的分析和表征的各种方法已为本领域技术人员所知且可用于接近 CDR 区。所述常用方法的实例包括 Kabat、Chothia、AbM 及接触定义。

Kabat 定义是基于序列可变性且为预测 CDR 区的最常用定义。Johnson 等人，Nucleic Acids Research, 28: 214-8 (2000)。Chothia 定义是基于结构性环区的位置。(Chothia 等人，J. Mol. Biol., 196: 901-17 [1986]; Chothia 等人，Nature, 342: 877-83 [1989])。AbM 定义是兼顾于 Kabat 与 Chothia 定义之间。AbM 为牛津大学分子组(Oxford Molecular Group)所生产的用于抗体结构建模的整套程序(Martin 等人，Proc. Natl. Acad. Sci (USA) 86:9268-9272 [1989]; Rees 等人，ABMTM(用于抗体可变区建模的电脑程序)，Oxford, UK; Oxford Molecular, Ltd.)。AbM 程序组是利用知识数据库与从头开始方法的组合由一级序列对抗体的三级结构建模。最近已引入另一种定义，称为接触定义。参见 MacCallum 等人，J. Mol. Biol., 5:732-45 (1996)。此定义是基于对可获得的复合物晶体结构的分析。

根据惯例，重链中的 CDR 结构域通常称为 H1、H2 及 H3，且依氨基末端向羧基末端移动的次序依次编号。轻链中的 CDR 区通常称为 L1、L2 及 L3 且依氨基末端向羧基末端移动的次序依次编号。

CDR-H1 长为约 10 至 12 个残基且根据 Chothia 和 AbM 定义，通常始于 Cys 后的第 4 个残基，或根据 Kabat 定义通常始于其后的第 5 个残基。H1 之后通常为 Trp，通常为 Trp-Val，而且可为 Trp-Ile 或 Trp-Ala。根据 AbM 定义，H1 的长度为约 10 至 12 个残基，而 Chothia

定义不包括最后的 4 个残基。

根据 Kabat 和 AbM 定义，CDR-H2 通常始于 H1 末端后的第 15 个残基。H2 之前的残基通常为 Leu-Glu-Trp-Ile-Gly，但可存在多种变异。H2之后通常为氨基酸序列 Lys/Arg-Leu/Ile/Val/Phe/Thr/Ala-Thr/Ser/Ile/Ala。根据 Kabat 定义，H2 的长度为约 16 至 19 个残基，其中 AbM 定义预测长度通常为 9 至 12 个残基。

CDR-H3 通常始于 H2 末端后的第 33 个残基，且通常在氨基酸序列 Cys-Ala-Arg 之后。H3 之后通常为氨基酸 Gly。H3 长度的范围为 3 至 25 个残基。

CDR-L1 通常始于约残基 24 且通常在 Cys 之后。CDR-L1 之后的残基通常为 Trp 且通常以如下序列中之一者开始：Trp-Tyr-Gln、Trp-Leu-Gln、Trp-Phe-Gln 或 Trp-Tyr-Leu。CDR-L1 的长度为约 10 至 17 个残基。

CDR-L2 始于 L1 末端后的约 16 个残基。其通常在以下残基之后：Ile-Tyr、Val-Tyr、Ile-Lys 或 Ile-Phe。CDR-L2 的长度为约 7 个残基。

CDR-L3 通常始于 L2 末端后的第 33 个残基且通常在 Cys 之后。L3 之后通常为氨基酸序列 Phe-Gly-XXX-Gly。L3 的长度为约 7 至 11 个残基。

用于修饰抗体的各种方法已描述于本领域中，包括例如人源化抗体的制备方法，其中人源化免疫球蛋白重链可变区框架的序列与供体免疫球蛋白重链可变区框架的序列具有 65%至 95%的一致性。各人源化免疫球蛋白链除包含 CDR 外，通常亦包含来自供体免疫球蛋白框架的氨基酸，所述氨基酸例如能够与 CDR 相互作用以实现结合亲和性，诸如一或多个与供体免疫球蛋白中的 CDR 紧邻的氨基酸或约 3 埃(angstrom)以内的那些氨基酸(如通过分子建模所预测)。重链和轻链可各自通过使用各种定位标准中的任一种或所有标准来设计。当组合成完整抗体时，人源化免疫球蛋白在人类中大体上具有非免疫原性且对抗原(诸如含有表位的蛋白或其他化合物)保留与供体免疫球蛋白大体上相同的亲和性。

在一实施例中，描述与亲本抗体具有类似结合特异性，但具有增

强的人类特征的抗体及抗体片段的制备方法。人源化抗体是通过利用例如噬菌体显示技术的链替换法及包含对所关注的抗原具特异性的非人类抗体的重链或轻链可变区的多肽而获得，接着将其与全部人类互补(轻链或重链)链可变区组合。鉴别对所关注的抗原具特异性的杂交对，且将所选对中的人类链与全部人类互补可变结构域(重链或轻链)组合。在另一实施方案中，将 CDR 中来自非人类抗体的组份与 CDR 中来自人类抗体的全部组成部分组合。自抗体多肽二聚体的所得文库选择杂交体，且其可用于第二人源化替换步骤中；或者，若该杂交体已具有足够的人类特征而具有治疗价值，则排除此第二步骤。增强人类特征的修饰方法于本领域中已知。

另一实施例为一种如下制备人源化抗体的方法：以 CDR 氨基酸序列取代对应的人类 CDR 氨基酸序列和/或以 FR 氨基酸序列取代对应的人类 FR 氨基酸序列。

另一实施例提供用于鉴别抗体可变结构域中可经修饰的氨基酸残基的方法，所述方法在减少其相对于异源物种的免疫原性的同时不减弱抗原结合结构域的天然亲和性；及制备所述适于给予异源物种的经修饰抗体可变区的方法。

通过本领域中已知的任何方法对免疫球蛋白(诸如抗体)的修饰可被设计以达成增加或减少的抗原结合亲和性和/或以减少抗体在受者体内的免疫原性和/或以调节效应活性程度。在一种方法中，人源化抗体可被修饰以排除糖基化位点从而增强抗体对其同源抗原的亲和性(Co 等人，Mol. Immunol. 30:1361-1367[1993])。诸如“重构”、“超嵌合”及“面饰/重修表面”的技术已制备具有更大治疗潜能的人源化抗体。Vaswami 等人，Annals of Allergy, Asthma, & Immunol 81:105 (1998); Roguska 等人，Prot. Engineer. 9:895-904 (1996)]。亦参见美国专利第 6,072,035 号，其描述抗体重构方法。尽管所述技术通过减少外来残基的数目而减弱抗体免疫原性，但在抗体重复给药之后，其无法阻止抗遗传型反应及抗同种异型反应。用于减少免疫原性的所述方法的替代方法是描述于 Gilliland 等人，J. Immunol. 62 (6):3663-71 (1999)中。

在很多情况下，人源化抗体导致抗原结合能力下降。因此，优选

使人源化抗体“回复突变”以包括一或多个初始(最常为啮齿动物)抗体中所存在的氨基酸残基，以试图恢复抗体的结合亲和性。参见例如Saldanha 等人，Mol. Immunol. 36:709-19 (1999)。

已证明免疫球蛋白的糖基化可影响效应功能、结构稳定性及自抗体生产细胞分泌的速率(参见 Leatherbarrow 等人，Mol. Immunol. 22:407 (1985)，该文献以引用方式并入本文中)。导致所述特性的碳水化合物群通常与抗体的恒定区相连接。例如，在 C_{H2} 结构域中的 Asn 297 处将 IgG 糖基化可促使 IgG 活化补体依赖性细胞溶解的全部能力(Tao 等人，J. Immunol. 143:2595 (1989))。在 C_{H3} 结构域中的 Asn 402 处将 IgM 糖基化例如可促成抗体的适当组装及细胞溶解活性(Muraoka 等人，J. Immunol. 142:695 (1989))。在 IgA 抗体的 C_{H1} 及 C_{H3} 结构域中的位置 162 及 419 处将糖基化位点移除导致细胞内降解及对分泌的至少 90%抑制作用(Taylor 等人，Wall, Mol. Cell. Biol. 8:4197 (1988))。因此，本发明的分子包括经突变方式改变的免疫球蛋白，所述免疫球蛋白因例如恒定亚区中的特异性残基发生突变而显示改变的糖基化模式，从而改变效应功能。参见 Co 等人，Mol. Immunol. 30:1361-1367 (1993), Jacquemon 等人，J. Thromb. Haemost. 4:1047-1055 (2006), Schuster 等人，Cancer Res. 65:7934-7941 (2005)，及 Warnock 等人，Biotechnol Bioeng. 92:831-842 (2005)，各文献以引用方式并入本文中。

本发明亦包括具有至少一个结合结构域的多价结合分子，该结合结构域在序列上与已知免疫球蛋白可变区序列具有至少 80%、优选 90%或 95%或 99%的一致性且具有至少一个不同于该免疫球蛋白可变区的残基，其中所改变的残基增加糖基化位点、改变一或多个糖基化位点的位置或优选相对于免疫球蛋白可变区移除糖基化位点。在某些实施方案中，该变化移除免疫球蛋白可变区框架中的 N-连接糖基化位点或移除存在于免疫球蛋白重链可变区框架中，在约氨基酸残基 65 至约氨基酸残基 85(使用 Co 等人，J. Immunol. 148: 1149, (1992)的编号规定)范围的区域内的 N-连接糖基化位点。

涵盖本领域中已知的用于制备相对于免疫球蛋白参照序列显示改变的糖基化模式的多价结合分子的任何方法。例如，可利用多种遗传

技术中的任何技术来改变一或多个特定残基。或者，用于制备的宿主细胞可经工程化以制备改变的糖基化模式。例如，本领域中已知的一种方法以增强 ADCC 的对分式非岩藻糖化变异体形式提供改变的糖基化作用。所述变异体因在含有寡糖修饰酶的宿主细胞表达而产生。或者，涵盖 BioWa/Kyowa Hakko 的 Potelligent 技术以减少本发明的糖基化分子中的岩藻糖含量。在一种已知方法中，提供经由产生 GDP-岩藻糖来修饰免疫球蛋白 F_C 区的糖基化模式的用于重组免疫球蛋白制备的 CHO 宿主细胞。该技术可用于修饰本发明的多价结合分子的恒定亚区的糖基化模式。

除修饰结合结构域(诸如免疫球蛋白的结合结构域)的结合特性外，和除诸如人源化的修饰外，本发明亦包括通过使造成效应功能(诸如恒定亚区的效应功能)的残基变化或突变来调节效应功能。所述修饰可利用本领域中已知的任何技术来实现，诸如 Presta 等人，Biochem. Soc. Trans. 30:487-490 (2001)中所揭示的方法，该文献以引用方式并入本文中。例示性方法将包括使用 Presta 等人所揭示的实验方案以修饰已知可影响对应于 FCγRI、FCγRII、FCγRIII、FCαR 及 FCεR 的一或多个恒定亚区中的结合的特异性残基。

在另一种方法中，Xencor XmAb 技术可用于使对应于 F_C 结构域的恒定亚区工程化以增强细胞杀死效应功能。参见 Lazar 等人，Proc. Natl. Acad. Sci. (USA) 103(11):4005-4010 (2006)，该文献以引用方式并入本文中。例如，利用该方法可形成使 $F_C\gamma R$ 特异性与结合性最优化的恒定亚区，从而增强细胞杀死效应功能。

制备具有效应功能的多价结合蛋白

可使用含有且表达本发明的多价结合蛋白(具有效应功能)的多种表达载体/宿主系统。所述系统包括但不限于微生物，诸如经重组噬菌体、质粒、粘粒或其他表达载体转化的细菌；经酵母表达或穿梭载体转化的酵母；经病毒表达载体(例如杆状病毒)感染的昆虫细胞系统；经病毒表达载体(例如花椰菜花叶病毒(cauliflower mosaic virus; CaMV)；烟草花叶病毒(tobacco mosaic virus; TMV))转染或经细菌性表达载体(例如 Ti 或 pBR322 质粒)转化的植物细胞系统；或动物细胞

系统。适用于重组多价结合蛋白制备的哺乳动物细胞包括但不限于：VERO 细胞、海拉细胞、中国仓鼠卵巢(CHO)细胞系、COS 细胞(诸如 COS-7)、W138、BHK、HepG2、3T3、RIN、MDCK、A549、PC12、K562 及 HEK293 细胞。本文中以下描述重组表达多价结合蛋白的示例性实验方案。

表达载体可包含转录单元，该转录单元包含以下各者的组合：(1)在基因表达中具有调节作用的一或多种遗传元件，例如启动子、增强子或因子特异性结合位点；(2)编码转录成 mRNA 且翻译成蛋白的结合剂的结构或序列；及(3)适当的转录起始和终止序列。欲用于酵母或真核表达系统中的结构单元优选包括使宿主细胞能够胞外分泌所翻译的蛋白的前导序列。或者，若在无前导序列或转运序列的情况下表达重组多价结合蛋白，则其可包括氨基末端甲硫氨酸残基。该残基随后可从所表达的重组蛋白中裂解或不从其裂解以提供最终多价结合蛋白。

例如，可使用商购表达系统(例如毕赤氏酵母表达系统(Pichia Expression System) (Invitrogen, San Diego, CA))，依照制造商说明书在酵母中重组表达多价结合蛋白。该系统亦依赖于前-原-α 序列以引导分泌，但插入物的转录是在甲醇诱导下由醇氧化酶(AOX1)启动子来驱动。所分泌的多价结合肽可通过以下方法自酵母生长培养基中纯化，例如通过用于自细菌性及哺乳动物细胞上清液中纯化肽的方法。

或者，编码多价结合肽的 cDNA 可被克隆入杆状病毒表达载体 pVL1393(PharMingen, San Diego, CA)中。该载体可根据制造商说明书(PharMingen)使用以感染无 SF9 蛋白的培养基中的草地贪夜蛾细胞及制备重组蛋白。多价结合蛋白可使用肝素-琼脂糖柱(Pharmacia, Piscataway, NJ)自培养基中纯化及浓缩。用于蛋白表达的昆虫系统(诸如 SF9 系统)已为本领域技术人员所熟知。在一种该系统中，苜蓿丫纹夜蛾核多角体病病毒(Autographa californica nuclear polyhedrosis virus; AcNPV)可用作在草地贪夜蛾细胞中或在粉纹夜蛾幼虫(Trichoplusia larvae)中表达外来基因的载体。可将多价结合肽编码序列克隆入病毒的非必需区(诸如多角体蛋白基因)且在多角体蛋白启动子的控制下安置。多价结合肽的成功插入将使得多角体蛋白基因具有非活性且产生

缺乏外壳蛋白的重组病毒。重组病毒可用于感染表达肽的草地贪夜蛾细胞或粉纹夜蛾幼虫(Smith 等人，J Virol 46: 584, 1983; Engelhard 等人，Proc Nat Acad Sci (USA) 91: 3224-7, 1994)。

在另一实施例中，编码多价结合肽的 DNA 序列可通过 PCR 扩增且克隆入适当载体内，例如 pGEX-3X(Pharmacia, Piscataway, NJ)。pGEX 载体被设计以制备由该载体编码的包含谷胱甘肽-S-转移酶(GST)的融合蛋白，及插入该载体的克隆位点中的由 DNA 片段编码的多价结合蛋白。可形成包括例如适当裂解位点的 PCR 引物。若多价结合蛋白融合部分仅用于促进表达，或无需作为所关注的肽的连接物，则重组多价结合蛋白融合体可自融合蛋白的 GST 部切割。将 pGEX-3X/多价结合肽构建体被转化至大肠杆菌 XL-1 Blue 细胞(Stratagene, La Jolla CA)中，且将个别转化体分离并培养。将来自个别转化体的质粒 DNA 纯化且可使用自动测序仪对其进行部分测序以证明存在具有适当取向的所要多价结合蛋白编码核酸插入物。

可作为不溶性包涵体于细菌中制备的融合多价结合蛋白可如下加以纯化。通过离心收集宿主细胞；于 0.15 M NaCl、10 mM Tris (pH 8)、1 mM EDTA 中洗涤；且在室温下用 0.1 mg/ml 溶菌酶(Sigma Chemical Co.)处理 15 分钟。通过超音波处理使溶菌产物澄清，且将细胞碎片通过以 12,000 倍重力离心 10 分钟而沉淀。可将含有多价结合蛋白融合体的沉淀再悬浮于 50 mM Tris (pH 8)和 10 mM EDTA 中，经 50%甘油分层，且以 6000 倍重力离心 30 分钟。可将沉淀再悬浮于不含 Mg^{++} 和 Ca^{++}的标准磷酸盐缓冲盐水溶液(PBS)中。多价结合蛋白融合体可通过将再悬浮的沉淀于变性 SDS 聚丙烯酰胺凝胶中分离而进一步纯化(Sambrook 等人)。将凝胶浸泡于 0.4 M KCl 中以显现蛋白，将其切除且于缺乏 SDS 的凝胶电泳缓冲剂中电洗脱。若 GST/多价结合肽融合蛋白是以可溶性蛋白的形式而于细菌中制备，则其可使用 GST 纯化模组(Pharmacia Biotech)来纯化。

优选使多价结合蛋白融合体经受消化以使 GST 自本发明的多价结合肽裂解。可将消化反应物(20-40 μg 融合蛋白、20-30 单位人类凝血酶(4000 U/mg(Sigma)于 0.5 ml PBS 中)在室温下孵育 16 至 48 小时且

加载于变性SDS-PAGE凝胶上以分离反应产物。可将凝胶浸泡于0.4 M KCl 中以显现蛋白带。对应于具有预期分子量的多价结合肽的蛋白带的特性可使用自动测序仪(Applied Biosystems Model 473A, Foster City, CA)通过氨基酸序列分析来证明。或者，可通过对所述肽进行 HPLC 和/或质谱分析来证明其特性。

或者，可将编码多价结合肽的 DNA 序列克隆入含有所要启动子及任选地含有前导序列的质粒内(参见例如 Better 等人，Science, 240:1041-43, 1988)。该构建体的序列可通过自动测序来证明。接着可使用对细菌进行$CaCl_2$孵育及热休克处理的标准程序(Sambrook等人)，将质粒转化至合适大肠杆菌菌株(诸如菌株 MC1061)中。经转化的细菌可培养于 LB 培养基中，该 LB 培养基补充有羧苄西林(carbenicillin)或如本领域中已知的其他合适形式的选择物，且通过培养于合适培养基中来诱导表达蛋白的产生。若存在，则前导序列可实现多价结合肽的分泌且在分泌期间裂解。所分泌的重组蛋白可通过本文中以下所述的方法自细菌培养基中纯化。

用于表达重组蛋白的哺乳动物宿主系统已为本领域技术人员所熟知且为优选系统。可选择具有加工所表达的蛋白或产生某些用于提供蛋白活性的翻译后修饰的特定能力的宿主细胞系。多肽的所述修饰包括但不限于：乙酰化、羧基化、糖基化、磷酸化、脂质化及酰化。不同宿主细胞(诸如 CHO、海拉、MDCK、293、WI38 及其类似宿主细胞)具有特异于所述翻译后活性的细胞机制及特征机制且可被选择以确保对外来蛋白进行正确的修饰及加工。

转化的细胞优选可用于长期高产率的蛋白生产，且因而需要稳定表达。所述细胞优选含有至少一个可选择标志物以及所要表达盒的载体转化后，将所述细胞于富集培养基中培养 1-2 天，之后转移至选择性培养基。可选择标志物被设计以赋予选择物抗性且其存在使得成功表达外来蛋白的细胞可生长及回收。稳定转化细胞的抗性团可利用适于该细胞的组织培养技术来增殖。

可使用多种选择系统来回收转化的细胞以用于产生重组蛋白。所述选择系统包括但不限于分别位于 tk-细胞、hgprt-细胞或 aprt-细胞中

的 HSV 胸苷激酶、次黄嘌呤-鸟嘌呤磷酸核糖转移酶及腺嘌呤磷酸核糖转移酶基因。亦可利用抗代谢物抗性作为以下各者的选择基础：赋予甲氨喋呤抗性的 dhfr；赋予霉酚酸抗性的 gpt；赋予氨基糖苷 G418 抗性及赋予氯磺隆(chlorsulfuron)抗性的 neo；及赋予潮霉素抗性的 hygro。可适用的其他可选择基因包括 trpB，其使得细胞可使用吲哚替代色氨酸；或 hisD，其使得细胞可使用组氨醇替代组氨酸。为鉴别转化体而提供目测指示的标志物包括花青素、β-葡糖醛酸酶及其底物 GUS，及萤光素酶和其底物萤光素。

蛋白纯化

蛋白纯化技术已为本领域技术人员所熟知。所述技术包括在某一级别将多肽与非多肽组分的粗分离。将多价结合多肽与至少一种其他蛋白分离，可纯化所关注的多肽，但通常需要利用层析技术及电泳技术进一步纯化以达成部分或完全纯化(或纯化为均质性)。尤其适于制备纯多价结合肽的分析方法为离子交换层析法、排阻层析法、聚丙烯酰胺凝胶电泳法及等电聚焦法。尤其有效的肽纯化方法为快速蛋白液相层析法及 HPLC。

本发明的某些方面是关于所编码的多价结合蛋白或肽的纯化和在特定实施方案中的大体上纯化。如本文中所使用的术语“纯化的多价结合蛋白或肽”欲指可与其他组份分离的组合物，其中多价结合蛋白或肽被纯化至任何相对于其天然可获得的状态的程度。因此，纯化的多价结合蛋白或肽亦是指其天然存在的环境中所不存在的多价结合蛋白或肽。

通常，“纯化”是指一种已经受分离以移除各种其他组份且该组合物大体上保留其所表达的生物活性的多价结合蛋白组合物。若使用术语“大体上纯化”，则此名称是指多价结合蛋白组合物，其中多价结合蛋白或肽形成组合物的主要组份，诸如在组合物中构成蛋白重量的约 50%、约 60%、约 70%、约 80%、约 90%、约 95%、约 99%或 99%以上。

本领域技术人员根据本公开将获知各种用于定量多价结合蛋白的纯化度的方法。所述方法包括例如测定活性组分的特异性结合活性或

通过 SDS/PAGE 分析来评估组分中多价结合多肽的量。用于评估多价结合蛋白组分纯度的优选方法为计算该组分的结合活性，将其与初始提取物的结合活性比较，及从而计算纯化度(本文中通过“纯化倍数”来评估)。用于表示结合活性的量的实际单位当然视选用于纯化的特定测定技术而定，而无论所表达的多价结合蛋白或肽是否显示可检测的结合活性。

适用于多价结合蛋白纯化的各种技术已为本领域技术人员所熟知。所述技术包括例如：经硫酸铵、PEG、抗体及其类似物的沉淀或通过热变性，继而离心；层析步骤，诸如离子交换、凝胶过滤、反相层析、羟基磷灰石及亲和层析法；等电聚焦；凝胶电泳；及所述技术与其他技术的组合。如本领域中通常所知，认为进行各种纯化步骤的次序可改变，或某些步骤可省略且仍不失为制备大体上纯化的多价结合蛋白的合适方法。

一般不需要总是以最纯化状态提供多价结合蛋白。实际上，预期在某些实施方案中可使用不是大体上纯化的多价结合蛋白。部分纯化可通过组合使用较少的纯化步骤来实现，或通过使用相同的常规纯化方案的不同形式来实现。例如，应了解使用 HPLC 装置所进行的阳离子交换柱层析通常比使用低压层析系统的相同技术产生更大的纯化。显示相对较低纯化度的方法在多价结合蛋白产物的回收总量方面或在维持所表达的多价结合蛋白的结合活性方面具有优势。

已知多肽的迁移在 SDS/PAGE 的不同条件下可变化，有时显著变化(Capaldi 等人，Biochem. Biophys. Res. Comm., 76:425, 1977)。因此应了解，在不同电泳条件下，纯化或部分纯化的多价结合蛋白表达产物的表观分子量可变化。

效应细胞

针对靶标细胞诱导例如 ADCC、ADCP(抗体依赖性细胞吞噬作用)及其类似者的效应细胞包括人类白细胞、巨噬细胞、单核细胞、活化嗜中性白细胞、活化天然杀伤(NK)细胞及嗜伊红细胞。效应细胞表达 $F_C\alpha R$(CD89)、FcγRI、FcγRII、FcγRIII 和/或 $F_C\varepsilon R1$ 且包括例如单核细胞及活化嗜中性白细胞。已发现例如 FcγRI 的表达可通过干扰素

γ(IFN-γ)来上调。此增强的表达可增强单核细胞和嗜中性白细胞防御靶标细胞的细胞毒性活性。因此，效应细胞可经(IFN-γ)或其他细胞因子(例如 TNF-α 或 β、群落刺激因子、IL-2)活化，以在与本发明的多价蛋白接触之前增加细胞表面上 FcγRI 的存在。

本发明的多价蛋白提供抗体效应功能，诸如抗体依赖性效应细胞介导的细胞毒性(ADCC)，以供用于防御靶标细胞。具有效应功能的多价蛋白可如本文中所教导单独给予，或与效应细胞偶联后给予，从而形成“活化效应细胞”。“活化效应细胞”为如本文中所定义的效应细胞，其与亦如本文中所定义的具有效应功能的多价蛋白连接，以便在给药之前有效提供具有靶向功能的效应细胞。

活化效应细胞可以细胞在生理学上可接受的溶液中的悬浮液形式被体内给药。所给予的细胞数目是在约 10^8-10^9 个的数量级上，但视治疗目的而有所变化。一般而言，该量将足以获得效应细胞在靶标细胞处的定位及在此位置提供所要程度的效应细胞功能，诸如通过 ADCC 和/或吞噬作用的细胞杀死作用。如本文中所使用的术语生理学上可接受的溶液意欲包括可使体内给予的靶向效应细胞稳定的任何载体溶液，包括例如盐水及缓冲水溶液、溶剂、抗菌剂及抗真菌剂、等渗剂及其类似物。

因此，本发明的其他方面提供一种在个体中针对细胞诱导特异性抗体效应功能(诸如 ADCC)的方法，该方法包含向个体给予于生理学上可接受的培养基中的多价蛋白(或编码核酸)或活化效应细胞。如本领域中所知，给药途径可变化且合适的给药途径将由本领域技术人员基于对病例特异性变量及常规程序的考虑来判定。

无细胞效应

无细胞效应亦由本发明的多价分子提供，例如通过提供 CDC 功能性。补体系统为免疫系统的生物化学级联，其有助于将外来物(诸如病原体)自有机体中清除。其是源自很多小的血浆蛋白，所述血浆蛋白共同运作，通过破坏靶标细胞的质膜而诱导靶标细胞的细胞溶解。补体系统由 35 个以上的可溶性且结合细胞的蛋白组成，其中 12 个直接涉及于补体通路中。所述蛋白在引起补体系统活化的三种生物化学通路

中具有活性：经典补体通路、替代性补体通路及甘露糖结合凝集素通路。抗体(尤其 IgG1 类抗体)亦可“结合”补体。所述通路在本领域中已获得详细了解且此处不再重复，但值得注意的是，补体依赖性细胞毒性不取决于结合分子与免疫系统的细胞(例如 B 细胞)的相互作用。亦值得注意的是补体系统通过补体调节蛋白来调节。所述蛋白以高于补体蛋白的浓度存在于血液血浆中。补体调节蛋白是存在于自身细胞表面上，提供防止自身细胞受补体蛋白靶向的机制。预期补体系统在具有免疫组份的若干种疾病中起作用，诸如巴-西二氏(Barraquer-Simon)综合征、阿兹海默氏病(Alzheimer's disease)、哮喘、红斑性狼疮症、各种形式的关节炎、自体免疫心脏病及多发性硬化症。缺乏末端通路使得个体易感染自体免疫疾病及传染病(尤其脑膜炎)。

疾病、病症和状态

本发明提供结合一或多个结合伙伴的具有效应功能的多价结合蛋白及其变异体和衍生物，且该结合事件适用于治疗、预防或改善与疾病、病症或病理性状态(优选使人类痛苦的疾病、病症或病理性状态)相关的症状。在所述方法的优选实施方案中，具有效应功能的多价(及多特异性)结合蛋白可使具有靶标(诸如肿瘤特异性细胞表面标志物)的细胞与效应细胞(诸如免疫系统中显示细胞毒活性的细胞)结合。在其他实施方案中，具有效应功能的多特异性、多价结合蛋白特异性地结合两种不同的疾病特异性细胞表面标志物、病症特异性细胞表面标志物或状态特异性细胞表面标志物，以确保正确靶标与效应细胞(诸如免疫系统的细胞毒性细胞)结合。此外，可使用具有效应功能的多价结合蛋白来诱导或增强抗原活性或抑制抗原活性。具有效应功能的多价结合蛋白亦适于组合疗法和姑息疗法。

在一方面中，本发明提供适用于治疗或预防以与细胞相关的抗原活性量异常为特征的疾病和状态的组合物及方法。所述疾病包括癌症及其他过度增殖性状态，诸如增生、牛皮癣、接触性皮炎、免疫性病症及不育症。包括实体肿瘤及白血病在内的多种癌症易受本文中所揭示的组合物及方法作用。可治疗的癌症类型包括但不限于：乳房、前列腺及结肠的腺癌；所有形式的肺支气管癌；髓样瘤；黑素瘤；肝瘤；

神经母细胞瘤；乳头状瘤；氨前体摄取与脱羧细胞瘤(apudoma)；迷芽瘤；鳃原性瘤；恶性类癌综合征；类癌心脏病；和癌(例如瓦尔克癌(Walker)、基底细胞癌、基底鳞状细胞癌、布-皮二氏(Brown-Pearce)氏癌、乳腺管瘤、艾利希氏(Ehrlich)瘤、Krebs 2 癌、默克尔(merkel)细胞癌、粘质肿瘤、非小细胞肺癌、燕麦细胞癌、乳突状癌、硬癌、细支气管癌、支气管癌、鳞状细胞癌及转移细胞癌)。可治疗的其他癌症类型包括：组织细胞病症；白血病；恶性组织细胞增生症；霍奇金氏(Hodgkin's)病；小免疫增殖症；非霍奇金氏淋巴瘤；浆细胞瘤；网状内皮细胞增生病；黑素瘤；软骨母细胞瘤；软骨瘤；软骨肉瘤；纤维瘤；纤维肉瘤；巨细胞瘤；组织细胞瘤；脂肪瘤；脂肪肉瘤；间皮瘤；粘液瘤；粘液肉瘤；骨瘤；骨肉瘤；脊索瘤；颅咽管瘤；无性细胞瘤；错构瘤；间质瘤；中肾瘤；肌肉瘤；釉母细胞瘤；牙骨质瘤；牙瘤；畸胎瘤；胸腺瘤；滋养叶瘤。此外，亦涵盖易受治疗作用的以下癌症类型：腺瘤；胆管瘤；胆脂瘤；圆柱瘤；囊腺癌；囊腺瘤；粒导细胞瘤；两性胚细胞瘤；肝瘤；汗腺腺瘤；胰岛细胞瘤；莱氏(Leydig)细胞瘤；乳头状瘤；塞氏(sertoli)细胞瘤；卵泡膜细胞瘤；平滑肌瘤；平滑肌肉瘤；肌母细胞瘤；肌瘤；肌肉瘤；横纹肌瘤；横纹肌肉瘤；室管膜瘤；节细胞神经瘤；神经胶质瘤；髓母细胞瘤；脑膜瘤；神经鞘瘤；神经母细胞瘤；神经上皮瘤；纤维神经瘤；神经瘤；副神经节瘤；非嗜铬性副神经节瘤。可治疗的癌症类型亦包括但不限于：血管角质瘤；血管淋巴样过度增生伴发嗜伊红血球增多症；硬化性血管瘤；血管瘤病；血管球瘤；血管内皮瘤；血管瘤；血管外皮细胞瘤；血管肉瘤；淋巴管瘤；淋巴管肌瘤；淋巴管肉瘤；松果体瘤；癌肉瘤；软骨肉瘤；叶状囊肉瘤；纤维肉瘤；血管肉瘤；平滑肌肉瘤；白血病性肉瘤；脂肪肉瘤；淋巴管肉瘤；肌肉瘤；粘液肉瘤；卵巢癌瘤；横纹肌肉瘤；肉瘤；赘瘤；神经纤维瘤；及子宫颈发育不良。本发明进一步提供适用于治疗其中细胞因抗原的表达异常高而变得永生化或过度增殖的其他状态的组合物及方法。

易受本发明的组合物及方法作用的各种过度增殖性病症例示为B-细胞癌症，包括B-细胞淋巴瘤(诸如各种形式的霍奇金氏病、非霍奇

金氏淋巴瘤(NHL)或中枢神经系统淋巴瘤)、白血病(诸如急性淋巴母细胞性白血病(ALL)、慢性淋巴细胞性白血病(CLL)、毛细胞白血病及慢性肌母细胞白血病)及骨髓瘤(诸如多发性骨髓瘤)。其他 B 细胞癌症包括小淋巴细胞性淋巴瘤、B-细胞前淋巴细胞性白血病、淋巴浆细胞淋巴瘤、脾边缘区淋巴瘤、浆细胞骨髓瘤、骨孤立性浆细胞肉瘤、骨外浆细胞瘤、粘膜相关淋巴组织(MALT)的节外边缘区 B-细胞淋巴瘤、节边缘区 B-细胞淋巴瘤、滤泡性淋巴瘤、套细胞淋巴瘤、扩散性大 B-细胞淋巴瘤、纵隔(胸腺)大 B-细胞淋巴瘤、血管内大 B-细胞淋巴瘤、原发性渗出性淋巴瘤、伯基特氏(Burkitt's)淋巴瘤/白血病、未定恶性潜能肿瘤的 B-细胞增殖、淋巴瘤样肉芽肿病及移植后淋巴组织增殖性病症。

以产生自身抗体为特征的病症通常视为自体免疫疾病。自体免疫疾病包括但不限于：关节炎、类风湿性关节炎、青少年类风湿性关节炎、骨关节炎、多软骨炎、牛皮癣性关节炎、牛皮癣、皮炎、多发性肌炎/皮肌炎、包涵体肌炎、炎性肌炎、中毒性表皮坏死溶解、全身性硬皮病及硬化症、肢端硬皮综合征(CREST syndrome)、炎性肠病伴随的反应、克罗恩氏(Crohn's)病、溃疡性结肠炎、呼吸窘迫综合征、成人呼吸窘迫综合征(ARDS)、脑膜炎、脑炎、葡萄膜炎、结肠炎、丝球体肾炎、过敏性状态、湿疹、哮喘、涉及 T 细胞渗入及慢性炎性反应的状态、动脉粥样硬化症、自体免疫心肌炎、白细胞粘附缺陷症、全身性红斑性狼疮症(SLE)、亚急性皮肤红斑性狼疮症、盘状狼疮症、狼疮骨髓炎、狼疮大脑炎、青少年发作型糖尿病、多发性硬化症、过敏性脑脊髓炎、视神经脊髓炎、风湿热、西登哈姆舞蹈病(Sydenham's chorea)、与由细胞因子及 T-淋巴细胞介导的急性及迟发型过敏相关的免疫反应、结核病、类肉瘤病、肉芽肿病(包括韦格纳氏(Wegener's)肉牙肿病及丘-施二氏(Churg-Strauss)疾病)、粒性白细胞缺乏症、脉管炎(包括过敏反应脉管炎/血管炎、ANCA 及类风湿性脉管炎)、再生障碍性贫血、戴-布二氏(Diamond Blackfan)贫血、免疫溶血性贫血(包括自体免疫溶血性贫血(AIHA))、恶性贫血、单纯红血球发育不全(PRCA)、凝血因子 VIII 缺乏症、A 型血友病、自体免疫嗜中性白血球减少症、

全血细胞减少症、白血球减少症、涉及白血球渗出的疾病、中枢神经系统(CNS)炎性病症、多器官损伤综合征、重症肌无力、抗原-抗体复合物介导的疾病、抗肾小球基底膜疾病、抗磷脂抗体综合征、过敏性神经炎、贝切特氏(Behcet)病、凯瑟曼氏(Castleman's)综合征、古德帕斯丘(Goodpasture)综合征、伊-兰二氏(Lambert-Eaton)重症肌无力综合征、雷诺氏(Reynaud's)综合征、修格兰氏(Sjorgen's)综合征、史蒂芬-强森氏(Stevens-Johnson)综合征、实体器官移植排斥反应、移植物抗宿主疾病(GVHD)、大疱性类天疱疮、天疱疮、自体免疫多内分泌腺病、血清反应阴性脊柱关节病、莱特尔氏(Reiter's)病、全身肌僵直综合征、巨细胞性动脉炎、免疫复合性肾炎、IgA 肾病、IgM 多发性神经病或 IgM 介导的神经病、特发性血小板减少性紫癜(ITP)、血栓性血小板减少性紫癜(TTP)、亨诺-许兰氏(Henoch-Schonlein)紫癜、自体免疫血小板减少症、睾丸及卵巢的自体免疫疾病(包括自体免疫睾丸炎及卵巢炎)、原发性甲状腺功能不足；自体免疫内分泌疾病，包括自体免疫甲状腺炎、慢性甲状腺炎(桥本氏(Hashimoto's)甲状腺炎)、亚急性甲状腺炎、特发性甲状腺功能不足、阿狄森氏(Addison's)病、格雷氏(Grave's)病、自体免疫多腺综合征(或多腺内分泌病综合征)、I 型糖尿病(亦称为胰岛素依赖性糖尿病(IDDM))及席汉氏(Sheehan's)综合征；自体免疫肝炎、淋巴间质性肺炎(HIV)、闭塞性细支气管炎(非移植)伴发性 NSIP、吉兰-巴雷氏(Guillain-Barré)综合征、大血管炎(包括风湿性多肌痛及巨细胞(高安氏(Takayasu's))动脉炎)、中血管炎(包括川崎氏(Kawasaki's)病及结节性多动脉炎)、结节性多动脉炎(PAN)、僵直性脊椎炎、伯杰氏(Berger's)病 (IgA 肾病)、快速进行性丝球体肾炎、原发性胆汁性肝硬化症、乳糜泻(麦胶性肠病)、冷球蛋白血症、冷球蛋白血症伴发肝炎、肌萎缩性侧索硬化症(ALS)、冠状动脉病、家族性地中海热、显微镜下多血管炎、科根氏(Cogan's)综合征、威斯科特-奥尔德里奇二氏(Whiskott-Aldrich)综合征及阻塞血栓性血管炎。

类风湿性关节炎(RA)为以关节的炎症为特征的慢性疾病，其导致肿胀、疼痛及功能丧失。长期患有 RA 的患者通常显示进行性关节损坏、畸形、失能及甚至早逝。除 RA 外，炎性疾病、病症及状态一般

而言均易受与炎症进程相关的症状(例如发热、疼痛、肿胀、充血)的治疗、预防或改善的作用，且本发明的组合物及方法有益于治疗、预防或改善异常或反常炎性过程，包括RA。

克罗恩氏病及相关疾病溃疡性结肠炎为属于称为炎性肠病(IBD)的疾病群的两类主要疾病。克罗恩氏病为导致消化道或胃肠(GI)道发炎的慢性病症。尽管其会牵涉口腔至肛门的胃肠道的任何区域，但其最常影响小肠和/或结肠。在溃疡性结肠炎中，GI的牵涉仅限于结肠。克罗恩氏病的特征为抵抗嗜中性白细胞抗原的抗体，亦即“核周抗嗜中性白细胞抗体”(pANCA)；及抗酿酒酵母的抗体，亦即“抗酿酒酵母抗体”(ASCA)。很多溃疡性结肠炎患者在其血液中具有pANCA抗体，但无ASCA抗体；而很多克罗恩氏病患者显示ASCA抗体且不显示pANCA抗体。评估克罗恩氏病的一种方法为使用克罗恩氏病活动指数(CDAI)，该活动指数是基于医师所收集的18种预测变量记分。150及150以下的CDAI值与静态疾病有关；150以上的值指示活动性疾病，且450以上的值被视为极其严重的疾病[Best等人，“Development of a Crohn's disease activity index (克罗恩氏病活动指数的发展).” Gastroenterology 70:439-444 (1976)]。然而，自研究开始起，有些研究者使用200至250的“主观值”作为健康记分。

全身性红斑性狼疮症(SLE)为自体免疫疾病，其因包括肾、皮肤及关节的多种器官中的血管发生反复性损伤所致。在患有SLE的患者中，T细胞与B细胞之间不适当的相互作用导致产生攻击细胞核的自身抗体。通常一致认为自身抗体导致SLE，因此耗尽B细胞谱系的新颖疗法(由于新B细胞由前驱细胞形成而使得免疫系统可复原)为SLE患者持久受益提供希望。

多发性硬化症(MS)亦为自体免疫疾病。其特征为中枢神经系统的炎症及髓鞘的破坏，使得脑、脊髓及身体内的神经细胞纤维隔绝。尽管MS的病因尚且未知，但广泛认为自体免疫T细胞为该疾病的发病机制的始作俑者。然而，MS患者的脑脊髓液中存在高含量的抗体，且有些理论预言导致抗体产生的B细胞反应对于介导该疾病起重要作用。

自体免疫甲状腺病是因产生刺激甲状腺而导致甲状腺功能亢进(格雷氏病)或破坏甲状腺而导致甲状腺功能减退(桥本氏甲状腺炎)的自身抗体所致。甲状腺的刺激是因自身抗体结合且活化促甲状腺激素(TSH)受体所致。甲状腺的破坏是因自身抗体与其他甲状腺抗原反应所致。

可受本发明的组合物及方法所提供的益处作用的其他疾病、病症及状态包括修格兰氏综合征，该综合征为以身体产生水分的腺体的破坏为特征的自体免疫疾病。此外，免疫性血小板减少性紫癜(ITP)是因自身抗体结合血小板且导致其破坏所致，且此状态适于应用本发明的物质及方法。重症肌无力(MG)(一种导致随意肌群无力的慢性自体免疫神经肌肉病症，其特征为自身抗体结合在神经肌肉接点处所表达的乙酰胆碱受体)为具有可使用本发明的组合物及方法治疗的症状的疾病，且预期本发明将有益于治疗和/或预防 MG。此外，使用本发明的组合物及方法，预期劳氏(Rous)肉瘤病毒感染可受至少一种症状的治疗或改善的作用。

本发明的另一方面是使用本发明的物质及方法来预防和/或治疗皮肤的任何过度增殖性状态，包括牛皮癣及接触性皮炎或其他过度增殖性疾病。牛皮癣的特征为皮肤的自体免疫性炎症且亦与关节炎(在30%的情况下)以及硬皮症、炎性肠病(包括克罗恩氏病及溃疡性结肠炎)相关。已证明患有牛皮癣及接触性皮炎的患者在所述病灶内具有增高的抗原活性(Ogoshi 等人，J. Inv. Dermatol., 110:818-23 [1998])。多特异性多价结合蛋白可将免疫系统的细胞毒性细胞直接递送至例如表达高含量抗原的病灶内的细胞。多价(例如多特异性)结合蛋白可经皮下给予或通过使用本文中所述的各种给药途径及本领域技术人员熟知的其他给药途径中的任何途径而给予至病灶附近。

亦涵盖特发性炎性肌病(IIM)(包括皮肌炎(DM)及多肌炎(PM))的治疗。炎性肌病可利用多种分类方案来加以分类。米勒氏分类法(Miller's classification schema)(Miller, Rheum Dis Clin North Am. 20:811-826, 1994)鉴别 2 种特发性炎性肌病(IIM): 多肌炎(PM)和皮肌炎(DM)。

多肌炎和皮肌炎为牵涉肌肉和(在 DM 情况下)皮肤的慢性、使衰弱的炎性疾病。所述病症为罕见病症，据报导，在美国，年发病率为每百万成人中每年约 5 至 10 例及每百万儿童中每年 0.6 至 3.2 例(Targoff, Curr Probl Dermatol. 1991, 3:131-180)。特发性炎性肌病与显著发病率及死亡率有关，已注意到患病成人中多达一半曾罹患严重损伤(Gottdiener 等人，Am J Cardiol. 1978, 41:1141-49)。米勒(Rheum Dis Clin North Am. 1994, 20:811-826 和 Arthritis and Allied Conditions (关节炎和有关状态), 第 75 章, Koopman 及 Moreland 编, Lippincott Williams 和 Wilkins, 2005)列举五类用于诊断 IIM 的标准，亦即，特发性炎性肌病标准(IIMC)评估，包括肌无力、肌肉退化的活组织检查迹象、肌肉相关酶的血清含量升高、肌病的电磁三联症、皮肌炎中皮疹的迹象，且亦包括自身抗体的迹象作为第二标准。

IIM 相关因子(包括肌肉相关酶及自身抗体)包括但不限于肌酸激酶(CK)、乳酸脱氢酶、醛缩酶、C-反应性蛋白、天冬氨酸氨基转移酶(AST)、丙氨酸氨基转移酶(ALT)及抗核自身抗体(ANA)、肌炎特异性抗体(MSA)及针对可提取核抗原的抗体。

易受本发明的方法作用的优选自体免疫疾病包括克罗恩氏病、吉兰-巴雷氏综合征(GBS；亦称为急性炎性脱髓鞘多发性神经病、急性特发性多神经根神经炎、急性特发性多发性神经炎及兰德里氏(Landry's)上行性麻痹)、红斑性狼疮症、多发性硬化症、重症肌无力、视神经炎、牛皮癣、类风湿性关节炎、甲状腺功能亢进(例如格雷氏病)、甲状腺功能不足(例如桥本氏病)、奥德氏(Ord's)甲状腺炎(类似于桥本氏病的甲状腺炎)、糖尿病(1 型)、再生障碍性贫血、莱特尔氏综合征、自体免疫肝炎、原发性胆汁性肝硬化症、抗磷脂抗体综合征(APS)、眼阵挛-肌阵挛综合征(OMS)、颞动脉炎(亦称为“巨细胞性动脉炎”)、急性播散性脑脊髓炎(ADEM)、古德帕斯丘综合征、韦格纳氏肉牙肿病、腹腔病、天疱疮、犬多发性关节炎、温热自体免疫性溶血性贫血。此外，本发明涵盖用于治疗或改善与以下疾病相关的症状的方法：子宫内膜异位、间质性膀胱炎、神经性肌强直、硬皮症、白斑病、外阴炎、导致扩大型心脏病(心肥大)的卡格氏(Chagas')病、类肉瘤病、慢性疲劳

综合征及自主神经障碍。

认为补体系统在具有免疫组份的很多疾病中起作用，诸如阿兹海默氏病、哮喘、红斑性狼疮症、各种形式的关节炎、自体免疫心脏病及多发性硬化症，所有所述疾病均作为易受治疗作用的疾病、病症或状态或利用本发明的方法可改善的症状而涵盖于本发明中。

由多价单链结合分子显示的一个或多个特定效应功能而定，优选为某些恒定亚区。例如，对于补体活化而言，IgG (IgG1、IgG2 或 IgG3)及 IgM 优选；对于调理作用及毒素中和作用而言，任何亚型的 IgG 均优选；对于病原体结合而言，IgA 优选；且对于诸如蠕虫的寄生虫的结合而言，IgE 优选。

例如,人类白细胞上存在作为三种不同类型的 Fcγ 受体的识别 IgG 抗体的恒定区的 F_CR，其可依据结构特性及功能特性以及依据 CD 单克隆抗体所检测的抗原结构来辨别。其称为 FcγRI、FcγRII 及 FcγRIII 且在白细胞的(重叠)子集上经差异性表达。

FcgRI (CD64)(在单核细胞、巨噬细胞、嗜中性白细胞、骨髓前体细胞及树突状细胞上经表达的高亲和性受体)包含同种型 Ia 及 Ib。FcgRI 对单体性人类 IgG1 及 IgG3 具有高亲和性。其对 IgG4 的亲和性为低约 10 倍，同时其不结合 IgG2。FcgRI 不显示遗传多态现象。

FcγRII (CD32)(包含同种型 IIa、IIb1、IIb2、IIb3 及 IIc)为分布最广泛的人类 FcγR 型，其被表达于大部分类型的血液白细胞上以及被表达于郎格罕氏(Langerhans)细胞、树突状细胞及血小板上。FcγRII 为仅结合所聚集的 IgG 的低亲和性受体。仅 FcγR 类能够结合 IgG2。FcγRIIa 显示遗传多态现象，分别产生两种不同的同种异型：FcγRIIa-H131 及 FcγRIIa-R131。该功能性多态现象可归因于位置 131 处的单一氨基酸差异：组氨酸(H)或精氨酸(R)残基，此对于 IgG 结合而言至关重要。FcγRIIa 易于结合人类 IgG 和 IgG3 且似乎不结合 IgG4。FcγRIIa-H131 比 FcγRIIa-R131 同种异型对复合 IgG2 具有高得多的亲和性。

FcγRIII (CD16)具有两种同种型或等位基因型，其两者均能够结合 IgG1 和 IgG3。对 IgG 具有中等亲和性的 FcγRIIa 是表达于巨噬细胞、单核细胞、天然杀伤(NK)细胞及 T 细胞子集上。FcγRIIIb 为对 IgG 具

低亲和性的受体，其在嗜中性白细胞上被选择性地表达。高移动性受体可与其他膜受体进行有效协作。对于骨髓瘤 IgG 二聚体的研究已证明，仅 IgG1 和 IgG3 结合 FcγRIIIb(以低亲和性结合)，而未发现 IgG2 和 IgG4 的结合。FcγRIIIb 具有共显性双等位基因多态现象，所述同种异型命名为 NA1(嗜中性白细胞抗原)和 NA2。

本发明的又一方面是使用本发明的物质及方法通过治疗、预防或减轻感染的效应来抗击因多种感染物中的任何感染物所致的感染。本发明的多价多特异性结合分子被设计以高效且有效地募集宿主有机体的免疫系统以抵抗源自外来有机体、外来细胞、外来病毒或外来无生命物体的感染。例如，多特异性结合分子可具有一个特异性地结合感染物上的靶标的结合结构域及另一个特异性地结合抗原呈递细胞上的靶标(诸如 CD40、CD80、CD86、DC-SIGN、DEC-205、CD83 及其类似者)的结合结构域。或者，多价结合分子的每一结合结构域均可特异性地结合感染物，从而更有效地中和感染物。此外，本发明涵盖特异性地结合感染物上的靶标且结合非细胞相关结合伙伴的多特异性多价结合分子，所述多特异性多价结合分子可结合多特异性结合分子的效应功能而有效治疗或预防因感染物所致的感染。

本发明所涵盖的感染性细胞包括任何已知的感染性细胞，包括但不限于各种细菌(例如病原性大肠杆菌(*E. coli*)、鼠伤寒沙门氏杆菌(*S. typhimurium*)、绿脓杆菌(*P. aeruginosa*)、炭疽芽孢杆菌(*B. anthracis*)、肉毒梭菌(*C. botulinum*)、难治梭状芽孢杆菌(*C. difficile*)、产气荚膜梭菌(*C. perfringens*)、幽门螺旋杆菌(*H. pylori*)、霍乱弧菌(*V. cholerae*)及其类似细菌)、分枝杆菌、支原体、真菌(包括酵母及霉菌)及寄生虫(包括原生虫纲(Protozoa)、吸虫纲(Trematoda)、多节绦虫纲(Cestoda)及线虫纲(Nematoda)的任何已知寄生性成员)中的任何者。感染性病毒包括但不限于真核病毒(例如腺病毒(adenovirus)、布尼亚病毒(bunyavirus)、疱疹病毒(herpesvirus)、乳多泡病毒(papovavirus)、副粘病毒(paramyxovirus)、小核糖核酸病毒(picornavirus)、痘病毒(poxvirus)、呼肠弧病毒(reovirus)、逆转录病毒(retrovirus)及其类似病毒)以及噬菌体。外来物包括(无论进入方式且无论是否存在有意伤害)进入有机体

(优选人类)内的物体。鉴于多耐药性感染物(例如细菌)(尤其医院感染的病原体)日益盛行，因此本发明的物质及方法可提供一种治疗方法以避免抗生素抗性增强所导致的困难。

与感染物相关且易受本文中所揭示的物质及方法(预防性或治疗性)治疗的疾病、状态或病症包括但不限于炭疽病、曲霉病、细菌性脑膜炎、细菌性肺炎(例如肺炎衣原体)、芽生菌病、肉毒中毒(botulism)、布鲁氏菌病(brucellosis)、念珠菌病(candidiasis)、霍乱、脑隐球菌病、隐球菌病、致泻性大肠杆菌、肠出血性大肠杆菌或肠毒性大肠杆菌、白喉、鼻疽、组织胞浆菌病、退伍军人病、麻风病、李氏杆菌病(listeriosis)、诺卡氏菌病(nocardiosis)、百日咳、沙门氏菌病(salmonellosis)、猩红热、孢子丝菌病、链球菌性喉炎、中毒性休克综合征、旅行者痢疾及伤寒。

本发明的其他方面及细节将由以下实施例显而易见，所述实施例旨在说明而非限制。实施例 1 描述免疫球蛋白重链和轻链可变区的重组性克隆。实施例 2 描述小模组免疫药物的构建。实施例 3 描述用于具有效应功能的多价结合蛋白的原型盒的构建。实施例 4 描述对该初始原型分子的结合及表达的研究。实施例 5 描述源自该初始原型分子的替代性构建体的构建，其中 EFD 与 BD2 之间连接子区的序列在长度与序列上已改变。此外，描述了替代形式，其中结合结构域 2 中 V 区的取向亦改变。实施例 6 描述随后对所述具有变异连接子形式的替代性构建体的结合及功能研究，从而在若干所述衍生形式中鉴别出连接子区的切割；及为解决该问题所研发的新颖序列变异体。实施例 7 描述多特异性多价融合蛋白(其中 BD1 与 BD2 均结合相同细胞类型上的抗原(CD20 和 CD37))或另一种多特异性融合蛋白(其中对于 BD2 具结合特异性的抗原已改变为人类 CD3 而非 CD28)的优选替代实施方案的构建。实施例 8 描述用 CD20-hIgG-CD37 多特异性构建体所进行的结合及功能研究。实施例 9 描述用 CD20-hIgG-CD3 多价融合蛋白构建体所进行的结合及功能研究。实施例 10 揭示具有连接子的多价结合分子，所述连接子是基于免疫球蛋白超家族成员的胞外结构域的特定区。实施例 11 揭示鉴别结合结构域的测定，所述结合结构域在多价结合分

子中预期可有效达成至少一种有益效应，该效应经鉴别与所述分子有关(例如疾病治疗)。

实施例 1

免疫球蛋白重链和轻链可变区的克隆

可使用本领域中已知的任何方法来诱发针对既定抗原靶标的抗体。此外，可使用本领域中已知的任何方法来克隆免疫球蛋白轻链和/或重链可变区，以及一或多种抗体的恒定亚区。以下方法提供例示性克隆方法。

A.总 RNA 的分离

为克隆免疫球蛋白重链和轻链可变区或恒定亚区，从分泌适当抗体的杂交瘤细胞中分离总 RNA。将获自杂交瘤细胞系的细胞(2×10^7)用 PBS 洗涤 1 次，且于 12×75 mm 圆底聚丙烯管(Falcon 第 2059 号)中经由离心来沉淀。向各管中添加 TRIzol™总 RNA 分离试剂(Gibco BRL, Life Technologies，目录号第 15596-018 号)(8 ml)且经由重复吸取将细胞溶解。在室温下将溶菌产物孵育 5 分钟，之后添加 1.6 ml (0.2 个体积)氯仿且剧烈震荡 15 秒钟。在室温下静置 3 分钟后，将溶菌产物于 4°C 预冷却的 Beckman JA-17 旋转器中以 9,000 rpm 离心 15 分钟，以便将水相与有机相分离。将上部水相(约 4.8 ml)移入新管中且与 4 ml 异丙醇轻缓地混合。在室温下孵育 10 分钟后，通过在 4°CJA-17 旋转器中以 9,000 rpm 离心 11 分钟使 RNA 沉淀，将 RNA 沉淀用 8 ml 的冰冷 75%乙醇洗涤，且于 4°C 下于 JA-17 旋转器中以 7,000×rpm 离心 7 分钟来再沉淀。将乙醇洗涤液轻轻倒出，且将 RNA 沉淀风干 10 分钟。将RNA沉淀再悬浮于 150 μl 经焦碳酸二乙酯(DEPC)处理的 ddH_2O 中，每 1 ml 经 DEPC 处理的 ddH_2O 中含有 1 μl 核糖核酸酶抑制剂(目录号第 799017 号；Boehringer Mannheim/Roche)。将所述沉淀通过轻缓吸取而再悬浮且在 55°C 下孵育 20 分钟。通过测量经稀释等分试样的 $OD_{260\ nm}$ (1.0 $OD_{260\ nm}$ 单位=40 μg/ml RNA)来对 RNA 样本进行定量。

B. cDNA 末端的快速扩增

进行 5' RACE 以扩增重链和轻链可变区或恒定亚区的末端。根据

制造商说明书，使用快速扩增 cDNA 末端试剂盒 2.0 版的 5' RACE 系统(Life Technologies，目录号第 18374-058 号)。使用寡核苷酸设计程序 Oligo 5.1 版(Molecular Biology Insights, Cascade CO)将简并 5' RACE 寡核苷酸引物设计成与例如两类常用小鼠免疫球蛋白重链(IgG1 和 IgG2b)的恒定区匹配。引物亦被设计成与小鼠 IgG κ 轻链的恒定区匹配。此为唯一类别的免疫球蛋白轻链，因此在引物设计中无需简并。引物的序列如下:

名称 序列	SEQ ID NO.
重链 GSP1 5'AGGTGCTGGAGGGGACAGTCACTGAGCTGC3'	7
嵌套重链 5'GTCACWGTCACTGRCTCAGGGAARTAGC3' (W=A 或 T; R=A 或 G)	8
轻链 GSP1 5'GGGTGCTGCTCATGCTGTAGGTGCTGTCTTTGC3'	9
嵌套轻链 5'CAAGAAGCACACGACTG AGGCACCTCCAGATG3'	10
5'Race 简化式锚定引物 5'GGCCACGCGTCGACTAGTACGG GNNGGGNNGGGNNG3'	11

为扩增小鼠免疫球蛋白重链组份，在 0.2 ml 薄壁 PCR 管中进行逆转录酶反应，该 PCR 管含有 2.5 皮摩尔(pmole)的重链 GSP1 引物(SEQ ID NO: 7)、4 μg 自合适杂交瘤克隆(例如克隆 4A5 或克隆 4B5)所分离的总 RNA 及 12 μl 经 DEPC 处理的 ddH_2O。同样，对于小鼠轻链组份而言，在 0.2 ml 薄壁 PCR 管中进行逆转录酶反应，该 PCR 管含有 2.5 皮摩尔的轻链 GSP1 引物(SEQ ID NO: 9)、4 μg 获自合适杂交瘤克隆(例如克隆 4A5 或克隆 4B5)的总 RNA 及 12 μl 经 DEPC 处理的 ddH_2O。

在 PTC-100 可程序化热循环器(MJ research Inc., Waltham, MA)中

进行反应。将混合物在 70°C 下孵育 10 分钟以使 RNA 变性且接着在湿冰上冷却 1 分钟。将所述管简短离心以便自管盖收集水份。随后，将以下组份添加至反应中：2.5 μl 的 10x PCR 缓冲剂(200 mM Tris-HCl (pH 8.4)、500 mM KCl)、2.5 μl 的 25 mM $MgCl_2$、1 μl 的 10 mM dNTP 混合物及 2.5 μl 的 0.1 M DTT。在通过轻缓吸取将各管混合后，将所述管置于 42°C 的 PTC-100 热循环器中历时 1 分钟以将混合物预温热。随后，将 1 μl(200 个单位)的 SuperScript™ II 逆转录酶(Gibco-BRL；目录号第 18089-011 号)添加至各管中，通过吸取轻缓地混合，且在 42°C 下孵育 45 分钟。使反应循环至 70°C 历时 15 分钟以终止反应，且接着循环至 37°C。接着将核糖核酸酶混合物(1 μl)添加至各反应管中，轻缓地混合，且在 37°C 下孵育 30 分钟。

将逆转录酶反应所产生的第一链 cDNA 用 GlassMAX DNA 分离旋转滤筒(Gibco-BRL)根据制造商说明书加以纯化。向各第一链反应中添加 120 μl 的 6 M NaI 结合溶液。接着将 cDNA/NaI 溶液转移至 GlassMAX 旋转滤筒中且以 13,000 倍重力离心 20 秒钟。将滤筒插入件小心移除且丢弃管中的流动部分。接着将旋转滤筒放回空管内且将 0.4 ml 冷(4°C)1×洗涤缓冲液添加至各旋转滤筒中。将所述管以 13,000 倍重力离心 20 秒钟且丢弃流动部分。将此洗涤步骤再重复三次。接着将 GlassMAX 滤筒用 0.4 ml 冷(4°C)70%乙醇洗涤 4 次。在丢弃最后 70%乙醇洗液的流动部分后，将滤筒放回管内且以 13,000 倍重力再离心 1 分钟以便使滤筒完全干燥。接着将旋转滤筒插入件转移至新鲜样本回收管，其中将 50 μl 经 65°C(预热)DEPC 处理的 ddH_2O 快速添加至各旋转滤筒中。将滤筒以 13,000 倍重力离心 30 秒钟以洗脱 cDNA。

C.末端去氧核苷酸转移酶(TdT)加尾

对于各第一链 cDNA 样本，将以下组份添加至 0.2 ml 薄壁 PCR 管内：6.5 μl 经 DEPC 处理的 ddH_2O、5.0 μl 的 5×加尾缓冲剂、2.5 μl 的 2 mM dCTP 及 10 μl 经 GlassMAX 纯化的适当 cDNA 样本。将各 24 μl 反应物于热循环器中在 94°C 下孵育 2-3 分钟以使 DNA 变性，且在湿冰上冷却 1 分钟。通过简单离心收集管的内含物。随后，将 1 μl 末端去氧核苷酸转移酶(TdT)添加至各管中。将所述管经由轻缓吸取来混合

且于 PTC-100 热循环器中在 37°C 下孵育 10 分钟。此 10 分钟孵育之后，将 TdT 通过循环至 65°C 历时 10 分钟而将其热灭活。将反应物于冰上冷却且将经 TdT 加尾的第一链 cDNA 储存于-20°C 下。

D.经 dC 加尾的第一链 cDNA 的 PCR

于 50 μl 体积中进行双重复 PCR 扩增(各经 dC 加尾的第一链 cDNA 样本进行两次独立的 PCR 反应)，该 50 μl 体积含有：200 μM dNTP、0.4 μM 的 5' RACE 简化式锚定引物(SEQ ID NO: 11)及 0.4 μM 嵌套重链 GSP2(SEQ ID NO: 8)或嵌套轻链 GSP2(SEQ ID NO: 10)中的任一者、10 mM Tris-HCl(pH 8.3)、1.5 mM $MgCl_2$、50 mM KCl、5 μl 的经 dC 加尾 cDNA 及 5 个单位的 Expand™ Hi-Fi DNA 聚合酶(Roche/Boehringer Mannheim GmbH, Germany)。采用"递减/递增"退火温度实验方案，在 PTC-100 可程序化热循环仪(MJ Research Inc.)中以如下条件扩增 PCR 反应物：初始 95°C 变性 40 秒钟，如下 5 个循环：94°C 下 20 秒，61°C-2°C/循环 20 秒钟，72°C 40 秒+1 秒/循环；接着如下 5 个循环：94°C 下 25 秒，53°C + 1°C/循环 20 秒，72°C 46 秒+1 秒/循环；接着如下 20 个循环：94°C 下 25 秒，55°C 20 秒，72°C 51 秒+1 秒/循环；及最后于 72°C 下孵育 5 分钟。

E. TOPO TA-克隆

使用 QIAQuick 凝胶纯化系统(QIAGEN Inc., Chatsworth, CA)将所得 PCR 产物经 1.0%琼脂糖凝胶进行凝胶纯化，并使用 TOPO TA Cloning®试剂盒(Invitrogen, San Diego, CA，目录号第 K4550-40 号)将其 TA-克隆进 pCR2.1，且根据制造商说明书将其转化至大肠杆菌 TOP10F'细胞(Invitrogen)中。根据制造商说明书，通过蓝/白筛检法鉴别具有插入物的克隆，其中白色克隆视为阳性克隆。将 3.5 ml 含有 50 μg/ml 氨苄西林的液体 Luria 肉汤(LB)的培养物用白色菌落接种，且以 225 rpm 震荡下于 37°C 下培养隔夜(约 16 小时)。

使用 QIAGEN 质粒微量制备试剂盒(QIAGEN Plasmid Miniprep Kit，QIAGEN Inc.，目录号第 12125 号)，根据制造商说明书纯化培养物中的质粒 DNA。将质粒 DNA 悬浮于 34 μl 的 1x TE 缓冲剂(pH 8.0)中且接着如前所述，通过萤光双去氧核苷酸测序和自动检测法，使用

ABI Big Dye Terminator 3.1 试剂以 1:4-1:8 的稀释度对阳性克隆测序，且使用 ABI 3100 DNA 测序仪进行分析。所用测序引物包括 T7(5'GTAATACGACTCACTATAGG3'; SEQ ID NO: 12)和 M13 反向(5'CAGGAAACAGCTATGACC3'; SEQ ID NO: 13)引物。测序结果将证明所述克隆对应于小鼠 IgG 序列。

F.利用重叠寡核苷酸延伸 PCR 的从头基因合成

该方法涉及使用重叠寡核苷酸引物和 PCR，采用高保真 DNA 聚合酶或聚合酶的混合物，合成免疫球蛋白 V-区或其他基因。起始于 V-区序列的中部，设计 40 至 50 个碱基引物，以便以任一方向使生长链扩增 20 至 30 个碱基，且邻接引物最少重叠 20 个碱基。各 PCR 步骤需要两个引物，一者引发反义链(正向引物或有义引物)且一者引发有义链(反向引物或反义引物)，以形成生长的双链 PCR 产物。在引物设计期间，可改变最终产物的核苷酸序列以形成限制性酶位点，破坏现有限制性酶位点，添加柔性连接子，改变、缺失或插入改变氨基酸序列的碱基，最优化整体 DNA 序列以增强引物合成且符合供预期用于表达合成基因的有机体的密码子使用规则。

将引物对合并且稀释以使得第一对为 5 μM，其后每一对具有 2 倍以上的浓度，最高为 80 μM。使用 Platinum PCR SuperMix-High Fidelity (Invitrogen, San Diego, CA，目录号第 12532-016 号)于 50 μL PCR 反应中将从所述引物混合物每一者中所取的 1 μL 扩增。初始在 94°C 下变性 2 分钟后，用如下的循环方案进行 PCR 的 30 个循环：94°C 下 20 秒、60°C 下 10 秒及 68°C 下 15 秒。使用 Qiaquick PCR 纯化柱(Qiagen Inc.，目录号第 28704 号)纯化 PCR 产物以移除过量引物及酶。接着利用如上所确切描述的 PCR 条件(例外之处为将每个循环至 68°C 的时间延长为 30 秒钟)，用下一组类似稀释的引物对将该 PCR 产物再扩增。将所得 PCR 产物经如上所述的引物和酶再次纯化，经 TOPO-TA 克隆，且如以上 E 部分中所确切描述进行测序。

实施例 2

小模组免疫药物(SMIP)的构建

以单链重组(鼠/人)scFv 的形式构建含有结合结构域 1 的具有效应功能的多特异性多价结合蛋白，将其命名为 2H7(VL-连接子-VH)。scFv 2H7 为特异性地识别 CD20 的小模组免疫药物(SMIP)。结合结构域是基于可公开获得的人类 CD20 抗体序列，对于 VH 而言 GenBank 登录号为 M17953 且对于 VL 而言 GenBank 登录号为 M17954。CD20-特异性 SMIP 是描述于共同拥有的美国专利公开 2003/133939、2003/0118592 和 2005/0136049 中，所述公开是以引用方式全文并入本文中。分隔 VL 和 VH 的肽连接子为编码以下序列的 15 氨基酸连接子: Asp-Gly_3Ser-$(Gly_4Ser)_2$。结合结构域 1 是位于多特异性结合蛋白的 N-末端，此结构域的 C-末端与恒定亚区的 N-末端直接连接，该恒定亚区含有铰链、C_{H2} 结构域和 C_{H3} 结构域(以氨基至羧基取向)。恒定亚区是源自 IgG1 抗体，其是通过从人类 PBMC 中 PCR 扩增人类 IgG1 而分离。铰链区的修饰通过: 用三个 Ser 残基取代野生型人类 IgG1 铰链结构域中所存在的三个 Cys 残基，该铰链区通过 15 氨基酸序列: EPKSCDKTHTCPPCP (SEQ ID NO: 14; 该三个经 Ser 残基置换的 Cys 残基以粗体指示)编码。在替代实施方案中，铰链区在一或多个半胱氨酸处被修饰，以便形成 SSS 及 CSC 型铰链。此外，最后脯氨酸有时被丝氨酸取代，以及半胱氨酸取代。

C_{H3} 结构域的 C-末端与一系列替代连接子结构域共价连接，所述连接子结构域并置于恒定亚区 C 末端与结合结构域 2 的氨基末端之间。视 BD2 的折叠特性而定，具有效应功能的优选多价结合蛋白将具有所述连接子的一者以使恒定亚区与结合结构域 2 相间隔，尽管该连接子并非本发明组合物的必要组份。对于某些特异性多价分子而言，该连接子对于结构域的分离可能是重要的，而对于其他者而言，该连接子可能是次要的。该连接子与特异性地识别 CD28 的 scFv 2E12 的 N-末端连接(V_H-连接子-V_L)。使多价结合分子的 scFv 2E12 部分的 VH 结构域与 VL 结构域分离的连接子为 20 氨基酸连接子$(Gly_4Ser)_4$，而非通常插于 scFv 的 V 结构域之间的标准$(Gy_4Ser)_3$ 连接子。经观测，较长连接子可改良 VH-VL 取向的 2e12 scFv 的结合特性。

所构建的多特异性多价结合分子含有结合结构域 1，该结合分子

包含：来自 SEQ ID NO: 171 的氨基酸 1-23 的 2E12 前导肽序列；显示于 SEQ ID NO: 171 的位置 24 的 2H7 鼠抗人类 CD20 轻链可变区；始于 SEQ ID NO: 171 的残基 130 的 Asp-Gly_3-Ser-$(Gly_4Ser)_2$ 连接子；2H7 鼠抗人类 CD20 重链可变区，其中在 VH 的可变结构域的残基 11 处的氨基酸由亮氨酸取代为丝氨酸(VHL11S)，且该结合分子在重链区末端具有单一丝氨酸残基(亦即 VTVS，其中规范序列为 VTVSS) (Genbank 登录号 M17953)，且插于两个结合结构域 BD1(2H7)与 BD2(2E12)之间者为人类 IgG1 恒定亚区，其包括包含“CSC”或“SSS”序列的经修饰铰链区，及野生型 C_{H2} 和 C_{H3} 结构域。具有效应功能的多价结合蛋白的核苷酸和氨基酸序列对于 CSC 形式而言分别以 SEQ ID NO: 228 和 229 显示，且对于 SSS 形式分别以 SEQ ID NO: 170 和 171 显示。

稳定表达细胞系的形成通过：经由电穿孔将未切或线性化的重组表达质粒转染至中国仓鼠卵巢细胞(CHO DG44 细胞)中，继而在含甲氨喋呤的培养基中选择。将产生最高量的多价结合蛋白的大量培养物和主孔在递增量的甲氨喋呤中扩增，且随后通过限制性稀释来克隆适应的培养物。在生物反应器或波纹袋中，使用获自 JRH Biosciences 的无血清培养基(Excell 302，目录号第 14324-1000M 号，补充有 4 mM 谷氨酰胺(Invitrogen，25030-081)、丙酮酸钠(Invitrogen 11360-070，稀释 1 倍)、非必需氨基酸(Invitrogen，11140-050，最终稀释 1 倍)、青霉素-链霉素 100 IU/ml(Invitrogen，15140-122)及 1 μg/mL 的重组胰岛素(Invitrogen，97-503311))培养产生多价结合蛋白的经转染 CHO 细胞。其他无血清的 CHO 基础培养基(诸如 CD-CHO 及其类似物)亦可用于制备。

通过蛋白 A 亲和层析法，从已消耗 CHO 培养物上清液纯化融合蛋白。使用一系列层析和过滤步骤(包括病毒减少过滤器)纯化多价结合蛋白。将细胞培养物上清液过滤，接着在 GE Healthcare XK 16/40 柱上经受蛋白 A 琼脂糖亲和层析。在蛋白结合柱之后，将柱以 dPBS 洗涤，接着以 1.0 M NaCl、20 mM 磷酸钠(pH 6.0)洗涤，且接着以 25 mM NaCl、25 mN NaOAc(pH 5.0)洗涤，以移除非特异性结合蛋白。以 100 mM 甘氨酸(Sigma)(pH 3.5)将所结合的蛋白自柱洗脱，且用 0.5 M 2-(N-

吗啉基)乙烷磺酸(MES)(pH 6.0)使 pH 值达到 5.0。将蛋白样本浓缩至 25 mg/mL，以备 GPC 纯化。使用 GE healthcare XK 柱及 Superdex 200 制备级(GE healthcare)，在 GE Healthcare AKTA Explorer 100 Air 装置上进行尺寸排阻层析。

接着将物质浓缩且与 20 mM 磷酸钠和 240 mM 蔗糖一起配制，所得 pH 值为 6.0。将组合物过滤，之后以各种浓度装填入无菌小瓶中，此视所回收物质的量而定。

实施例 3

构建蝎形分子表达盒

将含有如实施例 2 中所述的与恒定亚区连接的合成 2H7 scFv(抗-CD20; SEQ ID NO: 1)的核酸命名为 TRU-015。使用 TRU-015 核酸以及编码小模组免疫药物的合成 scFv 2E12(抗-CD28 VL-VH; SEQ ID NO: 3)和合成 scFv 2E12(抗-CD28 VH-VL; SEQ ID NO: 5)核酸作为用于 PCR 扩增蝎形分子盒的各种组份的模板。结合结构域 1 和恒定亚区的模板或骨架通过 TRU-015(编码与恒定亚区连接的 scFv 2H7(抗-CD20)的核酸)提供，且此模板是于表达载体 pD18 中建构。上述含有具有两种取向(V_L-V_H及 V_H-V_L)中任一取向的 scFv 2E12 的核酸为结合结构域 2 提供编码区。

用于 BD2/连接子插入物的 TRU 015 SSS 铰链 $C_{H2}C_{H3}$

使用含有 SSS 铰链的合成 2H7 scFv IgG1 型，通过充当添加 EcoRI 位点的模板以置换现有终止密码子和 XbaI 位点来形成蝎形分子盒。使用引物 9(SEQ ID NO: 23; 参见表 1)和引物 87(SEQ ID NO: 40; 参见表 1)以及 Platinum PCR 高保真混合物(Invitrogen)，通过 PCR 扩增该分子。将所得 1.5 Kbp 片段纯化且克隆至载体 pCR2.1-TOPO(Invitrogen)中，并被转化至大肠杆菌菌株 TOP10(Invitrogen)中，验证 DNA 序列。

表 1

编号	名称	序列 5'-3'	SEQ ID NO
	PCR 引物		
1	hVK3L-F3H3	GCGATAAAGCTTGCCGCCATGGAA GCACCAGCGCAGCTTCTCTTCC	15
2	hVK3L-F2	ACCAGCGCAGCTTCTCTTCCTCCTG CTACTCTGGCTCCCAGATACCACCG	16
3	hVK3L-F1-2H7VL	GGCTCCCAGATACCACCGGTCAAAT TGTTCTCTCCCAGTCTCCAG	17
4	2H7VH-NheF	GCGATAGCTAGCCAGGCTTATCTAC AGCAGTCTGG	18
5	G4S-NheR	GCGATAGCTAGCCCCACCTCCTCCA GATCCACCACCGCCCGAG	19
6	015VH-XhoR	GCGTACTCGAGGAGACGGTGACCGT GGTCCCTGTG	20
7	G1H-C-XHO	GCAGTCTCGAGCGAGCCCAAATCTTG TGACAAAACTC	21
8	G1H-S-XHO	GCAGTCTCGAGCGAGCCCAAATCTTC TGACAAAACTC	22
9	CH3R-EcoR1	GCGTGAGAATTCTTACCCGGAGACAGG GAGAGGCTC	23
10	G1-XBA-R	GCGACGTCTAGAGTCATTTACCCGGAG ACAGG	24
11	G4SLinkR1-S	AATTATGGTGGCGGTGGCTCGGGCGGT GGTGGATCTGGAGGAGGTGGGAGTGGG	25
12	G4SLinkR1-AS	AATTCCCACTCCCACCTCCTCCAGATCCA CCACCGCCCGAGCCACCGCCACCAT	26
13	2E12VLXbaR	GCGTGTCTAGATTAACGTTTGATTTCCAG CTTGGTG	27
14	2E12VLR1F	GCGATGAATTCTGACATTGTGCTCACCCA ATCTCC	28
15	2E12VHR1F	GCGATGAATTCTCAGGTGCAGCTGAAGGA GTCAG	29
16	2E12VHXbaR	GCGAGTCTAGATTAAGAGGAGACGGTGAC TGAGGTTC	30
17	2e12VHdXbaF1	GGGTCTGGAGTGGCTGGGAATGATATG	31
18	2e12VHdXbaR1	ATTCCCAGCCACTCCAGACCCTTTCCTG	32
19	IgBsrG1F	GAGAACCACAGGTGTACACCCTG	33
20	IgBsrG1R	GCAGGGTGTACACCTGTGGTTCTCG	34

编号	名称	序列 5'-3'	SEQ ID NO
	测序引物		
82	M13R	CAGGAAACAGCTATGAC	35
83	M13F	GTAAAACGACGGCCAGTG	36
84	T7	GTAATACGACTCACTATAGG	37
85	pD18F-17	AACTAGAGAACCCACTG	38
86	pD18F-20	GCTAACTAGAGAACCCACTG	39
87	pD18F-1	ATACGACTCACTATAGGG	40
88	pD18R-s	GCTCTAGCATTTAGGTGAC	41
89	CH3seqF1	CATGAGGCTCTGCACAAC	42
90	CH3seqF2	CCTCTACAGCAAGCTCAC	43
91	CH3seqR1	GGTTCTTGGTCAGCTCATC	44
92	CH3seqR2	GTGAGCTTGCTGTAGAGG	45

表 1: 寡核苷酸引物用于构建 CD20-CD28 蝎形分子盒。引物分成 2 组: PCR 组和测序组。PCR 引物是用于构建所述盒，测序引物是用于证明所有中间构建体和最终构建体的 DNA 序列。

n2H7 V_K 和人类 V_{K3} 前导序列融合体

使用表 1 中的引物 3 和 5，利用寡核苷酸引导的 PCR 诱变将 AgeI (ACCGGT)限制性位点引入 TRU 015 VK 的编码区的 5'端，将 Nhe I (GCTAGC)限制性位点引入(G4S)3 连接子的编码区的 3'端。由于引物 3 亦编码人类 VK3 前导(gb:X01668)的最后 6 个氨基酸，因此使用表 1 中的引物 1、2 和 5，利用重叠 PCR 依序增添前导的 N-末端序列(包括一致性 Kozak 盒和 HinDIII(AAGCTT)限制性位点)。

n2H7 IgG1 SSS 铰链-$C_{H2}C_{H3}$ 构建体

使用引物 4 和 6(分别为 SEQ ID NO: 18 和 20; 表 1)再扩增 TRU-015 V_H，其中使 NheI 位点 5'与 TRU-015 的 V_K 融合且使 Xho I (5'-CTCGAG-3')位点在 3'端与 IgG1 铰链-$C_{H2}C_{H3}$ 结构域接合。同样，使用表 1 中的引物 8 和 9 扩增 IgG1 铰链-C_{H2}-C_{H3} 区，从而引入 5' Xho I 位点，将现有 3'端用 EcoRI (5'-GAATTC-3')位点置换以供克隆，且破坏终止密码子以使得可翻译 CH3 结构域下游所连接的结合结构域 2。此蝎形分子盒型不同于上述字首为"n"的盒。

除上述多价结合蛋白外，本发明的蛋白可具有对应于免疫球蛋白

的单一可变区的结合结构域(结合结构域 1 或 2，或两者)。本发明的该方面的例示性实施方案将包括对应于骆驼科动物抗体的 V_H 结构域的结合结构域，或能够结合靶标抗原的其他物种抗体的经单一修饰或未经修饰的 V 区，尽管涵盖任何适用于本发明的蛋白中的单一可变结构域。

2E12 VL-VH 和 VH-VL 构建体

为制备与所述盒相容的 2E12 scFv，须使用表 1 中的重叠寡核苷酸引物 17 和 18 来破坏内部 Xba I (5'-TCTAGA-3')位点。使用该两个引物与引物对 14/16 (VL-VH)或 13/15 (VH-VL)的组合来扩增两个相反取向的结合结构域，以使得其在其 5'和 3'端分别携有 EcoRI 位点与 XbaI 位点。引物 13 和 16 亦编码 Xba I 位点前方紧邻的终止密码子(TAA)。

2H7 SSS IgG1 2e12 LH/HL 构建体

效应结构域-结合结构域 2 连接子添加(STD 连接子-STD1 和 STD2)

将表 1 中的互补性引物 11 和 12 组合，加热至 70°C，且缓慢冷却至室温以使该两链粘接。利用制造商实验方案，使用具有 1 mM ATP 的 1X 连接缓冲剂(Roche)中的 T4 聚核苷酸激酶(Roche)添加 5'磷酸基。接着使用 T4 DNA 连接酶(Roche)将所得双链连接子接合至位于 IgG1 C_{H3} 末端的编码区与结合结构域 2 的始端之间的 EcoRI 位点中。对所得 DNA 构建体筛检连接子-BD2 接合处的 EcoRI 位点的存在及 C_{H3}-连接子接合处的核苷酸序列 GAATTA 的存在。接着将适当 STD 1 连接子构建体与 EcoRI 一起再消化，且重复连接子连接以产生具有由两个(STD 2)一致性迭代的 Lx1 序列组成的连接子的分子。再次如上筛检 DNA 构建体。

实施例 4

表达研究

对上述编码具有效应功能的多价结合蛋白的核酸进行表达研究。将编码多价结合蛋白的核酸经瞬间转染至 COS 细胞中且将经转染的细胞维持在容许异源基因在所述细胞中受表达的熟知条件下。如上所述，使用 PEI 或 DEAE-葡聚糖将 DNA 瞬间转染至 COS 细胞中

(PEI=Boussif O.等人，PNAS 92: 7297-7301, (1995)，该文献以引用方式并入本文中；Pollard H.等人，JBC 273: 7507-7511, (1998)，该文献以引用方式并入本文中)。对各新颖分子进行多次独立转染以便测定各新形式的平均表达程度。若通过 PEI 转染，则将 COS 细胞涂覆于 DMEM/10% FBS 培养基中的 60 mm 组织培养板上且孵育隔夜，以使其在转染当日长满约 90%。将培养基换为不含抗生素的无血清 DMEM 且孵育 4 小时。转染培养基(4 毫升/板)含有具有 50 μg PEI 和 10-20 μg 所关注 DNA 质粒的无血清 DMEM。将转染培养基通过涡流混合，在室温下孵育 15 分钟，且在抽出现有培养基后添加至板中。将培养物孵育 3-7 天后再收集上清液。通过 SDS-PAGE、蛋白质印迹法测定培养物上清液的蛋白表达，通过流式细胞术验证结合且利用各种测定(包括 ADCC、CDC 和共培养物实验)来测定功能。

SDS-PAGE 分析和蛋白质印迹分析

由每孔含有 8 μg 蛋白的粗培养物上清液(通常 30 微升/孔)或纯化的蛋白等分试样制备样本，且将 2X Tris-甘氨酸 SDS 缓冲剂(Invitrogen)添加至 1 倍最终浓度。对 10 微升 SeeBlue 标记(Invitrogen, Carlsbad, CA)电泳以提供 MW 大小标准。使多价结合(融合)蛋白变异体在 4-20% Novex Tris-甘氨酸凝胶(Invitrogen, San Diego, CA)上经受 SDS-PAGE 分析。在 95°C 下加热 3 分钟后，在还原或非还原条件下使用 Novex Tris-甘氨酸 SDS 样本缓冲剂(2 X)装载样本，继而在 175 V 下电泳 60 分钟。使用 1X Novex Tris-甘氨酸 SDS 电泳缓冲剂(Invitrogen)进行电泳。

电泳后，使用半干式电印迹装置(Ellard, Seattle, WA)将蛋白在 100 mAmp 下转移至 PVDF 膜历时 1 小时。Western 转移缓冲剂包括存在于饱和 Whatman 滤纸上的如下三种缓冲剂：1 号含有 36.34 g/升 Tris (pH 10.4)和 20%甲醇；2 号含有 3.02 g/升 Tris (pH 10.4)和 20%甲醇；3 号含有 3.03 g/升 Tris (pH 9.4)、5.25 g/升 ε-氨基己酸和 20%甲醇。在搅拌下，将膜于 PBS 中的 BLOTTO(=5%脱脂乳)中阻断隔夜。将细胞膜与 5 μg/ml 的 HRP 结合的山羊抗人类 IgG (Fc 特异性，Caltag)一起于 BLOTTO 中孵育一小时，接着于 PBS-0.5% Tween 20 中洗涤 3 次，每次 15 分钟。将湿膜用 ECL 溶液孵育 1 分钟，继而暴露于 X-omat 膜历

时 20 秒钟。图 2 显示在 COS 细胞培养物上清液(30 微升/孔)中所表达的蛋白在非还原条件下电泳的蛋白质印迹。泳道用标记 1-9 指示且含有以下样本：泳道 1 (切去边 = See Blue 标记，kDa 指示印迹侧。泳道 2 = 2H7-sssIgG P238S/P331S-STD1-2e12 VLVH；泳道 3 = 2H7-sssIgG P238S/P331S-STD1-2e12 VHVL；泳道 4 = 2H7-sssIgG P238S/P331S-STD2-2e12 VLVH；泳道 5 = 2H7-sssIgG P238S/P331S-STD2-2e12 VHVL；泳道 6 = 2e12 VLVH SMIP；泳道 7 = 2e12 VHVL SMIP；泳道 8 = 2H7 SMIP。所述构建体中的 2H7 始终为 V_LV_H 取向，sssIgG 指示位于连接子位置 1 的铰链/连接子的特性(如图 5 中所示)，P238S/P331S 指示自野生型(所示第一 aa)突变为突变体(所示第二 aa)的人类 IgG1 型式及其存在于野生型人类 IgG1 C_{H2} 和 C_{H3} 结构域中的氨基酸位置，STD1 指示位于连接子位置 2 的 20 氨基酸(18+限制性位点)连接子(如图 5 中所示)，且 STD2 指示位于连接子位置 2 的 38 氨基酸(36+限制性位点)连接子(如图 6 中所示)。

结合研究

进行结合研究以评估 CD20/CD28 多特异性多价结合肽的双特异性结合特性。初始，将 WIL2-S 细胞添加至 96 孔板且离心为沉淀状细胞。在整块板上使用两倍滴定法(20 μg/ml 滴定至 0.16 μg/ml)，将 CD20/CD28 纯化蛋白添加至经接种的板中。亦将 TRU-015(结合结构域 1 的来源)纯化蛋白的两倍连续稀释液(TRU-015 浓度范围为 20 μg/ml 至 0.16 μg/ml)添加至接种的板孔中。一个不含蛋白的孔充当本底对照。

将含有蛋白的接种板在冰上孵育一小时。随后，将所述孔用 200 μl 1% FBS(于 PBS 中)洗涤一次。接着向各孔中添加以 1:100 经 FITC (Fc Sp)标记的山羊抗人类抗体且将所述板在冰上再孵育一小时。接着将所述板用 200 μl 1% FBS(于 PBS 中)洗涤一次，且将所述细胞再悬浮于 200 μl 1% FBS 中且通过 FACS 分析。

为评估抗-CD28 肽 2E12 V_HV_L 的结合特性，将表达 CD28 的 CHO 细胞通过接种于培养板的个别孔中来涂覆。接着使用两倍稀释方案(浓度范围为 20 μg/ml 至 0.16 μg/ml)将 CD20/CD28 纯化蛋白添加至个别

孔中。再次使用两倍稀释方案(亦即 20 μg/ml 至 0.16 μg/ml)，将 2E12IgG-VHVL SMIP 纯化蛋白添加至个别接种孔中。一个孔不接受蛋白以提供本底对照。接着将所述板在冰上孵育一小时，用 200 μl 1% FBS (于 PBS 中)洗涤一次，且向各孔中添加以 1:100 经 FITC(Fc Sp, CalTag, Burlingame, CA)标记的山羊抗人类抗体。再将所述板于冰上孵育一小时且随后用 200 μl 1% FBS(于 PBS 中)洗涤一次。将所述细胞再悬浮于 200 μl 1% FBS 中之后，进行 FACS 分析。结果显示具有 N-末端 CD20 结合结构域 1 的多价结合蛋白结合 CD20; 具有 N-V_H-V_L-C 取向的 C-末端 CD28 结合结构域 2 的那些蛋白亦结合 CD28。

流式细胞术(FACS)表明所表达的蛋白可结合 WIL-2S 细胞上所呈递的 CD20(参见图 3)且可结合 CHO 细胞所呈递的 CD28(参考图 3)，从而证明 BD1 或 BD2 可具有结合特异性靶标抗原的功能。图 3 的图形上所示的各数据显示，不同多价结合(融合)蛋白的连续稀释液在所指示的滴定范围内具有结合性。使用所述初始构建体所获得的数据指示，在当量浓度下，具有结合结构域 2 型(使用 VHVL 取向的 2e12)的多价结合(融合)蛋白比 VLVH 取向形式更佳地表达且更佳地结合 CD28。

图 4 显示用获自所述转染体/构建体中每一者的经纯化蛋白所进行的结合研究结果的图形表示。该图显示蛋白与表达 CD20 的 WIL-2S 细胞的结合概况，从而证明该多价分子结合 CD20，以及在相同浓度下的单一特异性 SMIP。图 5 的顶图和底图显示 BD2 特异性(2e12=CD28)结合 CD28 CHO 细胞的概况。对于结合结构域 2 与 CD28 的结合而言，V 区的取向影响 2e12 的结合。具有 VH-VL(HL)取向的 2e12 的 2H7-sss-hIgG-STD1-2e12 多价结合蛋白显示以相当于单一特异性 SMIP 的程度结合，而 2e12 LH 分子在相同浓度下显示低效结合。

实施例 5

多价融合蛋白的各种连接子形式的构建

该实施例描述图 6 所示表格中所列的不同连接子形式的构建。

C_{H3}-BD2 连接子 H1 至 H7 的构建

为探究 C_{H3}-BD2 连接子长度和组成对蝎形分子的表达与结合的影响，设计实验以将现有分子 2H7sssIgG1-Lx1-2e12HL 与具有不同连接子的更大组类似构建体相比较。使用 2H7sssIgG1-Lx1-2e12HL 作为模板，使用寡核苷酸表 2 中所列的引物进行一系列 PCR 反应，从而形成长度为 0 至 16 个氨基酸不等的连接子。所述连接子可构建为核酸片段，所述核酸片段跨越 BsrGI 位点处的 C_{H3} 编码区至 EcoRI 位点处的编码连接子-BD2 接点的核酸末端。

表 2

编号	名称	序列 5'-3'	SEQ ID NO.
	PCR 引物		
1	L1-11R	GCGATAGAATTCCCAGATCCACCACCGCCCGAGCCACCGCCACCATAATTC	46
2	L1-6R	GCGATAGAATTCCCAGAGCCACCGCCACCATAATTC	47
3	L3R	GCGTATGAATTCCCCGAGCCACCGCCACCCTTACCCGGAGACAGG	48
4	L4R	GCGTATGAATTCCCAGATCCACCACCGCCCGAGCCACCGCCACCCTTAC	49
5	L5R	GCGTATGAATTCCCGCTGCCTCCTCCCCCAGATCCACCACCGCC	50
6	IgBsrG1F	GAGAACCACAGGTGTACACCCTG	51
7	L-CPPCPR	GCGATAGAATTCGGACAAGGTGGACACCCCTTACCCGGAGACAGGGAGAG	52

表 2: 用于形成 CH3-BD2 连接子变异体的引物的序列

图 6 图示多价结合(融合)蛋白的示意性结构且显示各结合结构域的 V 区的取向，存在于连接子位置 1 的序列(仅列出 Cys 残基)及位于分子的连接子位置 2 的连接子的序列和识别符。

实施例 6

针对 2H7-IgG-2e12 原型多价融合蛋白的变异体连接子形式的结合及功能研究

该实施例显示对连接子位置 2 中存在各种连接子(H1-H7)的“原型”2H7-sssIgG-Hx-2e12 VHVL 构建体的一系列表达和结合研究的结果。如以上实施例中所述，使用蛋白 A 亲和层析法，通过分子的大规模 COS 瞬间转染和纯化来表达所述蛋白中的每一者。接着使经纯化的蛋白经受分析，包括 SDS-PAGE、蛋白质印迹法、通过流式细胞术对结合的分析研究及生物活性的功能性测定。

比较不同 BD2 取向的结合研究

如以上实施例中所述进行结合研究，例外之处在于使用经蛋白 A 纯化的物质，且对于所研究的各变异体使用恒定量的结合(融合)蛋白，亦即 0.72 μg/ml。图 7 显示比较各连接子变异体及 2e12 取向变异体对于 CD20 与 CD28 靶标细胞的结合特性的柱状图。H1-H6 是指具有 H1-H6 连接子及 VH-VL 取向的 2e12 的构建体。L1-L6 是指具有 H1-H6 连接子及 VL-VH 取向的 2e12 的构建体。数据表明，对 2e12 具特异性的结合结构域 2 在以 HL 取向(样本 H1-H6)存在时比以 LH 取向(样本 L1-L6)存在时可更有效地结合。如下一组图所示，发现连接子长度的影响复杂化，复杂之处在于具有较长连接子的分子含有某些单一特异性裂解分子，所述裂解分子在羧基末端缺乏 CD28 结合特异性。进行其他实验以确定所选连接子的结合，结果显示于图 10、12 及 13 中。

纯化的 H1-H7 连接子变异体的 SDS-PAGE 分析

由纯化的蛋白等分制备样本，每孔含有 8 μg 蛋白，且添加 2X Tris-甘氨酸 SDS 缓冲剂(Invitrogen)至 1X 最终浓度。对于还原型样本/凝胶，则将 10X 还原缓冲剂外加 Tris-甘氨酸 SDS 缓冲剂一起添加至样本中至 1X。对 10 μl SeeBlue 标记(Invitrogen, Carlsbad, CA)电泳以提供 MW 大小标准。使多价结合(融合)蛋白变异体在 4-20% Novex Tris-甘氨酸凝胶(Invitrogen, San Diego, CA)上进行 SDS-PAGE 分析。在 95°C 加热 3 分钟后，在还原或非还原条件下使用 Novex Tris-甘氨酸 SDS 样本缓冲剂(2X)装载样本，继而在 175 V 电泳 60 分钟。使用 1X Novex Tris-甘氨酸 SDS 对缓冲剂(Invitrogen)进行电泳。电泳后，在搅拌下凝胶于库马斯(Coomassie)SDS PAGE R-250 染色剂中染色 30 分钟，且脱色至少一小时。图 8 显示[2H7-sss-hIgG P238S/P331S-Hx-2e12 VHVL]多价

结合(融合)蛋白变异体以及作为对照样本的 TRU-015 和 2e12 HLSMIP 的非还原和还原的库马斯染色凝胶。当连接子长度增加时，在约 SMIP 大小(或 52 kDa)电泳的蛋白的量增加。此带中蛋白量的增加对应于上面在约 90 kDa 电泳的带中蛋白量的降低。凝胶数据显示全长分子在或邻近连接子处裂解，形成一个缺失 BD2 区的分子。较小 BD2 片段不存在，显示(1)连接子序列中的核苷酸序列可能形成一个隐蔽的剪接位点，使该较小片段自剪接的 RNA 转录物中移除；或(2)蛋白在全长多肽翻译后经蛋白水解方式裂解，该较小的 BD2 片段不稳定，亦即易受到蛋白水解处理。一些所述分子的蛋白质印迹法分析显示所述蛋白均含有 CD20 BD1 序列，但该较小的带缺乏 CD28 BD2 反应性。在任何凝胶或印迹上均未观测到较小带以“裸”scFv 大小(25-27 kDa)迁移，显示样本中不存在此较小的肽片段。

蛋白质印迹分析 BD1 和 BD2 与 2H7 特异性 Fab 或 CD28mIg 的结合

图 9 显示蛋白质印迹分析 2H7-sss-hIgG-H6 多价结合(融合)蛋白与各单一特异性 SMIP 相比的结果。

在非还原条件下进行电泳，且在装载之前不使样本沸腾。电泳后，使用半干式电印迹装置(Ellard, Seattle, WA)将蛋白在 100 mAmp 下转移至 PVDF 膜历时 1 小时。在搅拌下，将膜于 BLOTTO (PBS 中的 5% 脱脂乳)中阻断隔夜。图 9A: 在 BLOTTO 中将膜与 5 μg/mL 的 AbyD02429.2 (针对 2H7 抗体的 Fab)一起孵育一小时，接着于 PBS-0.5% Tween 20 中洗涤 3 次，每次 5 分钟。接着将膜以 0.5 μg/ml 的浓度于 6X His-HRP 中孵育一小时。将印迹于 PBST 中洗涤三次，每次 15 分钟。将湿膜用 ECL 溶液孵育 1 分钟，继而暴露于 X-omat 膜历时 20 秒钟。

图 9B: 将膜于 BLOTTO 中与 10 μg/ml 的 CD28Ig(Ancell, Bayport, MN)一起孵育，接着于 PBS-0.5% Tween 20 中洗涤三次，每次 15 分钟。接着将膜以 1:3000 的山羊抗小鼠 HRP 结合物(CalTag, Burlingame, CA)孵育于 BLOTTO 中。将膜洗涤三次，每次 15 分钟，接着于 ECL 溶液中孵育 1 分钟，继而暴露于 X-omat 膜历时 20 秒钟。蛋白质印迹分析结果指示，CD28 结合结构域是以多价“单体”组份(该组份在约 90 kDa

下迁移)存在且以更高级别的形式存在。在对于单一 SMIP 或裸 scFv 大小片段预期的位置未观测到带迁移。当使用 CD20 抗遗传型 Fab 时，检测到 SMIP 大小的片段，说明存在含有(2H7-sss-hIgG)且缺乏融合蛋白的 CD28 scFv BD2 部分的肽片段。

对所选连接子的结合研究

图 10 显示对经纯化的 2H7-sss-hIgG-Hx-2e12 融合蛋白所进行的结合研究的结果。进行结合研究以评估 CD20/CD28 多特异性结合肽的双特异性结合特性。初始时利用常规技术涂覆 WIL2-S 细胞。在整块板上使用两倍滴定法(20 μg/ml 至 0.16 μg/ml)将 CD20/CD28 纯化蛋白添加至经接种的板中。亦将 TRU-015(结合结构域 1 的来源)纯化蛋白的两倍连续稀释液(TRU-015 的浓度范围为 20 μg/ml 至 0.16 μg/ml)添加至经接种的板孔中。一个不含蛋白的孔充当本底对照。

将含有蛋白的经接种板在冰上孵育一小时。随后，将所述孔用 200 μl 1% FBS(于 PBS 中)洗涤一次。接着向各孔中添加以 1:100 经 FITC(Fc Sp)标记的山羊抗人类抗体且将所述板在冰上再孵育一小时。接着将所述板用 200 μl 1% FBS(于 PBS 中)洗涤一次，且将所述细胞再悬浮于 200 μl 1% FBS 中且通过 FACS 分析。

为评估抗-CD28 肽 2E12 V_HV_L 的结合特性，将表达 CD28 的 CHO 细胞通过接种于培养板的个别孔中来涂覆。接着使用两倍稀释方案(浓度范围为 20 μg/ml 至 0.16 μg/ml)将 CD20/CD28 纯化蛋白添加至个别孔中。再次使用两倍稀释方案(亦即 20 μg/ml 至 0.16 μg/ml)，将 2E12IgGvHvL SMIP 纯化蛋白添加至个别接种孔中。一个孔不接收蛋白以提供本底对照。接着将所述板在冰上孵育一小时，用 200 μl 1% FBS(于 PBS 中)洗涤一次，且向各孔中添加以 1:100 经 FITC(Fc Sp)标记的山羊抗人类抗体。再将所述板在冰上孵育一小时且随后用 200 μl 1% FBS(于 PBS 中)洗涤一次。在将所述细胞再悬浮于 200 μl 1% FBS 中之后，进行 FACS 分析。流式细胞术(FACS)表明经表达的蛋白可结合 WIL-2S 细胞上所呈递的 CD20(参见图 10A)且可结合 CHO 细胞上所呈递的 CD28(参考图 10B)，从而表明 BD1 或 BD2 可具有结合特异性靶标抗原的功能。此外，未发现所用连接子(H1-H6)对结合靶标抗原

的亲和力具有显著影响。

SEC分离多价结合(融合)蛋白。在GE Healthcare XK 16/40管柱上，通过蛋白 A 琼脂糖亲和层析法自细胞培养物上清液中纯化结合(融合)蛋白。在蛋白结合至管柱后，将管柱以 dPBS 洗涤，接着以 1.0 M NaCl、20 mM 磷酸钠(pH 6.0)洗涤，且接着以 25 mM NaCl、25 mN NaOAc(pH 5.0)洗涤，以移除非特异性结合蛋白。以 100 mM 甘氨酸(Sigma)(pH 3.5)将所结合的蛋白自管柱中洗脱，且用 0.5 M 2-(N-吗啉基)乙烷磺酸(MES)(pH 6.0)使其 pH 值达到 5.0。利用常规技术将蛋白样本浓缩至 25 mg/ml，以备 GPC 纯化。使用 GE healthcare XK 管柱及 Superdex 200 制备级(GE healthcare)，在 GE Healthcare AKTA Explorer 100 Air 装置上进行尺寸排阻层析(SEC)。

图 12 显示的表格概括不同结合(融合)蛋白的 SEC 分离的结果。随连接子长度增加，溶液中分子的复合物亦增加，使得难以通过 HPLC 自更高级别形式分离所关注的峰(或 POI)。H7 连接子似乎可将此大部分复合物在溶液中拆分为更均质的形式，以使得可溶性形式在约 172 kDa 下大部分以单一 POI 形式迁移。

其他结合研究

对更小子集的多价结合(融合)蛋白进行第二系列实验(参见图 12 和图 13)，此次比较连接子 H3、H6 及 H7。数据表明，与结合 CD20 相比，结合 CD28 的程度下降更明显，但当连接子长度增加时，两者均略微下降。此外，数据显示 H7 连接子对两种抗原均显示最高程度的结合。使用经蛋白 A 纯化，但未经 SEC 进一步纯化的多价结合(融合)蛋白获得所述数据，因此溶液中可存在多种形式的分子。结果亦指示，截短形式比实际多价多肽的稳定性更差，因为结合曲线似乎未完全反映溶液中所存在的连接子 H6 的大量单一特异性形式。

证明单一分子的多特异性结合

进行替代性结合测定(参见图 13)，其中用对 CD28 BD2 具特异性的试剂检测 CD20 与 WIL-2S 细胞表面的结合，从而表明同一多特异性结合(融合)蛋白上的 BD1 与 BD2 可同时发生与两靶标抗原的结合(参考图 12)。该测定证明蛋白的多特异性结合特性。

实施例 7

具有替代特异性的 BD2 的多特异性结合(融合)蛋白的构建

除原型 CD20-CD28 多特异性结合分子外,可制备具有替代性结合结构域 2 个区(包括 CD37 结合结构域和 CD3 结合结构域)的其他两种形式。亦可制备具有关于[2H7-sss-IgG-Hx/STDx-2e12 HL]多特异性结合(融合)蛋白所述的若干连接子结构域的分子。所述其他多特异性结合(融合)分子的构建描述如下。

抗-CD37 结合结构域构建体

表 3

编号	名称	序列	SEQ ID NO.
		ACTGCTGCAGCTGGACCGCGCT	
23	G281LH-NheR	AGCTCCGCCGCCACCCGAC	53
		GGCGGAGCTAGCGCGGTCCAGC	
24	G281LH-NheF	TGCAGCAGTCTGGACCTG	54
		GCGATCACCGGTGACATCCAGAT	
25	G281-LH-LPinF	GACTCAGTCTCCAG	55
		GCGATACTCGAGGAGACGGTGAC	
26	G281-LH-HXhoR	TGAGGTTCCTTGAC	56
		GCGATCGAATTCAGACATCCAGAT	
27	G281-LH-LEcoF	GACTCAGTCTCCAG	57
		GCGATTCTAGATTAGGAAGAGACG	
28	G281-LH-HXbaR	GTGACTGAGGTTCCTTGAC	58
		GCGATAACCGGTGCGGTCCAGCTG	
29	G281-HL-HF	CAGCAGTCTGGAC	59
		GACCCACCACCGCCCGAGCCACCG	
		CCACCAGAAGAGACGGTGACTGAGG	60
30	G281-HL-HR3	TTC	
		ACTCCCGCCTCCTCCTGATCCGCCG	
31	G281-HL-HR2	CCACCCGACCCACCACCGCCCGAG	61
		GAGTCATCTGGATGTCGCTAGCACTC	
32	G281-HL-HNheR	CCGCCTCCTCCTGATC	62
		ATCAGGAGGAGGCGGGAGTGCTAGC	
33	G281-HL-LNheF	GACATCCAGATGACTCAGTC	63
		GCGATACTCGAGCCTTTGATCTCCAG	
34	G281-HL-LXhoR	TTCGGTGCCTC	64

		GCGATATCTAGACTCAACCTTTGATCT	
35	G281-HL-LXbaR	CCAGTTCGGTGCCTC	65
		GCGATAGAATTCGCGGTCCAGCTGCA	
36	G281-HL-EcoF	GCAGTCTGGAC	66

表 3: 用于形成 SMIP 分子和蝎分子的 G28-1 抗-CD37 结合结构域的寡核苷酸引物。

使用上文表 X 中的引物，通过 PCR 将 G28-1 scFv (SEQ ID NO: 102) 转化为 G28-1 LH SMIP。将引物 23 和 25 与 10 ng G28-1 scFv 组合，使用 Platinum PCR Supermix Hi-Fidelity PCR 混合器(Invitrogen, Carlsbad, CA)于 ABI 9700 热循环器中扩增 VK (30 个如下循环：94°C 下历时 20 秒钟、58°C 下历时 15 秒钟、68°C 下历时 15 秒钟)。此 PCR 产物在 VK 的 5'端具有限制性位点 PinAI (AgeI)且在 scFv (G4S)3 连接子的末端具有 NheI。通过在 PCR 运作中，在与上文 VK 相同的条件下将引物 24 和 26 与 10 ng G28-1 scFv 组合来类似地改变 VH。此 PCR 产物在 VH 的 5'端具有限制性位点 NheI 且在 3'端具有 XhoI。由于重要的序列一致性重叠已工程化为引物 23 和 24，因此将 VK 和 VH 稀释 5 倍，接着以 1:1 比率使用侧翼引物 25 和 26 添加至 PCR 且通过将 68°C 持续时间由 15 秒钟延长至 45 秒钟来如上扩增全长 scFv。此 PCR 产物代表整个 G28-1 scFv 作为 PinAI-XhoI 片段，且可通过 MinElute 管柱(Qiagen)纯化来纯化以移除过量引物、酶及盐。将洗脱物在 37°C 下于 50 μL 体积的 1X H 缓冲剂(Roche)中用 PinAI (Invitrogen)和 XhoI (Roche)消化 4 小时直至完成。接着将经消化的 PCR 产物于 1%琼脂糖凝胶中电泳，将片段自凝胶中移除且使用缓冲剂 QG 在 MinElute 管柱上再纯化，且伴随间歇性混合将凝胶-缓冲剂混合物在 50°C 下孵育 10 分钟以溶解琼脂糖，之后如同引物移除后 PCR 在管柱上进行纯化。将 3 μL 经 PinAI-XhoI 消化的 G28-1 LH 与 1 μL 经 PinAI-XhoI 消化的 pD18-n2H7sssIgG1 SMIP 及 5 μL 2X LigaFast 连接缓冲剂(Promega, Madison, WI)及 1 μL T4 DNA 连接酶(Roche)组合成 10 μL 反应物，充分混合且在室温下孵育 10 分钟。接着利用制造商方案将 3 μL 该连接物转化到感受态 TOP 10 (Invitrogen)。将所述转化体涂覆于具有 100 μg/ml 羧苄西林(Teknova)的 LB 琼脂板上且在 37°C 下孵育隔夜。培养

18 小时后，拣选菌落且接种于深孔型 96 孔板中的含有 100 μg/ml 羧苄西林的 1 ml T-Broth (Teknova)中，且在 37°C 震荡恒温箱中培养隔夜。培养 18 至 24 小时后，使用 QIAprep 96 Turbo 试剂盒(Qiagen)在 BioRobot8000 (Qiagen)上自各隔夜培养物中分离 DNA。接着自各克隆取 10 μL，于 1X B 缓冲剂中以 15 μL 的反应体积用 HindIII 和 XhoI 限制性酶消化。将经消化的 DNA 在 1%琼脂糖 E-凝胶(Invitrogen, CA)上电泳以供限制性位点分析。对含有适当大小的 HindIII-XhoI 片段的克隆进行序列验证。通过使用上文表 X 中的引物 29、30、31 和 32 将 PinAI 位点安置于 5'端且使(G4S) 4 连接子端接于 G28-1 VH 的 Nhe I 位点，以类似方式来构建 G28-1 HL SMIP。使用表 X 中的引物 33 和 34 通过 PCR 来改变 VK，以便将 NheI 位点引入 VK 的 5'端且将 XhoI 引入 3'端。接着将所述 PCR 如上组合且用侧翼引物 29 和 34 扩增以产生 VH-VL 取向的完整 G28-1 scFv DNA，恰如同 G28-1 LH SMIP 将其克隆至经 PinAI-XhoI 消化的 pD18-(n2H7)sssIgG1 SMIP 中。

2H7sssIgG1-STD1-G28-1 LH/HL 构建体

使用 G28-1 LH 及 G28-1 HL SMIP 作为模板，通过 PCR 改变 LH 和 HL 抗-CD37 结合结构域，以使得其侧翼限制性位点与蝎形分子盒相容。使用引物 27 (LH)或 36 (HL)将 EcoRI 位点引入各 scFv 的 5'端，且使用引物 28 (LH)或 35 (HL)将终止密码子/XbaI 位点引入 3'端。将所得 DNA 克隆至经 EcoRI-XbaI 消化的 pD18-2H7sssIgG-STD1 中。

2H7sssIgG1-Hx-G28-1 HL 构建体

将 2H7sssIgG1-Hx-2e12 HL DNA 与 BsrGI 和 EcoRI 及由 IgG1 的 C-末端与连接子组成的 325 bp 片段一起消化。通过使用 BsrGI-EcoRI 移除 STD1 连接子将所述物用 2H7sssIgG1-STD1-G19-4 HL 中的等效区取代，且将其用 2H7sssIgG1-Hx-2e12 HL 克隆中的对应连接子置换。

抗-CD3 结合结构域构建体

表 4

编号	名称	序列	SEQ ID NO.
		GCGTATGAACCGGTGACATCCAGAT	
37	194-LH-LF1	GACACAGACTACATC	67
		ATCCAGATGACACAGACTACATCCTC	
38	194-LF2	CCTGTCTGCCTCTCTGGGAGACAG	68
		GTCTGCCTCTCTGGGAGACAGAGTCA	
39	194-LF3	CCATCAGTTGCAGGGCAAGTCAGGAC	69
		GTTGCAGGGCAAGTCAGGACATTCGC	
40	194-LF4	AATTATTTAAACTGGTATCAGCAG	70
		ATTTAAACTGGTATCAGCAGAAACCAG	
41	194-LF5	ATGGAACTGTTAAACTCCTGATC	71
		GAACTGTTAAACTCCTGATCTACTACA	
42	194-LF6	CATCAAGATTACACTCAGGAGTC	72
		CAAGATTACACTCAGGAGTCCCATCAA	
43	194-LF7	GGTTCAGTGGCAGTGGGTCTGGAAC	73
		CAGGTTGGCAATGGTGAGAGAATAATC	
44	194-LR7	TGTTCCAGACCCACTGCCACTGAAC	74
		GCAAAAGTAAGTGGCAATATCTTCTGGT	
45	194-LR6	TGCAGGTTGGCAATGGTGAGAG	75
		GAACGTCCACGGAAGCGTATTACCC	
46	194-LR5	TGTTGGCAAAAGTAAGTGGCAATATC	76
		CGTTTGGTTACCAGTTTGGTGCCTCCAC	
47	194-LR4	CGAACGTCCACGGAAGCGTATTAC	77
		ACCACCGCCCGAGCCACCGCCACC	
48	194-LR3	CCGTTTGGTTACCAGTTTGGTG	78
		GCTAGCGCTCCCACCTCCTCCAGATCCA	
49	194-LR2	CCACCGCCCGAGCCACCGCCAC	79
		GTTGCAGCTGGACCTCGCTAGCGCT	
50	194-LH-LR1	CCCACCTCCTCCAGATC	80
		GATCTGGAGGAGGTGGGAGCGCTAGC	
51	194-LH-HF1	GAGGTCCAGCTGCAACAGTCTGGACCTG	81
		AGCTGCAACAGTCTGGACCTGAACT	
52	194-HF2	GGTGAAGCCTGGAGCTTCAATGAAG	82
		AGCCTGGAGCTTCAATGAAGATTTCC	
53	194-HF3	TGCAAGGCCTCTGGTTACTCATTC	83
		GCAAGGCCTCTGGTTACTCATTCACT	
54	194-HF4	GGCTACATCGTGAACTGGCTGAAGCAG	84

		ATCGTGAACTGGCTGAAGCAGAGCC	
55	194-HF5	ATGGAAAGAACCTTGAGTGGATTGGAC	85
		GAACCTTGAGTGGATTGGACTTATTA	
56	194-HF6	ATCCATACAAAGGTCTTACTACCTAC	86
		AATGTGGCCTTGCCCTTGAATTTCTG	
57	194-HR6	GTTGTAGGTAGTAAGACCTTTGTATG	87
		CATGTAGGCTGTGCTGGATGACTTGT	
58	194-HR5	CTACAGTTAATGTGGCCTTGCCCTTG	88
		ACTGCAGAGTCTTCAGATGTCAGACTG	
59	194-HR4	AGGAGCTCCATGTAGGCTGTGCTGGATG	89
		ACCATAGTACCCAGATCTTGCACAG	
60	194-HR3	TAATAGACTGCAGAGTCTTCAGATGTC	90
		GCGCCCCAGACATCGAAGTACCAGTC	
61	194-HR2	CGAGTCACCATAGTACCCAGATCTTG	91
		GCGAATACTCGAGGAGACGGTGACCG	
62	194-LH-HR1	TGGTCCCTGCGCCCCAGACATCGAAG	92
		GCGTATGAACCGGTGAGGTCCAGC	
63	194-HL-HF1	TGCAACAGTCTGGACCTG	93
		ACCGCCACCAGAGGAGACGGTGACCGT	
64	194-HL-HR1	GGTCCCTGCGCCCCAGACATCGAAGTAC	94
		ACCTCCTCCAGATCCACCACCGCCCG	
65	194-HL-HR0	AGCCACCGCCACCAGAGGAGACGGTG	95
		GCGGGGGAGGTGGCAGTGCTAGCGA	
66	194-HL-LF1	CATCCAGATGACACAGACTACATC	96
		GCGAATACTCGAGCGTTTGGTTACCA	
67	194-HL-LR3Xho	GTTTGGTG	97
		GCGATATCTAGATTACCGTTTGGTTAC	
68	194-HL-LR3Xba	CAGTTTGGTG	98
		GCGTATGAGAATTCAGAGGTCCAGCTG	
69	194-HL-HF1R1	CAACAGTCTGGACCTG	99
		GCGTATGAGAATTCTGACATCCAGA	
70	194-LH-LF1R1	TGACACAGACTACATC	100
		GCGTATCTAGATTAGGAGACGGTGACC	
71	194-LH-HR1Xba	GTGGTCCCTGCGCCCCAGACATCGAAG	101

表 4: 用于形成 G19-4 scFv 序列的抗-CD3 结合结构域的寡核苷酸。

通过如上所述延长重叠寡核苷酸引物来合成 G19-4 结合结构域。轻链 PCR 是以两步骤进行，起始为将浓度分别为 5 μM、10 μM、20 μM

和 40 μM 的引物 43/44、42/45、41/46 和 40/47 于 Platinum PCR Supermix Hi-Fidelity 中组合进行如下的 30 个循环：94°C 下历时 20 秒钟、60°C 下历时 10 秒钟、68°C 下历时 15 秒钟。将 1 μL 所得 PCR 产物利用相同的 PCR 条件(例外之处为将 68°C 持续时间增加至 25 秒钟)，使用 39/48 (10 μM)、38/49 (20 μM)和 37/50 (40 μM)的引物混合物(对于 LH 取向)或 66/67 (40 μM)的引物混合物(对于 HL 取向)再扩增。具有 LH 取向的 VK 在 5'端由 PinAI 结合且在 3'端由 NheI 结合，而具有 HL 取向的 VK 在 5'端由 NheI 结合且在 3'端由 XhoI 结合。

为合成重链，通过在第一 PCR 步骤将引物 56/57、55/58、54/59 和 53/60 组合来制备具有与上文相同浓度的引物混合物。在第二 PCR 中，利用与轻链的第二 PCR 相同的 PCR 条件以来自第一 PCR 的 1 μl 来扩增引物 52/61 (20 μM)和 51/62 (50 μM)，以形成在 5'端具有 NheI 且在 3'端具有 XhoI 的 LH 取向。在第二 PCR 中，将引物 52/61 (10 μM)、63/64 (20 μM)、63 (20 μM)/65 (40 μM)和 63 (20 μM)/5 (80 μM)与来自前一 PCR 的 1 μL 组合，以形成在 5'端具有 PinAI 且在 3'端具有 NheI 的 HL 取向的重链。如同上述构建体，将充足重叠设计为以 NheI 位点为中心的引物，以便通过将具有 LH 取向的重链及轻链 PCR 组合且用侧翼引物 37 和 62 再扩增来合成 G19-4 LH，且通过将 HL PCR 组合且用引物 63 和 67 再扩增来合成 G19-4 HL。

将全长 G19-4 LH/HL PCR 产物通过琼脂糖凝胶电泳来分离，自凝胶中切除且用如上所述的 Qiagen MinElute 管柱纯化。接着将所述 DNA 经 TOPO 克隆至 pCR2.1 (Invitrogen)中，转化到 TOP10，且将菌落首先通过 EcoRI 片段大小筛检，接着通过 DNA 测序筛检。接着经由 PinAI-XhoI 将 G19-4 LH/HL 克隆至 pD18-IgG1 中以便在哺乳动物细胞中表达。

2H7sssIgG1-STD1-G19-4 LH/HL 构建体

使用 G19-4 LH 和 G19-4 HL SMIP 作为模板，通过 PCR 改变 LH 和 HL 抗-CD3 结合结构域，以使得其侧翼限制性位点与蝎形分子盒相容。使用引物 27 (LH)或 36 (HL)将 EcoRI 位点引入各 scFv 的 5'端，且使用引物 28 (LH)或 35 (HL)将终止密码子/XbaI 位点引入 3'端。将所

得 DNA 克隆至经 EcoRI-XbaI 消化的 pD18-2H7sssIgG-STD1 中。

2H7sssIgG1-Hx-G19-4 HL 构建体

将 2H7sssIgG1-Hx-2e12 HL DNA 与 BsrGI 和 EcoRI 及由 IgG1 的 C-末端与连接子组成的 325 bp 片段一起消化。通过使用 BsrGI-EcoRI 移除 STD1 连接子将所述物用 2H7sssIgG1-STD1-G19-4 HL 中的等效区取代，且将其用 2H7sssIgG1-Hx-2e12 HL 克隆中的对应连接子置换。

考虑本文中所揭示的各种多价结合蛋白后可显而易见易组合形成本发明的分子的所述分子的特征。该特征包括结合结构域 1、恒定亚区(包括铰链或铰链样结构域)、连接子结构域及结合结构域 2。所述新颖结合蛋白设计中的固有模组化使得本领域技术人员可直接在任何所要模组的 N-末端和/或 C-末端操控 DNA 序列，以使得该序列可插入几乎任何位置以形成显示与其所来源的亲本分子相比经修改或增强的功能性的新颖分子。例如，涵盖源自免疫球蛋白超家族的成员的任何结合结构域作为本发明的分子的结合结构域 1 或结合结构域 2。所述所衍生的结合结构域包括：具有与免疫球蛋白超家族成员的序列一一对应的氨基酸序列及甚至编码聚核苷酸序列的结构域，以及与免疫球蛋白超家族成员优选共有 80%、90%、95%、99%或 99.5%序列一致性的变异体及衍生物。所述结合结构域(1 及 2)优选经由如本文中其他处所述的具有不同序列及长度的连接子与本发明分子的其他模组相连接，其限制条件为所述连接子足以提供分子达成功能性三级结构所必需的任何空间及柔性。多价结合蛋白的另一模组为铰链区，其不仅可对应于免疫球蛋白超家族成员的铰链区，而且可为其变异体，诸如本文中所述的“CSC”或“SSS”铰链区。此外，恒定亚区包含本发明的蛋白的模组，其可对应于免疫球蛋白超家族成员的恒定区的亚区，如铰链-C_{H2}-C_{H3} 恒定亚区的结构所说明。亦涵盖恒定亚区的变异体及衍生物，其优选具有与免疫球蛋白超家族成员共有 80%、90%、95%、99%或 99.5%序列一致性的氨基酸序列。

具有所述分子特征的例示性一级结构显示于表 5 中，其揭示说明性结合结构域 1 和 2 的聚核苷酸和同源氨基酸序列，以及恒定亚区的一级结构(包括铰链或铰链样结构域)及可插入例如多价结合蛋白的恒

定亚区的 C-末端与结合结构域 2 区的 N-末端之间的连接子。本发明分子的其他范例包括上述特征，其中例如结合结构域 1 与结合结构域 2 中的任一者或两者包含源自免疫球蛋白超家族成员的 V_L 或 V_L 样结构域及源自免疫球蛋白超家族的相同或不同成员的 V_H 或 V_H 样结构域的结构域，所述结构域可由如本文中所揭示的任何连接子所代表的连接子隔开。涵盖其中所述结构域中 BD1 和/或 BD2 的取向为 V_L-V_H 或 V_H-V_L 的分子。具有本发明的多价结合分子的各种特征的一级结构的更完整表示是存在于本公开结尾所附的表格中。本发明进一步包括编码所述分子的聚核苷酸。

表 5

结合结构域	核苷酸序列	氨基酸序列	SEQ ID NOS. (氨基酸序列)
2H7 LH	atggattttcaagtgcagattttcagcttcctgcta atcagtgcttcagtcataatgtccagaggacaaa ttgttctctcccagtctccagcaatcctgtctgcat ctccaggggagaaggtcacaatgacttgcagg gccagctcaagtgtaagttacatgcactggtacc agcagaagccaggatcctccccaaaccctgg atttatgccccatccaacctggcttctggagtccc tgctcgcttcagtggcagtgggtctgggacctct tactctctcacaatcagcagagtggaggctgaa gatgctgccacttattactgccagcagtggagttt taacccacccacgttcggtgctgggaccaagct ggagctgaaagatggcggtggctcgggcggtg gtggatctggaggaggtgggagctctcaggctt atctacagcagtctggggctgagtcggtgaggc ctggggcctcagtgaagatgtcctgcaaggctt ctggctacacatttaccagttacaatatgcactgg gtaaagcagacacctagacagggcctggaatg gattggagctatttatccaggaaatggtgatactt cctacaatcagaagttcaagggcaaggccaca ctgactgtagacaaatcctccagcacagcctac atgcagctcagcagcctgacatctgaagactct gcggtctatttctgtgcaagagtggtgtactatag taactcttactggtacttcgatgtctggggcacag ggaccacggtcaccgtctct	mdfqvqifsfllisasvimsrgqivlsqspails aspgekvtmtcrasssvsymhwyqqkpgss pkpwiyapsnlasgvparfsgsgsgtsysltis rveaedaatyycqqwsfnpptfgagtklelkd gggsggggsggggssqaylqqsgaesvrpga svkmsckasgytftsynmhwvkqtprqgle wigaiypgngdtsynqkfkgkatltvdkssst aymqlssltsedsavyfcarvvyysnsywyf dvwgtgttvtvs	1 (2)

结合结构域	核苷酸序列	氨基酸序列	SEQ ID NOS. (氨基酸序列)
2e12 LH	atggattttcaagtgcagattttcagcttcctgcta atcagtgcttcagtcataatgtccagaggagtcg acattgtgctcacccaatctccagcttctttggct gtgtctctaggtcagagagccaccatctcctgca gagccagtgaaagtgttgaatattatgtcacaag tttaatgcagtggtaccaacagaaaccaggaca gccacccaaactcctcatctctgctgctagcaac gtagaatctggggtccctgccaggtttagtggca gtgggtctgggacagactttagcctcaacatcca tcctgtggaggaggatgatattgcaatgtatttct gtcagcaaagtaggaaggttccatggacgttcg gtggaggcaccaagctggaaatcaaacggggt ggcggtggatccggcggaggtgggtcgggtg gcggcggatctcaggtgcagctgaaggagtca ggacctggcctggtggcgccctcacagagcct gtccatcacatgcaccgtctcagggttctcattaa ccggctatggtgtaaactgggttcgccagcctc caggaaagggtctggagtggctgggaatgatat ggggtgatggaagcacagactataattcagctc tcaaatccagactatcgatcaccaaggacaact ccaagagccaagttttcttaaaaatgaacagtct gcaaactgatgacacagccagatactactgtgc ccgagatggttatagtaactttcattactatgttat ggactactggggtcaaggaacctcagtcaccgt ctcctct	MDFQVQIFSFLLISASVIMSRGV DIVLTQSPASLAVSLGQRATISC RASESVEYYVTSLMQWYQQKP GQPPKLLISAASNVESGVPARFS GSGSGTDFSLNIHPVEEDDIAM YFCQQSRKVPWTFGGGTKLEIK RGGGGSGGGGSGGGGSQVQLK ESGPGLVAPSQSLSITCTVSGFSL TGYGVNWVRQPPGKGLEWLG MIWGDGSTDYNSALKSRLSITK DNSKSQVFLKMNSLQTDDTAR YYCARDGYSNFHYYVMDYWG QGTSVTVSS	3 (4)
2e12 HL	atggattttcaagtgcagattttcagcttcctgcta atcagtgcttcagtcataatgtccagaggagtcc aggtgcagctgaaggagtcaggacctggcctg gtggcgccctcacagagcctgtccatca catgcaccgtctcagggttctcattaaccggcta tggtgtaaactgggttcgccagcctccaggaaa gggtctggagtggctgggaatgatatggggtga tggaagcacagactataattcagctctcaaatcc agactatcgatcaccaaggacaactccaagagc caagttttcttaaaaatgaacagtctgcaaactga tgacacagccagatactactgtgcccgagatgg ttatagtaactttcattactatgttatggactactgg ggtcaaggaacctcagtcaccgtctcctctggg ggtggaggctctggtggcggtggatccggcgg aggtgggtcgggtggcggcggatctgacattgt	MDFQVQIFSFLLISASVIMSRGV QVQLKESGPGLVAPSQSLSITCT VSGFSLTGYGVNWVRQPPGKG LEWLGMIWGDGSTDYNSALKS RLSITKDNSKSQVFLKMNSLQT DDTARYYCARDGYSNFHYYVM DYWGQGTSVTVSSGGGGSGGG GSGGGGSGGGGSDIVLTQSPAS LAVSLGQRATISCRASESVEYY VTSLMQWYQQKPGQPPKLLIS AASNVESGVPARFSGSGSGTDF SLNIHPVEEDDIAMYFCQQSRK VPWTFGGGTKLEIKR	5 (6)

结合结构域	核苷酸序列	氨基酸序列	SEQ ID NOS. (氨基酸序列)
	gctcacccaatctccagcttctttggctgtgtctct aggtcagagagccaccatctcctgcagagcca gtgaaagtgttgaatattatgtcacaagtttaatg cagtggtaccaacagaaaccaggacagccacc caaactcctcatctctgctgctagcaacgtagaa tctggggtccctgccaggtttagtggcagtgggt ctgggacagactttagcctcaacatccatcctgt ggaggaggatgatattgcaatgtatttctgtcag caaagtaggaaggttccatggacgttcggtgga ggcaccaagctggaaatcaaacgt		
G28-1 LH	accggtgacatccagatgactcagtctccagcc tccctatctgcatctgtgggagagactgtcacca tcacatgtcgaacaagtgaaaatgtttacagttat ttggcttggtatcagcagaaacagggaaaatctc ctcagctcctggtctcttttgcaaaaaccttagca gaaggtgtgccatcaaggttcagtggcagtgga tcaggcacacagttttctctgaagatcagcagcc tgcagcctgaagattctggaagttatttctgtcaa catcattccgataatccgtggacgttcggtggag gcaccgaactggagatcaaaggtggcggtggc tcgggcggtggtgggtcgggtggcggcggatc tgctagcgcagtccagctgcagcagtctggacc tgagctggaaaagcctggcgcttcagtgaagat ttcctgcaaggcttctggttactcattcactggcta caatatgaactgggtgaagcagaataatggaaa gagccttgagtggattggaaatattgatccttatt atggtggtactacctacaaccggaagttcaagg gcaaggccacattgactgtagacaaatcctcca gcacagcctacatgcagctcaagagtctgacat ctgaggactctgcagtctattactgtgcaagatc ggtcggccctatggactactggggtcaaggaa cctcagtcaccgtctcgag	DIQMTQSPASLSASVGETVTITC RTSENVYSYLAWYQQKQGKSP QLLVSFAKTLAEGVPSRFSGSGS GTQFSLKISSLQPEDSGSYFCQH HSDNPWTFGGGTELEIKGGGGS GGGGSGGGGSASAVQLQQSGP ELEKPGASVKISCKASGYSFTG YNMNWVKQNNGKSLEWIGNID PYYGGTTYNRKFKGKATLTVD KSSSTAYMQLKSLTSEDSAVYY CARSVGPMDYWGQGTSVTVS	102 (103)
G28-1 HL	accggtgaggtccagctgcaacagtctggacct gaactggtgaagcctggagcttcaatgaagattt cctgcaaggcctctggttactcattcactggcta catcgtgaactggctgaagcagagccatggaa agaaccttgagtggattggacttattaatccatac aaaggtcttactacctacaaccagaaattcaagg gcaaggccacattaactgtagacaagtcatcca gcacagcctacatggagctcctcagtctgacat	EVQLQQSGPELVKPGASMKISC KASGYSFTGYIVNWLKQSHGK NLEWIGLINPYKGLTTYNQKFK GKATLTVDKSSSTAYMELLSLTS EDSAVYYCARSGYYGDSDWYF DVWGAGTTVTVSSGGGGSGGG GSGGGGSGGGGSASDIQMTQT TSSLSASLGDRVTISCRASQDIR	104 (105)

结合结构域	核苷酸序列	氨基酸序列	SEQ ID NOS. (氨基酸序列)
	ctgaagactctgcagtctattactgtgcaagatct gggtactatggtgactcggactggtacttcgatg tctggggcgcagggaccacggtcaccgtctcct ctggtggcggtggctcgggcggtggtggatct ggaggaggtgggagcgggggaggtggcagt gctagcgacatccagatgacacagactacatcc tccctgtctgcctctctgggagacagagtcacca tcagttgcagggcaagtcaggacattcgcaatta tttaaactggtatcagcagaaaccagatggaact gttaaactcctgatctactacacatcaagattaca ctcaggagtcccatcaaggttcagtggcagtgg gtctggaacagattattctctcaccattgccaacc tgcaaccagaagatattgccacttacttttgccaa cagggtaatacgcttccgtggacgttcggtgga ggcaccaaactggtaaccaaacgctcgag	NYLNWYQQKPDGTVKLLIYYT SRLHSGVPSRFSGSGSGTDYSLT IANLQPEDIATYFCQQGNTLPW TFGGGTKLVTKRS	
G19-4 LH	accggtgacatccagatgacacagactacatcc tccctgtctgcctctctgggagacagagtcacca tcagttgcagggcaagtcaggacattcgcaatta tttaaactggtatcagcagaaaccagatggaact gttaaactcctgatctactacacatcaagattaca ctcaggagtcccatcaaggttcagtggcagtgg gtctggaacagattattctctcaccattgccaacc tgcaaccagaagatattgccacttacttttgccaa cagggtaatacgcttccgtggacgttcggtgga ggcaccaaactggtaaccaaacggggtggcg gtggctcgggcggtggtggatctggaggaggt gggagcgctagcgaggtccagctgcaacagtc tggacctgaactggtgaagcctggagcttcaat gaagatttcctgcaaggcctctggttactcattca ctggctacatcgtgaactggctgaagcagagcc atggaaagaaccttgagtggattggacttattaat ccatacaaaggtcttactacctacaaccagaaat tcaagggcaaggccacattaactgtagacaagt catccagcacagcctacatggagctcctcagtct gacatctgaagactctgcagtctattactgtgca agatctgggtactatggtgactcggactggtact tcgatgtctggggcgcagggaccacggtcacc gtctcctcgag	DIQMTQTTSSLSASLGDRVTISC RASQDIRNYLNWYQQKPDGTV KLLIYYTSRLHSGVPSRFSGSGS GTDYSLTIANLQPEDIATYFCQQ GNTLPWTFGGGTKLVTKRGGG GSGGGGSGGGGSASEVQLQQS GPELVKPGASMKISCKASGYSF TGYIVNWLKQSHGKNLEWIGLI NPYKGLTTYNQKFKGKATLTVD KSSSTAYMELLSLTSEDSAVYYC ARSGYYGDSDWYFDVWGAGT TVTVSS	106 (107)
G19-4 HL	accggtgaggtccagctgcaacagtctggacct gaactggtgaagcctggagcttcaatgaagattt	EVQLQQSGPELVKPGASMKISC KASGYSFTGYIVNWLKQSHGK	108 (109)

结合结构域	核苷酸序列	氨基酸序列	SEQ ID NOS. (氨基酸序列)
	cctgcaaggcctctggttactcattcactggcta catcgtgaactggctgaagcagagccatggaa agaaccttgagtggattggacttattaatccatac aaaggtcttactacctacaaccagaaattcaagg gcaaggccacattaactgtagacaagtcatcca gcacagcctacatggagctcctcagtctgacat ctgaagactctgcagtctattactgtgcaagatct gggtactatggtgactcggactggtacttcgatg tctggggcgcagggaccacggtcaccgtctcct ctggtggcggtggctcgggcggtggtggatct ggaggaggtgggagcgctagcgacatccagat gacacagactacatcctccctgtctgcctctctg ggagacagagtcaccatcagttgcagggcaag tcaggacattcgcaattatttaaactggtatcagc agaaaccagatggaactgttaaactcctgatcta ctacacatcaagattacactcaggagtcccatca aggttcagtggcagtgggtctggaacagattatt ctctcaccattgccaacctgcaaccagaagatat tgccacttacttttgccaacagggtaatacgcttc cgtggacgttcggtggaggcaccaaactggta accaaacgctcgag	NLEWIGLINPYKGLTTYNQKFK GKATLTVDKSSSTAYMELLSLTS EDSAVYYCARSGYYGDSDWYF DVWGAGTTVTVSSGGGGSGGG GSGGGGSASDIQMTQTTSSLSA SLGDRVTISCRASQDIRNYLNW YQQKPDGTVKLLIYYTSRLHSG VPSRFSGSGSGTDYSLTIANLQP EDIATYFCQQGNTLPWTFGGGT KLVTKRS	

铰链区	核苷酸序列	氨基酸序列	SEQ ID NO. (氨基酸序列)
sss(s)-hIgG1	gagcccaaatcttctgacaaaactcacacatct ccaccgagctca	EPKSSDKTHTSPPSS	230 (231)
csc(s)-hIgG1	gagcccaaatcttgtgacaaaactcacacatct ccaccgtgctca	EPKSCDKTHTSPPCS	232 (233)
ssc(s)-hIgG1	gagcccaaatcttctgacaaaactcacacatct ccaccgtgctca	EPKSSDKTHTSPPCS	110 (111)
scc(s)-hIgG1	gagcccaaatcttctgacaaaactcacacatgt ccaccgtgctca	EPKSSDKTHTCPPCS	112 (113)
css(s)-hIgG1	gagcccaaatcttgtgacaaaactcacacatct ccaccgagctca	EPKSCDKTHTSPPSS	114 (115)
scs(s)-hIgG1	gagcccaaatcttgtgacaaaactcacacatgt ccaccgagctca	EPKSSDKTHTCPPSS	116 (117)
ccc(s)-hIgG1	gagcccaaatcttgtgacaaaactcacacatgt ccaccgtgctca	EPKSCDKTHTSPPCS	118 (119)
ccc(p)-	Gagcccaaatcttgtgacaaaactcacacatgt	EPKSCDKTHTSPPCP	120 (121)

hIgG1	ccaccgtgccca		
sss(p)-hIgG1	gagcccaaatcttctgacaaaactcacacatct ccaccgagccca	EPKSSDKTHTSPPSP	122 (123)
csc(p)-hIgG1	gagcccaaatcttgtgacaaaactcacacatct ccaccgtgccca	EPKSCDKTHTSPPCP	124 (125)
ssc(p)-hIgG1	gagcccaaatcttctgacaaaactcacacatct ccaccgtgccca	EPKSSDKTHTSPPCP	126 (127)
scc(p)-hIgG1	gagcccaaatcttctgacaaaactcacacatgt ccaccgtgccca	EPKSSDKTHTCPPCP	128 (129)
css(p)-hIgG1	gagcccaaatcttgtgacaaaactcacacatct ccaccgagccca	EPKSCDKTHTSPPSP	130 (131)
scs(p)-hIgG1	gagcccaaatcttgtgacaaaactcacacatgt ccaccgagccca	EPKSSDKTHTCPPSP	132 (133)
scppcp	agttgtccaccgtgccca	SCPPCP	134 (135)
EFD	**核苷酸序列**	**氨基酸序列**	**序列识别符(氨基酸序列)**
hIgG1 (P238S) $C_{H2}C_{H3}$	gcacctgaactcctgggtggatcgtcagtcttc ctcttcccccaaaacccaaggacaccctcatg atctcccggaccctgaggtcacatgcgtggt ggtggacgtgagccacgaagaccctgaggtc aagttcaactggtacgtggacggcgtggaggt gcataatgccaagacaaagccgcgggaggag cagtacaacagcacgtaccgtgtggtcagcgt cctcaccgtcctgcaccaggactggctgaatg gcaaggagtacaagtgcaaggtctccaacaaa gccctcccagccccatcgagaaaacaatctc caaagccaaagggcagccccgagaaccaca ggtgtacaccctgccccatcccgggatgagc tgaccaagaaccaggtcagcctgacctgcctg gtcaaaggcttctatcccagcgacatcgccgtg gagtgggagagcaatgggcagccggagaac aactacaagaccacgcctcccgtgctggactc cgacggctccttcttcctctacagcaagctcac cgtggacaagagcaggtggcagcaggggaa cgtcttctcatgctccgtgatgcatgaggctctg cacaaccactacacgcagaagagcctctccc tgtctccgggtaaatga	APELLGGSSVFLFPPKPKD TLMISRTPEVTCVVVDVSH EDPEVKFNWYVDGVEVH NAKTKPREEQYNSTYRVV SVLTVLHQDWLNGKEYKC KVSNKALPAPIEKTISKAK GQPREPQVYTLPPSRDELT KNQVSLTCLVKGFYPSDIA VEWESNGQPENNYKTTPP VLDSDGSFFLYSKLTVDKS RWQQGNVFSCSVMHEAL HNHYTQKSLSLSPGK	142 (143)
hIgG1 (P331S) $C_{H2}C_{H3}$	gcacctgaactcctgggtggaccgtcagtcttc ctcttcccccaaaacccaaggacaccctcatg atctcccggaccctgaggtcacatgcgtggt ggtggacgtgagccacgaagaccctgaggtc aagttcaactggtacgtggacggcgtggaggt gcataatgccaagacaaagccgcgggaggag cagtacaacagcacgtaccgtgtggtcagcgt	APELLGGPSVFLFPPKPKD TLMISRTPEVTCVVVDVSH EDPEVKFNWYVDGVEVH NAKTKPREEQYNSTYRVV SVLTVLHQDWLNGKEYKC KVSNKALPASIEKTISKAK GQPREPQVYTLPPSRDELT	144 (145)

	cctcaccgtcctgcaccaggactggctgaatg gcaaggagtacaagtgcaaggtctccaacaaa gccctcccagcctccatcgagaaaacaatctc caaagccaaagggcagccccgagaaccaca ggtgtacaccctgcccccatcccgggatgagc tgaccaagaaccaggtcagcctgacctgcctg gtcaaaggcttctatcccagcgacatcgccgtg gagtgggagagcaatgggcagccggagaac aactacaagaccacgcctcccgtgctggactc cgacggctccttcttcctctacagcaagctcac cgtggacaagagcaggtggcagcaggggaa cgtcttctcatgctccgtgatgcatgaggctctg cacaaccactacacgcagaagagcctctccc tgtctccgggtaaatga	KNQVSLTCLVKGFYPSDIA VEWESNGQPENNYKTTPP VLDSDGSFFLYSKLTVDKS RWQQGNVFSCSVMHEAL HNHYTQKSLSLSPGK	
hIgG1 (P238S/ P331S) $C_{H2}C_{H3}$	gcacctgaactcctgggtggatcgtcagtcttc ctcttcccccaaaacccaaggacaccctcatg atctcccggacccctgaggtcacatgcgtggt ggtggacgtgagccacgaagaccctgaggtc aagttcaactggtacgtggacggcgtggaggt gcataatgccaagacaaagccgcgggaggag cagtacaacagcacgtaccgtgtggtcagcgt cctcaccgtcctgcaccaggactggctgaatg gcaaggagtacaagtgcaaggtctccaacaaa gccctcccagcctccatcgagaaaacaatctc caaagccaaagggcagccccgagaaccaca ggtgtacaccctgcccccatcccgggatgagc tgaccaagaaccaggtcagcctgacctgcctg gtcaaaggcttctatcccagcgacatcgccgtg gagtgggagagcaatgggcagccggagaac aactacaagaccacgcctcccgtgctggactc cgacggctccttcttcctctacagcaagctcac cgtggacaagagcaggtggcagcaggggaa cgtcttctcatgctccgtgatgcatgaggctctg cacaaccactacacgcagaagagcctctccct gtctccgggtaaatga	APELLGGSSVFLFPPKPKD TLMISRTPEVTCVVVDVSH EDPEVKFNWYVDGVEVH NAKTKPREEQYNSTYRVV SVLTVLHQDWLNGKEYKC KVSNKALPASIEKTISKAK GQPREPQVYTLPPSRDELT KNQVSLTCLVKGFYPSDIA VEWESNGQPENNYKTTPP VLDSDGSFFLYSKLTVDKS RWQQGNVFSCSVMHEAL HNHYTQKSLSLSPGK	146 (147)
连接子	**核苷酸序列**	**氨基酸序列**	**序列识别符**
STD1	aattatggtggcggtggctcgggcggtggtgg atctggaggaggtgggagtgggaattct	NYGGGGSGGGGSGGGGS GNS	148 (149)
STD2	aattatggtggcggtggctcgggcggtggtgg atctggaggaggtgggagtgggaattatggtg gcggtggctcgggcggtggtggatctggagg aggtgggagtgggaattct	NYGGGGSGGGGSGGGGS GNYGGGGSGGGGSGGGG SGNS	150 (151)
H1	aattct	NS	152 (153)
H2	ggtggcggtggctcggggaattct	GGGGSGNS	154 (155)

H3	aattatggtggcggtggctctgggaattct	NYGGGGSGNS	156 (157)
H4	ggtggcggtggctcgggcggtggtggatctg ggaattct	GGGGSGGGGSGNS	158 (159)
H5	aattatggtggcggtggctcgggcggtggtgg atctgggaattct	NYGGGGSGGGGSGNS	160 (161)
H6	ggtggcggtggctcgggcggtggtggatctg gggggaggaggcagcgggaattct	GGGGSGGGGSGGGGSGNS	162 (163)
H7	gggtgtccaccttgtccgaattct	GCPPCPNS	164 (165)
(G4S)3	ggtggcggtggatccggcggaggtgggtcgg gtggcggcggatct	GGGSGGGSGGGS	166 (167)
(G4S)4	ggtggcggtggctcgggcggtggtggatctg gaggaggtgggagcgggggaggtggcagt	GGGSGGGSGGGSGGGGS	168 (169)

表 5: 具有多价结合分子的例示性特征的一级结构(聚核苷酸和同源氨基酸序列)

实施例 8

对替代多特异性融合蛋白的结合及功能研究

对上述其他多价结合分子的每一者进行与上述原型 CD20-IgG-CD28 多特异性结合(融合)分子的实验的平行实验。一般而言，所述其他分子所获得的数据与对于原型分子所观测的结果类似。所述实验的某些显著结果揭示于下文中。图 14 显示对新颖分子中的一者所进行的阻断研究的结果，其中 BD1 与 BD2 均结合相同细胞或细胞类型上的靶标抗原(在此情况下为 CD20 与 CD37)。该多特异性多价结合(融合)蛋白经设计为具有结合 CD20 的结合结构域 1 (2H7; VLVH 取向)及结合 CD37 的结合结构域 2 (G28-1 VL-VH(LH)或 VH-VL(HL))。进行所述实验旨在证明蛋白的多特异性特性。

阻断研究：将 Ramos 或 BJAB B 淋巴母细胞样细胞($2.5x10^5$)与鼠抗-CD20 (25 μg/ml)抗体或鼠抗-CD37 (10 μg/ml)抗体一起于染色培养基(具有 2%小鼠血清的 PBS)中于 96 孔 V 形底板中预孵育，一起或单独将染色培养基在黑暗中置于冰上历时 45 分钟。将阻断抗体与细胞一起在室温下预孵育 10 分钟，之后以指定浓度范围(一般为 0.02 μg/ml 至 10 μg/ml)添加多特异性结合(融合)蛋白，且在黑暗中于冰上再孵育 45 分钟。将细胞于染色培养基中洗涤 2 次，且于染色培养基中在冰上与 Caltag (Burlingame, CA) FITC 山羊抗人类 IgG (1:100)一起孵育一小时，以检测多特异性结合(融合)蛋白与细胞的结合。接着将细胞用 PBS

洗涤 2 次且用 1%三聚甲醛(目录号第 19943 号，USB, Cleveland, Ohio)固定。使用 FACsCalibur仪器及 CellQuest软件(BD Biosciences, San Jose, CA)通过流式细胞术分析细胞。各数据组描绘 2H7-sss-hIgG-STD1-G28-1 HL 融合蛋白在 CD20、CD37 或 CD20 与 CD37 二者阻断抗体存在下的结合。虽然该实验使用裂解连接子中的一者，但两阻断抗体的存在完全排除与多特异性结合(融合)蛋白的结合，此证明大多数分子对 CD20 与 CD37 具有结合功能。在图 A 和图 B 中对两种细胞系 Ramos 与 BJAB 测试的数据类似，其中 CD20 阻断抗体比 CD37 阻断抗体更有效地降低所观测的与多特异性结合(融合)蛋白结合的程度。

ADCC 测定

图 15 显示对 CD20-CD37 多特异性结合(融合)蛋白进行的 ADCC 测定的结果。使用 BJAB 淋巴母细胞样 B 细胞作为靶标及人类 PBMC 作为效应细胞来进行 ADCC 测定。在 37°C 下，在 IMDM/10% FBS 中，将 BJAB 细胞用 500 μCi/ml ^{51}Cr 铬酸钠(250 μCi/μg)标记 2 小时。将经标记的细胞于 RPMI.10% FBS 中洗涤三次且以 4×10^5 个细胞/毫升再悬浮于 RPMI 中。肝素化人类全血是获自匿名实验室供者，且经由淋巴细胞分离介质(LSM, ICN Biomedical)梯度通过分级分离来分离 PBMC。收获血沉棕黄层(Buffy coats)且于 RPMI/10% FBS 中洗涤两次，之后以 5×10^6 个细胞/毫升的最终浓度再悬浮于 RPMI/10% FBS 中。在用于后续测定之前，通过锥虫蓝排除法使用血球计对细胞进行计数。将试剂样本以最终浓度 4 次添加至具有 10% FBS 的 RPMI 培养基中，且对各试剂制备三份 10 倍连续稀释液。接着将所述试剂以 50 微升/孔添加至 96 孔 U 型底板中直至所指示的最终浓度。将经 ^{51}Cr 标记的 BJAB 细胞以 50 微升/孔(2×10^4 个细胞/孔)添加至板中。接着将 PBMC 以 100 微升/孔(5×10^5 个细胞/孔)添加至板中，使得效应物(PBMC): 靶标(BJAB)的最终比率为 25:1。将效应物及靶标添加至单纯培养基中以测量本底杀死率。将经 ^{51}Cr 标记的细胞添加至单纯培养基中以测量 ^{51}Cr 的自发释放率且添加至具有 5% NP40(目录号第 28324 号，Pierce, Rockford, IL)的培养基中以量测 ^{51}Cr 的最大释放率。在 96 孔板的孔中重复三次进行反应。将多特异性结合(融合)蛋白以 0.01 μg/ml 至 10

μg/ml 范围内的最终浓度(如图所示)添加至孔中。各数据组描绘所述滴定范围内的不同多特异性结合(融合)蛋白或对应单一特异性 SMIP。在收获及计数之前，使反应在 37°C 下、在 5% CO_2 中进行 6 小时。接着将来自各孔的 25 μl 上清液转移至 Luma Plate 96 (目录号第 6006633 号，Perkin Elmer, Boston, Mass)中且在室温下干燥隔夜。在 Packard TopCounNXT 上测量所释放的 CPM。通过减去(样本的 cpm{三份样本的平均值}-cpm 自发释放率)/(cpm 最大释放率-cpm 自发释放率)x100 来计算特异性杀死百分率。以特异性杀死%相对蛋白浓度对数据作图。所述数据证明多特异性结合(融合)蛋白能够介导抵抗表达一个或多个靶标抗原的细胞的 ADCC 活性以及 SMIP 对 CD20 和/或 CD37 的单一特异性，但未显示此效应功能程度的增加。

共培养实验

图 16 显示着眼于此类型多特异性结合(融合)蛋白的其他特性所设计的实验的结果，其中通过经两个所结合的表面受体信号传导/结合，使两个抵抗靶标的结合结构域在相同细胞或细胞类型上得以表达可产生协同效应。使用如上文关于 ADCC 测定所述分离的 PBMC 进行共培养实验。将所述 PBMC 以 $2x10^6$ 个细胞/毫升，以 500 微升/孔的最终体积再悬浮于培养基中，且单独培养，或使用 H7 连接子[2H7-sss-IgG-H7-G28-1 HL]与 CD20、CD37、CD20+CD37 或多特异性结合(融合)蛋白的单一特异性 SMIP 一起孵育。以 20 μg/ml 的最终浓度添加各测试试剂。培养 24 小时后，未观察到培养物中 B 细胞的%有实际差异；然而，当所述细胞经受流式细胞术时，在含有多特异性结合(融合)蛋白的培养物的 FWD X 90 染色图案中可见细胞团块，此说明表达该两个靶标抗原的 B 细胞参与同型粘附。培养 72 小时后，多特异性结合(融合)蛋白使所存在的几乎所有 B 细胞死亡。两种单一特异性 SMIP 的组合亦剧烈降低 B 细胞的百分率，但对于多特异性结合分子未见该量下降。所述数据表明，将 CD20 的结合结构域与 CD37 的结合结构域均加在同一多特异性分子上，可使 B 细胞之间产生同型粘附，且亦可使 CD20 与 CD37 抗原结合相同细胞。尽管不希望受理论束缚，但清除靶标细胞的协同效应可归因于(1)经由结合结构域 1 及 2 结合相同细胞类型；

和/或(2)在 PBMC 培养物中，多价结合分子的效应功能结构域(恒定亚区)与单核细胞或其他细胞类型相互作用，使得杀死延迟。此杀死效应的动力学速度不快，完成需要 24 小时以上，说明其可为次级效应，需要产生细胞因子或其他分子后再观测所述效应。

细胞凋亡测定

图 17 显示为探究在用[2H7-sss-hIgG-H7-G28-1 HL]多特异性多价结合(融合)蛋白或单一特异性 CD20 和/或 CD37 SMIPS 单独及相互组合处理 B 细胞系后诱导细胞凋亡所设计的实验的结果。在 37°C 下，在 5% CO_2 中，在具有 10% FBS 及 5 μg/ml、10 μg/ml 或 20 μg/ml 融合蛋白的 Iscoves(Gibco)完全培养基中，将 Ramos 细胞(图 A; ATCC 第 CRL-1596 号)及 Daudi 细胞(图 B; ATCC 第 CCL-213 号)以 3×10^5 个细胞/毫升孵育隔夜(24 小时)。若以单一特异性 SMIP 进行组合实验，则所述蛋白是以如下浓度使用：TRU-015(针对 CD20 的 SMIP)=10 μg/ml，G28-1 LH(针对 CD37 的 SMIP)=5 μg/ml。或者，将 20 μg/ml 的 TRU-015 与 10 μg/ml 的 G28-1 LH 组合。接着使用 BD Pharmingen 细胞凋亡检测试剂盒(目录号第 556547 号)将细胞用膜联蛋白 V-FITC (Annexin V-FITC)及碘化丙啶染色且根据试剂盒说明进行处理。将细胞轻缓地旋转，在黑暗中于室温下孵育 15 分钟，且于 400 μl 结合缓冲剂中稀释，之后进行分析。在使用 Cell Quest 软件(Becton Dickinson)的 FACsCalibur (Becton Dickinson)仪器上通过流式细胞术分析样本。数据以描绘膜联蛋白 V/碘化丙啶阳性细胞的百分率相对于处理类型的关系的柱状图来显示。显然，与单一特异性试剂(即使一起使用时)相比，多特异性结合(融合)蛋白对两种细胞系均能够诱导显著更高程度的细胞凋亡。此增强的功能性活性反映 BD1 和 BD2(对 CD20 和 CD37 具特异性)受体在靶标细胞上的协调结合相互作用。

实施例 9

2H7-hIgG-G19-4 多特异性结合(融合)蛋白的结合及功能特性

此实施例描述 2H7-hIgG-G19-4 多特异性融合蛋白的结合及功能特性。所述分子的构建描述于实施例 7 中。表达及纯化是如以上实施

例中所述。

如对于以上分子如所述进行结合实验，例外之处在于用于测量CD3 结合的靶标细胞为在其表面上表达 CD3 的 Jurkat 细胞。参考图18，其中顶图显示使用由 20 μg/ml 连续稀释至 0.05 μg/ml 的纯化蛋白，关于 CD20-CD3 多特异性分子结合 Jurkat 细胞所获得的结合曲线。与 LH 取向相比，具有 HL 取向的 G19-4 与 CD3 抗原的特异性结合似乎更优。下图显示关于 BD1(识别 CD20 的结合结构域)所获得的结合曲线。所有分子均结合良好且以近似相当于与 CD20 具单一特异性的 SMIP 的程度结合。

ADCC 测定

对于图 19 中所显示的数据，如上述实施例中所述进行 ADCC 测定。在此情况下，融合蛋白均为 2H7-hIgG-G19-4 变异体或单一特异性 SMIP(对 CD20 具特异性的 2H7)或抗体(对 CD3 具特异性的 G19-4)的组合。此外，对于图 19 中的下图所显示的数据，在使用之前通过使用 MACS (Miltenyi Biotec, Auburn, CA)管柱分离装置及 NK 细胞特异性 CD16 磁性微珠粒(目录号第 130-045-701 号)的磁性珠粒耗尽法自 PBMC 中耗尽 NK 细胞。两图中所显示的数据证明，所有 CD20-hIgG-CD3 多特异性分子均介导 ADCC，而与 NK 细胞是否耗尽或总 PBMC 是否用于测定无关。对于 TRU 015 或 G19-4 与 TRU015 的组合而言，仅含有 NK 细胞的培养物可介导 ADCC。尽管 G19-4 可结合表达 CD3 的 NK T 细胞且在所示第一测定中活化所述细胞，但在针对 BJAB 靶标的任一测定中，G19-4 均运作不佳，其不表达 CD3。在下图中所观测的多特异性结合(融合)蛋白的杀死作用可能是经由与表达 CD20 抗原的 BJAB 靶标相反，通过结合 CD3 来活化 T 细胞群的细胞毒性来介导。在所用浓度范围内，此杀死活性对于分子的剂量似乎相对不敏感，且甚至在 0.01 μg/ml 的浓度下，仍明显不同于所测试的其他分子。

实施例 10

多价结合分子

其他实施方案包括源自免疫球蛋白的连接子结构域。更特定而言，所述连接子的源序列为如下所获得的序列：比较存在于免疫球蛋白超家族的其他成员的 V-样结构域之间或存在于 V-样结构域与 C-样结构域之间的区。因为所述序列通常作为细胞表面受体的细胞外结构域的部分来表达，因此预期其对于蛋白水解裂解更稳定且亦不应具有免疫原性。预期不适用作多价结合(融合)蛋白的连接子，但存在于 C-样结构域与跨膜结构域之间的插入区中的一种序列类型为在-Ig 超家族的表面表达成员上所表达的序列类型。经观测许多所述分子具有可溶性形式且在靠近细胞膜的所述插入区裂解，说明所述序列比其余分子更易受裂解影响。

如本文中所述，将上述连接子插入单一特异性 SMIP 中，插入结合结构域与效应功能结构域之间，或插入多价结合(融合)蛋白的两种可能性连接子位置的一者中。

随附本申请中所揭示的完整序列表，且该序列表以引用方式全文并入本文中。彩色编码指示特定聚核苷酸及多肽中的各种区或结构域的序列适用于鉴别本文中所揭示的任何分子的序列中的对应区或结构域。

实施例 11

用于靶向 B-细胞的蝎形分子候选者的筛检基质

引言

作为鉴别很可能产生抵抗靶标群体的适用且有效多价结合分子或蝎形分子的成对单克隆抗体结合结构域的组合的方法，在代表各种非霍奇金氏淋巴瘤的抗 B 细胞系组合基质中对一系列抵抗 B 细胞抗原的单克隆抗体进行测试。为确保已知或预期结合所关注细胞的抗体的所有可能逐对比较均得以测定，可使用既定细胞类型，使用抗体的二维基质来导引研究设计。已知的根据其簇名称(CD)的抵抗多种 B 细胞抗原的单克隆抗体记录于左列。将某些所述抗体(依其特异性结合的一个或多个抗原来命名)(亦即 CD19、CD20、CD21、CD22、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86

及 CL II (II 类 MHC))单独或与该单克隆抗体组的其他成员组合而与抗原阳性靶标细胞一起孵育。涵盖所述抗体的可变结构域作为多价结合分子的例示性实施方案中的结合结构域。本领域技术人员利用对本领域的了解及常规程序，能够在例如可公开获得的数据库中鉴别合适的抗体序列(核酸编码序列以及氨基酸序列)，以形成合适抗体或其片段(例如通过基于杂交的克隆、PCR、肽合成法及其类似者)及使用所述化合物构建多价结合分子。如本文中所述的获得结合结构域的例示性抗体的来源提供于表 6 中。通常将使用可获得抗体链的 CDR 区的克隆或合成策略，但仍涵盖保留特异性结合靶标抗原的能力的任何抗体、其片段或其衍生物。

更详细的，来自杂交瘤克隆抗体的重链和/或轻链可变区为本领域中的标准。不需要获知所关注可变区的序列以便利用常规克隆技术来获取此区。参见例如 Gilliland 等人，Tissue Antigens 47(1):1-20 (1996)。为制备包含可识别鼠或人类白细胞抗原的可变区的单链多肽，设计一种用于快速克隆及表达的方法，该方法可在从杂交瘤细胞分离 RNA 的两至三周内产生功能蛋白。可变区是通过 poly-G 加尾第一链 cDNA，继而由前置 poly-C 锚定引物及对恒定区序列具有特异性的反向引物锚定 PCT 来克隆。两种引物均含有用于插入 pUC19 内的侧翼限制性核酸内切酶位点。形成用于分离鼠、仓鼠和大鼠 V_L 和 V_H 基因的 PCR 引物组。在测定特异性 V_L 与 V_H 对的共有序列后，通过编码介入肽连接子(通常编码$(Gly_4Ser)_3$)的 DNA 使 V_L 与 V_H 基因连接，且将 V_L-连接子-V_H 基因盒转移至 pCDM8 哺乳动物表达载体中。将构建体转染至 COS 细胞中且从条件培养基上清液中回收 sFv。该方法已成功用于从产生鼠、大鼠或仓鼠抗体的杂交瘤中形成抗人类 CD2、CD3、CD4、CD8、CD28、CD40、CD45 的功能性 sFv 和抗鼠 CD3 和 gp39 的功能性 sFv。起初，sFv 被表达为具有人类 IgG1 的铰链-C_{H2}-C_{H3} 结构域的融合蛋白，以利于使用山羊抗人类 IgG 试剂或蛋白 A 进行快速表征及纯化。活性 sFv 亦可用小肽(例如标记)来表达或以无尾形式来表达。表达 CD3 (G19-4) sFv 无尾形式证明细胞信号传导活性增强且揭示 sFv 具有活化受体的潜能。

或者，可如下直接实现对单克隆抗体的可变结构域的一级氨基酸序列的鉴别：例如限制抗体的蛋白水解作用，继而例如使用 Edman 降解方法或通过断裂质谱法进行 N-末端肽测序。N-末端测序方法已为本领域中所熟知。在测定可变结构域的一级氨基酸序列后，通过合成性核酸合成方法(例如 PCR)，继而通过 scFv 形成法来组装编码该序列的 cDNA。所述必需或优选核酸操控方法为本领域中的标准。

亦涵盖如上所述的抗体的片段、衍生物及类似物作为合适的结合结构域。此外，涵盖上述任何恒定亚区，包括包含上述任何铰链区的恒定亚区。此外，该实施例中所述的多价单链结合分子可包括本文中所述的任何或全部连接子。

初始使单克隆抗体暴露于细胞且接着使用山羊抗小鼠第二步抗体交联(第 2 步)。任选地，可在使细胞与抗体接触之前使抗体交联(例如通过在溶液中使抗体交联)。作为另一替代性方法，可如下使单克隆抗体以固相交联：通过吸附至组织培养孔的塑料底层上或藉助于吸附至塑料的山羊抗小鼠抗体而“截留”于该塑料上，继而通过基于孔板的测定来评估例如生长中断或细胞存活。

磷脂酰基丝氨酸经由细胞膜的胞质侧逆转至此质膜的外部细胞表面为促细胞凋亡事件的可接受指标。进展至细胞凋亡导致细胞膜完整性丧失，此可通过细胞不透性插入染料(例如碘化丙啶(PI))的进入来检测。在细胞暴露于单独或组合的单克隆抗体后，进行双重促细胞凋亡测定，且对经处理的细胞群就细胞表面阳性膜联蛋白 V(ANN)和/或 PI 包涵物方面进行评分。

膜联蛋白 V 结合/碘化丙啶内化分析

细胞和细胞培养条件。实验旨在检测使抵抗所表达的靶标的两种不同单克隆抗体交联对四种人类 B-细胞系的影响。通过测定暴露后 ANN 和/或 PI 染色的程度来测量对细胞系的影响。在 37°C 下，在 5% CO_2 中，在具有 10% FBS 的 Iscoves(Gibco)完全培养基中将人类 B 细胞系 BJAB、Ramos (ATCC#CRL-1596)、Daudi (ATCC#CCL-213)及 DHL-4 (DSMZ#ACC495)孵育 24 小时。研究之前，使细胞密度维持在 2×10^5 至 8×10^5 个细胞/毫升且存活率通常>95%。

以 2×10^5 个细胞/毫升的细胞密度及 2 μg/ml 的各比较单克隆抗体，经由抵抗 B-细胞抗原的基质来进行实验。各比较单克隆抗体是以 2 μg/ml 单独添加，或亦以 2 μg/ml 个别地与各基质单克隆抗体组合添加。表 6 列出所述实验中所用的单克隆抗体的目录号及来源。对于使所述单克隆抗体在溶液中交联而言，则将山羊抗小鼠 IgG (Jackson Labs 目录号第 115-001-008 号)以 2:1 的浓度比(山羊抗小鼠抗体:各单克隆抗体)添加至各孔中，例如，仅具有一种单克隆抗体(2 μg/ml)的孔中可添加有最终浓度为 4 μg/ml 的山羊抗小鼠抗体，而具有比较单克隆抗体(2 μg/ml)与来自基质的单克隆抗体(2 μg/ml)的孔中可添加有 8 μg/ml 的山羊抗小鼠抗体。

细胞在 37°C 下于 5% CO_2 中孵育 24 小时后，使用 BD Pharmingen 膜联蛋白 V-FITC 细胞凋亡检测试剂盒 I (#556547)，将其用膜联蛋白 V-FITC 及碘化丙啶染色。简而言之，将细胞用冷 PBS 洗涤两次且以 1×10^6 个细胞/毫升再悬浮于"结合缓冲剂"中。接着将 100 微升于结合缓冲剂中的细胞用 5 μl 膜联蛋白 V-FITC 及 5 μl 碘化丙啶染色。将细胞轻缓地混合且在室温下于黑暗中孵育 15 分钟。接着将 400 微升结合缓冲剂添加至各样本中。接着在 FACsCalibur (Becton Dickinson)上对样本读数且使用 Cell Quest 软件(Becton Dickinson)进行分析。

表 6

名称	目录号	供应商
抗-CD19	#C2269-74	US Biological (Swampscott, MA)
抗-CD20	#169-820	Ancell Corp (Bayport, MN)
抗-CD21	#170-820	Ancell Corp (Bayport, MN)
抗-CD22	#171-820	Ancell Corp (Bayport, MN)
抗-CD23	#172-820	Ancell Corp (Bayport, MN)
抗-CD30	#179-820	Ancell Corp (Bayport, MN)
抗-CD37	#186-820	Ancell Corp (Bayport, MN)
抗-CD40	#300-820	Ancell Corp (Bayport, MN)
抗-CD70	#222-820	Ancell Corp (Bayport, MN)

名称	目录号	供应商
抗-CD72	#C2428-41B1	US Biological (Swampscott, MA)
抗-CD79a	#235-820	Ancell Corp (Bayport, MN)
抗-CD79b	#301-820	Ancell Corp (Bayport, MN)
抗-CD80	#110-820	Ancell Corp (Bayport, MN)
抗-CD81	#302-820	Ancell Corp (Bayport, MN)
抗-CD86	#307-820	Ancell Corp (Bayport, MN)
抗-CL II DR, DQ, DP	#131-820	Ancell Corp (Bayport, MN)

表 6: 该研究中所使用的抵抗 B 细胞抗原的抗体及其来源。

将交联抗体(例如山羊抗小鼠抗体)添加至单独的单克隆抗体 A 中使得细胞敏感性增强，说明用识别相同抗原的两个结合结构域所构建的多价结合分子或蝎形分子可有效增强细胞敏感性。不希望受理论束缚，此敏感性增强可归因于抗原聚集及经修改的信号传导。TNF 受体家族成员例如需要同型多聚化以便信号转导，且在分子各端具有等效结合结构域的蝎形分子可促进此相互作用。CD40 的聚集及随后的信号传导为该现象在 B 细胞系统的实例。

如图 20、21 和 22 中所示，添加抵抗不同抗原的单克隆抗体 A 和单克隆抗体 B 可对经处理的细胞产生加成的促细胞凋亡效应，或在某些组合中产生大于加成(亦即协同)的促细胞凋亡效应。在图 20 中，例如，抗-CD20 与抵抗其他 B 细胞抗原的单克隆抗体的组合均可使细胞敏感性得到不同程度的增强。然而，某些组合(诸如抗-CD20 与抗-CD19 的组合，或抗-CD20 与抗-CD21 的组合)产生大于加成的促细胞凋亡效应，说明由所述结合结构域组成的多价结合分子或蝎形分子可特别有效地清除转化的 B 细胞。参考图 20，当暴露于单独的抗-CD 20 抗体时，显示促细胞凋亡活性的细胞的百分比为约 33%(对应于“20”的垂直条形柱，亦即抗-CD20 抗体)；暴露于抗-CD19 抗体时，促细胞凋亡细胞的百分比为约 12%(图 20 中对应于“19”的垂直条形柱，亦即抗-CD19 抗体)；且暴露于抗-CD20 与抗-CD19 抗体时，促细胞凋亡细胞的百分比为约 73%(图 20 中对应于“19”的水平条形柱)。在暴露于两种抗体后，

促细胞凋亡细胞为 73%，显著大于各单独抗体所产生的效应总和 45%(33%+12%)，说明抗-CD19 与抗-CD 20 抗体对可产生协同效应。适用的多价结合分子包括其中两个结合结构域可对 B-细胞行为产生加成效应的分子以及其中两个结合结构域可对 B-细胞行为产生协同效应的多价结合分子。在某些实施方案中，一个结合结构域将对所测量的细胞行为参数不具有可检测的影响，其中成对结合结构域中的每一者促成多价结合分子活性的不同方面，诸如多特异性多价结合分子(例如结合结构域 A 结合靶标细胞且促进细胞凋亡，而结合结构域 B 结合可溶性治疗剂，诸如细胞毒素)。视多价结合分子的设计而定，两个结合结构域对靶标细胞产生的组合效应的类型(加成、协同或抑制效应)可不相关，此是因为结合结构域中的一者可对非细胞性(例如可溶性)结合伙伴具特异性或对细胞相关性结合伙伴具特异性，而对不同细胞类型却并非如此。

产生加成、协同或抑制效应的例示性结合结构域配对(如图 20-23 中所示)由表 7 和表 8 显而易见。表 7 提供自图 20-23 各图中所提取的关于 ANN 和/或 PI 染色呈阳性的细胞的百分比的定量数据。表 8 提供利用表 7 的数据所计算的数值，基于所述数值可判定给定抗体对的相互作用是否产生加成、协同或抑制效应，又如依据 ANN 和/或 PI 染色呈阳性的细胞的百分比所评估。

表 7

名称	**抗-CD20**	**抗-CD79b**	**抗-CL II**	**抗-CD22**
抗-CD19	13/73*	18/76/66	14/47/46	12/11
抗-CD20	33/NA	42/94/92	33/71/76	28/33
抗-CD21	14/75	22/50/76	18/24/40	11/11
抗-CD22	8/55	12/39/33	12/19/17	10/12
抗-CD23	8/41	12/63/55	14/22/17	10/12
抗-CD30	8/38	14/72/61	12/56/61	10/11
抗-CD37	15/45	19/92/86	20/60/62	19/20
抗-CD40	10/48	12/44/30	13/21/28	14/13
抗-CD70	9/40	12/56/39	15/21/15	10/10
抗-CD72	NA	16/60/64	30/78/63	17/17

名称	抗-CD20	抗-CD79b	抗-CL II	抗-CD22
抗-CD79a	21/66	43/42/50	28/55/51	14/14
抗-CD79b	46/88	70/70/68	45/80/76	26/16
抗-CD80	7/41	14/35/30	15/19/17	11/11
抗-CD81	14/65	25/86/83	25/54/43	19/20
抗-CD86	7/38	16/58/42	15/24/18	14/11
抗-CL II	53/77	52/96/98	47/52/43	72/57

*表 7 的第 2-4 列中的数值分别反映图 20-22 中的直方图柱的高度，各细胞的第一数值表示垂直条形柱的高度，第二数值表示水平条形柱的高度，且第三数值(若存在)反映点画柱的高度。在第 5 列中，第一数值反映实心柱的高度且第二数值反映图 23 中的斜纹柱的高度。

表 8

名称	抗-CD20	抗-CD79b	抗-CL II	抗-CD22
抗-CD19	S: 13+33=46*	A: 18+56=74 A: 18+43=61	S: 14+26=40 S: 14+18=32	I: 12+10=22
抗-CD20	NA	A: 42+56=98 A: 42+43=85	S: 33+26=59 S: 33+18=51	A/I:28+10=38
抗-CD21	S: 14+33=47	I: 22+56=78 S: 22+43=65	I: 18+26=44 A: 18+18=36	I: 11+10=21
抗-CD22	S: 8+33=41	I: 12+56=68 I: 12+43=55	I: 12+26=38 I: 12+18=30	NA
抗-CD23	A: 8+33=41	A: 12+56=68 A: 12+43=55	I: 14+26=40 I: 14+18=32	I: 10+10=20
抗-CD30	A: 8+33=41	A: 14+56=70 A: 14+43=57	S: 12+26=38 S: 12+18=30	I: 10+10=20
抗-CD37	A: 15+33=48	S: 19+56=75 S: 19+43=62	S: 20+26=46 S: 20+18=38	I: 19+10=29
抗-CD40	A/S:10+33=43	I: 12+56=68 I: 12+43=55	I: 13+26=39 A: 13+18=31	I: 14+10=24
抗-CD70	A: 9+33=42	I: 12+56=68 I: 12+43=55	I: 15+26=41 I: 15+18=33	I: 10+10=20
抗-CD72	NA	I: 16+56=72 A: 16+43=59	S: 30+26=56 S: 30+18=48	I: 17+10=27

名称	抗-CD20	抗-CD79b	抗-CL II	抗-CD22
抗-CD79a	S: 21+33=54	I: 43+56=99 I: 43+43=86	A: 28+26=54 A: 28+18=46	I: 14+10=24
抗-CD79b	S: 46+33=79	NA	S: 45+26=71 S: 45+18=63	I: 26+10=36
抗-CD80	A: 7+33=40	I: 14+56=70 I: 14+43=57	I: 15+26=41 I: 15+18=33	I: 11+10=21
抗-CD81	S: 14+33=47	A: 25+56=81 S: 25+43=68	A: 25+26=51 A: 25+18=43	I: 19+10=29
抗-CD86	A: 7+33=40	I: 16+56=72 I: 16+43=59	I: 15+26=41 I: 15+18=33	I: 14+11=25
抗-CL II	I: 53+33=86	A: 52+56=108 A: 52+43=95	NA	I: 72+10=82

“A”意谓观测到“加成”效应；

“S”意谓观测到“协同”效应；

“I”意谓观测到“抑制”效应；

*示意性方程式：A+B=C，其中“A”为因基质抗体单独所致的 ANN 和/或 PI 阳性细胞百分比，“B”为因共同抗体(图 20 为抗-CD20、图 21 为抗-CD79b、图 22 为抗-CLII 和图 23 为抗-CD22)所致的 ANN 和/或 PI 阳性细胞百分比，且“C”为预计加成效应。(参见以上表 7 中与图 20-23 对应的定量数据)。若单元格中存在两个方程式，则上方程式反映使用较高指示浓度的共同抗体所得的结果；下方程式反映使用较低指示浓度的共同抗体所得的结果。

在某些实施方案中，两个结合结构域在敏化(或去敏化)靶标细胞(诸如 B 细胞)中以抑制、加成或协同方式相互作用。图 23 显示将抗-CD22 抗体与强促细胞凋亡单克隆抗体(诸如抗-CD79b 抗体或抗-MHC II 类(亦即抗-CL II)抗体)组合所产生的保护或抑制效应。例如，图 23 和表 7 显示，单独抗-CD22 抗体诱导不超过约 10%的细胞显示促细胞凋亡行为(图 23 中对应于“22”的实心柱)且抗-CD79b 诱导约 26%的促细胞凋亡细胞(图 23 中对应于“CD79b”的实心柱)。然而，抗-CD22 与抗-CD79b 的组合诱导仅约 16%的促细胞凋亡细胞(图 23 中对应于“79b”的斜纹柱)。因此，所组合的抗体诱导 16%的促细胞凋亡细胞，其小于抗-CD22(12%)和抗-CD79b(26%)单独所产生的效应的总和

38%。利用此方法检验图 23 和/或表 7-8 揭示，抗-CD22 抗体和经由延长包含抗-CD22 结合结构域的多特异性多价结合分子，在与以下各抗体(或对应结合结构域)独立组合使用时将获得抑制性总体效应：抗-CD19、抗-CD20、抗-CD21、抗-CD23、抗-CD30、抗-CD37、抗-CD40、抗-CD70、抗-CD72、抗-CD79a、抗-CD79b、抗-CD80、抗-CD81、抗-CD86 和抗-MHC II 类抗体/结合结构域。

不希望受理论束缚，所述数据可作如下解释：说明抗-CD22 抗体或包含抗-CD22 结合结构域的多特异性多价结合分子可防止或缓和上文刚列出的任何抗体效应。更一般而言，包含抗-CD22 结合结构域的多特异性多价结合分子将抑制与 CD19、CD20、CD21、CD23、CD30、CD37、CD40、CD70、CD72、CD79a、CD79b、CD80、CD81、CD86 和 MHC II 类分子中任一者相互作用所产生的效应。在图 23 和表 8 中可见，抗-CD22 抗体和经由延长结合结构域(包含抗-CD22 结合结构域)将起抑制或缓和识别 B-细胞表面标志物(诸如 CD 抗原)的任何抗体/结合结构域的活性的作用。预期多价结合分子(包括多特异性多价结合分子)适用于各种疾病的改进治疗方案中，其中需要减弱或控制结合结构域的活性。

除与靶标细胞相互作用的两个结合结构域通常经由细胞表面配体的结合产生抑制、加成或协同组合效应外，本文中所揭示的实验结果证明，给定的结合结构域对可依据两个结合结构域的相对浓度而提供不同类型的组合效应，从而增强本发明的通用性。例如，表 8 揭示，抗-CD21 与抗-CD79b 在抗-CD79b 的测试浓度较高时以抑制方式相互作用，但在抗-CD79b 的测试浓度较低时，该两种抗体以协同方式相互作用。尽管某些实施方案将使用单一类型的多价结合分子，亦即包含例如单一 CD21 结合结构域和单一 CD79b 结合结构域的单特异性多价结合分子，但本发明包括多价结合分子的混合物，所述混合物将使得可调整相对结合结构域浓度以达成所要效应，诸如抑制、加成或协同效应。此外，本发明的方法包含将单一多价结合分子与另一种结合分子(诸如常规抗体分子)组合使用以调整或优化结合结构域的相对浓度。本领域技术人员将能够利用标准技术(例如通过设计两组实验性连

续稀释基质，每一结合结构域使用一组)来测定结合结构域的适用相对浓度。

不希望受理论束缚，应了解，一个配体的结合可诱导或调节第二配体在相同细胞类型上的表面表达，或其可改变第二配体的表面环境从而改变其与特异性结合分子(诸如抗体或多价结合分子)结合的敏感性。

尽管本文中以使用 B 细胞系和抗原为例，最佳测定有效多价结合分子(亦即蝎形分子)的所述方法适用于其他疾病背景和靶标细胞群体，包括其他正常细胞、其异常细胞对应物，包括慢性刺激性造血细胞、癌细胞及感染细胞。

在筛检单克隆抗体组合的直接效应的方法中，亦可使用其他信号传导表型(诸如 Ca^{2+}移动；酪氨酸磷酸化调节；半胱天冬酶(caspase)活化；NF-κB 活化；细胞因子、生长因子或趋化因子加工；或基因表达(例如在报道基因系统中))。

作为使用二级抗体以使一级抗体交联和模拟多价结合分子或蝎形分子结构的替代例，可使用结合抗体的 Fc 部分的其他分子(包括可溶性 Fc 受体、蛋白 A、补体组份(包括 C1q)、甘露糖结合凝集素、含有反应剂或交联剂的珠粒或基质、双官能化学交联剂及塑料吸附剂)使抵抗相同或不同抗原的多种单克隆抗体交联。

实施例 12

具有效应功能或蝎形分子的结构的多价结合蛋白

蝎形多肽的一般示意性结构为 H2N-结合结构域 1-蝎形连接子-恒定亚区-结合结构域 2。蝎形分子亦可具有安置于结合结构域 1 的 N-末端的铰链样区，通常为源自抗体铰链的肽区。在某些蝎形分子实施方案中，结合结构域 1 和结合结构域 2 是各自源自免疫球蛋白结合结构域，例如源自 V_L和 V_H。V_L与 V_H通常经连接子连接。实验旨在证明，蝎形多肽可具有不同于免疫球蛋白结合结构域的结合结构域，包括蝎形分子结合结构域所来源的 Ig 结合结构域，不同之处在于氨基酸序列差异，导致相对于来源 Ig 结合结构域通常小于 5%且优选小于 1%

的序列趋异。

通常，序列差异导致单一氨基酸变化，诸如取代。所述氨基酸变化的优选位置位于所对应的蝎形分子结合结构域的一或多个区中，或显示与该蝎形分子结合结构域所来源的 Ig 结合结构域的 Ig 互补性决定区(CDR)至少 80%且优选 85%或 90%的序列一致性。通过比较肽结合相同靶标(诸如 CD20)的模型来提供进一步指导。相对于 CD20，表位图谱已揭示，结合 CD20 的 2H7 抗体可识别 CD20 的 Ala-Asn-Pro-Ser (ANPS)基序，且预期结合 CD20 的蝎形分子亦可识别该基序。将 ANPS 基序深埋入由蝎形分子结合结构域区(对应于 Ig CDR)所形成的袋囊中所致的氨基酸序列变化预期为 CD20 的功能性结合剂。模型研究亦已揭示，对应于 CDR3 (V_L)、CDR1-3 (V_H)的蝎形区接触 CD20，且预期维持或促成所述接触的变化可产生结合 CD20 的蝎形分子。

除促进蝎形分子与其靶标的相互作用外，涵盖蝎形分子结合结构域中促进对应于 Ig V_L 和 V_H 结构域的蝎形分子结合结构域区之间相互作用的序列变化(相对于同源 Ig 结合结构域序列)。例如，在对应于 V_L 的结合 CD20 的蝎形分子区中，序列 SYIV 可通过用氨基酸(诸如 His)取代 Val (V33)而改变，产生序列 SYIH。此变化预期可改良对应于 V_L 和 V_H 结构域的蝎形分子区之间的相互作用。此外，预期在对应于 V_H-CDR3 的蝎形分子区的 N-末端添加残基将改变此蝎形分子区的取向，这可能影响其结合特征，因为 V_H-CDR3 的 N-末端 Ser 与 CD20 接触。常规测定揭示那些取向可使结合特征产生所要变化。亦预期，对应于 V_H-CDR2 和/或 V_H-CDR3 的蝎形分子区中的突变将形成与靶标(诸如 CD20)的潜在新颖接触。例如，基于模型研究，预期取代对应于 V_H-CDR3 的区中的 Y105 和 W106(存在于序列 NSYW 中)中任一个将以易经受常规测定的方式改变蝎形分子的结合特征以用于鉴别具有经修饰结合特征的蝎形分子。另外例如，预期对应于 Ig VL-CDR3 的蝎形分子结合结构域的序列(诸如序列 CQQW 中的 Trp (W))的改变将影响结合。通常，对于对靶标显示增强亲和性的那些蝎形分子，筛检对应于 Ig CDR 的蝎形分子区中的改变。

基于人源化 CD20 scFv 结合结构域 20-4 的模型结构，基于与 CD20

细胞外环结构相关的公开信息(Du 等人，J Biol. Chem. 282 (20):15073-80 (2007))，且基于由小鼠 2H7 抗体(其为人源化 20-4 scFv 结合结构域的 CDR 的来源)所识别的 CD20 结合表位，在 2Lm20-4x2Lm20-4 蝎形分子的 CDR 区中进行工程化突变，以便改良其与 CD20 结合的亲和性。首先，设计可影响 20-4 CDR 构型和促进与 CD20 细胞外环更有效结合的突变。其次，设计可在 2Lm20-4x2Lm20-4 蝎形分子与其靶标之间提供新型分子间相互作用的引入式变化。所述突变包括：VL CDR1 V33H (亦即在 VL 区的 CDR1 的位置 33 处用 His 取代 Val)、VL CDR3 W90Y、VH CDR2 D57E、VH CDR3 在残基 S99 后插入 V、VH CDR3 Y101K、VH CDR3 N103G、VH CDR3 N104G 及 VH CDR3 Y105D。由于预期将某些所述突变组合可产生协同效应，因此设计如表 9 中所示的将不同突变组合的 11 种突变体(由突变引入的残基字体加粗且加下划线)。

表 9

V_L CDR1	V_L CDR3	V_H CDR2	V_H CDR3
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	**SV**YYSNYWYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	**SV**YY**GG**YWYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SYYSNS**D**WYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SYYS**GGD**WYFDL
RASSSVSYIV	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SY**K**SNSYWYFDL
RASSSVSYIV	QQWSFNPPT	AIYPGNG**E**TSYNQKFKG	SYYSNSYWYFDL
RASSSVSYIV	QQ**Y**SFNPPT	AIYPGNGDTSYNQKFKG	SYYSNSYWYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNGDTSYNQKFKG	SY**K**SNS**D**WYFDL
RASSSVSYI**H**	QQWSFNPPT	AIYPGNG**E**TSYNQKFKG	SYYSNS**D**WYFDL
RASSSVSYI**H**	QQ**Y**SFNPPT	AIYPGNGDTSYNQKFKG	SYYSNS**D**WYFDL
RASSSVSYI**H**	QQ**Y**SFNPPT	AIYPGNG**E**TSYNQKFKG	SY**K**S**GGD**WYFDL

使用编码改变的序列区的引物，通过 PCR 诱变将突变引入 CD20xCD20 蝎形分子(2Lm20-4x2Lm20-4)的结合结构域中。序列验证

后，将编码具有对应突变的 2Lm20-4 scFv 片段的 DNA 片段克隆至含有蝎形分子的恒定亚区的编码区的常规表达载体中，从而产生含有新型 2Lm20-4x2Lm20-4 蝎形分子的完整 DNA 序列的聚核苷酸。具有 CDR 突变的 2Lm20-4x2Lm20-4 蝎形分子的变异体是通过在瞬间 COS 细胞系统中表达而产生，且经由蛋白 A 及尺寸排阻(SEC)层析法纯化。使用初级 B-细胞和 WIL2-S B-淋巴瘤细胞系，通过 FACS 分析来评估 2Lm20-4x2Lm20-4 蝎形分子变异体的结合特性。

亦已使用类似方法形成其他突变体以优化 CD20 结合结构域。命名为 TRU015 的 CD20 SMIP 充当形成突变体的底物，且除非相反说明，否则所有结构域均为人类结构域。发现以下突变体含有适用的功能性 CD20 结合结构域。018008 分子在 VL 中的 CDR1 的位置 27 处含有取代 S 的 Q (单字母氨基酸编码)，在 VH 的 CDR1 中的位置 28 处含有取代 T 的 S 且在 VH 的 CDR3 中的位置 102 处含有取代 V 的 L。依照蝎形分子的模组化设计，将以下对应于 IgG1 铰链中的 CCCP 序列的部分蝎形连接子序列独立地与突变的 VL 和 VH 组合：CSCS、SCCS 和 SCCP。018009 分子在 VL 的 CDR1 的位置 27 处含有取代 S 的 Q，在 VH 的 CDR1 的位置 28 处含有取代 T 的 S，且在 VH 的 CDR3 中，在位置 96 处含有取代 V 的 S、在位置 102 处含有取代 V 的 L 且在位置 95 处缺失 V。018009 中使用上述与 018008 中所使用的蝎形连接子中所发现的相同的蝎形连接子亚序列。018010 分子在 VL 的 CDR1 中，在位置 27 处含有取代 S 的 Q、在位置 33 处具有取代 M 的 I 且在位置 34 处含有取代 H 的 V，以及在 VH 的 CDR1 的位置 28 处含有取代 T 的 S，且在 VH 的 CDR3 的位置 102 处含有取代 V 的 L。由 CSCS 和 SCCS 亚序列定义的蝎形连接子适用于 018010。018011 在 VL 的 CDR1 和 VH 的 CDR1 中含有与如 018010 所述相同的突变，以及在 VH 的 CDR3 中，在位置 95 处缺失 V、在位置 96 处含有取代 V 的 S 且在位置 102 处含有取代 V 的 L。由 CSCS、SCCS 和 SCCP 亚序列定义的蝎形连接子用于 018011 分子中。018014 VL 为未突变的小鼠 VL，其中人类 VH 在 CDR1 的 28 处含有 S 取代 T 的变化且在 CDR3 的 102 处含有 L 取代 V 的变化。018015 亦含有未突变的小鼠 VL 以及人类 VH，

该人类 VH 在 CDR1 的 28 处含有 S 取代 T 的变化，且在 CDR3 中，在 95 处缺失 V、在 96 处含有取代 V 的 S 且在 102 处含有取代 V 的 L。2Lm5 分子在 VL 的 CDR1 中的 27 处具有取代 S 的 Q;在 VH 的 CDR1 中，在 27 处具有取代 Y 的 F 且在 30 处具有取代 T 的 S；以及在 VH 的 CDR3 中，在 95 处缺失 V、在 96 处具有取代 V 的 S 且在 102 处具有取代 V 的 L。由 CSCS、SCCS 和 SCCP 定义的蝎形连接子独立地用于 018014 和 018015 的每一者中。2Lm5-1 与 2Lm5 相同，例外之处为 2Lm5-1 在 VH 的 CDR1 中无突变，且仅使用由 CSSS 亚序列定义的蝎形连接子。2Lm6-1 具有 2Lm5 的突变且在 VL 的 CDR3 中，在 92 处具有取代 S 的 T 且在 93 处具有取代 F 的 S，且仅使用由 CSSS 亚序列定义的蝎形连接子。仅 2Lm16 中的突变为上文对 2Lm5-1 所列的 VH 的 CDR3 中的突变。由亚序列 CSCS、SCCS 和 SCCP 定义的蝎形连接子独立地用于 2Lm16 中。2Lm16-1 在 VL 的 CDR1 中的 27 处具有取代 S 的 Q；且在 VL 的 CDR3 中，在 92 处具有取代 S 的 T 且在 93 处具有取代 F 的 S；且 VH 的 CDR3 中，在 95 处缺失 V、在 96 处具有取代 V 的 S 且在 102 处具有取代 V 的 L；仅使用由 CSSS 亚序列定义的蝎形连接子。2Lm19-3 在 VL 的 CDR1 中，在 27 处具有取代 S 的 Q、在 33 处具有取代 M 的 I 且在 34 处具有取代 H 的 V，以及具有对 2Lm16-1 所列的 VH 的 CDR3 中的突变。由亚序列 CSCS、SCCS 和 SCCP 定义的蝎形连接子独立地用于 2Lm19-3 中。2Lm20-4 分子在 VL 的 CDR1 中，在 33 处含有取代 M 的 I 且在 34 处含有取代 H 的 V，以及含有对 2Lm16-1 所列的 VH 的 CDR3 中的突变。对于 2Lm5-1、2Lm6-1、2Lm16、2Lm16-1、2Lm19-3 和 2Lm20-4 而言，在 VH 的框架区的位置 11 处亦存在取代 L 的 S。由 CSCS、SCCS 和 SCCP 亚序列定义的蝎形连接子独立地用于 2Lm20-4 中。最后，以下突变体中在位置 331 处存在取代 P 的 S：其中蝎形连接子由 CSCS 定义的 018008；其中各蝎形连接子由 CSCS 和 SCCP 定义的 018009;其中蝎形连接子由 CSCS 定义的 018010；其中蝎形连接子由 SCCP 定义的 018011；其中蝎形连接子由 CSCS 定义的 018014;其中蝎形连接子由 CSCS 定义的 018015;其中蝎形连接子由 CSCS、SCCS 和 SCCP 中任一者定义的 2Lm16；其

中蝎形连接子由 CSCS 或 SCCP 定义的 2Lm19-3；及其中蝎形连接子由 CSCS 或 SCCP 定义的 2Lm20-4。

此外，涵盖使结合结构域的两个区(诸如蝎形分子结合结构域中对应于 Ig V_L 和 V_H 的区)连接的连接子长度的变化。例如，其中发现移除结构域间连接子的 C-末端 Asp 预期可影响蝎形分子的结合特征，如同用 Gly 取代 Asp。

亦涵盖具有相对于 Ig 的铰链区而言经延长的蝎形连接子(插入 C-末端与恒定亚区之间和 N-末端与结合结构域 2 之间)的蝎形分子，其中氨基酸残基被添加在蝎形分子中的 C-末端与任何半胱氨酸(对应于 Ig 铰链半胱氨酸)之间，其中该蝎形分子半胱氨酸能够形成链间二硫键。含有所述特征的蝎形分子已构建且于下文中表征。

已努力经由优化使恒定亚区与 C 末端安置的结合结构域 2 共价连接的蝎形连接子来改良蝎形分子的表达、稳定性及疗效。用于优化研究的原型蝎形分子含有使 N-末端与源自 IgG1 C_{H2} 和 C_{H3} 的恒定亚区融合的抗-CD20 scFV(结合结构域 1)和使 C-末端与此恒定亚区融合的二抗-CD20 scFv。该蝎形分子(如免疫球蛋白分子)预期经由恒定区(或亚区)结合以与由二硫键连接的肽链形成同型二聚体复合物。为获取对其 CD20 靶标具有高亲和性的四价稳定分子的高表达程度，恒定亚区与第二结合结构域之间的蝎形连接子必须满足以下考虑。首先，由两个 scFv 片段携有的同源结合结构域之间的空间位阻(两个蝎形分子单体的每一者上具有一个 scFv 片段)应最小化以利于维持各结合结构域的天然构象。其次，结合结构域的构型和取向应使得各结构域可生产性结合且各结合结构域与其靶标可高亲和性结合。第三，蝎形连接子本身应相对具有蛋白酶耐受性和非免疫原性。

在例示性 CD20xCD20 蝎形分子构建体 S0129 中，C_{H3} 的 C-末端与二抗-CD20 scFV 结构域经由 2H7 蝎形连接子(一种源自且对应于 IgG1 的天然人类铰链序列的片段的肽)连接。2H7 蝎形连接子充当设计研究的基础，设计研究使用旨在改良蝎形分子的表达和改良所表达分子的结合特征的计算机辅助模型。

为分析 2H7 蝎形连接子，使用 Insight II 软件对人类 IgG1 铰链的

二聚体形式的 3 维结构进行建模。具有 V_H-V_L 取向的抗-CD20 scFV 的晶体结构选作 20-4 结合结构域的参考结构(RCSB 蛋白数据库条目编码: 1A14)。在完整 IgG1 中，铰链使 C_{H1} 结构域的 C-末端与 C_{H2} 结构域的 N-末端连接，各结构域的构型应使得铰链半胱氨酸残基可成对以形成同型二聚体。在例示性蝎形分子中，源自铰链的 2H7 连接子使源自 IgG1 C_{H3} 结构域的蝎形分子结构域的 C-末端与源自 IgG1 V_{H2} 结构域的蝎形分子结合结构域 2 的此部分的 N-末端连接。利用 V_H-V_L scFV 的三维模型结构，预期 2H7 连接子的 C-末端之间的最佳距离受以下三种考虑因素的影响。第一，必须维持铰链稳定性，且稳定性可通过二聚体化(例如同型二聚体化)来支持，此意谓铰链半胱氨酸必须在两个折叠结合结构域的存在下能够成对。第二，两个结合结构域(例如 scFV)必须满足 2H7 连接子 C-末端不产生位阻干扰以使得蛋白可适当折叠。第三，各结合结构域的 CDR 应能够面对同一方向(如同在天然抗体中)，因为原型蝎形分子的各结合结构域可结合相同细胞表面上的邻近受体(CD20)。出于所述考虑，scFv 的两个 N-末端之间的距离预期为约 28 Å。在二聚体蝎形分子形式中，理论设计的 2H7 连接子的 C-末端之间的距离预期为约 16 Å。为满足预期为优化蝎形分子的效能所需的距离，在 2H7 连接子的 C-末端扩展至少 3 个氨基酸。该扩展预期使得在 2H7 连接子半胱氨酸残基之间可形成二硫键，从而使得 C-末端结合结构域 2 可适当正确折叠和促成 CDR 的正确取向。此外，在完整 IgG1 中，由于铰链与结合结构域之间存在 C_{H1} 和 V_{L1} 结构域，因此两链所携有的结合结构域之间的距离进一步增加且预期可进一步促进相同细胞表面上的邻近受体的交联。鉴于上述考虑，设计一组具有不同长度的连接子(表 10)。为使免疫原性最小化，使用 C_{H2} 结构域的 N-末端所存在的天然残基(Ala-Pro-Glu-Leu 或 APEL)，通过序列添加至蝎形连接子的 C-末端来延长 2H7 蝎形连接子。较长构建体含有一或多个已知具有蛋白酶耐受性和柔性的(Gly4Ser)连接子单元。

在恒定亚区的 C_{H3} 结构域与 C-末端 scFv 结合结构域之间含有经扩展的蝎形连接子的 CD20×CD20 蝎形分子构建体是利用 PCR 诱变法构建且被亚克隆至常规哺乳动物表达载体中。可通过在使用 COS 或

HEK293 细胞的瞬间表达实验中比较所分泌蛋白的产量来分析连接子长度对 CD20×CD20 蝎形分子表达的影响，或通过蛋白质印迹分析法或使用[35]S-标记的甲硫氨酸/半胱氨酸的脉冲追踪研究分析蛋白合成和积聚来分析连接子长度对 CD20×CD20 蝎形分子表达的影响。

表 10

构建体编号	蝎形连接子核心(2H7)序列	扩展序列	经扩展的蝎形连接子序列
1	GCPPCPNS	APEL	GCPPCPNSAPEL
2	GCPPCPNS	APELGGGGS	GCPPCPNSAPELGGGGS
3	GCPPCPNS	APELGGGGSGGGGS	GCPPCPNS APELGGGGSGGGGS
4	GCPPCPNS	APELGGGGSGGGGSGGGGS	GCPPCPNS APELGGGGSGGGGSGGGGS

亦涵盖糖基化蝎形分子，且在本上下文中，涵盖可在碳水化合物改良剂的存在下培养表达蝎形分子的宿主细胞，碳水化合物改良剂在本文中定义为优选具有小于 1000 道尔顿的分子量的小分子有机化合物，其可抑制作为连接至多肽的碳水化合物的部分的糖的添加、移除或修饰中所涉及的酶的活性，诸如在蛋白的 N 连接碳水化合物成熟期间所发生。糖基化为发生于内质网(“核心糖基化”)和高尔基(Golgi)体(“末端糖基化”)中的复杂过程。各种糖苷酶和/或甘露糖苷酶抑制剂提供以下一或多种所要效应：增强 ADCC 活性、增强 Fc 受体结合和改变的糖基化模式。例示性抑制剂包括(但不限于)栗精胺(castanospermine)和抑制 α-甘露糖苷酶(kifunensine)。以下实施例中揭示在至少一种该抑制剂的存在下表达蝎形分子的效应。

实施例 13

蝎形分子蛋白表达程度及表征

测定蝎形分子蛋白表达程度且表征所表达的蛋白以证明蛋白设计可产生具有实用益处的产物。利用常规技术，在培养液中的 CHO DG44

细胞中表达单特异性 CD20×CD20 蝎形分子和双特异性 CD20×CD37 蝎形分子。

观测 CD20×CD20 蝎形分子 S0129 (2lm20-4x2lm20-4)在各种补充饲料存在下所培养的 CHO DG44 细胞中的基准程度的稳定表达，如图 34 中所示。所有培养基均含有 50 nM 甲氨喋呤，该浓度可维持编码蝎形分子的聚核苷酸的拷贝数。聚核苷酸含有蝎形分子蛋白的编码区，该编码区未经密码子最优化以供在 CHO DG44 细胞中表达。使用 pD18 载体将聚核苷酸引入细胞中。由图 34 显然可见，获得约 7-46 μg/ml 的表达程度。

亦测定在扩增编码双特异性 CD20xCD37 蝎形分子的聚核苷酸后的表达程度。使用 pD18 载体克隆 CD20×CD37 蝎形分子编码区，且将质粒引入 CHO DG44 细胞中。使用本领域中已知的 *dhfr*-甲氨喋呤技术实现编码聚核苷酸的扩增，其中使用增加浓度的 MTX 来选择二氢叶酸还原酶基因(*dhfr*)的增加拷贝数，从而使所关注的紧密连接的聚核苷酸共同扩增。图 35 显示，观测到双特异性 CD20×CD37 蝎形分子的稳定表达程度通常为约 22-118 μg/ml。在不同条件(包括用于扩增的甲氨喋呤浓度)下观察到产量变化，但本领域技术人员易优化所述变化。亦使本文中所述的其他各种蝎形分子在 CHO 和/或 COS 细胞中经受表达分析，分析结果提供于下表 11 中。所述结果证明利用常规技术和常规扩增优化技术可获得高产量的蝎形分子蛋白。

所表达的蛋白亦通过 SDS-PAGE 分析来表征，以评估所表达蛋白的同源性程度和完整性且验证单体肽的分子量。在还原和非还原条件下对变性聚丙烯酰胺凝胶(4-20% Tris 甘氨酸)电泳。图 36 中所显示的结果揭示在还原条件下，具有预期单体分子量的 2Lm20-4 SCC SMIP 和 S1000 (CD20(2lm20-4)xCD20(2lm20-4)单特异性蝎形分子 S0126)中的每一者的单一蛋白带。所述数据说明 SMIP 和蝎形分子易以完整形式经受纯化。在非还原条件下，观察到痕量肽与预期尺寸的单体 SMIP 一致，其中大部分蛋白以与二聚体结构一致的单一清晰带形式出现。在所述非还原条件下，单特异性蝎形分子蛋白显示具有与二聚体结构一致的分子量的单一清晰带。SMIP 与蝎形分子的二聚体结构与其单体

结构一致，其各自含有铰链样蝎形连接子，该连接子含有至少一个能够参与二硫键形成的半胱氨酸。

亦评估蝎形连接子对蝎形分子的表达和完整性的影响，且结果显示于表 12 中。此表列出单特异性 CD20xCD20 (2Lm20-4x2Lm20-4) S0129蝎形分子和CD20xCD28 S0033蝎形分子(2H7sccpIgG1-H7-2e12)的蝎形连接子变异体，其完整性如同单链分子和其在 COS 细胞中相对于视需要具有H7连接子(设定为100%)的亲本蝎形分子S0129或S0033的瞬间表达程度。表 13 提供评估并入 CD20×CD20 蝎形分子中的蝎形连接子变异体所得的数据，以及针对 CD20×CD28 蝎形分子的类似数据。表 13 提供评估含有蝎形连接子的 S0129 变异体所得的数据，所述蝎形连接子并非含有至少一个能够形成二硫键的半胱氨酸的铰链样连接子；相反，所述分子中的蝎形连接子是源自 II 型 C-凝集素茎区。由表 13 中所显示的数据显而易见，在瞬间表达测定中，铰链样蝎形连接子可与以高于或低于未经修饰的亲本蝎形连接子的程度表达的蝎形分子结合。此外，某些连接子变异体显示比未经修饰的亲本连接子更大的蛋白水解裂解抗性，这关系到所有或几乎所有的经表达的蛋白。表 13 的数据显示在呈现与含有铰链样蝎形连接子的蝎形分子稍微不同的结合特征的蝎形分子中存在非铰链样连接子，诸如源自 II 型 C-凝集素的茎区的连接子。此外，含有非铰链样蝎形连接子的蝎形分子所显示的效应功能(ADCC)等同于或超过与具有铰链样蝎形连接子的蝎形分子相关的 ADCC。

表 11

连接子名称	上游(CH3)序列	S0129 (2Lm20-4 x 2Lm 20-4)连接子变异体-氨基酸序列 [1]	基于	#AAs	表达 COS[2]	裂解 [3]	表达 CHO[2]
H7	QKSLSLSPGK	GCPPCPNS	H7	18	100	-	100
H16	QKSLSLSPGK	LSVKADFLTPSIGNS	CD80-	*25*	174	+	
H18	QKSLSLSPGK	LSVLANFSQPEIGNS	CD86	25	165	++	
H19	QKSLSLSPGK	LSVLANFSQPEISCPPCPNA	CD86+H7	30	161	+	109
H26	QKSLSLSPGK	RIHQMNSELSVLANS	CD86	25	170	++	
H32	QKSLSLSPGK	RIHLNVSERPFPPNS	CD22	25	184	++	
H47	QKSLSLSPG	LSVKADFLTPSIGNS	H16	*24*	141	-	206

H48	QKSLSLSPG	KADFLTPSIGNS	H16	21	137	-	
H50	Q	LSVLANFSQPEIGNS	H18	16	21	-	
H51	QKS	LSVLANFSQPEIGNS	H18	18	110	-	
H52	QKSLSLSPG	SQPEIVPISNS	H18	20	95	-	
H53	QKSLSL	SQPEIVPISCPPCPNS	H19	26	96	-	
H54	Q	SVLANFSQPEISCPPCPNS	H19	21	72	+/-	
H55	QKSLSLSPG	RIHQMNSELSVLANS	H26	24	118	+	
H56	QKSLSLSPG	QMNSELSVLANS	H26	21	130	-	163
H57	QKSLSLSPG	VSERPFPPNS	H32	*19*	118	-	
H58	QKSLSLSPG	KPFFTCGSADTCPNS	CD72	*24*	103	-	
H59	CKSLS	KPFFTCGSADTCPNS	CD72	20	*94*	-	

[1]NFS 为糖基化一致基序

[2]相对于 SO129-H7(%)，于 COS(6W 板)或 CHO(单一烧瓶)中瞬间表达

[3]通过 SDS-PAGE/银染色法所观测的裂解产物：-=无，+=模糊带；++=主带，+++=>50%裂解)

表 12

连接子名称	S0129 (2Lm20-4 x 2Lm20-4) 连接子变异体-氨基酸序列	CH3 中的变化 [1]	基于以下的连接子序列	20x20 表达 [2]	20x20 裂解 [3]
H7	GCPPCPNS	N	H7	100	-
H8	GSPPSPNS	N	H7	107	+
H9	GSPPSPNS	Y	H7	142	-
H10	EPKSTDKTHTCPPCPNS	N	IgG1 铰链	98	-
H11	EPKSTDKTHTSPPSPNS	N	IgG1 铰链	126	+
H16	LSVKADFLTPSIGNS	N	CD80	174	+
H17	LSVKADFLTPSISCPPCPNS	N	CD80 + H7	113	+
H18	LSVLANFSQPEIGNS	N	CD86	165	++
H19	LSVLANFSQPEISCPPCPNS	N	CD86 + H7	161	+
H20	LKIQERVSKPKISNS	N	CD2	115	+++
H21	LKIQERVSKPKISCPPCPNS	N	CD2 + H7	90	+++
H22	LNVSERPFPPHIQNS	N	CD22	149	++
H23	LDVSERPFPPHIQSCPPCPNS	N	CD22 + H7	121	++
H24	REQLAEVTLSLKANS	N	CD80	145	++
H25	REQLAEVTLSLKACPPCPNS	N	CD80 + H7	98	+
H26	RIHQMNSELSVLANS	N	CD86	170	++
H27	RIHQMNSELSVLACPPCPNS	N	CD86 + H7	154	++
H28	DTKGKNVLEKIFSNS	N	CD2	153	+

H30	LPPETQESQEVTLNS	N	CD22	78	+
H32	RIHLNVSERPFPPNS	N	CD22	184	++
H33	RIHLNVSERPFPPCPPCPNS	N	CD22 + H7	74	+
H36	GCPPCPGGGGSNS	N	H7	110	+
H40	GCPPCPANS	Y	H7	110	+
H41	GCPPCPANS	Y	H7	102	-
H42	GCPPCPNS	Y	H7	99	-
H44	GGGASCPPCPGNS	Y	H7	108	+
H45	GGGASCPPCAGNS	Y	H7	107	-
H46	GGGASCPPCANS	Y	H7	98	-
H47	LSVKADFLTPSIGNS	Y	CD80	141	-
H48	ADFLTPSIGNS	N	CD80	137	-
H50	LSVLANFSQPEIGNS	Y	CD86	21	-
H51	LSVLANFSQPEIGNS	Y	CD86	110	-
H52	SQPEIVPISNS	Y	CD86	95	-
H53	SQPEIVPISCPPCPNS	Y	CD86 + H7	95	-
H54	SVLANFSQPEISCPPCPNS	Y	CD86 + H7	72	+/-
H55	RIHQMNSELSVLANS	Y	CD86	118	+
H56	QMNSELSVLANS	Y	CD86	130	-
H57	VSERPFPPNS	Y	CD22	118	-
H58	KPFFTCGSADTCPNS	Y	CD72	103	-
H59	KPFFTCGSADTCPNS	Y	CD72	94	-
H60	QYNCPGQYTFSMNS	Y	CD69	>100[5]	-
H61	EPAFTPGPNIELQKDSDCNS	Y	CD94	>100	-
H62	QRHNNSSLNTRTQKARHCNS	Y	NKG2A	>100	-
H63	NSLFNQEVQIPLTESYCNS	Y	NKG2D	>100	-

[1]CH3 末端的其他变化，诸如 1-9 氨基酸缺失和/或密码子优化

[2]相对于 S0129-H7 亲本(%)，于 COS(6W 板)中的瞬间表达

[3]通过 SDS-PAGE/银染色法所观测的裂解产物:

-=无，+=模糊带；++=主带，+++>50%裂解

[5]通过评估自 COS 消耗培养基中所纯化的蛋白的回收率来比较 H60-H63 变异体

表 13

蛋白	描述	产量(μg 纯化蛋白/ml 补充物)	POI%(以 Kd 表示的 M.wt,依据 MALS)	S0129 wt POI 的改良	结 合 Ramos	ADCC 测定	蝎形连接子的序列
S0129wt	H7 连接子	1.6	67 (167)	-	-	-	GCPPC
S0129-CD69	CD69 茎区	2.9	66 (167)	1.8	弱 于 S0129 wt	*稍优于 S0129wt POI	QYNCPGQY TFSM
S0129-CD72	CD72 截短茎区	2.0	69 (165)	1.2	类似于 S0129 wt	*稍优于 S0129wt POI	PFFTCGSA DTC
S0129-CD94	CD94 茎区	2.9	67 (171)	1.8	类似于 S0129 wt	*稍优于 S0129wt POI	EPAFTPGPNI ELQKDSDC
S0129-NKG2A	NKG2 A 茎区	2.5	**93 (170)**	2.2	稍优于 S0129 wt	类 似 于 S0129wt POI	QRHNNSSL NTRTQKAR HC
S0129-NKG2D	NKG2 D 茎区	1.9	70 (166)	1.2	类似于 S0129 wt	*稍优于 S0129wt POI	NSLFNQEV QIPLTESYC

如以上实施例中所述，涵盖通过在含有碳水化合物改良剂的培养液中表达来制备蝎形分子。在例示性实施方案中，将栗精胺(MW 189.21)添加至培养基中直至最终浓度为约 200 μM (对应于约 37.8 μg/mL)，或浓度范围大于约 10、20、30、40、50、60、70、80、90、100、110、120、130、140 或 150 μM，及高达约 300、275、250、225、200、175、150、125、100、75、60 或 50 μg/mL。例如，涵盖 10-50 μM 或 50-200 μM 或 50-300 μM 或 100-300 μM 或 150-250 μM 的范围。在其他例示性实施方案中，将 DMJ (例如 DMJ-HCl (MW 199.6))添加至培养基中直至最终浓度为约 200 μM (对应于约 32.6 μg DMJ/mL)或浓度范围大于约 10、20、30、40、50、60、70、80、90、100、110、120、130、140 或 150 μM，及高达约 300、275、250、225、200、175、150、125、100、75、60 或 50 μg/mL。例如，涵盖 10-50 μM 或 50-200 μM 或 50-300 μM 或 100-300 μM 或 150-250 μM 的范围。在其他例示性实施方案中，将抑制 α-甘露糖苷酶(MW 232.2)添加至培养基中直至最终

浓度为约 10 μM (对应于约 2.3 μg/mL)或浓度范围大于约 0.5、1、2、3、4、5、6、7、8、9 或 10 μM，及高达约 50、45、40、35、30、25、20、19、18、17、16、15、14、13、12 或 11 μM。例如，涵盖 1-10 μM 或 1-25 μM 或 1-50 μM 或 5-10 μM 或 5-25 μM 或 5-15 μM 的范围。

在一个实验中，用培养于 200 μM 栗精胺(S0129 CS200)或 10 μM(过量)抑制 α-甘露糖苷酶(S0129 KF 10)中的细胞来表达单特异性 CD20xCD20 蝎形分子(S0129)，且测量所表达的蝎形分子对 WIL2S 细胞的结合或染色，如图 42 中所示。此外，在比较结合研究中，糖基化 S0129 蝎形分子与 CD16 (FCγRIII)的结合比未糖基化 S0129 蝎形分子超过约三倍。

在另一研究中，探究人源化 CD20xCD20 蝎形分子(S0129)对 BJAB B-细胞的 ADCC 介导性杀死作用。图 43 中所示的结果证实，蝎形分子当在栗精胺或抑制 α-甘露糖苷酶的存在下所培养的细胞中表达时，导致在给定浓度的蝎形分子暴露下产生显著更强的 ADCC 介导性 BJAB B-细胞死亡。

实施例 14

蝎形分子结合

a. 结构域间距

双特异性蝎形分子能够利用位于分子的 N-末端和 C-末端处的结合结构域对而同时结合至少两个靶标。在此情况下，对于细胞表面靶标而言，该组合物可与所述靶标交联或促使所述靶标实体上协同趋近。本领域技术人员将了解，多种受体系统经此交联而活化，产生促使细胞表型改变的信号诱导。设计本文中所揭示的组合物部分地旨在使该信号传导最大化且控制所产生的表型。

在蝎形分子架构设计中应了解并考虑蝎形分子组合物中各结构域的近似尺寸以及对以结构域间角范围表示的结构域间柔性的预期值。对于使用 scFv 结合结构域(对于结合结构域 1 和 2 (BD1 和 BD2)而言)、IgG1 N-末端铰链(H1)和本文中所述的 H7 PIMS 连接子的蝎形分子而言，N-末端的结合结构域与 C-末端的结合结构域最大可相隔约

150-180 Å 且最小可相隔约 20-30 Å。N-末端的结合结构域最大可相隔约 90-100 Å 且最小可相隔约 10-20 Å (Deisenhofer 等人，1976, Hoppe-Seyler's Z. Physiol. Chem. Bd. 357, S. 435-445; Gregory 等人，1987, Mol. Immunol. 24(8):821-9.; Poljak 等人，1973, Proc. Natl. Acad. Sci., 1973, 70: 3305-3310; Bongini 等人，2004, Proc. Natl. Acad. Sci. 101: 6466-6471; Kienberger 等人，2004, EMBO Reports, 5: 579-583，各文献以引用方式并入本文中)。部分地选择所述尺寸可使得在蝎形分子所结合的受体复合物中，受体-受体距离小于约 50 Å，此是因为小于此距离的距离对于某些受体寡聚物的最大信号传导而言可为最佳的(Paar 等人，2002, J. Immunol., 169: 856-864，该文献以引用方式并入本文中)，同时使得可并入效应功能所需的 F_C 结构。

蝎形分子的 N-末端和 C-末端的结合结构域经设计成柔性结构以促进靶标结合且容许结合靶标的几何尺寸范围。本领域技术人员亦将了解，N-末端或 C-末端结合结构域(分别为 BD1 和 BD2)之间的柔性和分子的结合结构域与 F_C 结构域之间的柔性以及 BD1 和/或 BD2 所结合的受体之间的最大和最小距离，可例如通过选择 N-末端铰链结构域(H1)和通过结构性模拟更靠近 C-末端定位的蝎形连接子结构域(H2)来加以修饰。例如，IgG1、IgG2、IgG3、IgG4、IgE、IgA2 的铰链结构域、合成铰链及 IgM 的铰链样 C_{H2} 结构域显示不同程度的柔性以及不同长度。本领域技术人员将了解，H1 和蝎形连接子(H2)的最佳选择将视以下因素而定：蝎形分子被设计可与之相互作用的受体系统，以及蝎形分子结合所诱导的所需信号传导表型。

在某些实施方案中，蝎形分子具有蝎形连接子(H2)，该连接子为对应于 Ig 铰链(诸如 IgG1 铰链)的铰链样连接子。所述实施方案包括具有蝎形分子铰链的氨基酸序列的蝎形分子，该氨基酸序列为相对于例如 H7 序列或野生型 IgG1 铰链序列而被 N-末端扩展的序列。此类型的例示性蝎形连接子将具有被 H_2N-APEL$(x)_y$-CO_2H 进行 N-末端扩展的 H7 铰链序列，其中 x 为 Gly4Ser 连接子单元且 y 为 0 至 3 的数值。使用两种蝎形分子(双特异性 CD20xCD28 蝎形分子和单特异性 CD20xCD20 蝎形分子)例示性地研究蝎形连接子对蝎形分子稳定性的

影响。对于该两种蝎形分子设计中的每一者而言，插入各种蝎形连接子。特别地，蝎形连接子 H16 与 H17 的主要区别在于 H17 具有 H16 的序列，H7 的序列随附于 C-末端，且蝎形连接子 H18 与 H19 的主要区别类似地为，在形成 H19 中，H7 的序列随附于 H18 的 C-末端。对于两种蝎形分子主链(20×28 和 20×20)中的每一者而言，将上述四种蝎形连接子中的每一者插入适当位置。在 COS 细胞中获得对所述构建体的瞬间表达，且在蛋白 A/G-涂覆孔(Pierce SEIZE IP 试剂盒)上将存在于培养物上清液中的蝎形分子蛋白纯化。在 SDS-PAGE 凝胶上分离经纯化的蛋白且通过银染色法显现。检验图 44 揭示，蝎形连接子中的另一 H7 序列可增加各类型蝎形连接子和各类型蝎形分子蛋白的稳定性。换句话说，将 H7 附接至 H16 或 H18 的 C-末端可增加蝎形分子的稳定性，且无论该蝎形分子是否为 CD20xCD28 或 CD20xCD20，此观测均成立。在靶标结合方面，具有 CD20×CD20 架构的蝎形分子蛋白显示与亲本单特异性人源化 CD20×CD20 蝎形分子 S0129 类似的结合特性，如图 45 中所示。

然而，除上述实施方案外，需要防止所结合的受体彼此接近于约 50 Å 以内而有意形成次最大信号(Paar 等人，J. Immunol., 169: 856-864)。在此情况下，预期选择比上述那些连接子更短且柔性更小的 H1 和蝎形连接子(H2)为适当的。

相同间距考虑适用于非铰链样的蝎形连接子。所述蝎形连接子例如为具有 C-凝集素的茎区的氨基酸序列的肽类。包含 C-凝集素茎区的例示性蝎形铰链为源自 CD72 茎区、CD94 茎区和 NKG2A 茎区的蝎形铰链。含有所述蝎形铰链的蝎形分子是经构建且在表达、对裂解的敏感性和纯化适应性方面进行表征。数据显示于表 14 中。

表 14

连接子名称	G_4S 密码子优化 [1]	CH3 的末端	**S0129** 蝎形连接子变异体氨基酸序列	基于以下的连接子序列	表达 **(%S0129)** [2]	裂解 [3]	实验台纯化 **%POI**
H7	N	K	GCPPCPNS	H7	100	-	70
H60	Y(17)	K	GCPPCPNS	H7	114	-	ND
H61	Y(15)	K	GCPPCPNS	H7	90	-	66

H62	N	G	QRHNNSSLNTR TQKARHCPNS	NKG2A 茎区	129	-	89
H63	Y(17)	G	QRIINNSSLNTR TQKARHCPNS	NKG2A 茎区	100	-	85
H64	Y(15)	G	QRHNNSSLNTR TQKARHCPNS	NKG2A 茎区	81	-	93
H65	N	G	EPAFTPGPNIEL QKDSDCPNS	CD94 茎区	133	-	66
H66	Y(17)	G	EPAFTPGPNIEL QKDSDCPNS	CD94 茎区	200	-	64
H67	Y(15)	G	EPAFTPGPNIEL QKDSDCPNS	CD94 茎区	129	-	65
H68	N	G	RTRYLQVSQQL QQTNRVLEVTN SSLRQQLRLKIT QLGQSAEDLQG SRRELAQSQEA LQVEQRAHQAA EGQLQACQADR QKTKETLQSEE QQRRALEQKLS NMENRLKPFFT CGSADT	CD72 完整茎区	110	-	75

[1]Gly_4Ser 连接子的密码子优化，具有(17)限制性位点或不具有(15)限制性位点

[2]基于实验台纯化的蛋白回收率评估在 COS 中的表达

[3]通过 SDS-PAGE/库马斯蓝染色经纯化的蛋白所观测到的一个或多个裂解产物

b. N-末端和 C-末端结合结构域的结合

N-末端与 C-末端结构域参与靶标细胞结合

CD20 SMIP(TRU015)、CD37 SMIP(SMIP016)、CD20 与 CD37 SMIPS 的组合(TRU015+SMIP016)和 CD20xCD37 双特异性蝎形分子(015x016)的靶标细胞结合能力是通过测量所述分子各自阻断与相关靶标(CD37 或 CD20)特异性竞争结合的抗体的结合的能力来评估。竞争性抗体视需要为经 FITC 标记的单克隆抗-CD37 抗体或经 PE 标记的单克隆抗-CD20 抗体。Ramos B-细胞提供靶标。

将具有 5%小鼠血清的 PBS (#100-113, Gemini Bio-Products, West Sacramento, CA) (染色介质)中的 Ramos B-细胞(1.2×10^7 个细胞/毫升)

添加至 96 孔 V 形底板(25 微升/孔)中。将各种 SMIP 和蝎形分子于染色介质中稀释至 75 μg/ml，且 4 倍稀释至图 38 中所示的浓度。将经稀释的化合物添加至所涂板的细胞中(除单独用于对照孔的介质外)。将细胞用所述化合物孵育 10 分钟，且接着将 5 μg/ml 的 FITC 抗-CD37 抗体(#186-040, Ancell, Bayport, MN)和 3 μg/ml 的 PE 抗-CD20 抗体(#555623, BD Pharmingen, San Jose, CA)(纯)一起添加至孔中的 25 μl 染色介质中。将所述细胞在黑暗中于冰上孵育 45 分钟且接着用 PBS 洗涤 2.5 次。将细胞用 1%三聚甲醛(#19943 1 LT, USB Corp, Cleveland, OH)固定，且接着在 FACs Calibur(BD Biosciences, San Jose, CA)上运作。用 Cell Quest 软件(BD Biosciences, San Jose, CA)分析数据。图 38 中所示的结果证实，所有 SMIP、SMIP 组合和含有 CD20 结合位点的蝎形分子均成功地与经 PE 标记的抗-CD 20 抗体竞争结合至 Ramos B-细胞(上图)；所有 SMIP、SMIP 组合和含有 CD37 结合位点的蝎形分子均成功地与经 FITC 标记的抗-CD 37 抗体竞争结合至 Ramos B-细胞(下图)。因此，双特异性 CD20×CD37 蝎形分子被证明对于 B-细胞上的靶标具有可操作的 N-末端和 C-末端结合位点。

c. 细胞表面持久性

研究所结合的 SMIP 和蝎形分子(单特异性和双特异性)在 B-细胞表面上的细胞表面持久性揭示蝎形分子比 SMIP 显示更强的细胞表面持久性。将染色介质(PBS 中的 2.5%山羊血清、2.5%小鼠血清)中的 6×10^6 个细胞/毫升的 Ramos B-细胞(3×10^5 个细胞/孔)添加至 96 孔 V 形底板中。通过 5 倍连续稀释 500 nM 初始物质于染色介质中制备两倍最终浓度的测试试剂，且接着以 1:1 添加至 Ramos B-细胞中。此外，亦涂覆对照介质。将细胞在黑暗中于冰上孵育 45 分钟。接着将板用冷 PBS 洗涤 3.5 次。接着以 1:100 稀释度将第二试剂 FITC 山羊抗-人类 IgG (#H10501, Caltag/Invitrogen, Carlsbad, CA)添加于染色介质中。将细胞在黑暗中于冰上孵育 30 分钟。接着将细胞通过离心用冷 PBS 洗涤 2.5 次，用 1%三聚甲醛溶液(#19943 1 LT, USB Corp, Cleveland, OH)固定，且接着在 FACs Calibur (BD Biosciences, San Jose, CA)上运作。

用 CellQuest 软件(BD Biosciences, San Jose, CA)分析数据。数据分析结果显示于图 37 中，该图显示若干种 SMIP (单特异性 CD20×CD20 蝎形分子和双特异性 CD20×CD37 蝎形分子)与其靶标在 Ramos B 细胞上的结合。

将具有 Ramos B-细胞($7x10^5$个细胞/毫升)的两个管在冰上孵育 30 分钟，对两种化合物(亦即人源化 CD20 (2Lm20-4) SMIP 和人源化 CD20×CD20 (2Lm20-4x2Lm20-4)蝎形分子)中的每一者(各为 25 μg/ml，于具有 10% FBS 的 Iscoves 介质中)进行研究。孵育期结束时，通过离心将两管洗涤 3 次。接着将一管的细胞(于 150 μl Iscoves 介质中)以$2×10^5$个细胞/孔涂覆于 96 孔平底板中，接着将一板置于 37°C 恒温箱中且将另一板在冰上孵育。将第二管的各组细胞再悬浮于具有 2%小鼠血清及 1%叠氮化钠的冷 PBS(染色介质)中，且以$2×10^5$个细胞/孔涂覆于 96 孔 V 形底板中以便经第二抗体(亦即 FITC 山羊抗人类 IgG (#H10501, Caltag/Invitrogen, Carlsbad, CA))快速染色。将第二抗体以 1:100 最终稀释度添加于染色介质中且使细胞在黑暗中于冰上染色 30 分钟。接着将细胞用冷 PBS 洗涤 2.5 次且用 1%三聚甲醛(#19943 1 LT, USB Corp, Cleveland, OH)固定。

在图 39 中指定的时点，自 96 孔平底板中收获样本，在 37°C 下或在冰上孵育，且置于 96 孔 V 形底板($2×10^5$个细胞/孔)中。将细胞用冷染色介质洗涤一次，再悬浮，且将第二抗体以 1:100 的最终稀释度添加于染色介质中。将所述细胞在黑暗中于冰上孵育 30 分钟。接着通过离心将细胞于冷 PBS 中洗涤 2.5 次，且随后用 1%三聚甲醛固定。将样本在 FACS Calibur (BD Biosciences, San Jose, CA)上运作，且用 CellQuest 软件(BD Biosciences, San Jose, CA)分析数据。图 39 中所显示的结果证明，SMIP 和蝎形分子与 B-细胞表面的结合持续至少六小时，其中单特异性 hu CD20xCD20 (2Lm20-4x2Lm20-4)蝎形分子比 hu CD20 (2Lm20-4) SMIP 持续的程度更大。

实施例 15

单特异性和双特异性蝎形分子的直接细胞杀死作用

实验旨在评估单特异性和双特异性蝎形分子直接杀死淋巴瘤细胞的能力，亦即不涉及 ADCC 或 CDC 来杀死所述细胞的能力。特别地，使 Su-DHL-6 和 DoHH2 淋巴瘤细胞系单独地经受单特异性蝎形分子(亦即 CD20×CD20 蝎形分子或 CD37×CD37 蝎形分子)，或经受双特异性 CD20×CD37 蝎形分子。

利用常规技术建立 Su-DHL-6、DoHH2、Rec-1 和 WSU-NHL 淋巴瘤细胞的培养液，且接着使某些所述培养液个别地暴露于单特异性 CD20 SMIP、单特异性蝎形分子(CD20×CD20 或 CD37×CD37)或双特异性蝎形分子(CD20×CD37 或 CD19×CD37)。细胞暴露于 SMIP 或蝎形分子是在不导致交联的条件下进行。使细胞与所述分子保持接触 96 小时，之后如本领域中所知，通过检测 ATP 来测量其生长。自图 24 和表 15 中显而易见 CD20 SMIP 和 CD20xCD20 单特异性蝎形分子所产生的细胞杀死作用。自图 25 和表 15 中显而易见 CD37×CD37 单特异性蝎形分子的细胞杀死能力，自图 26 和表 15 中显而易见 CD20×CD37 双特异性蝎形分子杀死淋巴瘤细胞的能力且自图 27 和表 15 中显而易见 CD19×CD37 双特异性蝎形分子杀死淋巴瘤细胞的能力。将三次独立实验的数据组合，且各点表示平均值±标准误差(SEM)。如表 15 的符号所注，表 15 中的 IC_{50} 值通过图 24、图 25 和图 26 中的曲线确定，且定义为与未经处理的培养物相比产生 50%抑制作用的浓度。所述图和表格中的数据证明，蝎形分子杀死所述细胞系的效力与使用相同结合结构域的游离 SMIP 相比大出 10 倍以上。

表 15

	细胞系		
IC_{50} (nM)	SU-DHL-6	DoHH2	WSU-NHL
CD20 SMIP*	>100	60	NA
CD20xCD20 蝎形分子*	0.3	4.0	NA
CD37 SMIP**	>100	>100	NA
CD37xCD37 蝎形分子**	10	1.2	NA
CD20 SMIP 和 CD37 SMIP***	6	2	NA

	细胞系		
CD20×CD37 蝎形分子***	0.05	0.05	NA
CD19 SMIP 和 CD37 SMIP****	0.16	NA	0.40
CD19×CD37 蝎形分子****	0.005	NA	0.04

*数据源自图 24。

**数据源自图 25。

***数据源自图 26。

****数据源自图 27。

于以下细胞中进行人源化 CD20xCD20 蝎形分子 S0129 的其他实验：Su-DHL-4、Su-DHL-6、DoHH2、Rec-1 和 WSU-NHL 细胞。结果显示于图 46 和图 47 中。所述图中所提供的数据扩展上述结论，其证明蝎形分子具有直接杀死各种细胞系的能力。

上述结论可推广至其他单特异性和双特异性蝎形分子，各种蝎形分子证明具有直接杀死 B 细胞的能力。使 DoHH2 B-细胞体外暴露于单特异性 CD20×CD20 蝎形分子、单特异性 CD37×CD37 蝎形分子或双特异性 CD20×CD37 蝎形分子。图 48 中所显示的结果证明双特异性蝎形分子具有形式不同于单特异性蝎形分子的杀死率曲线。

在 70 nM CD20xCD20 蝎形分子(S0129)、CD20xCD37 蝎形分子或 CD37xCD37 蝎形分子的存在下培养 Su-DHL-6 细胞亦在体外环境中产生直接 B-细胞杀死作用(图 49)。同样，Su-DHL-6 细胞暴露于双特异性 CD19xCD37 蝎形分子或暴露于 Rituxan®产生直接细胞杀死作用，而双特异性蝎形分子在较低剂量下显示致死性，如图 50 中所揭示。

通过使 DHL-4 细胞暴露于识别 CD20 的四种独立的单特异性蝎形分子，可提供直接细胞杀死作用的另一证明。将 CD20xCD20 蝎形分子的两种型式设计成合并两个 20-4 结合结构域(20-4x20-4 和 S0129)且另外两种合并 011 与 20-4 结合结构域的杂交体。两种 CD20×CD20 蝎形分子设计的所有四种经独立构建且经纯化的型式((20-4×20-4 和 S0129)和杂交体(011×20-4 和 011×20-4ΔAsp))以直接方式有效杀死 DHL-4 细胞。对于该研究，用 1 μg/ml 指定蛋白将 DHL-4 细胞经体外

处理 24 小时。接着将细胞分别用膜联蛋白 V 和碘化丙啶(细胞死亡的早期和晚期标志物)染色，且通过 FACS 量化细胞群。如以黑柱所示的染色增强所证明，图 51 中所显示的结果证明各 CD20×CD20 构建体的直接杀死能力。此外，所述结果证明，相对于基于 20-4x20-4 的蝎形分子，杂交体 011x20-4 蛋白对于直接细胞杀死作用显示略微增强，尽管所述蝎形分子各自单特异性识别 CD20。在各别实验组中，通过 FACS 分析膜联蛋白 V 染色细胞群和碘化丙啶染色细胞群来测定四种独立蝎形分子构建体的剂量-反应。图 52 中所示的结果证明，在用各种独立蝎形分子构建体处理 DHL-4 细胞所导致的细胞死亡中，剂量-反应性增加。

实施例 16

蝎形分子所介导的附带功能(ADCC 和 CDC)

a. 蝎形分子依赖性细胞的细胞毒性

实验旨在判定蝎形分子是否介导 BJAB B 淋巴瘤细胞的杀死作用。经观测，BJAB B 淋巴瘤细胞可由 CD20 和/或 CD37 蝎形分子杀死。

初始时，在 37°C 下，在具有 10% FBS 的 Iscoves 介质中将 BJAB B-细胞(1×10^7 个/毫升)用 500 μCi/ml ^{51}Cr 铬酸钠(#CJS1, Amersham Biosciences, Piscataway, NJ)标记 2 小时。接着将载有 ^{51}Cr 的 BJAB B 细胞于具有 10% FBS 的 RPMI 介质中洗涤 3 次且以 4×10^5 个/毫升再悬浮于 RPMI 中。通过淋巴细胞分离介质(#50494, MP Biomedicals, Aurora, Oh)离心，从肝素化全血中分离实验室供者的外周血液单核细胞(PBMC)，用 RPMI 介质洗涤 2 次且以 5×10^6 个/毫升再悬浮于具有 10% FBS 的 RPMI 中。将试剂样本以 4 倍最终浓度添加至具有 10% FBS 的 RPMI 介质中，且对各试剂制备三份 10 倍连续稀释液。接着将所述试剂以 50 微升/孔添加至 96 孔 U 型底板中直至所指定的最终浓度。接着将经 ^{51}Cr 标记的 BJAB 细胞以 50 微升/孔(2×10^4 个细胞/孔)添加至板中。接着将 PBMC 以 100 微升/孔(5×10^5 个细胞/孔)添加至板中，使得效应物(PBMC): 靶标(BJAB)的最终比率为 25:1。将效应物和靶标添加至单独介质中以测量本底杀死率。将经 ^{51}Cr 标记的 BJAB 添加至单独介质

中以测量 ^{51}Cr 的自发性释放率且添加至具有 5% NP40(#28324, Pierce, Rockford, Ill)的介质中以测量 ^{51}Cr 的最大释放率。将所述板在 37°C 下于 5% CO_2 中孵育 6 小时。接着自各孔取 50 μl (25 μl 亦为合适的)上清液转移至 LumaPlate-96 (#6006633, Perkin Elmer, Boston, Mass)中且在室温下干燥隔夜。

干燥之后，在 Packard TopCount-NXT 上以 cpm 量化放射性发射。样本值为三份样本的平均值。使用以下方程式计算特异性杀死百分比：杀死%=((样本-自发性释放率)/(最大释放率-自发释放率))×100。图 30 中的曲线显示 BJAB B 细胞通过单特异性蝎形分子 CD20×CD20 和 CD37×CD37 杀死。CD20 SMIP 与 CD37 SMIP 的组合亦杀死 BJAB B 细胞。所述结果证明蝎形分子显示蝎形分子依赖性细胞的细胞毒性且预期该功能性通过提供 ADCC 活性的蝎形分子的恒定亚区提供。

b. 蝎形分子在补体依赖性细胞毒性中的作用

实验亦证明蝎形分子具有补体依赖性细胞毒性(CDC)活性。如下所述且如图 31 中所示，实验涉及使 Ramos B-细胞暴露于 CD19 和/或 CD37 SMIP 及蝎形分子。

实验初始为将 50 μl 的 Iscoves 介质(无 FBS)中的 5×10^5 至 2.5×10^5 个 Ramos B-细胞添加至 96 孔 V 形底板的孔中。将 Iscoves 中的测试化合物(或单独 Iscoves)以 50 μl 分两次以指定最终浓度添加至孔中。将细胞和试剂在 37°C 下孵育 45 分钟。将细胞用无 FBS 的 Iscoves 洗涤 2.5 次且以指定浓度再悬浮于 96 孔板中的具有人类血清的 Iscoves (#A113, Quidel, San Diego, CA)中。接着将细胞在 37°C 下孵育 90 分钟。将细胞通过离心来洗涤且再悬浮于 125 μl 冷 PBS 中。接着将细胞转移至 FAC 簇管(#4410, CoStar, Corning, NY)且添加 125 μl 具有 5 μg/ml 的碘化丙啶(#P-16063, Molecular Probes, Eugene, OR)的 PBS。将细胞在室温下于黑暗中用碘化丙啶孵育 15 分钟，且接着置放于冰上，量化，且在 FACsCalibur 上经 CellQuest 软件(Becton Dickinson)分析。图 31 中所显示的结果证明，CD19 SMIP (而非 CD37 SMIP)显示 CDC 活性，该两种 SMIP 的组合显示与单独 CD19 SMIP 大致相同的 CDC 活性程度。然而，CD19xCD37 蝎形分子显示比单独 SMIP 或 SMIP 组合显著

更强的 CDC 活性，此说明蝎形分子架构可提供比其他分子设计更大程度的补体依赖性细胞毒性。

c. CD20×CD20 单特异性蝎形分子的 ADCC/CDC 活性

对三种不同的 CD20xCD20 单特异性蝎形分子以及适当对照分子测试 ADCC 和 CDC 功能性。利用常规技术测定 ADCC，且结果显示于图 53 中。由图中显而易见与所测试的各 CD20×CD20 单特异性蝎形分子相关的显著(但非一致)的 ADCC 活性。

为评估 CDC，在 37°C 下将 Ramos B-细胞样本(4×10^5)用各 CD20×CD20 蝎形分子(0、0.5、5、50 和 500 nM)和血清(10%)孵育 3.5 小时。通过 7-AAD 染色和 FACS 分析来评估细胞死亡。结果显示于图 54 中，该图揭示蝎形分子显示某些 CDC 活性。在类似实验中，在 37°C 下将 Ramos B-细胞样本(4×10^5)用 CD20×CD20 蝎形分子蛋白(5、50、100 nM)和血清(10%)孵育 2 小时。将细胞洗涤 2 次且与抗人类 C1q FITC 抗体一起孵育。通过 FACS 分析来评估所结合的 C1q，且结果显示于图 55 中。所述结果与图 54 中所显示的结果一致，说明各 CD20xCD20 单特异性蝎形分子与某些 CDC 活性相关，但活性弱于与 CD20 SMIP 相关的活性。

d. 蝎形分子与 $F_C\gamma$RIII 的相互作用

ELISA 研究显示，在靶标细胞不存在的情况下，蝎形分子以增强程度结合低 FcγRIII(CD16) (低亲和性同种型或等位基因型)。初始利用常规技术，用低亲和性或高亲和性 CD16mIgG 涂覆 ELISA 板。评估此固定融合蛋白捕获 CD20 SMIP 或 CD20×CD20 单特异性蝎形分子的能力。用山羊抗人类 IgG(HRP)二级抗体检测所结合的 SMIP 和蝎形分子，且测定平均萤光强度(MFI)。单独 PBS(阴性对照物)以单点显示。结果显示于图 32A 中(通过 CD16 高亲和性同种型融合来捕获)和图 32B 中(通过 CD16 低亲和性同种型融合来捕获)。由对图 32A 和图 32B 的考虑显而易见，CD20 SMIP 与 CD20×CD20 单特异性蝎形分子显示与高亲和性和低亲和性 CD16 同种型融合体的结合增强，其中 CD20×CD20 蝎形分子显示与低亲和性同种型融合体的结合随蛋白浓度增加而显著增强。

亦评估在靶标细胞存在下蝎形分子与 FcγRIII 同种型的结合。数据显示，在靶标细胞存在下，蝎形分子与 FcγRIII(CD16)低亲和性同种型或等位基因型和高亲和性同种型或等位基因型的结合随蛋白浓度增加而增强。

在进行实验中，在容许 SMIP 或蝎形分子与 CD20 阳性靶标细胞结合的条件下，使 CD20 阳性靶标细胞暴露于 CD20 SMIP 或 CD20×CD20 单特异性蝎形分子。随后，使携有 SMIP 或蝎形分子的靶标细胞暴露于经小鼠 IgFc 标记的 CD16 高亲和性或低亲和性同种型。接着将经标记的山羊抗小鼠 Fc 作为二级抗体添加，以标记与小鼠 IgFc 连接的固定化 CD16。接着在 FACs Calibur(BD Biosciences, San Jose, CA)上利用流式细胞术检测细胞，且用 Cell Quest 软件(BD Biosciences, San Jose, CA)分析。如图 33 中所示，在靶标细胞存在下，增加 CD20 SMIP 和 CD20×CD20 单特异性蝎形分子每一者的浓度可使与 CD16 同种型的结合增强，其中 CD20×CD20 蝎形分子的结合性的增强比对于 CD20 SMIP 所见的结合性增强更为显著。

实施例 17

蝎形分子对靶标淋巴瘤细胞的细胞周期效应

通过使淋巴瘤细胞暴露于 SMIP、单特异性蝎形分子和双特异性蝎形分子来评估蝎形分子的细胞周期效应。更特别地，将 DoHH2 淋巴瘤细胞(0.5×10^6)用 0.4 nM 利妥昔单抗(rituximab)、CD20×CD37 蝎形分子、TRU-015(CD20 SMIP)+ SMIP-016 组合(各为 0.2 nM)、100 nM SMIP-016 或 100 nM CD37×CD37 蝎形分子处理 24 小时。在 96 小时生长抑制测定中，所述浓度显示的 IC_{50} 值比蝎形分子的 IC_{50} 值大出约 10 倍(参见图 24-27)。在 37°C 下，将培养物用 10 μM BrdU (溴去氧尿苷)标记 20 分钟。固定之后，将细胞用抗-BrdU-FITC 抗体染色且用碘化丙啶对比染色。图 28 中的数值为 2-3 个独立实验的 4 份复制式培养物的平均值+/-SD。同时分析所有样本数据且利用 BrdU 与 PI 合并点曲线来组合显示。曲线图证明，蝎形分子处理的主要效应为对处于 S-期的细胞的耗尽，以及 G_0/G_1 间隔的增加。

实施例 18

蝎形分子的生理效应

a. 线粒体电位

如 JC-1 测定中所揭示，CD20×CD20 蝎形分子诱导 DHL4 B-细胞的线粒体膜电位下降。JC-1 为显示线粒体的电位依赖性积聚的阳离子型碳菁染料(来自 Molecular Probes 的 Mitoprobe® JC-1 流式细胞术测定试剂盒)。与质膜相比，JC-1 对线粒体膜更具特异性，且用于测定线粒体膜电位的变化。线粒体的积聚可由绿色(529 nm)至红色(590 nm)的萤光偏移来指示。

在进行实验中，初始将 DHL-4 B-细胞(5×10^5 个细胞/毫升)培养于 24 孔板中且于标准组织培养恒温箱中，在 37°C 下于 5% CO_2 中用 1 μg/ml 的 CD20×CD20 蝎形分子、利妥昔单抗、IgG 对照抗体或 5 μM 星孢菌素(staurosporine)处理 24 小时。添加 JC-1 染料(10 μl/ml，2 μM 最终浓度)且在 37°C 下将细胞再孵育 30 分钟。将细胞通过离心(1200 rpm，5 分钟)来收获，用 1 ml PBS 洗涤且再悬浮于 500 μl PBS 中。通过流式细胞术(FACSCalibur, BD)经由 488 nM 激发及 530 nM 和 585 nM 发射滤光器来分析细胞。对于图 56 中所示的代表性散布图而言，在 Y 轴上测量红色萤光且在 X 轴上测量绿色萤光。如以阳性对照 CCCP(羰基氰化物 3-氯苯腙)(一种已知的线粒体膜电位扰乱剂)所见，线粒体膜的去极化是作为红色萤光的减弱来测量。为验证 JC-1 对膜电位变化具有反应，将 DHL-4 B-细胞在 37°C、5%CO_2 下用两种浓度的 CCCP (50 μM 和 250 μM)处理 5 分钟。另一种阳性对照物为用广谱激酶抑制剂星孢菌素处理以诱导细胞凋亡的细胞。图 56 中所示的结果为 10,000 点数的点曲线图，其中红色萤光描绘于 Y 轴上且绿色萤光描绘于 X 轴上。具有扰乱线粒体膜电位的细胞百分比的概括性直方图(扰乱 MMP: 黑柱)显示于图 56 中。所述结果证明，用 20-4×20-4 蝎形分子或 011×20-4 蝎形分子处理可引起与细胞死亡相关的线粒体膜电位的降低。

b. 钙流量

利用 Ca^{++}流动性(细胞信号传导的共同特征)作为量度，分析蝎形

分子对细胞信号传导途径的影响。将 SU-DHL-6 淋巴瘤细胞用钙 4 染料标记且用下文所鉴别的测试分子处理。对细胞读数历时 20 秒钟以测定本底萤光，且接着添加 SMIP/蝎形分子(图 28 中的第一虚线)，且测量萤光直至 600 秒钟。在 600 秒钟时，添加 8 倍过量的交联山羊抗人类 F(ab)'2，且再量测萤光 300 秒钟。图 28 中的图(A)显示，与未受刺激的细胞相比(蓝线)，由 CD20 SMIP 与 CD37 SMIP 的组合所获得的结果(红线)；由 CD20×CD37 双特异性蝎形分子所获得的结果(黑线)。在图 28 中的图 B 中，用单独 CD20 SMIP 处理细胞(红线)的结果导致 Ca^{++} 流动现象，但此并不如单特异性 CD20×CD20 蝎形分子所产生的信号(黑线)那样强。图 28 的 Ca^{++} 流动图表示获自经等摩尔量的蝎形分子和 SMIP/SMIP 组合处理的三个孔的萤光。

c. 半胱天冬酶 3、7 和 9

如通过膜联蛋白 V 和碘化丙啶染色增加和线粒体膜电位下降所证明，结合 CD20 的蝎形分子直接杀死 B-细胞的能力引起对结合 CD20 的蝎形分子对 B-细胞的其他细胞凋亡相关效应的进一步研究。所采用的方法是对暴露于 CD20×CD20 蝎形分子或适当对照物的 DHL-4 B-细胞进行 Apo1 测定。Apo1 测定是基于半胱天冬酶 3 和 7 的合成肽底物。测定组份可购自 Promega (Apo-ONE®同源半胱天冬酶-3/7 测定)。半胱天冬酶介导经标记的肽 Z-DEVD-若丹明 110 (Rhodamine 110)裂解释放萤光若丹明 110 标志物，使用 485 nm 激发和 530 nm 检测对其进行测量。

在实验中，将 100 μl DHL-4 B-细胞(1×10^6 个细胞/毫升)涂覆于黑色 96 孔平底组织培养板中且在标准组织培养恒温箱中，在 37°C、5% CO_2 下用 1 μg/ml 的 CD20×CD20 蝎形分子、利妥昔单抗、IgG 对照抗体或 5 μM 星孢菌素处理 24 或 48 小时。(星孢菌素为小分子广谱蛋白激酶抑制剂，在本领域中已知其为多种细胞类型的经典细胞凋亡的有效诱导剂)。24 或 48 小时后，将 100 μl 的 100 倍稀释底物添加至各孔中，在板式震荡器(300 rpm)上轻缓地混合一分钟且在室温下孵育两小时。使用 485 nM 激发和 527 nM 发射滤光器(Fluoroskan Ascent FL, Thermo Labsystems)测量萤光。图 57 中显示的图形显示 24 小时和 48

小时(对于星孢菌素，仅 24 小时)后，三次处理的平均萤光强度+/-标准偏差。所述结果证明，结合 CD20 的蝎形分子不经由涉及半胱天冬酶 3/7 活化的细胞凋亡途径来直接杀死 B-细胞。

Apo-1 测定中所获得的结果是通过蛋白质印迹分析法来验证，所述蛋白质印迹分析法经设计成检测导致半胱天冬酶活化的促半胱天冬酶裂解或检测 PARP(聚(ADP-核糖)聚合酶)(一种已知通过活化半胱天冬酶 3 而裂解的蛋白)的裂解。使 DHL-4 B-细胞暴露于结合 CD20 的蝎形分子或对照物历时 4、24 或 48 小时，且将细胞裂解物经 SDS-PAGE 分离并利用常规技术进行蛋白质印迹法分析。图 58 显示的结果为蛋白质印迹集合的形式。底部三处蛋白质印迹是使用抗半胱天冬酶抗体来检测反映蛋白水解活化的半胱天冬酶分子量的变化。对于半胱天冬酶 3、7 和 9，不存在半胱天冬酶受任何 CD20-结合分子活化的迹象。星孢菌素充当测定的阳性对照物且对半胱天冬酶 3、7 和 9 中的每一者诱导活性半胱天冬酶的促半胱天冬酶裂解。图 58 中所示的第四个蛋白质印迹显示，PARP(一种已知的活化半胱天冬酶 3 的底物)未裂解，此与结合 CD20 的蝎形分子未能活化半胱天冬酶 3 相一致。所有所述实验的结果均一致，说明半胱天冬酶 3 活化并非由结合 CD20 的蝎形分子所诱导的 DHL-4 B-细胞的直接细胞杀死作用的重要特征。

此外，时间连续研究旨在测定 CD20 结合蛋白(包括 CD20×CD20 蝎形分子)对半胱天冬酶 3 的效应。将 DoHH2 或 Su-DHL-6 B-细胞与 10 nM CD20 结合蛋白(S0129 蝎形分子、2Lm20-4 SMIP 或 Ritu×an®)+/-可溶性 CD16 Ig(40 nM)、单独可溶性 CD16 Ig 或介质一起孵育。将细胞以 3×10^5 个/孔/300 微升培养于具有 10% FBS 的完全 RPMI 中且在 4 小时、24 小时或 72 小时时收获。将 72 小时时间点的样本涂覆于 500 μl 测试剂中。将细胞用 PBS 洗涤，且接着使用 BD Pharmingen 半胱天冬酶 3(活性形式)、mAB 细胞凋亡试剂盒：FITC(目录号第 55048 号，BD Pharmingen, San Jose, CA)以细胞内活性半胱天冬酶-3 染色。简而言之，在以冷 PBS 再进行 2 次洗涤后，将细胞悬浮于冷 cytofix/cytoperm 溶液中且在冰上孵育 20 分钟。接着通过离心来洗涤细胞，抽吸，且在室温下用 Perm/Wash 缓冲剂洗涤两次。接着在室温下于黑暗中将样本于

100 μl 的 Perm-Wash 缓冲剂中用 20 μl FITC-抗半胱天冬酶 3 染色 30 分钟。接着将样本用 Perm-Wash 缓冲剂洗涤两次，且再悬浮于 500 μl 的 Perm-Wash 缓冲剂中。接着将所洗涤的细胞转移至 FACs 管中，且在 FACs Calibur(BD Biosciences, San Jose, CA)上运作，并用 Cell Quest 软件(BD Biosciences, San Jose, CA)分析。结果显示于表 16 中。

表 16

分子(10 nM)	半胱天冬酶-3 阳性细胞百分比			存活百分比		
	4 小时	24 小时	48 小时	4 小时	24 小时	48 小时
RT×N	7	25	7	75	53	56
与 CD16 hi (4×)	27	47	21	79	60	43
CD20 SMIP (2Lm 20-4)	5	5	10	89	85	81
与 CD16 hi	28	54	21	61	60	41
人源化 CD20×CD20 蝎形分子 (S00129)	7	13	14	69	68	61
与 CD16 hi	30	31	15	67	75	72
介质	7	5	9	89	82	80
与 CD16 hi	6	5	9	91	83	80

所有所述实验的结果均一致，说明在 CD16 不存在下，对半胱天冬酶 3 的活化有限，此并非意谓半胱天冬酶 3 活化为由结合 CD20 的蝎形分子所诱导的直接细胞杀死作用的重要特征。

d. DNA 断裂

诱导经典细胞凋亡信号传导途径最终导致染色体 DNA 的浓缩和断裂降解。为测定结合 CD20 的蝎形分子是否经由经典细胞凋亡机制直接杀死 B-细胞，在使细胞暴露于结合 CD20 的蝎形分子或对照物之后，检查 B-细胞染色体 DNA 的状态。初始时，用 CD20 结合分子(亦即单特异性 CD20×CD20(2Lm20-4×2Lm20-4)蝎形分子、CD20×CD20 (011×2Lm20-4)蝎形分子或利妥昔单抗)或用对照物体外处理 DHL-4 B-细胞历时 4 小时、24 小时或 48 小时。随后，将细胞溶解且利用常规技术将染色体 DNA 纯化。接着将染色体 DNA 通过凝胶电泳法进行分级分离。图 59 中所示的凝胶电泳图显示 DNA 断裂的缺乏，证明由结

合 CD20 的蝎形分子所引起的细胞死亡并非通过经典细胞凋亡途径所介导。经星孢菌素处理的细胞在所述测定用作阳性对照物。

e. SYK 磷酸化作用

SYK 为具有若干磷酸化位点的磷酸调节蛋白，其用作转录抑制子。SYK 位于细胞核中，但活化后能够快速再定位于细胞膜中。对于活化而言，SYK 须保留其细胞核定位序列。活化 SYK 具有抑制乳癌肿瘤的作用且 SYK 通过促细胞凋亡信号(诸如电离辐射、BCR 连接和 MHC II 类交联)活化。此外，已证明 SYK 可影响 PLC-γ 和 Ca^{++}通路。经由所述观测，研究结合 CD20 的蝎形分子影响 SYK 的能力。

使 DHL-6 B-细胞暴露于双特异性 CD20×CD37 蝎形分子历时 0、5、7 或 15 小时且使细胞溶解。用抗磷酸化酪氨酸抗体或用抗-SYK 抗体使溶胞物免疫沉淀通过凝胶电泳将免疫沉淀物分离，且结果显示于图 60 中。由检验图 60 显而易见，双特异性 CD20×CD37 蝎形分子未能诱发 SYK 的磷酸化作用，从而未能将其活化。与上述对半胱天冬酶活化及染色体 DNA 断裂的研究一致，结合 CD20 的蝎形分子似乎未利用经典细胞凋亡途径(诸如半胱天冬酶依赖性途径)直接杀死 B-细胞。尽管不希望受理论束缚，但预期结合 CD20 的蝎形分子是经由半胱天冬酶独立途径和 SYK 独立途径直接杀死 B-细胞，该途径不具有显著的染色体 DNA 断裂特征，至少在半胱天冬酶依赖性细胞凋亡期间发生断裂的相同时间段内不具有显著的染色体 DNA 断裂特征。

实施例 19

蝎形分子应用

a. 蝎形分子的体内活性

蝎形分子的活性亦利用小鼠模型来评估。对于蝎形分子体内活性的测量涉及给予 10-300 μg 蝎形分子且随后依时间次序测定此蝎形分子的血清浓度。所述研究的结果显示于图 40 中，其显示为对小鼠三周药物动力学研究的两种双特异性蝎形分子(亦即 S0033，CD20×CD27 蝎形分子和 CD20×CD37 蝎形分子)中每一者的血清浓度曲线。图 40 中的数据显示，给药后，在该两种双特异性蝎形分子中每一者的血清

含量回落至基线含量之前，需要耗费至少 500 小时。因此，蝎形分子显示血清稳定性和可再生性、持久的体内循环半衰期。

亦评估蝎形分子的体内功效。在 SMIP 对比历史性免疫球蛋白对照物的平行实验中使用侵袭性 Ramos 异种移植模型。图 41 中所提供的存活曲线揭示给予 10 μg 双特异性蝎形分子对存活具有可忽略不计的影响，但 100-300 μg 的给药对携有 Ramos 异种移植体的小鼠的存活具有重要的正面效应。

b. 组合疗法

预期蝎形分子将适用于预防、治疗或改善影响人类、其他哺乳动物及其他有机体的多种状态的症状。例如，预期结合 CD20 的蝎形分子可适用于治疗或预防与 B-细胞过量或异常相关的多种疾病。实际上，易经受治疗(涉及耗尽 B-细胞)的任何疾病均可经受结合 CD20 的蝎形分子的治疗。此外，蝎形分子(例如结合 CD20 的蝎形分子)可与其他治疗剂一起用于组合疗法中。为说明各种组合疗法的可行性，将单特异性 CD20×CD20 蝎形分子(S0129)与羟道诺红霉素(doxorubicin)、长春新碱(vincristine)或雷帕霉素(rapamycin)组合给予 Su-DHL-6 B-细胞。羟道诺红霉素为拓扑异构酶 II 型毒剂，其干扰 DNA 生物化学机制且属于针对抗癌治疗的药物类别。雷帕霉素(Sirolimus)为大环内酯抗生素，其抑制蛋白合成的起始且抑制免疫系统，适用于器官移植且可作为抗增殖剂配合冠状动脉血管内支架使用以抑制或预防再狭窄。长春新碱为抑制小管形成且已用于治疗癌症的长春花生物碱。

图 61 中所示的实验结果是显示为各组合在效应程度范围内的组合指数值。单特异性 CD20×CD20 蝎形分子 S0129 与各类药物的相互作用各不相同，但与 Rituxan®(RTXN)的曲线形式类似。由高浓度的羟道诺红霉素所见的效应可反映趋向单价结合的转变。所述数据说明，结合 CD20 的蝎形分子可与其他多种治疗剂组合使用且所述组合对于本领域技术人员而言在了解本公开后将显而易见。

具有效应功能的多价结合分子或蝎形分子的结构主题的变化对于本领域技术人员而言在回顾本公开后将显而易见，且所述变化结构属于本发明的范畴。

序 列 表

<110> 汤普森等

<120> 具有效应功能的单链多价结合蛋白

<130> 30906/41393C

<160> 379

<170> PatentIn version 3.3

<210> 1
<211> 795
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 抗-CD20 (2H7) LH (DNA)

<400> 1
atggattttc aagtgcagat ttcagcttc ctgctaatca gtgcttcagt cataatgtcc 60

agaggacaaa ttgttctctc ccagtctcca gcaatcctgt ctgcatctcc aggggagaag 120

gtcacaatga cttgcagggc cagctcaagt gtaagttaca tgcactggta ccagcagaag 180

ccaggatcct cccccaaacc ctggatttat gccccatcca acctggcttc tggagtccct 240

gctcgcttca gtggcagtgg gtctgggacc tcttactctc tcacaatcag cagagtggag 300

gctgaagatg ctgccactta ttactgccag cagtggagtt ttaacccacc cacgttcggt 360

gctgggacca agctggagct gaaagatggc ggtggctcgg gcggtggtgg atctggagga 420

ggtgggagct ctcaggctta tctacagcag tctggggctg agtcggtgag gcctggggcc 480

tcagtgaaga tgtcctgcaa ggcttctggc tacacattta ccagttacaa tatgcactgg 540

gtaaagcaga cacctagaca gggcctggaa tggattggag ctatttatcc aggaaatggt 600

gatacttcct acaatcagaa gttcaagggc aaggccacac tgactgtaga caaatcctcc 660

agcacagcct acatgcagct cagcagcctg acatctgaag actctgcggt ctatttctgt 720

gcaagagtgg tgtactatag taactcttac tggtacttcg atgtctgggg cacagggacc 780

acggtcaccg tctct 795

<210> 2
<211> 265
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 抗-CD20 (2H7) LH (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<400> 2

```
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1               5                  10                  15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
            20                  25                  30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
        35                  40                  45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
    50                  55                  60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220
```

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225 230 235 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
245 250 255

Gly Thr Gly Thr Thr Val Thr Val Ser
260 265

<210> 3
<211> 813
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 抗-CD28 (2e12) LH (DNA)

<400> 3
atggattttc aagtgcagat tttcagcttc ctgctaatca gtgcttcagt cataatgtcc 60

agaggagtcg acattgtgct cacccaatct ccagcttctt tggctgtgtc tctaggtcag 120

agagccacca tctcctgcag agccagtgaa agtgttgaat attatgtcac aagtttaatg 180

cagtggtacc aacagaaacc aggacagcca cccaaactcc tcatctctgc tgctagcaac 240

gtagaatctg gggtccctgc caggtttagt ggcagtgggt ctgggacaga ctttagcctc 300

aacatccatc ctgtggagga ggatgatatt gcaatgtatt tctgtcagca aagtaggaag 360

gttccatgga cgttcggtgg aggcaccaag ctggaaatca aacggggtgg cggtggatcc 420

ggcggaggtg ggtcgggtgg cggcggatct caggtgcagc tgaaggagtc aggacctggc 480

ctggtggcgc cctcacagag cctgtccatc acatgcaccg tctcagggtt ctcattaacc 540

ggctatggtg taaactgggt tcgccagcct ccaggaaagg gtctggagtg gctgggaatg 600

atatggggtg atggaagcac agactataat tcagctctca aatccagact atcgatcacc 660

aaggacaact ccaagagcca agttttctta aaaatgaaca gtctgcaaac tgatgacaca 720

gccagatact actgtgcccg agatggttat agtaactttc attactatgt tatggactac 780

tggggtcaag gaacctcagt caccgtctcc tct 813

<210> 4
<211> 271
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 抗-CD28 (2e12) LH (AA)

<220>
<221> misc_feature
<222> (1)..(23)
<223> 前导序列

<220>
<221> misc_feature
<222> (24)..(135)
<223> VL

<220>
<221> misc_feature
<222> (136)..(150)
<223> 连接子

<220>
<221> misc_feature
<222> (151)..(271)
<223> VH

<400> 4

```
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1               5                  10                  15

Val Ile Met Ser Arg Gly Val Asp Ile Val Leu Thr Gln Ser Pro Ala
            20                  25                  30

Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala
        35                  40                  45

Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln
    50                  55                  60

Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn
65                  70                  75                  80

Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr
                85                  90                  95

Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met
            100                 105                 110

Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly
        115                 120                 125

Thr Lys Leu Glu Ile Lys Arg Gly Gly Gly Gly Ser Gly Gly Gly Gly
    130                 135                 140

Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly
145                 150                 155                 160

Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly
                165                 170                 175

Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly
            180                 185                 190
```

Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp
195 200 205

Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser
210 215 220

Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr
225 230 235 240

Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr
245 250 255

Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
260 265 270

<210> 5
<211> 828
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 抗-CD28 (2e12) HL (DNA)

<400> 5
atggattttc aagtgcagat tttcagcttc ctgctaatca gtgcttcagt cataatgtcc 60

agaggagtcc aggtgcagct gaaggagtca ggacctggcc tggtggcgcc ctcacagagc 120

ctgtccatca catgcaccgt ctcagggttc tcattaaccg gctatggtgt aaactgggtt 180

cgccagcctc caggaaaggg tctggagtgg ctgggaatga tatggggtga tggaagcaca 240

gactataatt cagctctcaa atccagacta tcgatcacca aggacaactc caagagccaa 300

gttttcttaa aaatgaacag tctgcaaact gatgacacag ccagatacta ctgtgcccga 360

gatggttata gtaactttca ttactatgtt atggactact ggggtcaagg aacctcagtc 420

accgtctcct ctgggggtgg aggctctggt ggcggtggat ccggcggagg tgggtcgggt 480

ggcggcggat ctgacattgt gctcacccaa tctccagctt ctttggctgt gtctctaggt 540

cagagagcca ccatctcctg cagagccagt gaaagtgttg aatattatgt cacaagttta 600

atgcagtggt accaacagaa accaggacag ccacccaaac tcctcatctc tgctgctagc 660

aacgtagaat ctggggtccc tgccaggttt agtggcagtg ggtctgggac agactttagc 720

ctcaacatcc atcctgtgga ggaggatgat attgcaatgt atttctgtca gcaaagtagg 780

aaggttccat ggacgttcgg tggaggcacc aagctggaaa tcaaacgt 828

<210> 6

<211> 276
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 抗-CD28 (2e12) HL (AA)

<220>
<221> misc_feature
<222> (1)..(23)
<223> 前导序列

<220>
<221> misc_feature
<222> (24)..(144)
<223> VH

<220>
<221> misc_feature
<222> (145)..(164)
<223> 连接子

<220>
<221> misc_feature
<222> (165)..(276)
<223> VL

<400> 6

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Val Gln Val Gln Leu Lys Glu Ser Gly Pro
20 25 30

Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser
35 40 45

Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro
50 55 60

Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr
65 70 75 80

Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn
85 90 95

Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp
100 105 110

Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr
115 120 125

```
Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
    130                 135                 140

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
145                 150                 155                 160

Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala
                165                 170                 175

Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser
            180                 185                 190

Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro
        195                 200                 205

Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser
    210                 215                 220

Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser
225                 230                 235                 240

Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys
                245                 250                 255

Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu
            260                 265                 270

Glu Ile Lys Arg
        275
```

<210> 7
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 重链GSP1

<400> 7
aggtgctgga ggggacagtc actgagctgc 30

<210> 8
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>

<223> 嵌套重链

<400> 8
gtcacwgtca ctgrctcagg gaartagc 28

<210> 9
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 轻链 GSP1

<400> 9
gggtgctgct catgctgtag gtgctgtctt tgc 33

<210> 10
<211> 32
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 嵌套轻链

<400> 10
caagaagcac acgactgagg cacctccaga tg 32

<210> 11
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> Race 简化式锚定引物

<220>
<221> misc_feature
<222> (24)..(25)
<223> n 是 a, c, g 或 t

<220>
<221> misc_feature
<222> (29)..(30)
<223> n 是 a, c, g 或 t

<220>
<221> misc_feature
<222> (34)..(35)
<223> n 是 a, c, g 或 t

<400> 11
ggccacgcgt cgactagtac gggnngggnn gggnng 36

<210> 12
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> T7 引物

<400> 12
gtaatacgac tcactatagg 20

<210> 13
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> M13 反向引物

<400> 13
caggaaacag ctatgacc 18

<210> 14
<211> 15
<212> PRT
<213> 智人

<220>
<223> 人 IgG1 铰链结构域

<400> 14

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10 15

<210> 15
<211> 46
<212> DNA
<213> 智人

<220>
<223> hVK3L-F3H3

<400> 15
gcgataaagc ttgccgccat ggaagcacca gcgcagcttc tcttcc 46

<210> 16
<211> 50
<212> DNA
<213> 智人

<220>

<223> hVK3LF2

<400> 16
accagcgcag cttctcttcc tcctgctact ctggctccca gataccaccg 50

<210> 17
<211> 45
<212> DNA
<213> 智人

<220>
<223> hVK3LF12HVL

<400> 17
ggctcccaga taccaccggt caaattgttc tctcccagtc tccag 45

<210> 18
<211> 35
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 2H7VHNheF

<400> 18
gcgatagcta gccaggctta tctacagcag tctgg 35

<210> 19
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G4SNheR

<400> 19
gcgatagcta gccccacctc ctccagatcc accaccgccc gag 43

<210> 20
<211> 35
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 015VHXhoR

<400> 20
gcgtactcga ggagacggtg accgtggtcc ctgtg 35

<210> 21
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G1HC-Xho

<400> 21
gcagtctcga gcgagcccaa atcttgtgac aaaactc 37

<210> 22
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G1H-S-Xho

<400> 22
gcagtctcga gcgagcccaa atcttctgac aaaactc 37

<210> 23
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> CH3REcoR1

<400> 23
gcgtgagaat tcttacccgg agacagggag aggctc 36

<210> 24
<211> 32
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G1 XBZ-R

<400> 24
gcgacgtcta gagtcattta cccggagaca gg 32

<210> 25
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G4SLinkR1-S

<400> 25
aattatggtg gcggtggctc gggcggtggt ggatctggag gaggtgggag tggg 54

<210> 26
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G4SLinkR1-AS

<400> 26
aattcccact cccacctcct ccagatccac caccgcccga gccaccgcca ccat 54

<210> 27
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 2E12VLXbaR

<400> 27
gcgtgtctag attaacgttt gatttccagc ttggtg 36

<210> 28
<211> 35
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 2E12VLR1F

<400> 28
gcgatgaatt ctgacattgt gctcacccaa tctcc 35

<210> 29
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 2E12VHR1F

<400> 29
gcgatgaatt ctcaggtgca gctgaaggag tcag 34

<210> 30
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 2E12VHXbaR

<400> 30
gcgagtctag attaagagga gacggtgact gaggttc 37

<210> 31
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 2e12VHdXbaF1

<400> 31
gggtctggag tggctgggaa tgatatg 27

<210> 32
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 2e12VHdXbaR1

<400> 32
attcccagcc actccagacc ctttcctg 28

<210> 33
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> IgBsrG1F

<400> 33

gagaaccaca ggtgtacacc ctg 23

<210> 34
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> IgBsrG1R

<400> 34
gcagggtgta cacctgtggt tctcg 25

<210> 35
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> M13R

<400> 35
caggaaacag ctatgac 17

<210> 36
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> M13F

<400> 36
gtaaaacgac ggccagtg 18

<210> 37
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> T7

<400> 37
gtaatacgac tcactatagg 20

<210> 38

<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> pD18F-17

<400> 38
aactagagaa cccactg 17

<210> 39
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> pD18F-20

<400> 39
gctaactaga gaacccactg 20

<210> 40
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> pD18F-1

<400> 40
atacgactca ctataggg 18

<210> 41
<211> 19
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> pD18R-s

<400> 41
gctctagcat ttaggtgac 19

<210> 42
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> CH3seqF1

<400> 42
catgaggctc tgcacaac 18

<210> 43
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> CH3seqF2

<400> 43
cctctacagc aagctcac 18

<210> 44
<211> 19
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> CH3seqR1

<400> 44
ggttcttggt cagctcatc 19

<210> 45
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> CH3seqR2

<400> 45
gtgagcttgc tgtagagg 18

<210> 46
<211> 51
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>

<223> L1-11R

<400> 46
gcgatagaat tcccagatcc accaccgccc gagccaccgc caccataatt c 51

<210> 47
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> L1-6R

<400> 47
gcgatagaat tcccagagcc accgccacca taattc 36

<210> 48
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> L3R

<400> 48
gcgtatgaat tccccgagcc accgccaccc ttacccggag acagg 45

<210> 49
<211> 49
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> L4R

<400> 49
gcgtatgaat tcccagatcc accaccgccc gagccaccgc cacccttac 49

<210> 50
<211> 44
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> L5R

<400> 50
gcgtatgaat tcccgctgcc tcctcccccа gatccaccac cgcc 44

<210> 51
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> IgBsrG1F

<400> 51
gagaaccaca ggtgtacacc ctg 23

<210> 52
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> L-CPPCPR

<400> 52
gcgatagaat tcggacaagg tggacacccc ttacccggag acagggagag 50

<210> 53
<211> 41
<212> DNA
<213> 人工序列

<220>
<223> G281LH-NheR

<220>
<223> 合成引物

<400> 53
actgctgcag ctggaccgcg ctagctccgc cgccacccga c 41

<210> 54
<211> 40
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281LH-NheF

<400> 54
ggcggagcta gcgcggtcca gctgcagcag tctggacctg 40

<210> 55
<211> 37

<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-LH-LPinF

<400> 55
gcgatcaccg gtgacatcca gatgactcag tctccag 37

<210> 56
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-LH-HXhoR

<400> 56
gcgatactcg aggagacggt gactgaggtt ccttgac 37

<210> 57
<211> 38
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-LH-LEcoF

<400> 57
gcgatcgaat tcagacatcc agatgactca gtctccag 38

<210> 58
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-LH-HXbaR

<400> 58
gcgattctag attaggaaga gacggtgact gaggttcctt gac 43

<210> 59
<211> 37
<212> DNA
<213> 人工序列

<220>

<223> 合成引物

<220>
<223> G281-HL-HF

<400> 59
gcgataaccg gtgcggtcca gctgcagcag tctggac 37

<210> 60
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-HL-HR3

<400> 60
gacccaccac cgcccgagcc accgccacca gaagagacgg tgactgaggt tc 52

<210> 61
<211> 49
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-HL-HR2

<400> 61
actcccgcct cctcctgatc cgccgccacc cgacccacca ccgcccgag 49

<210> 62
<211> 42
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-HL-HNheR

<400> 62
gagtcatctg gatgtcgcta gcactcccgc ctcctcctga tc 42

<210> 63
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-HL-LNheF

<400> 63
atcaggagga ggcgggagtg ctagcgacat ccagatgact cagtc 45

<210> 64
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-HL-LXhoR

<400> 64
gcgatactcg agcctttgat ctccagttcg gtgcctc 37

<210> 65
<211> 42
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-HL-LXhoR

<400> 65
gcgatatcta gactcaacct ttgatctcca gttcggtgcc tc 42

<210> 66
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> G281-HL-EcoF

<400> 66
gcgatagaat tcgcggtcca gctgcagcag tctggac 37

<210> 67
<211> 40
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LH-LF1

<400> 67
gcgtatgaac cggtgacatc cagatgacac agactacatc 40

<210> 68
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LF2

<400> 68
atccagatga cacagactac atcctccctg tctgcctctc tgggagacag 50

<210> 69
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LF3

<400> 69
gtctgcctct ctgggagaca gagtcaccat cagttgcagg gcaagtcagg ac 52

<210> 70
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LF4

<400> 70
gttgcagggc aagtcaggac attcgcaatt atttaaactg gtatcagcag 50

<210> 71
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LF5

<400> 71
atttaaactg gtatcagcag aaaccagatg gaactgttaa actcctgatc 50

<210> 72
<211> 50

<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LF6

<400> 72
gaactgttaa actcctgatc tactacacat caagattaca ctcaggagtc 50

<210> 73
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LF7

<400> 73
caagattaca ctcaggagtc ccatcaaggt tcagtggcag tgggtctgga ac 52

<210> 74
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LR7

<400> 74
caggttggca atggtgagag aataatctgt tccagaccca ctgccactga ac 52

<210> 75
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LR6

<400> 75
gcaaaagtaa gtggcaatat cttctggttg caggttggca atggtgagag 50

<210> 76
<211> 51
<212> DNA
<213> 人工序列

<220>

<223> 合成引物

<220>
<223> 194-LR5

<400> 76
gaacgtccac ggaagcgtat taccctgttg gcaaaagtaa gtggcaatat c 51

<210> 77
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LR4

<400> 77
cgtttggtta ccagtttggt gcctccaccg aacgtccacg gaagcgtatt ac 52

<210> 78
<211> 46
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LR3

<400> 78
accaccgccc gagccaccgc cacccgttt ggttaccagt ttggtg 46

<210> 79
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LR2

<400> 79
gctagcgctc ccacctcctc cagatccacc accgcccgag ccaccgccac 50

<210> 80
<211> 42
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LH-LR1

<400> 80
gttgcagctg gacctcgcta gcgctcccac ctcctccaga tc 42

<210> 81
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LH-HF1

<400> 81
gatctggagg aggtgggagc gctagcgagg tccagctgca acagtctgga cctg 54

<210> 82
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HF2

<400> 82
agctgcaaca gtctggacct gaactggtga agcctggagc ttcaatgaag 50

<210> 83
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HF3

<400> 83
agcctggagc ttcaatgaag atttcctgca aggcctctgg ttactcattc 50

<210> 84
<211> 53
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HF4

<400> 84
gcaaggcctc tggttactca ttcactggct acatcgtgaa ctggctgaag cag 53

<210> 85
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HF5

<400> 85
atcgtgaact ggctgaagca gagccatgga aagaaccttg agtggattgg ac 52

<210> 86
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HF6

<400> 86
gaaccttgag tggattggac ttattaatcc atacaaaggt cttactacct ac 52

<210> 87
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HR6

<400> 87
aatgtggcct tgcccttgaa tttctggttg taggtagtaa gacctttgta tg 52

<210> 88
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HR5

<400> 88
catgtaggct gtgctggatg acttgtctac agttaatgtg gccttgccct tg 52

<210> 89
<211> 55
<212> DNA

<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HR4

<400> 89
actgcagagt cttcagatgt cagactgagg agctccatgt aggctgtgct ggatg 55

<210> 90
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HR3

<400> 90
accatagtac ccagatcttg cacagtaata gactgcagag tcttcagatg tc 52

<210> 91
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HR2

<400> 91
gcgccccaga catcgaagta ccagtccgag tcaccatagt acccagatct tg 52

<210> 92
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LH-HR1

<400> 92
gcgaatactc gaggagacgg tgaccgtggt ccctgcgccc cagacatcga ag 52

<210> 93
<211> 42
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HL-HF1

<400> 93
gcgtatgaac cggtgaggtc cagctgcaac agtctggacc tg 42

<210> 94
<211> 55
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HL-HR1

<400> 94
accgccacca gaggagacgg tgaccgtggt ccctgcgccc cagacatcga agtac 55

<210> 95
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HL-HR0

<400> 95
acctcctcca gatccaccac cgcccgagcc accgccacca gaggagacgg tg 52

<210> 96
<211> 49
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HL-LF1

<400> 96
gcgggggagg tggcagtgct agcgacatcc agatgacaca gactacatc 49

<210> 97
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HL-LR3Xho

<400> 97
gcgaatactc gagcgtttgg ttaccagttt ggtg 34

<210> 98
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HL-LR3Xba

<400> 98
gcgatatcta gattaccgtt tggttaccag tttggtg 37

<210> 99
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-HL-HF1R1

<400> 99
gcgtatgaga attcagaggt ccagctgcaa cagtctggac ctg 43

<210> 100
<211> 41
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LH-LF1R1

<400> 100
gcgtatgaga attctgacat ccagatgaca cagactacat c 41

<210> 101
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 194-LH-HR1Xba

<400> 101
gcgtatctag attaggagac ggtgaccgtg gtccctgcgc cccagacatc gaag 54

<210> 102
<211> 725
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> G28-1 VLVH (DNA)

<400> 102
accggtgaca tccagatgac tcagtctcca gcctccctat ctgcatctgt gggagagact 60

gtcaccatca catgtcgaac aagtgaaaat gtttacagtt atttggcttg gtatcagcag 120

aaacagggaa aatctcctca gctcctggtc tcttttgcaa aaaccttagc agaaggtgtg 180

ccatcaaggt tcagtggcag tggatcaggc acacagtttt ctctgaagat cagcagcctg 240

cagcctgaag attctggaag ttatttctgt caacatcatt ccgataatcc gtggacgttc 300

ggtggaggca ccgaactgga gatcaaaggt ggcggtggct cgggcggtgg tgggtcgggt 360

ggcggcggat ctgctagcgc agtccagctg cagcagtctg gacctgagct ggaaaagcct 420

ggcgcttcag tgaagatttc ctgcaaggct tctggttact cattcactgg ctacaatatg 480

aactgggtga agcagaataa tggaaagagc cttgagtgga ttggaaatat tgatccttat 540

tatggtggta ctacctacaa ccggaagttc aagggcaagg ccacattgac tgtagacaaa 600

tcctccagca cagcctacat gcagctcaag agtctgacat ctgaggactc tgcagtctat 660

tactgtgcaa gatcggtcgg ccctatggac tactggggtc aaggaacctc agtcaccgtc 720

tcgag 725

<210> 103
<211> 239
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G-28-1 VLVH (AA)

<220>
<221> misc_feature
<222> (1)..(107)
<223> VL

<220>
<221> misc_feature
<222> (108)..(124)
<223> 连接子

<220>
<221> misc_feature
<222> (125)..(239)
<223> VH

<400> 103

Asp Ile Gln Met Thr Gln Ser Pro Ala Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Glu Thr Val Thr Ile Thr Cys Arg Thr Ser Glu Asn Val Tyr Ser Tyr
20 25 30

Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys Ser Pro Gln Leu Leu Val
35 40 45

Ser Phe Ala Lys Thr Leu Ala Glu Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Gln Phe Ser Leu Lys Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Ser Gly Ser Tyr Phe Cys Gln His His Ser Asp Asn Pro Trp
85 90 95

Thr Phe Gly Gly Gly Thr Glu Leu Glu Ile Lys Gly Gly Gly Gly Ser
100 105 110

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser Ala Val Gln Leu
115 120 125

Gln Gln Ser Gly Pro Glu Leu Glu Lys Pro Gly Ala Ser Val Lys Ile
130 135 140

Ser Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr Asn Met Asn Trp
145 150 155 160

Val Lys Gln Asn Asn Gly Lys Ser Leu Glu Trp Ile Gly Asn Ile Asp
165 170 175

Pro Tyr Tyr Gly Gly Thr Thr Tyr Asn Arg Lys Phe Lys Gly Lys Ala
180 185 190

Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln Leu Lys
195 200 205

Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg Ser Val
210 215 220

Gly Pro Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser
225 230 235

<210> 104
<211> 767

<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> G28-1 VHVL (DNA)

<400> 104
accggtgagg tccagctgca acagtctgga cctgaactgg tgaagcctgg agcttcaatg 60

aagatttcct gcaaggcctc tggttactca ttcactggct acatcgtgaa ctggctgaag 120

cagagccatg gaaagaacct tgagtggatt ggacttatta atccatacaa aggtcttact 180

acctacaacc agaaattcaa gggcaaggcc acattaactg tagacaagtc atccagcaca 240

gcctacatgg agctcctcag tctgacatct gaagactctg cagtctatta ctgtgcaaga 300

tctgggtact atggtgactc ggactggtac ttcgatgtct ggggcgcagg gaccacggtc 360

accgtctcct ctggtggcgg tggctcgggc ggtggtggat ctggaggagg tgggagcggg 420

ggaggtggca gtgctagcga catccagatg acacagacta catcctccct gtctgcctct 480

ctgggagaca gagtcaccat cagttgcagg gcaagtcagg acattcgcaa ttatttaaac 540

tggtatcagc agaaaccaga tggaactgtt aaactcctga tctactacac atcaagatta 600

cactcaggag tcccatcaag gttcagtggc agtgggtctg gaacagatta ttctctcacc 660

attgccaacc tgcaaccaga agatattgcc acttactttt gccaacaggg taatacgctt 720

ccgtggacgt tcggtggagg caccaaactg gtaaccaaac gctcgag 767

<210> 105
<211> 253
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G28-1 VHVL (AA)

<220>
<221> misc_feature
<222> (1)..(121)
<223> VH

<220>
<221> misc_feature
<222> (122)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(253)
<223> VL

<400> 105

Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala
1 5 10 15

Ser Met Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr
20 25 30

Ile Val Asn Trp Leu Lys Gln Ser His Gly Lys Asn Leu Glu Trp Ile
35 40 45

Gly Leu Ile Asn Pro Tyr Lys Gly Leu Thr Thr Tyr Asn Gln Lys Phe
50 55 60

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Leu Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Gly Tyr Tyr Gly Asp Ser Asp Trp Tyr Phe Asp Val Trp
100 105 110

Gly Ala Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly
115 120 125

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser
130 135 140

Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly
145 150 155 160

Asp Arg Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Arg Asn Tyr
165 170 175

Leu Asn Trp Tyr Gln Gln Lys Pro Asp Gly Thr Val Lys Leu Leu Ile
180 185 190

Tyr Tyr Thr Ser Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly
195 200 205

Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ala Asn Leu Gln Pro
210 215 220

Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly Asn Thr Leu Pro Trp
225 230 235 240

Thr Phe Gly Gly Gly Thr Lys Leu Val Thr Lys Arg Ser
245 250

<210> 106
<211> 749
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> G19-4 VLVH (DNA)

<400> 106
accggtgaca tccagatgac acagactaca tcctccctgt ctgcctctct gggagacaga 60

gtcaccatca gttgcagggc aagtcaggac attcgcaatt atttaaactg gtatcagcag 120

aaaccagatg gaactgttaa actcctgatc tactacacat caagattaca ctcaggagtc 180

ccatcaaggt tcagtggcag tgggtctgga acagattatt ctctcaccat tgccaacctg 240

caaccagaag atattgccac ttacttttgc caacagggta tacgcttcc gtggacgttc 300

ggtggaggca ccaaactggt aaccaaacgg ggtggcggtg gctcgggcgg tggtggatct 360

ggaggaggtg ggagcgctag cgaggtccag ctgcaacagt ctggacctga actggtgaag 420

cctggagctt caatgaagat ttcctgcaag gcctctggtt actcattcac tggctacatc 480

gtgaactggc tgaagcagag ccatggaaag aaccttgagt ggattggact tattaatcca 540

tacaaaggtc ttactaccta caaccagaaa ttcaagggca aggccacatt aactgtagac 600

aagtcatcca gcacagccta catggagctc ctcagtctga catctgaaga ctctgcagtc 660

tattactgtg caagatctgg gtactatggt gactcggact ggtacttcga tgtctggggc 720

gcagggacca cggtcaccgt ctcctcgag 749

<210> 107
<211> 247
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G-19-4 VLVH (AA)

<220>
<221> misc_feature
<222> (1)..(108)
<223> VL

<220>
<221> misc_feature
<222> (109)..(125)
<223> 连接子

<220>
<221> misc_feature

<222> (126)..(247)
<223> VH

<400> 107

Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly
1 5 10 15

Asp Arg Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Arg Asn Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Asp Gly Thr Val Lys Leu Leu Ile
35 40 45

Tyr Tyr Thr Ser Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ala Asn Leu Gln Pro
65 70 75 80

Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly Asn Thr Leu Pro Trp
85 90 95

Thr Phe Gly Gly Gly Thr Lys Leu Val Thr Lys Arg Gly Gly Gly Gly
100 105 110

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser Glu Val Gln
115 120 125

Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala Ser Met Lys
130 135 140

Ile Ser Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr Ile Val Asn
145 150 155 160

Trp Leu Lys Gln Ser His Gly Lys Asn Leu Glu Trp Ile Gly Leu Ile
165 170 175

Asn Pro Tyr Lys Gly Leu Thr Thr Tyr Asn Gln Lys Phe Lys Gly Lys
180 185 190

Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Glu Leu
195 200 205

Leu Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg Ser
210 215 220

Gly Tyr Tyr Gly Asp Ser Asp Trp Tyr Phe Asp Val Trp Gly Ala Gly
225 230 235 240

Thr Thr Val Thr Val Ser Ser
245

<210> 108
<211> 752
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> G19-4 VHVL (DNA)

<400> 108
accggtgagg tccagctgca acagtctgga cctgaactgg tgaagcctgg agcttcaatg 60

aagatttcct gcaaggcctc tggttactca ttcactggct acatcgtgaa ctggctgaag 120

cagagccatg gaaagaacct tgagtggatt ggacttatta atccatacaa aggtcttact 180

acctacaacc agaaattcaa gggcaaggcc acattaactg tagacaagtc atccagcaca 240

gcctacatgg agctcctcag tctgacatct gaagactctg cagtctatta ctgtgcaaga 300

tctgggtact atggtgactc ggactggtac ttcgatgtct ggggcgcagg gaccacggtc 360

accgtctcct ctggtggcgg tggctcgggc ggtggtggat ctggaggagg tgggagcgct 420

agcgacatcc agatgacaca gactacatcc tccctgtctg cctctctggg agacagagtc 480

accatcagtt gcagggcaag tcaggacatt cgcaattatt taaactggta tcagcagaaa 540

ccagatggaa ctgttaaact cctgatctac tacacatcaa gattacactc aggagtccca 600

tcaaggttca gtggcagtgg gtctggaaca gattattctc tcaccattgc caacctgcaa 660

ccagaagata ttgccactta cttttgccaa cagggtaata cgcttccgtg gacgttcggt 720

ggaggcacca aactggtaac caaacgctcg ag 752

<210> 109
<211> 248
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G19-4 VHVL (AA)

<220>
<221> misc_feature
<222> (1)..(122)
<223> VH

<220>
<221> misc_feature
<222> (123)..(139)

<223> 连接子

<220>
<221> misc_feature
<222> (140)..(248)
<223> VL

<400> 109

Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala
1 5 10 15

Ser Met Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr
20 25 30

Ile Val Asn Trp Leu Lys Gln Ser His Gly Lys Asn Leu Glu Trp Ile
35 40 45

Gly Leu Ile Asn Pro Tyr Lys Gly Leu Thr Thr Tyr Asn Gln Lys Phe
50 55 60

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Leu Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Gly Tyr Tyr Gly Asp Ser Asp Trp Tyr Phe Asp Val Trp
100 105 110

Gly Ala Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly
115 120 125

Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser Asp Ile Gln Met Thr
130 135 140

Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly Asp Arg Val Thr Ile
145 150 155 160

Ser Cys Arg Ala Ser Gln Asp Ile Arg Asn Tyr Leu Asn Trp Tyr Gln
165 170 175

Gln Lys Pro Asp Gly Thr Val Lys Leu Leu Ile Tyr Tyr Thr Ser Arg
180 185 190

Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr
195 200 205

Asp Tyr Ser Leu Thr Ile Ala Asn Leu Gln Pro Glu Asp Ile Ala Thr
210 215 220

Tyr Phe Cys Gln Gln Gly Asn Thr Leu Pro Trp Thr Phe Gly Gly Gly
225 230 235 240

Thr Lys Leu Val Thr Lys Arg Ser
245

<210> 110
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> ssc(s)-hIgG1 (DNA)

<400> 110
gagcccaaat cttctgacaa aactcacaca tctccaccgt gctca 45

<210> 111
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> ssc(s)-hIgG1 (AA)

<400> 111

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Ser Pro Pro Cys Ser
1 5 10 15

<210> 112
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> scc(s)-hIgG1 (DNA)

<400> 112
gagcccaaat cttctgacaa aactcacaca tgtccaccgt gctca 45

<210> 113
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> scc(s)-hIgG1 (AA)

<400> 113

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Ser
1 5 10 15

<210> 114
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> css(s)-hIgG1 (DNA)

<400> 114
gagcccaaat cttgtgacaa aactcacaca tctccaccga gctca 45

<210> 115
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> css(s)-hIgG1 (AA)

<400> 115

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Ser Pro Pro Ser Ser
1 5 10 15

<210> 116
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<221> misc_feature
<223> scs(s)-hIgG1 (DNA)

<400> 116

gagcccaaat cttgtgacaa aactcacaca tgtccaccga gctca 45

<210> 117
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> scs(s)-hIgG1 (AA)

<400> 117

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Ser Ser
1 5 10 15

<210> 118
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<400> 118
gagcccaaat cttgtgacaa aactcacaca tgtccaccgt gctca 45

<210> 119
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<400> 119

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Ser
1 5 10 15

<210> 120
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<400> 120
gagcccaaat cttgtgacaa aactcacaca tgtccaccgt gccca 45

<210> 121
<211> 15

<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<400> 121

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10 15

<210> 122
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<400> 122
gagcccaaat cttctgacaa aactcacaca tctccaccga gccca 45

<210> 123
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<400> 123

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Ser Pro Pro Ser Pro
1 5 10 15

<210> 124
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> csc(p)-hIgG1 (DNA)

<400> 124
gagcccaaat cttgtgacaa aactcacaca tctccaccgt gccca 45

<210> 125
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> csc(p)-hIgG1 (AA)

<400> 125

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Ser Pro Pro Cys Pro
1 5 10 15

<210> 126
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> ssc(p)-hIgG1 (DNA)

<400> 126
gagcccaaat cttctgacaa aactcacaca tctccaccgt gccca 45

<210> 127
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> ssc(p)-hIgG1 (AA)

<400> 127

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Ser Pro Pro Cys Pro
1 5 10 15

<210> 128
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> scc(p)-hIgG1 (DNA)

<400> 128
gagcccaaat cttctgacaa aactcacaca tgtccaccgt gccca 45

<210> 129
<211> 15

<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> scc(p)-hIgG1 (AA)

<400> 129

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10 15

<210> 130
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> css(p)-hIgG1 (DNA)

<400> 130
gagcccaaat cttgtgacaa aactcacaca tctccaccga gccca 45

<210> 131
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> css(p)-hIgG1 (AA)

<400> 131

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Ser Pro Pro Ser Pro
1 5 10 15

<210> 132
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<221> misc_feature
<223> ssc(p)-hIgG1 (DNA)

<400> 132
gagcccaaat cttgtgacaa aactcacaca tgtccaccga gccca 45

<210> 133
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> ssc(p)-hIgG1 (DNA)

<400> 133

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Ser Pro
1 5 10 15

<210> 134
<211> 18
<212> DNA
<213> 人工序列

<220>
<221> misc_feature
<223> ssc(p) (DNA)

<220>
<223> 合成引物

<400> 134
agttgtccac cgtgccca 18

<210> 135
<211> 6
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<400> 135

Ser Cys Pro Pro Cys Pro
1 5

<210> 136
<211> 7
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<400> 136
tgatcag 7

<210> 137
<211> 2
<212> PRT

<213> 人工序列

<220>
<223> 合成肽

<400> 137

Asp Gln
1

<210> 138
<211> 7
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<400> 138
ctcgagt 7

<210> 139
<211> 2
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<400> 139

Ser Ser
1

<210> 140
<211> 654
<212> DNA
<213> 智人

<220>
<223> hIgG1 野生型

<400> 140
gcacctgaac tcctgggtgg accgtcagtc ttcctcttcc ccccaaaacc caaggacacc 60

ctcatgatct cccggacccc tgaggtcaca tgcgtggtgg tggacgtgag ccacgaagac 120

cctgaggtca agttcaactg gtacgtggac ggcgtggagg tgcataatgc caagacaaag 180

ccgcgggagg agcagtacaa cagcacgtac cgtgtggtca gcgtcctcac cgtcctgcac 240

caggactggc tgaatggcaa ggagtacaag tgcaaggtct ccaacaaagc cctcccagcc 300

cccatcgaga aaacaatctc caaagccaaa gggcagcccc gagaaccaca ggtgtacacc 360

ctgcccccat cccgggatga gctgaccaag aaccaggtca gcctgacctg cctggtcaaa 420

ggcttctatc ccagcgacat cgccgtggag tgggagagca atgggcagcc ggagaacaac 480

tacaagacca cgcctcccgt gctggactcc gacggctcct tcttcctcta cagcaagctc 540

accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag 600

gctctgcaca accactacac gcagaagagc ctctccctgt ctccgggtaa atga 654

<210> 141
<211> 217
<212> PRT
<213> 智人

<220>
<223> hIgG1 野生型

<400> 141

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
20 25 30

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
35 40 45

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
50 55 60

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
65 70 75 80

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
85 90 95

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
100 105 110

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
115 120 125

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
130 135 140

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
145 150 155 160

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
165 170 175

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
180 185 190

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
195 200 205

Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215

<210> 142
<211> 654
<212> DNA
<213> 智人

<220>
<223> hIgG1 (P238S)

<400> 142
gcacctgaac tcctgggtgg atcgtcagtc ttcctcttcc ccccaaaacc caaggacacc 60

ctcatgatct cccggacccc tgaggtcaca tgcgtggtgg tggacgtgag ccacgaagac 120

cctgaggtca agttcaactg gtacgtggac ggcgtggagg tgcataatgc caagacaaag 180

ccgcgggagg agcagtacaa cagcacgtac cgtgtggtca gcgtcctcac cgtcctgcac 240

caggactggc tgaatggcaa ggagtacaag tgcaaggtct ccaacaaagc cctcccagcc 300

cccatcgaga aaacaatctc caaagccaaa gggcagcccc gagaaccaca ggtgtacacc 360

ctgcccccat cccgggatga gctgaccaag aaccaggtca gcctgacctg cctggtcaaa 420

ggcttctatc ccagcgacat cgccgtggag tgggagagca atgggcagcc ggagaacaac 480

tacaagacca cgcctcccgt gctggactcc gacggctcct tcttcctcta cagcaagctc 540

accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag 600

gctctgcaca accactacac gcagaagagc ctctccctgt ctccgggtaa atga 654

<210> 143
<211> 217
<212> PRT
<213> 智人

<220>
<223> hIgG1 (P238S)

<400> 143

Ala Pro Glu Leu Leu Gly Gly Ser Ser Val Phe Leu Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
20 25 30

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
35 40 45

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
50 55 60

```
Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
65                  70                  75                  80

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
                85                  90                  95

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
            100                 105                 110

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
        115                 120                 125

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
    130                 135                 140

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
145                 150                 155                 160

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
                165                 170                 175

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
            180                 185                 190

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
        195                 200                 205

Lys Ser Leu Ser Leu Ser Pro Gly Lys
    210                 215
```

<210> 144
<211> 654
<212> DNA
<213> 智人

<220>
<223> hIgG1 (P331S)

<400> 144

```
gcacctgaac tcctgggtgg accgtcagtc ttcctcttcc ccccaaaacc caaggacacc     60

ctcatgatct cccggacccc tgaggtcaca tgcgtggtgg tggacgtgag ccacgaagac    120

cctgaggtca agttcaactg gtacgtggac ggcgtggagg tgcataatgc caagacaaag    180

ccgcgggagg agcagtacaa cagcacgtac cgtgtggtca gcgtcctcac cgtcctgcac    240

caggactggc tgaatggcaa ggagtacaag tgcaaggtct ccaacaaagc cctcccagcc    300

tccatcgaga aaacaatctc caaagccaaa gggcagcccc gagaaccaca ggtgtacacc    360

ctgcccccat cccgggatga gctgaccaag aaccaggtca gcctgacctg cctggtcaaa    420

ggcttctatc ccagcgacat cgccgtggag tgggagagca atgggcagcc ggagaacaac    480
```

tacaagacca cgcctcccgt gctggactcc gacggctcct tcttcctcta cagcaagctc 540

accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag 600

gctctgcaca accactacac gcagaagagc ctctccctgt ctccgggtaa atga 654

<210> 145
<211> 217
<212> PRT
<213> 智人

<220>
<223> hIgG1 (P331S)

<400> 145

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
20 25 30

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
35 40 45

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
50 55 60

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
65 70 75 80

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
85 90 95

Ala Leu Pro Ala Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
100 105 110

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
115 120 125

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
130 135 140

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
145 150 155 160

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
165 170 175

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
180 185 190

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
195 200 205

Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215

<210> 146
<211> 654
<212> DNA
<213> 智人

<220>
<223> hIgG1 (P238S/P331S)

<400> 146
gcacctgaac tcctgggtgg atcgtcagtc ttcctcttcc ccccaaaacc caaggacacc 60

ctcatgatct cccggacccc tgaggtcaca tgcgtggtgg tggacgtgag ccacgaagac 120

cctgaggtca agttcaactg gtacgtggac ggcgtggagg tgcataatgc caagacaaag 180

ccgcgggagg agcagtacaa cagcacgtac cgtgtggtca gcgtcctcac cgtcctgcac 240

caggactggc tgaatggcaa ggagtacaag tgcaaggtct ccaacaaagc cctcccagcc 300

tccatcgaga aaacaatctc caaagccaaa gggcagcccc gagaaccaca ggtgtacacc 360

ctgcccccat cccgggatga gctgaccaag aaccaggtca gcctgacctg cctggtcaaa 420

ggcttctatc ccagcgacat cgccgtggag tgggagagca atgggcagcc ggagaacaac 480

tacaagacca cgcctcccgt gctggactcc gacggctcct tcttcctcta cagcaagctc 540
accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag 600

gctctgcaca accactacac gcagaagagc ctctccctgt ctccgggtaa atga 654

<210> 147
<211> 217
<212> PRT
<213> 智人

<220>
<223> hIgG1 (P238S/P331S)

<400> 147

Ala Pro Glu Leu Leu Gly Gly Ser Ser Val Phe Leu Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
20 25 30

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
35 40 45

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
50 55 60

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
65 70 75 80

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
85 90 95

Ala Leu Pro Ala Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
100 105 110

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
115 120 125

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
130 135 140

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
145 150 155 160

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
165 170 175

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
180 185 190

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
195 200 205

Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215

<210> 148
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> STD1 (DNA)

<400> 148
aattatggtg gcggtggctc gggcggtggt ggatctggag gaggtgggag tgggaattct 60

<210> 149
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> STD1 (AA)

<400> 149

Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
1 5 10 15

Ser Gly Asn Ser
20

<210> 150
<211> 114
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> STD2 (DNA)

<400> 150
aattatggtg gcggtggctc gggcggtggt ggatctggag gaggtgggag tgggaattat 60

ggtggcggtg gctcgggcgg tggtggatct ggaggaggtg ggagtgggaa ttct 114

<210> 151
<211> 38
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> STD2 (AA)

<400> 151

Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
1 5 10 15

Ser Gly Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
20 25 30

Gly Gly Ser Gly Asn Ser
35

<210> 152
<211> 6
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H1 (PN)

<400> 152
aattct 6

<210> 153
<211> 2
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H1 (AA)

<400> 153

Asn Ser
1

<210> 154
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H2 (PN)

<400> 154
ggtggcggtg gctcggggaa ttct 24

<210> 155
<211> 8
<212> PRT
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H2 (AA)

<400> 155

Gly Gly Gly Gly Ser Gly Asn Ser
1 5

<210> 156
<211> 30
<212> DNA

<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H3 (PN)

<400> 156
aattatggtg gcggtggctc tgggaattct 30

<210> 157
<211> 10
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H3 (AA)

<400> 157

Asn Tyr Gly Gly Gly Gly Ser Gly Asn Ser
1 5 10

<210> 158
<211> 39
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H4 (PN)

<400> 158
ggtggcggtg gctcgggcgg tggtggatct gggaattct 39

<210> 159
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H4 (AA)

<400> 159

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Asn Ser
1 5 10

<210> 160
<211> 45
<212> DNA

<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H5 (PN)

<400> 160
aattatggtg gcggtggctc gggcggtggt ggatctggga attct 45

<210> 161
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H5 (AA)

<400> 161

Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Asn Ser
1 5 10 15

<210> 162
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H6 (PN)

<400> 162
ggtggcggtg gctcgggcgg tggtggatct gggggaggag gcagcgggaa ttct 54

<210> 163
<211> 18
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H6 (AA)

<400> 163

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15

Asn Ser

<210> 164
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子 H7 (PN)

<400> 164
gggtgtccac cttgtccgaa ttct 24

<210> 165
<211> 8
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H7 (AA)

<400> 165

Gly Cys Pro Pro Cys Pro Asn Ser
1 5

<210> 166
<211> 45
<212> DNA
<213> 人工序列
<220>
<223> 合成引物

<220>
<223> 连接子(G4S)3

<400> 166
ggtggcggtg gatccggcgg aggtgggtcg ggtggcggcg gatct 45

<210> 167
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子(G4S)3

<400> 167

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
1 5 10

<210> 168
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<220>
<223> 连接子(G4S)4

<400> 168
ggtggcggtg gctcgggcgg tggtggatct ggaggaggtg ggagcggggg aggtggcagt 60

<210> 169
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子(G4S)4

<400> 169

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly
1 5 10 15

Ser

<210> 170
<211> 2337
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-STD1-2e12HL (DNA)

<400> 170
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagctcagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attatggtgg cggtggctcg ggcggtggtg gatctggagg aggtgggagt 1560

gggaattctc aggtgcagct gaaggagtca ggacctggcc tggtggcgcc ctcacagagc 1620

ctgtccatca catgcaccgt ctcagggttc tcattaaccg gctatggtgt aaactgggtt 1680

cgccagcctc caggaaaggg tctggagtgg ctgggaatga tatggggtga tggaagcaca 1740

gactataatt cagctctcaa atccagacta tcgatcacca aggacaactc caagagccaa 1800

gttttcttaa aaatgaacag tctgcaaact gatgacacag ccagatacta ctgtgctcga 1860

gatggttata gtaactttca ttactatgtt atggactact ggggtcaagg aacctcagtc 1920

accgtctcct ctgggggtgg aggctctggt ggcggtggat ccggcggagg tgggtcgggt 1980

ggcggcggat ctgacattgt gctcacccaa tctccagctt ctttggctgt gtctctaggt 2040

cagagagcca ccatctcctg cagagccagt gaaagtgttg aatattatgt cacaagttta 2100

atgcagtggt accaacagaa accaggacag ccacccaaac tcctcatctc tgctgctagc 2160

aacgtagaat ctggggtccc tgccaggttt agtggcagtg ggtctgggac agactttagc 2220

ctcaacatcc atcctgtgga ggaggatgat attgcaatgt atttctgtca gcaaagtagg 2280

aaggttccat ggacgttcgg tggaggcacc aagctggaaa tcaaacgtta atctaga 2337

<210> 171
<211> 772
<212> PRT

<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-STD1-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(281)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(519)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (520)..(640)
<223> VH2

<220>
<221> misc_feature
<222> (641)..(660)
<223> 连接子 2

<220>
<221> misc_feature
<222> (661)..(772)
<223> VL2

<400> 171

```
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1               5                   10                  15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
            20                  25                  30


Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
        35                  40                  45
```

```
Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
    50                  55                  60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300
```

```
Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro
        515                 520                 525

Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser
    530                 535                 540

Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro
545                 550                 555                 560
```

```
Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr
                565                 570                 575

Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn
            580                 585                 590

Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp
        595                 600                 605

Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr
    610                 615                 620

Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
                645                 650                 655

Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala
            660                 665                 670

Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser
        675                 680                 685

Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro
    690                 695                 700

Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser
705                 710                 715                 720

Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser
                725                 730                 735

Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys
            740                 745                 750

Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu
        755                 760                 765

Glu Ile Lys Arg
    770
```

<210> 172
<211> 772
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1(P238S/P331S)-STD1-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(519)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (520)..(640)
<223> VH2

<220>
<221> misc_feature
<222> (641)..(660)
<223> 连接子 2

<220>
<221> misc_feature
<222> (661)..(772)
<223> VL2

<400> 172

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

```
Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Ser Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320
```

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
325 330 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
340 345 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
355 360 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Ser Ile
370 375 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385 390 395 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
405 410 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
420 425 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
435 440 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
450 455 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465 470 475 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
485 490 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
500 505 510

Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro
515 520 525

Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser
530 535 540

Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro
545 550 555 560

Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr
565 570 575

```
Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn
            580                 585                 590

Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp
        595                 600                 605

Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr
    610                 615                 620

Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
                645                 650                 655

Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala
            660                 665                 670

Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser
        675                 680                 685

Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro
    690                 695                 700

Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser
705                 710                 715                 720

Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser
                725                 730                 735

Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys
            740                 745                 750

Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu
        755                 760                 765

Glu Ile Lys Arg
    770
```

<210> 173
<211> 2322
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-STD1-2e12LH (DNA)

<400> 173
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attatggtgg cggtggctcg ggcggtggtg gatctggagg aggtgggagt 1560

gggaattctg acattgtgct cacccaatct ccagcttctt tggctgtgtc tctaggtcag 1620

agagccacca tctcctgcag agccagtgaa agtgttgaat attatgtcac aagtttaatg 1680

cagtggtacc aacagaaacc aggacagcca cccaaactcc tcatctctgc tgctagcaac 1740

gtagaatctg gggtccctgc caggtttagt ggcagtgggt ctgggacaga ctttagcctc 1800

aacatccatc ctgtggagga ggatgatatt gcaatgtatt tctgtcagca aagtaggaag 1860

gttccatgga cgttcggtgg aggcaccaag ctggaaatca aacggggtgg cggtggatcc 1920

ggcggaggtg ggtcgggtgg cggcggatct caggtgcagc tgaaggagtc aggacctggc 1980

ctggtggcgc cctcacagag cctgtccatc acatgcaccg tctcagggtt ctcattaacc 2040

ggctatggtg taaactgggt tcgccagcct ccaggaaagg gtctggagtg gctgggaatg 2100

atatggggtg atggaagcac agactataat tcagctctca aatccagact atcgatcacc 2160

aaggacaact ccaagagcca agttttctta aaaatgaaca gtctgcaaac tgatgacaca 2220

gccagatact actgtgctcg agatggttat agtaactttc attactatgt tatggactac 2280

tggggtcaag gaacctcagt caccgtctcc tcttaatcta ga 2322

<210> 174
<211> 767
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-STD1-2e12LH (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(519)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (520)..(631)
<223> VL2

<220>
<221> misc_feature

<222> (632)..(646)
<223> 连接子 2

<220>
<221> misc_feature
<222> (647)..(767)
<223> VH2

<400> 174

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145 150 155 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
165 170 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
180 185 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
195 200 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
210 215 220

```
Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480
```

```
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Gly Gly Gly Ser Gly Asn Ser Asp Ile Val Leu Thr Gln Ser Pro Ala
        515                 520                 525

Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala
    530                 535                 540

Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln
545                 550                 555                 560

Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn
                565                 570                 575

Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr
            580                 585                 590

Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met
        595                 600                 605

Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly
    610                 615                 620

Thr Lys Leu Glu Ile Lys Arg Gly Gly Gly Gly Ser Gly Gly Gly Gly
625                 630                 635                 640

Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly
                645                 650                 655

Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly
            660                 665                 670

Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly
        675                 680                 685

Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp
    690                 695                 700

Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser
705                 710                 715                 720

Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr
                725                 730                 735
```

Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr
740 745 750

Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
755 760 765

<210> 175
<211> 767
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1(P238S/P331S)-STD1-2e12LH (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(519)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (520)..(631)
<223> VL2

<220>
<221> misc_feature
<222> (632)..(646)
<223> 连接子 2

<220>
<221> misc_feature
<222> (647)..(767)
<223> VH2

<400> 175

```
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1               5                  10                  15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
            20                  25                  30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
        35                  40                  45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
    50                  55                  60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255
```

```
Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Ser Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Ser Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510
```

```
Gly Gly Gly Ser Gly Asn Ser Asp Ile Val Leu Thr Gln Ser Pro Ala
        515                 520                 525

Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala
    530                 535                 540

Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln
545                 550                 555                 560

Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn
                565                 570                 575

Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr
            580                 585                 590

Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met
        595                 600                 605

Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly
    610                 615                 620

Thr Lys Leu Glu Ile Lys Arg Gly Gly Gly Gly Ser Gly Gly Gly Gly
625                 630                 635                 640

Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly
                645                 650                 655

Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly
            660                 665                 670

Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly
        675                 680                 685

Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp
    690                 695                 700

Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser
705                 710                 715                 720

Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr
                725                 730                 735

Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr
            740                 745                 750

Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
        755                 760                 765
```

<210> 176
<211> 2376
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-STD2-2e12LH (DNA)

<400> 176
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attatggtgg cggtggctcg ggcggtggtg gatctggagg aggtgggagt 1560

gggaattatg gtggcggtgg ctcgggcggt ggtggatctg gaggaggtgg gagtgggaat 1620

tctgacattg tgctcaccca atctccagct tctttggctg tgtctctagg tcagagagcc 1680

accatctcct gcagagccag tgaaagtgtt gaatattatg tcacaagttt aatgcagtgg 1740

taccaacaga aaccaggaca gccacccaaa ctcctcatct ctgctgctag caacgtagaa 1800

tctggggtcc ctgccaggtt tagtggcagt gggtctggga cagactttag cctcaacatc 1860

catcctgtgg aggaggatga tattgcaatg tatttctgtc agcaaagtag gaaggttcca 1920

tggacgttcg gtggaggcac caagctggaa atcaaacggg gtggcggtgg atccggcgga 1980

ggtgggtcgg gtggcggcgg atctcaggtg cagctgaagg agtcaggacc tggcctggtg 2040

gcgccctcac agagcctgtc catcacatgc accgtctcag ggttctcatt aaccggctat 2100

ggtgtaaact gggttcgcca gcctccagga aagggtctgg agtggctggg aatgatatgg 2160

ggtgatggaa gcacagacta taattcagct ctcaaatcca gactatcgat caccaaggac 2220

aactccaaga gccaagtttt cttaaaaatg aacagtctgc aaactgatga cacagccaga 2280

tactactgtg ctcgagatgg ttatagtaac tttcattact atgttatgga ctactggggt 2340

caaggaacct cagtcaccgt ctcctcttaa tctaga 2376

<210> 177
<211> 785
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-STD2-2e12LH (w/2e12 前导序列)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(129)
<223> VL

<220>
<221> misc_feature
<222> (130)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)

<223> 铰链

<220>
<221> misc_feature
<222> (500)..(537)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (538)..(649)
<223> VL2

<220>
<221> misc_feature
<222> (650)..(664)
<223> 连接子 2

<220>
<221> misc_feature
<222> (665)..(785)
<223> VH2

<400> 177

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145 150 155 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
165 170 175

```
Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430
```

```
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Gly Gly Gly Ser Gly Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly
        515                 520                 525

Ser Gly Gly Gly Gly Ser Gly Asn Ser Asp Ile Val Leu Thr Gln Ser
    530                 535                 540

Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys
545                 550                 555                 560

Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp
                565                 570                 575

Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala
            580                 585                 590

Ser Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser
        595                 600                 605

Gly Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile
    610                 615                 620

Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly
625                 630                 635                 640

Gly Gly Thr Lys Leu Glu Ile Lys Arg Gly Gly Gly Gly Ser Gly Gly
                645                 650                 655

Gly Gly Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Lys Glu Ser Gly
            660                 665                 670

Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val
        675                 680                 685
```

Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro
690 695 700

Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser
705 710 715 720

Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp
725 730 735

Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp
740 745 750

Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His
755 760 765

Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser
770 775 780

Ser
785

<210> 178
<211> 785
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1(P238S/P331S)-STD2-2e12LH (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(281)
<223> 铰链

<220>

<221> misc_feature
<222> (500)..(537)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (538)..(649)
<223> VL2

<220>
<221> misc_feature
<222> (650)..(664)
<223> 连接子 2

<220>
<221> misc_feature
<222> (665)..(785)
<223> VH2

<400> 178

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145 150 155 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
165 170 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
180 185 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
195 200 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
210 215 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225 230 235 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
245 250 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
260 265 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
275 280 285

Gly Ser Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
290 295 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305 310 315 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
325 330 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
340 345 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
355 360 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Ser Ile
370 375 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385 390 395 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
405 410 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
420 425 430

```
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Gly Gly Gly Ser Gly Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly
        515                 520                 525

Ser Gly Gly Gly Gly Ser Gly Asn Ser Asp Ile Val Leu Thr Gln Ser
    530                 535                 540

Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys
545                 550                 555                 560

Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp
                565                 570                 575

Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala
            580                 585                 590

Ser Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser
        595                 600                 605

Gly Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile
    610                 615                 620

Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly
625                 630                 635                 640

Gly Gly Thr Lys Leu Glu Ile Lys Arg Gly Gly Gly Gly Ser Gly Gly
                645                 650                 655

Gly Gly Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Lys Glu Ser Gly
            660                 665                 670

Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val
        675                 680                 685
```

```
Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro
    690                 695                 700

Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser
705                 710                 715                 720

Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp
                725                 730                 735

Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp
            740                 745                 750

Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His
        755                 760                 765

Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser
    770                 775                 780

Ser
785
```

<210> 179
<211> 2391
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-STD2-2e12HL (DNA)

<400> 179

```
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca     60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct    120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg    180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct    240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc    300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca    360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt    420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg    480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac    540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat    600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta    660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg    720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg    780
```

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attatggtgg cggtggctcg ggcggtggtg gatctggagg aggtgggagt 1560

gggaattatg gtggcggtgg ctcgggcggt ggtggatctg gaggaggtgg gagtgggaat 1620

tctcaggtgc agctgaagga gtcaggacct ggcctggtgg cgccctcaca gagcctgtcc 1680

atcacatgca ccgtctcagg gttctcatta accggctatg gtgtaaactg ggttcgccag 1740

cctccaggaa agggtctgga gtggctggga atgatatggg gtgatggaag cacagactat 1800

aattcagctc tcaaatccag actatcgatc accaaggaca actccaagag ccaagttttc 1860

ttaaaaatga acagtctgca aactgatgac acagccagat actactgtgc tcgagatggt 1920

tatagtaact ttcattacta tgttatggac tactggggtc aaggaacctc agtcaccgtc 1980

tcctctgggg gtggaggctc tggtggcggt ggatccggcg gaggtgggtc gggtggcggc 2040

ggatctgaca ttgtgctcac ccaatctcca gcttctttgg ctgtgtctct aggtcagaga 2100

gccaccatct cctgcagagc cagtgaaagt gttgaatatt atgtcacaag tttaatgcag 2160

tggtaccaac agaaaccagg acagccaccc aaactcctca tctctgctgc tagcaacgta 2220

gaatctgggg tccctgccag gtttagtggc agtgggtctg ggacagactt tagcctcaac 2280

atccatcctg tggaggagga tgatattgca atgtatttct gtcagcaaag taggaaggtt 2340

ccatggacgt tcggtggagg caccaagctg gaaatcaaac gttaatctag a 2391

<210> 180
<211> 790
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-STD2-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(537)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (538)..(658)
<223> VH2

<220>
<221> misc_feature
<222> (659)..(678)
<223> 连接子 2

<220>
<221> misc_feature
<222> (679)..(790)
<223> VL2

<400> 180
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

```
Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320
```

```
Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Gly Gly Gly Ser Gly Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly
        515                 520                 525

Ser Gly Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser
    530                 535                 540

Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr
545                 550                 555                 560

Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln
                565                 570                 575
```

```
Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly
            580                 585                 590

Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys
        595                 600                 605

Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr
    610                 615                 620

Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe
625                 630                 635                 640

His Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val
                645                 650                 655

Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
            660                 665                 670

Ser Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser
        675                 680                 685

Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser
    690                 695                 700

Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln
705                 710                 715                 720

Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val
                725                 730                 735

Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
            740                 745                 750

Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr
        755                 760                 765

Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr
    770                 775                 780

Lys Leu Glu Ile Lys Arg
785                 790
```

<210> 181
<211> 790
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>

<223> 2H7sssIgG1(P238S/P331S)-STD2-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(537)
<223> EFD-BD2 连接子
<220>
<221> misc_feature
<222> (538)..(658)
<223> VH2

<220>
<221> misc_feature
<222> (659)..(678)
<223> 连接子 2

<220>
<221> misc_feature
<222> (679)..(790)
<223> VL2

<400> 181

```
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1               5                  10                  15


Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
            20                  25                  30


Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
        35                  40                  45


Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
    50                  55                  60
```

```
Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Ser Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335
```

```
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Ser Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Gly Gly Gly Ser Gly Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly
        515                 520                 525

Ser Gly Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser
    530                 535                 540

Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr
545                 550                 555                 560

Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln
                565                 570                 575

Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly
            580                 585                 590
```

Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys
595 600 605

Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr
610 615 620

Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe
625 630 635 640

His Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val
645 650 655

Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
660 665 670

Ser Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser
675 680 685

Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser
690 695 700

Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln
705 710 715 720

Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val
725 730 735

Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
740 745 750

Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr
755 760 765

Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr
770 775 780

Lys Leu Glu Ile Lys Arg
785 790

<210> 182
<211> 2283
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H1-2e12HL (DNA)

<400> 182
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attctcaggt gcagctgaag gagtcaggac ctggcctggt ggcgccctca 1560

cagagcctgt ccatcacatg caccgtctca gggttctcat taaccggcta tggtgtaaac 1620

tgggttcgcc agcctccagg aaagggtctg gagtggctgg gaatgatatg gggtgatgga 1680

agcacagact ataattcagc tctcaaatcc agactatcga tcaccaagga caactccaag 1740

agccaagttt tcttaaaaat gaacagtctg caaactgatg acacagccag atactactgt 1800

gctcgagatg gttatagtaa ctttcattac tatgttatgg actactgggg tcaaggaacc 1860

tcagtcaccg tctcctctgg gggtggaggc tctggtggcg gtggatccgg cggaggtggg 1920

tcgggtggcg gcggatctga cattgtgctc acccaatctc cagcttcttt ggctgtgtct 1980

ctaggtcaga gagccaccat ctcctgcaga gccagtgaaa gtgttgaata ttatgtcaca 2040

agtttaatgc agtggtacca acagaaacca ggacagccac ccaaactcct catctctgct 2100

gctagcaacg tagaatctgg ggtccctgcc aggtttagtg gcagtgggtc tgggacagac 2160

tttagcctca acatccatcc tgtggaggag gatgatattg caatgtattt ctgtcagcaa 2220

agtaggaagg ttccatggac gttcggtgga ggcaccaagc tggaaatcaa acgttaatct 2280

aga 2283

<210> 183
<211> 754
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H1-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(281)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(501)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (502)..(622)
<223> VH2

<220>
<221> misc_feature
<222> (623)..(642)
<223> 连接子 2

<220>
<221> misc_feature
<222> (643)..(754)
<223> VL2

<400> 183

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145 150 155 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
165 170 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
180 185 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
195 200 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
210 215 220

```
Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480
```

```
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu
            500                 505                 510

Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe
        515                 520                 525

Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys
    530                 535                 540

Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr
545                 550                 555                 560

Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys
                565                 570                 575

Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala
            580                 585                 590

Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val
        595                 600                 605

Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly
    610                 615                 620

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
625                 630                 635                 640

Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser
                645                 650                 655

Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser Val Glu
            660                 665                 670

Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln
        675                 680                 685

Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser Gly Val
    690                 695                 700

Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn
705                 710                 715                 720

Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln
                725                 730                 735
```

Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
740 745 750

Lys Arg

<210> 184
<211> 2301
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H2-2e12HL (DNA)

<400> 184
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320
gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaagg gtggcggtgg ctcggggaat tctcaggtgc agctgaagga gtcaggacct 1560

ggcctggtgg cgccctcaca gagcctgtcc atcacatgca ccgtctcagg gttctcatta 1620

accggctatg gtgtaaactg ggttcgccag cctccaggaa agggtctgga gtggctggga 1680

atgatatggg gtgatggaag cacagactat aattcagctc tcaaatccag actatcgatc 1740

accaaggaca actccaagag ccaagttttc ttaaaaatga acagtctgca aactgatgac 1800

acagccagat actactgtgc tcgagatggt tatagtaact ttcattacta tgttatggac 1860

tactggggtc aaggaacctc agtcaccgtc tcctctgggg gtggaggctc tggtggcggt 1920

ggatccggcg gaggtgggtc gggtggcggc ggatctgaca ttgtgctcac ccaatctcca 1980

gcttctttgg ctgtgtctct aggtcagaga gccaccatct cctgcagagc cagtgaaagt 2040

gttgaatatt atgtcacaag tttaatgcag tggtaccaac agaaaccagg acagccaccc 2100

aaactcctca tctctgctgc tagcaacgta gaatctgggg tccctgccag gtttagtggc 2160

agtgggtctg ggacagactt tagcctcaac atccatcctg tggaggagga tgatattgca 2220

atgtatttct gtcagcaaag taggaaggtt ccatggacgt tcggtggagg caccaagctg 2280

gaaatcaaac gttaatctag a 2301

<210> 185
<211> 760
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H2-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(127)
<223> VL

<220>
<221> misc_feature
<222> (128)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(507)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (508)..(628)
<223> VH2

<220>
<221> misc_feature
<222> (629)..(648)
<223> 连接子 2

<220>
<221> misc_feature
<222> (649)..(760)
<223> VL2

<400> 185

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arq Pro Gly Ala
145 150 155 160

```
Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415
```

```
Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Gly Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys
            500                 505                 510

Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr
        515                 520                 525

Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val
    530                 535                 540

Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly
545                 550                 555                 560

Asp Gly Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile
                565                 570                 575

Thr Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu
            580                 585                 590

Gln Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser
        595                 600                 605

Asn Phe His Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val
    610                 615                 620

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
625                 630                 635                 640

Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro
                645                 650                 655

Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg
            660                 665                 670
```

Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr
675 680 685

Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser
690 695 700

Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly
705 710 715 720

Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala
725 730 735

Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly
740 745 750

Gly Thr Lys Leu Glu Ile Lys Arg
755 760

<210> 186
<211> 2307
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H3-2e12HL (DNA)

<400> 186
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120
ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attatggtgg cggtggctct gggaattctc aggtgcagct gaaggagtca 1560

ggacctggcc tggtggcgcc ctcacagagc ctgtccatca catgcaccgt ctcagggttc 1620

tcattaaccg gctatggtgt aaactgggtt cgccagcctc caggaaaggg tctggagtgg 1680

ctgggaatga tatggggtga tggaagcaca gactataatt cagctctcaa atccagacta 1740

tcgatcacca aggacaactc caagagccaa gttttcttaa aaatgaacag tctgcaaact 1800

gatgacacag ccagatacta ctgtgctcga gatggttata gtaactttca ttactatgtt 1860

atggactact ggggtcaagg aacctcagtc accgtctcct ctgggggtgg aggctctggt 1920

ggcggtggat ccggcggagg tgggtcgggt ggcggcggat ctgacattgt gctcacccaa 1980

tctccagctt ctttggctgt gtctctaggt cagagagcca ccatctcctg cagagccagt 2040
gaaagtgttg aatattatgt cacaagttta atgcagtggt accaacagaa accaggacag 2100

ccacccaaac tcctcatctc tgctgctagc aacgtagaat ctggggtccc tgccaggttt 2160

agtggcagtg ggtctgggac agactttagc ctcaacatcc atcctgtgga ggaggatgat 2220

attgcaatgt atttctgtca gcaaagtagg aaggttccat ggacgttcgg tggaggcacc 2280

aagctggaaa tcaaacgtta atctaga 2307

<210> 187
<211> 762
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H3-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL
<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(509)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (510)..(630)
<223> VH2

<220>
<221> misc_feature
<222> (631)..(650)
<223> 连接子 2

<220>
<221> misc_feature
<222> (651)..(762)
<223> VL2

<400> 187

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

```
Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365
```

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
370 375 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385 390 395 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
405 410 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
420 425 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
435 440 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
450 455 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465 470 475 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
485 490 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Asn Ser Gln Val Gln
500 505 510

Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser
515 520 525

Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn
530 535 540

Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile
545 550 555 560

Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu
565 570 575

Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn
580 585 590

Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly
595 600 605

Tyr Ser Asn Phe His Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr
610 615 620

```
Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln
                645                 650                 655

Ser Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser
            660                 665                 670

Cys Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln
        675                 680                 685

Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala
    690                 695                 700

Ala Ser Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly
705                 710                 715                 720

Ser Gly Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp
                725                 730                 735

Ile Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe
            740                 745                 750

Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
        755                 760
```

<210> 188
<211> 2316
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H4-2e12HL (DNA)

<400> 188

```
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca     60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct    120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg    180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct    240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc    300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca    360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt    420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg    480
```

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840
acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaagg gtggcggtgg ctcgggcggt ggtggatctg ggaattctca ggtgcagctg 1560

aaggagtcag gacctggcct ggtggcgccc tcacagagcc tgtccatcac atgcaccgtc 1620

tcagggttct cattaaccgg ctatggtgta aactgggttc gccagcctcc aggaaagggt 1680

ctggagtggc tgggaatgat atggggtgat ggaagcacag actataattc agctctcaaa 1740

tccagactat cgatcaccaa ggacaactcc aagagccaag ttttcttaaa aatgaacagt 1800

ctgcaaactg atgacacagc cagatactac tgtgctcgag atggttatag taactttcat 1860

tactatgtta tggactactg gggtcaagga acctcagtca ccgtctcctc tgggggtgga 1920

ggctctggtg gcggtggatc cggcggaggt gggtcgggtg gcggcggatc tgacattgtg 1980

ctcacccaat ctccagcttc tttggctgtg tctctaggtc agagagccac catctcctgc 2040

agagccagtg aaagtgttga tattatgtc acaagtttaa tgcagtggta ccaacagaaa 2100

ccaggacagc cacccaaact cctcatctct gctgctagca acgtagaatc tggggtccct 2160

gccaggttta gtggcagtgg gtctgggaca gactttagcc tcaacatcca tcctgtggag 2220

gaggatgata ttgcaatgta tttctgtcag caaagtagga aggttccatg gacgttcggt 2280

ggaggcacca agctggaaat caaacgttaa tctaga 2316

<210> 189
<211> 765

<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H4-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(512)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (513)..(633)
<223> VH2

<220>
<221> misc_feature
<222> (634)..(653)
<223> 连接子 2

<220>
<221> misc_feature
<222> (654)..(765)
<223> VL2

<400> 189

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

```
Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
        35                  40                  45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
    50                  55                  60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285
```

```
Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Asn Ser
            500                 505                 510

Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln
        515                 520                 525

Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr
    530                 535                 540

Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu
545                 550                 555                 560
```

Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser Ala Leu Lys
565 570 575

Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln Val Phe Leu
580 585 590

Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala
595 600 605

Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val Met Asp Tyr Trp Gly
610 615 620

Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly
625 630 635 640

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Val
645 650 655

Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala
660 665 670

Thr Ile Ser Cys Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser
675 680 685

Leu Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu
690 695 700

Ile Ser Ala Ala Ser Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser
705 710 715 720

Gly Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu
725 730 735

Glu Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro
740 745 750

Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
755 760 765

<210> 190
<211> 2322
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H5-2e12HL (DNA)

<400> 190
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attatggtgg cggtggctcg ggcggtggtg gatctgggaa ttctcaggtg 1560

cagctgaagg agtcaggacc tggcctggtg gcgccctcac agagcctgtc catcacatgc 1620

accgtctcag ggttctcatt aaccggctat ggtgtaaact gggttcgcca gcctccagga 1680

aagggtctgg agtggctggg aatgatatgg ggtgatggaa gcacagacta taattcagct 1740

ctcaaatcca gactatcgat caccaaggac aactccaaga gccaagtttt cttaaaaatg 1800

aacagtctgc aaactgatga cacagccaga tactactgtg ctcgagatgg ttatagtaac 1860

tttcattact atgttatgga ctactggggt caaggaacct cagtcaccgt ctcctctggg 1920

ggtggaggct ctggtggcgg tggatccggc ggaggtgggt cgggtggcgg cggatctgac 1980

attgtgctca cccaatctcc agcttctttg gctgtgtctc taggtcagag agccaccatc 2040

tcctgcagag ccagtgaaag tgttgaatat tatgtcacaa gtttaatgca gtggtaccaa 2100

cagaaaccag gacagccacc caaactcctc atctctgctg ctagcaacgt agaatctggg 2160

gtccctgcca ggtttagtgg cagtgggtct gggacagact ttagcctcaa catccatcct 2220

gtggaggagg atgatattgc aatgtatttc tgtcagcaaa gtaggaaggt tccatggacg 2280

ttcggtggag gcaccaagct ggaaatcaaa cgttaatcta ga 2322

<210> 191
<211> 767
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H5-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(514)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (515)..(635)
<223> VH2

<220>
<221> misc_feature

<222> (636)..(655)
<223> 连接子 2

<220>
<221> misc_feature
<222> (656)..(767)
<223> VL2

<400> 191

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145 150 155 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
165 170 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
180 185 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
195 200 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
210 215 220

```
Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480
```

```
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro
        515                 520                 525

Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr
    530                 535                 540

Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu
545                 550                 555                 560

Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser Ala
                565                 570                 575

Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln Val
            580                 585                 590

Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr Tyr
        595                 600                 605

Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val Met Asp Tyr
    610                 615                 620

Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp
                645                 650                 655

Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly Gln
            660                 665                 670

Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val
        675                 680                 685

Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys
    690                 695                 700

Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser Gly Val Pro Ala Arg
705                 710                 715                 720

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn Ile His Pro
                725                 730                 735
```

Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys
740 745 750

Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
755 760 765

<210> 192
<211> 2331
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H6-2e12HL (DNA)

<400> 192
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgagcccagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaagg gtggcggtgg ctcgggcggt ggtggatctg ggggaggagg cagcgggaat 1560

tctcaggtgc agctgaagga gtcaggacct ggcctggtgg cgccctcaca gagcctgtcc 1620

atcacatgca ccgtctcagg gttctcatta accggctatg gtgtaaactg ggttcgccag 1680

cctccaggaa agggtctgga gtggctggga atgatatggg gtgatggaag cacagactat 1740

aattcagctc tcaaatccag actatcgatc accaaggaca actccaagag ccaagttttc 1800

ttaaaaatga acagtctgca aactgatgac acagccagat actactgtgc tcgagatggt 1860

tatagtaact ttcattacta tgttatggac tactggggtc aaggaacctc agtcaccgtc 1920

tcctctgggg gtggaggctc tggtggcggt ggatccggcg gaggtgggtc gggtggcggc 1980

ggatctgaca ttgtgctcac ccaatctcca gcttctttgg ctgtgtctct aggtcagaga 2040

gccaccatct cctgcagagc cagtgaaagt gttgaatatt atgtcacaag tttaatgcag 2100

tggtaccaac agaaaccagg acagccaccc aaactcctca tctctgctgc tagcaacgta 2160

gaatctgggg tccctgccag gtttagtggc agtgggtctg ggacagactt tagcctcaac 2220

atccatcctg tggaggagga tgatattgca atgtatttct gtcagcaaag taggaaggtt 2280

ccatggacgt tcggtggagg caccaagctg gaaatcaaac gttaatctag a 2331

<210> 193
<211> 770
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H6-2e12HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)

<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(517)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (518)..(638)
<223> VH2

<220>
<221> misc_feature
<222> (639)..(658)
<223> 连接子 2

<220>
<221> misc_feature
<222> (659)..(770)
<223> VL2

<400> 193

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

```
Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415
```

```
Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495
                .

Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
            500                 505                 510

Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu
        515                 520                 525

Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe
    530                 535                 540

Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys
545                 550                 555                 560

Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr
                565                 570                 575

Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys
            580                 585                 590

Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala
        595                 600                 605

Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val
    610                 615                 620

Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly
625                 630                 635                 640

Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
                645                 650                 655

Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser
            660                 665                 670
```

Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser Val Glu
675 680 685

Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln
690 695 700

Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser Gly Val
705 710 715 720

Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn
725 730 735

Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln
740 745 750

Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
755 760 765

Lys Arg
770

<210> 194
<211> 2301
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sscIgG1-H7-2e12HL (DNA)

<400> 194
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgtgcccagc acctgaactc ctgggggggat cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa accatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatcca agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaagg ggtgtccacc ttgtccgaat tctcaggtgc agctgaagga gtcagggcct 1560

ggcagcgtgg cgccctcaca gagcctgtcc atcacatgca ccgtctcagg gttctcatta 1620

accggctatg gtgtaaactg ggttcgccag cctccaggaa agggtctgga gtggctggga 1680

atgatatggg gtgatggaag cacagactat aattcagctc tcaaatccag actatcgatc 1740

accaaggaca actccaagag ccaagttttc ttaaaaatga acagtctgca aactgatgac 1800

acagccagat actactgtgc tcgagatggt tatagtaact ttcattacta tgttatggac 1860

tactggggtc aaggaacctc agtcaccgtc tcctctgggg gtggaggctc tggtggcggt 1920

ggatccggcg gaggtgggtc gggtggcggc ggatctgaca ttgtgctcac ccaatctcca 1980

gcttctttgg ctgtgtctct aggtcagaga gccaccatct cctgcagagc cagtgaaagt 2040

gttgaatatt atgtcacaag tttaatgcag tggtaccaac agaaaccagg acagccaccc 2100

aagctcctca tctctgctgc tagcaacgta gaatctgggg tccctgccag gtttagtggc 2160

agtgggtctg ggacagactt tagcctcaac atccatcctg tggaggagga tgatattgca 2220

atgtatttct gtcagcaaag taggaaggtt ccatggacgt tcggtggagg caccaagctg 2280

gaaatcaaac gttaatctag a 2301

<210> 195
<211> 760
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sscIgG1-H7-2e12HL (w/2e12 连接子) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(507)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (508)..(628)
<223> VH2

<220>
<221> misc_feature
<222> (629)..(648)
<223> 连接子 2

<220>
<221> misc_feature
<222> (649)..(760)
<223> VL2

<400> 195

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

```
Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Ser Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335
```

```
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Gly Cys Pro Pro Cys Pro Asn Ser Gln Val Gln Leu Lys
            500                 505                 510

Glu Ser Gly Pro Gly Ser Val Ala Pro Ser Gln Ser Leu Ser Ile Thr
        515                 520                 525

Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val
    530                 535                 540

Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly
545                 550                 555                 560

Asp Gly Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile
                565                 570                 575

Thr Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu
            580                 585                 590
```

Gln Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser
595 600 605

Asn Phe His Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val
610 615 620

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
625 630 635 640

Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro
645 650 655

Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg
660 665 670

Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr
675 680 685

Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser
690 695 700

Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly
705 710 715 720

Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala
725 730 735

Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly
740 745 750

Gly Thr Lys Leu Glu Ile Lys Arg
755 760

<210> 196
<211> 2283
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H7-G194 HL (DNA)

<400> 196
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgtgcccagc acctgaactc ctgggggggat cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa accatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatcca agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaagg ggtgtccacc ttgtccgaat tctgaggtcc agctgcaaca gtctggacct 1560

gaactggtga agcctggagc ttcaatgaag atttcctgca aggcctctgg ttactcattc 1620

actggctaca tcgtgaactg gctgaagcag agccatggaa agaaccttga gtggattgga 1680

cttattaatc catacaaagg tcttactacc tacaaccaga aattcaaggg caaggccaca 1740

ttaactgtag acaagtcatc cagcacagcc tacatggagc tcctcagtct gacatctgaa 1800

gactctgcag tctattactg tgcaagatct gggtactatg gtgactcgga ctggtacttc 1860

gatgtctggg gcgcagggac cacggtcacc gtctcctctg gtggcggtgg ctcgggcggt 1920

ggtggatctg gaggaggtgg gagcgctagc gacatccaga tgacacagac tacatcctcc 1980

ctgtctgcct ctctgggaga cagagtcacc atcagttgca gggcaagtca ggacattcgc 2040

aattatttaa actggtatca gcagaaacca gatggaactg ttaaactcct gatctactac 2100

acatcaagat tacactcagg agtcccatca aggttcagtg gcagtgggtc tggaacagat 2160

tattctctca ccattgccaa cctgcaacca gaagatattg ccacttactt ttgccaacag 2220

ggtaatacgc ttccgtggac gttcggtgga ggcaccaaac tggtaaccaa acggtaatct 2280

aga 2283

<210> 197
<211> 754
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H7-G194 HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH1

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(507)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (508)..(629)
<223> VH2

<220>
<221> misc_feature
<222> (630)..(646)
<223> 连接子 2

<220>
<221> misc_feature
<222> (647)..(754)
<223> VL2

<400> 197

```
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1               5                  10                  15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
            20                  25                  30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
        35                  40                  45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
    50                  55                  60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65                  70                  75                  80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
                85                  90                  95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255
```

```
Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Ser Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Gly Cys Pro Pro Cys Pro Asn Ser Glu Val Gln Leu Gln
            500                 505                 510
```

```
Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala Ser Met Lys Ile Ser
        515                 520                 525

Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr Ile Val Asn Trp Leu
    530                 535                 540

Lys Gln Ser His Gly Lys Asn Leu Glu Trp Ile Gly Leu Ile Asn Pro
545                 550                 555                 560

Tyr Lys Gly Leu Thr Thr Tyr Asn Gln Lys Phe Lys Gly Lys Ala Thr
                565                 570                 575

Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Glu Leu Leu Ser
            580                 585                 590

Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg Ser Gly Tyr
        595                 600                 605

Tyr Gly Asp Ser Asp Trp Tyr Phe Asp Val Trp Gly Ala Gly Thr Thr
    610                 615                 620

Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
625                 630                 635                 640

Gly Gly Gly Ser Ala Ser Asp Ile Gln Met Thr Gln Thr Thr Ser Ser
                645                 650                 655

Leu Ser Ala Ser Leu Gly Asp Arg Val Thr Ile Ser Cys Arg Ala Ser
            660                 665                 670

Gln Asp Ile Arg Asn Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Asp Gly
        675                 680                 685

Thr Val Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val
    690                 695                 700

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr
705                 710                 715                 720

Ile Ala Asn Leu Gln Pro Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln
                725                 730                 735

Gly Asn Thr Leu Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Val Thr
            740                 745                 750

Lys Arg
```

<210> 198

<211> 2298
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H7-G281 HL (DNA)

<400> 198
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttctga caaaactcac 840

acatccccac cgtgcccagc acctgaactc ctgggggggat cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa accatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatcca agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaagg ggtgtccacc ttgtccgaat tctgaggtcc agctgcaaca gtctggacct 1560

gaactggtga agcctggagc ttcaatgaag atttcctgca aggcctctgg ttactcattc 1620

actggctaca tcgtgaactg gctgaagcag agccatggaa agaaccttga gtggattgga 1680

cttattaatc catacaaagg tcttactacc tacaaccaga aattcaaggg caaggccaca 1740

ttaactgtag acaagtcatc cagcacagcc tacatggagc tcctcagtct gacatctgaa 1800

gactctgcag tctattactg tgcaagatct gggtactatg gtgactcgga ctggtacttc 1860

gatgtctggg gcgcagggac cacggtcacc gtctcctctg gtggcggtgg ctcgggcggt 1920

ggtggatctg gaggaggtgg gagcgggggа ggtggcagtg ctagcgacat ccagatgaca 1980

cagactacat cctccctgtc tgcctctctg ggagacagag tcaccatcag ttgcagggca 2040

agtcaggaca ttcgcaatta tttaaactgg tatcagcaga aaccagatgg aactgttaaa 2100

ctcctgatct actacacatc aagattacac tcaggagtcc catcaaggtt cagtggcagt 2160

gggtctggaa cagattattc tctcaccatt gccaacctgc aaccagaaga tattgccact 2220

tacttttgcc aacagggtaa tacgcttccg tggacgttcg gtggaggcac caaactggta 2280

accaaacggt aatctaga 2298

<210> 199
<211> 759
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7sssIgG1-H7-G281 HL (w/2e12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (23)..(128)
<223> VL1

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH1

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>

<221> misc_feature
<222> (500)..(507)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (508)..(629)
<223> VH2

<220>
<221> misc_feature
<222> (630)..(651)
<223> 连接子 2

<220>
<221> misc_feature
<222> (652)..(759)
<223> VL2

<400> 199

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
100 105 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
115 120 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
130 135 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145 150 155 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
165 170 175

```
Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Ser
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Ser Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445
```

```
Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Gly Cys Pro Pro Cys Pro Asn Ser Glu Val Gln Leu Gln
            500                 505                 510

Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala Ser Met Lys Ile Ser
        515                 520                 525

Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr Ile Val Asn Trp Leu
    530                 535                 540

Lys Gln Ser His Gly Lys Asn Leu Glu Trp Ile Gly Leu Ile Asn Pro
545                 550                 555                 560

Tyr Lys Gly Leu Thr Thr Tyr Asn Gln Lys Phe Lys Gly Lys Ala Thr
                565                 570                 575

Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Glu Leu Leu Ser
            580                 585                 590

Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg Ser Gly Tyr
        595                 600                 605

Tyr Gly Asp Ser Asp Trp Tyr Phe Asp Val Trp Gly Ala Gly Thr Thr
    610                 615                 620

Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
625                 630                 635                 640

Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser Asp Ile Gln Met Thr
                645                 650                 655

Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly Asp Arg Val Thr Ile
            660                 665                 670

Ser Cys Arg Ala Ser Gln Asp Ile Arg Asn Tyr Leu Asn Trp Tyr Gln
        675                 680                 685

Gln Lys Pro Asp Gly Thr Val Lys Leu Leu Ile Tyr Tyr Thr Ser Arg
    690                 695                 700
```

Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr
705 710 715 720

Asp Tyr Ser Leu Thr Ile Ala Asn Leu Gln Pro Glu Asp Ile Ala Thr
725 730 735

Tyr Phe Cys Gln Gln Gly Asn Thr Leu Pro Trp Thr Phe Gly Gly Gly
740 745 750

Thr Lys Leu Val Thr Lys Arg
755

<210> 200
<211> 1533
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2e12-sss-IgG1 HL SMIP (DNA)

<400> 200
atggattttc aagtgcagat tttcagcttc ctgctaatca gtgcttcagt cataatgtcc 60

agaggagtcc aggtgcagct gaaggagtca ggacctggcc tggtggcgcc ctcacagagc 120

ctgtccatca catgcaccgt ctcagggttc tcattaaccg gctatggtgt aaactgggtt 180

cgccagcctc caggaaaggg tctagagtgg ctgggaatga tatggggtga tggaagcaca 240

gactataatt cagctctcaa atccagacta tcgatcacca aggacaactc caagagccaa 300

gttttcttaa aaatgaacag tctgcaaact gatgacacag ccagatacta ctgtgctcga 360

gatggttata gtaactttca ttactatgtt atggactact ggggtcaagg aacctcagtc 420

accgtctcct ctgggggtgg aggctctggt ggcggtggat ccggcggagg tgggtcgggt 480

ggcggcggat ctgacattgt gctcacccaa tctccagctt ctttggctgt gtctctaggt 540

cagagagcca ccatctcctg cagagccagt gaaagtgttg aatattatgt cacaagttta 600

atgcagtggt accaacagaa accaggacag ccacccaaac tcctcatctc tgctgctagc 660

aacgtagaat ctggggtccc tgccaggttt agtggcagtg ggtctgggac agactttagc 720

ctcaacatcc atcctgtgga ggaggatgat attgcaatgt atttctgtca gcaaagtagg 780

aaggttccat ggacgttcgg tggaggcacc aagctggaaa tcaaacgtga tcaggagccc 840

aaatcttctg acaaaactca cacatcccca ccgtccccag cacctgaact cctgggggga 900

tcgtcagtct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggacccct 960

gaggtcacat gcgtggtggt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 1020

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 1080

agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 1140

gagtacaagt gcaaggtctc caacaaagcc ctcccagcct ccatcgagaa aacaatctcc 1200

aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggatgag 1260

ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc 1320

gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1380

ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 1440

cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1500

cagaagagcc tctccctgtc tccgggtaaa tga 1533

<210> 201
<211> 510
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2e12-sss-IgG1 HL SMIP (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature
<222> (24)..(144)
<223> VH

<220>
<221> misc_feature
<222> (145)..(164)
<223> 连接子

<220>
<221> misc_feature
<222> (165)..(276)
<223> VL

<220>
<221> misc_feature
<222> (279)..(293)
<223> 铰链

<400> 201

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Val Gln Val Gln Leu Lys Glu Ser Gly Pro
20 25 30

Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser
35 40 45

```
Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro
    50                  55                  60

Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr
65                  70                  75                  80

Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn
                85                  90                  95

Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp
            100                 105                 110

Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr
        115                 120                 125

Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
    130                 135                 140

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
145                 150                 155                 160

Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala
                165                 170                 175

Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser
            180                 185                 190

Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro
        195                 200                 205

Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser
    210                 215                 220
Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser
225                 230                 235                 240

Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys
                245                 250                 255

Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu
            260                 265                 270

Glu Ile Lys Arg Asp Gln Glu Pro Lys Ser Ser Asp Lys Thr His Thr
        275                 280                 285

Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Ser Ser Val Phe
    290                 295                 300
```

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro
305 310 315 320

Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
325 330 335

Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr
340 345 350

Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val
355 360 365

Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
370 375 380

Lys Val Ser Asn Lys Ala Leu Pro Ala Ser Ile Glu Lys Thr Ile Ser
385 390 395 400

Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
405 410 415

Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
420 425 430

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
435 440 445

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
450 455 460

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
465 470 475 480

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
485 490 495

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
500 505 510

<210> 202
<211> 1518
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2e12-sss-IgG1 LH SMIP (DNA)

<400> 202
atggattttc aagtgcagat tttcagcttc ctgctaatca gtgcttcagt cataatgtcc 60

agaggagtcg acattgtgct cacccaatct ccagcttctt tggctgtgtc tctaggtcag 120

agagccacca tctcctgcag agccagtgaa agtgttgaat attatgtcac aagtttaatg 180

cagtggtacc aacagaaacc aggacagcca cccaaactcc tcatctctgc tgctagcaac 240

gtagaatctg gggtccctgc caggtttagt ggcagtgggt ctgggacaga ctttagcctc 300

aacatccatc ctgtggagga ggatgatatt gcaatgtatt tctgtcagca aagtaggaag 360

gttccatgga cgttcggtgg aggcaccaag ctggaaatca aacggggtgg cggtggatcc 420

ggcggaggtg ggtcgggtgg cggcggatct caggtgcagc tgaaggagtc aggacctggc 480

ctggtggcgc cctcacagag cctgtccatc acatgcaccg tctcagggtt ctcattaacc 540

ggctatggtg taaactgggt tcgccagcct ccaggaaagg gtctagagtg gctgggaatg 600

atatggggtg atggaagcac agactataat tcagctctca aatccagact atcgatcacc 660

aaggacaact ccaagagcca agttttctta aaaatgaaca gtctgcaaac tgatgacaca 720

gccagatact actgtgctcg agatggttat agtaactttc attactatgt tatggactac 780

tggggtcaag gaacctcagt caccgtctcc tctgatcagg agcccaaatc ttctgacaaa 840

actcacacat ccccaccgtc cccagcacct gaactcctgg ggggaccgtc agtcttcctc 900

ttcccccaa aacccaagga caccctcatg atctcccgga cccctgaggt cacatgcgtg 960

gtggtggacg tgagccacga agaccctgag gtcaagttca actggtacgt ggacggcgtg 1020

gaggtgcata atgccaagac aaagccgcgg gaggagcagt acaacagcac gtaccgtgtg 1080

gtcagcgtcc tcaccgtcct gcaccaggac tggctgaatg gcaaggagta caagtgcaag 1140

gtctccaaca aagccctccc agcccccatc gagaaaacaa tctccaaagc caaagggcag 1200

ccccgagaac cacaggtgta caccctgccc ccatcccggg atgagctgac caagaaccag 1260

gtcagcctga cctgcctggt caaaggcttc tatcccagcg acatcgccgt ggagtgggag 1320

agcaatgggc agccggagaa caactacaag accacgcctc ccgtgctgga ctccgacggc 1380

tccttcttcc tctacagcaa gctcaccgtg gacaagagca ggtggcagca ggggaacgtc 1440

ttctcatgct ccgtgatgca tgaggctctg cacaaccact acacgcagaa gagcctctcc 1500

ctgtctccgg gtaaatga 1518

<210> 203
<211> 505
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2e12-sss-IgG1 LH SMIP (AA)

<220>

<221> misc_feature
<222> (1)..(23)
<223> 前导序列

<220>
<221> misc_feature
<222> (24)..(135)
<223> VL

<220>
<221> misc_feature
<222> (136)..(150)
<223> 连接子

<220>
<221> misc_feature
<222> (151)..(271)
<223> VH

<220>
<221> misc_feature
<222> (274)..(288)
<223> 铰链

<400> 203

```
Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1               5                   10                  15

Val Ile Met Ser Arg Gly Val Asp Ile Val Leu Thr Gln Ser Pro Ala
            20                  25                  30

Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala
        35                  40                  45

Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln
    50                  55                  60

Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn
65                  70                  75                  80

Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr
                85                  90                  95

Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met
            100                 105                 110

Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly
        115                 120                 125

Thr Lys Leu Glu Ile Lys Arg Gly Gly Gly Gly Ser Gly Gly Gly Gly
    130                 135                 140

Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly
145                 150                 155                 160
```

```
Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly
                165                 170                 175

Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly
            180                 185                 190

Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp
        195                 200                 205

Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser
    210                 215                 220

Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr
225                 230                 235                 240

Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr
                245                 250                 255

Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Asp
            260                 265                 270

Gln Glu Pro Lys Ser Ser Asp Lys Thr His Thr Ser Pro Pro Ser Pro
        275                 280                 285

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
    290                 295                 300

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
305                 310                 315                 320

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
                325                 330                 335

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
            340                 345                 350

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
        355                 360                 365

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
    370                 375                 380

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
385                 390                 395                 400

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
                405                 410                 415
```

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
420 425 430

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
435 440 445

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
450 455 460

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
465 470 475 480

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
485 490 495

Lys Ser Leu Ser Leu Ser Pro Gly Lys
500 505

<210> 204
<211> 1498
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> G28-1 LH SMIP (DNA)

<400> 204
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtgacatcca gatgactcag tctccagcct ccctatctgc atctgtggga 120

gagactgtca ccatcacatg tcgaacaagt gaaaatgttt acagttattt ggcttggtat 180

cagcagaaac agggaaaatc tcctcagctc ctggtctctt ttgcaaaaac cttagcagaa 240

ggtgtgccat caaggttcag tggcagtgga tcaggcacac agttttctct gaagatcagc 300

agcctgcagc ctgaagattc tggaagttat ttctgtcaac atcattccga taatccgtgg 360

acgttcggtg gaggcaccga actggagatc aaaggtggcg gtggctcggg cggtggtggg 420

tcgggtggcg gcggatctgc tagcgcagtc cagctgcagc agtctggacc tgagctggaa 480

aagcctggcg cttcagtgaa gatttcctgc aaggcttctg gttactcatt cactggctac 540

aatatgaact gggtgaagca gaataatgga aagagccttg agtggattgg aaatattgat 600

ccttattatg gtggtactac ctacaaccgg aagttcaagg gcaaggccac attgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcaagagtc tgacatctga ggactctgca 720

gtctattact gtgcaagatc ggtcggccct atggactact ggggtcaagg aacctcagtc 780

accgtctcga gcgagcccaa atcttctgac aaaactcaca catgcccacc gtgcccagca 840

cctgaactcc tgggtggacc gtcagtcttc ctcttccccc caaaacccaa ggacaccctc 900

atgatctccc ggacccctga ggtcacatgc gtggtggtgg acgtgagcca cgaagaccct 960

gaggtcaagt tcaactggta cgtggacggc gtggaggtgc ataatgccaa gacaaagccg 1020

cgggaggagc agtacaacag cacgtaccgt gtggtcagcg tcctcaccgt cctgcaccag 1080

gactggctga atggcaagga gtacaagtgc aaggtctcca acaaagccct cccagccccc 1140

atcgagaaaa ccatctccaa agccaaaggg cagccccgag aaccacaggt gtacaccctg 1200

cccccatccc gggatgagct gaccaagaac caggtcagcc tgacctgcct ggtcaaaggc 1260

ttctatccaa gcgacatcgc cgtggagtgg gagagcaatg ggcagccgga gaacaactac 1320

aagaccacgc ctcccgtgct ggactccgac ggctccttct tcctctacag caagctcacc 1380

gtggacaaga gcaggtggca gcaggggaac gtcttctcat gctccgtgat gcatgaggct 1440

ctgcacaacc actacacgca gaagagcctc tccctgtctc cgggtaagtg actctaga 1498

<210> 205
<211> 492
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G28-1 LH SMIP (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(127)
<223> VL
<220>
<221> misc_feature
<222> (128)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(260)
<223> VH

<220>
<221> misc_feature
<222> (261)..(275)
<223> 铰链

<400> 205

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

```
Asp Thr Thr Gly Asp Ile Gln Met Thr Gln Ser Pro Ala Ser Leu Ser
            20                  25                  30

Ala Ser Val Gly Glu Thr Val Thr Ile Thr Cys Arg Thr Ser Glu Asn
        35                  40                  45

Val Tyr Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys Ser Pro
    50                  55                  60

Gln Leu Leu Val Ser Phe Ala Lys Thr Leu Ala Glu Gly Val Pro Ser
65                  70                  75                  80

Arg Phe Ser Gly Ser Gly Ser Gly Thr Gln Phe Ser Leu Lys Ile Ser
                85                  90                  95

Ser Leu Gln Pro Glu Asp Ser Gly Ser Tyr Phe Cys Gln His His Ser
            100                 105                 110

Asp Asn Pro Trp Thr Phe Gly Gly Gly Thr Glu Leu Glu Ile Lys Gly
        115                 120                 125

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser
    130                 135                 140

Ala Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Glu Lys Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr
                165                 170                 175

Asn Met Asn Trp Val Lys Gln Asn Asn Gly Lys Ser Leu Glu Trp Ile
            180                 185                 190

Gly Asn Ile Asp Pro Tyr Tyr Gly Gly Thr Thr Tyr Asn Arg Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Lys Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
225                 230                 235                 240

Ala Arg Ser Val Gly Pro Met Asp Tyr Trp Gly Gln Gly Thr Ser Val
                245                 250                 255

Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro
            260                 265                 270
```

```
Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe
        275                 280                 285

Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val
    290                 295                 300

Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe
305                 310                 315                 320

Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro
                325                 330                 335

Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr
            340                 345                 350

Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val
        355                 360                 365

Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala
    370                 375                 380

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg
385                 390                 395                 400

Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly
                405                 410                 415

Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro
            420                 425                 430

Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser
        435                 440                 445

Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln
    450                 455                 460

Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His
465                 470                 475                 480

Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490
```

<210> 206
<211> 1522
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>

<223> G28-1 HL SMIP (DNA)

<400> 206
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtgcggtcca gctgcagcag tctggacctg agtcggaaaa gcctggcgct 120

tcagtgaaga tttcctgcaa ggcttctggt tactcattca ctggctacaa tatgaactgg 180

gtgaagcaga ataatggaaa gagccttgag tggattggaa atattgatcc ttattatggt 240

ggtactacct acaaccggaa gttcaagggc aaggccacat tgactgtaga caaatcctcc 300

ggcacagcct acatgcagct caagagtctg acatctgagg actctgcagt ctattactgt 360

gcaagatcgg tcggccctat ggactactgg ggtcaaggaa cctcagtcac cgtctcttct 420

ggtggcggtg gctcgggcgg tggtgggtcg ggtggcggcg gatcaggagg aggcgggagt 480

gctagcgaca tccagatgac tcagtctcca gcctccctat ctgcatctgt gggagagact 540

gtcaccatca catgtcgaac aagtgaaaat gtttacagtt atttggcttg gtatcagcag 600

aaacagggaa aatctcctca gctcctggtc tcttttgcaa aaaccttagc agaaggtgtg 660

ccatcaaggt tcagtggcag tggatcaggc acacagtttt ctctgaagat cagcagcctg 720

cagcctgaag attctggaag ttatttctgt caacatcatt ccgataatcc gtggacgttc 780

ggtggaggca ccgaactgga gatcaaaggc tcgagcgagc ccaaatcttc tgacaaaact 840

cacacatgcc caccgtgccc agcacctgaa ctcctgggtg gaccgtcagt cttcctcttc 900

cccccaaaac ccaaggacac cctcatgatc tcccggaccc ctgaggtcac atgcgtggtg 960

gtggacgtga gccacgaaga ccctgaggtc aagttcaact ggtacgtgga cggcgtggag 1020

gtgcataatg ccaagacaaa gccgcgggag gagcagtaca acagcacgta ccgtgtggtc 1080

agcgtcctca ccgtcctgca ccaggactgg ctgaatggca aggagtacaa gtgcaaggtc 1140

tccaacaaag ccctcccagc ccccatcgag aaaaccatct ccaaagccaa agggcagccc 1200

cgagaaccac aggtgtacac cctgccccca tcccgggatg agctgaccaa gaaccaggtc 1260

agcctgacct gcctggtcaa aggcttctat ccaagcgaca tcgccgtgga gtgggagagc 1320

aatgggcagc cggagaacaa ctacaagacc acgcctcccg tgctggactc cgacggctcc 1380

ttcttcctct acagcaagct caccgtggac aagagcaggt ggcagcaggg gaacgtcttc 1440

tcatgctccg tgatgcatga ggctctgcac aaccactaca cgcagaagag cctctccctg 1500

tctccgggta agtgactcta ga 1522

<210> 207
<211> 500
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G28-1 HL SMIP (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(136)
<223> VH

<220>
<221> misc_feature
<222> (137)..(158)
<223> 连接子

<220>
<221> misc_feature
<222> (159)..(266)
<223> VL

<220>
<221> misc_feature
<222> (268)..(283)
<223> 铰链

<400> 207

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Ala Val Gln Leu Gln Gln Ser Gly Pro Glu Ser Glu
20 25 30

Lys Pro Gly Ala Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser
35 40 45

Phe Thr Gly Tyr Asn Met Asn Trp Val Lys Gln Asn Asn Gly Lys Ser
50 55 60

Leu Glu Trp Ile Gly Asn Ile Asp Pro Tyr Tyr Gly Gly Thr Thr Tyr
65 70 75 80

Asn Arg Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser
85 90 95

Gly Thr Ala Tyr Met Gln Leu Lys Ser Leu Thr Ser Glu Asp Ser Ala
100 105 110

Val Tyr Tyr Cys Ala Arg Ser Val Gly Pro Met Asp Tyr Trp Gly Gln
115 120 125

Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly
130 135 140

```
Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser Asp Ile
145                 150                 155                 160

Gln Met Thr Gln Ser Pro Ala Ser Leu Ser Ala Ser Val Gly Glu Thr
                165                 170                 175

Val Thr Ile Thr Cys Arg Thr Ser Glu Asn Val Tyr Ser Tyr Leu Ala
            180                 185                 190

Trp Tyr Gln Gln Lys Gln Gly Lys Ser Pro Gln Leu Leu Val Ser Phe
        195                 200                 205

Ala Lys Thr Leu Ala Glu Gly Val Pro Ser Arg Phe Ser Gly Ser Gly
    210                 215                 220

Ser Gly Thr Gln Phe Ser Leu Lys Ile Ser Ser Leu Gln Pro Glu Asp
225                 230                 235                 240

Ser Gly Ser Tyr Phe Cys Gln His His Ser Asp Asn Pro Trp Thr Phe
                245                 250                 255

Gly Gly Gly Thr Glu Leu Glu Ile Lys Gly Ser Ser Glu Pro Lys Ser
            260                 265                 270

Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu
        275                 280                 285

Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu
    290                 295                 300

Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser
305                 310                 315                 320

His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu
                325                 330                 335

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr
            340                 345                 350

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn
        355                 360                 365

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro
    370                 375                 380

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln
385                 390                 395                 400
```

Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val
405 410 415

Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val
420 425 430

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro
435 440 445

Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr
450 455 460

Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val
465 470 475 480

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
485 490 495

Ser Pro Gly Lys
500

<210> 208
<211> 1522
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> G19-4 LH SMIP (DNA)

<400> 208
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtgacatcca gatgacacag actacatcct ccctgtctgc ctctctggga 120

gacagagtca ccatcagttg cagggcaagt caggacattc gcaattattt aaactggtat 180

cagcagaaac cagatggaac tgttaaactc ctgatctact acacatcaag attacactca 240

ggagtcccat caaggttcag tggcagtggg tctggaacag attattctct caccattgcc 300

aacctgcaac cagaagatat tgccacttac ttttgccaac agggtaatac gcttccgtgg 360

acgttcggtg gaggcaccaa actggtaacc aaacggggtg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag cgctagcgag gtccagctgc aacagtctgg acctgaactg 480

gtgaagcctg gagcttcaat gaagatttcc tgcaaggcct ctggttactc attcactggc 540

tacatcgtga actggctgaa gcagagccat ggaaagaacc ttgagtggat tggacttatt 600

aatccataca aaggtcttac tacctacaac cagaaattca agggcaaggc cacattaact 660

gtagacaagt catccagcac agcctacatg gagctcctca gtctgacatc tgaagactct 720

gcagtctatt actgtgcaag atctgggtac tatggtgact cggactggta cttcgatgtc 780

tggggcgcag ggaccacggt caccgtctcc tcgagcgagc ccaaatcttc tgacaaaact 840

cacacatgcc caccgtgccc agcacctgaa ctcctgggtg gaccgtcagt cttcctcttc 900

cccccaaaac ccaaggacac cctcatgatc tcccggaccc ctgaggtcac atgcgtggtg 960

gtggacgtga gccacgaaga ccctgaggtc aagttcaact ggtacgtgga cggcgtggag 1020

gtgcataatg ccaagacaaa gccgcgggag gagcagtaca acagcacgta ccgtgtggtc 1080

agcgtcctca ccgtcctgca ccaggactgg ctgaatggca aggagtacaa gtgcaaggtc 1140

tccaacaaag ccctcccagc ccccatcgag aaaaccatct ccaaagccaa agggcagccc 1200

cgagaaccac aggtgtacac cctgccccca tcccgggatg agctgaccaa gaaccaggtc 1260

agcctgacct gcctggtcaa aggcttctat ccaagcgaca tcgccgtgga gtgggagagc 1320

aatgggcagc cggagaacaa ctacaagacc acgcctcccg tgctggactc cgacggctcc 1380

ttcttcctct acagcaagct caccgtggac aagagcaggt ggcagcaggg gaacgtcttc 1440

tcatgctccg tgatgcatga ggctctgcac aaccactaca cgcagaagag cctctccctg 1500

tctccgggta agtgactcta ga 1522

<210> 209
<211> 500
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G19-4 LH SMIP (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(145)
<223> 连接子

<220>
<221> misc_feature
<222> (146)..(267)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<400> 209

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser
20 25 30

Ala Ser Leu Gly Asp Arg Val Thr Ile Ser Cys Arg Ala Ser Gln Asp
35 40 45

Ile Arg Asn Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Asp Gly Thr Val
50 55 60

Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val Pro Ser
65 70 75 80

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ala
85 90 95

Asn Leu Gln Pro Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly Asn
100 105 110

Thr Leu Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Val Thr Lys Arg
115 120 125

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala
130 135 140

Ser Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly
145 150 155 160

Ala Ser Met Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly
165 170 175

Tyr Ile Val Asn Trp Leu Lys Gln Ser His Gly Lys Asn Leu Glu Trp
180 185 190

Ile Gly Leu Ile Asn Pro Tyr Lys Gly Leu Thr Thr Tyr Asn Gln Lys
195 200 205

Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala
210 215 220

Tyr Met Glu Leu Leu Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr
225 230 235 240

Cys Ala Arg Ser Gly Tyr Tyr Gly Asp Ser Asp Trp Tyr Phe Asp Val
245 250 255

Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ser Ser Glu Pro Lys Ser
260 265 270

Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu
275 280 285

Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu
290 295 300

Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser
305 310 315 320

His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu
325 330 335

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr
340 345 350

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn
355 360 365

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro
370 375 380

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln
385 390 395 400

Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val
405 410 415

Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val
420 425 430

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro
435 440 445

Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr
450 455 460

Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val
465 470 475 480

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
485 490 495

Ser Pro Gly Lys
500

<210> 210
<211> 1525
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> G19-4 HL SMIP (DNA)

<400> 210
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtgaggtcca gctgcaacag tctggacctg aactggtgaa gcctggagct 120

tcaatgaaga tttcctgcaa ggcctctggt tactcattca ctggctacat cgtgaactgg 180

ctgaagcaga gccatggaaa gaaccttgag tggattggac ttattaatcc atacaaaggt 240

cttactacct acaaccagaa attcaagggc aaggccacat tgactgtaga caagtcatcc 300

agcacagcct acatggagct cctcagtctg acatctgaag actctgcagt ctattactgt 360

gcaagatctg ggtactatgg tgactcagac tggtacttcg atgtctgggg cgcagggacc 420

acagtcaccg tctcctctgg cggcggtggc tcgggcggtg gtggatctgg aggaggtggg 480

agcgctagcg acatccagat gacacagact acatcctccc tgtctgcctc tctgggagac 540

agagtcacca tcagttgcag ggcaagtcag gacattcgca attatttgaa ctggtatcag 600

cagaaaccag atggaactgt taaactcctg atctactaca catcaagact acactcagga 660

gtcccatcaa ggttcagtgg cagtgggtct ggaacagatt attctctcac cattgccaac 720

ctgcaaccag aagatattgc cacttacttt tgccaacagg gtaatacact tccctggacg 780

ttcggtggag ggaccaaact ggtgaccaaa cgctcgagcg agcccaaatc ttctgacaaa 840

actcacacat gcccaccgtg cccagcacct gaactcctgg gtggaccgtc agtcttcctc 900

ttcccccaa aacccaagga caccctcatg atctcccgga cccctgaggt cacatgcgtg 960

gtggtggacg tgagccacga agaccctgag gtcaagttca actggtacgt ggacggcgtg 1020

gaggtgcata atgccaagac aaagccgcgg gaggagcagt acaacagcac gtaccgtgtg 1080

gtcagcgtcc tcaccgtcct gcaccaggac tggctgaatg gcaaggagta caagtgcaag 1140

gtctccaaca aagccctccc agcccccatc gagaaaacca tctccaaagc caaagggcag 1200

ccccgagaac cacaggtgta caccctgccc ccatcccggg atgagctgac caagaaccag 1260

gtcagcctga cctgcctggt caaaggcttc tatccaagcg acatcgccgt ggagtgggag 1320

agcaatgggc agccggagaa caactacaag accacgcctc ccgtgctgga ctccgacggc 1380

tccttcttcc tctacagcaa gctcaccgtg gacaagagca ggtggcagca ggggaacgtc 1440

ttctcatgct ccgtgatgca tgaggctctg cacaaccact acacgcagaa gagcctctcc 1500

ctgtctccgg gtaagtgact ctaga 1525

<210> 211
<211> 501
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> G19-4 HL SMIP (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(142)
<223> VH

<220>
<221> misc_feature
<222> (143)..(159)
<223> 连接子

<220>
<221> misc_feature
<222> (160)..(267)
<223> VL

<220>
<221> misc_feature
<222> (270)..(284)
<223> 铰链

<400> 211

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val
20 25 30

Lys Pro Gly Ala Ser Met Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser
35 40 45

Phe Thr Gly Tyr Ile Val Asn Trp Leu Lys Gln Ser His Gly Lys Asn
50 55 60

Leu Glu Trp Ile Gly Leu Ile Asn Pro Tyr Lys Gly Leu Thr Thr Tyr
65 70 75 80

Asn Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser
85 90 95

```
Ser Thr Ala Tyr Met Glu Leu Leu Ser Leu Thr Ser Glu Asp Ser Ala
            100                 105                 110

Val Tyr Tyr Cys Ala Arg Ser Gly Tyr Tyr Gly Asp Ser Asp Trp Tyr
        115                 120                 125

Phe Asp Val Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ser Gly Gly
    130                 135                 140

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser Asp
145                 150                 155                 160

Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly Asp
                165                 170                 175

Arg Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Arg Asn Tyr Leu
            180                 185                 190

Asn Trp Tyr Gln Gln Lys Pro Asp Gly Thr Val Lys Leu Leu Ile Tyr
        195                 200                 205

Tyr Thr Ser Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser
    210                 215                 220

Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ala Asn Leu Gln Pro Glu
225                 230                 235                 240

Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly Asn Thr Leu Pro Trp Thr
                245                 250                 255

Phe Gly Gly Gly Thr Lys Leu Val Thr Lys Arg Ser Ser Glu Pro Lys
            260                 265                 270

Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu
        275                 280                 285

Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
    290                 295                 300

Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
305                 310                 315                 320

Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val
                325                 330                 335

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser
            340                 345                 350

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
        355                 360                 365
```

```
Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala
    370                 375                 380

Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
385                 390                 395                 400

Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln
                405                 410                 415

Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
            420                 425                 430

Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
        435                 440                 445

Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu
    450                 455                 460

Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser
465                 470                 475                 480

Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
                485                 490                 495

Leu Ser Pro Gly Lys
            500
```

<210> 212
<211> 2328
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-STD1-2e12HL (DNA)

<400> 212

```
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca     60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg    120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag    180

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggcttctgga    240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga    300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg    360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct    420
```

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggacccct gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

aattatggtg gcggtggctc gggcggtggt ggatctggag gaggtgggag tgggaattct 1560

caggtgcagc tgaaggagtc aggacctggc ctggtggcgc cctcacagag cctgtccatc 1620

acatgcaccg tctcagggtt ctcattaacc ggctatggtg taaactgggt tcgccagcct 1680

ccaggaaagg gtctggagtg gctgggaatg atatggggtg atggaagcac agactataat 1740

tcagctctca aatccagact atcgatcacc aaggacaact ccaagagcca agttttctta 1800

aaaatgaaca gtctgcaaac tgatgacaca gccagatact actgtgctcg agatggttat 1860

agtaactttc attactatgt tatggactac tggggtcaag gaacctcagt caccgtctcc 1920

tctgggggtg gaggctctgg tggcggtgga tccggcggag gtgggtcggg tggcggcgga 1980

tctgacattg tgctcaccca atctccagct tctttggctg tgtctctagg tcagagagcc 2040

accatctcct gcagagccag tgaaagtgtt gaatattatg tcacaagttt aatgcagtgg 2100

taccaacaga aaccaggaca gccacccaaa ctcctcatct ctgctgctag caacgtagaa 2160

tctggggtcc ctgccaggtt tagtggcagt gggtctggga cagactttag cctcaacatc 2220

catcctgtgg aggaggatga tattgcaatg tatttctgtc agcaaagtag gaaggttcca 2280

tggacgttcg gtggaggcac caagctggaa atcaaacgtt aatctaga 2328

<210> 213
<211> 769
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-STD1-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(127)
<223> VL

<220>
<221> misc_feature
<222> (128)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>
<221> misc_feature
<222> (497)..(516)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (517)..(637)
<223> VH2

<220>
<221> misc_feature
<222> (638)..(657)
<223> 连接子 2

<220>
<221> misc_feature
<222> (658)..(769)
<223> VL2

<400> 213

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
20 25 30

```
Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
        35                  40                  45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
    50                  55                  60

Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65                  70                  75                  80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
                85                  90                  95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
            100                 105                 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
        115                 120                 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
    130                 135                 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145                 150                 155                 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
                165                 170                 175

His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
            180                 185                 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
        195                 200                 205

Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
    210                 215                 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225                 230                 235                 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
                245                 250                 255

Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
            260                 265                 270

His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
        275                 280                 285
```

```
Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
    290                 295                 300

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
            500                 505                 510

Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val
        515                 520                 525

Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser
    530                 535                 540
```

```
Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys Gly
545                 550                 555                 560

Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr Asn
                565                 570                 575

Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys Ser
            580                 585                 590

Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Arg
        595                 600                 605

Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val Met
    610                 615                 620

Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly
625                 630                 635                 640

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
                645                 650                 655

Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu
            660                 665                 670

Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser Val Glu Tyr
        675                 680                 685

Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln Pro
    690                 695                 700

Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser Gly Val Pro
705                 710                 715                 720

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn Ile
                725                 730                 735

His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln Ser
            740                 745                 750

Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
        755                 760                 765

Arg
```

<210> 214
<211> 2313
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-STD2-2e12LH (DNA)

<400> 214
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg 120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag 180

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggctctgga 240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga 300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg 360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct 420

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggacccct gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

aattatggtg gcggtggctc gggcggtggt ggatctggag gaggtgggag tgggaattct 1560

gacattgtgc tcacccaatc tccagcttct ttggctgtgt ctctaggtca gagagccacc 1620

atctcctgca gagccagtga aagtgttgaa tattatgtca caagtttaat gcagtggtac 1680

caacagaaac caggacagcc acccaaactc ctcatctctg ctgctagcaa cgtagaatct 1740

ggggtccctg ccaggtttag tggcagtggg tctgggacag actttagcct caacatccat 1800

cctgtggagg aggatgatat tgcaatgtat ttctgtcagc aaagtaggaa ggttccatgg 1860

acgttcggtg gaggcaccaa gctggaaatc aaacggggtg gcggtggatc cggcggaggt 1920

gggtcgggtg gcggcggatc tcaggtgcag ctgaaggagt caggacctgg cctggtggcg 1980

ccctcacaga gcctgtccat cacatgcacc gtctcagggt tctcattaac cggctatggt 2040

gtaaactggg ttcgccagcc tccaggaaag ggtctggagt ggctgggaat gatatggggt 2100

gatggaagca cagactataa ttcagctctc aaatccagac tatcgatcac caaggacaac 2160

tccaagagcc aagttttctt aaaaatgaac agtctgcaaa ctgatgacac agccagatac 2220

tactgtgctc gagatggtta tagtaacttt cattactatg ttatggacta ctggggtcaa 2280

ggaacctcag tcaccgtctc tcttaatct aga 2313

<210> 215
<211> 764
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-STD2-2e12LH (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(126)
<223> VL

<220>
<221> misc_feature
<222> (127)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>
<221> misc_feature
<222> (497)..(516)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (517)..(628)
<223> VL2

<220>
<221> misc_feature
<222> (629)..(643)
<223> 连接子

<220>
<221> misc_feature
<222> (644)..(764)
<223> VH2

<400> 215

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
20 25 30

Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
35 40 45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
50 55 60

Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65 70 75 80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
85 90 95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
100 105 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
130 135 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145 150 155 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
165 170 175

His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
180 185 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
195 200 205

Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
210 215 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225 230 235 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
245 250 255

Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
260 265 270

His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
275 280 285

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
290 295 300

```
Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
            500                 505                 510

Ser Gly Asn Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala
        515                 520                 525

Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser
    530                 535                 540

Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro
545                 550                 555                 560

Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser
                565                 570                 575
```

```
Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser
            580                 585                 590

Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys
        595                 600                 605

Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu
    610                 615                 620

Glu Ile Lys Arg Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
625                 630                 635                 640

Gly Gly Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val Ala
                645                 650                 655

Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu
            660                 665                 670

Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu
        675                 680                 685

Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser
    690                 695                 700

Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln
705                 710                 715                 720

Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr
                725                 730                 735

Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val Met Asp
            740                 745                 750

Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
        755                 760
```

<210> 216
<211> 2048
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-H1-2e12HL (DNA)

<400> 216
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg 120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag 180

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggcttctgga 240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga 300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg 360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct 420

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggaccccc gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

aattctcagg tgcagctgaa ggagtcagga cctggcctgg tggcgccctc acagagcctg 1560

tccatcacat gcaccgtctc agggttctca ttaaccggct atggtgtaaa ctgggttcgc 1620

cagcctccag gaaagggtct ggagtggctg ggaatgatat ggggtgatgg aagcacagac 1680

tataattcag ctctcaaatc cagactatcg atcaccaagg acaactccaa gagccaagtt 1740

ttcttaaaaa tgaacagtct gcaaactgat gacacagcca gatactactg tgctcgagat 1800

ggttatagta actttcatta ctatgttatg gactactggg gtcaaggaac ctcagtcacc 1860

gtctcctctg ggggtggagg ctctggtggc ggtggatccg gcggaggtgg gtcgggtggc 1920

ggcggatctg acattgtgct cacccaatct ccagcttctt tggctgtgtc tctaggtcag 1980

agagccacca tctcctgcag agccagtgaa agtgttgaat attatgtcac aagtttaatg 2040

cagtggta 2048

<210> 217
<211> 751
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-H1-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(126)
<223> VL

<220>
<221> misc_feature
<222> (127)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>
<221> misc_feature
<222> (497)..(498)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (499)..(619)
<223> VH2

<220>
<221> misc_feature
<222> (620)..(639)
<223> 连接子 2

<220>
<221> misc_feature
<222> (640)..(751)
<223> VL2

<400> 217
Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

```
Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
            20                  25                  30

Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
        35                  40                  45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
    50                  55                  60

Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65                  70                  75                  80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
                85                  90                  95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
            100                 105                 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
        115                 120                 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
    130                 135                 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145                 150                 155                 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
                165                 170                 175

His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
            180                 185                 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
        195                 200                 205

Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
    210                 215                 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225                 230                 235                 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
                245                 250                 255

Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
            260                 265                 270
```

```
His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
        275                 280                 285

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
    290                 295                 300

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro
            500                 505                 510

Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr
        515                 520                 525
```

```
Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu
    530                 535                 540

Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser Ala
545                 550                 555                 560

Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln Val
                565                 570                 575

Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr Tyr
            580                 585                 590

Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val Met Asp Tyr
        595                 600                 605

Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser
    610                 615                 620

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp
625                 630                 635                 640

Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly Gln
                645                 650                 655

Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val
            660                 665                 670

Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys
        675                 680                 685

Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser Gly Val Pro Ala Arg
    690                 695                 700

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn Ile His Pro
705                 710                 715                 720

Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys
                725                 730                 735

Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
            740                 745                 750
```

<210> 218
<211> 2292
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-H2-2e12HL (DNA)

<400> 218
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg 120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag 180

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggcttctgga 240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga 300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg 360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct 420

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggacccct gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

ggtggcggtg gctcggggaa ttctcaggtg cagctgaagg agtcaggacc tggcctggtg 1560

gcgccctcac agagcctgtc catcacatgc accgtctcag ggttctcatt aaccggctat 1620

ggtgtaaact gggttcgcca gcctccagga aagggtctgg agtggctggg aatgatatgg 1680

ggtgatggaa gcacagacta taattcagct ctcaaatcca gactatcgat caccaaggac 1740

aactccaaga gccaagtttt cttaaaaatg aacagtctgc aaactgatga cacagccaga 1800

tactactgtg ctcgagatgg ttatagtaac tttcattact atgttatgga ctactggggt 1860

caaggaacct cagtcaccgt ctcctctggg ggtggaggct ctggtggcgg tggatccggc 1920

ggaggtgggt cgggtggcgg cggatctgac attgtgctca cccaatctcc agcttctttg 1980

gctgtgtctc taggtcagag agccaccatc tcctgcagag ccagtgaaag tgttgaatat 2040

tatgtcacaa gtttaatgca gtggtaccaa cagaaaccag gacagccacc caaactcctc 2100

atctctgctg ctagcaacgt agaatctggg gtccctgcca ggtttagtgg cagtgggtct 2160

gggacagact ttagcctcaa catccatcct gtggaggagg atgatattgc aatgtatttc 2220

tgtcagcaaa gtaggaaggt tccatggacg ttcggtggag gcaccaagct ggaaatcaaa 2280

cgttaatcta ga 2292

<210> 219
<211> 757
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-H2-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(126)
<223> VL

<220>
<221> misc_feature
<222> (127)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>
<221> misc_feature
<222> (497)..(504)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (505)..(625)
<223> VH2

<220>
<221> misc_feature
<222> (626)..(645)
<223> 连接子 2

<220>
<221> misc_feature
<222> (646)..(757)
<223> VL2

<400> 219

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
20 25 30

Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
35 40 45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
50 55 60

Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65 70 75 80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
85 90 95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
100 105 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
130 135 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145 150 155 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
165 170 175

His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
180 185 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
195 200 205

```
Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
    210                 215                 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225                 230                 235                 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
                245                 250                 255

Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
            260                 265                 270

His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
        275                 280                 285

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
    290                 295                 300

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460
```

```
Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Gly Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser Gly
            500                 505                 510

Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val
        515                 520                 525

Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro
    530                 535                 540

Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser
545                 550                 555                 560

Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp
                565                 570                 575

Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp
            580                 585                 590

Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His
        595                 600                 605

Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser
    610                 615                 620

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu
                645                 650                 655

Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu
            660                 665                 670

Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys
        675                 680                 685

Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu
    690                 695                 700

Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
705                 710                 715                 720
```

Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe
725 730 735

Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys
740 745 750

Leu Glu Ile Lys Arg
755

<210> 220
<211> 2298
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-H3-2e12HL (DNA)

<400> 220
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg 120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag 180

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggcttctgga 240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga 300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg 360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct 420

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggacccct gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

aattatggtg gcggtggctc tgggaattct caggtgcagc tgaaggagtc aggacctggc 1560

ctggtggcgc cctcacagag cctgtccatc acatgcaccg tctcagggtt ctcattaacc 1620

ggctatggtg taaactgggt tcgccagcct ccaggaaagg gtctggagtg gctgggaatg 1680

atatggggtg atggaagcac agactataat tcagctctca aatccagact atcgatcacc 1740

aaggacaact ccaagagcca agttttctta aaaatgaaca gtctgcaaac tgatgacaca 1800

gccagatact actgtgctcg agatggttat agtaactttc attactatgt tatggactac 1860

tggggtcaag gaacctcagt caccgtctcc tctgggggtg gaggctctgg tggcggtgga 1920

tccggcggag gtgggtcggg tggcggcgga tctgacattg tgctcaccca atctccagct 1980

tctttggctg tgtctctagg tcagagagcc accatctcct gcagagccag tgaaagtgtt 2040

gaatattatg tcacaagttt aatgcagtgg taccaacaga aaccaggaca gccacccaaa 2100

ctcctcatct ctgctgctag caacgtagaa tctggggtcc ctgccaggtt tagtggcagt 2160

gggtctggga cagactttag cctcaacatc catcctgtgg aggaggatga tattgcaatg 2220

tatttctgtc agcaaagtag gaaggttcca tggacgttcg gtggaggcac caagctggaa 2280

atcaaacgtt aatctaga 2298

<210> 221
<211> 759
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-H3-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(126)
<223> VL

<220>
<221> misc_feature
<222> (127)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>
<221> misc_feature
<222> (497)..(506)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (507)..(627)
<223> VH2

<220>
<221> misc_feature
<222> (628)..(647)
<223> 连接子 2

<220>
<221> misc_feature
<222> (648)..(759)
<223> VL2

<400> 221

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
20 25 30

Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
35 40 45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
50 55 60

Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65 70 75 80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
85 90 95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
100 105 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
115 120 125

```
Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
    130                 135                 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145                 150                 155                 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
                165                 170                 175

His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
            180                 185                 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
        195                 200                 205

Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
    210                 215                 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225                 230                 235                 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
                245                 250                 255

Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
            260                 265                 270

His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
        275                 280                 285

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
    290                 295                 300

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380
```

```
Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Asn Tyr Gly Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu
            500                 505                 510

Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys
        515                 520                 525

Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg
    530                 535                 540

Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp
545                 550                 555                 560

Gly Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr
                565                 570                 575

Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln
            580                 585                 590

Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn
        595                 600                 605

Phe His Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr
    610                 615                 620

Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
625                 630                 635                 640

Gly Ser Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala
                645                 650                 655
```

Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala
660 665 670

Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln
675 680 685

Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn
690 695 700

Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr
705 710 715 720

Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met
725 730 735

Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly
740 745 750

Thr Lys Leu Glu Ile Lys Arg
755

<210> 222
<211> 2307
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-H4-2e12HL (DNA)

<400> 222
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg 120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag 180

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggcttctgga 240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga 300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg 360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct 420

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggaccccc gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

ggtggcggtg gctcgggcgg tggtggatct gggaattctc aggtgcagct gaaggagtca 1560

ggacctggcc tggtggcgcc ctcacagagc ctgtccatca catgcaccgt ctcagggttc 1620

tcattaaccg gctatggtgt aaactgggtt cgccagcctc caggaaaggg tctggagtgg 1680

ctgggaatga tatggggtga tggaagcaca gactataatt cagctctcaa atccagacta 1740

tcgatcacca aggacaactc caagagccaa gttttcttaa aaatgaacag tctgcaaact 1800

gatgacacag ccagatacta ctgtgctcga gatggttata gtaactttca ttactatgtt 1860

atggactact ggggtcaagg aacctcagtc accgtctcct ctgggggtgg aggctctggt 1920

ggcggtggat ccggcggagg tgggtcgggt ggcggcggat ctgacattgt gctcacccaa 1980

tctccagctt ctttggctgt gtctctaggt cagagagcca ccatctcctg cagagccagt 2040

gaaagtgttg aatattatgt cacaagttta atgcagtggt accaacagaa accaggacag 2100

ccacccaaac tcctcatctc tgctgctagc aacgtagaat ctggggtccc tgccaggttt 2160

agtggcagtg ggtctgggac agactttagc ctcaacatcc atcctgtgga ggaggatgat 2220

attgcaatgt atttctgtca gcaaagtagg aaggttccat ggacgttcgg tggaggcacc 2280

aagctggaaa tcaaacgtta atctaga 2307

<210> 223
<211> 762
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-H4-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(126)
<223> VL

<220>
<221> misc_feature
<222> (127)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>
<221> misc_feature
<222> (497)..(509)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (510)..(630)
<223> VH2

<220>
<221> misc_feature
<222> (631)..(650)
<223> 连接子 2

<220>
<221> misc_feature
<222> (651)..(762)
<223> VL2

<400> 223

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
20 25 30

Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
35 40 45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
50 55 60

```
Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65                  70                  75                  80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
                85                  90                  95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
            100                 105                 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
        115                 120                 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
    130                 135                 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145                 150                 155                 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
                165                 170                 175

His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
            180                 185                 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
        195                 200                 205

Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
    210                 215                 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225                 230                 235                 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
                245                 250                 255

Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
            260                 265                 270

His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
        275                 280                 285

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
    290                 295                 300

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320
```

```
Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Asn Ser Gln Val Gln
            500                 505                 510

Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser
        515                 520                 525

Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn
    530                 535                 540

Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile
545                 550                 555                 560

Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu
                565                 570                 575

Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn
            580                 585                 590
```

```
Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly
        595                 600                 605

Tyr Ser Asn Phe His Tyr Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr
    610                 615                 620

Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Val Leu Thr Gln
                645                 650                 655

Ser Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser
            660                 665                 670

Cys Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu Met Gln
        675                 680                 685

Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala
    690                 695                 700

Ala Ser Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly Ser Gly
705                 710                 715                 720

Ser Gly Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu Asp Asp
                725                 730                 735

Ile Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp Thr Phe
            740                 745                 750

Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
        755                 760
```

<210> 224
<211> 2313
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-H5-2e12HL (DNA)

<400> 224

```
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca     60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg    120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag    180
```

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggcttctgga 240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga 300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg 360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct 420

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggacccct gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

aattatggtg gcggtggctc gggcggtggt ggatctggga attctcaggt gcagctgaag 1560

gagtcaggac ctggcctggt ggcgccctca cagagcctgt ccatcacatg caccgtctca 1620

gggttctcat taaccggcta tggtgtaaac tgggttcgcc agcctccagg aaagggtctg 1680

gagtggctgg gaatgatatg gggtgatgga agcacagact ataattcagc tctcaaatcc 1740

agactatcga tcaccaagga caactccaag agccaagttt tcttaaaaat gaacagtctg 1800

caaactgatg acacagccag atactactgt gctcgagatg gttatagtaa ctttcattac 1860

tatgttatgg actactgggg tcaaggaacc tcagtcaccg tctcctctgg gggtggaggc 1920

tctggtggcg gtggatccgg cggaggtggg tcgggtggcg gcggatctga cattgtgctc 1980

acccaatctc cagcttcttt ggctgtgtct ctaggtcaga gagccaccat ctcctgcaga 2040

gccagtgaaa gtgttgaata ttatgtcaca agtttaatgc agtggtacca acagaaacca 2100

ggacagccac ccaaactcct catctctgct gctagcaacg tagaatctgg ggtccctgcc 2160

aggtttagtg gcagtgggtc tgggacagac tttagcctca acatccatcc tgtggaggag 2220

gatgatattg caatgtattt ctgtcagcaa agtaggaagg ttccatggac gttcggtgga 2280

ggcaccaagc tggaaatcaa acgttaatct aga 2313

<210> 225
<211> 764
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-H5-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(126)
<223> VL

<220>
<221> misc_feature
<222> (127)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>
<221> misc_feature
<222> (497)..(511)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (512)..(632)
<223> VH2

<220>
<221> misc_feature
<222> (633)..(652)
<223> 连接子 2

<220>
<221> misc_feature
<222> (653)..(764)
<223> VL2

<400> 225

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
20 25 30

Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
35 40 45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
50 55 60

Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65 70 75 80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
85 90 95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
100 105 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
130 135 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145 150 155 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
165 170 175

His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
180 185 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
195 200 205

Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
210 215 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225 230 235 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
245 250 255

```
Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
            260                 265                 270

His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
        275                 280                 285

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
    290                 295                 300

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Asn Ser Gln
            500                 505                 510
```

```
Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser
        515                 520                 525

Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly
    530                 535                 540

Val Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly
545                 550                 555                 560

Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser Ala Leu Lys Ser
                565                 570                 575

Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys
            580                 585                 590

Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg
        595                 600                 605

Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val Met Asp Tyr Trp Gly Gln
    610                 615                 620

Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly
625                 630                 635                 640

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Val Leu
                645                 650                 655

Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg Ala Thr
            660                 665                 670

Ile Ser Cys Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val Thr Ser Leu
        675                 680                 685

Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile
    690                 695                 700

Ser Ala Ala Ser Asn Val Glu Ser Gly Val Pro Ala Arg Phe Ser Gly
705                 710                 715                 720

Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn Ile His Pro Val Glu Glu
                725                 730                 735

Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys Val Pro Trp
            740                 745                 750

Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
        755                 760
```

<210> 226
<211> 2322
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> n2H7sssIgG1-H6-2e12HL (DNA)

<400> 226
aagcttgccg ccatggaagc accagcgcag cttctcttcc tcctgctact ctggctccca 60

gataccaccg gtcaaattgt tctctcccag tctccagcaa tcctgtctgc atctccaggg 120

gagaaggtca caatgacttg cagggccagc tcaagtgtaa gttacatgca ctggtaccag 180

cagaagccag gatcctcccc caaaccctgg atttatgccc catccaacct ggcttctgga 240

gtccctgctc gcttcagtgg cagtgggtct gggacctctt actctctcac aatcagcaga 300

gtggaggctg aagatgctgc cacttattac tgccagcagt ggagttttaa cccacccacg 360

ttcggtgctg ggaccaagct ggagctgaaa gatggcggtg gctcgggcgg tggtggatct 420

ggaggaggtg gagctagcca ggcttatcta cagcagtctg gggctgagct ggtgaggcct 480

ggggcctcag tgaagatgtc ctgcaaggct tctggctaca catttaccag ttacaatatg 540

cactgggtaa agcagacacc tagacagggc ctggaatgga ttggagctat ttatccagga 600

aatggtgata cttcctacaa tcagaagttc aagggcaagg ccacactgac tgtagacaaa 660

tcctccagca cagcctacat gcagctcagc agcctgacat ctgaagactc tgcggtctat 720

ttctgtgcaa gagtggtgta ctatagtaac tcttactggt acttcgatgt ctggggcaca 780

gggaccacgg tcaccgtctc gagcgagccc aaatcttctg acaaaactca cacatcccca 840

ccgagcccag cacctgaact cctgggggga ccgtcagtct tcctcttccc cccaaaaccc 900

aaggacaccc tcatgatctc ccggacccct gaggtcacat gcgtggtggt ggacgtgagc 960

cacgaagacc ctgaggtcaa gttcaactgg tacgtggacg gcgtggaggt gcataatgcc 1020

aagacaaagc cgcgggagga gcagtacaac agcacgtacc gtgtggtcag cgtcctcacc 1080

gtcctgcacc aggactggct gaatggcaag gagtacaagt gcaaggtctc caacaaagcc 1140

ctcccagccc ccatcgagaa aacaatctcc aaagccaaag ggcagccccg agaaccacag 1200

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 1260

ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 1320

gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 1380

agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 1440

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaag 1500

ggtggcggtg gctcgggcgg tggtggatct gggggaggag gcagcgggaa ttctcaggtg 1560

cagctgaagg agtcaggacc tggcctggtg gcgccctcac agagcctgtc catcacatgc 1620

accgtctcag ggttctcatt aaccggctat ggtgtaaact gggttcgcca gcctccagga 1680

aagggtctgg agtggctggg aatgatatgg ggtgatggaa gcacagacta taattcagct 1740

ctcaaatcca gactatcgat caccaaggac aactccaaga gccaagtttt cttaaaaatg 1800

aacagtctgc aaactgatga cacagccaga tactactgtg ctcgagatgg ttatagtaac 1860

tttcattact atgttatgga ctactggggt caaggaacct cagtcaccgt ctcctctggg 1920

ggtggaggct ctggtggcgg tggatccggc ggaggtgggt cgggtggcgg cggatctgac 1980

attgtgctca cccaatctcc agcttctttg gctgtgtctc taggtcagag agccaccatc 2040

tcctgcagag ccagtgaaag tgttgaatat tatgtcacaa gtttaatgca gtggtaccaa 2100

cagaaaccag gacagccacc caaactcctc atctctgctg ctagcaacgt agaatctggg 2160

gtccctgcca ggtttagtgg cagtgggtct gggacagact ttagcctcaa catccatcct 2220

gtggaggagg atgatattgc aatgtatttc tgtcagcaaa gtaggaaggt tccatggacg 2280

ttcggtggag gcaccaagct ggaaatcaaa cgttaatcta ga 2322

<210> 227
<211> 767
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> n2H7sssIgG1-H6-2e12HL (AA)

<220>
<221> misc_feature
<222> (1)..(20)
<223> 前导序列

<220>
<221> misc_feature
<222> (21)..(126)
<223> VL

<220>
<221> misc_feature
<222> (127)..(142)
<223> 连接子

<220>
<221> misc_feature
<222> (143)..(264)
<223> VH

<220>
<221> misc_feature
<222> (265)..(279)
<223> 铰链

<220>

<221> misc_feature
<222> (497)..(514)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (515)..(635)
<223> VH2

<220>
<221> misc_feature
<222> (636)..(655)
<223> 连接子 2

<220>
<221> misc_feature
<222> (656)..(767)
<223> VL2

<400> 227

Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro
1 5 10 15

Asp Thr Thr Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile Leu Ser
20 25 30

Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser Ser Ser
35 40 45

Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys
50 55 60

Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro Ala Arg
65 70 75 80

Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Arg
85 90 95

Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Phe
100 105 110

Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Asp Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ala Ser Gln Ala
130 135 140

Tyr Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ala Ser Val
145 150 155 160

Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asn Met
165 170 175

```
His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile Gly Ala
            180                 185                 190

Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys Gly
        195                 200                 205

Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
    210                 215                 220

Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
225                 230                 235                 240

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp Gly Thr
                245                 250                 255

Gly Thr Thr Val Thr Val Ser Ser Glu Pro Lys Ser Ser Asp Lys Thr
            260                 265                 270

His Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
        275                 280                 285

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
    290                 295                 300

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
305                 310                 315                 320

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
                325                 330                 335

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
            340                 345                 350

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
        355                 360                 365

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
    370                 375                 380

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
385                 390                 395                 400

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
                405                 410                 415

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
            420                 425                 430
```

```
Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
        435                 440                 445

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
    450                 455                 460

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
465                 470                 475                 480

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                485                 490                 495

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro
        515                 520                 525

Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr
    530                 535                 540

Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu
545                 550                 555                 560

Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr Asp Tyr Asn Ser Ala
                565                 570                 575

Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn Ser Lys Ser Gln Val
            580                 585                 590

Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Arg Tyr Tyr
        595                 600                 605

Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr Tyr Val Met Asp Tyr
    610                 615                 620

Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp
                645                 650                 655

Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly Gln
            660                 665                 670

Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser Val Glu Tyr Tyr Val
        675                 680                 685
```

Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys
690 695 700

Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser Gly Val Pro Ala Arg
705 710 715 720

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser Leu Asn Ile His Pro
725 730 735

Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys Gln Gln Ser Arg Lys
740 745 750

Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
755 760 765

<210> 228
<211> 2337
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 2H7sssIgG1-H7-G281 HL (DNA)

<400> 228
aagcttgccg ccatggattt tcaagtgcag attttcagct tcctgctaat cagtgcttca 60

gtcataatgt ccagaggaca aattgttctc tcccagtctc cagcaatcct gtctgcatct 120

ccaggggaga aggtcacaat gacttgcagg gccagctcaa gtgtaagtta catgcactgg 180

taccagcaga agccaggatc ctcccccaaa ccctggattt atgccccatc caacctggct 240

tctggagtcc ctgctcgctt cagtggcagt gggtctggga cctcttactc tctcacaatc 300

agcagagtgg aggctgaaga tgctgccact tattactgcc agcagtggag ttttaaccca 360

cccacgttcg gtgctgggac caagctggag ctgaaagatg gcggtggctc gggcggtggt 420

ggatctggag gaggtgggag ctctcaggct tatctacagc agtctggggc tgagtcggtg 480

aggcctgggg cctcagtgaa gatgtcctgc aaggcttctg gctacacatt taccagttac 540

aatatgcact gggtaaagca gacacctaga cagggcctgg aatggattgg agctatttat 600

ccaggaaatg gtgatacttc ctacaatcag aagttcaagg gcaaggccac actgactgta 660

gacaaatcct ccagcacagc ctacatgcag ctcagcagcc tgacatctga agactctgcg 720

gtctatttct gtgcaagagt ggtgtactat agtaactctt actggtactt cgatgtctgg 780

ggcacaggga ccacggtcac cgtctctgat caggagccca aatcttgtga caaaactcac 840

acatccccac cgtgctcagc acctgaactc ctgggggggac cgtcagtctt cctcttcccc 900

ccaaaaccca aggacaccct catgatctcc cggacccctg aggtcacatg cgtggtggtg 960

gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt acgtggacgg cgtggaggtg 1020

cataatgcca agacaaagcc gcgggaggag cagtacaaca gcacgtaccg tgtggtcagc 1080

gtcctcaccg tcctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc 1140

aacaaagccc tcccagcccc catcgagaaa acaatctcca aagccaaagg gcagccccga 1200

gaaccacagg tgtacaccct gcccccatcc cgggatgagc tgaccaagaa ccaggtcagc 1260

ctgacctgcc tggtcaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat 1320

gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga cggctccttc 1380

ttcctctaca gcaagctcac cgtggacaag agcaggtggc agcaggggaa cgtcttctca 1440

tgctccgtga tgcatgaggc tctgcacaac cactacacgc agaagagcct ctccctgtct 1500

ccgggtaaga attatggtgg cggtggctcg ggcggtggtg gatctggagg aggtgggagt 1560

gggaattctc aggtgcagct gaaggagtca ggacctggcc tggtggcgcc ctcacagagc 1620

ctgtccatca catgcaccgt ctcagggttc tcattaaccg gctatggtgt aaactgggtt 1680

cgccagcctc caggaaaggg tctggagtgg ctgggaatga tatggggtga tggaagcaca 1740

gactataatt cagctctcaa atccagacta tcgatcacca aggacaactc caagagccaa 1800

gttttcttaa aaatgaacag tctgcaaact gatgacacag ccagatacta ctgtgctcga 1860

gatggttata gtaactttca ttactatgtt atggactact ggggtcaagg aacctcagtc 1920

accgtctcct ctgggggtgg aggctctggt ggcggtggat ccggcggagg tgggtcgggt 1980

ggcggcggat ctgacattgt gctcacccaa tctccagctt ctttggctgt gtctctaggt 2040

cagagagcca ccatctcctg cagagccagt gaaagtgttg aatattatgt cacaagttta 2100

atgcagtggt accaacagaa accaggacag ccacccaaac tcctcatctc tgctgctagc 2160

aacgtagaat ctggggtccc tgccaggttt agtggcagtg ggtctgggac agactttagc 2220

ctcaacatcc atcctgtgga ggaggatgat attgcaatgt atttctgtca gcaaagtagg 2280

aaggttccat ggacgttcgg tggaggcacc aagctggaaa tcaaacgtta atctaga 2337

<210> 229
<211> 772
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 2H7cscIgG1-STD1-2e12HL (w/2E12 前导序列) (AA)

<220>
<221> misc_feature
<222> (1)..(22)
<223> 前导序列

<220>
<221> misc_feature

<222> (23)..(128)
<223> VL

<220>
<221> misc_feature
<222> (129)..(144)
<223> 连接子

<220>
<221> misc_feature
<222> (145)..(265)
<223> VH

<220>
<221> misc_feature
<222> (268)..(282)
<223> 铰链

<220>
<221> misc_feature
<222> (500)..(519)
<223> EFD-BD2 连接子

<220>
<221> misc_feature
<222> (520)..(640)
<223> VH2

<220>
<221> misc_feature
<222> (641)..(660)
<223> 连接子 2

<220>
<221> misc_feature
<222> (661)..(772)
<223> VL2

<400> 229

Met Asp Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser
1 5 10 15

Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
20 25 30

Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
35 40 45

Ser Ser Val Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Ser Ser
50 55 60

Pro Lys Pro Trp Ile Tyr Ala Pro Ser Asn Leu Ala Ser Gly Val Pro
65 70 75 80

Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
85 90 95

```
Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
            100                 105                 110

Ser Phe Asn Pro Pro Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
        115                 120                 125

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
    130                 135                 140

Gln Ala Tyr Leu Gln Gln Ser Gly Ala Glu Ser Val Arg Pro Gly Ala
145                 150                 155                 160

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
                165                 170                 175

Asn Met His Trp Val Lys Gln Thr Pro Arg Gln Gly Leu Glu Trp Ile
            180                 185                 190

Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe
        195                 200                 205

Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
    210                 215                 220

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
225                 230                 235                 240

Ala Arg Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val Trp
                245                 250                 255

Gly Thr Gly Thr Thr Val Thr Val Ser Asp Gln Glu Pro Lys Ser Cys
            260                 265                 270

Asp Lys Thr His Thr Ser Pro Pro Cys Ser Ala Pro Glu Leu Leu Gly
        275                 280                 285

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
    290                 295                 300

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
305                 310                 315                 320

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
                325                 330                 335

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
            340                 345                 350

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
        355                 360                 365
```

```
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
    370                 375                 380

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
385                 390                 395                 400

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
                405                 410                 415

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
            420                 425                 430

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
        435                 440                 445

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
    450                 455                 460

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
465                 470                 475                 480

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
                485                 490                 495

Pro Gly Lys Asn Tyr Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
            500                 505                 510

Gly Gly Gly Ser Gly Asn Ser Gln Val Gln Leu Lys Glu Ser Gly Pro
        515                 520                 525

Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser
    530                 535                 540

Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro
545                 550                 555                 560

Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly Asp Gly Ser Thr
                565                 570                 575

Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile Thr Lys Asp Asn
            580                 585                 590

Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp
        595                 600                 605

Thr Ala Arg Tyr Tyr Cys Ala Arg Asp Gly Tyr Ser Asn Phe His Tyr
    610                 615                 620
```

```
Tyr Val Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
625                 630                 635                 640

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
                645                 650                 655

Gly Gly Gly Ser Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala
            660                 665                 670

Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Glu Ser
        675                 680                 685

Val Glu Tyr Tyr Val Thr Ser Leu Met Gln Trp Tyr Gln Gln Lys Pro
    690                 695                 700

Gly Gln Pro Pro Lys Leu Leu Ile Ser Ala Ala Ser Asn Val Glu Ser
705                 710                 715                 720

Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Ser
                725                 730                 735

Leu Asn Ile His Pro Val Glu Glu Asp Asp Ile Ala Met Tyr Phe Cys
            740                 745                 750

Gln Gln Ser Arg Lys Val Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu
        755                 760                 765

Glu Ile Lys Arg
    770
```

<210> 230
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> scs(s) -hIgG1 (DNA)

<400> 230
gagcccaaat cttctgacaa aactcacaca tctccaccga gctca 45

<210> 231
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>

<223> scs(s) -hIgG1 (AA)

<400> 231

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Ser Pro Pro Ser Ser
1 5 10 15

<210> 232
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> ccc(s) (DNA)

<400> 232
gagcccaaat cttgtgacaa aactcacaca tctccaccgt gctca 45

<210> 233
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> ccc(s) (AA)

<400> 233

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Ser Pro Pro Cys Ser
1 5 10 15

<210> 234
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H8 (PN)

<400> 234
gggtctccac cttctccgaa ttct 24

<210> 235
<211> 8
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<400> 235

Gly Ser Pro Pro Ser Pro Asn Ser
1 5

<210> 236
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H9 (PN)

<400> 236
tctccacctt ctccgaattc t 21

<210> 237
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H9 (AA)

<400> 237

Ser Pro Pro Ser Pro Asn Ser
1 5

<210> 238

<400> 238
000

<210> 239

<400> 239
000

<210> 240
<211> 51
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H11 (PN)

<400> 240
gagcccacat ctaccgacaa aactcacaca tctccaccca gcccgaattc t 51

<210> 241
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H11 (AA)

<400> 241

Glu Pro Thr Ser Thr Asp Lys Thr His Thr Ser Pro Pro Ser Pro Asn
1 5 10 15

Ser

<210> 242
<211> 51
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H12 (PN)

<400> 242
gagcccacat ctaccgacaa aactcacaca tctccaccca gcccgaattc t 51

<210> 243
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H12 (AA)

<400> 243

Glu Pro Thr Ser Thr Asp Lys Thr His Thr Cys Pro Pro Cys Pro Asn
1 5 10 15

Ser

<210> 244
<211> 51
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>

<223> 连接子 H13 (PN)

<400> 244
gagcccacat ctaccgacaa aactcacaca tctccaccca gcccgaattc t 51

<210> 245
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H13 (AA)

<400> 245

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Asn
1 5 10 15

Ser

<210> 246
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H15 (PN)

<400> 246
ggcggtggtg gctcctgtcc accttgtccg aattct 36

<210> 247
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H15 (AA)

<400> 247

Gly Gly Gly Gly Ser Cys Pro Pro Cys Pro Asn Ser
1 5 10

<210> 248
<211> 45
<212> DNA
<213> 人工序列

<220>

<223> 合成聚核苷酸

<220>
<223> 连接子 H16 (PN)

<400> 248
ctgtctgtga aagctgactt cctcactcca tccatcggga attct 45

<210> 249
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H16 (AA)

<400> 249

Leu Ser Val Lys Ala Asp Phe Leu Thr Pro Ser Ile Gly Asn Ser
1 5 10 15

<210> 250
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H17 (PN)

<400> 250
ctgtctgtga aagctgactt cctcactcca tccatctcct gtccaccttg cccgaattct 60

<210> 251
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H17 (AA)

<400> 251

Leu Ser Val Lys Ala Asp Phe Leu Thr Pro Ser Ile Ser Cys Pro Pro
1 5 10 15

Cys Pro Asn Ser
20

<210> 252
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H18 (PN)

<400> 252
ctgtctgtgc tcgctaactt cagtcagcca gagatcggga attct 45

<210> 253
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H18 (AA)

<400> 253

Leu Ser Val Leu Ala Asn Phe Ser Gln Pro Glu Ile Gly Asn Ser
1 5 10 15

<210> 254
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H19 (PN)

<400> 254
ctgtctgtgc tcgctaactt cagtcagcca gagatctcct gtccaccttg cccgaattct 60

<210> 255
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H19 (AA)

<400> 255

Leu Ser Val Leu Ala Asn Phe Ser Gln Pro Glu Ile Ser Cys Pro Pro
1 5 10 15

Cys Pro Asn Ser
20

<210> 256
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H20 (PN)

<400> 256
ctgaaaatcc aggagagggt cagtaagcca aagatctcga attct 45

<210> 257
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H20 (AA)

<400> 257

Leu Lys Ile Gln Glu Arg Val Ser Lys Pro Lys Ile Ser Asn Ser
1 5 10 15

<210> 258
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 连接子 H21 (PN)

<400> 258
ctgaaaatcc aggagagggt cagtaagcca aagatctcct gtccaccttg cccgaattct 60

<210> 259
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H21 (AA)

<400> 259

Leu Lys Ile Gln Glu Arg Val Ser Lys Pro Lys Ile Ser Cys Pro Pro
1 5 10 15

Cys Pro Asn Ser
20

<210> 260
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成多肽

<220>
<223> 连接子 H22 (PN)

<400> 260
ctggatgtga gtgagaggcc ttttcctcca cacatccaga attct 45

<210> 261
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H22 (AA)

<400> 261

Leu Asp Val Ser Glu Arg Pro Phe Pro Pro His Ile Gln Asn Ser
1 5 10 15

<210> 262
<211> 63
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H23 (PN)

<400> 262
ctggatgtga gtgagaggcc ttttcctcca cacatccagt cctgtccacc ttgcccgaat 60

tct 63

<210> 263
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H23 (AA)

<400> 263

Leu Asp Val Ser Glu Arg Pro Phe Pro Pro His Ile Gln Ser Cys Pro
1 5 10 15

Pro Cys Pro Asn Ser
20

<210> 264
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H24 (DNA)

<400> 264
cgggaacagc tggcagaggt cactttgagc ttgaaagcga attct 45

<210> 265
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H24 (AA)

<400> 265

Arg Glu Gln Leu Ala Glu Val Thr Leu Ser Leu Lys Ala Asn Ser
1 5 10 15

<210> 266
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H25 (PN)

<400> 266
cgggaacagc tggcagaggt cactttgagc gtgaaagctt gtccaccctg cccgaattct 60

<210> 267
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H25 (AA)

<400> 267

Arg Glu Gln Leu Ala Glu Val Thr Leu Ser Leu Lys Ala Cys Pro Pro
1 5 10 15

Cys Pro Asn Ser
20

<210> 268
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H26 (PN)

<400> 268
cggattcacc agatgaactc cgagttgagc gtgctcgcga attct 45

<210> 269
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H26 (AA)

<400> 269

Arg Ile His Gln Met Asn Ser Glu Leu Ser Val Leu Ala Asn Ser
1 5 10 15

<210> 270
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H27 (PN)

<400> 270
cggattcacc agatgaactc cgagttgagc gtgctcgctt gtccaccctg cccgaattct 60

<210> 271
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H27 (AA)

<400> 271

Arg Ile His Gln Met Asn Ser Glu Leu Ser Val Leu Ala Cys Pro Pro
1 5 10 15

Cys Pro Asn Ser
20

<210> 272
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H28 (PN)

<400> 272
gataccaaag ggaagaacgt cctcgagaag atcttctcga attct 45

<210> 273
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H28 (AA)

<400> 273

Asp Thr Lys Gly Lys Asn Val Leu Glu Lys Ile Phe Ser Asn Ser
1 5 10 15

<210> 274
<211> 63
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H29 (PN)

<400> 274
gataccaaag ggaagaacgt cctcgagaag atcttcgact cctgtccacc ttgcccgaat 60

tct 63

<210> 275
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H29 (AA)

<400> 275

Asp Thr Lys Gly Lys Asn Val Leu Glu Lys Ile Phe Asp Ser Cys Pro
1 5 10 15

Pro Cys Pro Asn Ser
20

<210> 276
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H30 (PN)

<400> 276
ctgccacctg agacacagga gagtcaagaa gtcaccctga attct 45

<210> 277

<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H30 (AA)

<400> 277

Leu Pro Pro Glu Thr Gln Glu Ser Gln Glu Val Thr Leu Asn Ser
1 5 10 15

<210> 278
<211> 63
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H31 (PN)

<400> 278
ctgccacctg agacacagga gagtcaagaa gtcaccctgt cctgtccacc ttgcccgaat 60

tct 63

<210> 279
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H31 (AA)

<400> 279

Leu Pro Pro Glu Thr Gln Glu Ser Gln Glu Val Thr Leu Ser Cys Pro
1 5 10 15

Pro Cys Pro Asn Ser
20

<210> 280
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H32 (PN)

<400> 280
cggattcacc tgaacgtgtc cgagaggccc tttcctccga attct 45

<210> 281
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H32 (AA)

<400> 281

Arg Ile His Leu Asn Val Ser Glu Arg Pro Phe Pro Pro Asn Ser
1 5 10 15

<210> 282
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H33 (PN)

<400> 282
cggattcacc tgaacgtgtc cgagaggccc tttcctccct gtccaccctg cccgaattct 60

<210> 283
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H33 (AA)

<400> 283

Arg Ile His Leu Asn Val Ser Glu Arg Pro Phe Pro Pro Cys Pro Pro
1 5 10 15

Cys Pro Asn Ser
20

<210> 284

<400> 284
000

<210> 285

<400> 285
000

<210> 286

<400> 286
000

<210> 287

<400> 287
000

<210> 288
<211> 39
<212> DNA
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H36 (PN)

<400> 288
gggtgtccac cttgtccagg cggtggtgga tcgaattct 39

<210> 289
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H36 (AA)

<400> 289

Gly Cys Pro Pro Cys Pro Gly Gly Gly Gly Ser Asn Ser
1 5 10

<210> 290

<400> 290
000

<210> 291

<400> 291
000

<210> 292

<400> 292
000

<210> 293

<400> 293
000

<210> 294

<400> 294
000

<210> 295

<400> 295
000

<210> 296
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H40 (PN)

<400> 296
ggatgtccac cttgtcccgc gaattct 27

<210> 297
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H40 (AA)

<400> 297

Gly Cys Pro Pro Cys Pro Ala Asn Ser
1 5

<210> 298

<400> 298
000

<210> 299
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<221> misc_feature
<223> 连接子 H41

<400> 299

Gly Cys Pro Pro Cys Pro Ala Asn Ser
1 5

<210> 300

<400> 300
000

<210> 301
<211> 8
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H42(AA)

<400> 301

Gly Cys Pro Pro Cys Pro Asn Ser
1 5

<210> 302

<400> 302
000

<210> 303

<400> 303
000

<210> 304
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H44 (PN)

<400> 304
ggaggagcta gttgtccacc ttgtcccggg aattct 36

<210> 305

<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H44 (AA)

<400> 305

Gly Gly Ala Ser Cys Pro Pro Cys Pro Gly Asn Ser
1 5 10

<210> 306
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H45 (PN)

<400> 306
ggaggagcca gttgtccacc ttgtgccggg aattct 36

<210> 307
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H45 (AA)

<400> 307

Gly Gly Ala Ser Cys Pro Pro Cys Ala Gly Asn Ser
1 5 10

<210> 308
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<223> 连接子 H46

<400> 308
ggaggagcca gttgtccacc ttgtgcgaat tct 33

<210> 309
<211> 11
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H46

<400> 309
Gly Gly Ala Ser Cys Pro Pro Cys Ala Asn Ser
1 5 10

<210> 310
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H47

<400> 310

Leu Ser Val Lys Ala Asp Phe Leu Thr Pro Ser Ile Gly Asn Ser
1 5 10 15

<210> 311
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H48

<400> 311

Lys Ala Asp Phe Leu Thr Pro Ser Ile Gly Asn Ser
1 5 10

<210> 312

<400> 312
000

<210> 313
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>

<223> 连接子 H50

<400> 313

Leu Ser Val Leu Ala Asn Phe Ser Gln Pro Glu Ile Gly Asn Ser
1 5 10 15

<210> 314
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H51

<400> 314

Leu Ser Val Leu Ala Asn Phe Ser Gln Pro Glu Ile Gly Asn Ser
1 5 10 15

<210> 315
<211> 11
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H52

<400> 315

Ser Gln Pro Glu Ile Val Pro Ile Ser Asn Ser
1 5 10

<210> 316
<211> 16
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H53

<400> 316

Ser Gln Pro Glu Ile Val Pro Ile Ser Cys Pro Pro Cys Pro Asn Ser
1 5 10 15

<210> 317
<211> 19
<212> PRT

<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H54

<400> 317

Ser Val Leu Ala Asn Phe Ser Gln Pro Glu Ile Ser Cys Pro Pro Cys
1 5 10 15

Pro Asn Ser

<210> 318
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H55

<400> 318

Arg Ile His Gln Met Asn Ser Glu Leu Ser Val Leu Ala Asn Ser
1 5 10 15

<210> 319
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H56

<400> 319

Gln Met Asn Ser Glu Leu Ser Val Leu Ala Asn Ser
1 5 10

<210> 320
<211> 10
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>

<223> 连接子 H57

<400> 320

Val Ser Glu Arg Pro Phe Pro Pro Asn Ser
1 5 10

<210> 321
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H58

<400> 321

Lys Pro Phe Phe Thr Cys Gly Ser Ala Asp Thr Cys Pro Asn Ser
1 5 10 15

<210> 322
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H59

<400> 322

Lys Pro Phe Phe Thr Cys Gly Ser Ala Asp Thr Cys Pro Asn Ser
1 5 10 15

<210> 323
<211> 14
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H60

<400> 323

Gln Tyr Asn Cys Pro Gly Gln Tyr Thr Phe Ser Met Asn Ser
1 5 10

<210> 324
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H61

<400> 324

Glu Pro Ala Phe Thr Pro Gly Pro Asn Ile Glu Leu Gln Lys Asp Ser
1 5 10 15

Asp Cys Pro Asn Ser
20

<210> 325
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H62

<400> 325

Gln Arg His Asn Asn Ser Ser Leu Asn Thr Arg Thr Gln Lys Ala Arg
1 5 10 15

His Cys Pro Asn Ser
20

<210> 326
<211> 19
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H63

<400> 326

Asn Ser Leu Phe Asn Gln Glu Val Gln Ile Pro Leu Thr Glu Ser Tyr
1 5 10 15

Cys Asn Ser

<210> 327
<211> 21
<212> PRT

<213> 人工序列

<220>
<223> 合成肽

<400> 327

Gln Arg His Asn Asn Ser Ser Leu Asn Thr Arg Thr Gln Lys Ala Arg
1 5 10 15

His Cys Pro Asn Ser
20

<210> 328
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H65

<400> 328

Glu Pro Ala Phe Thr Pro Gly Pro Asn Ile Glu Leu Gln Lys Asp Ser
1 5 10 15

Asp Cys Pro Asn Ser
20

<210> 329
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H66

<400> 329

Glu Pro Ala Phe Thr Pro Gly Pro Asn Ile Glu Leu Gln Lys Asp Ser
1 5 10 15

Asp Cys Pro Asn Ser
20

<210> 330
<211> 21
<212> PRT
<213> 人工序列

<220>

<223> 合成肽

<220>
<223> 连接子 H67

<400> 330

Glu Pro Ala Phe Thr Pro Gly Pro Asn Ile Glu Leu Gln Lys Asp Ser
1 5 10 15

Asp Cys Pro Asn Ser
20

<210> 331
<211> 118
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 连接子 H68

<400> 331

Arg Thr Arg Tyr Leu Gln Val Ser Gln Gln Leu Gln Gln Thr Asn Arg
1 5 10 15

Val Leu Glu Val Thr Asn Ser Ser Leu Arg Gln Gln Leu Arg Leu Lys
20 25 30

Ile Thr Gln Leu Gly Gln Ser Ala Glu Asp Leu Gln Gly Ser Arg Arg
35 40 45

Glu Leu Ala Gln Ala Gln Glu Ala Leu Gln Val Glu Gln Arg Ala His
50 55 60

Gln Ala Ala Glu Gly Gln Leu Gln Ala Cys Gln Ala Asp Arg Gln Lys
65 70 75 80

Thr Lys Glu Thr Leu Gln Ser Glu Glu Gln Gln Arg Arg Ala Leu Glu
85 90 95

Gln Lys Leu Ser Asn Met Glu Asn Arg Leu Lys Pro Phe Phe Thr Cys
100 105 110

Gly Ser Ala Asp Thr Cys
115

<210> 332
<211> 10
<212> PRT

<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD-20 VL CDR1

<400> 332

Arg Ala Ser Ser Ser Val Ser Tyr Ile His
1 5 10

<210> 333
<211> 10
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD-20 VL CDR1

<400> 333

Arg Ala Ser Ser Ser Val Ser Tyr Ile Val
1 5 10

<210> 334
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD-20 VL CDR3

<400> 334

Gln Gln Trp Ser Phe Asn Pro Pro Thr
1 5

<210> 335
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD-20 VL CDR3

<400> 335

Gln Gln Tyr Ser Phe Asn Pro Pro Thr
1 5

<210> 336
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD-20 VH CDR2

<400> 336

Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn Gln Lys Phe Lys
1 5 10 15

Gly

<210> 337
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD-20 VH CDR2

<400> 337

Ala Ile Tyr Pro Gly Asn Gly Glu Thr Ser Tyr Asn Gln Lys Phe Lys
1 5 10 15

Gly

<210> 338
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3

<400> 338
Ser Val Tyr Tyr Ser Asn Tyr Trp Tyr Phe Asp Leu
1 5 10

<210> 339
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3

<400> 339

Ser Val Tyr Tyr Gly Gly Tyr Trp Tyr Phe Asp Leu
1 5 10

<210> 340
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3

<400> 340

Ser Tyr Tyr Ser Asn Ser Asp Trp Tyr Phe Asp Leu
1 5 10

<210> 341
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3

<400> 341

Ser Tyr Tyr Ser Gly Gly Asp Trp Tyr Phe Asp Leu
1 5 10

<210> 342
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3

<400> 342

Ser Tyr Lys Ser Asn Ser Tyr Trp Tyr Phe Asp Leu
1 5 10

<210> 343
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3

<400> 343

Ser Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Leu
1 5 10

<210> 344
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD-20 VH CDR3

<400> 344

Ser Tyr Lys Ser Asn Ser Asp Trp Tyr Phe Asp Leu
1 5 10

<210> 345
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3

<400> 345

Ser Tyr Lys Ser Gly Gly Asp Trp Tyr Phe Asp Leu
1 5 10

<210> 346
<211> 8
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 蝎形连接子核心(2H7)序列

<400> 346

Gly Cys Pro Pro Cys Pro Asn Ser
1 5

<210> 347
<211> 4
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 扩展序列

<400> 347

Ala Pro Glu Leu
1

<210> 348
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 扩展序列

<400> 348

Ala Pro Glu Leu Gly Gly Gly Gly Ser
1 5

<210> 349
<211> 14
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>

<223> 扩展序列

<400> 349

Ala Pro Glu Leu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10

<210> 350
<211> 19
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 扩展序列

<400> 350

Ala Pro Glu Leu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
1 5 10 15

Gly Gly Ser

<210> 351
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 经扩展的蝎形连接子

<400> 351

Gly Cys Pro Pro Cys Pro Asn Ser Ala Pro Glu Leu
1 5 10

<210> 352
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 经扩展的蝎形连接子

<400> 352

Gly Cys Pro Pro Cys Pro Asn Ser Ala Pro Glu Leu Gly Gly Gly Gly
1 5 10 15

Ser

<210> 353
<211> 22
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 经扩展的蝎形连接子

<400> 353

Gly Cys Pro Pro Cys Pro Asn Ser Ala Pro Glu Leu Gly Gly Gly Gly
1 5 10 15

Ser Gly Gly Gly Gly Ser
20

<210> 354
<211> 27
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 经扩展的蝎形连接子

<400> 354

Gly Cys Pro Pro Cys Pro Asn Ser Ala Pro Glu Leu Gly Gly Gly Gly
1 5 10 15

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25

<210> 355
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VL CDR1-X-VL CDR3-X- VH CDR2-X-VH CDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>

<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 355

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Val Tyr Tyr Ser Asn Tyr Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 356
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 356

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Val Tyr Tyr Gly Gly Tyr Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 357
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 357

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Tyr Ser Asn Ser Asp Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 358
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 358

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Tyr Ser Gly Gly Asp Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 359
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 359
Arg Ala Ser Ser Ser Val Ser Tyr Ile Val Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Lys Ser Asn Ser Tyr Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 360
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 360

Arg Ala Ser Ser Ser Val Ser Tyr Ile Val Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Glu Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Tyr Ser Asn Ser Tyr Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 361
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 361

Arg Ala Ser Ser Ser Val Ser Tyr Ile Val Xaa Gln Gln Tyr Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Tyr Ser Asn Ser Tyr Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 362
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature

<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 362

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Lys Ser Asn Ser Asp Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 363
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 363

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Trp Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Glu Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Tyr Ser Asn Ser Asp Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 364
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 364

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Tyr Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Tyr Ser Asn Ser Asp Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 365
<211> 51
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> VLCDR1-X-VLCDR3-X- VHCDR2-X-VHCDR3

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa = 介于 VL CDR1 和 VL CDR3 之间的氨基酸范围

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa = 介于 VL CDR3 和 VH CDR2 之间的氨基酸范围

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa = 介于 VH CDR2 和 VH CDR3 之间的氨基酸范围

<400> 365

Arg Ala Ser Ser Ser Val Ser Tyr Ile His Xaa Gln Gln Tyr Ser Phe
1 5 10 15

Asn Pro Pro Thr Xaa Ala Ile Tyr Pro Gly Asn Gly Glu Thr Ser Tyr
20 25 30

Asn Gln Lys Phe Lys Gly Xaa Ser Tyr Lys Ser Gly Gly Asp Trp Tyr
35 40 45

Phe Asp Leu
50

<210> 366
<211> 10
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 部分 CH3 序列

<400> 366

Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
1 5 10

<210> 367
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 部分 CH3 序列

<400> 367

Gln Lys Ser Leu Ser Leu Ser Pro Gly
1 5

<210> 368
<211> 1
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 部分 CH3 序列

<400> 368

Gln
1

<210> 369
<211> 3
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 部分 CH3 序列

<400> 369

Gln Lys Ser
1

<210> 370
<211> 6
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 部分 CH3 序列

<400> 370

Gln Lys Ser Leu Ser Leu
1 5

<210> 371
<211> 5
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>

<223> 部分 CH3 序列

<400> 371

Gln Lys Ser Leu Ser
1 5

<210> 372
<211> 5
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 蝎形连接子

<400> 372

Gly Cys Pro Pro Cys
1 5

<210> 373
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 蝎形连接子

<400> 373

Gln Tyr Asn Cys Pro Gly Gln Tyr Thr Phe Ser Met
1 5 10

<210> 374
<211> 11
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 蝎形连接子

<400> 374

Pro Phe Phe Thr Cys Gly Ser Ala Asp Thr Cys
1 5 10

<210> 375
<211> 18

<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 蝎形连接子

<400> 375

Glu Pro Ala Phe Thr Pro Gly Pro Asn Ile Glu Leu Gln Lys Asp Ser
1 5 10 15

Asp Cys

<210> 376
<211> 18
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 蝎形连接子

<400> 376

Gln Arg His Asn Asn Ser Ser Leu Asn Thr Arg Thr Gln Lys Ala Arg
1 5 10 15

His Cys

<210> 377
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 蝎形连接子

<400> 377

Asn Ser Leu Phe Asn Gln Glu Val Gln Ile Pro Leu Thr Glu Ser Tyr
1 5 10 15

Cys

<210> 378
<211> 10
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VL CDR1 (TRU-015)

<400> 378

Arg Ala Ser Ser Ser Val Ser Tyr Met His
1 5 10

<210> 379
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> 合成肽

<220>
<223> 抗-CD20 VH CDR3 (TRU-015)

<400> 379

Val Val Tyr Tyr Ser Asn Ser Tyr Trp Tyr Phe Asp Val
1 5 10



图 1

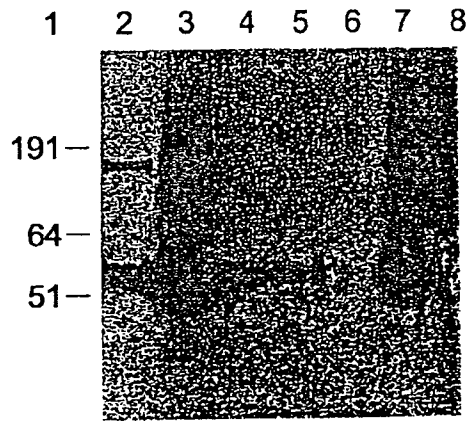


图 2

COS上清液中表达的蛋白与表达靶抗原的细胞结合

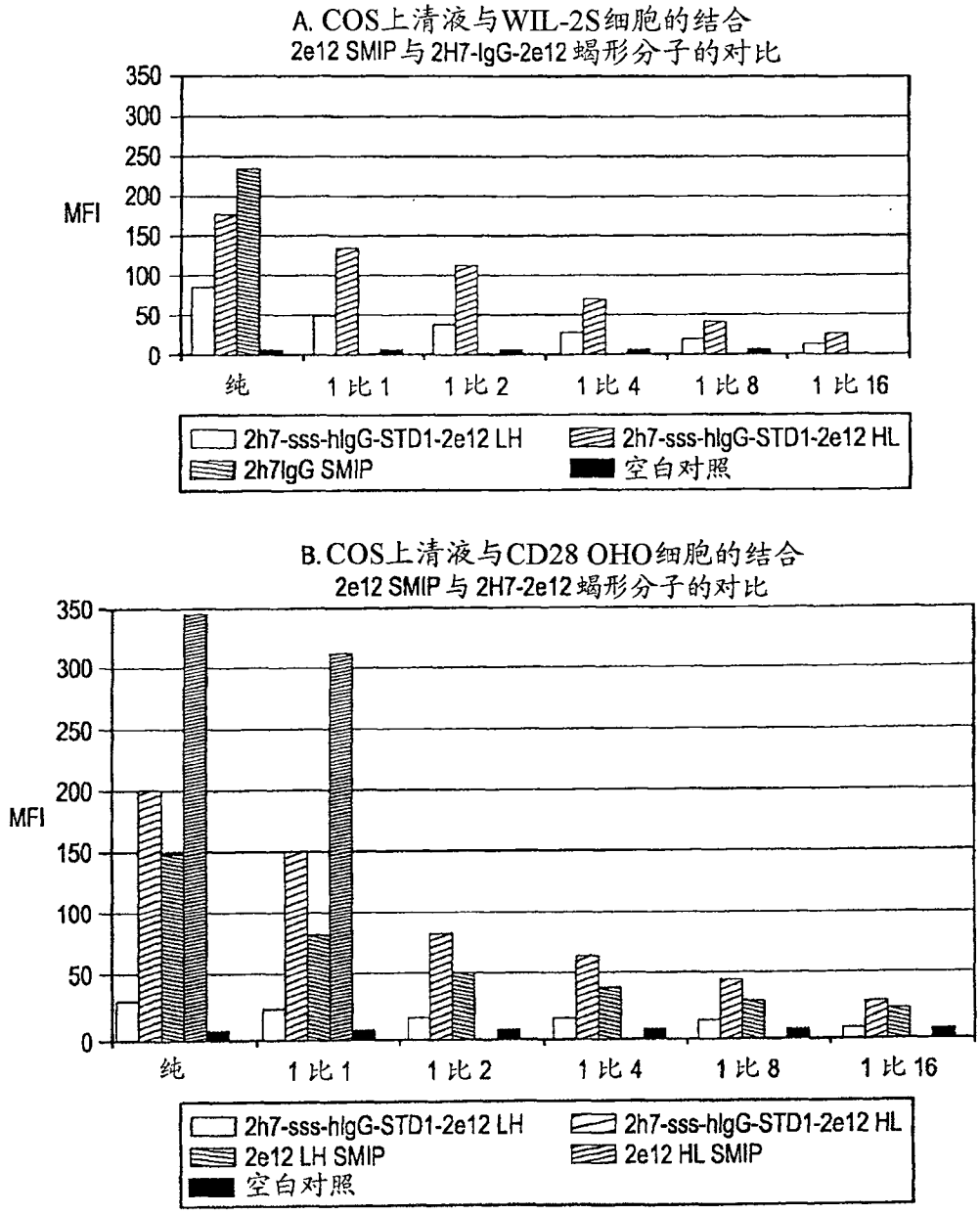


图 3

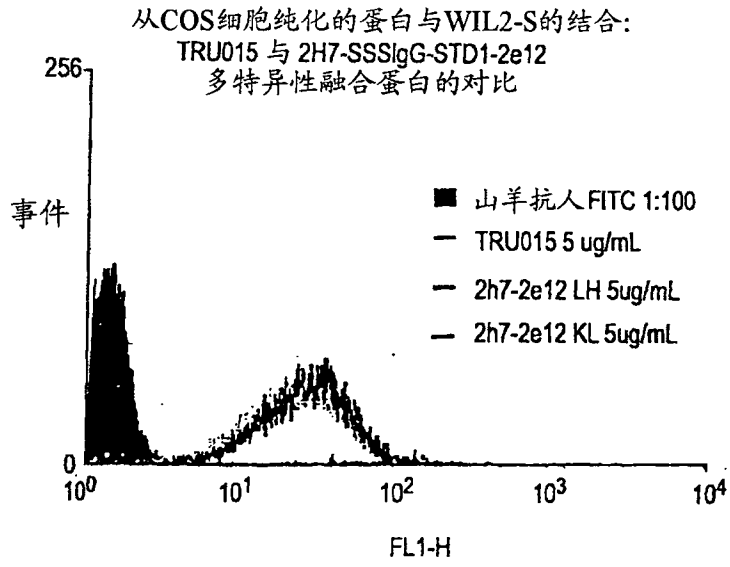

从COS细胞纯化的蛋白与WIL2-S的结合:
TRU015 与 2H7-SSSIgG-STD1-2e12
多特异性融合蛋白的对比
256
事件
0
山羊抗人FITC 1:100
TRU015 5 ug/mL
2h7-2e12 LH 5ug/mL
2h7-2e12 KL 5ug/mL
10^0
10^1
10^2
10^3
10^4
FL1-H

图 4

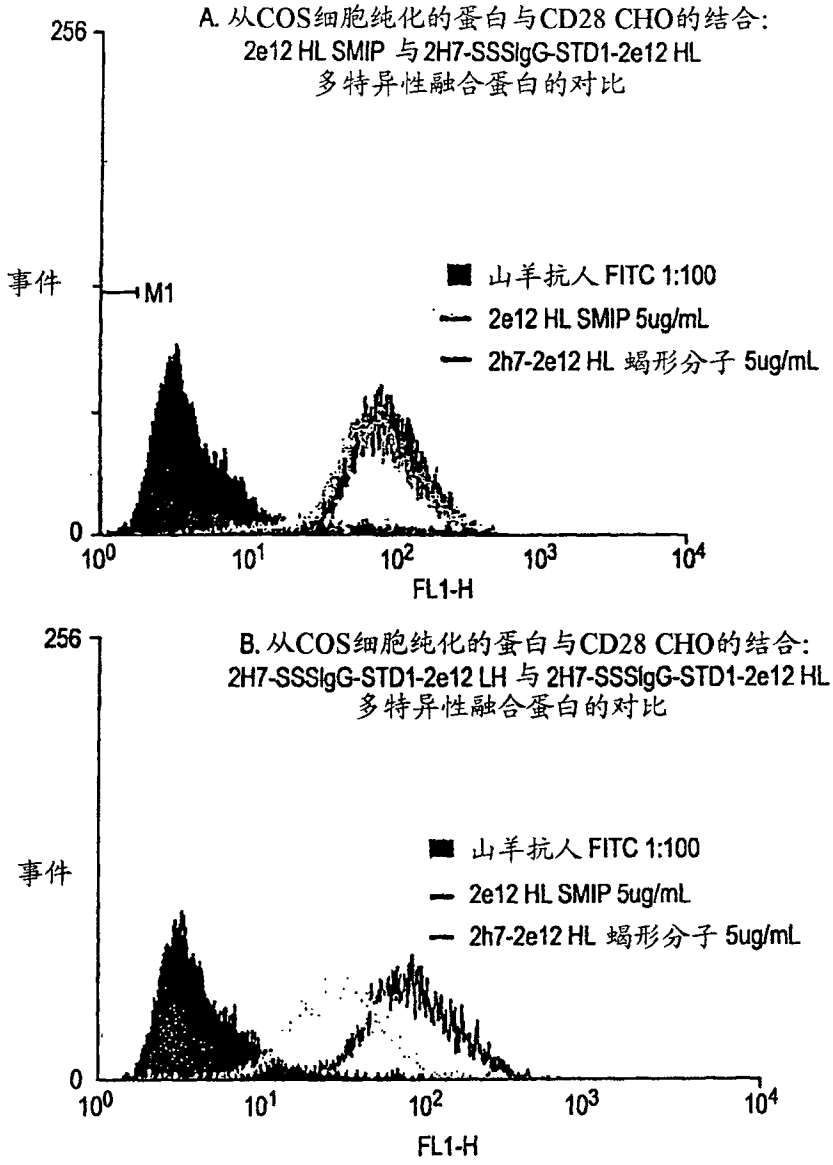

A. 从COS细胞纯化的蛋白与CD28 CHO的结合:
2e12 HL SMIP 与2H7-SSSIgG-STD1-2e12 HL
多特异性融合蛋白的对比
256
事件
M1
0
山羊抗人 FITC 1:100
2e12 HL SMIP 5ug/mL
2h7-2e12 HL 蝎形分子 5ug/mL
10^0
10^1
10^2
10^3
10^4
FL1-H
B. 从COS细胞纯化的蛋白与CD28 CHO的结合:
2H7-SSSIgG-STD1-2e12 LH 与 2H7-SSSIgG-STD1-2e12 HL
多特异性融合蛋白的对比
256
事件
0
山羊抗人 FITC 1:100
2e12 HL SMIP 5ug/mL
2h7-2e12 HL 蝎形分子 5ug/mL
10^0
10^1
10^2
10^3
10^4
FL1-H

图 5

标识多特异性融合蛋白的功能性元件的表格



BD1 结合结构域1	连接子1	EFD 效应功能结构域	连接子2 ID	连接子长度(aa) 连接子+限制性位点	CH3 COOH	连接子2序列 (-限制性位点)	与BD2 (L2 RS--QVQ 2e12) 的融合接合	BD2 结合结构域2
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	STD	20	PGK	NYGGGGSGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	STD2	38	PGK	NYGGGGSGGGGSGGGGSG NYGGGGSGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	H1 L1=2e12 V_L-V_H	2 (RS)	PGK	-	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	H2 L2=2e12 V_L-V_H	8	PGK	GGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	H3 L3=2e12 V_L-V_H	10	PGK	NYGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	H4 L4=2e12 V_L-V_H	13	PGK	GGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	H5 L5=2e12 V_L-V_H	15	PGK	NYGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	H6 L6=2e12 V_L-V_H	18	PGK	GGGGSGGGGSGGGGSG	NS QVQ	2e12 V_H-V_L
2H7 V_L-V_H	经修饰的 hIgGl 铰链 CCC->SSS	hIgGl-CH2-CH3	H7	8	PGK	GCPPCP	NS QVQ	2e12 V_H-V_L

图 6A

构建体

名称	BD1	铰链	EFD	连接子	BD2
2H7-sss-hIgG-STD1-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD1	2e12 (VL-VH)
2H7-sss-hIgG-STD1-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD1	2e12 (VH-VL)
2H7-sss-hIgG-STD2-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD2	2e12 (VL-VH)
2H7-sss-hIgG-STD2-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD2	2e12 (VH-VL)
2H7-csc-hIgG-STD1-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD1	2e12 (VL-VH)
2H7-csc-hIgG-STD2-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD2	2e12 (VL-VH)
2H7-csc-hIgG-STD1-2e12 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD1	2e12 (VL-VH)
2H7-csc-hIgG-STD2-2e12 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD2	2e12 (VH-VL)
2H7-sss-hIgG-H1-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H1	2e12 (VL-VH)
2H7-csc-hIgG-H1-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H1	2e12 (VL-VH)
2H7-sss-hIgG-H1-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H1	2e12 (VH-VL)
2H7-csc-hIgG-H1-2e12 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H1	2e12 (VH-VL)
2H7-sss-hIgG-H2-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H2	2e12 (VH-VL)
2H7-csc-hIgG-H2-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H2	2e12 (VL-VH)
2H7-sss-hIgG-H2-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H2	2e12 (VH-VL)
2H7-csc-hIgG-H2-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H2	2e12 (VH-VL)
2H7-sss-hIgG-H3-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H3	2e12 (VL-VH)
2H7-csc-hIgG-H3-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H3	2e12 (VL-VH)
2H7-sss-hIgG-H3-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H3	2e12 (VH-VL)
2H7-csc-hIgG-H3-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H3	2e12 (VH-VL)
2H7-sss-hIgG-H4-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H4	2e12 (VL-VH)
2H7-csc-hIgG-H4-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H4	2e12 (VL-VH)
2H7-sss-hIgG-H4-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H4	2e12 (VH-VL)
2H7-csc-hIgG-H4-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H4	2e12 (VH-VL)

图 6B-1

名称	BD1	铰链	EFD	连接子	BD2
2H7-sss-hIgG-H5-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H5	2e12 (VL-VH)
2H7-csc-hIgG-H5-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H5	2e12 (VL-VH)
2H7-sss-hIgG-H5-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H5	2e12 (VH-VL)
2H7-csc-hIgG-H5-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H5	2e12 (VH-VL)
2H7-sss-hIgG-H6-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H6	2e12 (VL-VH)
2H7-csc-hIgG-H6-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H6	2e12 (VL-VH)
2H7-sss-hIgG-H6-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H6	2e12 (VH-VL)
2H7-csc-hIgG-H6-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H6	2e12 (VH-VL)
2H7-sss-hIgG-H7-2e12 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H7	2e12 (VL-VH)
2H7-csc-hIgG-H7-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H7	2e12 (VL-VH)
2H7-sss-hIgG-H7-2e12 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H7	2e12 (VH-VL)
2H7-csc-hIgG-H7-2e12 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H7	2e12 (VH-VL)
2H7-sss-hIgG-STD1-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-sss-hIgG-STD1-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-sss-hIgG-STD2-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD2	G28-1 (VL-VH)
2H7-sss-hIgG-STD2-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD2	G28-1 (VH-VL)
2H7-csc-hIgG-STD1-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-csc-hIgG-STD2-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD2	G28-1 (VH-VL)
2H7-csc-hIgG-STD1-G28-1 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD1	G28-1 (VL-VH)
2H7-csc-hIgG-STD2-G28-1 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD2	G28-1 (VH-VL)
2H7-sss-hIgG-H1-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H1	G28-1 (VL-VH)
2H7-csc-hIgG-H1-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H1	G28-1 (VH-VL)
2H7-sss-hIgG-H1-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H1	G28-1 (VL-VH)
2H7-csc-hIgG-H1-G28-1 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H1	G28-1 (VH-VL)
2H7-sss-hIgG-H2-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H2	G28-1 (VL-VH)

图 6B-2

名称	BD1	铰链	EFD	链接子	BD2
2H7-csc-hIgG-H2-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H2	G28-1 (VH-VL)
2H7-sss-hIgG-H2-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H2	G28-1 (VL-VH)
2H7-csc-hIgG-H2-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H2	G28-1 (VH-VL)
2H7-sss-hIgG-H3-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H3	G28-1 (VL-VH)
2H7-csc-hIgG-H3-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H3	G28-1 (VH-VL)
2H7-sss-hIgG-H3-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H3	G28-1 (VL-VH)
2H7-csc-hIgG-H3-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H3	G28-1 (VH-VL)
2H7-sss-hIgG-H4-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H4	G28-1 (VL-VH)
2H7-csc-hIgG-H4-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H4	G28-1 (VH-VL)
2H7-sss-hIgG-H4-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H4	G28-1 (VL-VH)
2H7-csc-hIgG-H4-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H4	G28-1 (VH-VL)
2H7-sss-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-sss-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-sss-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-csc-hIgG-H5-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H5	G28-1 (VH-VL)
2H7-sss-hIgG-H6-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H6	G28-1 (VL-VH)
2H7-csc-hIgG-H6-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H6	G28-1 (VL-VH)
2H7-sss-hIgG-H6-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H6	G28-1 (VH-VL)
2H7-csc-hIgG-H6-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H6	G28-1 (VH-VL)
2H7-sss-hIgG-H7-G28-1 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H7	G28-1 (VL-VH)
2H7-csc-hIgG-H7-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H7	G28-1 (VL-VH)
2H7-sss-hIgG-H7-G28-1 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	H7	G28-1 (VH-VL)
2H7-csc-hIgG-H7-G28-1 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG 或hIgG (P238S/P331S)	H7	G28-1 (VH-VL)
2H7-sss-hIgG-STD1-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	STD1	G19-4 (VL-VH)
2H7-sss-hIgG-STD1-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG 或hIgG (P238S/P331S)	STD1	G19-4 (VH-VL)

图 6B-3

名称	BD1	铰链	EFD	连接子	BD2
2H7-sss-hIgG-STD2-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD2	G19-4 (VL-VH)
2H7-sss-hIgG-STD2-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	STD2	G19-4 (VH-VL)
2H7-csc-hIgG-STD1-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD1	G19-4 (VL-VH)
2H7-csc-hIgG-STD2-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD2	G19-4 (VH-VL)
2H7-csc-hIgG-STD1-G19-4 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD1	G19-4 (VL-VH)
2H7-csc-hIgG-STD2-G19-4 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	STD2	G19-4 (VH-VL)
2H7-sss-hIgG-H1-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H1	G19-4 (VL-VH)
2H7-csc-hIgG-H1-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H1	G19-4 (VH-VL)
2H7-sss-hIgG-H1-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H1	G19-4 (VL-VH)
2H7-csc-hIgG-H1-G19-4 HL	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H1	G19-4 (VH-VL)
2H7-sss-hIgG-H2-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-csc-hIgG-H2-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-sss-hIgG-H2-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-csc-hIgG-H2-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H2	G19-4 (VL-VH)
2H7-sss-hIgG-H3-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H3	G19-4 (VL-VH)
2H7-csc-hIgG-H3-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H3	G19-4 (VH-VL)
2H7-sss-hIgG-H3-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H3	G19-4 (VL-VH)
2H7-csc-hIgG-H3-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H3	G19-4 (VH-VL)
2H7-sss-hIgG-H4-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H4	G19-4 (VL-VH)
2H7-csc-hIgG-H4-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H4	G19-4 (VH-VL)
2H7-sss-hIgG-H4-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H4	G19-4 (VL-VH)
2H7-csc-hIgG-H4-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H4	G19-4 (VH-VL)
2H7-sss-hIgG-H5-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H5	G19-4 (VL-VH)
2H7-csc-hIgG-H5-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H5	G19-4 (VH-VL)
2H7-sss-hIgG-H5-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H5	G19-4 (VL-VH)

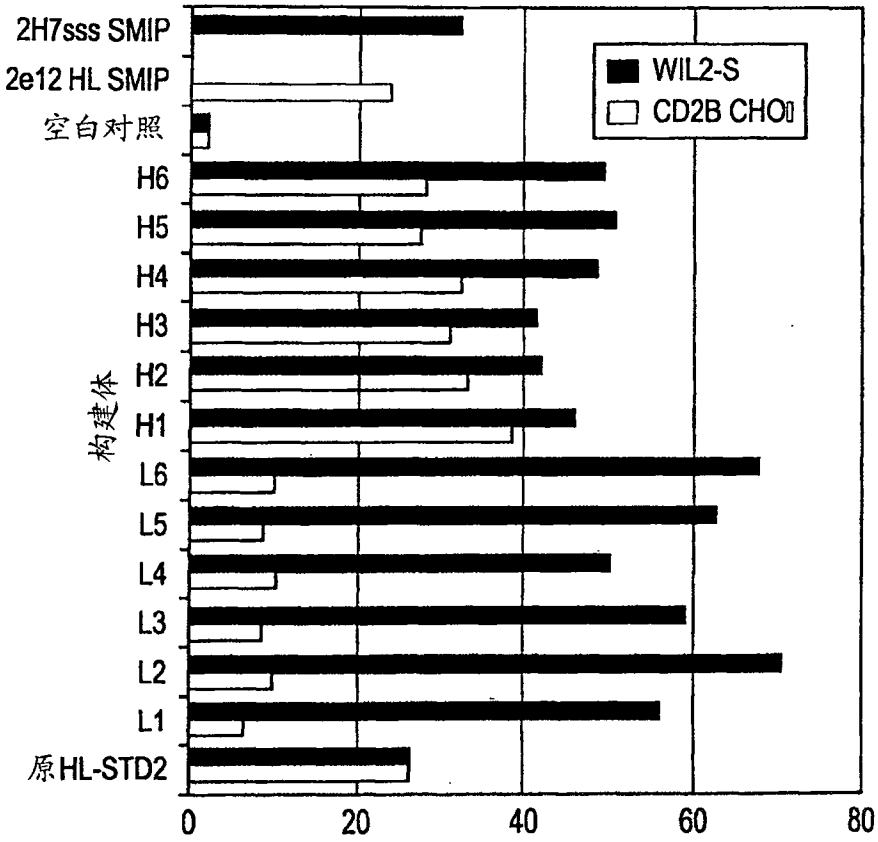
图 6B-4

名称	BD1	铰链	EFD	连接子	BD2
2H7-csc-hIgG-H5-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H5	G19-4 (VH-VL)
2H7-sss-hIgG-H6-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H6	G19-4 (VL-VH)
2H7-csc-hIgG-H6-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H6	G19-4 (VL-VH)
2H7-sss-hIgG-H6-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H6	G19-4 (VH-VL)
2H7-csc-hIgG-H6-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H6	G19-4 (VH-VL)
2H7-sss-hIgG-H7-G19-4 LH	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H7	G19-4 (VL-VH)
2H7-csc-hIgG-H7-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H7	G19-4 (VL-VH)
2H7-sss-hIgG-H7-G19-4 HL	2H7 (VL-VH)	hIgG1 -SSS	hIgG或hIgG (P238S/P331S)	H7	G19-4 (VH-VL)
2H7-csc-hIgG-H7-G19-4 LH	2H7 (VL-VH)	hIgG1 -csc	hIgG或hIgG (P238S/P331S)	H7	G19-4 (VH-VL)

图 6B-5

修改EFD后的连接子和BD2中V区
取向影响多特异性融合蛋白的结合效率

具有各种连接子的2H7-SSShIgG/Hx-2e12
多特异性融合蛋白与WIL2-S和CD28 CHO的结合



** 所有融合蛋白由瞬间COS上清液产生，
经蛋白A纯化，以0.72 μg/mL使用

图 7

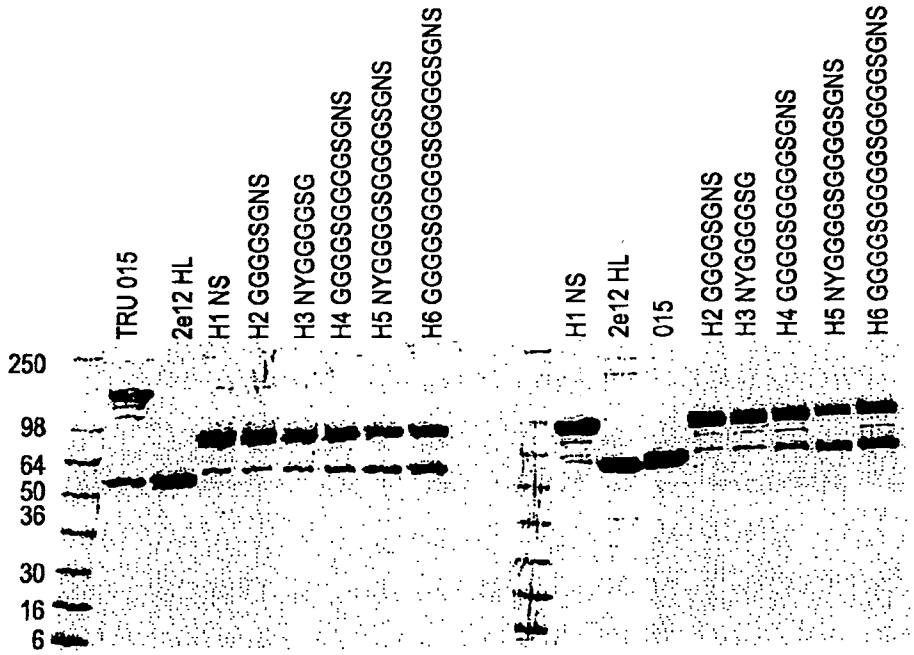

250
98
64
50
36
30
16
6
TRU 015
2e12 HL
H1 NS
H2 GGGGSGNS
H3 NYGGGGSG
H4 GGGGSGGGGSGNS
H5 NYGGGGSGGGGSGNS
H6 GGGGSGGGGSGGGGSGNS
H1 NS
2e12 HL
015
H2 GGGGSGNS
H3 NYGGGGSG
H4 GGGGSGGGGSGNS
H5 NYGGGGSGGGGSGNS
H6 GGGGSGGGGSGGGGSGNS
非还原的
还原的

图 8

对具有H6连接子的多特异性融合蛋白的蛋白印记

A.不存在SMIP或较小的CD28可检测形式

B.存在使用CD20抗-id Fab的SMIP尺寸化形式

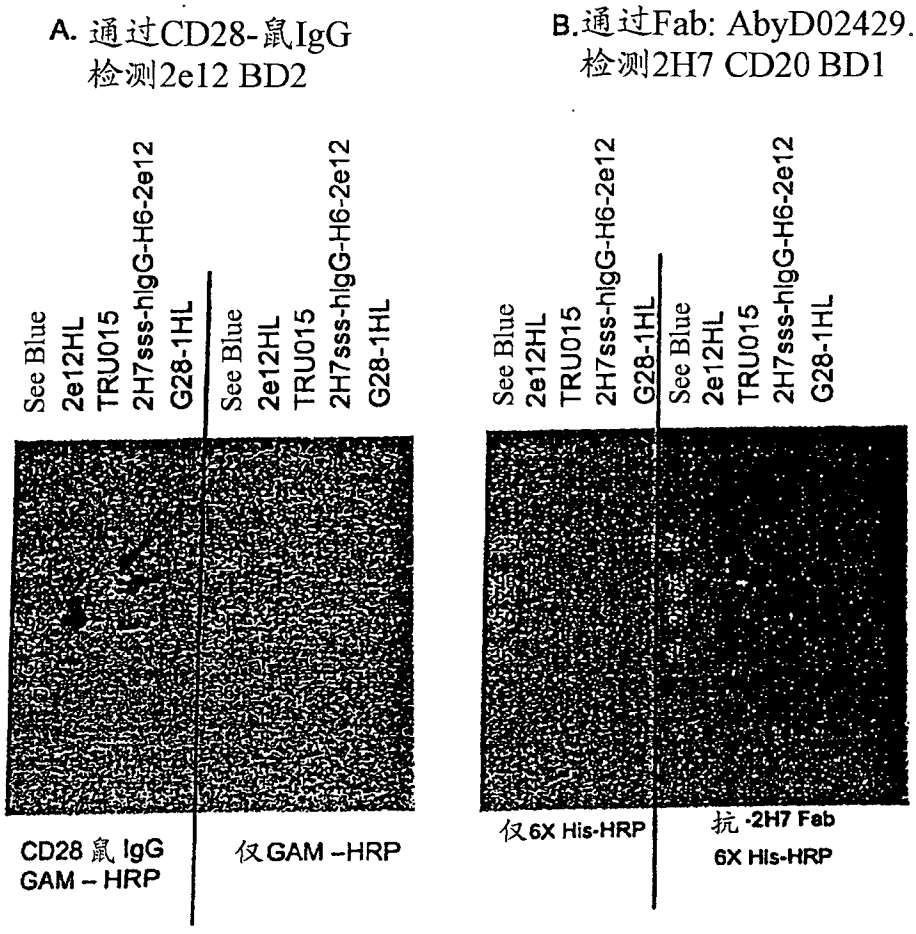
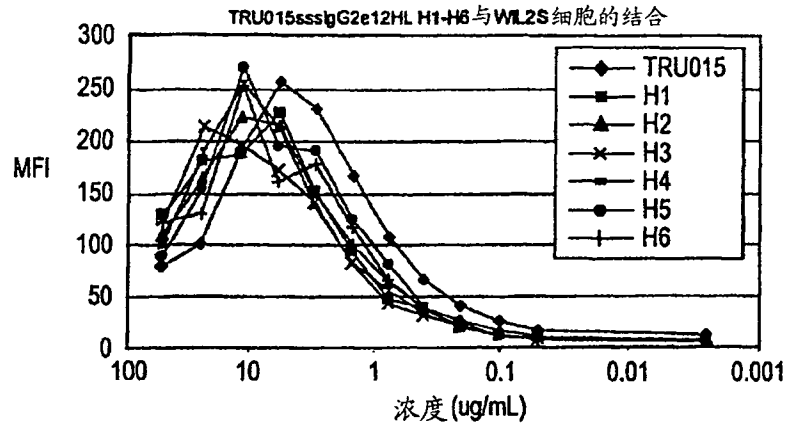


图 9

A.具有变异体连接子的多特异性融合蛋白
与WIL-2S细胞的结合



B.具有变异体连接子的多特异性融合蛋白
与CD28 CHO细胞的结合

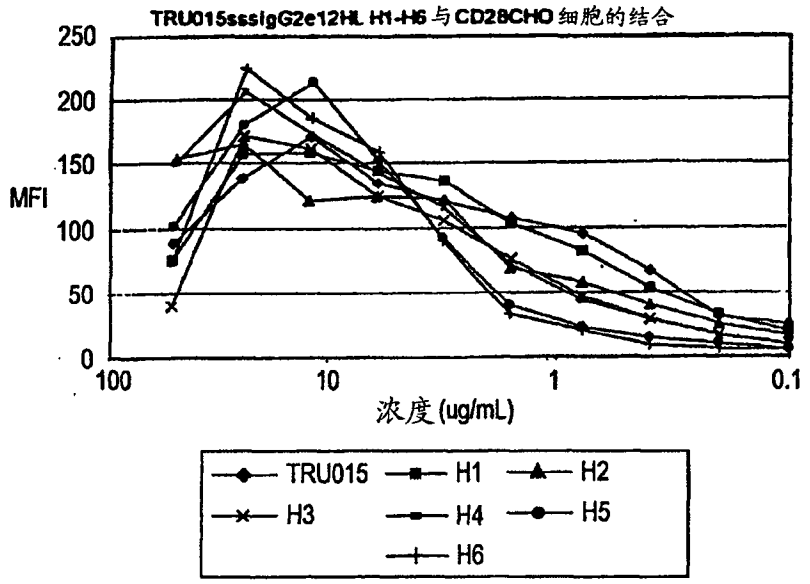


图 10

具有变异体连接子的2H7-sss-IgG-2e12-HL
多特异性融合蛋白与SEC分离的概述

连接子ID	POI的滞留时间	% POI	% 其他形式	在EFD/BD2结点处裂解的SDS-PAGE分析
H1	7.620	70.9	29.0	(是)
H2	7.589	67.5	32.5	(是)
H3	7.605	68.7	31.4	是
H4	7.622 (肩峰)	68.0	32.0	是
H5	7.933/7.680 (带有肩峰的双重峰)	32.42/33.94	33.63	是
H6	7.901/7.69 (带有肩峰的双重峰)	34.5/29.3	36.1	是
H7	7.788	84.2	15.8 (HMW)	否

图 11

具有不同连接子的[2H7-sss-hIgG-Hx-2e12HL]融合蛋白
与表达BD1或BD2的靶标抗原的细胞结合

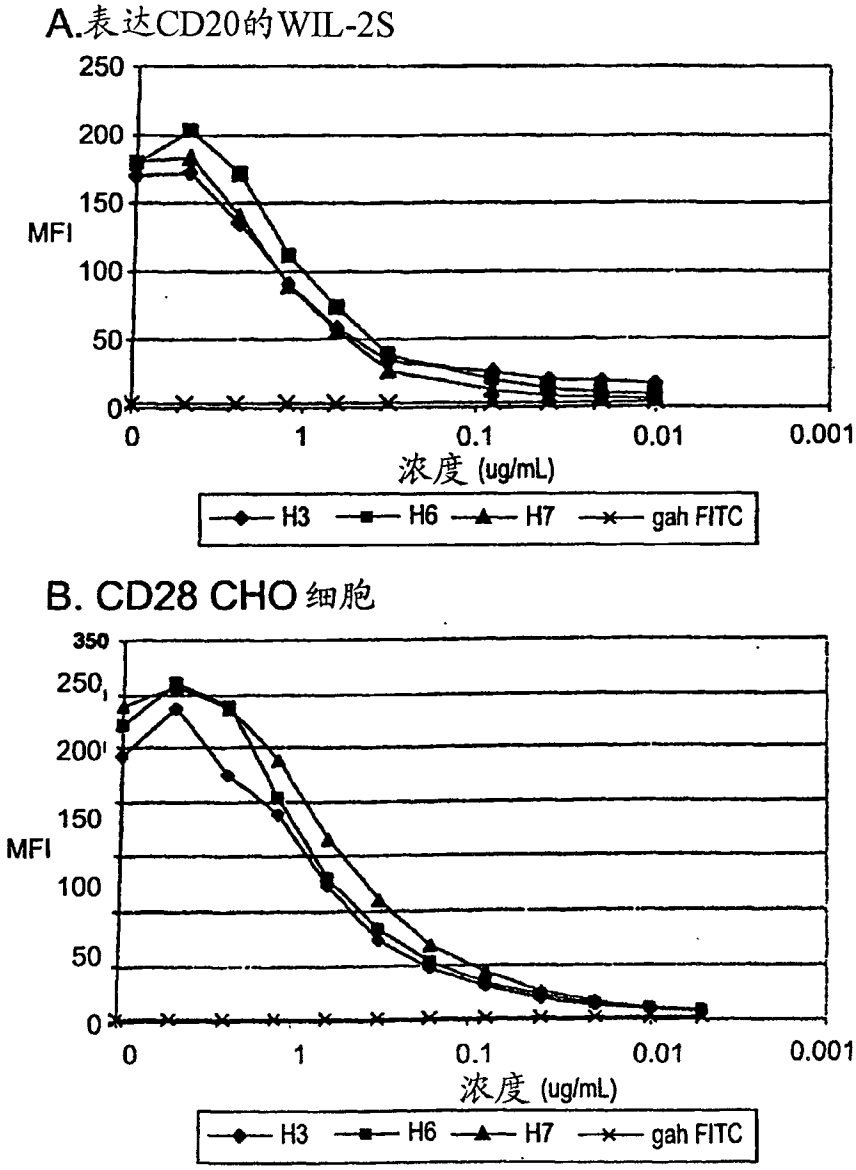


图 12

BD1和BD2的同时结合:

具有H3、H6和H7连接子的[2H7-sss-hIgG-Hx-2e12HL]
融合蛋白与WIL-2S细胞的结合可由CD28mIgG + FITC抗小鼠来检测

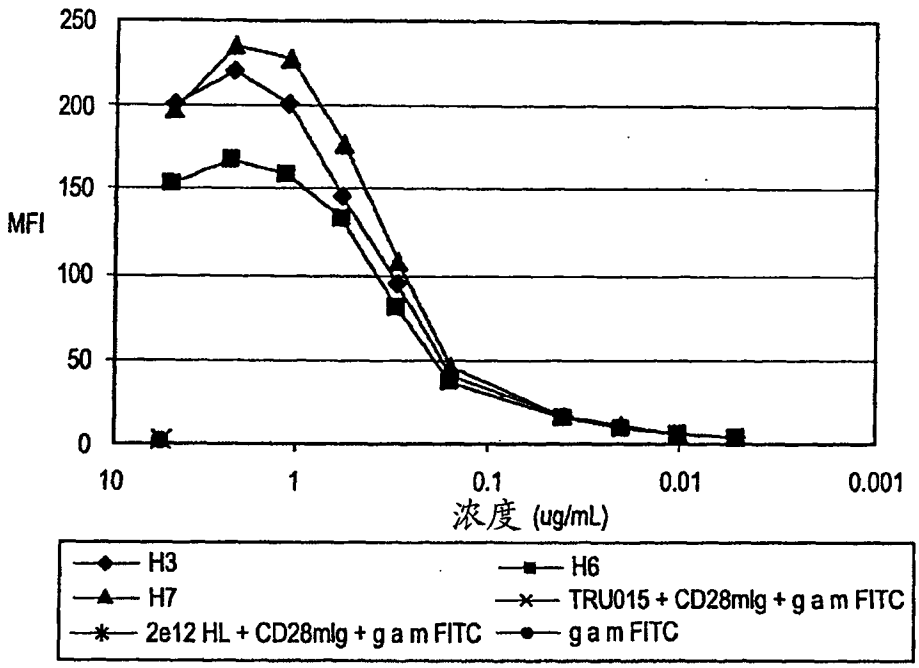
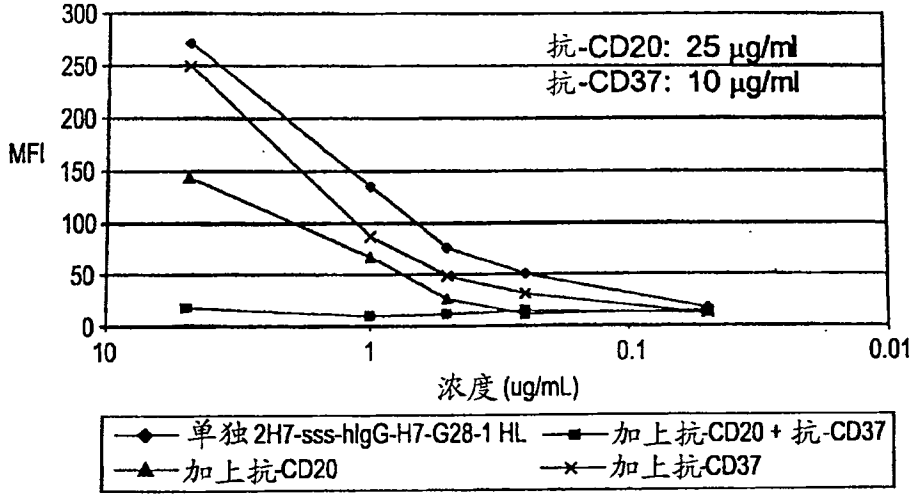


图 13

A.靶向CD20和/或CD37的抗体阻断2H7-sss-hIgG-H7-G28-1 HL蛋白和Ramos细胞的结合



B.靶向CD20和/或CD37的抗体阻断2H7-sss-hIgG-H7-G28-1 HL和BJAB细胞的结合

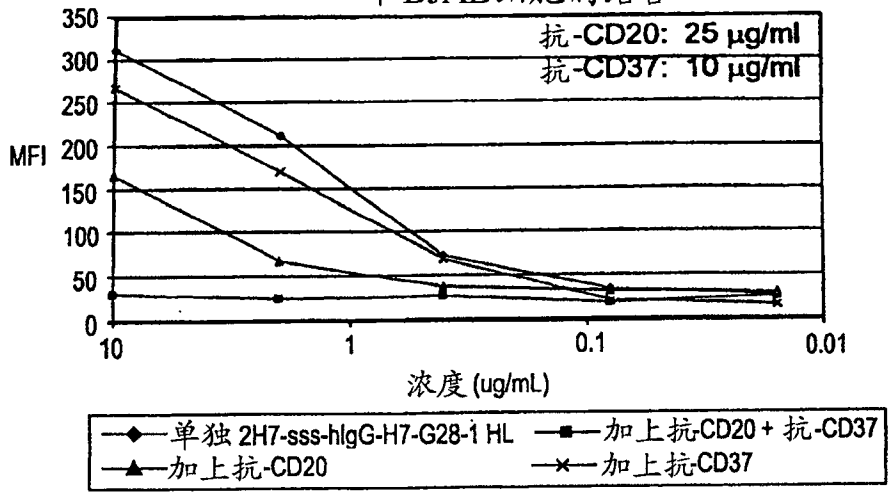


图 14

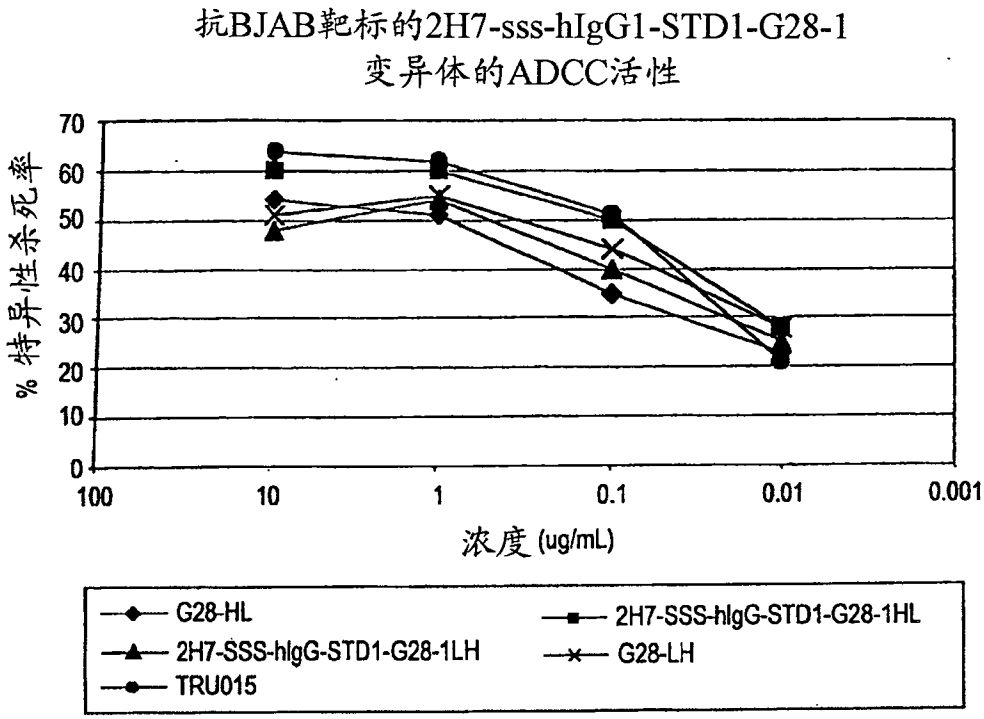

抗BJAB靶标的2H7-sss-hIgG1-STD1-G28-1
变异体的ADCC活性
% 特异性杀死率
70
60
50
40
30
20
10
0
100
10
1
0.1
0.01
0.001
浓度 (ug/mL)
G28-HL
2H7-SSS-hIgG-STD1-G28-1HL
2H7-SSS-hIgG-STD1-G28-1LH
G28-LH
TRU015

图 15

用TRU-015、G28-1 SMIP、TRU-015+G28-1 SMIP或
2H7-sss-hIgG1-H7-G28-1 HL孵育后，
培养液中存在的CD3、CD19和CD40阳性淋巴细胞的百分比

24小时	$CD3^+$	$CD19^+$	$CD40^+$
培养基	77%	8%	7%
TRU015	76%	8.5%	7.9%
G28-1 SMIP	78%	8.9%	7.8%
TRU015+G28-1 SMIP	76.3%	8.1%	7.5%
2H7-ssshIgG1-G28-1HL (H7)	78.4%	8.9%	7.9%

72小时	$CD3^+$	$CD19^+$	$CD40^+$
培养基	78.5%	7.9%	7.7%
TRU015	85.2%	2.3%	2.7%
G28-1 SMIP	86.7%	2.2%	2.3%
TRU015+G28-1 SMIP	87.1%	1.1%	1.1%
2H7-ssshIgG1-G28-1 HL (H7)	96.7%	.08%	.77%

图 16

经单特异性或多特异性融合蛋白孵育24小时后，
培养液中膜联蛋白和/或PI阳性B细胞的百分比

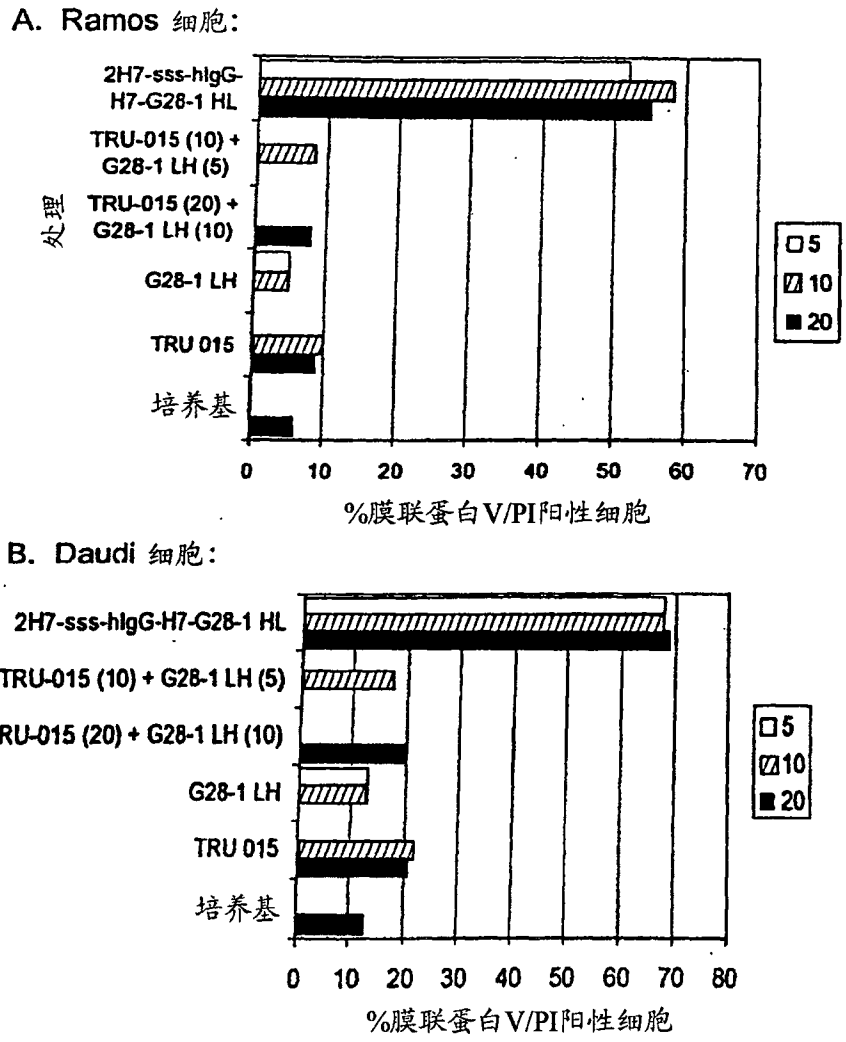
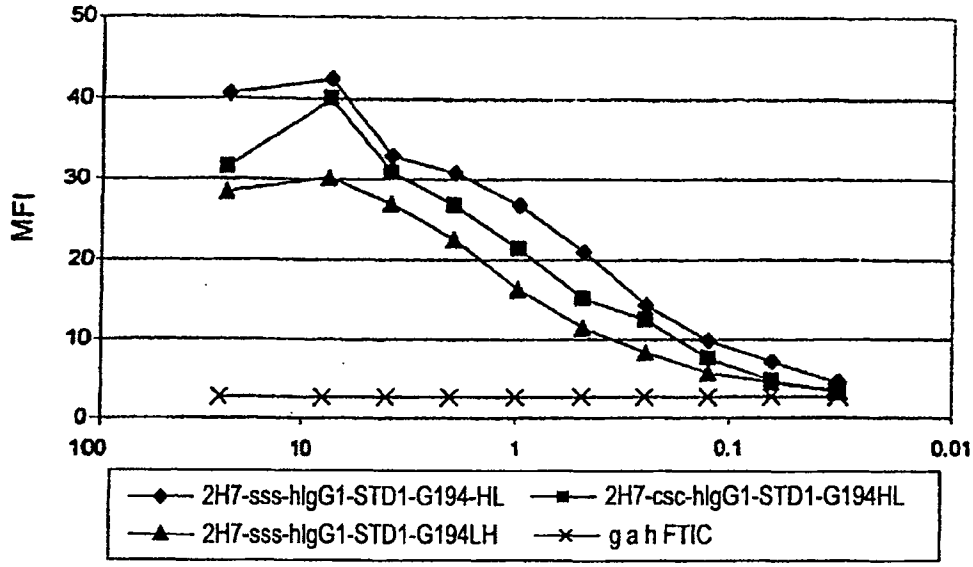


图 17

A. 2H7-hIgG-G19-4融合蛋白与Jurkat细胞的结合



B. 2H7-hIgG-G19-4与WIL2S细胞的结合

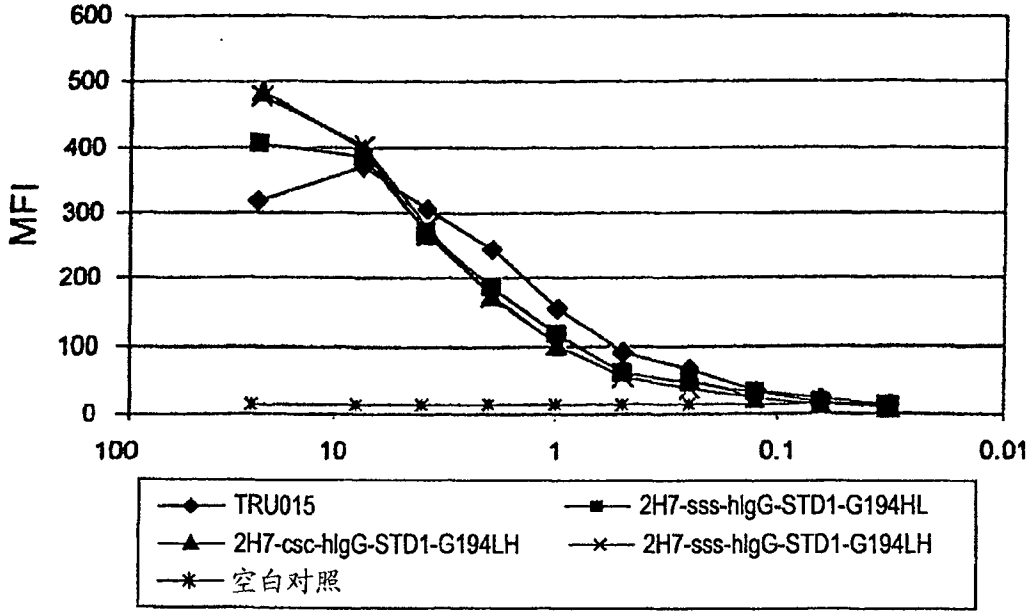
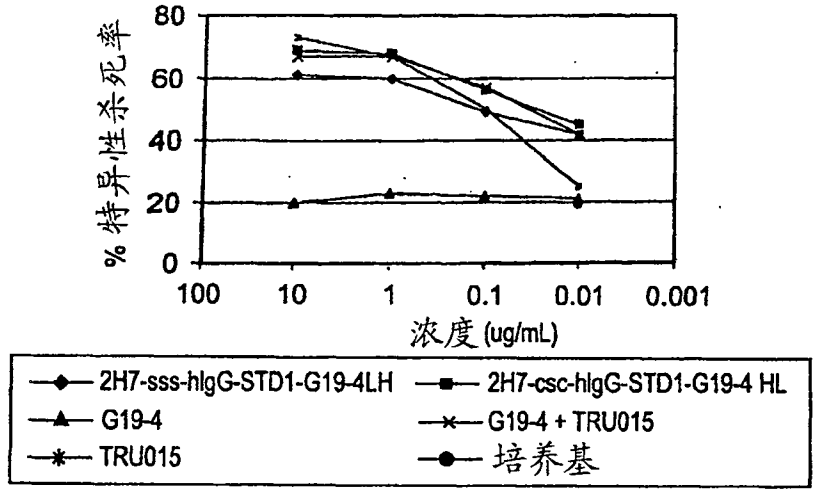


图 18

A. 以25:1使用人PBMC效应物时抗BJAB靶标的2H7-G19-4多特异性融合蛋白的ADCC活性



B. 使用BJAB靶标(30:1, E:T)和2H7-G19-4融合蛋白时NK细胞耗尽的PBMC效应物培养液中的ADCC活性

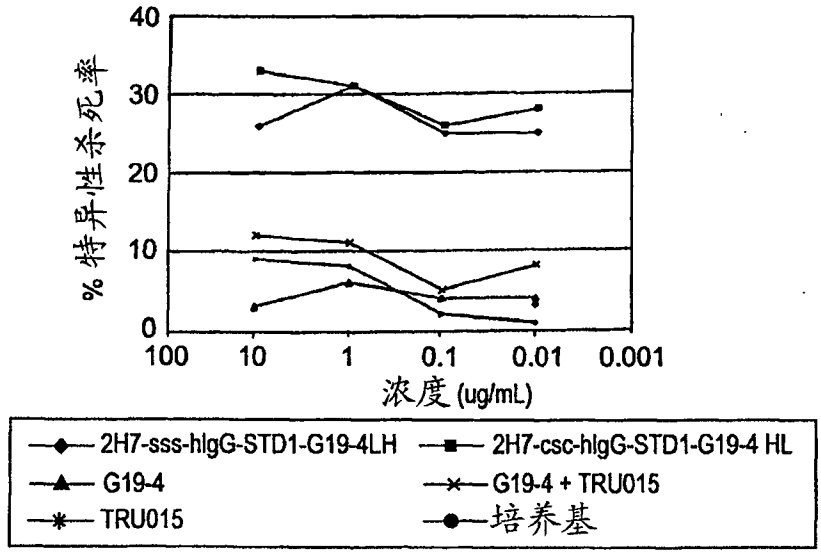


图 19

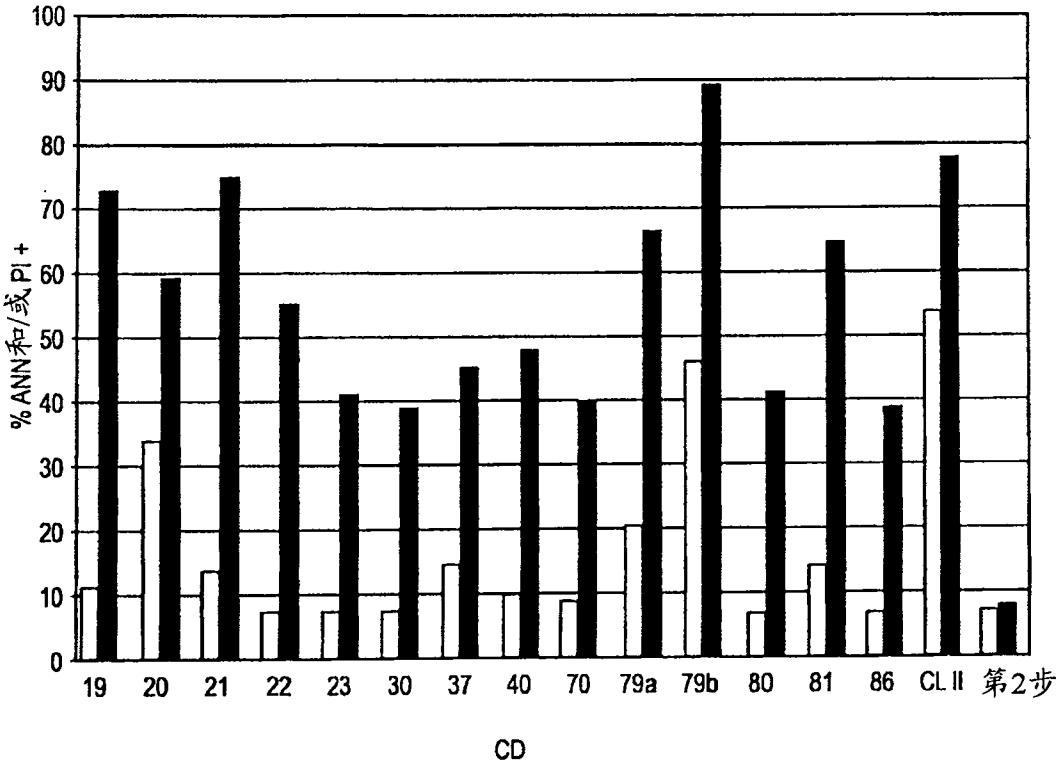

100
90
80
70
60
50
40
30
20
10
0
% ANN和/或 PI +
19
20
21
22
23
30
37
40
70
79a
79b
80
81
86
CL II
第2步
CD

图 20

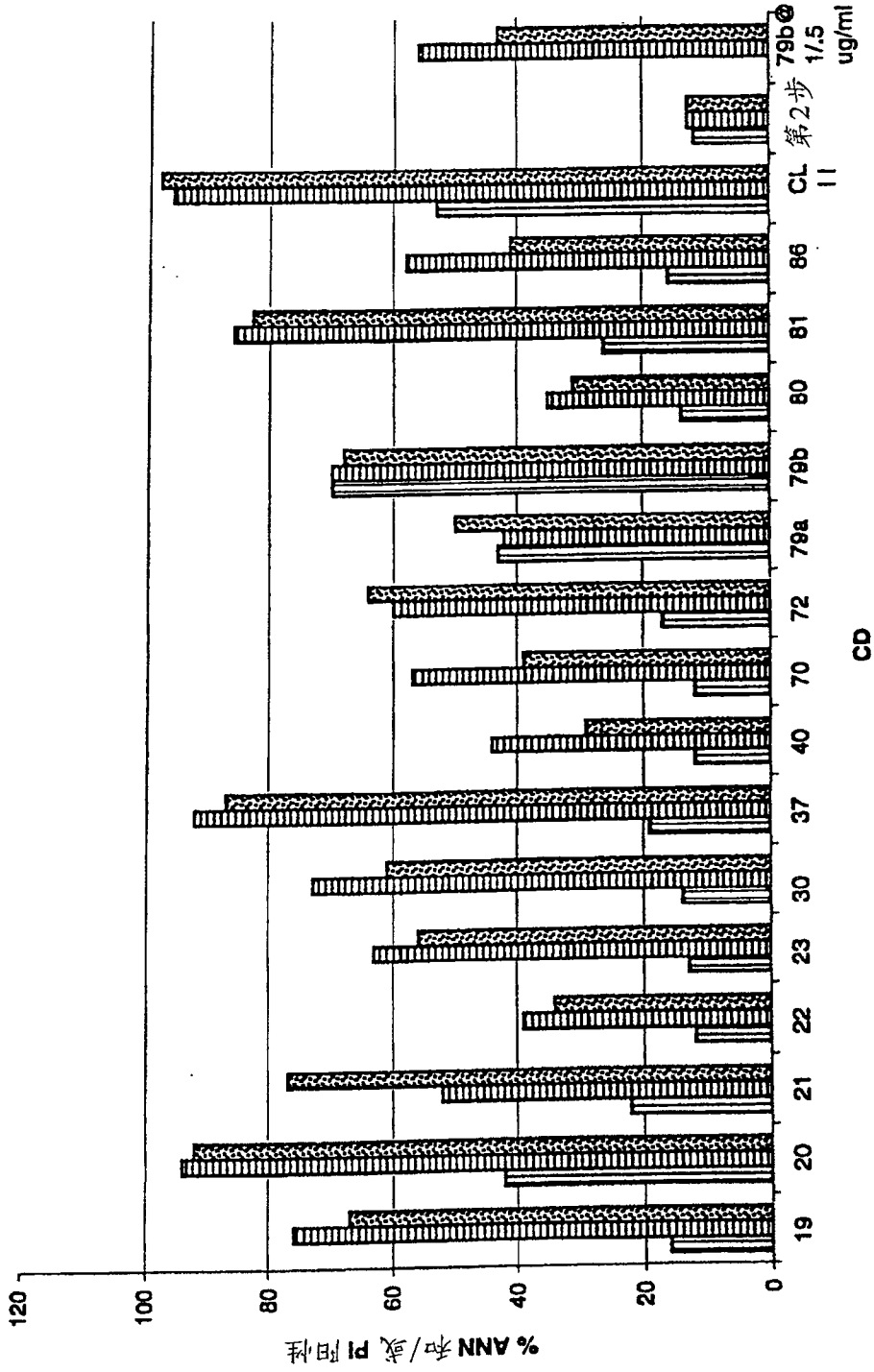

% ANN 和/或 PI 阳性
120
100
80
60
40
20
0
19
20
21
22
23
30
37
40
70
72
79a
79b
80
81
86
CL 11
第2步 1/5 ug/ml
79b@
CD

图 21

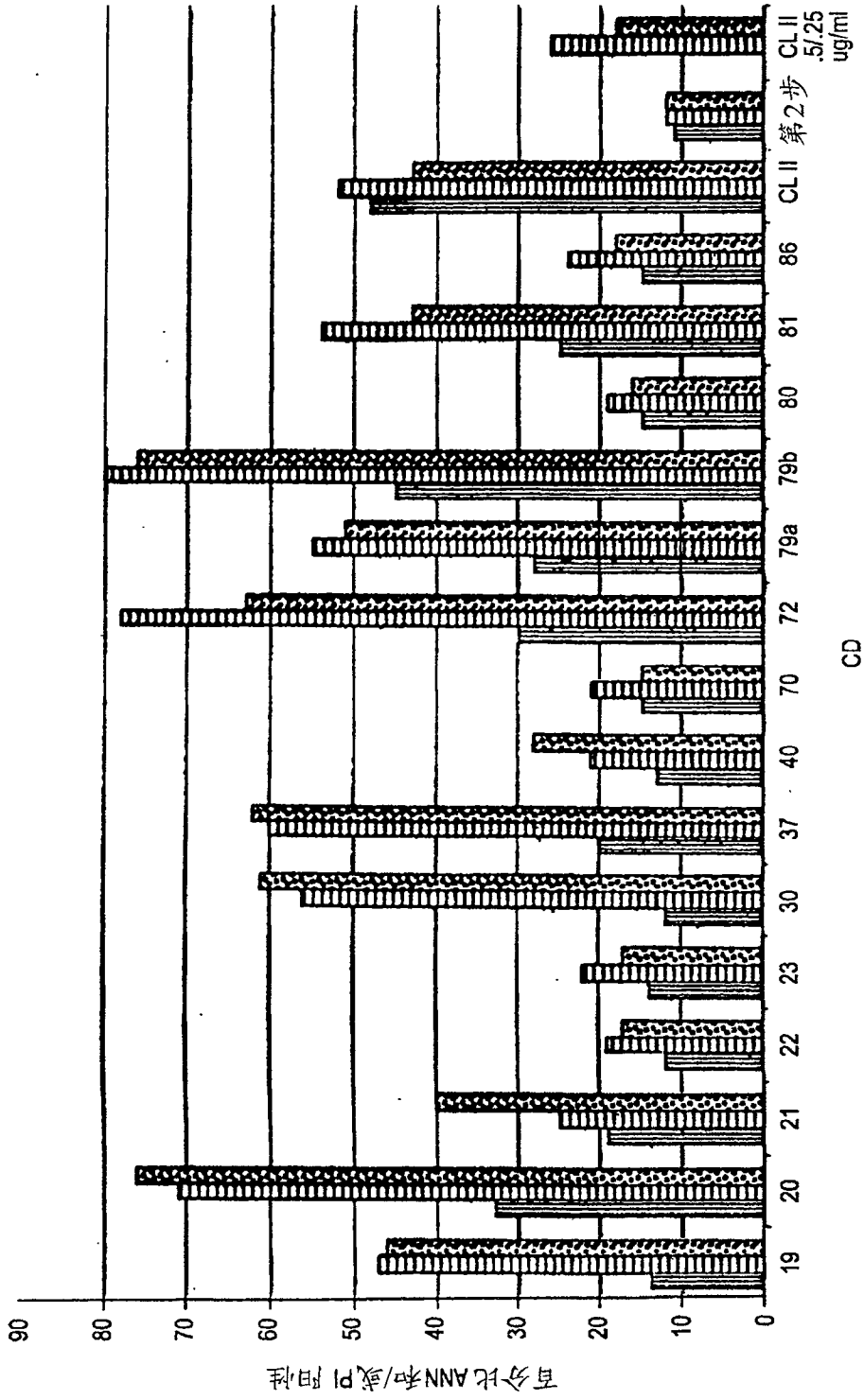

百分比ANN和/或PI 阳性
0
10
20
30
40
50
60
70
80
90
19
20
21
22
23
30
37
40
70
72
79a
79b
80
81
86
CL II
第2步
CL II .5/.25 ug/ml
CD

图 22

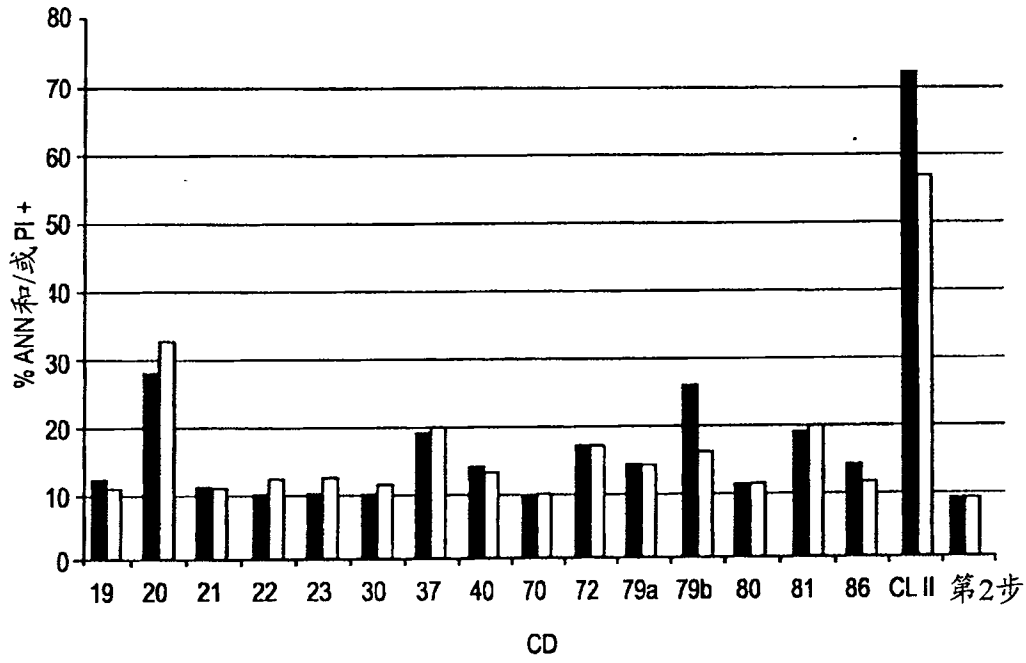

80
70
60
50
40
30
20
10
0
% ANN和/或 PI +
19
20
21
22
23
30
37
40
70
72
79a
79b
80
81
86
CL II
第2步
CD

图 23

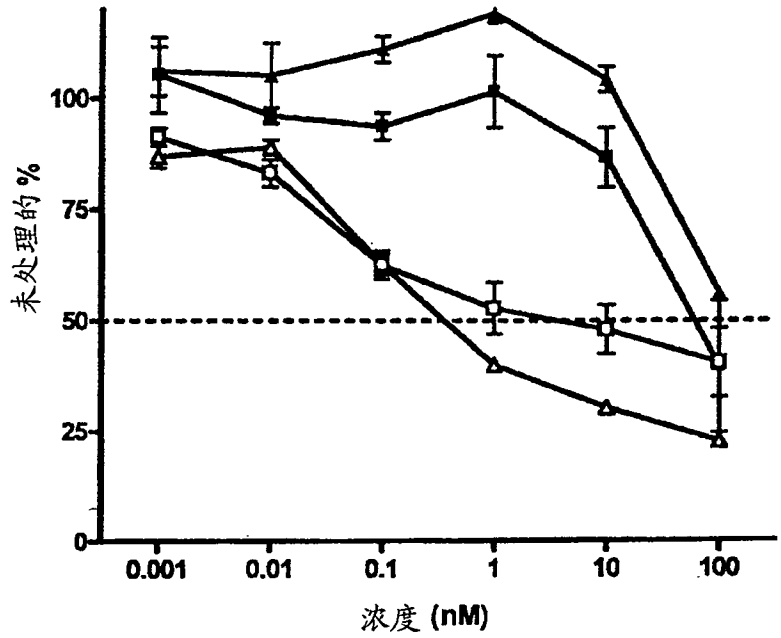


图 24

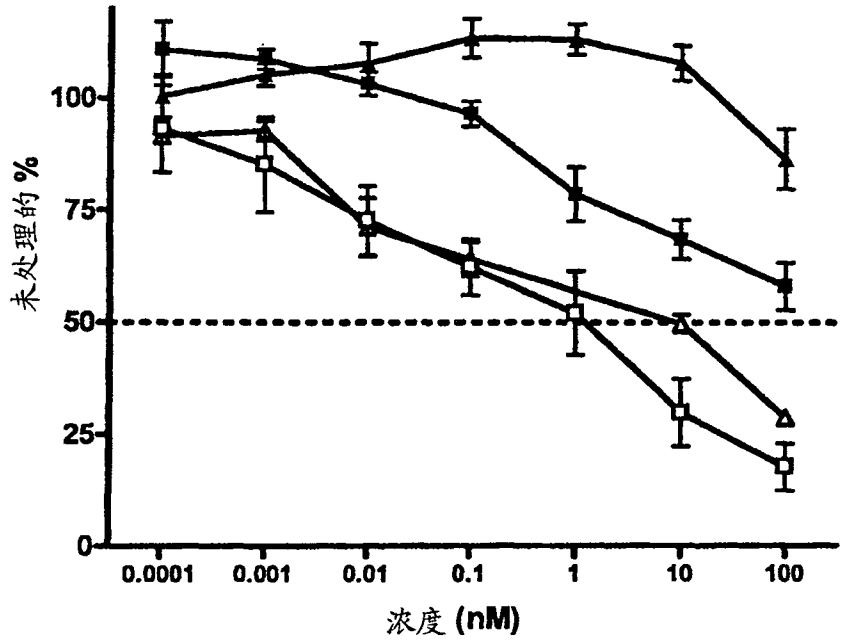


图 25

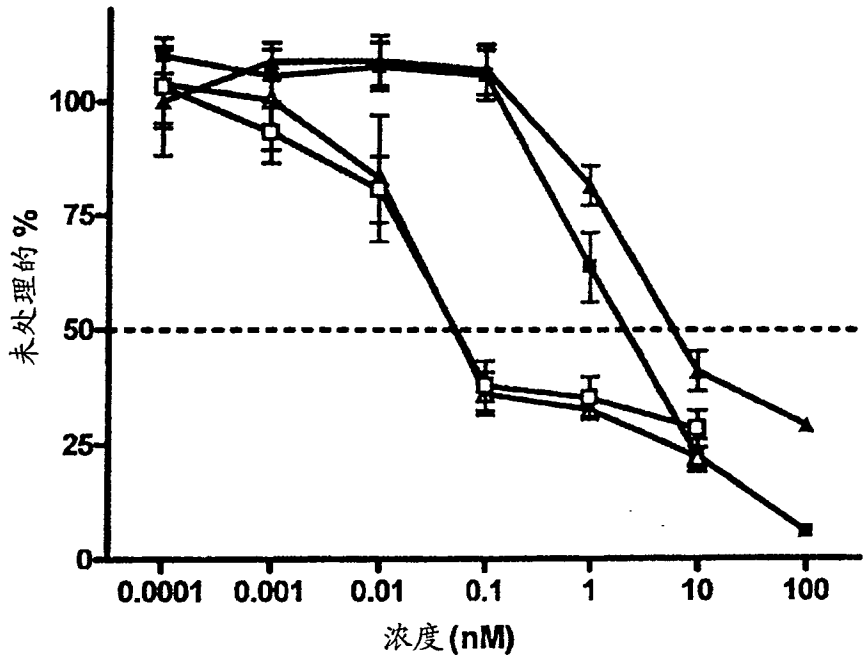


图 26

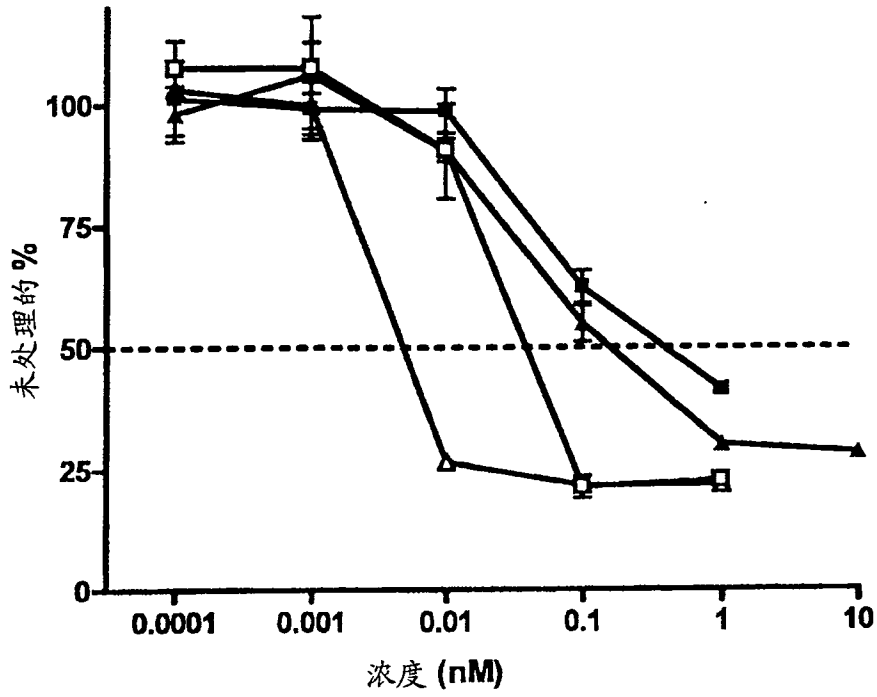


图 27

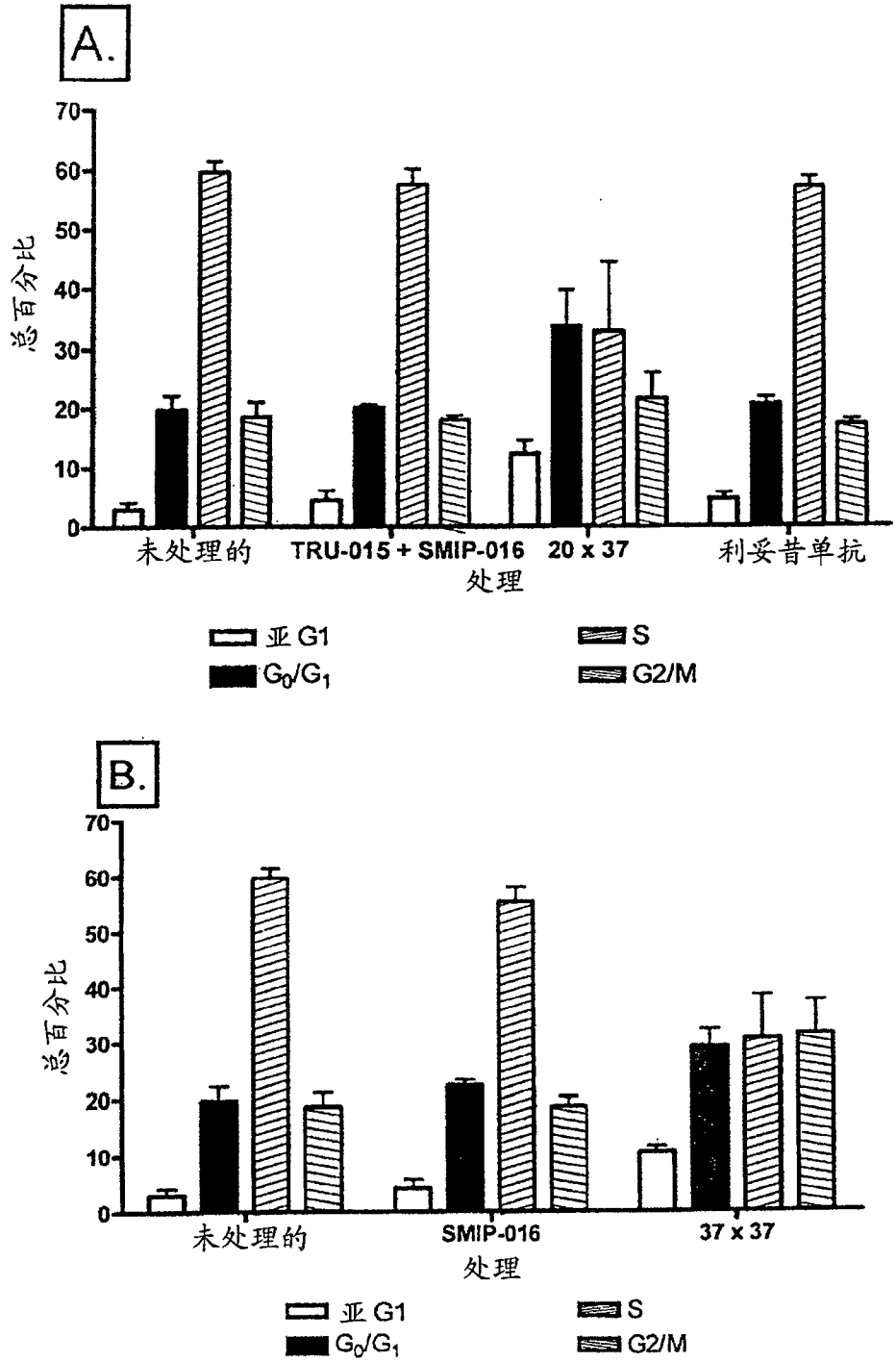

A.
总百分比
70
60
50
40
30
20
10
0
未处理的
TRU-015 + SMIP-016
20 x 37
利妥昔单抗
处理
亚 G1
S
G_0/G_1
G2/M
B.
总百分比
70
60
50
40
30
20
10
0
未处理的
SMIP-016
37 x 37
处理
亚 G1
S
G_0/G_1
G2/M

图 28

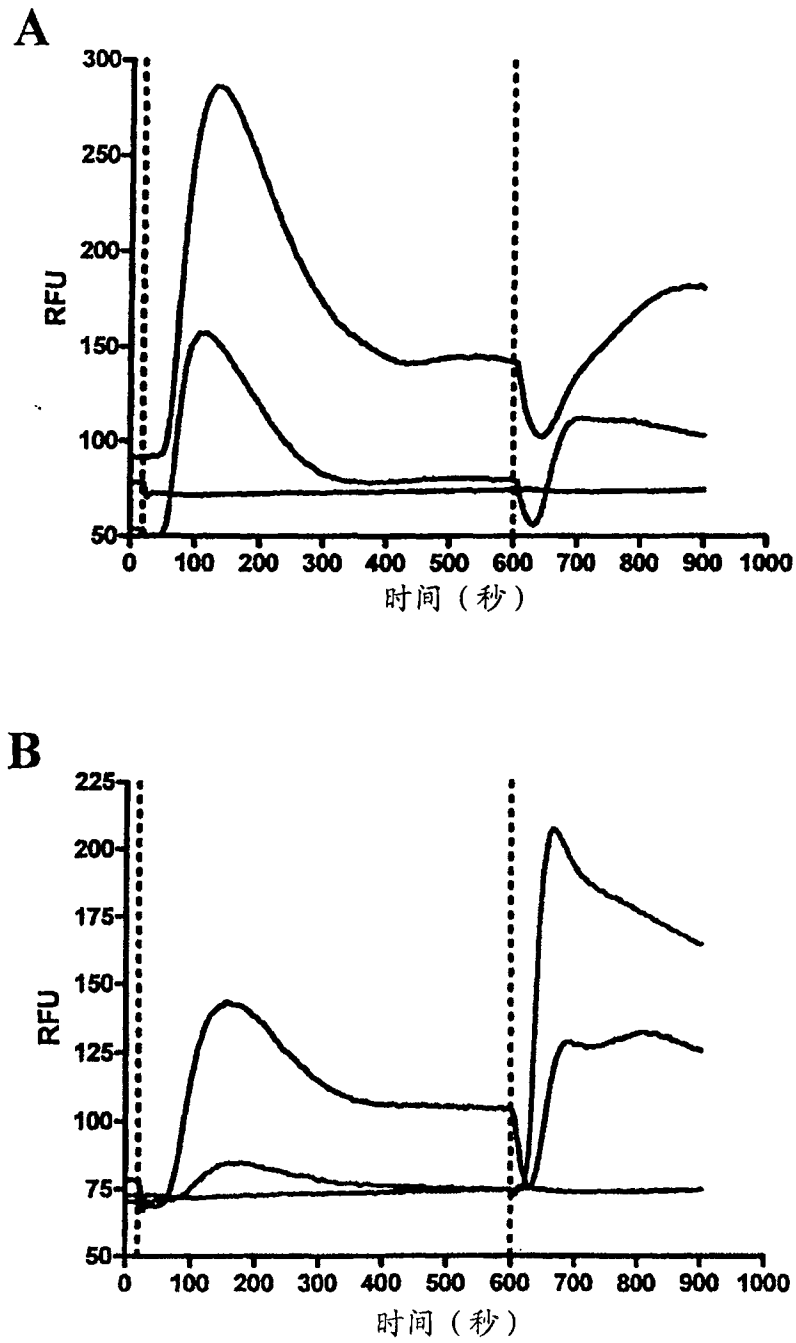

A
300
250
200
150
100
50
RFU
0 100 200 300 400 500 600 700 800 900 1000
时间（秒）
B
225
200
175
150
125
100
75
50
RFU
0 100 200 300 400 500 600 700 800 900 1000
时间（秒）

图 29

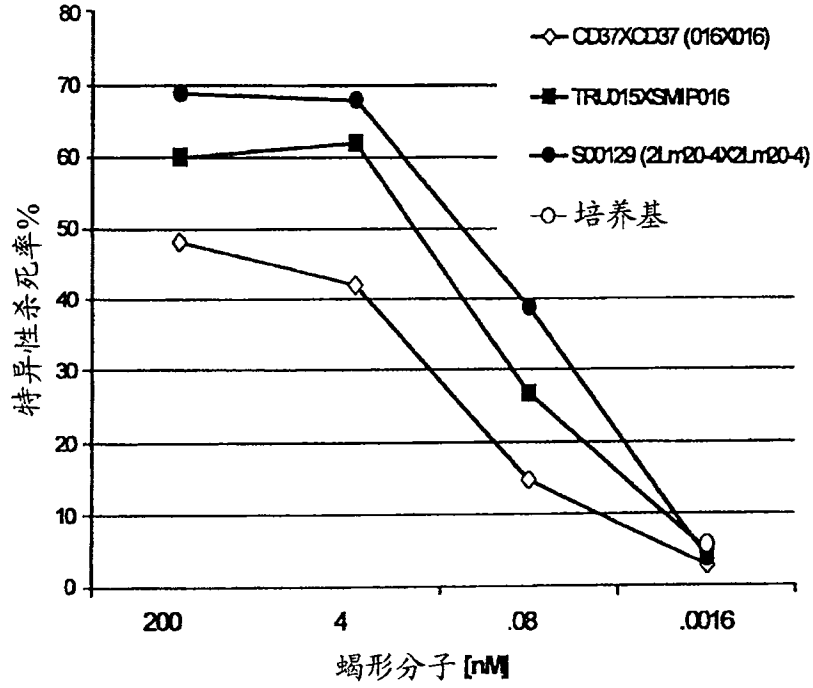

80
70
60
50
40
30
20
10
0
特异性杀死率%
200
4
.08
.0016
蝎形分子 [nM]
CD37XCD37 (016X016)
TRU015XSMIP016
S00129 (2Lm20-4X2Lm20-4)
培养基

图 30

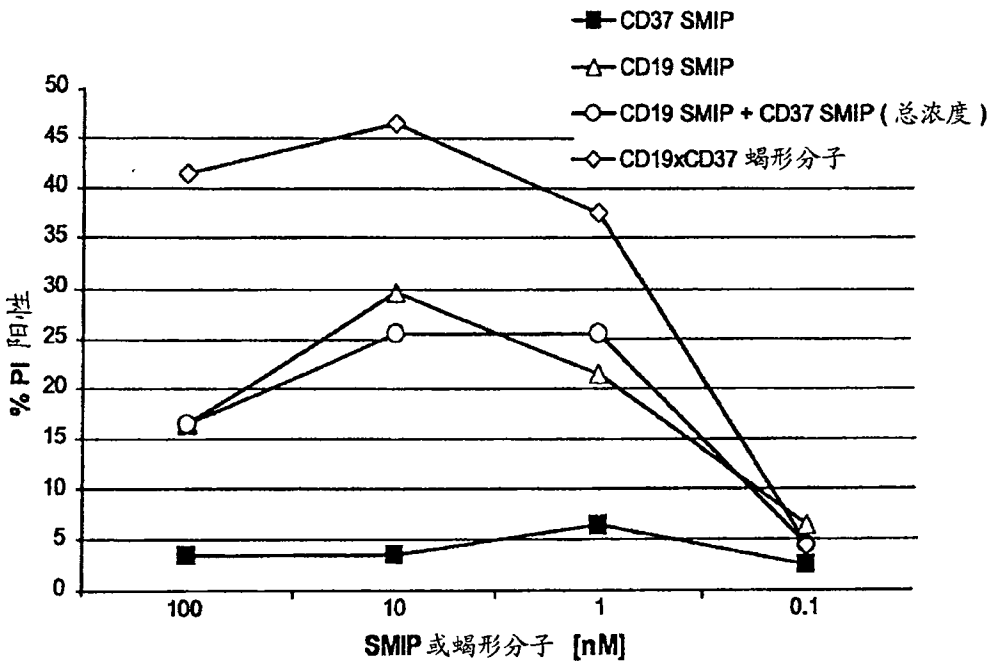

CD37 SMIP
CD19 SMIP
CD19 SMIP + CD37 SMIP (总浓度)
CD19xCD37 蝎形分子
% PI 阳性
50
45
40
35
30
25
20
15
10
5
0
100
10
1
0.1
SMIP 或蝎形分子 [nM]

图 31

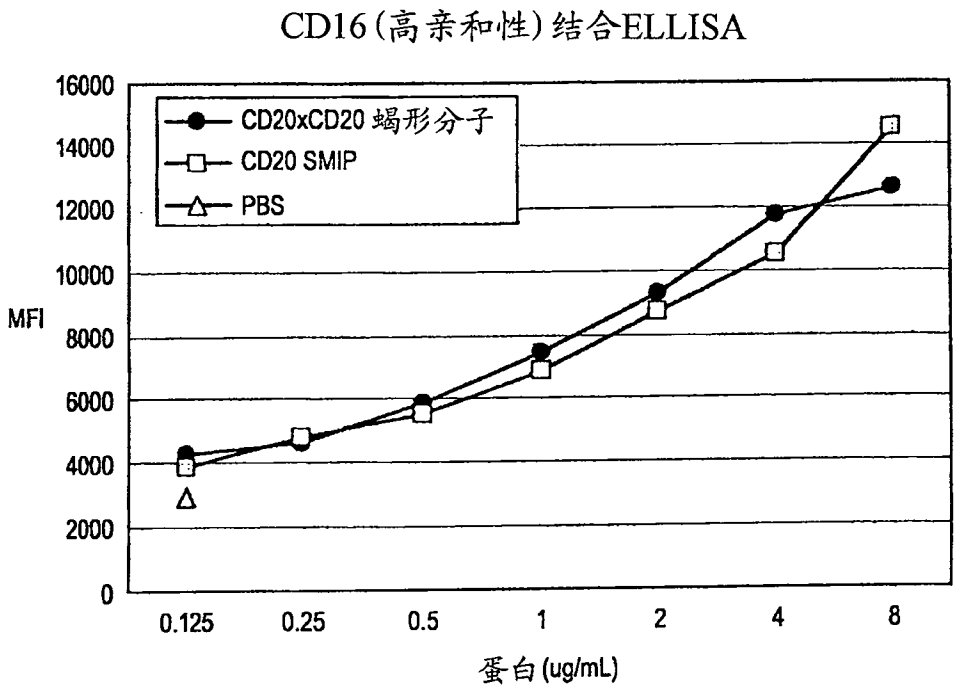

CD16(高亲和性)结合ELLISA
CD20xCD20 蝎形分子
CD20 SMIP
PBS
MFI
16000
14000
12000
10000
8000
6000
4000
2000
0
0.125
0.25
0.5
1
2
4
8
蛋白(ug/mL)

图 32A

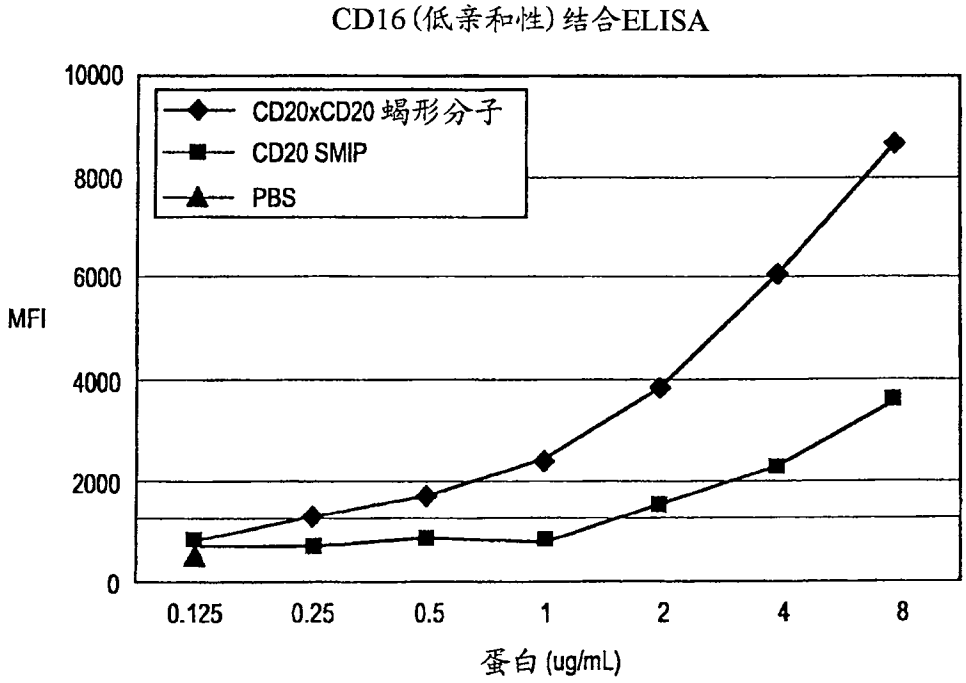
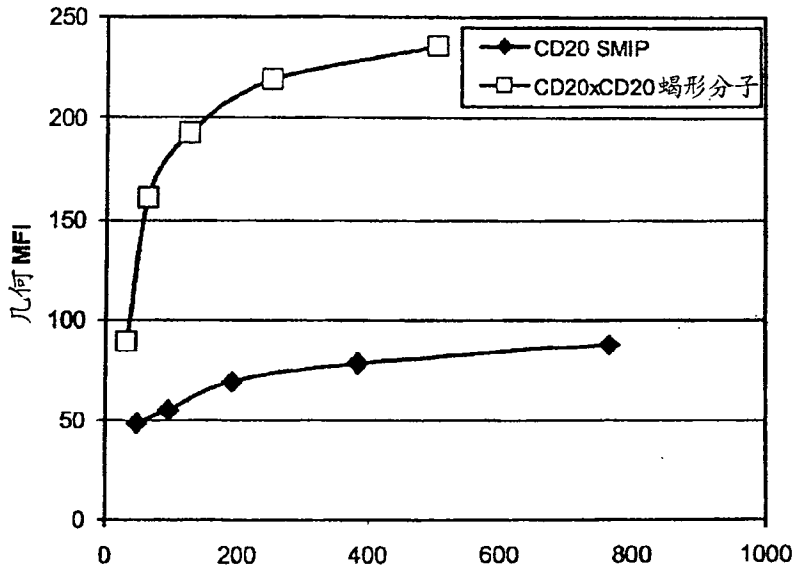

CD16(低亲和性)结合ELISA
10000
8000
6000
4000
2000
0
MFI
CD20xCD20 蝎形分子
CD20 SMIP
PBS
0.125
0.25
0.5
1
2
4
8
蛋白(ug/mL)

图 32B

CD16与结合SMIP/蝎形分子的WIL2S的低亲和性结合



CD16与结合SMIP/蝎形分子的WIL2S的高亲和性结合

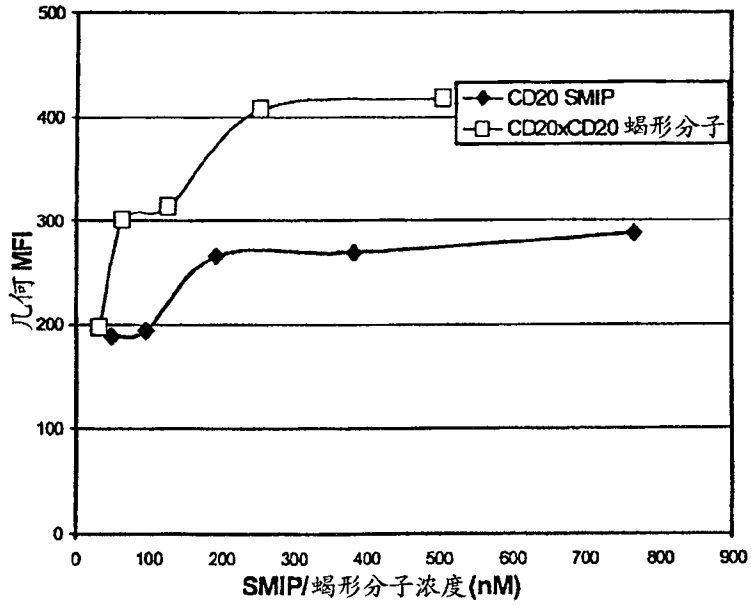


图 33

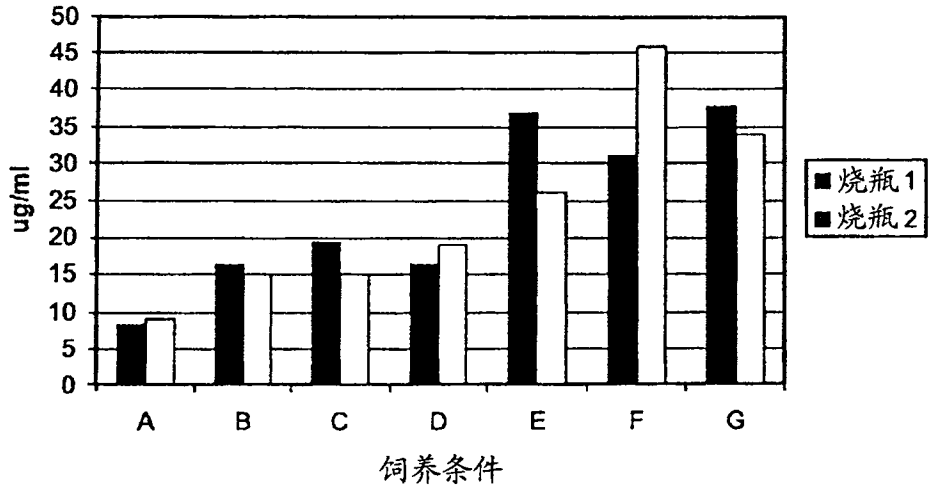


图 34

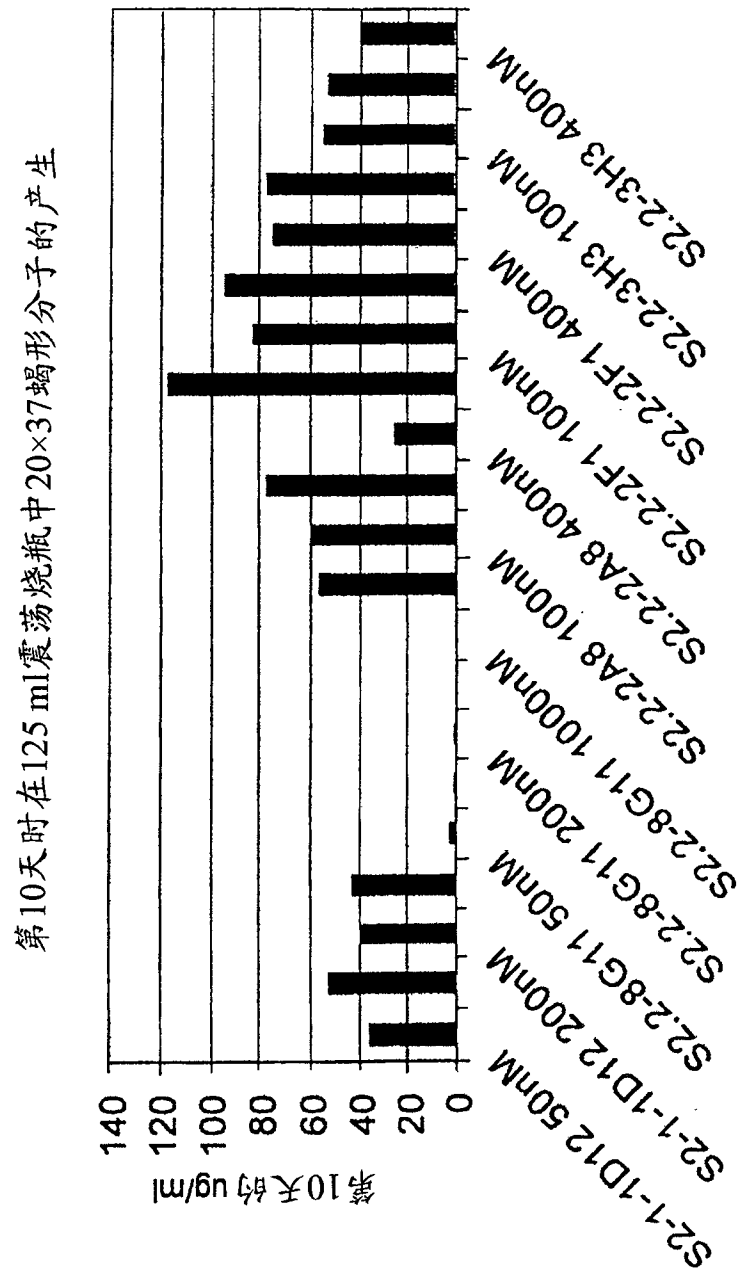

第10天时在125 ml震荡烧瓶中20×37蝎形分子的产生
第10天的 ug/ml
140
120
100
80
60
40
20
0
S2-1-1D12 50nM
S2-1-1D12 200nM
S2.2-8G11 50nM
S2.2-8G11 200nM
S2.2-8G11 1000nM
S2.2-2A8 100nM
S2.2-2A8 400nM
S2.2-2F1 100nM
S2.2-2F1 400nM
S2.2-3H3 100nM
S2.2-3H3 400nM
细胞系/MTX浓度

图 35

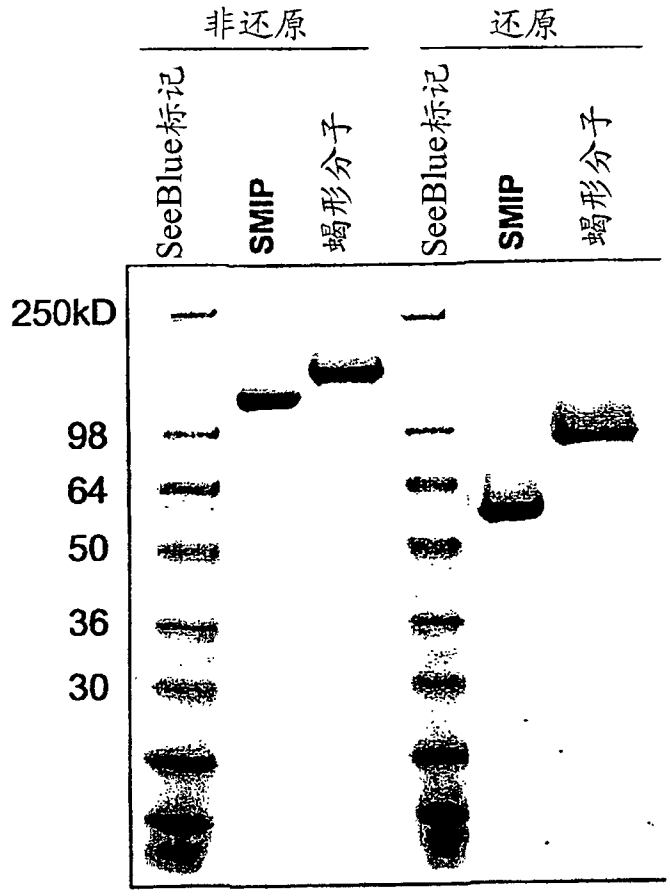

非还原
还原
SeeBlue标记
SMIP
蝎形分子
SeeBlue标记
SMIP
蝎形分子
250kD
98
64
50
36
30

图 36

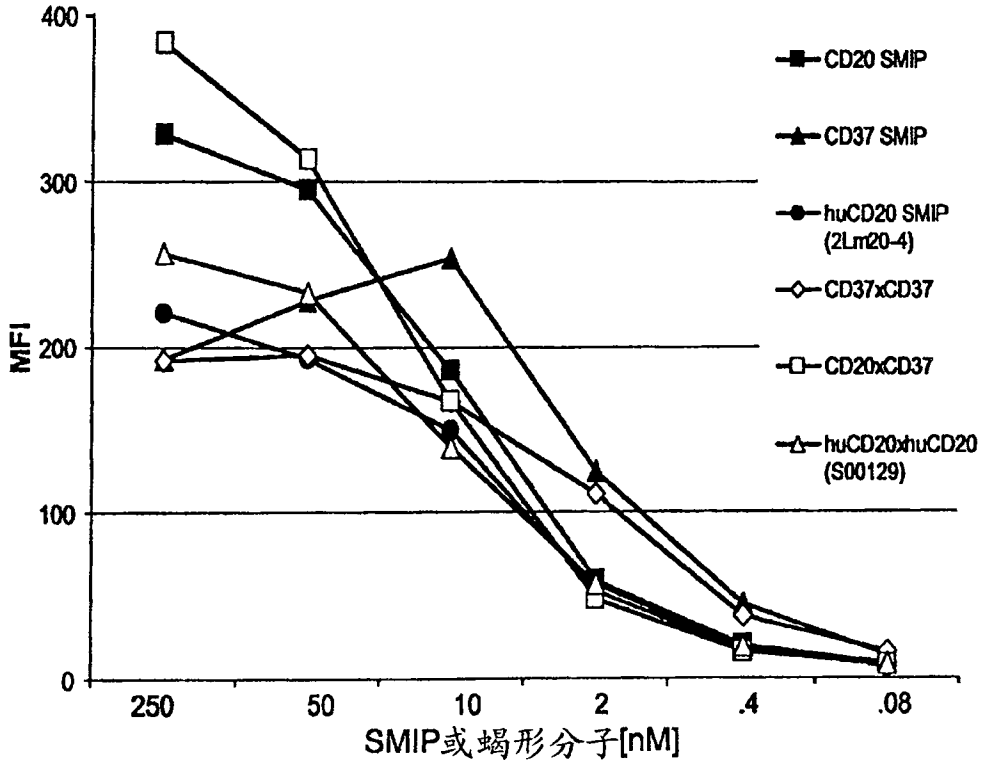

400
300
200
100
0
MFI
250
50
10
2
.4
.08
SMIP或蝎形分子[nM]
CD20 SMIP
CD37 SMIP
huCD20 SMIP (2Lm20-4)
CD37xCD37
CD20xCD37
huCD20xhuCD20 (S00129)

图 37

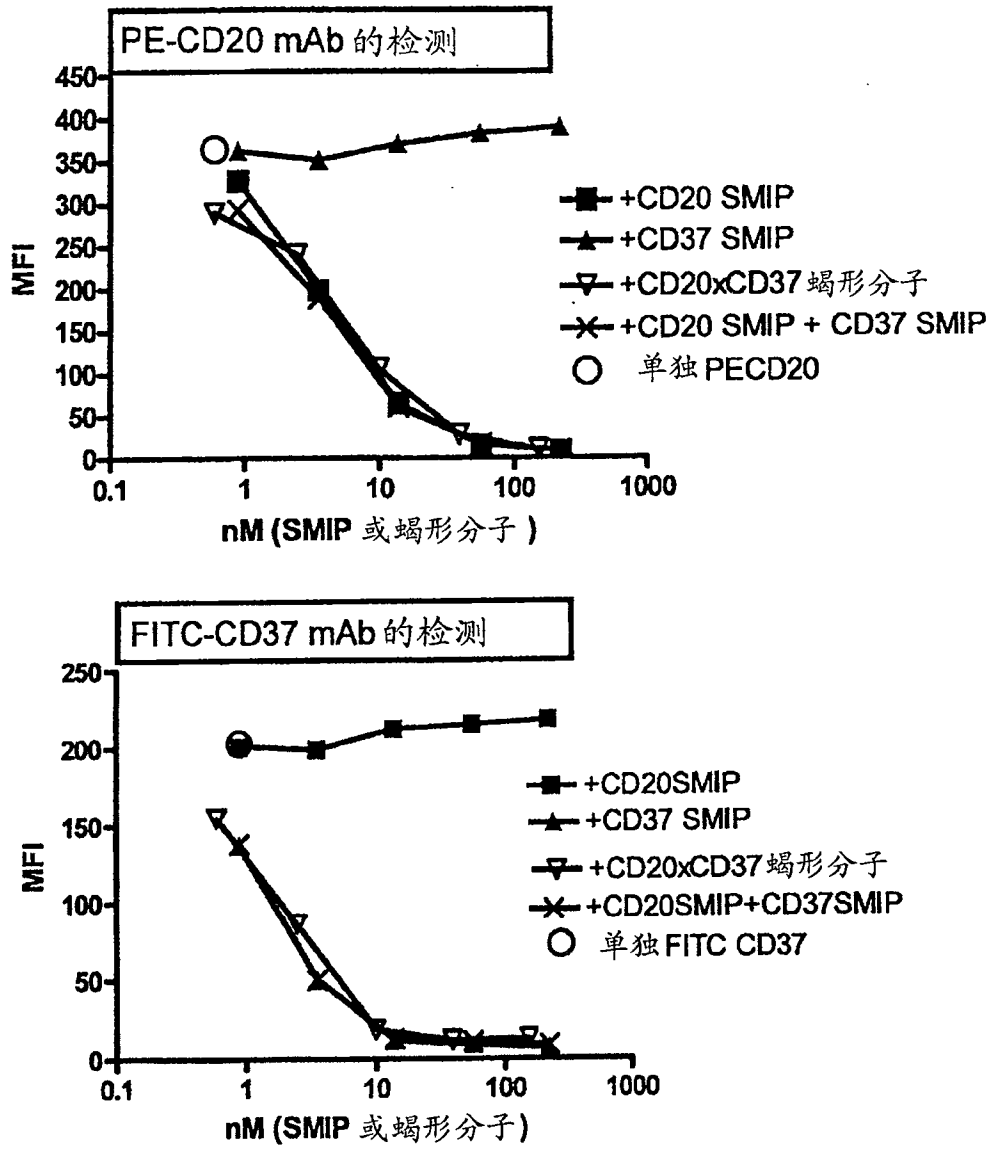

PE-CD20 mAb 的检测
MFI
450
400
350
300
250
200
150
100
50
0
0.1
1
10
100
1000
nM (SMIP 或蝎形分子)
+CD20 SMIP
+CD37 SMIP
+CD20xCD37蝎形分子
+CD20 SMIP + CD37 SMIP
单独PECD20
FITC-CD37 mAb 的检测
MFI
250
200
150
100
50
0
0.1
1
10
100
1000
nM (SMIP 或蝎形分子)
+CD20SMIP
+CD37 SMIP
+CD20xCD37蝎形分子
+CD20SMIP+CD37SMIP
单独FITC CD37

图 38

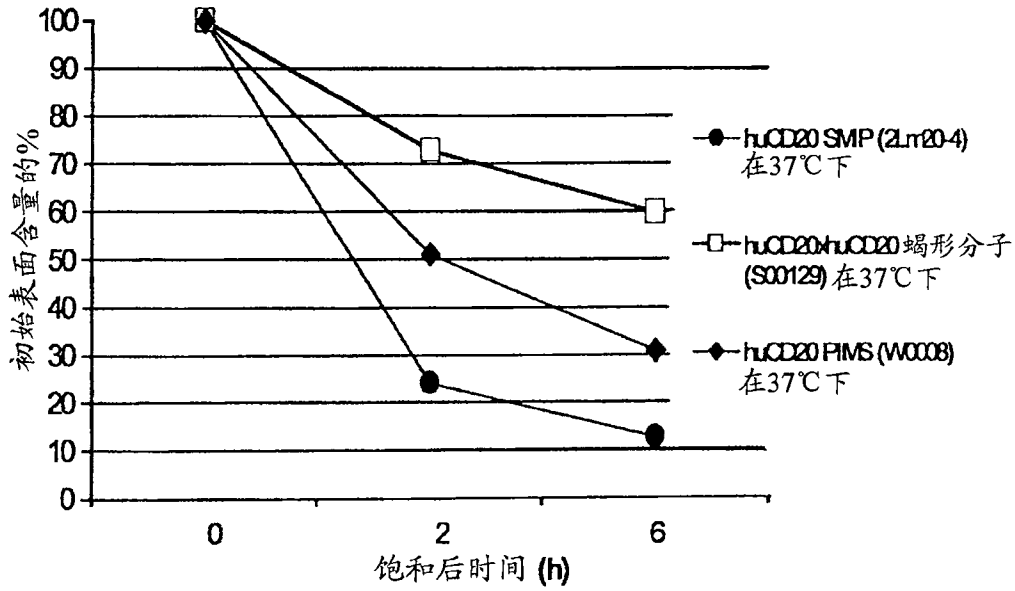

100
90
80
70
60
50
40
30
20
10
0
初始表面含量的%
0
2
6
饱和后时间 (h)
huCD20 SMP (2Lm20-4) 在37℃下
huCD20/huCD20 蝎形分子 (S00129) 在37℃下
huCD20 PIMS (W0008) 在37℃下

图 39

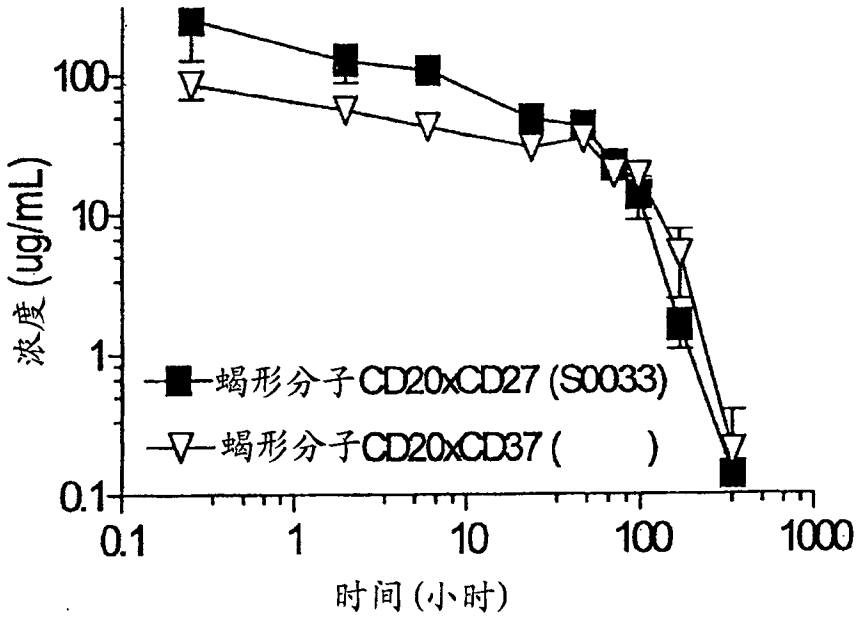

100
10
1
0.1
浓度 (ug/mL)
0.1
1
10
100
1000
时间 (小时)
蝎形分子CD20xCD27 (S0033)
蝎形分子CD20xCD37 ( )

图 40

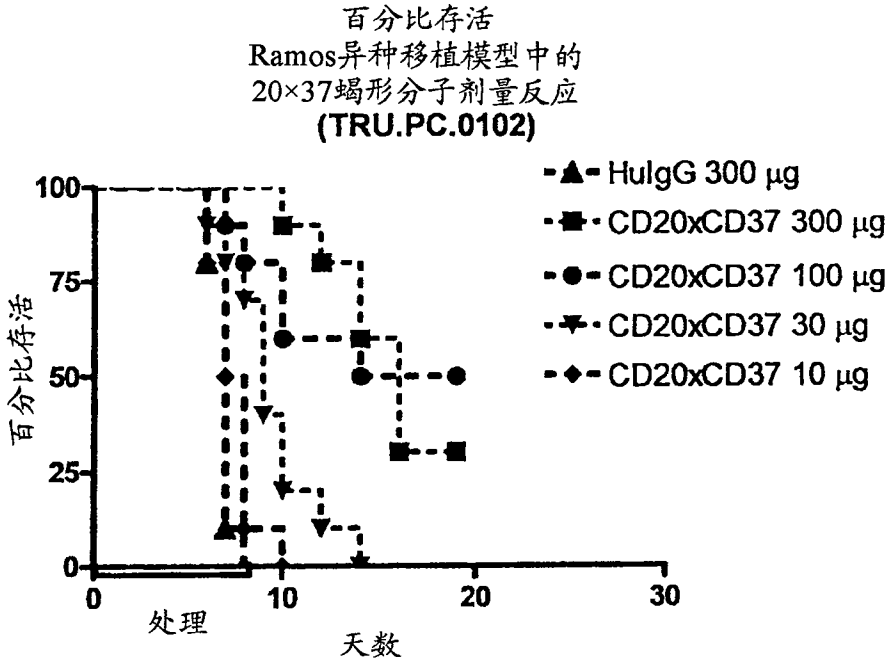
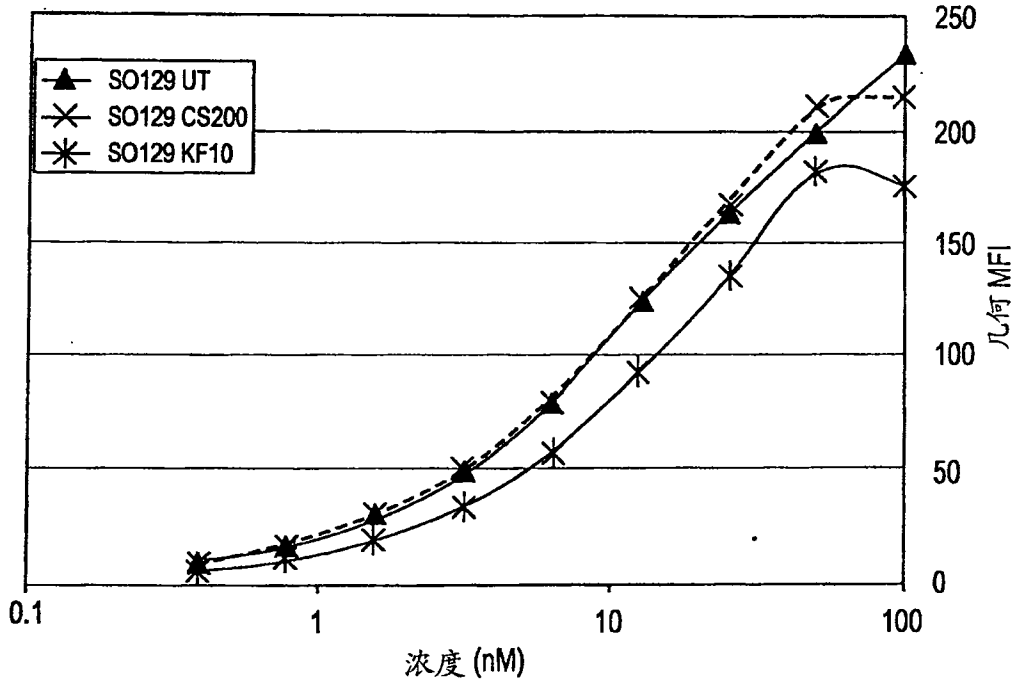

百分比存活
Ramos异种移植模型中的
20×37蝎形分子剂量反应
(TRU.PC.0102)
HulgG 300 μg
CD20xCD37 300 μg
CD20xCD37 100 μg
CD20xCD37 30 μg
CD20xCD37 10 μg
100
75
50
25
0
百分比存活
0
10
20
30
处理
天数

图 41

SO129亲本分子和糖变异体的靶标细胞结合



经2lm20-4染色的WIL2S，SO129具有或没有10 μM碳水化合物改良剂α-甘露糖苷酶(过量)、200 μM栗精胺

图 42

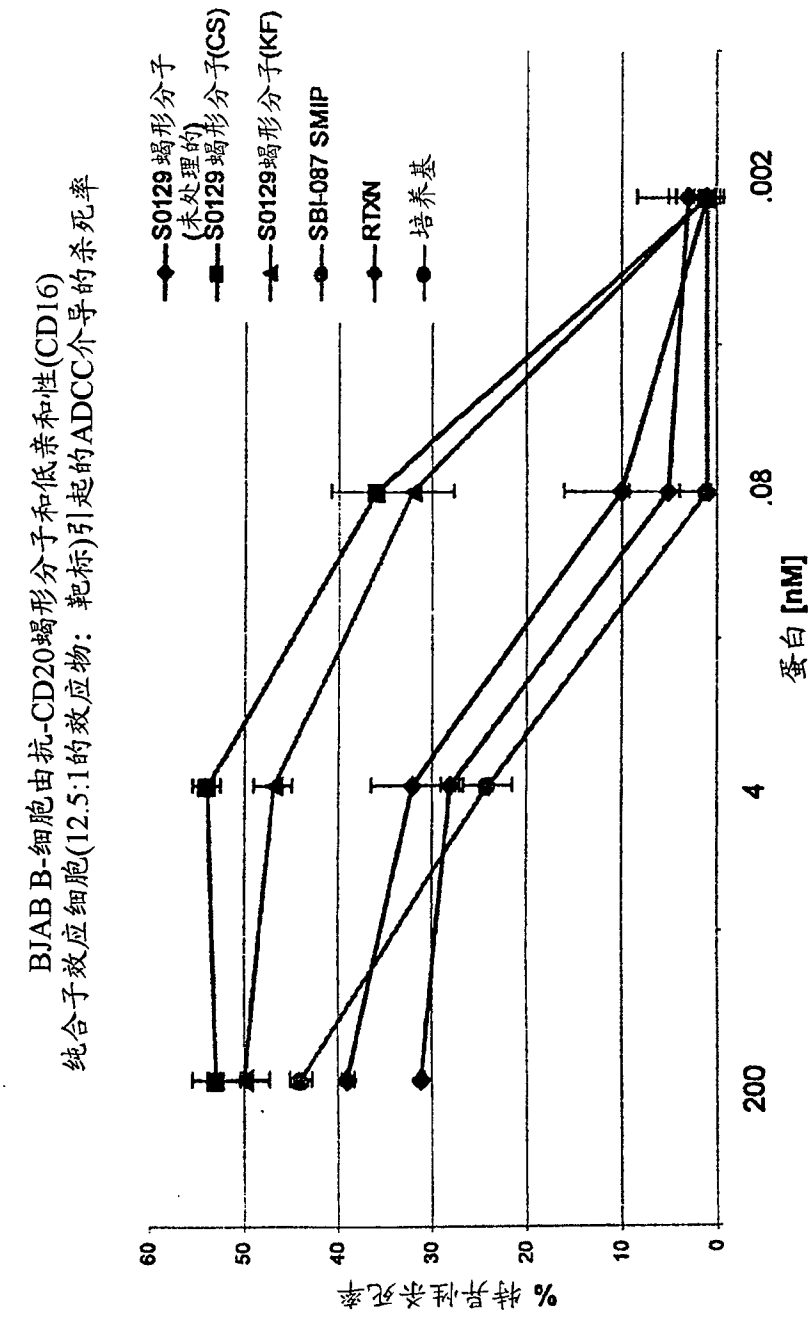

BJAB B-细胞由抗-CD20蝎形分子和低亲和性(CD16)
纯合子效应细胞(12.5:1的效应物：靶标)引起的ADCC介导的杀死率
S0129蝎形分子
(未处理的)
S0129蝎形分子(CS)
S0129蝎形分子(KF)
SBI-087 SMIP
RTXN
培养基
% 特异性杀死率
0
10
20
30
40
50
60
200
4
.08
.002
蛋白 [nM]

图 43

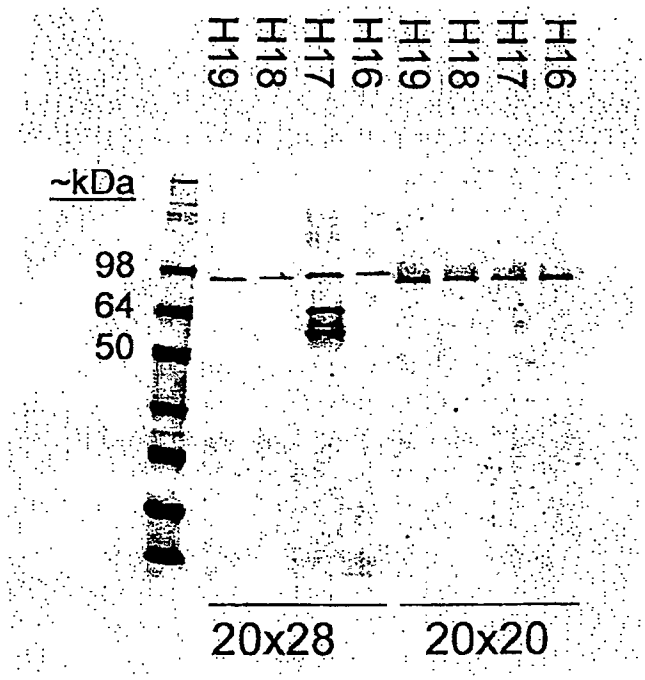


图 44

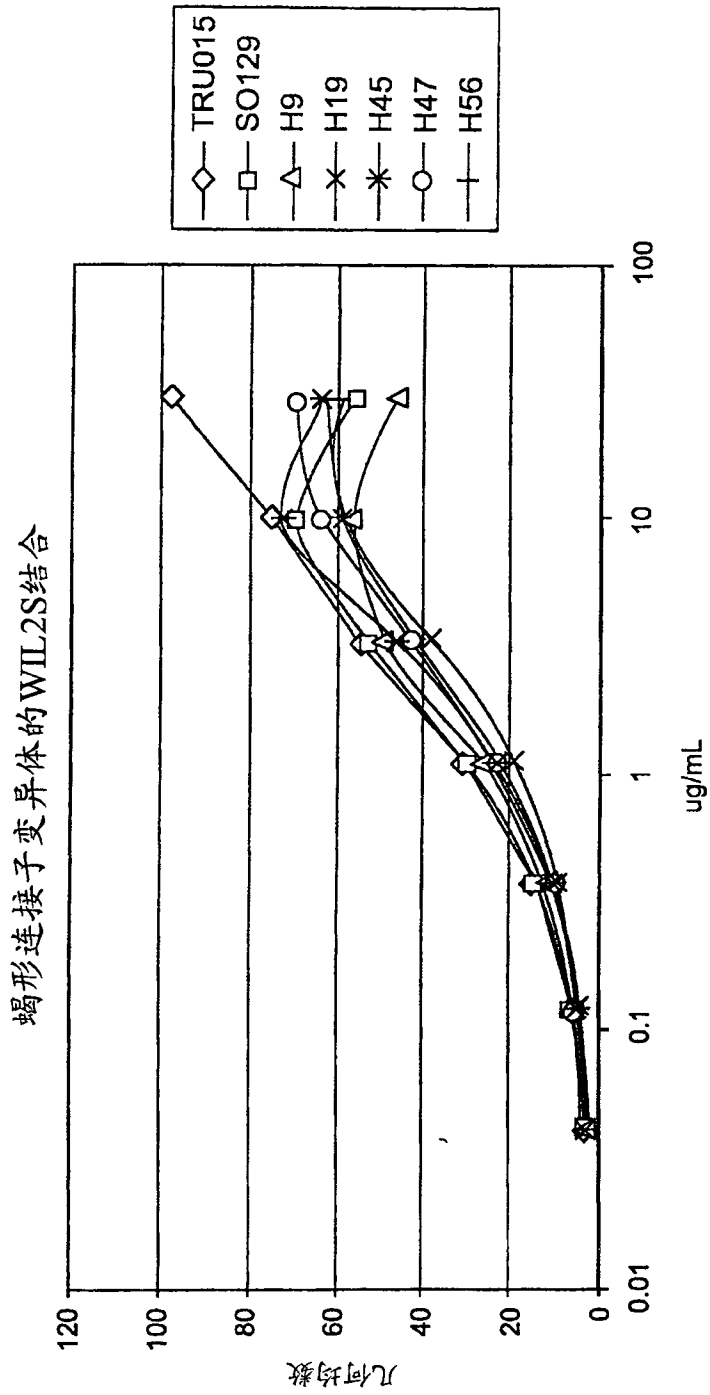

蝎形连接子变异体的WIL2S结合
几何均数
ug/mL
TRU015
SO129
H9
H19
H45
H47
H56
120
100
80
60
40
20
0
0.01
0.1
1
10
100

图 45

20×20蝎形分子活性(实心)与SMIP活性(空心)的对比

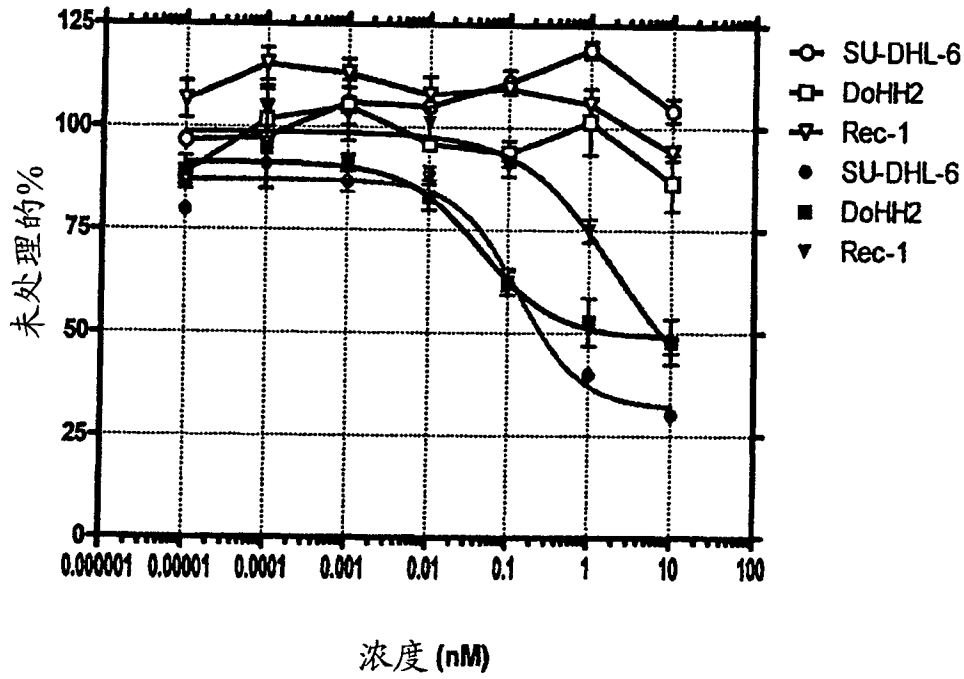


图 46

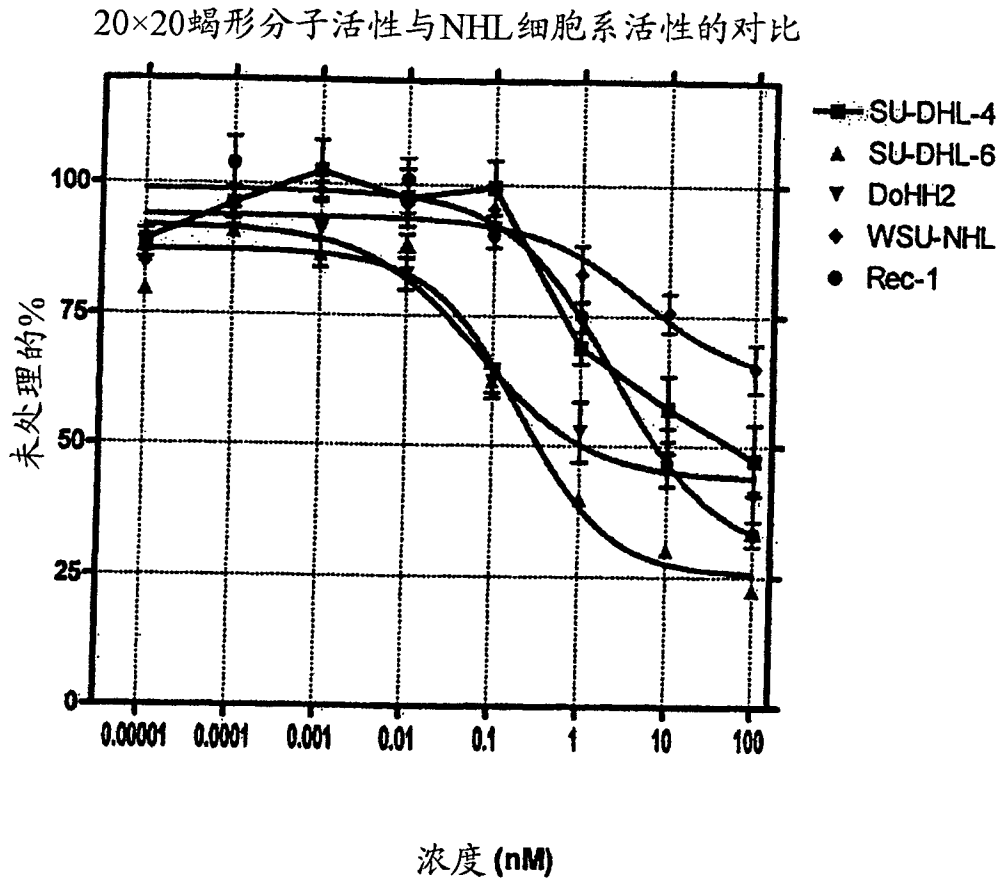

20×20蝎形分子活性与NHL细胞系活性的对比
SU-DHL-4
SU-DHL-6
DoHH2
WSU-NHL
Rec-1
未处理的%
100
75
50
25
0
0.00001
0.0001
0.001
0.01
0.1
1
10
100
浓度 (nM)

图 47

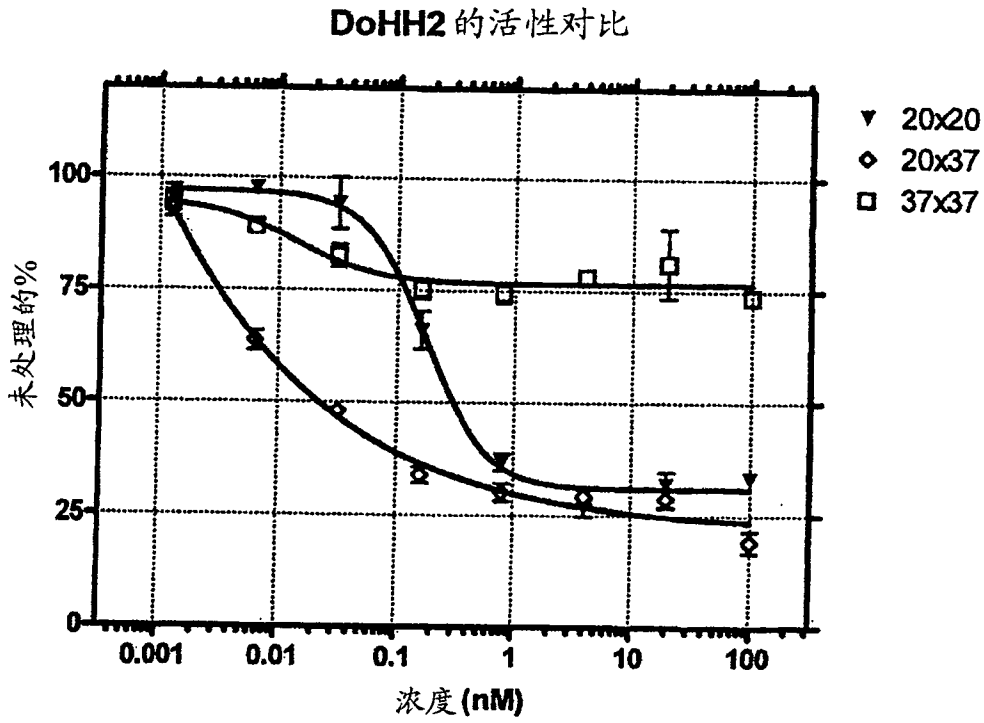

DoHH2的活性对比
20x20
20x37
37x37
未处理的%
100
75
50
25
0
0.001
0.01
0.1
1
10
100
浓度(nM)

图 48

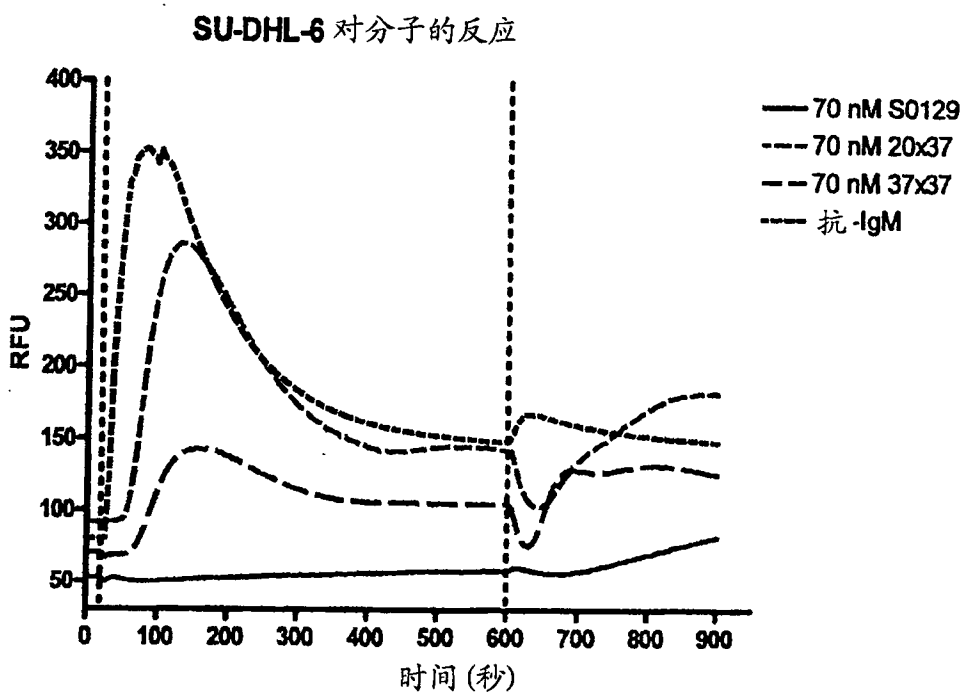

SU-DHL-6 对分子的反应
RFU
400
350
300
250
200
150
100
50
0
100
200
300
400
500
600
700
800
900
时间 (秒)
70 nM S0129
70 nM 20x37
70 nM 37x37
抗 -IgM

图 49

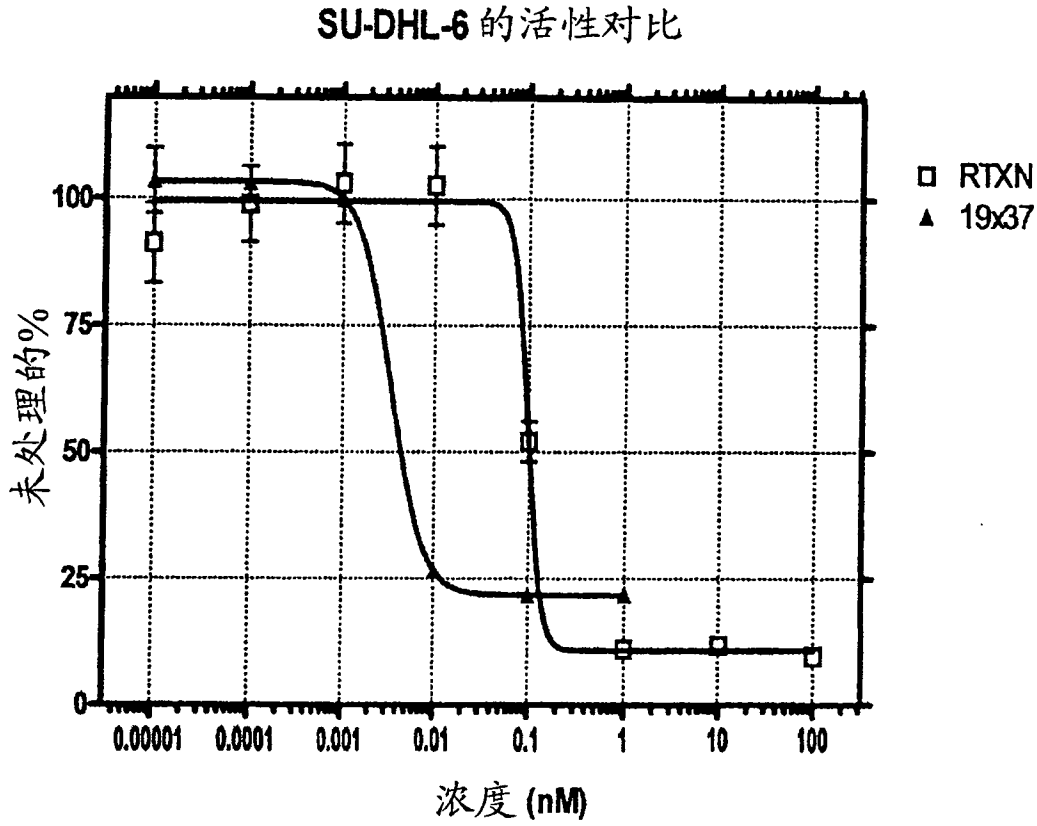

SU-DHL-6 的活性对比
未处理的%
100
75
50
25
0
0.00001
0.0001
0.001
0.01
0.1
1
10
100
浓度 (nM)
RTXN
19x37

图 50

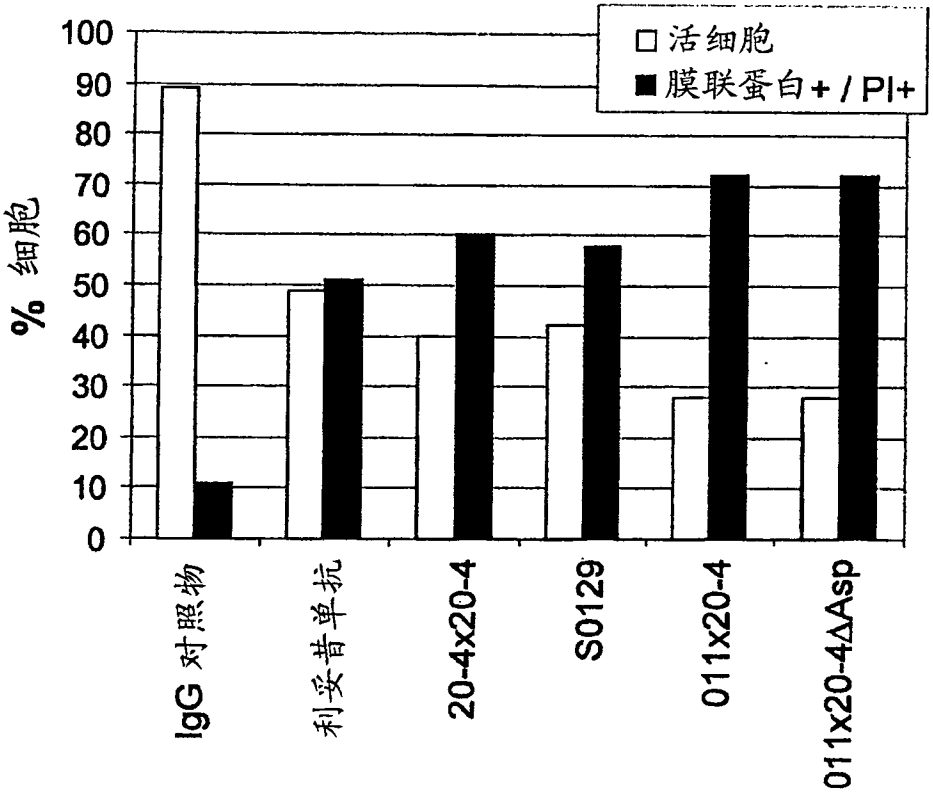

□活细胞
■膜联蛋白+ / PI+
% 细胞
100
90
80
70
60
50
40
30
20
10
0
IgG 对照物
利妥昔单抗
20-4x20-4
S0129
011x20-4
011x20-4ΔAsp

图 51

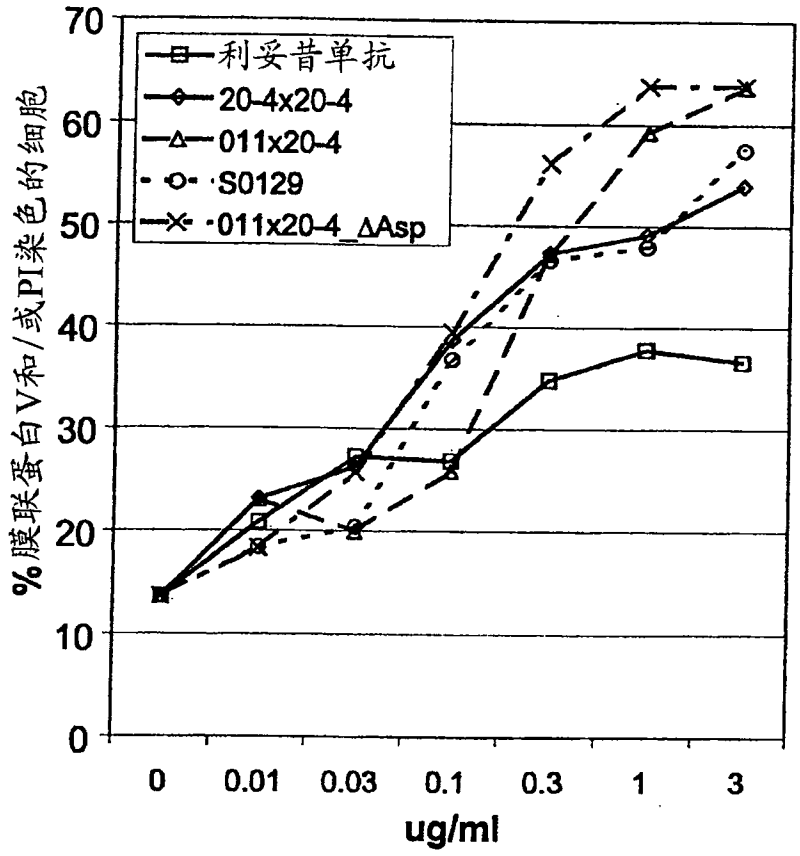

70
60
50
40
30
20
10
0
%膜联蛋白V和/或PI染色的细胞
利妥昔单抗
20-4x20-4
011x20-4
S0129
011x20-4_ΔAsp
0
0.01
0.03
0.1
0.3
1
3
ug/ml

图 52

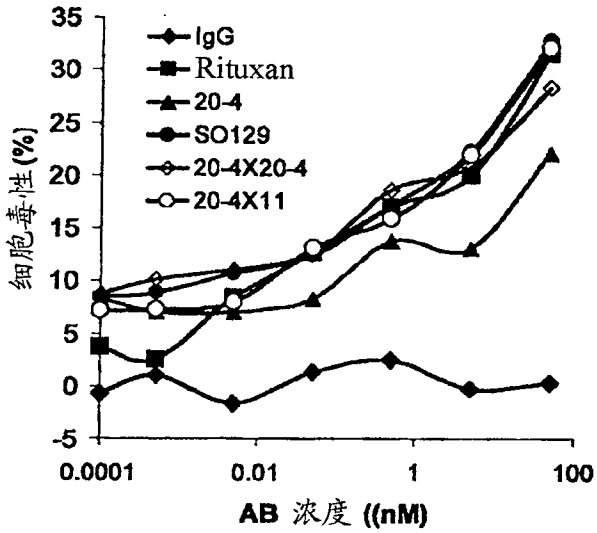

35
30
25
20
15
10
5
0
-5
细胞毒性(%)
IgG
Rituxan
20-4
SO129
20-4X20-4
20-4X11
0.0001
0.01
1
100
AB 浓度 ((nM)

图 53

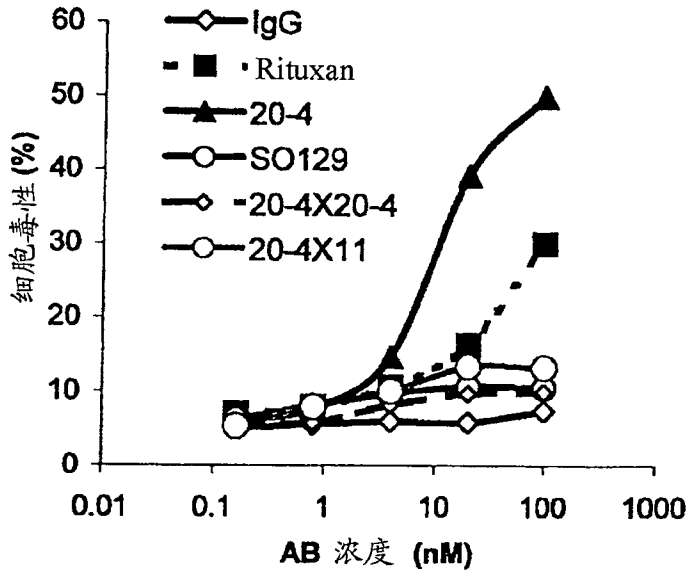

60
50
40
30
20
10
0
细胞毒性 (%)
IgG
Rituxan
20-4
SO129
20-4X20-4
20-4X11
0.01
0.1
1
10
100
1000
AB 浓度 (nM)

图 54

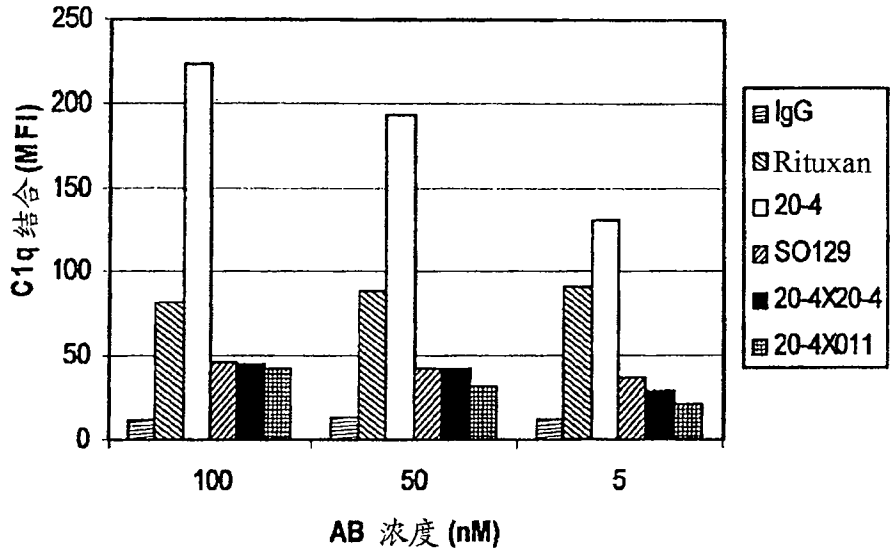

250
200
150
100
50
0
C1q 结合(MFI)
100
50
5
AB 浓度 (nM)
IgG
Rituxan
20-4
SO129
20-4X20-4
20-4X011

图 55

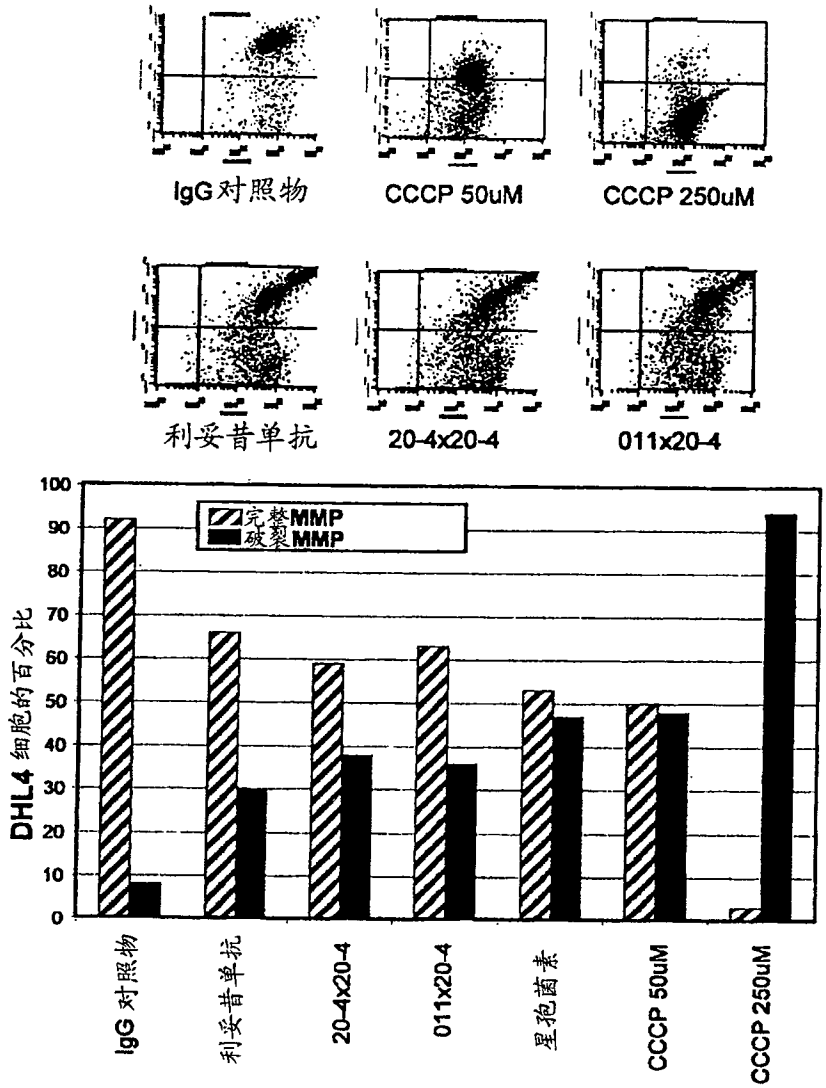

IgG对照物
CCCP 50uM
CCCP 250uM
利妥昔单抗
20-4x20-4
011x20-4
完整MMP
破裂MMP
DHL4 细胞的百分比
100
90
80
70
60
50
40
30
20
10
0
IgG 对照物
利妥昔单抗
20-4x20-4
011x20-4
星孢菌素
CCCP 50uM
CCCP 250uM

图 56

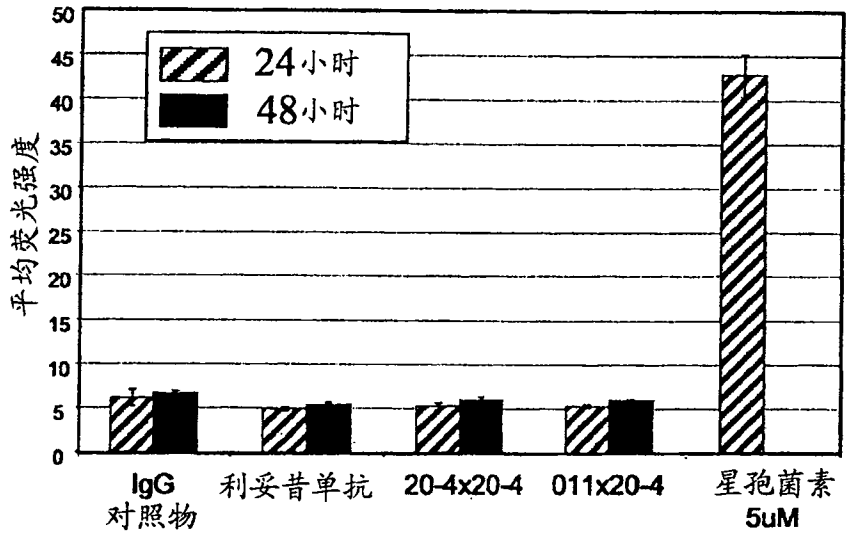

50
45
40
35
30
25
20
15
10
5
0
平均荧光强度
24小时
48小时
IgG
对照物
利妥昔单抗
20-4x20-4
011x20-4
星孢菌素
5uM

图 57

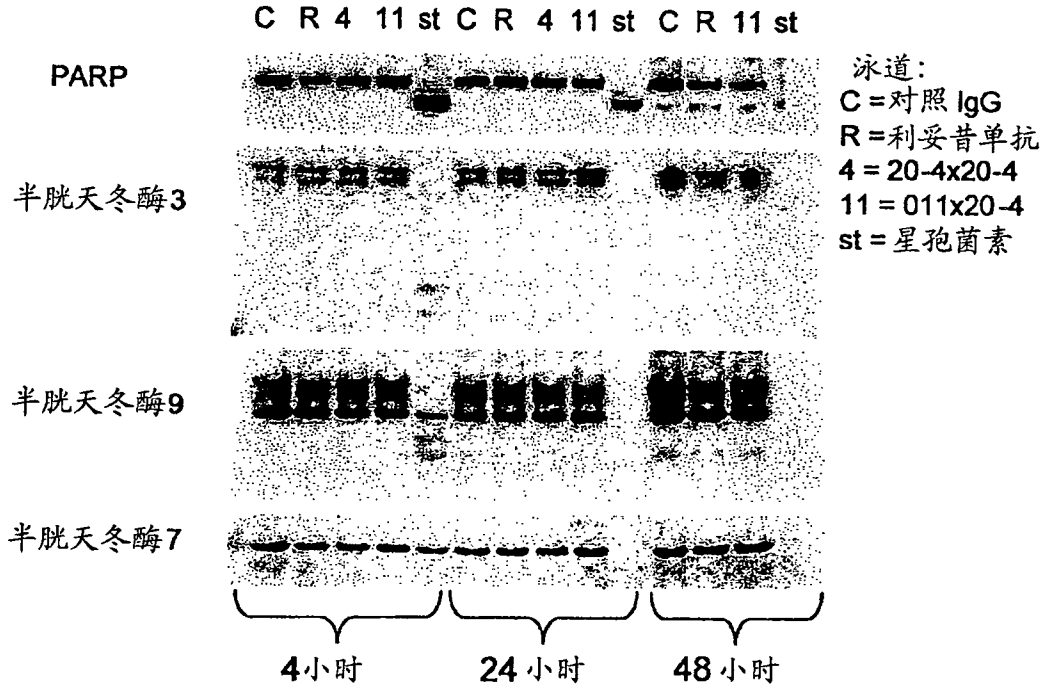

C R 4 11 st C R 4 11 st C R 11 st
PARP
半胱天冬酶3
半胱天冬酶9
半胱天冬酶7
4小时
24小时
48小时
泳道:
C=对照 IgG
R=利妥昔单抗
4=20-4x20-4
11=011x20-4
st=星孢菌素

图 58

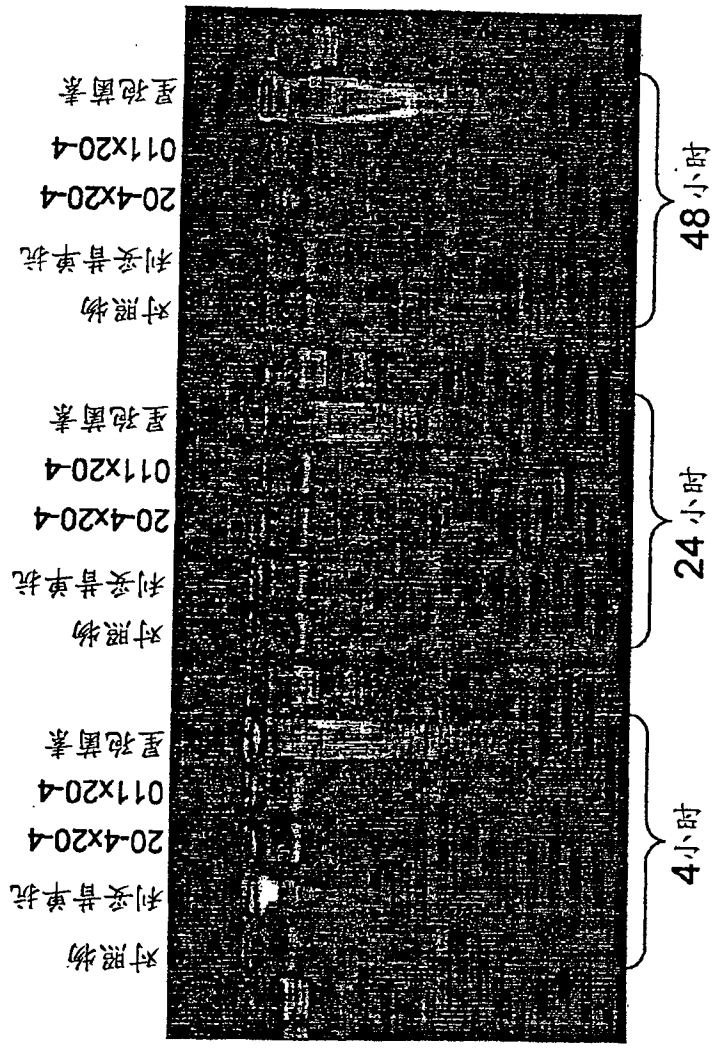

对照物
利妥昔单抗
20-4x20-4
011x20-4
星孢菌素
4小时
对照物
利妥昔单抗
20-4x20-4
011x20-4
星孢菌素
24小时
对照物
利妥昔单抗
20-4x20-4
011x20-4
星孢菌素
48小时

图 59

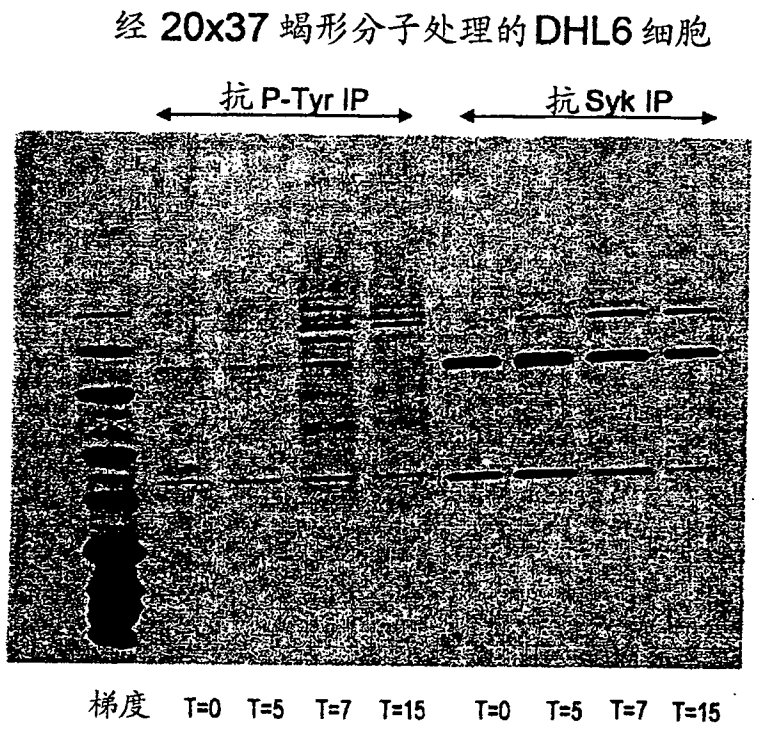

经 20x37 蝎形分子处理的DHL6 细胞
抗 P-Tyr IP
抗 Syk IP
梯度 T=0 T=5 T=7 T=15 T=0 T=5 T=7 T=15

图 60

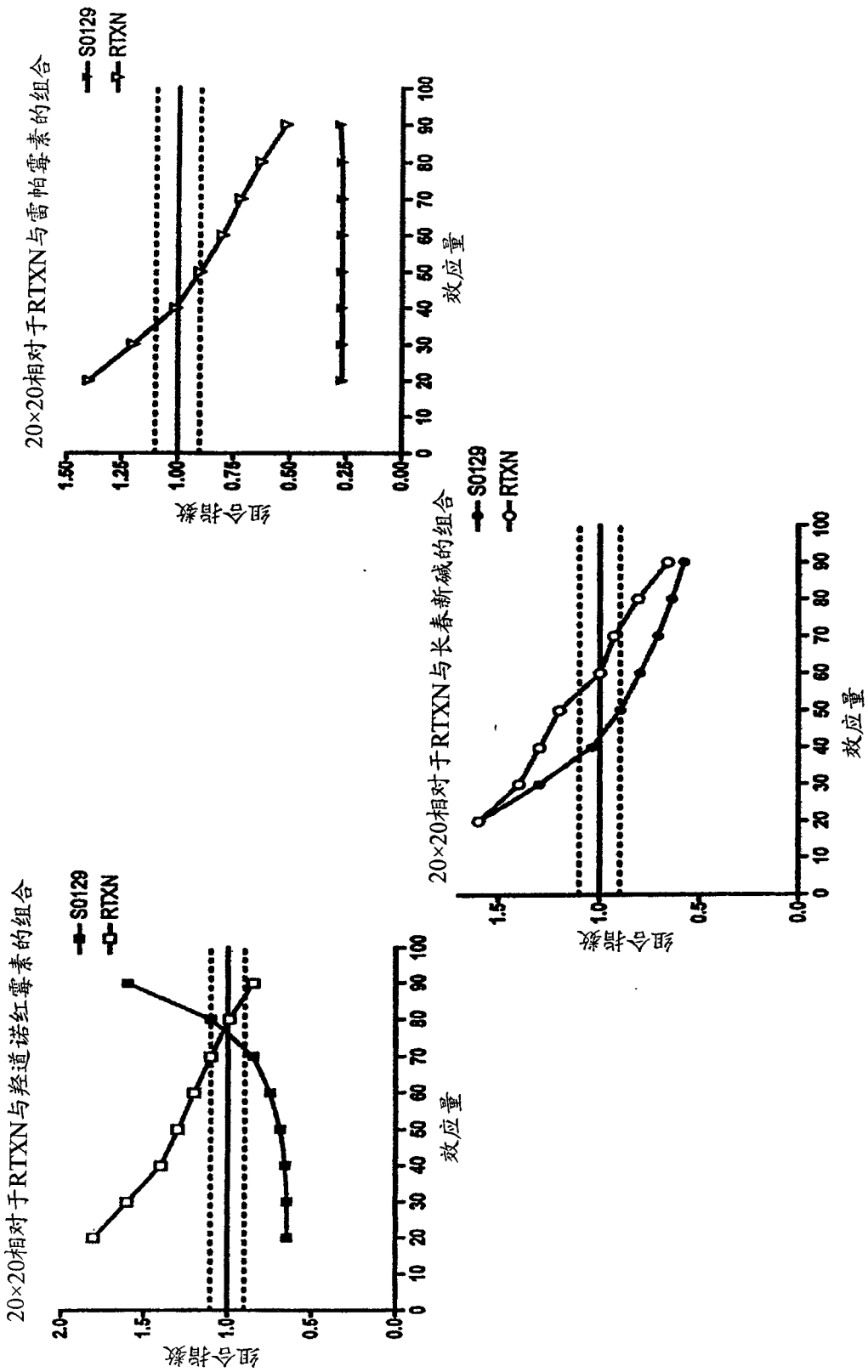

20×20相对于RTXN与雷帕霉素的组合
S0129
RTXN
组合指数
效应量
20×20相对于RTXN与长春新碱的组合
S0129
RTXN
组合指数
效应量
20×20相对于RTXN与羟道诺红霉素的组合
S0129
RTXN
组合指数
效应量

图 61