



(43) International Publication Date
6 January 2005 (06.01.2005)

PCT

(10) International Publication Number
WO 2005/000902 A1

- (51) **International Patent Classification⁷:** C07K 19/00,
16/00, 14/21
- (21) **International Application Number:**
PCT/KR2004/001595
- (22) **International Filing Date:** 30 June 2004 (30.06.2004)
- (25) **Filing Language:** Korean
- (26) **Publication Language:** English
- (30) **Priority Data:**
10-2003-0043599 30 June 2003 (30.06.2003) KR
- (71) **Applicant and**
- (72) **Inventor: CHOE, Mu-Hyeon** [KR/KR]; 15-1205, Dong-sung 3 cha Apt., Sinnae 1-dong, Jungnang-gu, Seoul 131-131 (KR).
- (72) **Inventors; and**
- (75) **Inventors/Applicants** (*for US only*): **CHOI, Seong-Hyeok** [KR/KR]; 769, Sanjeong-ri 3-gu, Sinpung-myeon, Gongju-si, Chungcheongnam-do 314-882 (KR). **LEE, Yong-Chan** [KR/KR]; 138, Eumnae-dong, Seosan-si, Chungcheongnam-do 356-805 (KR). **KWON, Hye-Won** [KR/KR]; 101-1101 Samho Apt., Sangdo 5-dong, Dongjak-gu, Seoul 156-030 (KR). **WON, Jae-Seon** [KR/KR]; 1-164, Bugahyeon 3-dong, Seodaemun-gu, Seoul 120-193 (KR). **YU, Mi-Hyun** [KR/KR]; 101-1001, Mokdong 2-cha Samsungraemian, Sinjeong 7-dong, Yangcheon-gu, Seoul 158-077 (KR). **SONG, Jeong-Hwa** [KR/KR]; A-205, Daewun villa, 171-3.

Bangbae 1-dong, Seocho-gu, Seoul 137-061 (KR). **KIM, Yong-Jae** [KR/KR]; 104-1504, Yangjidaerim Apt., Jung-gye-dong, Nowon-gu, Seoul 139-923 (KR).

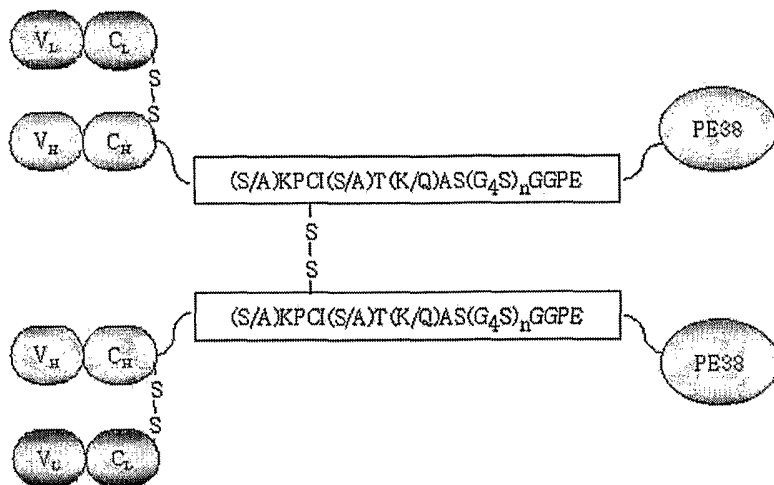
- (74) **Agent:** LEE, Duck-Rog; YEiL Patent & Trademark International, YEiLPAT Bldg., Yorksam-dong, Kangnam-ku, Seoul 135-915 (KR).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, 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, IT, LU, MC, 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

[Continued on next page]

- (54) Title:** THE DIMER OF CHIMERIC RECOMBINANT BINDING DOMAIN-FUNCTIONAL GROUP FUSION FORMED VIA DISULFIDE-BOND-BRIDGE AND THE PROCESSES FOR PRODUCING THE SAME



- (57) Abstract:** The present invention relates to a method producing dimer of chimeric recombinant binding domain-heterogeneous functional group fusion ([B-F fusion]₂) by using covalent disulfide-bond-bridge connecting the two monomers of chimeric recombinant binding domain-heterogeneous functional group fusion (B-F fusion). The dimer of chimeric recombinant binding domain(B)-heterogeneous functional group(F) fusion was the first to be formed by using covalent disulfide-bond-bridge to connect monomers to have double binding valency of the monomer. It has higher functional efficiency to its targets and the production yield is high by containing said extension peptide



— *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.

THE DIMER OF CHIMERIC RECOMBINANT BINDING DOMAIN-
FUNCTIONAL GROUP FUSION FORMED VIA DISULFIDE-BOND-
BRIDGE AND THE PROCESSES FOR PRODUCING THE SAME.

5

Technical Field

This invention is related to the dimer of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) by using disulfide-bond-bridge connecting the two of a monomer of chimeric recombinant binding domain-functional group fusion(B-F fusion) and method for producing the said dimer.

10

Background Art

The dimer of chimeric recombinant binding domain(B)-functional group(F) fusion was the first to be formed by using disulfide-bond-bridge to connect monomers having twice number of binding domain, and it has higher efficiency for targeting its functional group to the targets than the monomer and the production yield is high containing said extension peptide chain(LFA, lath flexible amino acid).

15

20

The fusions of binding domain and heterogeneous functional group have been made of various kinds of binding domains and heterogeneous functional groups.

25

Antibody is a typical use of binding domain[Reference: Hall, Walter A., Immunotoxin Method and Protocols, *Method in Molecular Biology* Vol 166, Humana Press, Totowa, New Jersey]. Antibody has been studied with changing its binding region through recombination and modification maintaining its binding affinity and binding specificity. For examples, there are scFv, pFv, dsFv, Fab, L(using one light chain), LL(using two light chains), H(using one heavy chain), HH(using two heavy chains), diabody, triabody, tetrabody, double headed

antibody and others[Reference: Brinkmann, U., *et al.*, *J. Mol. Biol.* 268, 101~111, 1997, Chaudbary, V.K., *et al.*, *Nature* 339, 394~397, 1989, Webber, K.O., *et al.*, *Mol. Immunol.* 4, 249~258, 1995, Yokota, T., *et al.*, *Cancer Res.* 52, 3402~3408, 1992, Kreitman R.J., *et al.*, *Leukemia* 7(4), 553~562, 1993, Pluckthun A. and Pack P., *Immunotechnology* 3, 83, 1997, Hollinger, P., *et al.*, *Protein Eng.* 9, 299~305, 1996, Atwell J., *et al.*, *Protein Eng.* 12, 597~604, 1999, Iliades P., *et al.*, *FEBS Lett.* 409, 437, 1997]. These fragments of antibody binding regions have been used independently without fusing to functional groups or used as a fusion with functional groups which provokes physiological responses to target cells to deliver the functional group specifically.

Also, examples for use as binding domain are, many kinds of ligands or fragments which have ligand binding affinity, for example, TGF alpha, TGF beta, IL2, IL6, TNF, GMSCF and more. And they include many kinds of ligand receptors or fragments which have receptor binding affinity, for example, TBP1, TBP2, IFN alpha or beta receptor, gonadotropin receptor and other receptors.

There are many functional groups that have been used in fusion with binding domains [Reference: Hall, Walter A., *Immunotoxin Method and Protocols*, *Method in Molecular Biology* Vol 166, Humana Press, Totowa, New Jersey]. For example, enzymes that have functions in prodrug transformation, material detection, decomposition, formation, proteins containing cytotoxic functional group and other functional group, organisms including the viruses for gene therapy, compounds that form cationic tail for delivering DNA, drug compounds, liposomes for drug delivery, biosensors for detecting real time target molecule and many others are used as functional group for fusion[Reference: (Hudson, P.J., *Curr Opin Immunol* 11(5), 548~5, 1999)(Bagshawe, K.D., *et al.*, *Curr Opin Immunol* 11(5), 579~83, 1999)].

Antibody-toxin functional group fusion is a molecule which has cytotoxic

factor connected via chemical or genetic method with antibodies as specific cell binding-domain[Reference: Cobb, P.W., *et al.*, *Semin Hematol* 29, 6~13, 1992]. Antibody-toxin functional group fusion was expected to be successful in cancer treatment after the development of the antibody recognizing cancer cell.

5 Antibody-toxin functional group fusion in its early stages was made by connecting two proteins via protein chemical cross-linking reaction, but in accordance with the development of recombinant DNA technology, it was produced in various forms of recombinant protein through genetic fusion. The incipient of antibody-toxin fusion(mAb-toxin) made by protein chemical cross-
10 linking showed high stability of the fusion protein in blood and it exterminated the cancer cells at clinical demonstration[Reference: Pai, L.H.,*et al.*, *Cancer Res.* 52, 3189~93, 1992] but, the damages on antibody caused during chemical reaction and inactivate molecules produced by chemical side reactions remained as problems.

15 These problems were solved mostly through recombinant DNA technology. Genetic engineering allowed genetically fusing the essential elements for antibody-toxin fusion to make molecules in purity and homogenously, and also it allowed small molecular weighted-proteins to be designed and produced[Reference: Pai, L.H.,*et al.*, *Proc Natl Acad Sci USA* 88, 3358~3362,
20 1991]. For the minimum domain for antibody-toxin functional group fusion, the variable region of the antibody for binding(except constant region) and toxic enzymatic region of toxin (except cell-binding domain of toxin)[Reference: Kondo, T., *et al.*, *J Biol Chem* 263, 9470~9475, 1988] was used. However, nowadays binding domains and toxic domains themselves are modified to be
25 made as derivatives for better activity[Reference: (Pastan, I., *et al.*, *Science* 254, 1173~1177, 1991)(Pastan, I., *et al.*, *Proc Natl Acad Sci USA* 88, 3358~3362, 1991)(Vitetta, E.S., *et al.*, *Cell Biology* 2, 47~58, 1991)(Allured,V.S., *et al.*, *Proc.*

Natl. Acad. Sci. USA 83, 1320~1324, 1986)(Hwang, J et al., *Cell* 48, 129~136, 1987)].

The modified antibody binding region produced by genetic recombination can be classified in 4 types. These are scFv(single chain Fv form) characterized in
5 connecting the minimum binding unit of antibody V_H and V_L with 15 amino acid polypeptide linker(Gly₄Ser₄)[Reference: Buchner, J., et al., *Anal Biochem* 205, 263~70, 1992], dsFv(disulfide-stabilized Fv form) characterized by connecting V_H and V_L via disulfide bonds, pFv(permutated Fv form) characterized in connecting V_H and V_L with base loop and Fab form etc. scFv-toxin functional group form
10 has the smallest molecular weight from antibody binding domain produced and for this it was expected to have good penetration ability into cancer tissues showing good cytotoxicity. However, the low productions yield[Reference: (Buchner, J., et al., *Anal Biochem* 205, 263~70, 1992)(Brinkmann. U., et al., *Proc. Natl. Acad. Sci. USA* 88, 8616~8620, 1991)] and short half life in animal blood
15 circulation[Reference: Brinkmann. U., et al., *Proc. Natl. Acad. Sci. USA* 89, 3065~3069, 1992] were problems and there were no effects observed from results of clinical demonstration.

dsFv-toxin functional group has similar size with scFvtoxin functional group and made to have high stability in animal blood circulation. This type of
20 antibody-toxin functional group was more stable than scFv toxin in blood circulation but the cytotoxicity test with cultured cell *in vitro* showed similar results[Reference: (Pastan, I., et al., *Science* 254, 1173~1177, 1991)(Pastan, I., et al., *Cancer Research* 51, 3781~3787, 1991)]. Results on dsFv distribution tests in animal with radionuclide labeled dsFv-toxin functional group[Reference: Choi,
25 C., et al., *Cancer Res.* 55, 5323~9, 1995] showed that dsFv disappeared from the blood circulation through excretion more fast than to bind with cancer cells and accumulate.

pFv is made by connecting the β -strand between 3 and 3b on V_L and β -strand between 3 and 3b on V_H [Reference: Brinkmann, U., *et al.*, *J Mol Biol* 268, 107~17, 1997]. However, this form of antibody-toxin functional group showed short half-life rate, low production yield, and no improved cytotoxicity effect.

5 Recombinant Fab-toxin functional group fusion was made to overcome the problem mentioned above[Reference (Ghetie, M.A., *et al.*, 1991)(Kreitman, R.J., *et al.*, *Cancer Res* 53, 819~25, 1993)(Choe, M., *et al.*, *Cancer Res* 54,3406~7, 1994)(Kreitman, R.J *et al.*, *Int J Cancer* 57,856~64, 1994)]. This molecule showed similar half life of activity in blood circulation as the incipient
10 antibody-toxin functional group chemical fusion(mAb-toxin) although it was a recombinant antibody-toxin functional group fusion and was more stable than scFv-toxin functional group, dsFv toxin functional group, pFv-toxin functional group in structure[Reference: (Choe, M., *et al.*, *Cancer Res.* 54, 3460~7, 1994)(Kreitman, R.J., *et al.*, *Int J Cancer* 57, 856~64, 1994)]. Also, the
15 production yield of refolding was 10 times higher in maximum[Reference: (Buchner, J., *et al.*, *Bio/Technology* 9, 157~162, 1991)(Buchner, J., *et al.*, *Anal Biochem* 205, 263~70, 1992)]. Even though the Fab-toxin functional group in blood circulation was more active than scFv, the therapeutic efficacy towards cancer cells in animal model was good or had no big difference according to
20 antibody type. The reason it had no big difference was assumed to be that Fab-toxin functional group has a bigger binding domain that each of the quaternary structure of Fab didn't form properly during refolding or it has two more intrachain disulfide bond to make disulfide bond incomplete or make disulfide bond scrambled or the cysteins for interchain disulfide bond between heavy and
25 light chain did not form disulfide bond completely or got scrambled.. Therefore, the complete structure formation of molecule was disturbed to have abnormal binding affinity against antigen and these may have mixed with the normal

molecules to lower the efficacy[Reference: Choe, M., *et al.*, *Cancer Res* 54. 3460~7, 1994].

In this way, the Fab-toxin functional group has weak binding affinity because the isolation and removal of the inactivated molecules and molecules with disulfide bond scrambled or incomplete formed is very difficult. According to these results, the recombinant antibody-toxin functional groups with the best stability in structure and appropriate half-life are Fab-toxin functional group form. To explain this biochemically, C_{H1} and C_L and the disulfide bond between light and heavy chain give high stability to quaternary structure, and the resistance against degradation and clearance will be strong. The most important issue of the antibody-toxin functional group is stability and affinity.

Therefore, development of new antibody-toxin functional group is needed with improved stability and affinity to have high efficacy for targets and high productivity. To increase structural stability, the number of disulfide bond can be increased but if wrong bonds form or scrambling happen molecules with inactive structures are produced and the production yield lowers or doesn't produce it at all.

In addition to the described molecules mentioned above, there are toxin functional group-antibody fusions, which have toxin functional group on the amino terminal and antibody binding domain on the carboxyl terminal and they showed similar results.

Also, antibody derivatives with multiple binding domains were produced to increase binding affinity and they are diabody, triabody, tetrabody, double headed antibody and others[Reference: Takemura S., *et al.*, *Cancer Immunol Immunother.* 51(1): 33~44, 2002]. Increasing the numbers of binding domains two, three, four times and with the affinities between the chains composing binding domains, multiple binding domain derivatives are made. They showed that the binding affinity increased by the increase of binding domain numbers and there

was no report before about these molecules being fused with heterogeneous functional groups to produce chimeric fusion. It is a difficult matter to connect binding domain to functional group without hindrance to each other and between the functional group themselves in forming dimeric form of the chimeric fusion
5 and manufacturers concerned will know that this is the key to successful production. If the extension chain, which connects binding domain and functional group, disturbs the refolding of the big chains of binding domain or functional group or some hindrance occurs between the big chains, no production will occur.

Also, when the Cysteine lacking intrachain disulfide bond counterpart
10 cysteine is added in the extension sequence connecting binding domain and functional group this may form disulfide bond with other Cysteine, which has its right counterpart, and this may lead to disulfide bond structure scrambling. In this case the uncoupled Cysteine that has no natural disulfide bond counterpart will form wrong disulfide bonds to ruin sterical structure of the molecule and lose its
15 activity. And also, it is natural to think that if the extension chain containing uncoupled Cysteine has lots of flexible amino acids to be long making the extension chain have no regular structure, it is easier for the uncoupled Cysteine to intermix with other naturally coupled Cysteine in neighboring big structure. The inventors experienced these kinds of failures in producing active dimer
20 molecules and it will be thought as a special and incidental case when the manufacturers concerned produce one of these types of molecule.

If the problems and limits mentioned above are overcome by new findings about forming dimers with disulfide bond generally, the dimer of binding domain(B)- functional group(F) fusion by disulfide bond will be possible by the
25 manufacturers concerned.

On the other side, there were reports about compared studies of various structures of molecules[Reference: Bera, T.K., *et al.*, *Bioconjug Chem* 9(6), 736~4,

1998] but no reports were found about producing dimer[antibody-toxin functional group]₂ by using disulfide bonding between antibody-toxin functional groups fusion to form molecules with double binding valency. Therefore, it has been expected that the manufacturers will be able to know how extension peptide chain to be put in between big sequences of antibody and toxin functional group and how the uncoupled Cysteine to be put in to extension peptide chain to form disulfide bonds for dimerization after the understanding about the refolding process forming the tertiary and quaternary structure of the protein.

10 **Brief description of the drawings**

Figure 1 is a structure of the dimer of chimeric recombinant binding domain(B)-functional group(F) fusion with double binding valency produced from pMH21, 22, 23 and pMHS22.

Figure 2 shows the construction procedure of pMH21, 22, 23, pMHS22
15 by PCR.

Figure 3 shows the plasmid map of pMH21, 22, 23, pMHS22.

Figure 4 is a graph showing the production yield according to the position of Cysteine at position 1, 4 and 15 on extension peptide chain.

Figure 5 shows antibody-toxin finally purified.

Figure 6 shows the structure of [B3(Fab)-Ext(15CL14FA13)-PE38]₂ produced from pCW1.

Figure 7 shows the result profile of Superdex 200 column chromatography and SDS-PAGE analysis.

Figure 8 is a graph showing representative cytotoxicity assay results of
25 [B3 (Fab)-Ext-PE38]₂ and control molecule scFv-PE40.

Figure 9 shows the structure of [B3(FabH1)-PE38]₂ (= [B3(Fab)-Ext(1CL13FA7)-PE38]₂)

Figure 10 shows the TSK-GEL G3000SW column chromatography result of B3 (FabH1)-PE38 and [B3 (FabH1)-PE38]₂.

Figure 11 shows the structure of BMH and BM[PEO]₄.

Figure 12 shows the structure of antibody-toxin fusion produced from
5 pLSC52, 32, 22.

Figure 13 shows the construction procedure of pLSC52 by PCR.

Figure 14 shows the construction procedure of pLSC32 by PCR.

Figure 15 shows the profile of Source-Q column chromatography and SDS-PAGE results of [B3(Fab)-h(H123-CH3)-PE38R]₂.

10 Figure 16 shows the profile of Superdex 200 column chromatography and SDS-PAGE results of [B3(Fab)-h(H124-Fc)-PE38R]₂.

Figure 17 shows the profile of Superdex 200 column chromatography and SDS-PAGE results of [B3(Fab)-h(H124-CH3)-PE38R]₂.

Figure 18 shows the profile of Superdex 200 column chromatography and
15 SDS-PAGE results of [B3(Fab)-h(H124-CH2)-PE38R]₂.

Disclosure

Technical Problem

The technical task of this invention is fusing binding domain(B) with
20 heterogeneous functional group(F) to form a monomer of chimeric recombinant binding domain-functional group fusion(B-F fusion) and connecting two of these with covalent disulfide-bond-bridge to make a dimer of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double binding valency of the monomer.

25

Technical Solution

This invention is related to a method producing dimer of chimeric

recombinant binding domain-heterogeneous functional group fusion([B-F fusion]₂) with double binding valency of the monomer by using covalent disulfide-bond-bridge connecting the two of a monomer of chimeric recombinant binding domain- heterogeneous functional group fusion(B-F fusion).

5 In a concrete way, two monomer is connected to become a dimer by disulfide-bond-bridge formed by the oxidation reaction between the two uncoupled Cys that is on any of 1~45 amino acid position of extension peptide chain (Ext) which extends from binding domain to functional group for fusion.

At this time, said extension peptide chain(Ext, extension amino acid
10 sequence) is firstly composed of peptide linker(L) from the last of the uncoupled Cys to functional group(end of Ext) and composed of 1~50 amino acids from the last of the uncoupled Cys to functional group. Secondly, said extension peptide chain(Ext, extension amino acid sequence) is composed of peptide linker(L) from the last of the uncoupled Cys to functional group(end of Ext), and said peptide
15 linker(L) is a peptide linker containing an affinity domain(LAD) which has homomeric self affinity or heteromeric affinity making the domain to assemble and leading to the formation of homomeric multimer or heteromeric multimer, and the amino acid sequence from the end of said affinity domain(AD) to functional group(end of Ext) is composed of 1~50 amino acids. Thirdly, said extension
20 peptide chain(Ext, extension amino acid sequence) is composed of peptide linker(L) from the last of the uncoupled Cys to functional group(end of Ext), and said peptide linker(L) is a flexible amino acid sequence peptide linker(LFA) which contains non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D), and the
25 amino acid sequence from the last of the uncoupled Cys to functional group is composed of 1~50 amino acids. Fourthly, said extension peptide chain(Ext, extension amino acid sequence) is composed of peptide linker(L) from the last of

the uncoupled Cys to functional group (end of Ext), and said peptide linker(L) is a peptide linker containing an affinity domain(LAD) which has homomeric self affinity or heteromeric affinity making the domain to assemble and leading to the formation of homomeric multimer or heteromeric multimer, and said peptide
5 linker(L) is also a flexible amino acid sequence peptide linker(LFA) which contains non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D), and the amino acid sequence from the end of said affinity domain(AD) to functional group(end of Ext) is composed of 1~50 amino acids.

10 On the other side, the binding domain on the chimeric recombinant binding domain-heterogeneous functional group fusion is composed of multiple chains(B1,B2,...,Bn) and one of the chain(B1) of binding domain is connected to extension peptide chain that has the cysteine for the disulfide-bond-bridge between two monomer and another chain(B2) is connected to heterogeneous
15 functional group.

At this time, said extension peptide chain(Ext) which is connected to one of the chain(B1) comprises firstly, uncoupled Cys on any of 1~45 amino acids position on the extension peptide chain. Secondly said extension peptide chain comprises uncoupled Cys on any of 1~45 amino acids position and comprises
20 peptide linker(L) from the last uncoupled Cys to the end of extension peptide chain. Thirdly, said extension peptide chain comprises uncoupled Cys on any of 1~45 amino acids position and comprises peptide linker(L) from the last uncoupled Cys to the end of extension peptide chain and said peptide linker(L) has homomeric self affinity or heteromeric affinity domain(AD) making the
25 domain to assemble and leading to the formation of homomeric multimer or heteromeric multimer. Fourthly, said extension peptide chain comprises uncoupled Cys on any of 1~45 amino acids position and comprises peptide linker(L) from

the last uncoupled Cys to the end of extension peptide chain and said peptide linker(L) has homomeric self affinity or heteromeric affinity domain(AD) making the domain to assemble and leading to the formation of homomeric multimer or heteromeric multimer and said peptide linker(L) comprises flexible amino acid
5 sequence with non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D).

At this time, another chain(B2) firstly, has functional group(F) at the end of the chain. Secondly, has functional group(F) at the end of the peptide linker(L) connected to the end of the chain and said peptide linker comprises 1~50 amino
10 acids. Thirdly, has functional group(F) at the end of the peptide linker(L) connected to the end of the chain and said peptide linker(L) is flexible amino acid sequence peptide linker(LFA) which comprises Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) and said peptide linker(L) comprises 1~50 amino acids.

15 Binding domain(B) is adhesive protein or amino acid sequence having adhesive fragment for instance, antibody, fragment of antibody like scFv, pFv, dsFv, Fab, L(using 1 light chain), LL(using 2 light chains), H(using 1 heavy chain), HH(using 2 heavy chains), diabody, triabody, tetrabody, double-headed antibody, ligands for example, TGF alpha, TGF beta, IL2, IL6, TNF, GMSCF or
20 some fragments having ligand's affinity, all kinds of ligand receptors for example, insulin receptor, TBP1, TBP2, IFN alpha or beta receptor, gonadotropin receptor or some fragments having receptor's affinity and sequences having binding affinity.

Functional group(F) is a functional group with all kinds of physiological
25 functions including enzymes used in prodrug transformation, detection, decomposition, formation of materials and proteins containing toxin-functional group which has cytotoxicity, organisms like viruses for gene therapy, compounds

with cationic tail for delivering DNA, drug compounds, liposome for drug delivery, biosensor for detecting real time target molecule.

The dimer of chimeric recombinant binding domain(B)-heterogeneous functional group(F) fusion was the first to be formed by using disulfide-bond-
5 bridge to connect monomers having twice as high binding valency, and it has higher functional efficiency to its targets and the production yield is high containing said extension peptide chain.

Namely, from one point of view, this invention offers dimer of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double
10 binding valency by using disulfide-bond-bridge connecting the two of a monomer of chimeric recombinant binding domain-functional group fusion([B-F fusion]).

In another point of view, this invention offers polypeptides used in producing dimer of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double binding valency by using disulfide-bond-bridge
15 connecting the two of a monomer of chimeric recombinant binding domain-functional group fusion([B-F fusion]).

In another point of view, this invention offers recombinant plasmid containing genes coding polypeptides used in producing dimer of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double
20 binding valency by using disulfide-bond-bridge connecting the two of a monomer of chimeric recombinant binding domain-functional group fusion([B-F fusion]).

Something to make clear in this invention is concluding not only the amino acid sequence of binding domain(B)-functional group(F) protein but concluding all the nucleotide sequences coding the same protein. Actually, all the
25 nucleotide sequences means all the different DNA sequences which codes the same amino acid using different codons to code same amino acids.

PCR is mostly used to produce chimeric recombinant binding domain(B)-

functional domain(F) fusion coding sequences by using oligonucleotide primers made from clones which has genes coding the protein we want.

The invention chimeric recombinant binding domain(B)-functional domain(F) fusion can be expressed in high class eukaryotic cell(for example, yeast, insects or mammalian cells) or prokaryotic cell of microorganism using
5 appropriate expression vector. Any method known in this field can be used.

For example, chimeric recombinant binding domain(B)-functional domain(F) fusion protein coding DNA produced from said any method referred above can be inserted into appropriate expression vector. Double strand cDNA is
10 inserted into the vector by using synthesized DNA linker or blunt-ended DNA ligation or restriction enzyme treated DNA ligation or terminal polymerase treated DNA ligation. DNA ligase is used in DNA molecules ligation and alkaline phosphatase is used in removing the phosphate group.

Expression vector must have specific nucleotides elements for mRNA
15 transcription and translation into protein from the coding DNA sequence. Promoter is needed for the RNA polymerase to recognize for transcription. RNA polymerase binds to this promoter to start transcription. Promoters used are various on general but has different efficiency.

In another point of view, this invention offers host cells including
20 recombinant vectors cloned with genes coding polypeptides used in producing dimer of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) by using disulfide-bond-bridge connecting the two of a monomer of chimeric recombinant binding domain-functional group fusion([B-F fusion]).

According to eukaryotic hosts, it varies with control sequences for
25 transcription and translation. These are regulatory signals related to highly expressed genes and can be gained from viruses like adenovirus, cow papillomavirus, anthropoid virus. For example, there are TK promoter from

Herpes virus, SV40 early promoter, gal4 gene promoter from yeast and more. There are transcription start control signals that can both inhibit or promote the expression of the gene to be selected for use.

Transformation, transfection, conjugation, protoplast fusion,
5 electroporation, calcium phosphate-precipitation, direct microinjection and others can insert the expression vector into host cell.

Host cells are prokaryotic or eukaryotic cells. Eukaryotic cells are preferable for example, mammalian cells like human, monkey, mouse, Chinese hamster ovarian cell(CHO) because these cells offer correct folding or
10 modification of proteins like glycosylation. Also, in yeast cells modification including glycosylation takes place after protein expression. Yeast cells provide high copy numbers of recombinant vectors in the cell and transcription efficiency is high. The yeast recognizes the guidance sequence of protein secretion signal of the cloned mammalian gene and secretes the peptide(precursor protein) which has
15 the signal.

In another point of view, this invention offers method to produce chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) through host cells including recombinant vectors cloned with genes coding polypeptides used in producing dimer of chimeric recombinant binding domain-functional group
20 fusion([B-F fusion]₂) by using disulfide-bond-bridge connecting the two of a monomer of chimeric recombinant binding domain-functional group fusion([B-F fusion]).

In another point of view, this invention offers pharmaceutical compounds containing chimeric recombinant binding domain-functional group fusion([B-F
25 fusion]₂).

The dimeric form of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double binding valency is indicated as a formula

below.

[Binding domain(B)-Ext-F]₂(I)

The binding sequence(B) in said formula(I) above is adhesive protein or adhesive fragment containing amino acid sequence ;

5 Ext is extension peptide chain(extension sequence) extended from binding domain(B) to functional group(F) to fuse them, and disulfide-bond-bridge is formed between the monomers by the oxidation of two uncoupled Cys on the extension peptide chain(Ext) to produce dimer, or said extension peptide chain comprises peptide linker(L) from the last uncoupled Cys to the end of extension
10 peptide chain or said peptide linker is an affinity domain(AD) containing peptide linker(LAD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain, or said peptide linker(L) is flexible amino acid(FA) sequence peptide linker(LFA) which has non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or
15 Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) or said peptide linker having affinity domain(AD) with homomeric self affinity or heteromeric affinity comprising multidimer domains and having non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) to be a flexible amino acid sequence
20 peptide linker with affinity domain(LADFA) so to be a sequence as [B-Ext or Ext(L) or Ext(LAD) or Ext(LFA) or Ext(LADFA)-F] in its monomer form ;

F is for physiological functional group like enzyme, protein with physiological (e.g. toxic) function, organisms like viruses, compounds, drug compounds for treatment, liposome, biosensor, pro-drug and more[Reference:
25 (Farah, R.A., et al., *Crit. Rev. Eukaryot. Gene Expr.* 8, 321~356, 1998) (Trail, P.A., et al., *Science* 261, 212~215, 1993)(Hinman, L.M., et al., *Cancer res.* 53, 3336~3342, 1993)(Pastan, I. *Biochem. Biophys. Acta* 1333, C1~C6,

1997)(Kreitman, P.J., *et al.*, *J. Clin. Oncol.* 18, 1622~1636, 2000)(Zalutsky, M.R & Vaidyanathan, G. *Curr.Pharm.Des.* 6, 1433~1455, 2000)(Goldenburg, D.M. in *Clinical Uses of Antibodies* (eds Baum, R.P., *et al.*) 1~13 (Kluwer academic, The Netherlands, 1991)(Lode, H.N. & Reisfield, R.A. *Immunol.Res.* 21, 279~288, 2000)
 5 (Penichet, M.L. & Morrison, S.L. *J.Immunol.Methods* 248, 91~101, 2001)(Lasic, D.D. & Papahadjopoulos, D. *Science* 267, 1275~1276, 1995) (Park, J.W., *et al.*, *Proc. Natl Acad. Sci. USA* 92, 1327~1331)(Niculescu-Davaz, I., *et al.*, *Anticancer Drug Des.* 14, 517~538, 1999)(SToldt, H.S., *et al.*, *Eur. J. Cancer* 33, 186~192, 1997))]

10 In another mode, the dimeric form of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double binding valency is indicated as a formula below.

[Antibody binding domain-Ext- F]₂....(II)

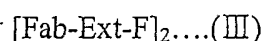
The antibody binding domain in formula(II) is antibody or amino acid
 15 sequence comprising adhesive fragment of antibody.

Ext is extension peptide chain(extension sequence) extended from binding domain(B) to functional group(F) to fuse them, and disulfide-bond-bridge is formed between the monomers by the oxidation of two uncoupled Cys on the extension peptide chain(Ext) to produce dimer, or said extension peptide chain
 20 comprises peptide linker(L) from the last uncoupled Cys to the end of extension peptide chain or said peptide linker is an affinity domain(AD) containing peptide linker(LAD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain, or said peptide linker(L) is flexible amino acid(FA) sequence peptide linker(LFA) which
 25 has non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) or said peptide linker having affinity domain(AD) with homomeric self affinity or

heteromeric affinity comprising multimer domains and having non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) to be a flexible amino acid sequence peptide linker with affinity domain(LADFA) so to be a sequence as [B-Ext or
5 Ext(L) or Ext(LAD) or Ext(LFA) or Ext(LADFA)-F] in its monomer form ;

F is for physiological functional group like enzyme, protein with physiological (e.g. toxic) function, organisms like viruses, compounds, drug compounds for treatment, liposome, biosensor, pro-drug and more.

In another mode, the dimeric form of chimeric recombinant binding
10 domain-functional group fusion([B-F fusion]₂) with double binding valency is indicated as a formula below



Fab in the said formula(III) is Fab fragment of antibody.

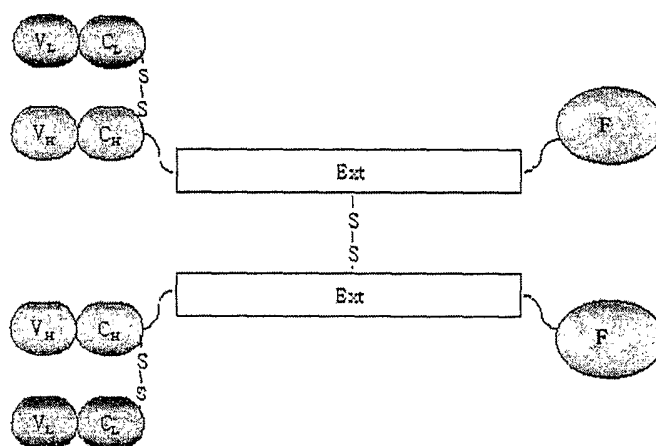
Ext is extension peptide chain (extension sequence) extended from
15 binding domain(B) to functional group(F) to fuse them, and disulfide-bond-bridge is formed between the monomers by the oxidation of two uncoupled Cys on the extension peptide chain(Ext) to produce dimer, or said extension peptide chain comprises peptide linker(L) from the last uncoupled Cys to the end of extension peptide chain or said peptide linker is an affinity domain(AD) containing peptide
20 linker(LAD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain, or said peptide linker(L) is flexible amino acid (FA) sequence peptide linker(LFA) which has non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) or said
25 peptide linker having affinity domain(AD) with homomeric self affinity or heteromeric affinity comprising multimer domains and having non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic

acid(E) or Asparagine(N), Aspartic acid(D) to be a flexible amino acid sequence peptide linker with affinity domain(LADFA) so to be a sequence as [B-Ext or Ext(L) or Ext(LAD) or Ext(LFA) or Ext(LADFA)-F] in its monomer form ;

F is for physiological functional group like enzyme, protein with
5 physiological(e.g. toxic) function, organisms like viruses, compounds, drug compounds for treatment, liposome, biosensor, pro-drug and more.

In another mode, the dimeric form of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double binding valency is indicated as a formula below

10 [Reference figure 1]



According to the formula above,

V_L is variable region on light chain of monoclonal antibody;

C_L is constant region on light chain of monoclonal antibody;

15 V_H is variable region on heavy chain of monoclonal antibody;

C_H is constant region on heavy chain of monoclonal antibody;

Ext is extension peptide chain(extension sequence) extended from binding domain(B) to functional group(F) to fuse them, and disulfide-bond-bridge is formed between the monomers by the oxidation of two uncoupled Cys on the
20 extension peptide chain(Ext) to produce dimer, or said extension peptide chain

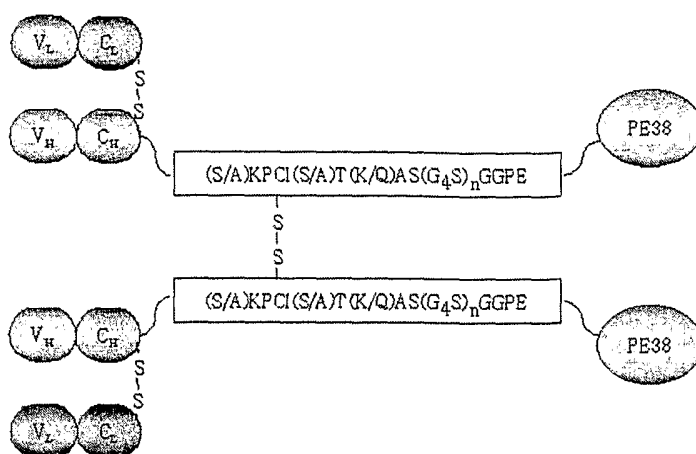
comprises peptide linker(L) from the last uncoupled Cys to the end of extension peptide chain or said peptide linker is an affinity domain(AD) containing peptide linker(LAD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain, or said
5 peptide linker(L) is flexible amino acid (FA) sequence peptide linker(LFA) which has non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) or said peptide linker having affinity domain (AD) with homomeric self affinity or heteromeric affinity comprising multimer domains and having non-bulky amino
10 acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) to be a flexible amino acid sequence peptide linker with affinity domain(LADFA) so to be a sequence as [B-Ext or Ext(L) or Ext(LAD) or Ext(LFA) or Ext(LADFA)-F] in its monomer form ;

F is for physiological functional group like enzyme, protein with
15 physiological(e.g. toxic) function, organisms like viruses, compounds, drug compounds for treatment, liposome, biosensor, pro-drug and more.

More specifically, the dimeric form of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) with double binding valency is indicated as a formula below

20 [Reference figure 2]

21



According to the formula above,

V_L is variable region on light chain of monoclonal antibody;

C_L is constant region on light chain of monoclonal antibody;

5 V_H is variable region on heavy chain of monoclonal antibody;

C_H is constant region on heavy chain of monoclonal antibody;

A is Alanine;

C is Cysteine;

E is Glutamic acid;

10 G is Glycine;

I is Isoleucine;

K is Lysine;

P is Proline;

Q is Glutamine;

15 S is Serine;

T is Threonine;

n is 1 or 2 or 3;

20 PE38 is 38kDa exotoxin derivative of *Pseudomonas aeruginosa* which has been truncated of amino acid sequences of domain I and II that are not required for cytotoxicity.

The inventors used the previous recombinant binding domain(B)-functional domain(F) fusion model to invent [B3(Fab)-(S/A)KPCI(S/A)T(K/Q)AS(G4S)_nGGPE – toxin-functional group]₂ having twice the adhesion valency and using monoclonal antibody B3 as a model binding domain. Antibody-toxin functional group fusion made in the meantime was monovalent but the invented dimeric antibody-toxin functional group fusion is divalent so it will display much more cytotoxicity towards cultured cancer cells. This molecule is produced by modifying Fd-CKPSISTKASGGPE-toxin functional group chain among Fd-CKPSISTKASGGPE-toxin functional group chain and L chain which are used in composing monovalent Fab-CKPSISTKASGGPE-toxin functional group formed monomer. The modified chain was made by changing the position of uncoupled Cysteine with Serine coming after Cysteine. (S/A) means Serine or Alanine and (K/Q) means Lysine or Glutamine. (G4S)_n peptide chain is (GGGGS) or (GGGGSGGGGS) or (GGGGSGGGSGGGGS) and it's in the middle of the KASGGPE to have Fd-(S/A)KPCI(S/A)T(K/Q)AS(G4S)_nGGPE-PE38 structure and mixing and refolding this with L chain makes an activated molecule. The divalent dimer [Fab-(S/A)KPCI(S/A)T(K/Q)AS(G4S)_nGGPE – PE38]₂ is formed by disulfide-bond-bridge connecting two Fab-(S/A)KPCI(S/A)T(K/Q)AS(G4S)_nGGPE – PE38 by the fourth amino acid , uncoupled Cysteine. This molecule keeps the stability of the Fab-toxin functional group structure and targeting efficiency too.

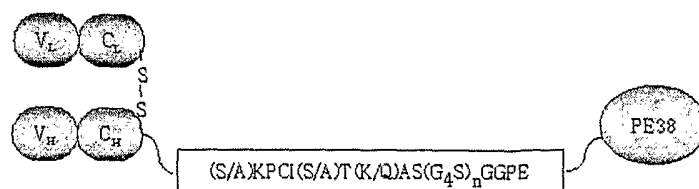
The invention dimer of chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) by using disulfide-bond-bridge connecting the two of a monomer of chimeric recombinant binding domain-functional group fusion(B-F fusion) has extension peptide chain(extension amino acid sequence, Ext) extended from binding domain(B) to functional group(F) to fuse them, and disulfide-bond-bridge is formed between the monomers by two uncoupled

oxidized Cys on the extension peptide chain(Ext) to produce dimer, or said extension peptide chain comprises peptide linker(L) from the last uncoupled Cys to the end of extension peptide chain or said peptide linker is an affinity domain containing peptide linker(LAD) which has homomeric self affinity or heteromeric affinity comprising homomeric multimer or heteromeric multimer domain, affinity domain(AD) or said peptide linker(L) is flexible amino acid (FA) sequence peptide linker(LFA) which has non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) or said peptide linker having affinity domain (AD) with homomeric self affinity or heteromeric affinity comprising multimer domains and having non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) to be a flexible amino acid sequence peptide linker with affinity domain (LADFA) so to be a sequence as [B-Ext or Ext(L) or Ext(LAD) or Ext(LFA) or Ext(LADFA)-F] in its monomer form. And the dimer has decreased three-dimensional hindrance between the two functional groups while being produced and they are produced through mixing and refolding the polypeptide obtained from the host cells containing recombinant plasmid with the Ext sequences for dimerization.

20 In the following, the double binding valency dimer of chimeric recombinant binding domain(B)-functional domain(F) fusion will be described in details and through this description the manufacturers concerned will be able to produce any related dimer of chimeric recombinant binding domain(B)-functional domain(F) fusion.

25 In another mode, this invention offers polypeptides for producing dimeric chimeric recombinant binding domain-functional group fusion([B-F fusion]₂) and it is indicated as a formula below.

[Reference figure 3]



In another mode, this invention offers recombinant plasmid that expresses said polypeptide.

5 In another mode, this invention offers methods to produce said divalent recombinant antibody-toxin functional group fusion by culturing host cells with recombinant plasmid expressing said polypeptide gene and other host cells with plasmid expressing light chain which includes V_L and C_L and joining them together and refolding them.

10 The divalent recombinant antibody-toxin functional group fusion specifically binding its antibody to cancer cell and connecting the toxin functional group is anti-cancer therapeutic agent and it kills the cancer cells without damaging other normal cells.

B3(Fab)-PE38 is produced from the fusion of Fab of monoclonal
15 antibody B3 and PE38 which is truncated type of *Pseudomonas aeruginosa* exotoxin and it doesn't have uncoupled Cysteine which means it can only be a monomer. The [B3(Fab)-Ext-PE38]₂ which belongs to divalent chimeric recombinant binding domain(B)-functional domain(F) fusion is derived from B3(Fab)-PE38 and the Fab of antibody is connected to functional group with
20 (S/A)KPCI(S/A)T(K/Q)AS(G₄S)_nGGPE.

The monoclonal antibody B3 used in this invention binds directly to LeY type carbohydrate antigen found not only in mucinous cancer like colon carcinoma, stomach cancer, ovarian cancer, breast cancer and lung cancer but also various epidermoid carcinomas.

Pseudomonas aeruginosa exotoxin PE –derived PE38 was used as toxin functional group in this invention. PE is composed of three structural domains. Domain 1 at the amino terminal binds with the cell, domain 2 enables the transport of the protein into the cell, domain 3 is at the carboxy terminal having cytotoxic enzymatic activity. Truncating needless amino acids not used for cytotoxic enzymatic activity from domain 1, 2 produced the 38kDa PE38.

The B3 monoclonal antibody is dimer form in nature as [Fab-Fc]₂ because of the three Cysteine in the hinge. The inventors used 1 uncoupled Cysteine, or 3 uncoupled Cysteines although 3 uncoupled Cysteines have much higher risk of mixing disulfide bonding while refolding and showed that using multiple numbers of uncoupled Cysteine can also produce physiologically active dimers. Manufacturers concerned know the dimer produced with multiple disulfide bonds using multiple uncoupled Cysteins will be highly thermodynamically stable.

The extension peptide chain(Ext) (S/A)KPCI(S/A)T(K/Q)AS(G4S)nGGPE has a flexible peptide linker(LFA) I(S/A)T(K/Q) AS(G4S)nGGPE including GASQEND. Also the thiol group of the uncoupled Cysteine on Ext enables disulfide bonding to form divalent chimeric recombinant antibody-toxin functional group fusion[B3(Fab)-Ext-PE38]₂ from two polypeptide[B3(Fab)-Ext-PE38].

The peptide linker following the uncoupled Cysteine in the extension peptide chain reduces the three-dimensional hindrance between the two toxin functional group PE38 helping dimerization. Also under same condition, when the number n increases from 1 to 3, the production yield increased. This means while the dimer has 50kDa Fab sequence and 38kDa PE38, until the number of GASQEND which is in flexible peptide linker(LFA) increases to 21, the dimerization will increase.

The inventors realized that dimer is not only formed by specific numbers and location of the uncoupled Cysteine and specific linker(L) amino acid sequence but, multiple numbers and certain range of locations of Cys, certain range of amino acid sequence of linker(L) can make dimers too. The manufacturers concerned can easily predict that if the size of the binding domain(B) differs and the functional domain(F) also differs, the trend will differ too. So to speak, in case of small sized binding domain and functional domain fusion, dimerization will occur even under peptide linker with only 1 amino acid and if there are too many GASQEND of the flexible peptide chain(LFA) in the extension peptide chain, it will rather disturb refolding of the binding domain and functional domain because of the easy mutual interference and result in decrease of the active molecule yield. Therefore, it is important to find out the range of the position of the uncoupled Cys and the range of the length of the amino acid sequence(the length of L,LFA or the length between AD and F in LAD,LADFA) needed in the peptide linker of extension peptide chain that makes dimerization possible without disulfide bond scrambling. Naturally, the big sized binding domain(B) and functional domain(F) fusion will need more than 1 flexible amino acid of GASQEND for dimerization and longer length can be predicted easily among manufacturers concerned.

The divalent chimeric recombinant antibody-toxin functional group fusion dimer[B3(Fab)-Ext-PE38]₂ is produced by applying fusion of the antibody gene and toxin functional group gene together and expressing it in *E.coli*.(Reference:Allured *et al.*,1986; Brinkmann *et al.*, 1992; Pai *et al.*,1991: Roscoe *et al.*,1997)

In detail, recombinant vector expressing polypeptide B3(Fd)-Ext-PE38 is produced in this invention. The polypeptide B3(Fd)-Ext-PE38 has a polypeptide chain (S/A)KPCI(S/A)T(K/Q)AS(G4S)nGGPE between Fd and PE38. Therefore,

in this case, flexible amino acid peptide linker(LFA) of extension peptide chain(Ext) is "I(S/A)T(K/Q)AS(G4S)_nGGPE" and the constitution of B3(Fd)-Ext-PE38 is same as B3(VH)-B3(CH1)-(S/A)KPCI(S/A)T(K/Q)AS(G4S)_nGGPE-PE38. The uncoupled Cysteine on extension peptide chain forms disulfide bond-
5 bridge between two Fab-Ext-PE38. The decrease of three dimensional hindrance between two PE38 enables dimerization of B3(Fab)-Ext-PE38 easily by flexible amino acid following uncoupled Cysteine on the extension chain.

The extension peptide chain is for amicable, independent refolding without disturbances between antibody and PE toxin functional group in
10 producing recombinant protein antibody-toxin functional group fusion.

The Proline(P) on the end of extension peptide chain (S/A)KPCI(S/A)T(K/Q)AS(G4S)_nGGPE cuts off the continuation of refolding motifs to prevent each refolding motif from interfering neighbor refolding domain during the process of refolding of antibody and toxin functional group.

15 The divalent recombinant antibody-toxin functional group fusion is produced by culturing the host cells having recombinant vectors cloned with B3(Fd)-Ext-PE38 and light chain of B3 respectively to gain polypeptides of these and then mixing these to refold. The cloned gene is expressed and gained as inclusion body in *E. coli*. Polypeptide is refolded and the active products are
20 separated.

The method for refolding can follow well-known procedures and some references describe the methods[Reference: Buchner,J.,*et al.*, *Anal .Biochem.* 205(2), 263~70 (1992), Buchner, J., *et al.*, *Biotechnonlogy*, 9:157~162 (1991)]

Typical ID50 value of divalent recombinant antibody-toxin functional
25 group fusion[B3(Fab)-Ext-PE38]₂ was measured on A431 adenosquamous carcinoma cell line at 4ng/ml, CRL1739 stomach cancer cell line at 1ng/ml, MCF-7 breast cancer cell line at 5ng/ml. The divalent recombinant antibody-toxin

functional group fusion[B3(Fab)-Ext-PE38]₂ showed nearly 12 times higher cytotoxicity than the monovalent antibody-toxin functional group B3(scFv)-PE40.

This invention dimer of chimeric recombinant binding domain(B)-functional group(F) fusion offers pharmaceutical drug formulation which includes
5 excipients like carriers or diluents. Pharmaceutical drug formulation containing dimer of chimeric recombinant binding domain(B)-functional group(F) fusion can be produced by commonly used method and prescribed in an appropriate way.

Dimer of chimeric recombinant binding domain(B)-functional group(F) fusion can be used as a drug alone or with radiotherapy, chemotherapy(cell
10 growth inhibitor, cytotoxic agents, antibiotic agents, alkylating agents, anti-metabolic agents, hormones, immunological agents, interferon agents, cyclooxygenase inhibitor(e.g. COX-2 inhibitor), metalomatrixprotease inhibitor, telomerase inhibitor, tyrosine kinase inhibitor, anti-growth factor receptor agents, anti-HER agents, anti-EGFR agents, anti-angiogenesis agents, farnesyl transferase
15 inhibitor, ras-raf signal transduction inhibitor, cell cycle inhibitor, cdk inhibitors, tubulin binders, topoisomerase I inhibitor, topoisomerase II inhibitor etc.) and others.

For example, dimer of chimeric recombinant binding domain(B)-functional group(F) fusion can be prescribed with liposomal agents containing
20 more than one chemical agents (e.g. taxane, taxane-derivatives, capsulated taxane, CPT-11, camptothecin-derivatives, anthracycline glycoside for example, idarubicin, epirubicin, etoposide, navelbine, vinblastine, carboplatine, cisplatin, estramustine, celecoxib, Sugen SU-5416, Sugen SU-6668, Herceptin etc.) used in chemotherapy.

25 If the mixed formulation of dimer of chimeric recombinant binding domain(B)-functional group(F) fusion and chemicals is not appropriate, they can be used on sequential order.

Dimer of chimeric recombinant binding domain(B)-functional group(F) fusion is administered through common routes and the dosage is decided according to the age, weight, condition and medication route of the drug. For example, the suitable dosage for dimer of chimeric recombinant binding domain(B)-functional group(F) fusion can be 10 ~2000mg , once to 5 times a day. Dimer of chimeric recombinant binding domain(B)-functional group(F) fusion can be medicated in various forms like oral delivery system as tablets, capsules, sugar-coated tablets, film-coated tablet, solution, suspension or suppositories systems or parenteral methods as injection into muscles, vein and/or central canal and/or spinal cord.

Solid drug for oral application contains active components with diluents (e.g. lactose, dextrose, sucrose, cellulose, corn starch, potato starch), lubricant (e.g. silica, talc, stearate, magnesium, calcium stearate and/or polyethylene glycol), bonding agent(e.g. starch, Arabic gum, gelatin methylcellulose, carboxymethyl cellulose, poly(N-vinyl pyrrolidone), disintegrants(e.g. starch, alginate, sodium starch, glycolate), formal mixture, dyes, sweetenings, wetting agents(e.g. lecithin, polysorbate, lauryl sulfate) and pharmaceutically inactive agents commonly used in pharmaceuticals. Using common methods, for example mixing, granulation, tablet formation, sugar or film coating can produce these drugs.

Liquid dispersion for oral application is for example, syrup, emulsion, and suspension. Emulsion and suspension can contain carriers and there are natural gum, agar, sodium alginate, pectin, methylcellulose, carboxymethylcellulose and polyvinyl alcohol, which are used. Suspension or solution for injection into the muscle may contain active components with pharmaceutically allowed carriers like distilled water, olive oil, ethyl oleate, glycol(e.g. propylene glycol) and appropriate dosage of lidocaine hydrochloride if needed. Carriers used for injections or infusions are distilled water and preferably

they are sterilized, hydrogenous, isotonic solution or they may contain propylene glycol as a carrier. Suppositories use carriers like cocoa butter, polyethylene glycol, polyoxyethylene sorbitan, fatty acid ester surface-active agent, lecithin.

The following will describe this invention more minutely in detail using
5 examples but it does not limit the boundary of the claim.

Advantageous Effects

The invention dimer of chimeric recombinant binding domain(B)-functional group(F) fusion formed via disulfide-bond-bridge between monomeric
10 chimeric recombinant binding domain(B)-functional group(F) fusion is valuable for the pharmaco-medical industry since it is the first dimer formed by covalent disulfide bond to have twice the binding valency and high structural stability and has excellent functional efficiency towards its targets and has high production yield by containing said flexible peptide chain (LFA).

15

Best mode for carrying out the invention

**Example 1: Production of dimers which have Ext (4CLFA5X = 4CL15FA11, 4CL20FA16, 4CL25FA21, AQ4CL20FA16) with fixed uncoupled Cysteine at fourth amino acid position and additional increase of five flexible
20 amino acids in its flexible linker sequence (LFA)**

[B3(Fab)-Ext(4CLFA5X = 4CL15FA11, 4CL20FA16, 4CL25FA21, AQ4CL20FA16)-toxin]₂ has uncoupled Cysteine at fixed 4th position and has an extension peptide chain for the formation of the dimer with five amino acid
25 additional increase in flexible amino acids linker(LFA).

On this example, the uncoupled Cysteine in the fourth location from the Fab of B3 antibody is fixed to induce dimer formation and with the uncoupled

Cysteine, different numbers of flexible amino acids on extension peptide chain were put in between Fd and toxin domain to compare production yield fluctuation caused by three-dimensional steric space difference. 11, 16, 21 flexible amino acids which make the length of Linker to be 15, 20, 25 were used on this test.

5 Alanine was used instead of Serine on the position of cysteine of natural antibody hinge sequence. The OH group of two serines takes more three-dimensional space than the natural cystine disulfide bond in dimer form. The reduced number of cysteine in extension sequence that is derived from antibody hinge sequence prevents scrambling of disulfide bonds.

10 The presented four types of [B3(Fab)-Ext(4CLFA5X)-PE38]₂ molecules had same disulfide bond formed between uncoupled Cysteines fixed at 4th position on Ext of both sides of monomers to make its interference on the formation of three dimensional structure and intermixing of disulfide bonds equal. And the amino acids following the uncoupled Cysteines were left the same except
15 that the numbers of flexible amino acids were increased to see the effect of the sterical tolerance given by the increased length of the LFA. Adding flexible amino acids led to 12~17 times higher production yield of these molecules than molecules produced before.

Therefore, this example compared with the accumulated previous results
20 shows the disulfide bond between uncoupled Cys on Ext of the two fusion monomer of large molecular weight 50kD B3 antibody and 38 kD Fab domain can be formed to produce the dimer on any position in certain range of the position of Ext and not on a specific point of Ext without interfering the large neighboring domains. So, this example shows definitely that the position of the uncoupled
25 Cysteine on Ext can be chosen generally in certain range and the increase of the number of flexible amino acids in the range of 11~21, which make the length of LFA to be 15~25, helps the disulfide bond formation for the dimerization of two

monomers.

(Apparatus and Methods)

E.coli BL21(DE3) was used for protein expression. Cancer cell lines used
 5 in cytotoxicity test were A431, CRL1739, MCF7, KB3-1. pMH21, 22, 23,
 pMHS22 is plasmid coding B3(Fd)-Ext(4CL15FA11, 4CL20FA16, 4CL25FA21,
 AQ4CL20FA16)-PE38 and plasmid pMC74, pCE2 is used for construction of the
 plasmids and plasmid which offers the light chain that matches with pMH21, 22,
 23, pMHS22 is pMCH75. The description of the plasmid, cell line and primer
 10 used in the experiment is on table 1, 2 and 3 respectively.

Table 1

Plasmid Name	Coding Polypeptide Chain
pMC74	B3(Fd)-SKPSISTKASGGPE-PE38REDLK
pCE2	B3(Fd)-SKPCISTKASGGPE- PE38REDLK
pMCH75	H6-B3(L)
pMH21	B3(Fd)-Ext(4CL15FA11)-PE38 =B3(Fd)-SKPCISTKAS(G4S)1GGPE- PE38REDLK
pMH22	B3(Fd)-Ext(4CL20FA16)-PE38 =B3(Fd)-SKPCISTKAS(G4S)2GGPE- PE38REDLK
pMH23	B3(Fd)-Ext(4CL25FA21)-PE38 =B3(Fd)-SKPCISTKAS(G4S)3GGPE- PE38REDLK
pMHS22	B3(Fd)-Ext(AQ4CL20FA16)-PE38 =B3(Fd)-AKPCLATQAS(G4S)2GGPE- PE38REDLK

Table 2

cell line	cell type	B3 antigen expression	Used Media
A431	Epidermoid	+++	RPMI 1640, 10% FBS
CRL-1739	Gastric	+	RPMI 1640, 10% FBS
MCR-7	Breast adenocarcinoma	+++	RPMI 1640, 5% FBS
KB3-1	Epidermoid cervix	-	DMEM, 5% FBS

Table 3

Name of the primer	Sequence 5'-3'
Primer MH-1	TAA TAC GAC TCA CTA TAG GGA GA
Primer MH-2	AGA TCC GCC ACC ACC AGA AGC TTT TGT ACT TAT GCT
Primer MH-3	CCA GAT CCG CCA CCA CCA CTT CCC CCT CCC CCG GAA GCT TTT GTA CTT ATG CTA GGC TTA CT
Primer MH-4	TGC TGG TGG CGG ATC TGG AGG TCC CGA GGG CGG CAAG C
Primer MH-5	TGG TGG TGG CGG ATC TGG AGG TGG CGG AAG CGG AGG TCC CGA GGG CGG CAG C
Primer MH-6	GCC GCG GGT GCT GAA GCT GAC GTC GCC GCC GTC
Primer MH-7	GGG AAT TCA TTA AGC TTG TGT AGC TAT GCA AGG CTT AGC ACC ACA

(Used reagents)

Tryptone, yeast extract, agar(Difco co.), minerals(Junsei co.) was used in
 5 culture medium for bacterial culture. Ampicilin (Sigma chemical co.) 200 μ g/mL
 was used for bacterial selection as final concentration added to the medium.

Nde I , *Hind*III, *Sal* I used for plasmid construction was products from
 NEB co. Ex *Taq* polymerase for PCR amplification and T4 DNA ligase for
 ligation were products from TaKaRa co. Coomassie Plus Protein Assay Reagent
 10 and BSA standard protein from Pierce co. were used for protein analysis. Buffers
 needed for dialysis, denaturation, refolding, protein purification were from Sigma.
 Isopropyl- β -thiogalctopyranoside(IPTG) for protein induction and urea for
 dialysis were products from Duchefa co. Q-sepharose(Pharmacia Biotech) and
 Source Q(Pharmacia Biotech) were used as column for purification of protein
 15 which is anion exchange chromatography and Superdex-200(Pharmacia Biotech)
 was used for size exclusion chromatography.

(Plasmid construction)

Construction of plasmid expressing B3(Fd)-Ext-PE38 recombinant protein was made by splicing PCR using pMC74(Fd-PE38) as a template to insert (G4S)_n inside of the KASGGPE in existing Fd-PE38. This (G4S)_n-inserted coding
5 sequence was exchanged with appropriate part of plasmid pCE2(vector which expresses B3(Fd)-SKPCISTKASGGPE-PE38) using *Hind* III, *Eco*R I .

Plasmid pMHS22(B3(Fd)-Ext(AQ4CL20FA16)-PE38) which has exchanged the Cysteine on the natural hinge sequence to Alanine was constructed by PCR using pMH22(B3(Fd)-Ext(4CL20FA16)-PE38) as a template to get a
10 fragment and then exchanged with appropriate part of pMH22.

pMH21, 22, 23 which expresses [B3(Fd)-Ext(4CL15FA11, 4CL20FA16, 4CL25FA21)-PE38]₂ and pMHS22 which expresses B3(Fd)-Ext(AQ4CL20FA16)-PE28 are shown in figure 1. Figure 2 shows the construction procedure for plasmid.

15

(Protein expression and isolation of inclusion body)

Proteins are expressed from plasmid pMH21, 22, 23, pMHS22, pMCH75 in *E.coli* BL21(DE3)(Studier *et al.*,1986). Bacteria was cultured at 37°C in superbroth(Tryptone 10g, yeast extract 5g, sodium chloride 10g) to which was
20 added 0.05% MgSO₄, 2% glucose, ampicilin 150µg/mL per liter. IPTG induction was at OD₆₀₀ 1.5~2.0 and it was cultured for 3 more hours until OD₆₀₀ 3. Cells were harvested at 3500rpm, 4°C for 20 minutes and collected in pellet. They were resuspended in 200mL ice-chilled sucrose solution and they were collected again at 8,000g, 4°C for 20 minute. The cell were resuspended again in ice-chilled
25 water 200mL to give osmotic shock and centrifuged at 15,000g, 4°C for 20 minutes to gain pellets. This pellet was resuspended in TE buffer(50mM Tris-Cl pH8.0, 20mM EDTA, pH8.0) and treated with lysozyme to remove peptidoglycan

layer. The highly concentrated salt 5M NaCl and 25% Triton X-100 was added and mixed evenly with tissuemizer, and incubated for 1 hour and centrifuged at 25,000g for 30 minutes to gain inclusion body pellet. To remove periplasmic protein 25% Triton X-100 was added and mixed evenly with tissuemizer, and
5 centrifuged at 25000g for 30 minutes and Triton X-100 treatment was repeated once again. The pellet was washed with 4M urea buffer(4M urea, 0.1M Tris-Cl). Residual Triton X-100 and urea was removed by resuspending and washing the pellet in TE buffer(50mM Tris-Cl pH7.4, 20mM EDTA pH7.4) and centrifuging at 25,000g, RT for 30 minutes for 3 times. The whole protein quantity of inclusion
10 body gained was analyzed using Coomassie Plus Protein Assay reagent and the antibody-toxin protein quantity was analyzed using Tina 2.0 program. These were stored in -70°C freezer before the refolding procedure.

(Refolding process and isolation of protein)

15 Each inclusion body was dissolved in solubilizing buffer solution(6M Guanidine-HCl, 0.1M Tris-Cl, pH8.0, 2mM EDTA pH8.0) and the 5mL solution of the 1:1 molar ratio mixture of B3(Fd)-Ext-PE38 and B3(L) was prepared to be 40mg in 5mL final volume with solubilizing buffer(use the dissolving buffer solution mixing with 40mg antibody-toxin protein and final volume to be 5mL to
20 make.). 0.06mM Dithiotreitol(DTT) was mixed into this for reduction procedure. Refolding procedure was taken in 500mL refolding buffer solution and started with rapidly diluting 5mL of above inclusion body mixture in solubilizing buffer solution, which is a 1:100 dilution ratio. The sample was incubated at 10°C for 48 hours. The quantity of antibody-toxin protein used in this procedure was 80mg
25 and 1L for refolding buffer solution. The refolded protein was taken into dialysis process and isolated through Q-sepharose, source Q, Superdex200 column chromatography(Choe et al., 1994).

(Analysis of cytotoxicity effect against cells)

The cytotoxicity effect of isolated protein[B3(Fab)-Ext-PE38]₂ was analyzed as described below. Antigen LeY expressing cell lines which are A431, CRL1739, MCF-7 and non-expressing cell line which is KB3-1 were diluted to be 1×10^5 cells/mL, aliquoted into 96-well plate in $180 \mu\text{l}$ /each well and cultured in a CO₂ incubator for 24 hours at 37°C. Purified antibody-toxins are serially diluted to be 10000ng/mL, 1000ng/mL, 100ng/mL, 10ng/mL, 1ng/mL, 0.1ng/mL, 0.01ng/mL in PBS with 0.2% BSA. The each of diluted antibody-toxin are added to 3 wells in $20 \mu\text{l}$ volume each well and incubated for 24 hours. 1 μCi of [³H]-leucine(NEN) was put into each well, cultured for 14 hours. The cells were put in to -70°C freezer and put them out to melt to detached the cells from the plate. The quantity of [³H] incorporated into living cell was analyzed using 1450 Microbeta TriLux Liquid Scintillation Counter(Wallac EG & Co.). The cells are shown in table 2.

(Construction of expression vectors and isolation of inclusion body.)

The nucleotide sequence of pMH21, pMH22, and pMH23, pMHS22 made from PCR was confirmed by nucleotide sequence analysis.

The amounts of proteins in inclusion body form produced by T7 polymerase system was 100~120mg/L culture for Fd-Ext-toxin, and 80~100mg/L culture for light chain. To isolate antibody-toxin protein from whole protein, osmotic pressure was firstly given to the cells to remove outer membrane which releases periplasmic protein and the pellet was cleaned by washing with 25% Triton X-100, 5M NaCl once, washing with 25% Triton X-100 once, washing with 4M Urea(4M urea, 0.1M Tris-Cl) once, 3 times of washing with TE buffer. Through this process, proteins except inclusion body were removed. The purity of antibody-toxin was analyzed on PAGE gel with densitometry(TINA 2.0), and the

purity of Fd-Ext-toxin and light chain was about 30%.

(Refolding of [B3(Fab)-Ext(4CLFA5X)-PE38]₂ molecule)

Each inclusion body, B3(Fd)-Ext-PE38 and B3(L) was used in 1:1 molar
5 ratio for refolding process. [Fab-toxin]₂ antibody-toxin with uncoupled Cysteine
for disulfide-bond-bridge at the first amino acid position on Ext(molecule from
pCE1) which was made in previous experiments, showed very low production
yield after the refolding(Choi *et al.*, 2001). The reason for this low productivity is
because of the uncoupled Cysteine is too close to the Fab region causing
10 scrambling with the internal Cysteines of Fab domain and another reason for it is
because of the number of flexible amino acid in Ext is too small(total FA is 7,
peptide linker LFA length is 13 in total) the three-dimensional hindrance between
the two big PE38 molecules couldn't be overcome. To solve this problem, the
uncoupled Cysteine was transferred to the 15th position of Ext chain to get far
15 from Fab region and the number of flexible amino acid (FA) of LFA chain was
increased to 13(total amino acid number of LFA is 14, the dimer from pCW1) to
give spaces and rotational freedom for effective refolding and dimerization. In this
case the disulfide bond is formed between the cysteines on the inserted extension
peptide. When the flexible extension peptide is inserted, the productivity of
20 [B3(Fab)-Ext(15CL14FA13)-PE38]₂ has increased to 0.06% from 0.014% of the
previous [B3(FabH1)-PE38]₂ (= [B3(Fab)-Ext(1CL13FA7)-PE38]₂). This is
because the uncoupled Cysteine transferred to the 15th position doesn't cause
disulfide bond scrambling with either Cysteines inside of Fab or on Ext and this
does not cause any interference to the formation of the three dimensional structure.
25 Also the increase of flexible amino acid in LFA makes the three dimensional
hindrance between PE38 reduced.

Therefore, in this example the inventors tried to confirm that it is not true

that the uncoupled Cysteine only at a specific position allows the formation of dimer without interfering interactions with Fab and/or PE38 functional group but that it is true that the uncoupled Cysteine on any point in some range of the position enables dimerization too. Also, it was confirmed with an uncoupled amino acid fixed at the middle 4th position and increasing the numbers of flexible amino acid in LFA was serially by 5 flexible amino acid that the number and the sequence of amino acid of LFA can be in certain range to enable the formation of dimer with various production yield and that it do not need to be of specific number and sequence as that of the previous molecules.

10 The inventors held experiments as below.

1) Comparing the production yield of [B3(Fab)-Ext(4CLFA5X)-PE38]₂ according to the LFA chain length and number of flexible amino acids when the Ext chain of Ext(4CL15FA11,4CL20FA16, 4CL25FA16) has 11, 16, 21 flexible amino acids.

15 2) Examining the effects of [B3(Fab)-Ext(4CLFA5X)-PE38]₂ having the disulfide bond formed by uncoupled Cysteine not locating in the 1st or 15th but 4th position.

3) Examining the effects of [B3(Fab)-Ext(AQ4CL20FA16)-PE38]₂ having AKPCIATQ instead of SKPCISTK in Ext sequence which is derived from the hinge of antibody.

The results are as followings.

1) When 11,16,21 flexible amino acids were inserted, the maximum production yield of [B3(Fab)-Ext(4CLFA5X)-PE38]₂ increased to 0.17~0.25%. This is 12.1~17.8 times higher than [B3(FabH1)-PE38]₂(=[B3(Fab)-Ext(1CL13FA7)-PE38]₂) and 2.8~4.1 times higher than [B3(Fab)-Ext(15CL14FA13)-PE38]₂.

Table 4

Plasmid name	Structure of Ext	Yield of dimer(%)	Position of Cys in Ext	Length of L (distance between Cys and F)	Number of GASQEND in L
pCE1	1CL13FA7	0.016	1	13	7
pCW1	15CL14FA13	0.06	15	14	13
pMH21	4CL15FA11	0.18	4	15	11
pMH22	4CL20FA16	0.23	4	20	16
pMH23	4CL25FA21	0.25	4	25	21
pMHS22	AQ4CL20FA16	0.24	4	20	17

According to table 4, one can find the examples of the molecules with LFA peptide, which has 13 to 25 amino acids between the uncoupled Cysteine and PE38. Among these, the production yield of $[B3(FabH1)-PE38]_2 (= [B3(Fab)-Ext(1CL13FA7)-PE38]_2)$ which has 13 amino acids is 0.014%, which is nearly no production. But using 14, 15 amino acids which has insignificant difference with the above 13, the maximum production yield of dimer increased to 0.17~0.25%. This shows that for preventing the hindrance between PE38, the flexible amino acid containing LFA peptide is essential although peptides between uncoupled Cysteine and PE38 are similar. When inserting 11,16,21 amino acids containing LFA peptide, the production yield increases from 0.17% to 0.25% though it is in small quantities and this indicates L peptides with flexible amino acid are needed for preventing the hindrance between two PE38. Also, the case of pMH23(B3(Fd)-Ext(4CL25FA21)-PE38) indicates that, even though the 25 amino acid LFA having 21 non-bulky flexible amino acids exists between disulfide bond and PE38, this long length of LFA didn't allow PE38 functional group to disturb Fab refolding or binding activity and active dimer $[B3(Fab)-Ext(4CLFA5X)-PE38]_2$ was formed.

2) The production yield of $[B3(Fab)-Ext(15CL14FA13)-PE38]_2$ was

0.06% when the disulfide bond is transferred to the 15th location of extension chain but, when the disulfide bond is on the 4th location the maximum production yield increased to 0.17~0.25%. This indicates that the uncoupled Cysteine forming disulfide bond doesn't have to take specific position and the possible range is very wide to form appropriate disulfide bonds without disturbing the formation of tertiary and quaternary structure of neighboring binding domain and functional group which are massive structure. This example wants to shows whether the sequence, which has derived from natural antibody hinge region, on extension chain helps disulfide bonding for dimerization. For this, there are facts from the preceding experiment that the production yield of [B3(FabH1)-PE38]₂(=[B3(Fab)-Ext(1CL13FA7)-PE38]₂) having Ext with modified sequence derived from antibody hinge and uncoupled Cysteine was 0.016%, and the production yield of [B3(Fab)-Ext(15CL14FA13)-PE38]₂ having Ext with modified sequence derived from antibody hinge and uncoupled Cysteine not on the hinge derived sequence but on irregular flexible sequence after the hinge derived sequence showed the increased yield to 0.06%. But, if when the uncoupled Cysteine was put into the middle of the 1st and 15th position, the production yield was higher than both of the previous 1st and 15th position. Therefore, the conclusion for whether the hinge sequence derived from antibody has effects on dimerization is not clear.

3) For the construction of plasmid pMHS22, the Serine(S) on SKPCISTK sequence which is derived from hinge was exchanged with Alanine(A) that has no -OH group. This is because the -OH group on Serine may make the space sterically too crowded and disturb three-dimensional structure to decrease production yield of dimer. Also, uncoupled Cysteine was put into the 4th location and exchanged bulky amino acid Lysine(K) which takes big volume to flexible amino acid Glutamine(Q) to make AKPCIAATQ. The productivity of [B3(Fab)-

Ext(4CL20FA16)-PE38]₂ by pMH22 having SKPCISTK and [B3(Fab)-Ext(AQ4CL20FA16)-PE38]₂ by pMHS22 having AKPCIATQ was similar which is nearly 0.2%.

Through this experiment, the conclusion for the dimerization of antibody-toxin which is a kind of fusion of binding domain and functional group is that the location of uncoupled Cysteine in extension chain doesn't have to be specific like naturally coupled Cysteines for intra- or inter-chain disulfide bond and it can have wide range for dimerization. The relationship between uncoupled Cysteine and dimerization is shown in figure 4.

Also, to decrease three dimensional hindrance between the big functional group PE38, the flexible amino acids are needed following uncoupled Cysteine, and according to the experiment the productivity increased with the increased numbers of flexible amino acids. There are 21 flexible amino acids in LFA(flexible chain) for independent refolding of PE38 and Fab and it did not cause inter-domain hindrance between Fab and PE38 though it is a 25 amino acid long length of flexible linker chain, and the formation of [B3(Fab)-Ext(4CL25FA21)-PE38]₂ is allowed with appropriate disulfide bond between uncoupled Cysteine.

(Purification of [B3(Fab)-Ext(4CLFA5X)-PE38]₂)

To purify refolded [B3(Fab)-Ext(4CLFA5X)-PE38]₂, the fact that PE has net negative charge was exploited and anion exchange resin was used. First, the one liter of dialyzed sample was loaded into Q-sepharose column and the sample was washed and eluted with buffer A and B. The eluted sample was analyzed on 12% reducing and 8% nonreducing SDS-PAGE electrophoresis. The fractions of divalent antibody-toxin was pooled and were gel filtrated using superdex 200 column and eluted with PBS. The final column result is on figure 5.

(Cytotoxicity assay)

Using ^3H -incorporation method, the cytotoxicity effect of antibody-toxin was examined on four cancer cell lines. The ADP-ribosylation ability of third domain of PE38 inactivates elongation factor II to inhibit protein synthesis. scFv-
5 PE38 was used as reference molecule for comparison. KB3-1 cancer cell line was used as negative control cell. ID_{50} is the concentration of antibody-toxin when $[^3\text{H}]$ -leucine-incorporation into cancer cell decreases to 50%. The cytotoxicity assay was held three times each in triplicate samples and the triplicate values were taken average to evaluate the result. The source of error in triplicate cytotoxicity
10 assay might be from sensitive cell condition, mix of active and inactive antibody-toxins obtained through *in vitro* refolding and purification, and contaminant proteins. But, although the three values are a little different from each other, they are in the range of experimental errors. The results showed same ID_{50} values as that of the reference monovalent molecule and also lower values than that of the
15 reference molecule, which means that the dimer has higher cytotoxicity than the reference. This is due to antigen density and structural conditions on the cell surface as observed in the previous case. More knowledge on cell surface antigen structure will make the explanation possible.

If cytotoxicity from divalent molecule is superior to monovalent
20 molecules the treatment effect of divalent molecule will be higher at the same dosage. Also, the lowest limit of the therapeutic window that is determined by the minimum dosage showing the effect of the drug can be lowered by using the dimer as it has higher binding efficiencies than the monovalent and the same therapeutic effect can be obtained with less amount of dose. This means side
25 effect caused by the functional groups with physiological activities can be overcome by the use of dimer as it can give same therapeutic effect with lower dose.

Example 2: The dimerization and effects of antibody-toxin fusion that has extension peptide chain(Ext) with uncoupled Cysteine at the 15th position and 14 amino acids linker(LFA) containing 13 flexible amino acids.

5 In this example, the B3(Fd)-Ext(15CL14FA13)-PE38 was newly constructed. The flexible peptide G4C(G4S)₂ was inserted following the fourth Serine on extension chain SKPSISTKASGGPE of B3(Fd)-PE38. The Cysteine on sequence derived from antibody hinge sequence was all changed into Serine that does not form disulfide bond. The composition of B3(Fd)-Ext(15CL14FA13)-
10 PE38 is same to B3(VH)-B3(CH1)-SKPSISTKASGGGGCGGGGSGGGGSG GPE-PE38. The uncoupled Cys on Ext forms disulfide binds between two B3(Fab)-Ext(15CL14FA13)-PE38 monomers. The uncoupled Cys was located at 15th position of extension chain compared to the previous example which uses uncoupled Cys on 1st position which is [B3(FabH1)-PE38]₂(=[B3(Fab)-
15 Ext(1CL13FA7)-PE38]₂)(protein from pCE1) to place Cys far from binding domain Fab and this showed whether the position of Cys doesn't have to be specific for disulfide bonding. Also, the inventors wanted to find out whether dimerization of B3(Fab)-Ext(15CL14FA13)-PE38 is enhanced by putting 13 flexible amino acids in peptide linker following uncoupled Cys which will give
20 more spaces between PE38 and decrease three dimensional hindrances between them during dimerization.

The B3(Fab)-PE38 derived divalent immunotoxin is expected to have more merits than monovalent immunotoxin. First, the binding affinity will be stronger because it's divalent binding valency. Second, cytotoxicity against cancer
25 cells will be better. Third, the stability in circulation of blood will be better. The longer flexible extension peptide chain was applied expecting that [B3(Fab)-Ext(15CL14FA13)-PE38]₂ which has longer flexible extension peptide chain has

higher productivity and better refolding yield than [B3(Fab)-Ext(1CL13FA7)-PE38]₂ having about half number of flexible amino acid in liker L chain.(Choi. *et al.*, 2001)

5 (Apparatus and Methods)

The construction of B3(Fd)-Ext(15CL14FA13)-PE38 fusion protein was performed by splicing PCR using four primers of the template pMC74(Fd-SKPSISTKASGGPE-PE38 protein expressing vector), G4C(G4S)₂ which follows after the fourth Serine on SKPSISTKASGGPE sequence between Fd and PE38.

10 The Fd and PE38 fragments from PCR were purified and splicing PCR was performed using the appropriate primers. The products from splicing PCR was purified and they were exchanged with appropriate part of pMC74(Fd-PE38 expressing vector) which was cut with *Nde* I and *Sac* II.

The expression of protein, isolation of inclusion body, refolding, isolation
15 of protein and cytotoxicity assays were performed the same as in example 1.

(Construction of pCW1 Expression vector and isolation of inclusion body)

The structure of pCW1 which expresses B3(Fd)-Ext(15CL14FA13)-PE38 chain composing [B3(Fab)-Ext(15CL14FA13)-PE38]₂ is shown on figure 6.

20 B3(Fd)-Ext(15CL14FA13)-PE38 is a modified form of [B3(Fd)-SKPSISTKASGGPE-PE38] using PCR. The peptide chain G4C(G4S)₂ was inserted between S and G on KASGGPE of [B3(Fd)-SKPSISTKASGGPE-PE38]. The plasmids used are on table 5.

Table 5

Name	Coding Proteins
pMC74	B3(Fd)-SKPSIST-KASGGPE-PE38REDLK
pMC75	B3(L)
pCW1	B3(Fd)-SKPSIST-KASG4C(G4S)2GGPE -PE38REDLK

The nucleotide sequence of pCW1 was confirmed through sequencing. The polypeptide was gained from the inclusion body of *E. coli*. The protein purity of inclusion body was measured as 40~60% by densitometry(TINA2.0).

5 (The refolding of [B3(Fab)-Ext(15CL14FA13)-PE38]₂ molecule)

The refolding procedure was performed by mixing in the inclusion bodies B3(Fd)-Ext(15CL14FA13)-PE38 and B3(L) in 1:1 molar ratio. The [B3(Fab)-CKPSISTKASGGPE-PE38]₂ (= [B3(Fab)-Ext(1CL13FA7)-PE38]₂) immunotoxin which was reported previously was a form having the uncoupled Cysteine at the
 10 1st position of the extension chain to form disulfide bond between two monomer, and it showed low production yield after refolding(Choi. et al.,2001). This is because of the location of uncoupled Cysteine that is too close to the Fab region, and also because of the three dimensional hindrance between the two big PE38. The solution for this is to transfer the uncoupled Cysteine far from Fab and
 15 increase the flexible amino acids in LFA to decrease the three-dimensional collision of PE. The designed molecule have increased number of flexible amino acids in LFA and increased spaces for the rotational freedom of PE that will help dimerization. The uncoupled Cysteine in extension peptide chain forms disulfide bond to form dimer[Fab-Toxin]₂, and LFA contains 14 amino acids of GASQEND
 20 group. The dimer([B3(Fab)-Ext(1CL13FA7)-PE38]₂) containing Fd-CKPSISTKASGGPE-PE38 from pCE1 showing low productivity is composed of 13 amino acids from the point of disulfide bond and the start point of PE. [B3(Fab)-Ext(1CL13FA7)-PE38]₂ (=Fd-SKPSISTKASGGGGCGGGSGGGGS GGPE-PE38) also has 14 amino acids in the same area but the productivity by
 25 refolding procedure was examined as 0.06% which is 4~5 times bigger than the dimer[B3(Fab)-Ext(1CL13FA7)-PE38]₂ (=Fd-CKPSISTKASGGPE-PE38) made from pCE1. Therefore, the insert of flexible amino acids increase refolding

efficiency by allowing flexible movement of the chain and decreasing three dimensional collision at dimerization with each B3(Fab)-Ext(15CL14FA13)-PE38.

(Purification of [B3(Fab)-Ext(15CL14FA13)-PE38]₂)

- 5 The same method was used as example 1. The final column result is shown on figure 7.

(Cytotoxicity assay)

- 10 The experiment was performed in triplicate as same as example 1. The ID₅₀ of divalent immunotoxin[B3(Fab)-Ext(15CL14FA13)-PE38]₂ was measured as 4ng/mL with A431 cell line, 1ng/mL with CRL1739 cell line, 5ng/mL with MCF-7 cell line. The ID₅₀ of monovalent immunotoxin B3(scFv)-PE40 was measured highly as 5ng/mL with A431, 12ng/mL with CRL1739, 10g/mL with MCF-7. The result for monovalent immunotoxin B3(scFv)-PE40 as a control was
- 15 exactly the same as reported previously(Brinkmann *et al.*,1991). The ID₅₀ of B3(scFv)-PE40 is appropriate for using it as a control because the value was measured many times, and if the known value is gained is obtained in the assay it means that the error in the cytotoxicity assay is negligible. The results are shown on figure 8 and table 6.

20 Table 6

Cell line	Cytotoxicity (ID ₅₀ ; ng/ml) on cell lines of B3 antigen	
	scFv-PE40	[B3(Fab)-Ext(15CL14FA13)-PE38] ₂
A431	5	4
CRL1739	12	1
MCF7	10	5
KB3-1	>1000	>1000

This results show that the divalent immunotoxin[B3(Fab)-Ext(15CL14FA13)-PE38]₂ has 12 times the higher cytotoxicity than the

monovalent B3(scFv)-PE40 with CRL1739 cell line.

The CRL1739 cell line that showed 12 times differences in cytotoxicity assay is a stomach cancer cell line. This means that the divalent immunotoxin has more cytotoxicity depending on the surface structure against cancer cells that express same LeY antigen. If LeY is on very long and flexible structure and binding of divalent immunotoxin to two LeY antigens at same time is easy, divalent immunotoxin will bind two of LeY simultaneously showing more binding affinity than monovalent immunotoxin. If LeY is on non-flexible structure or LeY is too far apart from each other, then the binding will be the same for both divalent and monovalent immunotoxin as the binding of the immunotoxin to antigen is through only one binding domain to one antigen even though the molecule has divalent binding domain. In the case of cell line CRL1739, it seems like that it has LeY on a very long and flexible polysaccharide structure and allows [B3(Fab)-Ext(15CL14FA13)-PE38]₂ to bind two antigens simultaneously. The binding of antibody with antigen on cell surface depending on their interaction can stimulate the growth of the cell. Some cell growth was observed at very high concentration of divalent immunotoxin (data not shown) but not for monovalent immunotoxin. The previously reported monovalent B3 immunotoxin haven't been yet reported that it stimulates cell growth by interaction of antigen-antibody. The differences between monovalent and divalent immunotoxin on cell growth are remained to be investigated.

Example 3: The dimerization of antibody-toxin fusion that has extension peptide chain(Ext) with uncoupled Cysteine at the 1st position and 13 amino acids linker(LFA) containing 7 flexible amino acids

The B3(Fab) and PE38 are used in constructing divalent immunotoxin

[B3(FabH1)-PE38]₂ (= [B3(Fab)-CKPSISTKASGGPE-PE38]₂) (= [B3(Fab)-Ext(1CL13FA7)-PE38]₂). The divalent immunotoxin B3 has about 174.4kDa of molecular weight and comprises two light chain of B3 and two chains composed of B3(Fd) fused with PE38. There are three Cysteines on the hinge sequence of B3 antibody. The rear two Cysteines are changed to Serines and only the front Cysteine is used to make the Ext chain. It forms disulfide bond between monovalent immunotoxin to produce dimer.

(Materials and methods)

The method used are same as example 1. MTT(3-[4,5-Dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) from Sigma Co. was used at 1.5mg/mL final concentration. Cy5TM AutoReadTM sequencing kit for sequence analysis was a product from Pharmacia Biotech. The plasmids used are shown on table 7.

Table 7

Plasmid Name	Coding Protein
pMC75	B3(L)
pCE1	B3(Fd)-CKPSISTKASGGPE-PE38
pMC74	B3(Fd)-SKPSISTKASGGPE-PE38
pMC76	B3(L)-KASGGPE-PE38

The method using chromatography column was the same as example 1.

(Construction of plasmid pCE1 expressing B3(FdH1)-PE38=B3(Fd)-Ext(1CL13FA7)-PE38)

Plasmid pCE1(expressing B3(FdH1)-PE38) is a modified form of pMC74(expressing B3(Fd)-PE38=B3(Fd)-SKPSISTKASGGPE-PE38). PCR was performed to change the sequence to pCE1. (pCE1 expresses B3(FdH1)-PE38=B3(Fd)-CKPSISTKASGGPE-PE38).

The constructed expression system was confirmed by DNA sequence analysis. Primers used were designed appropriately to be used with each template DNA and to code designed amino acid sequence. B3(L) is expressed from pMC75.

5 (Preparation of protein)

The methods for expression of B3(FdH1)-PE38 and B3L, purification of spheroplast and inclusion body, quantity analysis of inclusion body, refolding procedure, purification of B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂ were same as example 1.

10

(Measurement of [B3(FabH1)-PE38]₂ formation according to temperature change)

[B3(FabH1)-PE38]₂ formation from B3(FabH1)-PE38 monomer isolated by Mono-Q was measured at 37°C, 40°C, 45°C, 53°C incubating for 24hours. [B3(FabH1)-PE38]₂ was purified with Mono-Q and proteins of each fraction were
15 analyzed with electrophoresis.

(Measurement of [B3(FabH1)-PE38]₂ formation using cross-linkers)

B3(FabH1)-PE38 obtained from TSK-GEL G3000SW and cross-linkers, which are bis-maleimido-hexane(BMH) and 1,11-bis-
20 maleimidotetraethyleneglycol (BM[PEO]₄), were reacted together in molar ratio of 6:1, 5:1, 4:1, 3:1, 2:1, 1:1, 1:3 in reaction buffer to connect the Thiol(-SH) groups of Cysteine of two B3(FabH1)-PE38. The proteins were analyzed by electrophoresis.

25 (Cytotoxicity assay of purified proteins against cancer cells)

Same method with example 1 was used except following description. The viability of cancer cells was measured according to decomposition of MTT to

analyze cytotoxic effect of purified B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂. Cancer cells were cultured with immunotoxin and 20 μ l of 5mg/mL MTT was added. The samples were wrapped with aluminum foil and left alone at 37 °C, 5% CO₂ incubator for 10 hours. The reduced MTT-formazan was centrifuged at
5 3000rpm for 4~15 minutes to make MTT-formazan crystal to pellet and 200 μ l of supernatant was removed. 0.016N acidic isopropanol 100 μ l was added. Microfilter plate shaker was used at 300rpm to melt MTT-formazan and the optical density was measured using ELISA READER at 570nm. Results were average values of three samples.

10

(Purification of B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂)

Same method as example 1 was used.

The quantity of B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂ was 4.3mg and 16.5mg respectively which is 3.8% and 0.016% of total quantity of
15 B3(FdH1)-PE38 and B3L used in refolding. The reason for the low productivity of dimer [B3(FabH1)-PE38]₂ is that lots of B3(FdH1)-PE38 and B3L coagulate and precipitate to disappear during refolding and dialysis. Also the reason for the low productivity of [B3(FabH1)-PE38]₂ in spite of 3.8% B3(FabH1)-PE38 produced, is that Cysteine is located in a place which is difficult for disulfide bond
20 to form due to unfavorable interactions of Fab and/or PE38 or being in a wrong orientation or being hid inside of the three-dimensional structure. But the small quantity of [B3(FabH1)-PE38]₂ obtained in this example was assumed to be through the rare disulfide bond formation before complete refolding between small amounts of B3(FdH1)-PE38, which can happen before the disturbance on
25 disulfide bond formation by Fab and PE38, and later the disulfide bond is formed with B3L and the dimers are produced.

(Measurement of [B3(FabH1)-PE38]₂ formation according to temperature change)

Because the productivity of [B3(FabH1)-PE38]₂ is very low, the formation by disulfide bond between B3(FabH1)-PE38 was checked with heating the monomer B3(FabH1)-PE38 to relax the structure of the monomer and to free the Cysteines from the unfavorable interactions of Fab and/or PE38 and to allow the oxidation for disulfide bond formation. Productivity of [B3(FabH1)-PE38]₂ was highest at 45°C heating observed with B3(FabH1)-PE38 monomer isolated from Mono-Q. But many other kinds of side products besides [B3(FabH1)-PE38]₂ was formed too.

10

(Measurement of [B3(FabH1)-PE38]₂ formation using cross-linkers)

To increase productivity of [B3(FabH1)-PE38]₂, the cross-linkers(BMH and BM[PEO]₄) was used to connect the -SH group between two B3(FabH1)-PE38. The formation of [B3(FabH1)-PE38]₂ was hardly observed when the molar ratio of B3(FabH1)-PE38 and cross-linker was used at 6:1, 5:1, 4:1, 3:1, 2:1, 1:1, 1:3 in the reaction. This result shows when the refolding with cross-linker was performed, the Cysteines for disulfide bond still can not get close each other to be linked by the cross-linker because of unfavorable interactions with Fab and/or PE38, or because of being in a wrong orientation, or because of being hid in three-dimensional structure, or because the Lysine which follows after Cysteine is too big and interrupts the bonding between cross-linkers and Cysteines leading to the failure of forming a dimer between two monomers.

15
20

(Cytotoxic assay of purified proteins on cancer cells)

The cytotoxicity effect on cancer cells of B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂ can be observed by measuring viability of cancer cells which is indicated by the amount of MTT-tetrazolium being reduced to MTT-formazan

25

by the mitochondrial dehydrogenase. The cancer cells used were A431, MCF7, CRL1739, which have B3 antigens, and KB3-1 as a negative control, which doesn't have B3 antigen. B3(scFv)-PE40 is a single chain immunotoxin, which is a monovalent molecule, and it was used as reference molecule. The ID₅₀ which shows cytotoxic effect of B3(scFv)-PE40, B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂ was 61.3ng/mL, 30ng/mL, 10.3ng/mL respectively. However, MTT-tetrazolium has to be delivered to the mitochondria in the MTT assay and it may introduce errors depending on the conditions of each cancer cells. But B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂ had 2 times and 6 times higher effects than B3(scFv)-PE40 respectively. However, the reference molecule showed very low cytotoxic activity compared to the previous reports reported as 2~5ng/mL ID₅₀. This is because cytotoxic assay is very sensitive to conditions of the cells and the immunotoxins were not purified enough and contain impurities. As the errors of MTT assay is relatively large the reconfirmation of the results is preferable.

In conclusion,

1) B3(FdH1)-PE38 and B3L occupied 17~25% of the total expressed protein in the cell when they were over-expressed.

2) Production yield of B3(FabH1)-PE38 was 4.3mg which is 3.8% and [B3(FabH1)-PE38]₂ was 16.5 μ g which is 0.016% when prepared with B3(FdH1)-PE38 and B3L in 82mg and 32.2mg amounts respectively in the refolding procedure.

3) The productivity of proteins with correct conformation formed through 100-fold rapid dilution refolding procedure is very low, and 96% of the protein were incorrectly refolded and aggregated.

4) The highest productivity is obtained at 45 °C when [B3(FabH1)-PE38]₂ was formed by heating B3(FabH1)-PE38 monomer.

5) B3(FabH1)-PE38 doesn't form dimers by connecting the monomer

with long cross-linkers like BMH and BM[PEO]₄.

6) The Cysteins used in disulfide bonding between B3(FabH1)-PE38 is placed in an environment difficult to form dimers.

7) The purified B3(FabH1)-PE38 and [B3(FabH1)-PE38]₂ prepared from
5 large amount of materials showed 2 times and 6 times higher cytotoxicity respectively than B3(scFv)-PE40 reference molecule.

8) For high purity preparation of [B3(FabH1)-PE38]₂, large amount of materials are needed and the contaminating proteins have to be removed.

10 **Example 4: The dimerization of antibody-toxin that has extension peptide chain(Ext) with uncoupled Cysteine at 1st, 4th, 6th position and affinity domain containing flexible liker peptide (LADFA) containing CH2, CH3 self-affinity domain of antibody and (G4S)₂ sequence.**

15 On this experiment, three uncoupled Cysteines were positioned at the frontal region of extension chain, and they form very thermodynamically stable dimer with three disulfide bonds. In other words, this is to test whether the big sequence Fab and big group PE38, which has natural internal Cysteines, can be refolded into right conformation without being scrambled with three uncoupled
20 Cysteines on extension chain. In another words, this is to confirm whether binding domain-functional group fusion dimer can be formed in triple disulfide bonded thermodynamically highly stable structure while the extension chain connecting binding domain and functional group has multiple uncoupled Cysteines. Predicting whether the multiple uncoupled Cysteines will find their disulfide bond
25 couple correctly without bothering or intermixing with neighboring big sequences to form dimer is not easy for manufacturers concerned.

[B3(Fab)-h(H123-CH2/CH3/Fc)-PE38R]₂ having Ext(LADFA) was

produced which has Fc, CH2, CH3 domain in LFA sequence having self-affinity for easy meeting of the Cysteines for dimerization and flexible amino acid sequence following the 'self-affinity domain'. CH3 domain was reported not to disturb antigen binding and induces homodimerization (Acpua *et al.*, 1998, 5 Ridgway *et al.*, 1996). Also, the molecule with CH3 domain has similar antigen binding affinity and homodimerization ability as those with Fc domain which has both CH2 and CH3 domain (Alt *et al.*, 1999). Therefore, the merit of stability of Fab-toxin in blood circulation is saved and also production yield of dimer[Fab-toxin]₂ has been increased through the insertion of Fc, CH2, CH3 domain that 10 causes homodimerization(Wu *et al.*, 2001). Also, binding affinity of divalent molecule and monovalent molecule was compared (Gall *et al.*, 1999).

(Materials and Methods)

E.coli BL21(DE3) was used for protein expression system.

15 For the construction of plasmid having Fab-h(H123-CH2/CH3/Fc)-PE38 chain gene, Fab-PE38 was obtained from pMC74 as a template and Fc region was obtained from human hinge(including three uncoupled Cysteines) and Fc containing pcDNA3Cγ1 as a template. For the light chain, 5'-end 6xHis tagged chain from pMCH75 was used. Each name and construction procedure of the 20 plasmid is on figure 12,13,14 and table 8. Mediums, reagents, enzymes, columns are used the same as example 1.

Table 8

Name	Coding Proteins
pLSC52	B3(Fd)-human(H123-Fc)-(G4S)2-KASGGPE-PE38REDLK
pLSC32	B3(Fd)-human(H123-CH3)-(G4S)2-KASGGPE-PE38REDLK
pLSC22	B3(Fd)-human(H123-CH2)-(G4S)2-KASGGPE-PE38REDLK
pMC74	B3(Fd)-SKPSIST-KASGGPE-PE38REDLK
pMCH75	H6-B3(L)

(Construction of plasmid and Preparation of protein)

To construct pLSC52 which contains all of the Fc, PCR of the region from hinge to Fc using pcDNA3Cγ1 as a template was performed. And PCR of Fd region from pMC74 was performed and splicing PCR was performed with
5 these two fragments. Produced fragments and pMC74 were digested with *Nde* I and *Hind*III for insertion ligation. Plasmids with CH2, CH3 were constructed same way and called pLSC22 and pLSC32 respectively.

Methods for expression and isolation of protein and protein refolding procedure were the same as example 1. The molar ratio of LSC52 or LSC32 or
10 LSC22 to MCH75 was 1:1 and the quantity of IT protein added was 40mg/5mL for the condition of 500mL refolding and the method was same as example 1. Method for purification of refolded antibody-toxin was the same as example 1.

The method for cytotoxicity assay on 4 cell lines with [B3(Fab)-h(H123-Fc)-PE38]₂, [B3(Fab)-h(H123-CH3)-PE38]₂, [B3(Fab)-h(H123-CH2)-PE38]₂,
15 [B3(Fab)-h(H123-CH2)-PE38] isolated from Superdex 200 was same as example 1.

(Production and purification of IB protein)

To produce [B3(Fab)-h(H123-CH2/CH3/Fc)-PE38]₂, plasmids pLSC22, pLSC32, pLSC52 encoding [B3(Fd)-h(H123-CH2/CH3/Fc)-PE38]₂ was
20 constructed and pMCH75 was used for light chain.

The inclusion body protein preparation from T7 polymerase mass production system was repeated 12 times to get 38.2mg/liter culture in average. Measured by densitometry analysis(Tina2.0) the purity of protein on the PAGE
25 gel was 34.2% for heavy chain which was the average of 9 measurements and 40.9% for light chain which was the average of 3 measurements. Inclusion body was analyzed through SDS-PAGE. To enhance the purity of inclusion body,

additional washing procedure treating with 4M Urea buffered by Tris-Cl pH7.4 for 3 hours was performed in addition. The purity of inclusion body was over 30% and the protein folding products were observed.

5 (Refolding of proteins)

Disulfide bond during refolding procedure can be formed in two different kinds. One is disulfide bond inside the domain chain, intra-chain disulfide bond, and the other is disulfide bond between the chains, inter-chain disulfide bond. B3(Fd)-h(H123-CH2/CH3/Fc)-PE38 from pLSC22, pLSC32, pLSC52 and light
10 chain from pMVH75 forms disulfide bond between Fd and L in refolding solution and composes Fab domain. Disulfide bonds are formed between the H123 Cysteine on two of B3(Fd)-h(H123-CH2/CH3/Fc)-PE38 monomer to make a dimer molecule.

15 (Purification of refolded antibody-toxin)

After the purification, the yield of [B3(Fab)-h(H123-Fc)-PE38R]₂ which has Fc was the lowest value because the molecule size is relatively big(divalent: 228kd) and the refolding is easily disturbed. [B3(Fab)-h(H123-Fc)-PE38R]₂ which has both the CH2 and CH3 domains had a dragging bands other than the
20 right size refolded band compared to molecules which has only one domain. This is because the size of [B3(Fab)-h(H123-Fc)-PE38R]₂ is big to make other false molecules during refolding.

[B3(Fab)-h(H123-CH3)-PE38R]₂ which has CH3 didn't show the peaks of divalent and monovalent molecule separated on Superdex 200 column. This is
25 because they exist as divalent form while being isolated through superdex200 column because of the strong self-affinity of CH3 domain. But during SDS-PAGE, the dimer molecules that is formed only by CH3 affinity but not by

covalent disulfide bond falls apart and are observed as monomers on the SDS-PAGE. According to band intensity, the dimers formed by disulfide bonds on the Ext or formed by CH3 affinity are in similar quantity. Therefore the CH3 region is effective for dimerization of Fab molecule keeping them in close proximity, and
 5 refolding with dimer formation will take place though the numbers of uncoupled Cysteine is 3. The immaturation of disulfide bonds within affinity domain held dimers could be matured by long exposure to oxygen in the air. The immaturation of disulfide bonds are well known to manufacturers concerned who are experienced.

10 Finally, the quantity of [B3(Fab)-h(H123-CH2)-PE38R]₂ which has CH2 is small but that of monovalent is big. This shows that CH3 self affinity is much stronger than CH2 self affinity to form more dimer. Through the known fact that self affinity of CH2 is very weak and through the experimentally observed fact that dimer which only has CH2 domain forms in very small quantity, it can be
 15 deduced that the dimer can be formed through disulfide bond even in the case of the three uncoupled Cysteines on the extension chain with small helps from the affinity domain.

(Cytotoxicity assay)

20 Cytotoxicity effect of purified immunotoxin[B3(Fab)-h(H123-CH2/CH3/Fc)-PE38]₂ were tested on 4 cancer cell lines. According to cell types and culture conditions, the effect of divalents and monovalents differed.

(Productivity of dimerization)

25 Table 9

Refolding	[B3(Fab)-h(H123-Fc)-PE38] ₂	[B3(Fab)-h(H123-CH3)-PE38] ₂	[B3(Fab)-h(H123-CH2)-PE38] ₂
1st (240mg/3L)	148g 0.06%	455g 0.19%	198g 0.08%

2nd (80mg/1L)	44g 0.05%	146g 0.18%	78g 0.10%
3rd 80mg/1L)	47g 0.06%	162g 0.20%	59g 0.07%

In the case of [B3(Fab)-Ext(15CL14FA13)-PE38]₂, which has extension peptide chain SKPSISTKASG4C(G4S)2GGPE between Fab and PE38, the Ext(15CL14FA13) has decreased the steric hindrance between PE molecules, and
5 productivity of this molecule was 0.06%.

In this experiment, dimerization was induced using CH3 domain affinity and the productivity was 0.19%. This is three time higher compared to [B3(Fab)-Ext(15CL14FA13)-PE38]₂. The productivity using CH2 domain was 0.08%, and using Fc domain was 0.06%. These two results are because of the difficulties in
10 refolding of self affinity region and weak affinity of CH2. These are similar result to that of [B3(Fab)-Ext(15CL14FA13)-PE38]₂, and it means that these two cases didn't help the dimer production. However the manufacturers concerned can predict they will get good productivity if they use smaller size and stronger affinity domain. Especially, in the case of [B3(Fab)-h(H123-Fc)-PE38R]₂ which
15 has CH3 domain, productivity was 0.06% because the size was too big for the proper folding.

According to the result above, putting self-affinity domain in LFA induces assembly between uncoupled Cysteine, and the disulfide bond formation even with three Cysteines can take place normally without disturbing the sterical
20 conformation of binding domain and functional group domain. And also, this triple disulfide bonded dimer is more stable at room temperature when it is left alone and more resistant to damage by proteinase than single disulfide bonded dimer.

25 **Example 5: Dimerization of [B3(Fab)-cytosine deaminase]₂ which has Fab of B3 antibody as a binding domain and cytosine deaminase as**

functional group.

Up to now, monoclonal antibodies, monoclonal antibody treatments, toxins and radionuclide complexes and others have been studied for many years.

5 In recent years, many treatments based on monoclonal antibodies were clinically approved (Rituxan, Herceptin, Panorex). And some other treatments are in clinical demonstration.

To use big molecules as successful therapeutic agent for solid tumor, transmission hindrance into tumor mass, foreign antigen, efficacy of complex
10 treatment and efficient transfer from monoclonal antibody into cancer cell and others must be overcome.

Alternative plan for therapeutic agent delivery is to undergo two steps for cancer treatment. Once the complex binds to cancer cell and be removed from blood circulation, an enzyme activates anti-cancer prodrug to become activated
15 drug. The isolated drug can be transmitted into tumor mass and eliminates both monoclonal antibody binding cells and non-binding neighboring cells.

These treatments are called Antibody-Directed Enzyme Prodrug Therapy(ADEPT) which is a new method for selective cancer treatment. It introduces enzyme-antibody complex to change harmless prodrug at cancer cell
20 surface into cytotoxic compound selectively. Bagshawe first reported this idea at 1987.

The first step of ADEPT is that antibody-enzyme complexes accumulate on cancer cells. As time passes, the complex gets degraded in blood and normal tissues. The second step is to introduce harmless prodrug and it changes into
25 cytotoxic drug by enzymes from complexes.

Through this method, the cancer specificity increases and can deliver higher dose of drugs than direct administration. Also, one molecule of enzyme can

amplify cytotoxicity by catalyzing many prodrugs into activated drugs and activated drugs have small molecular weight to be easily diffused near to cancer cells. As a bystander effect, a unique character of ADEPT, the cells near the cancer cells also get killed.

5 There are various monoclonal antibodies, enzymes and prodrugs that can be applied to ADEPT. Among enzymes there is yeast cytosine deaminase which catalyze exchange of cytosine into uracil and it can change anti-mold reagent 5-Fluorocytosine(5-FC) into 5-Fluorouracil(5-FU) as an anti cancer agent. Especially the yeast cytosine deaminase is reported to have more remedial value
10 than bacterial enzymes used in enzyme-prodrug treatment. In addition, yeast cytosine deaminase prevents prodrug activity by restriction enzymes in blood or cells because it's from non-mammal, which has no homology to mammals. And it can be used easily with large amount because it has no translational modifications on it.

15 Pharmaceutical efficacy and toxicity must be clinically demonstrated before use and this is the dormant problem of applying ADEPT. However, 5-FU, which is used present clinically, can be used at colon carcinoma, for which it is hard to use other chemical agents. And cytotoxicity of drug and prodrug was confirmed at H2981 human lung adenocarcinoma cells. It is proved that there is
20 no cytotoxicity in 5-FC having no effect at 200uM concentration compared to 5-FU which has $ID_{50}=20\mu M$.

 On this experiment, recombinant protein molecule [B3(Fab)-cytosine deaminase]₂ for ADEPT application was constructed. This molecule was made by modification of Fab to be a divalent molecule and keeps the structural stability
25 and the turnover rate of Fab. The second disulfide bond in the hinge was used for dimerizing monovalent Fab-enzyme to divalent molecule. Since this molecule is divalent, the binding affinity will be more than 2 times stronger and since the

enzyme quantity has increased two times, the cytotoxicity will increase more. Also, although the IgG didn't have complete free rotation ability due to the three disulfide bonds of Fab region on the hinge, [B3(Fab)-enzyme]₂ has free rotation ability on its binding domain, and it will lead to stronger binding to antigens spread over the cells.

Though [B3(Fab)-cytosine deaminase]₂ is a big molecule, it has strong equilibrium binding affinity, fast binding reaction rate, increment of enzyme administration and long activity turn over rate in blood circulation system, and it will show higher efficacy than Fv derivatives.

10

(Apparatus and methods)

The same method was used as example 1. For the construction of plasmid containing B3(Fd)-yCD gene, chromosomal DNA from *Saccharomyces cerevisiae* was used for template and pMC74 was used as template for Fab region. For the light chain, pMCH75, which is 6x His tagged on 3'-end was used as a template. Mediums, reagents, enzymes and columns were used the same as example 1. Constructed plasmids are shown on table 10.

Table 10

Plasmid	Protein Sequence
pKL1	H6-CDase
pKL2	B3(Fd-SKPCIST-KAS-(G4S)2-GGPE-CDase-H6
pKL3	B3(Fd-SKPCIST-KAS-(G4S)2-GGPE-CDase
pKL4	H6-B3(Fd-AKPCIAT-QAS-(G4S)2-GGPE-CDase

20 (Construction of plasmids)

Plasmid which expresses B3(Fd)-yCD fusion protein was constructed from PCR using CDase containing region of *Saccharomyces cerevisiae* chromosomal DNA as a template and two primers below. The CDase was purified

and recombination with pMC74 led to recombinant plasmid pKL1.

Primer 1: 5'-GGC-CCA-TAT-GCA-TCA-CCA-TCA-CCA-TCA-CGT-
GAC-AGG-GGG-AAT-G-3'

Primer 2: 5'-TTG-GTT-TGA-AGA-TAT-TGG-TGA-GTA-GGA-ATT-
5 CGG-CC-3'

Primer 1 binds to 5'-end of CDase. Primer 2 binds to 3'-end of CDase to be used in CDase purification by PCR.

From PCR using pKL1 as a template with two primers below, CDase fragment was gained and it was recombined with PE of pMC22 to produce pKL2.

10 Primer 3: 5'-GGC-CCC-CGA-GGT-GAC-AGG-GGG-AAT-G-3'

Primer 4: 5'-GAA-GAT-ATT-GGT-GAG-CAT-CAC-CAT-CAC-CAT-
CAC-TAG-GAA-TTC-GGC-C-3'

Primer 3 binds to 5' end of CDase. And primer 4 having six Histidines binds to 3' end of CDase to be used in CDase purification by PCR.

15 pKL2 has B3(Fd)-SKPCISTKAS-GGGSGGGGS-GGPE-CDase-6His.

CDase region from pKL1 and CDase region from pKL2 was recombined to produce pKL3. pKL3 is B3(Fd)-SKPCISTKAS-GGGSGGGGS-GGPE-CDase and doesn't have six His at the 3' end.

pKL4 was constructed as following. PCR was performed by using
20 template pMC74 and two primers below to gain Fd fragment which has six His on 5'-end and recombined with pKL3 to produce pKL4.

Primer 5: 5'-GGC-CCA-TAT-GCA-TCA-CCA-TCA-CCA-TCA-CGA-
TGT-GAA-GCT-GGT-GGA-GTC-T-3'

Primer 6: 5'-GGG-AAT-TCA-TTA-AGC-TTG-TGT-AGC-TAT-GCA-
25 AGG-CTT-AGC-ACC-ACA-3'

Primer 5 has six His and binds to Fd region. Primer 6 has AKPCIATQAS and binds to 3' end of Fd.

The nucleotide sequence of plasmid pKL4 was confirmed by ALFexpress Dedeoxy Sequencing Kit(Amersham Pharmacia).

The method for protein expression and isolation, protein refolding, purification of refolded antibody-toxin was the same as example 1.

5 Cytotoxicity assay was performed on 4 kinds of cell lines with [B3(Fab)-CDase]₂ isolated by Superdex200 and the same method was used as example 1.

Through preceding methods, molecules purified had normal binding activity and functional group. And the binding activity of binding domain and cytosine activity of functional domain of the purified molecule was confirmed to
10 be normal.

Example 6: The dimerization and increase of the length of flexible chain(LFA) with the uncoupled Cysteine fixed at the 4th position on extension peptide chain(Ext).

15

The inventors increased the length of peptide linker(L) which is from uncoupled Cys on Ext to functional group to find out the relations with dimerization. The peptide linker(L) must be long enough to give spaces for the refolding to occur without disturbance from the sterical collision between two
20 functional groups when dimerization occurs between two monovalent binding domain-functional domain fusion. Also, the peptide linker must have good flexibilities being composed of flexible amino acids to allow the functional groups to get separated widely enough in an appropriate angle. But if the peptide linker is too long, the uncoupled Cysteine on Ext will have too much freedom to move and
25 meet the other naturally coupled Cysteines of intra-chain and inter-chain disulfide bond to get scrambled leading to inactivation of the molecule. And also since the peptide linker does not have a definite structure, it might penetrate into structures

of binding domain or functional group while refolding to disturb their correct refolding.

To find out the relations between the length of the peptide linker and dimerization, plasmids were constructed as below.

- 5 pMH28: B3(Fd)-SKPCG-----PE38; Ext(4CL1FA1)
- pMH29: B3(Fd)-SKPCKASPE-----PE38; Ext(4CL5FA3)
- pMH30: B3(Fd)-SKPCISTKASGGPE-----PE38; Ext(4CL10FA6)
- pMH34: B3(Fd)-SKPCISTKAS(GGGGS)4GGPE--PE38; Ext(4CL30FA26)
- pMH35: B3(Fd)-SKPCISTKAS(GGGGS)5GGPE--PE38; Ext(4CL35FA31)
- 10 pMH36: B3(Fd)-SKPCISTKAS(GGGGS)6GGPE--PE38; Ext(4CL40FA36)
- pMH37: B3(Fd)-SKPCISTKAS(GGGGS)7GGPE--PE38; Ext(4CL45FA41)
- pMH38: B3(Fd)-SKPCISTKAS(GGGGS)8GGPE--PE38; Ext(4CL50FA46)

The protein chains above have Cys on LFA on 4th position of Ext and the
 15 peptide linker(L) has 1,5,10,30,35,40,45,50 amino acids each containing
 1,3,6,26,31,36,41,46 non-bulky flexible amino acids.

The eight molecules of [B3(Fab)-Ext(4CLxxFAxx)-PE38]₂ type in this
 example were made to have the same possibilities of steric hindrance and
 intermixing that can be caused by the uncoupled Cys by fixing the Cys at 4th
 20 position on the Ext chain. And the amino acids following the uncoupled Cys were
 kept the same except the numbers of flexible amino acid was increased in five
 amino acids step to increase the peptide linker length and the sterical space was
 expanded with the same step. Among these molecules, the molecule with the
 shortest L having 1 amino acids which is Ext(4CL1FA1) and the longest L having
 25 50 amino acids which is Ext(4CL50FA46) showed very low productivity.

Therefore, this example shows that when the peptide linker(L) has over
 50 amino acids, it is impossible to get dimer made through the disulfide bond

bridge between the two fusion monomers of 50kD Fab domain of B3 antibody and 38kD functional group PE38.

(Materials and methods)

5 The same materials and methods were used as example 1.

(Construction of plasmids)

 The expression plasmids were constructed by PCR using primers which are appropriate for each DNA sequence. Used templates were same to example 1
10 and the same plasmids were used as example 1. The procedure for inserting GGGGS sequence repeatedly was done by restriction and ligation.

(Protein expression and isolation of inclusion body)

 Same methods were used as example 1.

15

(Refolding procedure and isolation of proteins)

 Same methods were used as example 1.

(Construction of expression vector and isolation of inclusion body)

20 The coding nucleotide sequences obtained through PCR and cloned in expression plasmids was confirmed by sequence analysis.

 The proteins in the form of inclusion body was analyzed by densitometry analysis(TINA2.0) on PAGE sample and the purity of protein chain was in 25~30%.

25

(Refolding of [B3(Fab)-Ext(4CLFA5X)-PE38]₂ molecule)

Table 11

Plasmid name	Structure of Ext	Yield of dimer (%)	Position of Cys in Ext	Length of L(distance bet. Cys and F)	Number of GASQEND in L	Note
pMH28	Ext(4CL1FA1)	<<0.01	4	1	1	
pMH29	Ext(4CL5FA3)	0.01	4	5	3	
pMH30	Ext(4CL10FA6)	0.04	4	10	6	
pCE1	Ext(1CL13FA7)	0.016	1	13	7	Example3
pCW1	Ext(15CL14FA13)	0.06	15	14	13	Example2
pMH21	Ext(4CL15FA11)	0.18	4	15	11	Example1
pMH22	Ext(4CL20FA16)	0.23	4	20	16	Example1
pMH23	Ext(4CL25FA21)	0.25	4	25	21	Example1
pMHS22	Ext(AQ4CL20FA16)	0.24	4	20	17	Example1
pMH34	Ext(4CL30FA26)	0.32	4	30	26	
pMH35	Ext(4CL35FA31)	0.17	4	35	31	
pMH36	Ext(4CL40FA36)	0.21	4	40	36	
pMH37	Ext(4CL45FA41)	0.08	4	45	41	
pMH38	Ext(4CL50FA46)	<<0.01	4	50	46	

According to table 11 above, molecules with very long peptide linker and very short ones have very small production yield and it is possible to detect them but not enough for testing. It shows that the length of the peptide linker do not have certain relationship with production yield of dimer and that it is possible to obtain dimer in some range of the peptide linker length. It can be assumed that the big sized Fab and PE38 may not get disturbed in forming its dimeric structural conformation during refolding even if the length of the peptide linker is long, but this example shows that if the amino acids on peptide linker is increased more than 50, the dimers can be detected but it is not practical to prepare them.

This means that using more than 50 amino acids on peptide linker is not realistic. If the protein production technology develops, the purification of 50

amino acid peptide linker dimer may be possible.

(Purification of B3(Fab)-Ext(4CLFA5X)-PE38)₂)

It was analyzed the same as example 1 and the purity was confirmed on
5 PAGE.

(Cytotoxicity assay)

Cytotoxicity assay was performed in triplicate. The result was that some
molecules had same ID₅₀ value as monovalent molecule and some had higher
10 cytotoxicity effect than monovalent molecules like in the previous example. This
is because of the density and the structural conditions of antigen on cell surface
according to the cell type.

**Example 7: Dimerization when the uncoupled Cysteine is on the 25th,
15 35th, 45th position on the extension chain(Ext)**

The relationship of uncoupled Cys position at 25th, 35th, 45th position on
Ext and dimerization was observed. The uncoupled Cysteine on Ext between
binding domain and functional domain can react and be oxidized to form disulfide
20 bond with the naturally coupled Cys of intra-chain and inter-chain disulfide bond
of binding domain and functional domain and can cause scrambling of the
disulfide bond and ruining the structures of domains to inactivate them. The
position of uncoupled Cys is very important factor that plays a role in scrambling
the disulfide bond with the naturally coupled Cysteine of binding domain and
25 functional group. Depending on the uncoupled Cysteine position, the disulfide
bond scrambling can happen.

Plasmids below were constructed to find out effects on dimerization when

uncoupled Cysteines are in different position from previous examples.

pMH42;B3(Fd)-SKPSISTKAS(GGGGS)2GGGGC(GGGGS)3GGPE-
PE38; Ext(25CL19FA18)

pMH44;B3(Fd)-SKPSISTKAS(GGGGS)4GGGGC(GGGGS)3GGPE-
5 PE38; Ext(35CL19FA18)

pMH46;B3(Fd)-SKPSISTKAS(GGGGS)6GGGGC(GGGGS)3GGPE-
PE38; Ext(45CL19FA18)

The protein chains above have fixed L19FA18 as LFA sequence on
10 Ext(LFA) and the position of Cys is changed to 25th, 35th, 45th. Therefore, it shows
whether there are any limits on the position of Cysteine compared to dimers
formed previously. When the uncoupled Cys was on the 45th position, detection of
dimers was possible but not enough to purify. Accordingly, this example shows
that dimerization is impossible if the uncoupled Cysteine is on over the 45th
15 position.

(Apparatus and method)

The same apparatus and method was used as example 1.

20 (Materials and Methods)

The same materials and methods were used as example 1.

(Construction of plasmids)

Protein expression plasmids were constructed by using primers and PCR.
25 Plasmids for templates used were the same as in example 2 and example 6. For
the repeated GGGGS sequence insertion, restriction and ligation was performed
using restriction enzymes and ligase.

(Protein expression and isolation of inclusion body)

The same method was used as example 1.

(Refolding procedure and isolation of protein)

5 The same method was used as example 1.

(Construction of expression vector and isolation of inclusion body)

The cloned nucleotide sequence into expression plasmids made by PCR was confirmed by sequence analysis.

10 The proteins in the form of inclusion body was analyzed by densitometry analysis(TINA2.0) on PAGE gel and the wanted protein chain was in 21~27% ratio.

(Refolding of [B3(Fab)-Ext(10XC L19FA18)-PE38]₂ L19FA18

15 The production yield of dimer was calculated by the analysis of the sample on PAGE after refolding procedure and ion exchange chromatography. The compared dimer production yield is shown on table 12.

Table 12

Plasmid name	Structure of Ext	Yield of dimer (%)	Position of Cys in Ext	Length of L(distance bet. Cys and F)	Number of GASQEND in L	Note
pCE1	Ext(1CL13FA7)	0.016	1	13	7	Example3
pMH21	Ext(4CL15FA11)	0.18	4	15	11	Example1
pMH22	Ext(4CL20FA16)	0.23	4	20	16	Example1
pCW1	Ext(15CL14FA13)	0.06	15	14	13	Example2
pMH42	Ext(25CL19FA18)	0.12	25	19	18	
pMH44	Ext(35CL19FA18)	0.04	35	19	18	
pMH46	Ext(45CL19FA18)	<<0.01	45	19	18	

According to table 12, when the Cysteine is at the 1st position (Ext(1CL13FA7)) the productivity is the lowest and when it is at the 4th position (Ext(4CL20FA16)), the productivity is the highest. When the Cys is positioned at the farthest (Ext(45CL19FA18)) position, the dimer can be detected but can't be
5 purified because the quantity is too small. If the protein production technology develops, it can be possible.

This example shows there is no definite relation between the position of Cys on extension chain with production yield, and dimerization is possible when the uncoupled Cys is in certain range of position. This means that Cys positioned
10 over 45th position is impractical for the production of dimers.

(Purification of [B3(Fab)-Ext(10XC L19FA18)-PE38]₂)

It was analyzed the same as example 1 and the different purity was confirmed with PAGE.

Claims

[1] A chimeric recombinant binding domain-heterogeneous functional group fusion wherein the chimeric recombinant binding domain-extension peptide chain-functional group fusion [B-Ext-F], which is characterized by

5 adding functional group(F) at the end of extension peptide chain(Ext) connected to binding domain(B) ;

 having uncoupled Cysteine on said extension peptide chain(extension chain) at any of position 1~45 for disulfide bonding between two monomers to produce dimer.

10

[2] A chimeric recombinant binding domain-heterogeneous functional group fusion wherein the chimeric recombinant binding domain-extension peptide chain-functional group fusion [B-Ext(L)-F], which is characterized by

 adding functional group(F) at the end of extension peptide chain(Ext)
15 connected to binding domain(B) ;

 having uncoupled Cysteine on said extension peptide chain(extension chain) at any of position 1~45 for disulfide bonding between two monomers to produce dimer ;

 having the peptide linker(L) that comprises the amino acid sequence
20 between the last Cysteine of said uncoupled Cysteines on extension chain and functional group(F)(the end of extension chain(Ext)) ;

 having 1~50 amino acids between the last Cysteine of said uncoupled Cysteines on extension chain and functional group(F).

25 [3] A chimeric recombinant binding domain-heterogeneous functional group fusion wherein the chimeric recombinant binding domain-extension peptide chain-functional group fusion [B-Ext(LAD)-F], which is characterized by

adding functional group(F) at the end of extension peptide chain(Ext)
connected to binding domain(B) ;

having uncoupled Cysteine on said extension peptide chain(extension
chain) at any of position 1~45 for disulfide bonding between two monomers to
5 produce dimer ;

having the peptide linker(L) that comprises the amino acid sequence
between the last Cysteine of said uncoupled Cysteines on extension chain and
functional group(F)(the end of extension chain(Ext)) ;

said peptide linker(L) is a peptide linker containing an affinity domain
10 (LAD) which has homomeric self affinity or heteromeric affinity making the
domain to assemble and leading to the formation of homomeric multimer or
heteromeric multimer domain ;

having 1~50 amino acids between the affinity domain(AD) in peptide
linker and functional group(F).

15

[4] A chimeric recombinant binding domain-heterogeneous functional group
fusion wherein the chimeric recombinant binding domain-extension peptide
chain-functional group fusion [B-Ext(LFA)-F], which is characterized by

adding functional group(F) at the end of extension peptide chain(Ext)
20 connected to binding domain(B) ;

having uncoupled Cysteine on said extension peptide chain(extension
chain) at any of position 1~45 for disulfide bonding between two monomers to
produce dimer ;

having the peptide linker(L) that comprises the amino acid sequence
25 between the last Cysteine of said uncoupled Cysteines on extension chain and
functional group(F)(the end of extension chain(Ext)) ;

said peptide linker(L) is a flexible amino acid sequence peptide

linker(LFA) which contains non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) ;

having 1~50 amino acids between the last Cysteine of said uncoupled
5 Cysteines on extension chain and functional group(F).

[5] A chimeric recombinant binding domain-heterogeneous functional group fusion wherein the chimeric recombinant binding domain-extension peptide chain-functional group fusion [B-Ext(LADFA)-F], which is characterized by

10 adding functional group(F) at the end of extension peptide chain(Ext) connected to binding domain(B) ;

having uncoupled Cysteine on said extension peptide chain(extension chain) at any of position 1~45 for disulfide bonding between two monomers to produce dimer ;

15 having the peptide linker(L) that comprises the amino acid sequence between the last Cysteine of said uncoupled Cysteines on extension chain and functional group(F)(the end of extension chain(Ext)) ;

said peptide linker(L) is a peptide linker containing an affinity domain (LAD) which has homomeric self affinity or heteromeric affinity making the
20 domain to assemble and leading to the formation of homomeric multimer or heteromeric multimer domain ;

said peptide linker(L) is also a flexible amino acid sequence peptide linker(LFA) which contains non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic
25 acid(D) ;

having 1~50 amino acids between the affinity domain(AD) in peptide linker and functional group(F).

[6] A chimeric recombinant binding domain-heterogeneous functional group fusion which has binding domain(B) composed of multiple chains(B1,B2,B3,,,Bn) wherein the chimeric recombinant binding domain-extension peptide chain-functional group fusion [B1-Ext or Ext(L) or Ext(LAD) or Ext(LADFA), B2-F, B3,,,Bn], which is characterized by

having one of the chain(B1) of binding domain connected to extension peptide chain(Ext) ;

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer{B1-Ext};

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain {B1-Ext(L)} ;

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain, and said peptide linker(L) has an affinity domain(AD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain{B1-Ext(LAD)} ;

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain, and said peptide linker(L) has an affinity domain(AD) which has homomeric self

affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain and said peptide linker(L) also comprises flexible amino acid sequence with non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D){B1-Ext(LADFA)} ;

having functional group(F) added to another chain(B2) which composes the binding domain{B2-F}.

[7] A chimeric recombinant binding domain-heterogeneous functional group fusion which has binding domain(B) composed of multiple chains(B1,B2,B3,,,Bn) wherein the chimeric recombinant binding domain-extension peptide chain-functional group fusion [B1-Ext or Ext(L) or Ext(LAD) or Ext(LADFA), B2-L-F, B3,,,Bn], which is characterized by

having one of the chain(B1) of binding domain connected to extension peptide chain(Ext) ;

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer{B1-Ext};

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain{B1-Ext(L)} ;

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain, and

said peptide linker(L) has an affinity domain(AD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain{B1-Ext(LAD)} ;

said extension peptide chain(extension chain) has uncoupled Cysteine at
 5 any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain, and said peptide linker(L) has an affinity domain(AD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or
 10 heteromeric multimer domain and said peptide linker(L) also comprises flexible amino acid sequence with non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D){B1-Ext(LADFA)} ;

having functional group(F) added to peptide linker(L) connected from the
 15 end of another chain(B2) which composes binding domain and said peptide linker(L) comprises 1~50 amino acids{B2-L-F}.

[8] A chimeric recombinant binding domain-heterogeneous functional group fusion which has binding domain(B) composed of multiple chains(B1,B2,B3,,,Bn)
 20 wherein the chimeric recombinant binding domain-extension peptide chain-functional group fusion[B1-Ext or Ext(L) or Ext(LAD) or Ext(LADFA), B2-LFA-F, B3,,,Bn], which is characterized by

having one of the chain(B1) of binding domain connected to extension peptide chain(Ext) ;

25 said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer{B1-Ext};

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain{B1-
5 Ext(L)} ;

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain, and
10 said peptide linker(L) has an affinity domain(AD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain{B1-Ext(LAD)} ;

said extension peptide chain(extension chain) has uncoupled Cysteine at any of position 1~45 for disulfide bonding between two monomers to produce
15 dimer and contains peptide linker(L) that comprises the amino acid sequence from the last Cysteine of said uncoupled Cysteines to the end of extension chain, and said peptide linker(L) has an affinity domain(AD) which has homomeric self affinity or heteromeric affinity leading to the formation of homomeric multimer or heteromeric multimer domain and said peptide linker(L) also comprises flexible
20 amino acid sequence with non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D){ B1-Ext(LADFA)} ;

having functional group(F) added to peptide linker(L) connected from the end of another chain(B2) which composes binding domain and said peptide
25 linker(L) is LFA comprising flexible amino acid sequence with non-bulky amino acids like Glycine(G) or Alanine(A) or Serine(S) or Glutamine(Q) or Glutamic acid(E) or Asparagine(N), Aspartic acid(D) and said peptide linker(L) comprises

1~50 amino acids {B2-LFA-F}.

[9] A dimer of chimeric recombinant binding domain-heterogeneous functional group fusion characterized in that the uncoupled Cysteine on the extension chain
5 (Ext) of chimeric recombinant binding domain-extension chain-functional group fusion (B-Ext-F) as defined in any one of claims 1 to 8 forms disulfide-bond-bridge to make dimers by oxidation with the uncoupled Cysteine on extension chain of binding domain-extension chain fusion (B-Ext) as defined in any one of claims 1 to 8 or with uncoupled Cysteine on the extension chain of chimeric
10 recombinant binding domain- extension chain-functional group fusion(B-Ext-F) as defined in any one of claims 1 to 8.

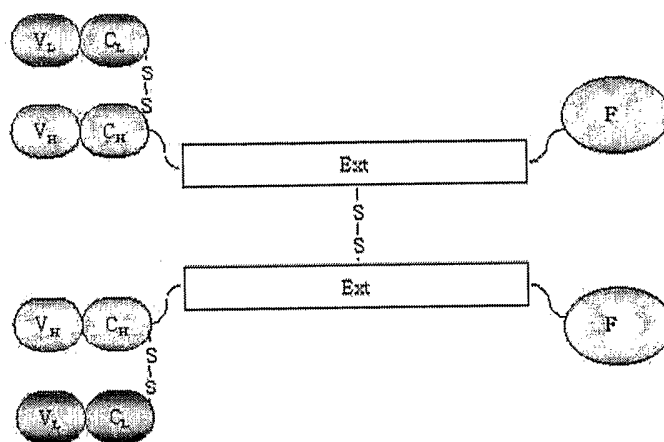
[10] The chimeric recombinant binding domain-heterogeneous functional group fusion according to any one of claims 1 to 8, characterized in that the order of
15 protein domain from amino terminal to carboxyl terminal is binding domain-extension chain-functional group(B-Ext-F).

[11] The chimeric recombinant binding domain-heterogeneous functional group fusion according to any one of claims 1 to 8, characterized in that the order of
20 protein domain from amino terminal to carboxyl terminal is functional group-extension chain-binding domain (F-Ext-B).

[12] The chimeric recombinant binding domain-heterogeneous functional group fusion according to any one of claims 1 to 8, characterized in that the binding
25 domain is amino acid sequence which compose antibody or fragments of antibody or fragments of antibody affinity.

[13] The chimeric recombinant binding domain-heterogeneous functional group fusion according to any one of claims 1 to 8, characterized in that binding domain is Fab fragment of antibody.

- 5 [14] The chimeric recombinant binding domain-heterogeneous functional group fusion according to any one of claims 1 to 8, characterized in that binding domain has structure as below,

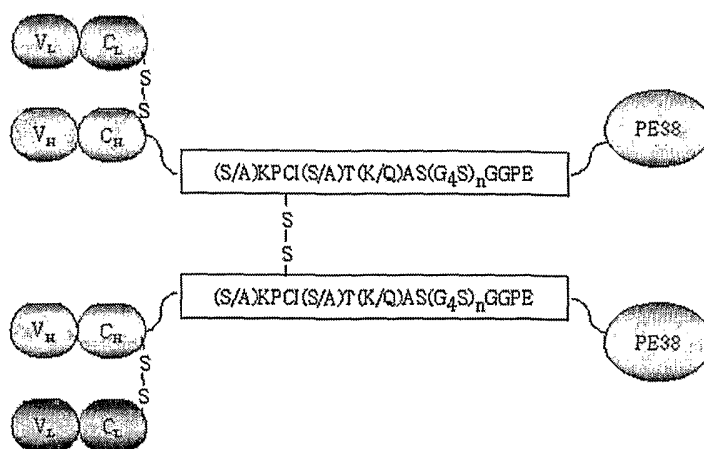


According to the structural formula of protein above,

- 10 V_L is variable region on light chain of monoclonal antibody;
 C_L is constant region on light chain of monoclonal antibody;
 V_H is variable region on heavy chain of monoclonal antibody;
 C_H is constant region 1 on heavy chain of monoclonal antibody.

- 15 [15] The chimeric recombinant binding domain-heterogeneous functional group fusion according to any one of claims 1 to 8, characterized in that binding domain has structure as below,

80



According to the structural formula of protein above,

V_L is variable region on light chain of monoclonal antibody;

C_L is constant region on light chain of monoclonal antibody;

5 V_H is variable region on heavy chain of monoclonal antibody;

C_H is constant region 1 on heavy chain of monoclonal antibody;

A is Alanine;

C is Cysteine;

E is Glutamic acid;

10 G is Glycine;

I is Isoleucine;

K is Lysine;

P is Proline;

Q is Glutamine;

15 S is Serine;

T is Threonine;

n is 1 or 2 or 3;

PE38 is 38kDa exotoxin derivative of *Pseudomonas aeruginosa* which

has been truncated of amino acid sequences of domain I and II that

20 are not required for cytotoxicity

[16] A polypeptide used in production of chimeric recombinant binding domain-heterogeneous functional group fusion as defined in any one of claims 1 to 8.

- 5 [17] A recombinant plasmid which codes polypeptide used in production of chimeric recombinant binding domain-heterogeneous functional group fusion as defined in any one of claims 1 to 8.

- [18] A host cell having recombinant plasmid which codes polypeptide used in
10 production of chimeric recombinant binding domain-heterogeneous functional group fusion as defined in any one of claims 1 to 8.

- [19] A pharmaceutical composition which contains chimeric recombinant binding domain-heterogeneous functional group fusion as defined in any one of claims 1
15 to 8.

[20] A method for producing chimeric recombinant binding domain-heterogeneous functional group fusion as defined in any one of claims 1 to 8.

1/13

Fig. 1

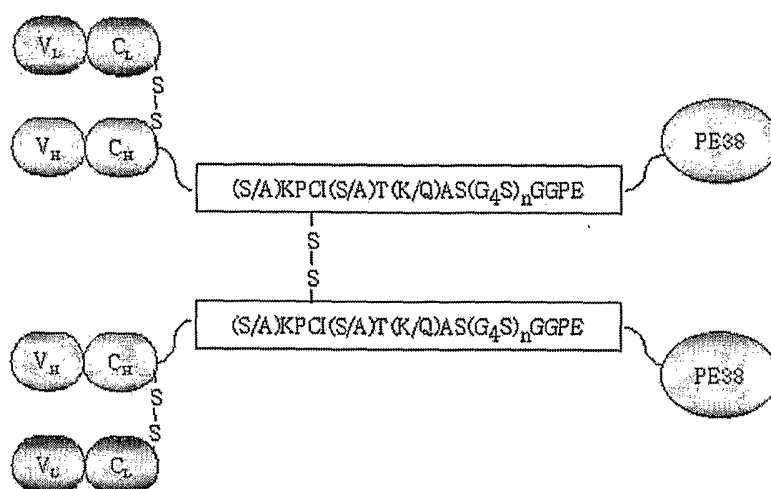


Fig. 2

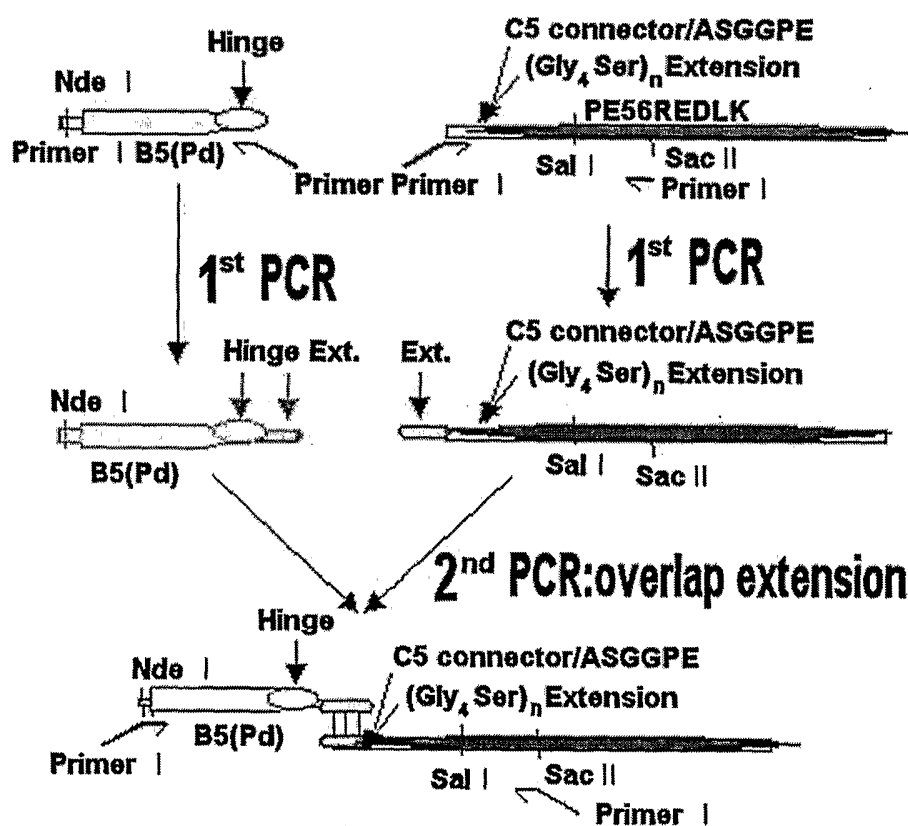


Fig. 3

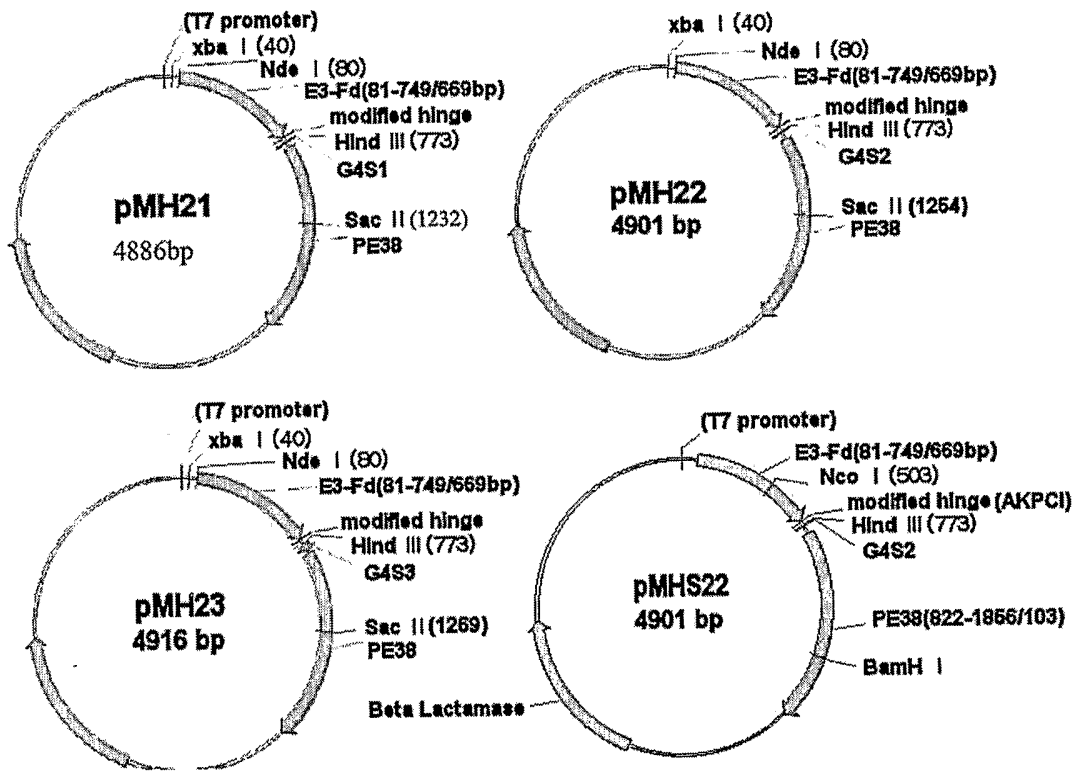
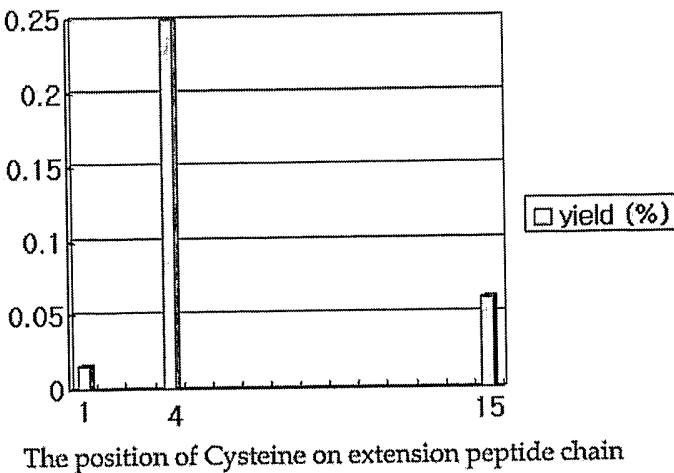


Fig. 4



3/13

Fig. 5

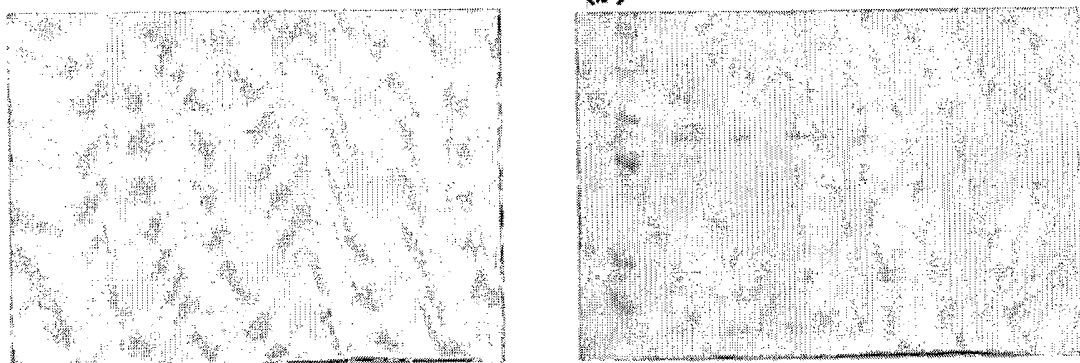


Fig. 6

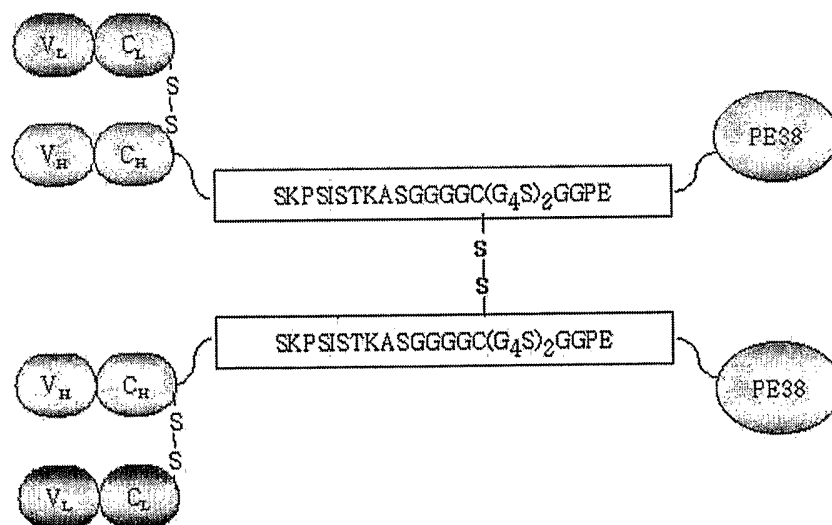


Fig 7

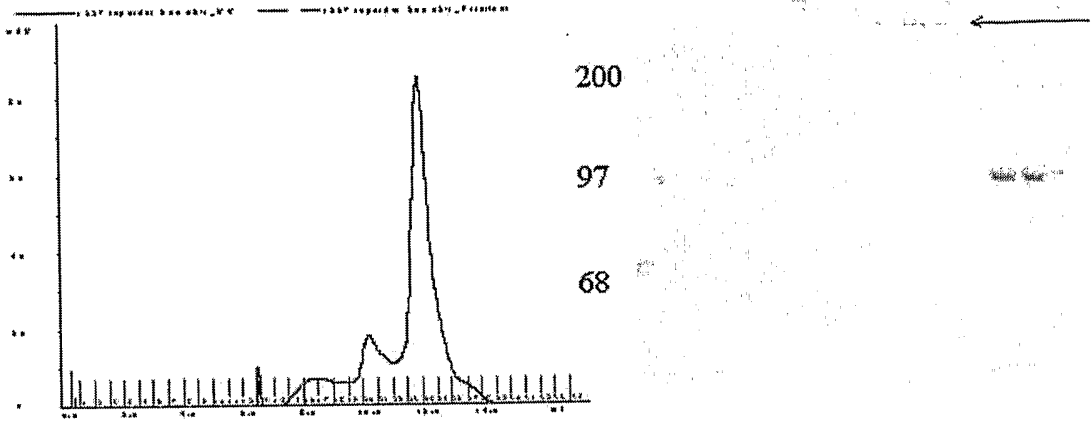


Fig. 8

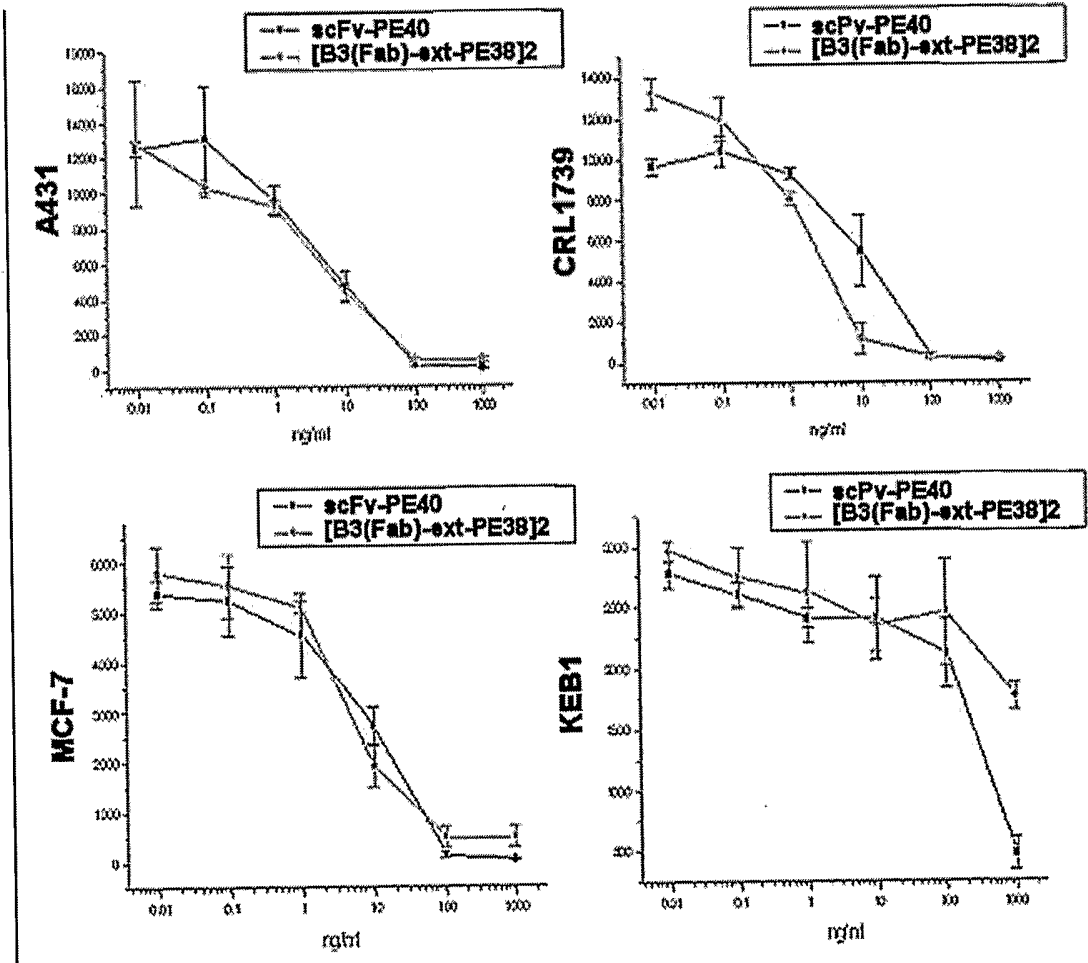


Fig. 9

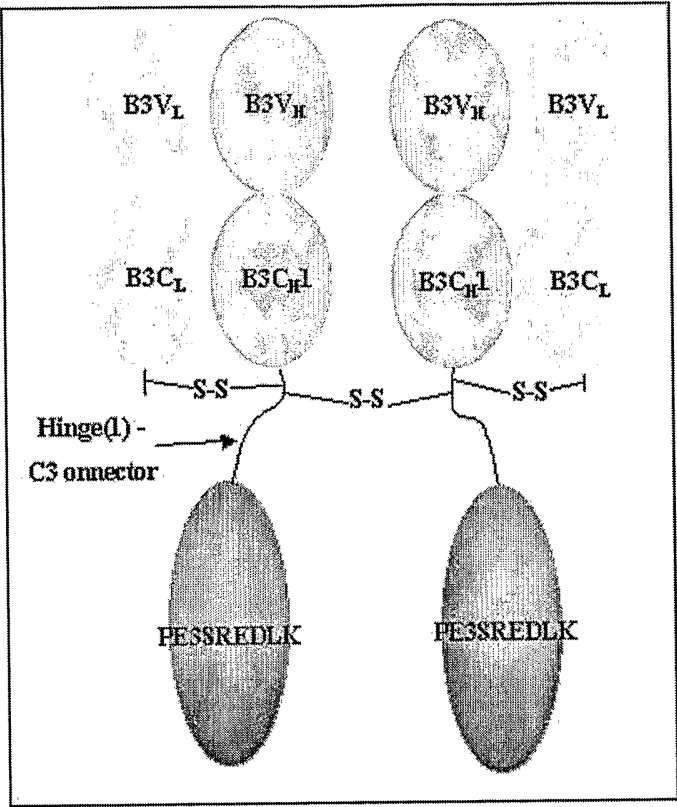
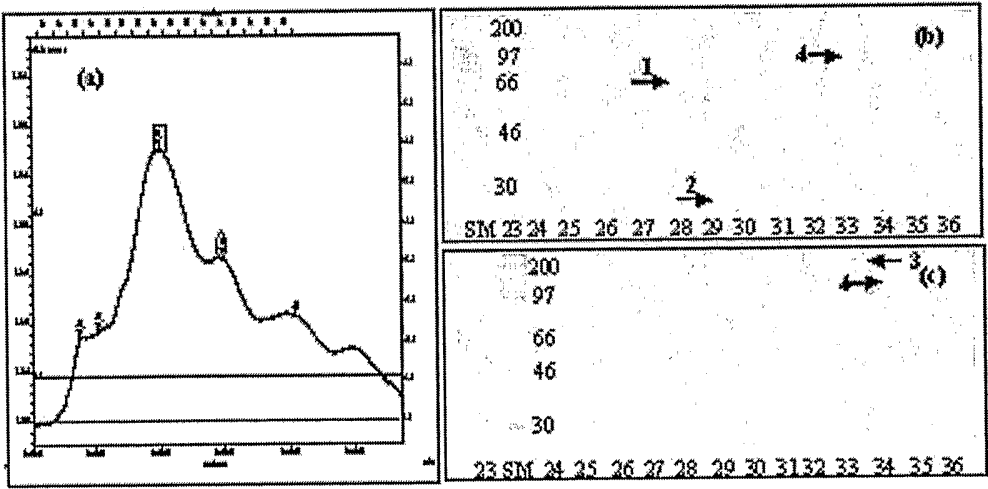
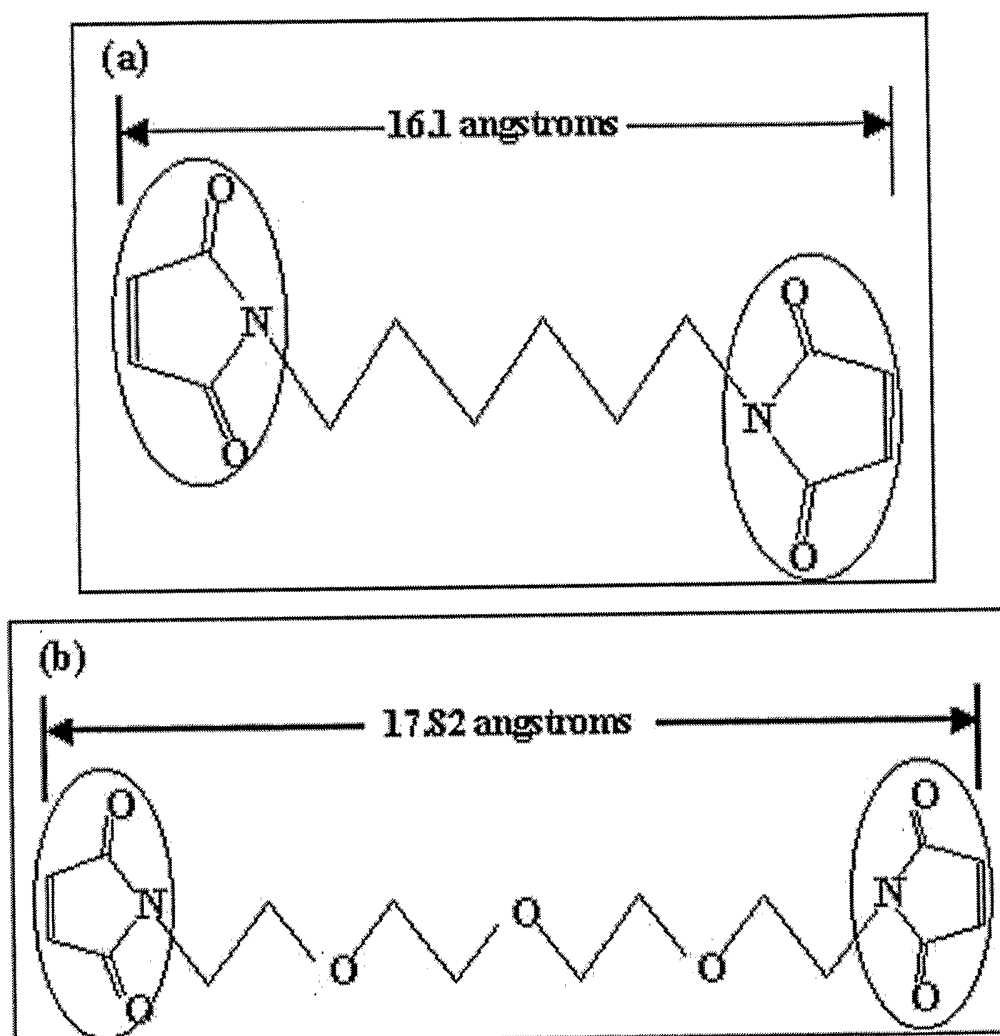


Fig. 10



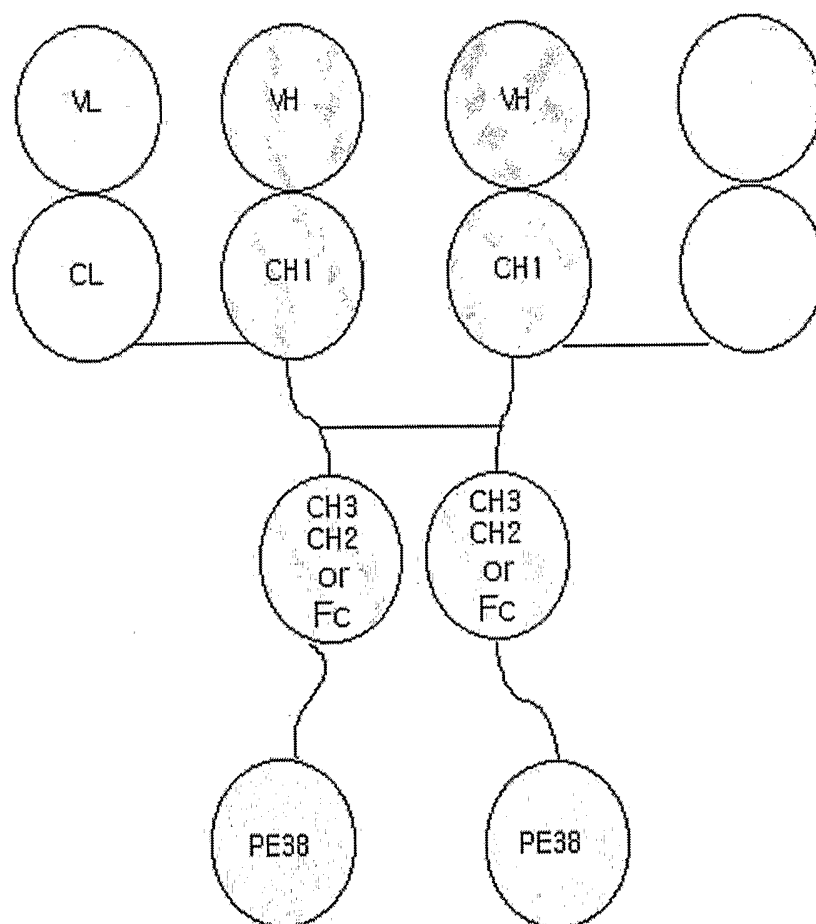
6/13

Fig. 11



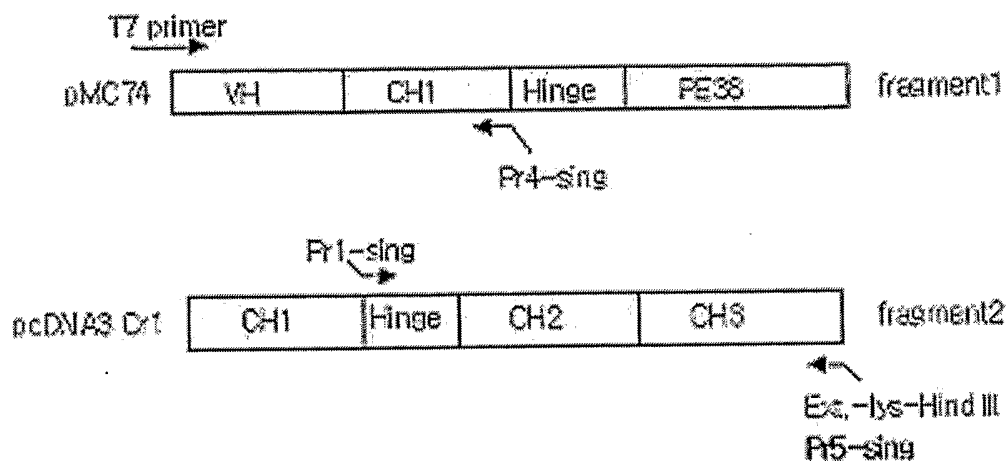
7/13

Fig. 12

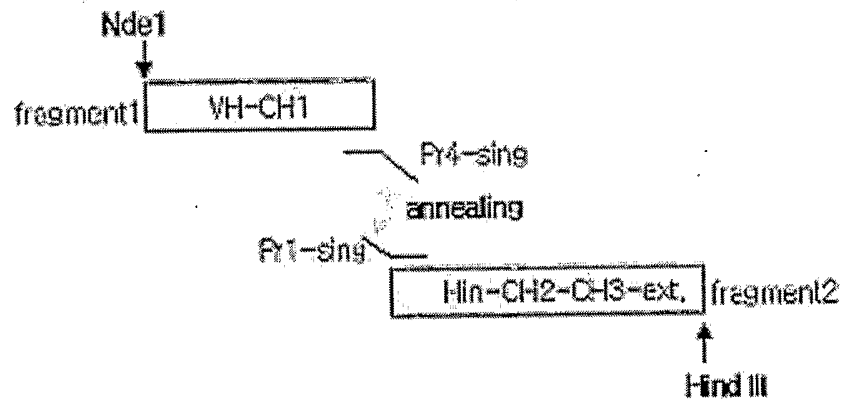


8/13

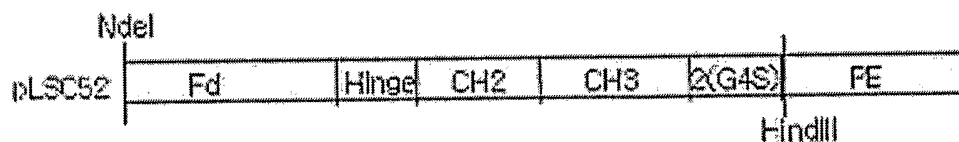
Fig. 13

(a) Fd, Fc fragment PCR from pMC74 and pcDNA3C_γ1

(b) Splicing PCR amplification by overlap extension

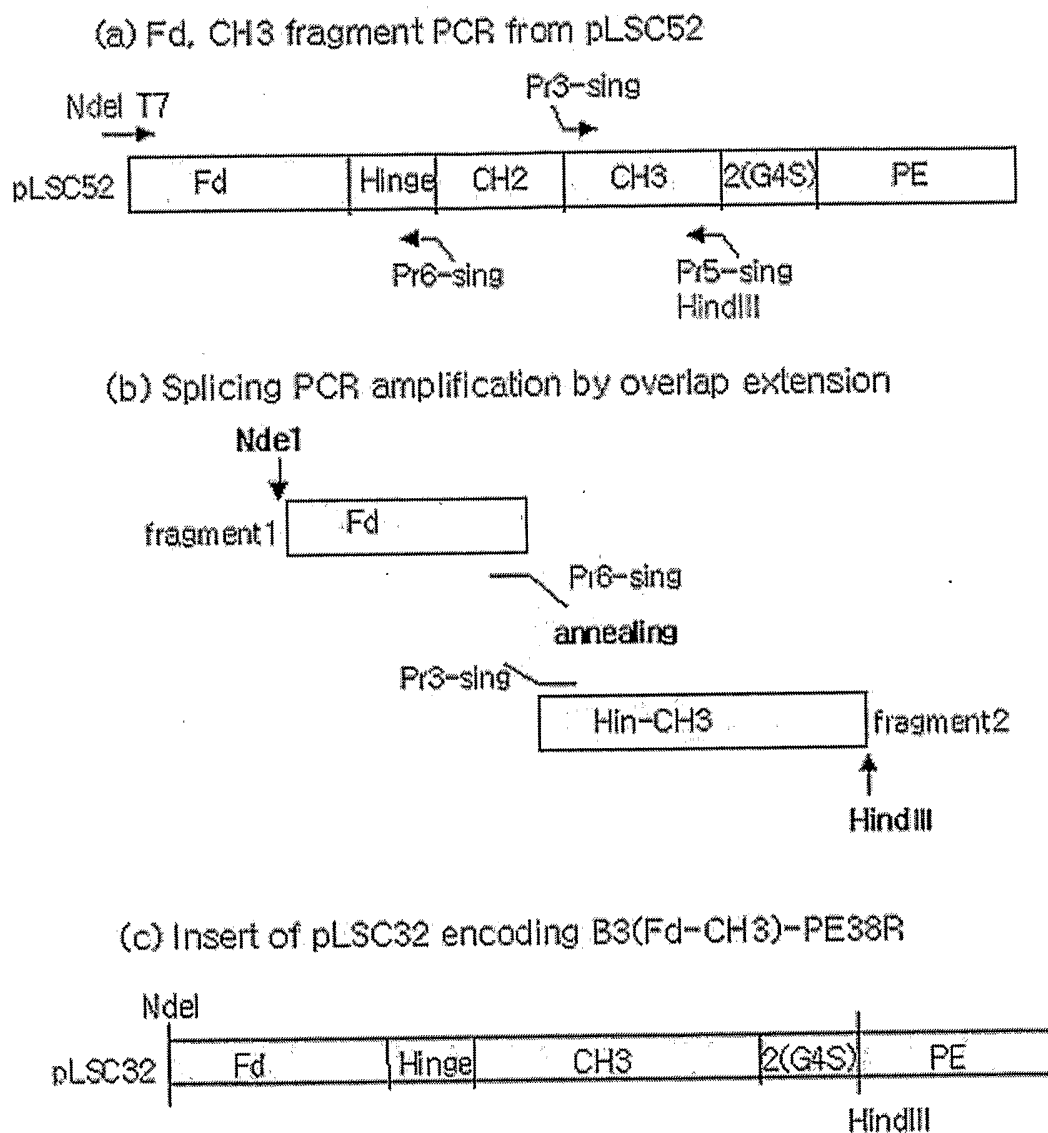


(c) Insert of pLSC52 encoding B3(Fd-Fc)-PE38R



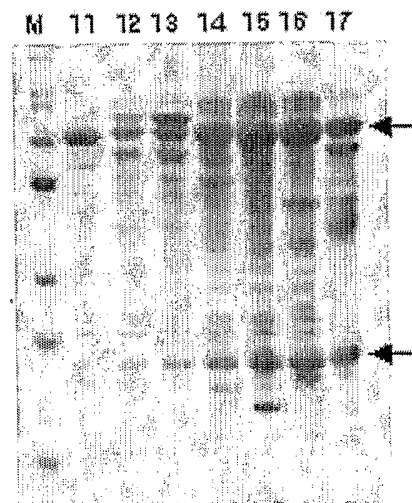
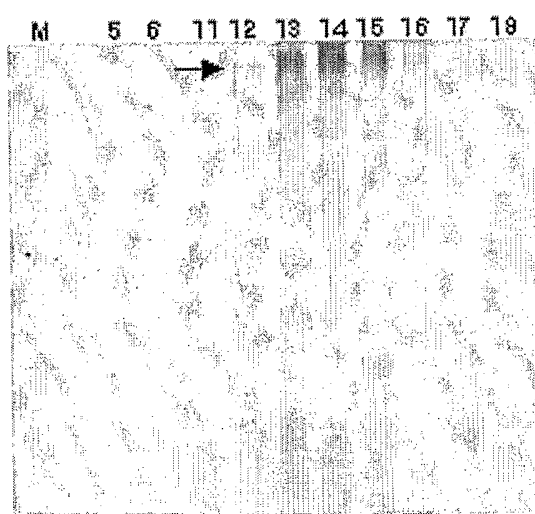
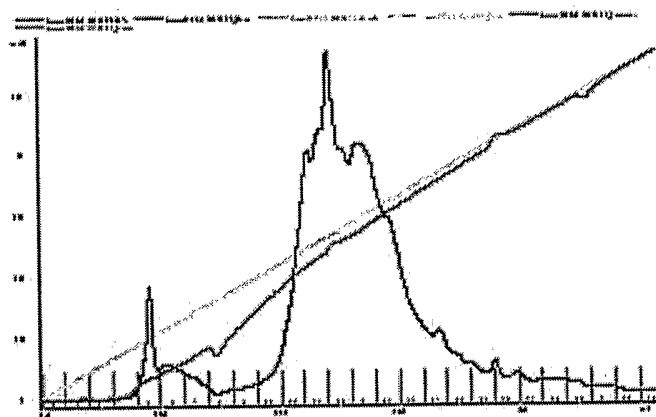
9/13

Fig. 14



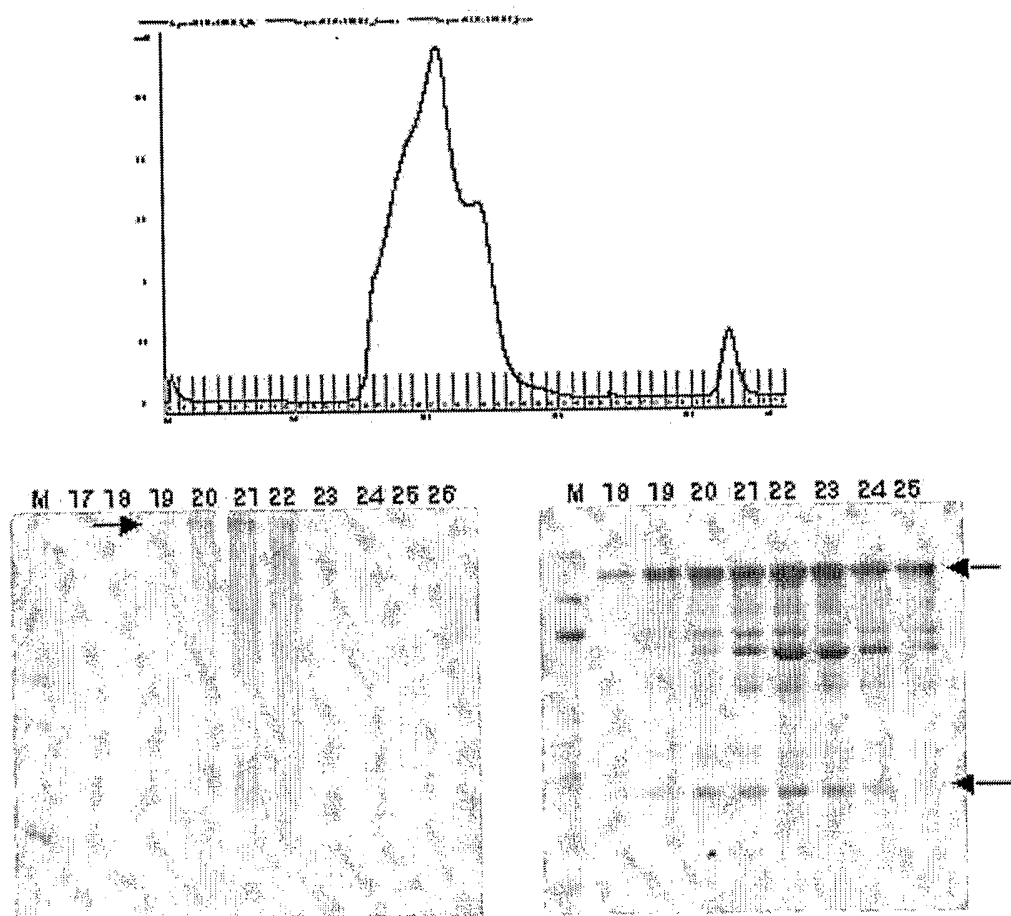
10/13

Fig. 15



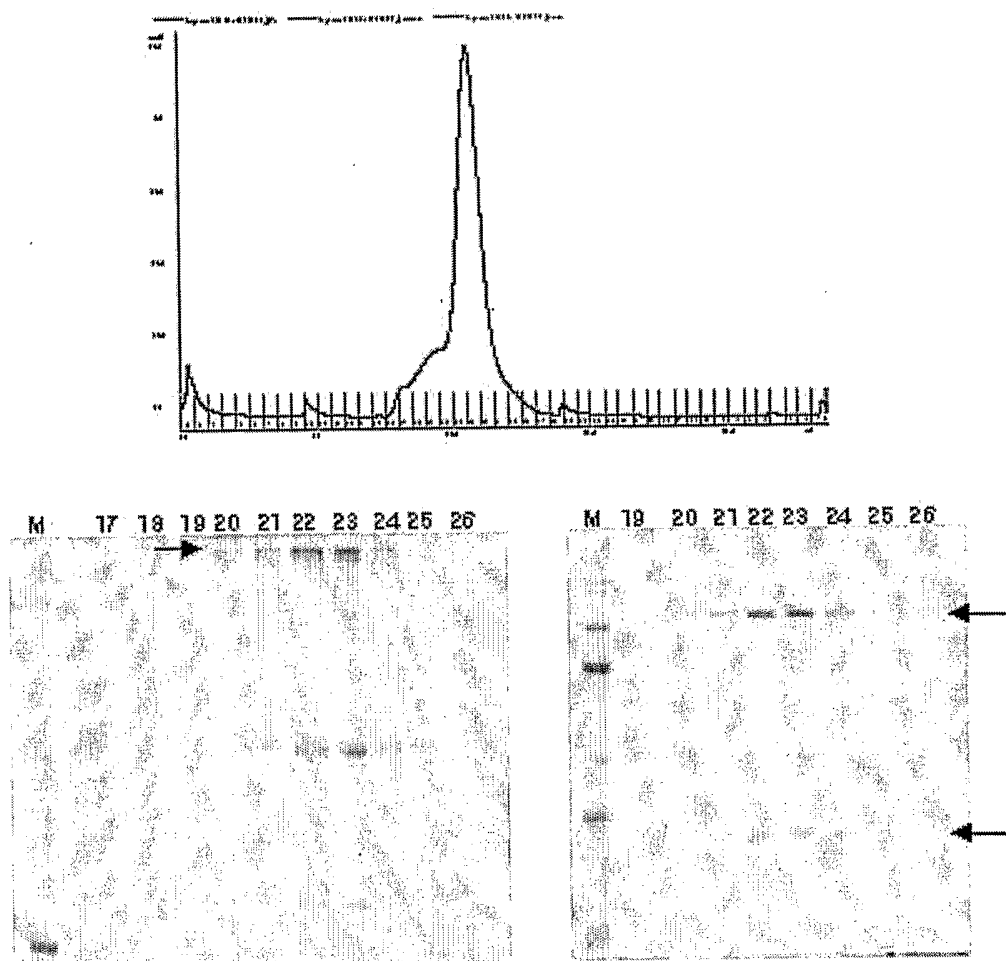
11/13

Fig. 16



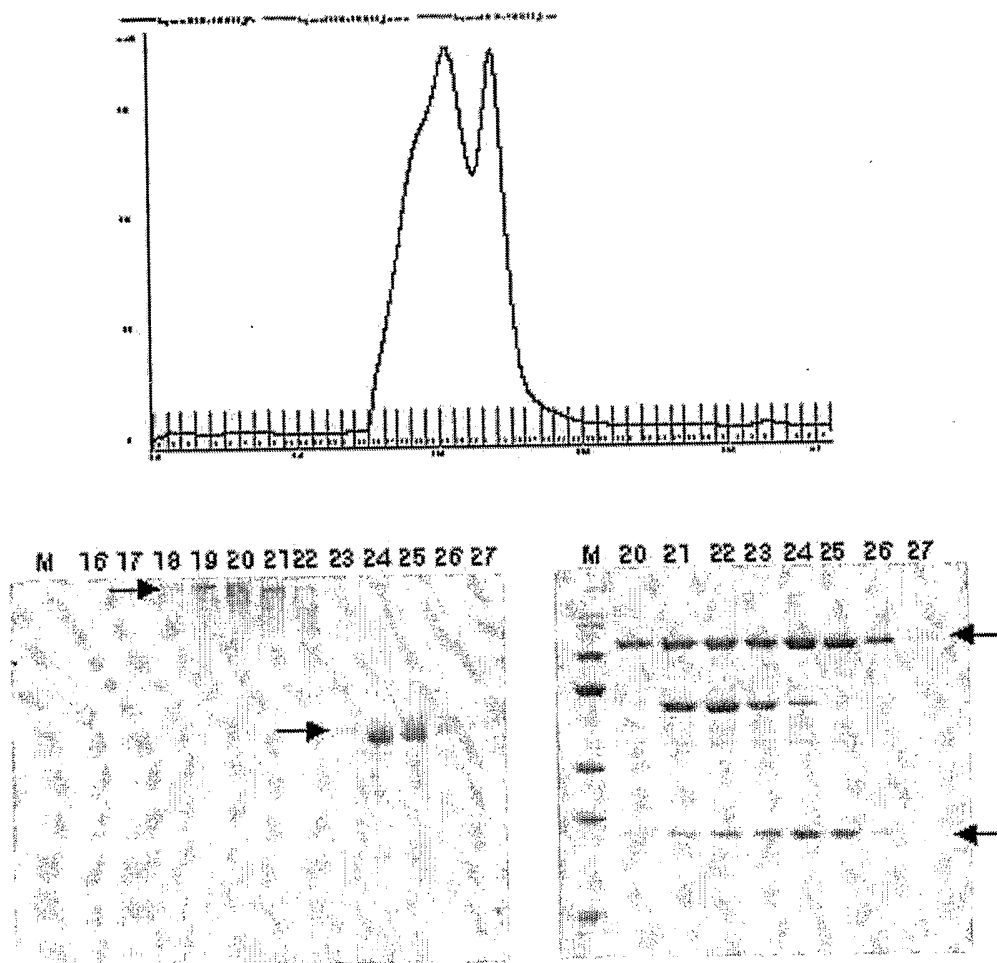
12/13

Fig. 17



13/13

Fig. 18



Sequence Listing

<110> CHOE, Mu-Hyeon

<120> THE DIMER OF CHIMERIC RECOMBINANT BINDING DOMAIN-FUNCTIONAL GROUP
FUSION FORMED VIA DISULFIDE-BOND-BRIDGE AND THE PROCESSES FOR
PRODUCING THE SAME

<130> YL04011PCT

<140> PCT/KR2004/001595

<141> 2004-06-30

<150> KR2003-0043599

<151> 2003-06-30

<160> 12

<170> KopatentIn 1.71

<210> 1

<211> 1749

<212> DNA

<213> Artificial Sequence

<220>

<223> pMC74 plasmid coding sequence

<400> 1

atggatgtga agctggtgga atctggagga ggcttagtgc agcctggagg gtccctgaaa	60
ctctcctgtg caacctctgg attcactttc agtgactatt acatgtattg ggttcgccag	120
actccagaga agaggctgga gtgggtcgca tacattagta atgatgatag ttccgccgct	180
tattcagaca ctgtaaaggg ccggttcacc atctccagag acaatgccag gaacaccctc	240
tacctgcaaa tgagccgtct gaagtctgag gacacagcca tatattcctg tgcaagagga	300
ctggcctggg gagcctgggt tgcttactgg ggccaaggga ctctggtcac tgtctctgca	360

Sequence Listing

gccaaaacga ccccccatc tgtctatcca ctggcccctg gatctgctgc ccaaactaac	420
tccatggtga ccctgggatg cctggtcaag ggctatttcc ctgagccagt gacagtgacc	480
tggaactctg gatccctgtc cagcgggtgtg cacaccttcc cagctgtcct gcagtctgac	540
ctctacactc tgagcagctc agtgactgtc ccctccagca cctggcccag cgagaccgtc	600
acctgcaacg ttgccacccc ggccagcagc accaagggtg acaagaaaat tgtgcccagg	660
gattgtggta gtaagcctag cataagtaca aaagcttccg gaggtcccga gggcggcagc	720
ctggccgcgc tgaccgcgca ccaggttgc cacctgccgc tggagacttt caccggtcat	780
cgcagccgc gcggtggga acaactggag cagtgcggct atccggtgca gcggtgtgtc	840
gccctctacc tggcggcgcg gctgtcgtgg aaccaggtcg accaggtgat ccgcaacgcc	900
ctggccagcc ccggcagcgg cggcgacctg ggcgaaagca tccgcgagca gccggagcag	960
gcccgtctgg ccctgaccct ggccgcgcc gagagcgagc gcttcgtccg gcagggcacc	1020
ggcaacgacg aggcgggcgc ggccaacggc cggcgcgaca gcggcgacgc cctgctggag	1080
cgcaactatc ccaactggcg ggagttctc ggcgacggcg gcgacgtcag cttcagcacc	1140
cgcggcacgc agaactggac ggtggagcgg ctgctccagg cgcaccgcca actggaggag	1200
cgcggctatg tgttcgtcgg ctaccacggc accttcctcg aagcggcgca aagcatcgtc	1260
ttcggcgggg tgcgcgcgcg cagccaggac ctcgacgca tctggcgcg tttctatatc	1320
gccggcgatc cggcgctggc ctacggctac gcccaggacc aggaaccga cgcacgcggc	1380
cggatccgca acggtgccct gctgcgggtc tatgtgccgc gctcgagcct gccgggcttc	1440
taccgcacca gcctgaccct ggccgcgcgc gaggcggcgg gcgaggtcga acggtgatc	1500
ggccatccgc tgccgctgcg cctggacgcc atcacgggcc ccgaggagga aggcgggcgc	1560

Sequence Listing

ctggagacca ttctcggctg gccgctggcc gagcgaccg tggtgattcc ctcgcgatc 1620

cccaccgacc cgcgcaacgt cggcggcgac ctcgaccgt ccagcatccc cgacaaggaa 1680

caggcgatca gcgcctgccc ggactacgcc agccagccc gcaaaccgcc gcgcgaggac 1740

ctgaagtaa 1749

<210> 2

<211> 1764

<212> DNA

<213> Artificial Sequence

<220>

<223> pMH21 plasmid coding sequence

<400> 2

atggaggatga agctgggtga atctggagga ggcttagtgc agcctggagg gtccctgaaa 60

ctctcctgtg caacctctgg attcactttc agtgactatt acatgtattg ggttcgccag 120

actccagaga agaggctgga gtgggtcgca tacattagta atgatgatag ttccgccgct 180

tattcagaca ctgtaaaggc cgggttcacc atctccagag acaatgccag gaacaccctc 240

tacctgcaaa tgagccgtct gaagtctgag gacacagcca tatattcctg tgcaagagga 300

ctggcctggg gagcctggtt tgcttactgg ggccaaggga ctctgggtcac tgtctctgca 360

gccaaaacga caccctcatc tgtctatcca ctggcccctg gatctgctgc ccaaactaac 420

tccatggtga ccctgggatg cctgggtcaag ggctatttcc ctgagccagt gacagtgacc 480

tggaactctg gatccctgtc cagcgggtgtg cacaccttcc cagctgtcct gcagtctgac 540

ctctacactc tgagcagctc agtgactgtc ccctccagca cctggcccag cgagaccgtc 600

acctgcaacg ttgccacccc ggccagcagc accaagggtgg acaagaaaat tgtgcccagg 660

Sequence Listing

gattgtggta gtaagccttg cataagtaca aaagcttctg gtgggtggcgg atctggaggt	720
cccgagggcg gcagcctggc cgcgctgacc gcgcaccagg cttgccacct gccgctggag	780
actttcaccg gtcacgcca gccgcgcggc tgggaacaac tggagcagtg cggtatccg	840
gtgcagcggc tggtcgccct ctacctggcg gcgcggctgt cgtggaacca ggtcgaccag	900
gtgatccgca acgccctggc cagccccggc agcggcggcg acctgggcga agcgatccgc	960
gagcagccgg agcaggcccg tctggccctg acctggccg ccgccgagag cgagcgcttc	1020
gtccggcagg gcaccggcaa cgacgaggcc ggcgcgcca acggcccggc ggacagcggc	1080
gagccctgc tggagcgcaa ctatcccact ggcgcgaggt tcctcggcga cggcggcgac	1140
gtcagcttca gcaccgcgg cagcagaac tggacggtgg agcggctgct ccaggcgcac	1200
cgccaactgg aggagcgcgg ctatgtgttc gtcggctacc acggcacctt cctcgaagcg	1260
gcgcaaagca tcgtcttcgg cggggtgcgc gcgcgcagcc aggacctcga cgcgatctgg	1320
cgcggtttct atatcgccgg cgatccggcg ctggcctacg gctacgcca ggaccaggaa	1380
cccgacgcac gcggccggat ccgcaacggt gccctgctgc gggcttatgt gccgcgctcg	1440
agcctgccgg gcttctaccg caccagcctg acctggccg cgccggaggc ggcgggcgag	1500
gtcgaacggc tgatcgcca tcgctgccg ctgcgcttg acgcatcac cggccccgag	1560
gaggaaggcg ggcgcctgga gaccattctc ggctggccgc tggccgagcg caccgtggtg	1620
attccctcgg cgatccccac cgaccgcgc aacgtcggcg gcgacctcga ccggtccagc	1680
atccccgaca aggaacaggc gatcagcgcc ctgccggact acgccagcca gcccgcaaa	1740
ccgcgcgcg aggacctgaa gtaa	1764

Sequence Listing

<210> 3
 <211> 1749
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pCE2 plasmid coding sequence

<400> 3
 atggatgtga agctggtgga atctggagga ggcttagtgc agcctggagg gtccctgaaa 60
 ctctcctgtg caacctctgg attcactttc agtgactatt acatgtattg ggttcgccag 120
 actccagaga agaggctgga gtgggtcgca tacattagta atgatgatag ttccgccgct 180
 tattcagaca ctgtaaaggg ccggttcacc atctccagag acaatgccag gaacaccctc 240
 tacctgcaaa tgagcgtctt gaagtctgag gacacagcca tatattcctg tgcaagagga 300
 ctggcctggg gagcctgggtt tgcttactgg ggccaaggga ctctggtcac tgtctctgca 360
 gccaaaacga caccgccatc tgtctatcca ctggcccctg gatctgctgc ccaaactaac 420
 tccatggtga ccctgggatg cctgggtcaag ggctatttcc ctgagccagt gacagtgacc 480
 tggaactctg gatccctgtc cagcgggtgtg cacaccttcc cagctgtcct gcagtctgac 540
 ctctacactc tgagcagctc agtgactgtc ccctccagca cctggcccag cgagaccgtc 600
 acctgcaacg ttgccacccc ggccagcagc accaagggtg acaagaaaat tgtgcccagg 660
 gattgtggtg gtaagccttg cataagtaca aaagcttccg gaggtcccga gggcggcagc 720
 ctggccgcgc tgaccgcgca ccaggcttgc cacctgccgc tggagacttt caccgtcat 780
 cgccagccgc gcggctggga acaactggag cagtgcggct atccggtgca gcggctggtc 840
 gccctctacc tggcggcgcg gctgtcgtgg aaccaggctg accagggtgat ccgcaacgcc 900

Sequence Listing

ctggccagcc ccggcagcgg cggcgacctg ggcgaagcga tccgcgagca gccggagcag	960
gcccgtcttg ccctgaccct ggccgccgcc gagagcgagc gcttcgtccg gcagggcacc	1020
ggcaacgacg aggccggcgc ggccaacggc ccggcggaca gcggcgacgc cctgctggag	1080
cgcaactatc ccaactggcg ggagttcctc ggcgacggcg gcgacgtcag cttcagcacc	1140
cgcggcacgc agaactggac ggtggagcgg ctgctccagg cgcaccgcca actggaggag	1200
cgcggctatg tgttcgtcgg ctaccacggc accttcctcg aagcggcgca aagcatogtc	1260
ttcggcgggg tgcgcgcgcg cagccaggac ctgcacgca tctggcgcgg ttctatatatc	1320
gccggcgatc cggcgctggc ctacggctac gccaggacc aggaaccgca cgcacgcggc	1380
cggatccgca acggtgccct gctgcgggtc tatgtgccgc gctcgagcct gccgggcttc	1440
taccgcacca gcctgaccct ggccgcgcgc gaggcggcgg gcgaggtcga acggtgatc	1500
ggccatccgc tgccgctgcg cctggacgcc atcaccggcc ccgaggagga aggcgggcgc	1560
ctggagacca ttctcggctg gccgctggcc gagcgacccg tggtgattcc ctccgcgatc	1620
cccaccgacc cgcgcaacgt cggcggcgac ctcgaccgt ccagcatccc cgacaaggaa	1680
caggcgatca gcgccctgcc ggactacgcc agccagcccg gcaaaccgcc gcgcgaggac	1740
ctgaagtaa	1749

<210> 4

<211> 672

<212> DNA

<213> Artificial Sequence

<220>

<223> pMC75 plasmid coding sequence

Sequence Listing

<400> 4

atggatgtgc tgatgaccca gtctccattg agtttacctg tcagtcttgg agatcaagcc	60
tccatctctt gcagatctag tcagatcatt gtacatagta atggaaacac ctatttagaa	120
tggtagctgc agaaaccagg ccagtctcca aagctcctga tctacaaagt ttccaaccga	180
ttttctgggg tcccagacag gttcagtggc agtggatcag ggacagattt cacactcaag	240
atcagcagag tggaggctga ggatctggga gtttattact gctttcaagg ttcacatgtt	300
ccattcacgt tcggctcggg gacaaagttg gaaataaaac gggctgatgc tgcaccaact	360
gtatccatct tcccaccatc cagtggagcag ttaacatctg gaggtgcctc agtcgtgtgc	420
ttcttgaaca acttctaccc caaagacatc aatgtcaagt ggaagattga tggcagtga	480
cgacaaaatg gcgtcctgaa cagttggact gatcaggaca gcaaagacag cacctacagc	540
atgagcagca ccctcacgtt gaccaaggac gagtatgaac gacataacag ctatacctgt	600
gaggccactc acaagacatc aacttcaccc attgtcaaga gcttcaacag gaatgagtgt	660
ggtaaagctt aa	672

<210> 5

<211> 2454

<212> DNA

<213> Artificial Sequence

<220>

<223> pLSC52 plasmid coding sequence

<400> 5

atggatgtga agctgggtga atctggagga ggcttagtgc agcctggagg gtccctgaaa	60
ctctcctgtg caacctctgg attcactttc agtgactatt acatgtattg ggttcgccag	120

Sequence Listing

actccagaga agagggctgga gtgggtcgca tacattagta atgatgatag ttccgccgct	180
tattcagaca ctgtaaaggg ccggttcacc atctccagag acaatgccag gaacaccctc	240
tacctgcaaa tgagccgtct gaagtctgag gacacagcca tatattcctg tgcaagagga	300
ctggcctggg gagcctggtt tgcttactgg ggccaaggga ctctgggtcac tgtctctgca	360
gccaaaacga ccccccatc tgtctatcca ctggccctg gatctgctgc ccaaactaac	420
tccatggtga ccctgggatg cctggtcaag ggctatttcc ctgagccagt gacagtgacc	480
tggaactctg gatccctgtc cagcgggtg cacaccttc cagctgtcct gcagtctgac	540
ctctacactc tgagcagctc agtgactgtc ccctccagca cctggcccag cgagaccgtc	600
acctgcaacg ttgccacccc ggccagcagc accaaggtgg acaagaaaat tgtgcccagg	660
gatttgtgtg agcccaaadc ttgtgacaaa actcacacat gccaccgtg cccagcacct	720
gaactcctgg ggggaccgtc agtcttcctc ttcccccaa aaccaagga caccctcatg	780
atctcccgga ccctgaggt cacatgcgtg gtggtggacg tgagccacga agaccctgag	840
gtcaagttca actggtacgt ggacggcgtg gaggtgcata atgccaagac aaagccgcgg	900
gaggagcagt acaacagcac gtaccgtgtg gtcagcgtcc tcaccgtcct gcaccaggac	960
tggctgaatg gcaaggagta caagtgaag gtctccaaca aagccctccc agccccatc	1020
gagaaaacca tctccaaagc caaagggcag ccccgagaac cacaggtgta caccctgcc	1080
ccatcccggg atgagctgac caagaaccag gtcagcctga cctgcctggt caaaggcttc	1140
tatcccagcg acatcgccgt ggagtgggag agcaatgggc agccggagaa caactacaag	1200
accacgctc ccgtgctgga ctccgacggc tcctttcttc tctacagcaa gtcaccgtg	1260
gacaagagca ggtggcagca ggggaacgtc ttctcatgct ccgtgatgca tgaggctctg	1320

Sequence Listing

cacaaccact acacgcagaa gagcctctcc ctgtctccgg gtaaaggcgg aggcggatcc	1380
ggtggtggcg gttctaaagc ttccggaggt cccgagggcg gcagcctggc cgcgctgacc	1440
gcgcaccagg ctgtccacct gccgctggag actttcaccc gtcacgcgca gccgcgcggc	1500
tgggaacaac tggagcagtg cggctatccg gtgcagcggc tggtcgccct ctacctggcg	1560
gcgcggctgt cgtggaacca ggtcgaccag gtgatccgca acgccctggc cagccccggc	1620
agcggcggcg acctgggcga agcgatccgc gagcagccgg agcaggcccg tctggccctg	1680
acctggccg ccgccgagag cgagcgcttc gtccggcagg gcaccggcaa cgacgaggcc	1740
ggcgcggcca acggccccgc ggacagcggc gacgccctgc tggagcgcaa ctatcccact	1800
ggcgcggagt tcctcggcga cggcggcgac gtcagcttca gcaccgcgg cacgcagaac	1860
tggacggtgg agcggctgct ccaggcgcac cgccaactgg aggagcgcgg ctatgtgttc	1920
gtcggctacc acggcacctt cctcgaagcg gcgcaaagca tcgtcttcgg cggggtgcgc	1980
gcgcgcagcc aggacctcga cgcgatctgg cgcggtttct atatcgccgg cgatccggcg	2040
ctggcctacg gctacgcccc ggaccaggaa cccgacgcac gcggccggat ccgcaacggt	2100
gccctgctgc gggcttatgt gccgcgctcg agcctgccgg gcttctaccg caccagcctg	2160
acctggccg cgccggaggc ggcgggcgag gtcgaacggc tgatcggcc tccgctgccg	2220
ctgcgcctgg acgccatcac cggccccgag gaggaaggcg ggcgcctgga gaccattctc	2280
ggctggccgc tggccgagcg caccgtggtg attccctcgg cgatccccac cgaccgcgc	2340
aacgtcggcg gcgacctcga cccgtccagc atccccgaca aggaacaggc gatcagcgcc	2400
ctgccggact acgccagcca gcccggaaca ccgccgcgcg aggacctgaa gtaa	2454

<210> 6

Sequence Listing

<211> 1233
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pKL4 plasmid coding sequence

<400> 6
 atgcatcacc atcaccatca cgatgtgaag ctggtggaat ctggaggagg cttagtgcag 60
 cctggagggt ccttgaaact ctctgtgca acctctggat tcactttcag tgactattac 120
 atgtattggg ttgcgcagac tccagagaag aggctggagt gggtcgcata cattagtaat 180
 gatgatagtt ccgccgctta ttcagacact gtaaagggcc gggtcaccat ctccagagac 240
 aatgccagga acaccctcta cctgcaaatt agccgtctga agtctgagga cacagccata 300
 tattcctgtg caagaggact ggcttgggga gcctgttttg cttactgggg ccaagggact 360
 ctggtcactg tctctgcagc caaaacgaca ccccatctg tctatccact ggcccttga 420
 tctgtgtccc aaactaactc catggtgacc ctgggatgcc tggtaagggt ctatttccct 480
 gagccagtga cagtgcctg gaactctgga tccctgtcca gcggtgtgca caccttccca 540
 gctgtcctgc agtctgacct ctacactctg agcagctcag tgactgtccc ctccagcacc 600
 tggcccagcg agaccgtcac ctgcaacgtt gccaccgg ccagcagcac caaggtggac 660
 aagaaaattg tgcccaggga ttgtggtgct aagccttga tagctacaca agcttccggt 720
 ggtggcggat ctggaggtgg cggaagcgga ggtcccagg tgacaggggg aatggcaagc 780
 aagtgggatc agaagggtat ggacattgcc tatgaggagg cggccttagg ttacaaagag 840
 ggtggtgttc ctattggcgg atgtcttatc aataacaaag acggaagtgt tctcggtcgt 900
 ggtcacaaca tgagatttca aaagggatcc gccacactac atggtgagat ctccactttg 960

Sequence Listing

gaaaactgtg ggagattaga gggcaaagtg tacaaagata ccactttgta tacgacgctg	1020
tctccatgcg acatgtgtac aggtgccatc atcatgtatg gtattccacg ctgtgtttgc	1080
ggtgagaacg ttaatttcaa aagtaagggc gagaaatatt tacaaactag aggtcacgag	1140
gttggtgttg ttgacgatga gaggtgtaaa aagatcatga aacaatttat cgatgaaaga	1200
cctcaggatt ggtttgaaga tattggtgag tag	1233

<210> 7
 <211> 4871
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> pMC74 plasmid full sequence

<400> 7 taatacgact cactataggg agaccacaac ggtttccctc tagaaataat ttgttttaac	60
tttaagaagg agatatacat atggatgtga agctggtgga atctggagga ggcttagtgc	120
agcctggagg gtccctgaaa ctctcctgtg caacctctgg attcactttc agtgactatt	180
acatgtattg ggttcgccag actccagaga agaggctgga gtgggtcgca tacattagta	240
atgatgatag ttccgccgct tattcagaca ctgtaaaggg ccggttcacc atctccagag	300
acaatgccag gaacaccctc tacctgcaaa tgagccgtct gaagtctgag gacacagcca	360
tatattcctg tgcaagagga ctggcctggg gagcctgggt tgcttactgg ggccaaggga	420
ctctggtcac tgtctctgca gccaaaacga ccccccatc tgtctatcca ctggcccctg	480
gatctgctgc ccaaactaac tccatggtga cctgggatg cctggtcaag ggctatttcc	540

Sequence Listing

ctgagccagt gacagtgacc tggaactctg gatccctgtc cagcgggtgtg cacaccttcc	600
cagctgtcct gcagtctgac ctctacactc tgagcagctc agtgactgtc ccctccagca	660
cctggcccag cgagaccgtc acctgcaacg ttgccacccc ggccagcagc accaaggtgg	720
acaagaaaat tgtgcccagg gatttgtgta gtaagcctag cataagtaca aaagcttccg	780
gaggtcccga gggcggcagc ctggcgcgcg tgaccgcgca ccaggcttgc cacctgcgcg	840
tggagacttt caccgcgtcat cgccagccgc gcggctggga acaactggag cagtgcggct	900
atccggtgca gcggctggtc gccctctacc tggcggcgcg gctgtcgtgg aaccaggtcg	960
accaggtgat ccgcaacgcc ctggccagcc ccggcagcgg cggcgacctg ggcaagcga	1020
tccgcgagca gccggagcag gcccgctctg ccctgacctt ggccgcgcgc gagagcgagc	1080
gcttcgtccg gcagggcacc ggcaacgacg aggccggcgc ggccaacggc ccggcggaca	1140
gcggcgacgc cctgctggag cgcaactatc ccactggcgc ggagttcctc ggcgacggcg	1200
gcgacgtcag cttcagcacc cgcggcacgc agaactggac ggtggagcgg ctgtccagg	1260
cgcacgcgca actggaggag cgcggtatg tgttcgtcgg ctaccacggc accttcctcg	1320
aagcggcgca aagcatcgtc ttcggcgggg tgcgcgcgcg cagccaggac ctcgacgca	1380
tctggcgcgg tttctatata gccggcgatc cggcgtggc ctacggctac gccaggacc	1440
aggaaccgca cgcacgcggc cggatccgca acggtgcctt gctgcgggtc tatgtgccg	1500
gctcgagcct gccgggcttc taccgcacca gcctgacctt ggccgcgcgc gaggcggcgg	1560
gcgaggtcga acggtgatc ggccatccgc tgccgctgcg cctggacgcc atcaccggcc	1620
ccgaggagga aggcgggcgc ctggagacca ttctcggctg gccgctggcc gagcgacccg	1680
tggtgattcc ctggcgatc cccaccgacc cgcgcaacgt cggcggcgac ctcgacctg	1740

Sequence Listing

ccagcatccc cgacaaggaa caggcgatca gcgccctgcc ggactacgcc agccagcccg	1800
gcaaaccgcc gcgcgaggac ctgaagtaac tgccgcgacc ggccggctcc cttcgcagga	1860
gccggccttc tcggggcctg gccatacatc aggttttcct gatgccagcc caatcgaata	1920
tgaattcggc tgctaacaaa gcccgaaagg aagctgagtt ggctgctgcc accgctgagc	1980
aataactagc ataaccctt gggcctctaa acgggtcttg aggggttttt tgctgaaagg	2040
aggaactata tccggatcgg agatcaattc tggcgtaata gcgaagaggc ccgcaccgat	2100
cgcccttccc aacagttgcg tagcctgaat ggcgaatggg acgcgccctg tagcggcgca	2160
ttaagcgagg cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta	2220
gcgccgctc ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttcccgt	2280
caagctctaa atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac	2340
cccaaaaaac ttgattaggg tgatggttca cgtagtgggc catcgccctg atagacggtt	2400
tttcgccctt tgacgttggg gtccacgttc tttaatagtg gactcttggt ccaaactgga	2460
acaacactca accctatctc ggtctattct tttgatttat aagggtttt gccgatttcg	2520
gcctattggt taaaaaatga gctgatttaa caaaaattta acgcgaattt taacaaaata	2580
ttaacgttta caatttcagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt	2640
ttatttttct aaatacatto aaatatgtat ccgctcatga gacaataacc ctgataaatg	2700
cttcaataat attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt	2760
cccttttttg cggcattttg ccttctgtt tttgctcacc cagaaacgct ggtgaaagta	2820
aaagatgctg aagatcagtt ggtgacaga gtgggttaca tcgaactgga tctcaacagc	2880
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cactttttaa	2940

Sequence Listing

gtttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc	3000
cgcatacact attctcagaa tgacttggtt gactactcac cagtcacaga aaagcatctt	3060
acggatggca tgacagtaag agaattatgc agtgctgcc aagcatgag tgataacact	3120
gcgccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc tttttttcac	3180
aacatggggg atcatgtaac tcgccttgat cgttggaac cggagctgaa tgaagccata	3240
ccaaacgacg agcgtgacac cacgatgcct gtagcaatgg caacaacgtt gcgcaacta	3300
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg	3360
gataaagttg caggaccact tctgcgctcg gcccttcggt ctggctggtt tattgctgat	3420
aaatctggag ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt	3480
aagccctccc gtatcgtagt tatctacacg acgggcagtc aggcaactat ggatgaacga	3540
aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa	3600
gtttactcat atatacttta gattgattta aaacttcatt ttttaatttaa aaggatctag	3660
gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttcac	3720
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc	3780
gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg tttgccggat	3840
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat	3900
actgtccttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct	3960
acatacctcg ctctgctaata cctgttacca gtggctgctg ccagtggcga taagtctgt	4020
cttaccgggt tggactcaag acgatagtta ccgataagg cgcagcggtc gggctgaacg	4080
gggggttcgt gcacacagcc cagcttgag cgaacgacct acaccgaact gagataccta	4140

Sequence Listing

```

cagcgtgagc attgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg      4200
gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg gaacgcctgg      4260
tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc      4320
tcgtcagggg ggccgagcct atggaaaaac gccagcaacg cggccttttt acggttcctg      4380
gccttttgct ggccttttgc tcacatgttc tttcctgcgt tatcccctga ttctgtggat      4440
aaccgtatta ccgcctttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc      4500
agcgagtcag tgagcgagga agcggaagag cgctgatgc ggtattttct ccttacgcat      4560
ctgtgcggta tttcacaccg catatatggt gcactctcag tacaatctgc tctgatgccg      4620
catagttaag ccagtataca ctccgctatc gctacgtgac tgcaaggaga tggcgcccaa      4680
cagtcccccg gccacggggc ctgccaccat acccacgccg aaacaagcgc tcatgagccc      4740
gaagtggcga gccgatcctt ccccatcggt gatgtcggcg atataggcgc cagcaaccgc      4800
acctgtggcg ccggtgatgc cggccacgat gcgtccggcg tagaggatct tgagatctcg      4860
atccgcgaaa t                                                                4871

```

<210> 8

<211> 4886

<212> DNA

<213> Artificial Sequence

<220>

<223> pMH21 plasmid full sequence

<400> 8

```

taatacgact cactataggg agaccacaac ggtttccttc tagaaataat tttgtttaac      60
ttaagaagg agatatacat atggagggtga agctgggtgga atctggagga ggcttagtgc      120

```

Sequence Listing

agcctggagg gtccctgaaa ctctcctgtg caacctctgg attcactttc agtgactatt	180
acatgtattg ggttcgccag actccagaga agaggctgga gtgggtcgca tacattagta	240
atgatgatag ttccgccgct tattcagaca ctgtaaaggg cgggttcacc atctccagag	300
acaatgccag gaacaccctc tacctgcaaa tgagccgtct gaagtctgag gacacagcca	360
tatattcctg tgcaagagga ctggcctggg gagcctggtt tgcttactgg ggccaaggga	420
ctctggtcac tgtctctgca gccaaaacga ccccccatc tgtctatcca ctggcccctg	480
gatctgctgc ccaaactaac tccatggtga ccctgggatg cctgggtcaag ggctatttcc	540
ctgagccagt gacagtgacc tggaactctg gatccctgtc cagcgggtgtg cacaccttcc	600
cagctgtcct gcagtctgac ctctacactc tgagcagctc agtgactgtc ccctccagca	660
cctggcccag cgagaccgtc acctgcaacg ttgccacccc ggccagcagc accaagggtg	720
acaagaaaat tgtgcccagg gattgtggta gtaagccttg cataagtaca aaagcttctg	780
gtggtggcgg atctggaggt cccgagggcg gcagcctggc cgcgctgacc gcgcaccagg	840
cttgccacct gccgctggag actttcaccg gtcacgcca gccgcgcggc tgggaacaac	900
tggagcagtg cggctatccg gtgcagcggc tggtcgccct ctacctggcg gcgcggctgt	960
cgtggaacca ggtcgaccag gtgatccgca acgcccctggc cagccccggc agcggcggcg	1020
acctgggcga agcgatccgc gagcagccgg agcaggcccg tctggccctg acctggccg	1080
ccgccgagag cgagcgcttc gtccggcagg gcaccggcaa cgacgaggcc ggcgcgcca	1140
acggcccggc ggacagcggc gacgccctgc tggagcgcaa ctatcccact ggcgcgaggt	1200
tcctcggcga cggcggcgac gtcagcttca gcaccgcgg cagcagaac tggacggtg	1260
agcggctgct ccaggcgcac cgccaactgg aggagcgcg ctatgtgttc gtcggctacc	1320

Sequence Listing

acggcacctt cctcgaagcg gcgcaaagca tcgtcttcgg cggggtgcmc gcgcgcagcc	1380
aggacctcga cgcgatctgg cgcggtttct atatcgccgg cgatccggcg ctggcctacg	1440
gctacgcccc ggaccaggaa cccgacgcac gcggccggat ccgcaacggt gccctgctgc	1500
gggtctatgt gccgcgctcg agcctgccgg gcttctaccg caccagcctg accctggccg	1560
cgccggaggc ggcgggcgag gtcgaacggc tgatcgccca tccgtgccc ctgcgcctgg	1620
acgccatcac cggccccgag gaggaaggcg ggcgctgga gaccattctc ggctggccgc	1680
tggccgagcg caccgtggtg attccctcgg cgatccccac cgacccgcgc aacgtcggcg	1740
gcgacctcga cccgtccagc atccccgaca aggaacaggc gatcagcgcc ctgccggact	1800
acgccagcca gcccggaaca ccgcccgcgc aggacctgaa gtaactgccg cgaccggccg	1860
gtccctctcg caggagccgg ctttctcggg gcctggccat acatcagggt ttcctgatgc	1920
cagcccaatc gaatatgaat tcggctgcta acaaagcccg aaaggaagct gagtggctg	1980
ctgccaccgc tgagcaataa ctagcataac cccttgggcc tctaaacggg tcttgagggg	2040
ttttttgctg aaaggaggaa ctatatccgg atcggagatc aattctggcg taatagcgaa	2100
gaggcccgca ccgatcgccc ttccaacag ttgcgtagcc tgaatggcga atgggacgcg	2160
ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca	2220
cttgccagcg ccctagcgcc cgctcctttc gctttcttcc ctctctttct cgccacgttc	2280
gccggctttc ccgctcaagc tctaaatcgg gggctccctt tagggttccg atttagtgct	2340
ttacggcacc tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg	2400
ccctgataga cggttttctg ccctttgacg ttggagtcca cgttctttta tagtggactc	2460
ttgttccaaa ctggaacaac actcaaccct atctcggctc attcttttga tttataaggg	2520

Sequence Listing

atTTTgCCga tTtCGgccta tTggttaaaa aatgagctga' tTtaacaaaa atttaacgCG	2580
aattTTaaca aaatattaac gTTtacaatt tCaGgtggCa cTtttCGggg aaatgtGcgC	2640
ggaaccCcta tTtgTttatt tTtctaaata cattcaaata tGtatcGct catgagacaa	2700
taaccctgat aaatgcttCa ataattattga aaaaggaaga gTatgagtat tCaacatttC	2760
cgtgtGgccc tTattccctt tTttgCGgCa tTttgCcttC ctgttttttGc tCaccagaa	2820
acGctggtga aagtaaaaga tGctgaagat cagttgggtG cagGagtggg tTacatcgaa	2880
ctGgatctCa acagCGgtaa gatCcttgag agTtttCGcc cCGaagaacg tTttccaatG	2940
atgagcactt tTaaagttct gctatgtggC gcggtattat cCGtattga cGccgggcaa	3000
gagcaactcg gTcgcCGcat aCactattct cagaatgact tGgttgagta cTcaccagtC	3060
acagaaaagc atCttacgga tggcatgaca gTaagagaat tatGcagtGc tGccataagc	3120
atgagtGata aCactgCGgc caacttactt ctGacaacga tCGgaggacc gaaggagcta	3180
accgcttttt tTcacaacat gggggatcat gTaactCGcc tTgatcgttG ggaaccggag	3240
ctgaatgaag cCataccaaa cGacgagcgt gacaccagCa tGcctgtagc aatggcaaca	3300
acgttGcgCa aactattaac tggcgaacta ctTactctag ctTccCGgCa acaattaata	3360
gactggatgg aggcggataa agttgcagga cCacttctGc gctCGgccct tCGgctggc	3420
tGgtttattG ctGataaatC tggagccggt gagcgtgggt cTcgCGgtat cattGcagCa	3480
ctggggccag atGgtaagcc cTcccgTatC gtagttatct acacgacggg cagtcaggCa	3540
actatggatG aacgaaatag acagatCGct gagataggTg cTcactgat taagcattgg	3600
taactgtcag accaagttta cTcatatata ctTtagattG atttaaaact tCatttttaa	3660
tttaaaagga tctaggTgaa gatccTtttT gataatctCa tGaccaaaat cCcttaacgt	3720

Sequence Listing

gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat	3780
cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg	3840
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga	3900
gcgagatac caaatactgt ctttctagt tagcctagt taggccacca cttcaagaac	3960
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt	4020
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag	4080
cggtcgggct gaacggggggg ttcgtgcaca cagcccagct tggagcgaac gacctacac	4140
gaactgagat acctacagcg tgagcattga gaaagcgcca cgcttcccga agggagaaag	4200
gcgagacagg atccggtaag cggcagggc ggaacaggag agcgacgag ggagcttcca	4260
ggggggaacg cctgggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt	4320
cgatttttgt gatgctcgtc aggggggccc agcctatgga aaaacgccag caacgcggcc	4380
tttttacggt tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc	4440
cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc	4500
cgaacgaccg agcgacgca gtcagttagc gaggaagcgg aagagcgct gatgcggtat	4560
tttctcctta cgcctctgtg cggatattca caccgcatat atggtgcact ctgagtacaa	4620
tctgctctga tgccgcatag ttaagccagt atacactccg ctatcgctac gtgactgcaa	4680
ggagatggcg cccaacagtc ccccgccac ggggcctgcc accataccca cgccgaaaca	4740
agcgctcatg agcccgaagt ggcgagccc atcttcccca tcggtgatgt cggcgatata	4800
ggcgccagca accgcacctg tggcgccggt gatgccggcc acgatgcgtc cggcgtagag	4860
gatcttgaga tctcgatccg cgaaat	4886

Sequence Listing

<210> 9
 <211> 4871
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pCE2 plasmid full sequence

<400> 9
 taatacgact cactataggg agaccacaac ggtttcctc tagaaataat ttgtttaac 60
 ttaagaagg agatatacat atggatgtga agctgggga atctggagga ggcttagtgc 120
 agcctggagg gtccctgaaa ctctcctgtg caacctctgg attcactttc agtgactatt 180
 acatgtattg ggttogccag actccagaga agaggctgga gtgggtcgca tacattagta 240
 atgatgatag ttccgcccgt tattcagaca ctgtaaaggg ccggttcacc atctccagag 300
 acaatgccag gaacaccctc tacctgcaaa tgagccgtct gaagtctgag gacacagcca 360
 tatattcctg tgcaagagga ctggcctggg gagcctggtt tgcttactgg ggccaaggga 420
 ctctggtcac tgtctctgca gccaaaacga ccccccatc tgtctatcca ctggcccctg 480
 gatctgctgc ccaaactaac tccatggtga ccctgggatg cctggtcaag ggctatttcc 540
 ctgagccagt gacagtgacc tggaactctg gatccctgtc cagcgggtgtg cacaccttcc 600
 cagctgtcct gcagtctgac ctctacactc tgagcagctc agtgactgtc ccctccagca 660
 cctggcccag cgagaccgtc acctgcaacg ttgccacccc ggccagcagc accaaggtgg 720
 acaagaaaat tgtgcccagg gattgtggta gtaagccttg cataagtaca aaagcttccg 780
 gaggtcccga gggcggcagc ctggccgcgc tgaccgcgca ccaggcttgc cacctgccgc 840

Sequence Listing

tggagacttt caccggtcat cgccagccgc gcggtggga acaactggag cagtgcggct	900
atccggtgca gcggtggtc gccctctacc tggcggcgcg gctgtcgtgg aaccaggtcg	960
accaggtgat ccgcaacgcc ctggccagcc ccggcagcgg cggcgacctg ggcgaagcga	1020
tccgcgagca gccggagcag gcccgctctgg ccctgaccct ggccgccgcc gagagcgagc	1080
gcttcgtccg gcagggcacc ggcaacgacg aggccggcg gcgcaacggc ccggcggaca	1140
gcggcgacgc cctgctggag cgcaactatc ccactggcg gcggttcctc ggcgacggcg	1200
gcgacgtcag cttcagcacc cgcgccacgc agaactggac ggtggagcgg ctgctccagg	1260
cgcaccccca actggaggag cgcggtatg tgttcgtcgg ctaccacggc accttctcg	1320
aagcggcgca aagcatcgtc ttccgggggg tgccgcgcgc cagccaggac ctcgacgcga	1380
tctggcgcgg tttctatata gccggcgatc cggcgctggc ctacggctac gccaggacc	1440
aggaaccoga cgcacgcggc cggatccgca acggtgccct gctgcgggtc tatgtgccgc	1500
gctcgagcct gccgggcttc taccgcacca gctgaccct ggccgcgcgc gaggcggcg	1560
gcgaggtcga acggtgatc ggccatccgc tgccgctgcg cctggacgcc atcaccggcc	1620
ccgaggagga aggcgggcgc ctggagacca ttctcggtg gccgctggcc gagcgcaccg	1680
tggtgattcc ctcgcgatc cccaccgacc cgcgcaacgt cggcggcgac ctcgaccgt	1740
ccagcatccc cgacaaggaa caggcgatca gcgccctgcc ggactacgcc agccagccc	1800
gcaaaccgcc gcgcgaggac ctgaagtaac tgccgcgacc ggccggctcc cttcgcagga	1860
gccggccttc tcggggcctg gccatacatc aggttttctt gatgccagcc caatcgaata	1920
tgaattcggc tgctaacaaa gcccgaaagg aagctgagtt ggctgctgcc accgctgagc	1980
aataactagc ataaccctt ggccctctaa acgggtcttg aggggttttt tgctgaaagg	2040

Sequence Listing

aggaactata tccggatcgg agatcaattc tggcgtaata gcgaagaggc ccgcaccgat	2100
cgcccttccc aacagttgcg tagcctgaat ggcgaaatggg acgcgccctg tagcggcgca	2160
ttaagcgcg ggggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta	2220
gcgcccgcctc ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttcccgt	2280
caagctctaa atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac	2340
cccaaaaaac ttgattaggg tgatggttca cgtagtgggc catcgccctg atagacggtt	2400
tttcgccctt tgacgttgga gtccacgttc tttaatagtg gactcttggt ccaaactgga	2460
acaacactca accctatctc ggtctattct tttgatttat aagggtttt gccgatttcg	2520
gcctattggt taaaaaatga gctgatttaa caaaaattta acgcgaattt taacaaaata	2580
ttaacgttta caatttcagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt	2640
ttatTTTTCT aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaatg	2700
cttcaataat attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt	2760
ccctTTTTTG cggcattttg ccttcctggt tttgctcacc cagaaacgct ggtgaaagta	2820
aaagatgctg aagatcagtt ggggtgcacga gtgggttaca tcgaactgga tctcaacagc	2880
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa	2940
gttctgctat gtggcgcggt attatcccggt attgacgccg ggcaagagca actcggtcgc	3000
cgcatacact attctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt	3060
acggatggca tgacagtaag agaattatgc agtgctgcca taagcatgag tgataacact	3120
gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc tttttttcac	3180
aacatggggg atcatgtaac tcgccttgat cgttggaac cggagctgaa tgaagccata	3240

Sequence Listing

cctaaacgacg agcgtgacac cactgatgcct gtagcaatgg caacaacgtt gcgcaaacta	3300
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg	3360
gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctgggt tattgctgat	3420
aaatctggag ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt	3480
aagccctccc gtatcgtagt tatctacacg acgggcagtc aggcaactat ggatgaacga	3540
aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa	3600
gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa aaggatctag	3660
gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac	3720
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt tttctgcgc	3780
gtaatctgct gcttgcaaac aaaaaacca ccgtaccag cggtggtttg tttgccgat	3840
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat	3900
actgtccttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct	3960
acatacctcg ctctgctaata cctgttacca gtggctgctg ccagtggcga taagtcgtgt	4020
cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg	4080
gggggttcgt gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta	4140
cagcgtgagc attgagaaag cgccacgctt cccgaaggga gaaaggcgga caggatatcg	4200
gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg gaacgcctgg	4260
tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc	4320
tcgtcagggg ggccgagcct atggaaaaac gccagcaacg cggccttttt acggttcctg	4380
gccttttgct ggccttttgc tcacatgttc tttcctgcgt tatccctga ttctgtggat	4440

Sequence Listing

aaccgtatta ccgcctttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc 4500
agcgagtcag tgagcgagga agcggaagag cgcctgatgc ggtattttct ctttacgcat 4560
ctgtgcggta ttccacaccg catatatggt gcactctcag tacaatctgc tctgatgccg 4620
catagttaag ccagtataca ctccgctatc gctacgtgac tgcaaggaga tggcgcccaa 4680
cagtcccccg gccacggggc ctgccaccat acccagccg aaacaagcgc tcatgagccc 4740
gaagtggcga gcccgatctt ccccatcggt gatgtcggcg atataggcgc cagcaaccgc 4800
acctgtggcg ccggtgatgc cggccacgat gcgtccggcg tagaggatct tgagatctcg 4860
atccgcgaaa t 4871

<210> 10
<211> 3703
<212> DNA
<213> Artificial Sequence

<220>
<223> pMC75 plasmid full sequence

<400> 10
taatacgact cactataggg agaccacaac ggtttccctc tagaaataat tttgtttaac 60

ttaaagaagg agatatacat atggatgtgc tgatgacca gtctccattg agtttacctg 120

tcagtcttgg agatcaagcc tccatctctt gcagatctag tcagatcatt gtacatagta 180

atggaaacac ctatttagaa tggtagctgc agaaaccagg ccagtctcca aagctcctga 240

tctacaaagt ttccaaccga ttttctgggg tcccagacag gttcagtggc agtggatcag 300

ggacagattt cacactcaag atcagcagag tggaggctga ggatctggga gtttattact 360

gctttcaagg ttcatatgtt ccattcacgt tcggctcggg gacaaagttg gaaataaac 420

Sequence Listing

gggctgatgc tgcaccaact gtatccatct tcccaccatc cagtgcagcag ttaacatctg	480
gaggtgcctc agtcgtgtgc ttcttgaaca acttctaccc caaagacatc aatgtcaagt	540
ggaagattga tggcagtgaa cgacaaaatg gcgtcctgaa cagttggact gatcaggaca	600
gcaaagacag cacctacagc atgagcagca ccctcacgtt gaccaaggac gagtatgaac	660
gacataacag ctatacctgt gagggcactc acaagacatc aacttcaccc attgtcaaga	720
gcttcaacag gaatgagtgt ggtaaagctt aatgaattcg gctgctaaca aagcccgaag	780
ggaagctgag ttggctgctg ccaccgctga gcaataacta gcataacccc ttgggcctct	840
aaacgggtct tgaggggttt ttgctgaaa ggaggaacta tatccgcatc ggagatcaat	900
tctggcgtaa tagcgaagag gcccgacccg atcgcccttc ocaacagttg cgtagcctga	960
atggcgaatg ggacgcgccc tgtagcggcg cattaagcgc ggcggtgtg gtggttacgc	1020
gcagcgtgac cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttcctt	1080
cctttctcgc cacgttcgcc ggtttcccc gtcaagctct aaatcggggg ctcccttag	1140
ggttccgatt tagtgcttta cggcacctcg acccaaaaa acttgattag ggtgatggtt	1200
cacgtagtgg gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt	1260
tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt	1320
cttttgattt ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt	1380
aacaaaaatt taacgcgaat tttaacaaaa tattaacgtt tacaatttca ggtggcactt	1440
ttcggggaaa tgtgcgcgga acccctattt gttattttt ctaaatacat tcaaatatgt	1500
atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta	1560
tgagtattca acatttcctg gtgcgcccta ttcccttttt tgcggcattt tgccctcctg	1620

Sequence Listing

tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac	1680
gagtgggtta catcgaactg gatctcaaca gcggtaaagat ccttgagagt tttcgccccg	1740
aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc	1800
gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg	1860
ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat	1920
gcagtgtctgc cataagcatg agtgataaca ctgcggccaa cttacttctg acaacgatcg	1980
gaggaccgaa ggagctaacc gctttttttc acaacatggg ggatcatgta actcgccttg	2040
atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc	2100
ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt	2160
cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct	2220
cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc	2280
gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca	2340
cgacgggcag tcaggcaact atggatgaac gaaatagaca gatcgtgag ataggtgcct	2400
cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt	2460
taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga	2520
ccaaaatccc ttaacgtgag ttttggttcc actgagcgtc agaccccgta gaaaagatca	2580
aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac	2640
caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg	2700
taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag	2760
gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atoctgttac	2820

Sequence Listing

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt	2880
taccggataa ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg	2940
agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gcattgagaa agcgccacgc	3000
ttcccgaagg gagaaaggcg gacaggtatc cggttaagcgg cagggtcgga acaggagagc	3060
gcacgagggg gcttccaggg gggaacgcct ggtatcttta tagtcctgtc gggtttcgcc	3120
acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggccgagc ctatggaaaa	3180
acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt	3240
tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg	3300
ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggag	3360
agcgctgat gcggtatttt ctcttacgc atctgtgcgg tatttcacac cgcatatatg	3420
gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagtata cactccgcta	3480
tcgctacgtg actgcaagga gatggcgccc aacagtcccc cggccacggg gcctgccacc	3540
ataccacgc cgaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg	3600
gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gccggccacg	3660
atgcgtccgg cgtagaggat cttgagatct cgatccgcga aat	3703

<210> 11

<211> 5576

<212> DNA

<213> Artificial Sequence

<220>

<223> pLSC52 plasmid full sequence

Sequence Listing

<400> 11

taatacgact cactataggg agaccacaac ggtttccctc tagaaataat ttgtttaac	60
tttaagaagg agatatacat atggatgtga agctggtgga atctggagga ggcttagtgc	120
agcctggagg gtccctgaaa ctctcctgtg caacctctgg attcactttc agtgactatt	180
acatgtattg ggttcgccag actccagaga agaggctgga gtgggtcgca tacattagta	240
atgatgatag ttccgcccgt tattcagaca ctgtaaaggg ccggttcacc atctccagag	300
acaatgccag gaacaccctc tacctgcaaa tgagccgtct gaagtctgag gacacagcca	360
tatattcctg tgcaagagga ctggcctggg gagcctggtt tgcttactgg ggccaaggga	420
ctctggtcac tgtctctgca gccaaaacga ccccccatc tgtctatcca ctggcccctg	480
gatctgctgc ccaaactaac tccatggtga ccctgggatg cctgggtcaag ggctatttcc	540
ctgagccagt gacagtgacc tggaactctg gatccctgtc cagcgggtgtg cacaccttcc	600
cagctgtcct gcagtctgac ctctacactc tgagcagctc agtgactgtc ccctccagca	660
cctggcccag cgagaccgtc acctgcaacg ttgccacccc ggccagcagc accaagggtg	720
acaagaaaat tgtgcccagg gatttgtgtg agcccaaatac ttgtgacaaa actcacacat	780
gccacccgtg ccagcacct gaactcctgg ggggaccgtc agtcttctc ttcccccaa	840
aaccaagga caccctcatg atctcccga ccctgaggt cacatgcgtg gtggtggacg	900
tgagccacga agaccctgag gtcaagttca actggtacgt ggacggcgtg gaggtgcata	960
atgccaagac aaagccgagg gaggagcagt acaacagcac gtaccgtgtg gtcagcgtcc	1020
tcaccgtcct gcaccaggac tggctgaatg gcaaggagta caagtgaag gtctccaaca	1080
aagccctccc agccccatc gagaaaacca tctccaaagc caaagggcag ccccgagaac	1140

Sequence Listing

cacaggtgta caccctgccc ccatcccggg atgagctgac caagaaccag gtcagcctga	1200
cctgcctggt caaaggcttc tatcccagcg acatcgccgt ggagtgggag agcaatgggc	1260
agccggagaa caactacaag accacgcctc ccgtgctgga ctccgacggc tcctttctcc	1320
ttacagcaa gctcaccgtg gacaagagca ggtggcagca ggggaacgtc ttctcatgct	1380
ccgtgatgca tgaggctctg cacaaccact acacgcagaa gacgctctcc ctgtctccgg	1440
gtaaaggcgg aggcggatcc ggtggtggcg gttctaaagc ttccggaggt cccgagggcg	1500
gcagcctggc cgcgctgacc gcgcaccagg cttgccacct gccgctggag actttcacc	1560
gtcatcgcca gccgcggcgc tgggaacaac tggagcagtg cggctatccg gtgcagcggc	1620
tggtcgccct ctacctggcg gcgcggctgt cgtggaacca ggtcgaccag gtgatccgca	1680
acgccctggc cagccccggc agcggcggcg acctgggcga agcgatccgc gagcagccgg	1740
agcaggcccg tctggccctg accctggccg ccgccgagag cgagcgcttc gtccggcagg	1800
gcaccggcaa cgacgaggcc ggcgcggcca acggccccgc ggacagcggc gacgccctgc	1860
tggagcgcaa ctatcccact ggcgcggagt tcctcggcga cggcggcgac gtcagcttca	1920
gcaccgcggc cagcagaac tggacggtgg agcggctgct ccaggcgac cgccaactgg	1980
aggagcgcg ctagtgttgc gtcggctacc acggcacctt cctcgaagcg gcgcaaagca	2040
tcgtcttcgg cggggtgccc gcgcgcagcc aggaacctga cgcgatctgg cgcggtttct	2100
atatcgccgg cgatccggcg ctggcctacg gctacgccca ggaccaggaa cccgacgcac	2160
gcggccggat ccgcaacggt gccctgctgc ggtctatgt gccgcgctcg agcctgccgg	2220
gcttctaccg caccagcctg accctggccg cgcggaggc ggcgggagag gtogaacggc	2280
tgatcggcca tccgctgccg ctgcgcctgg acgccatcac cggccccgag gaggaaggcg	2340

Sequence Listing

ggcgcctgga gaccattctc ggctggccgc tggccgagcg caccgtggtg attccctcgg	2400
cgatccccac cgaccgcgc aacgtcggcg gcgacctga cccgtccagc atccccgaca	2460
aggaacaggc gatcagcgcc ctgccggact acgccagcca gcccggcaaa ccgccgcgcg	2520
aggacctgaa gtaactgccg cgaccggccg gctcccttcg caggagccgg ccttctcggg	2580
gcctggccat acatcagggtt ttctgatgc cagcccaatc gaatatgaat tcggctgcta	2640
acaaagcccg aaaggaagct gagttggctg ctgccaccgc tgagcaataa ctagcataac	2700
cccttggggc tctaaacggg tcttgagggg ttttttctg aaaggaggaa ctatatccgg	2760
atcgagatc aattctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag	2820
ttgcgtagcc tgaatggcga atgggacgcg ccctgtagcg gcgcattaag cgcggcgggt	2880
gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc	2940
gctttcttcc cttcctttct cgccacgttc gcgggcttcc cccgtcaagc tctaaatcgg	3000
gggctccctt tagggttccg atttagtgct ttacggcacc tcgaccccaa aaaacttgat	3060
tagggtagtg gttcacgtag tgggccatcg ccctgataga cggtttttcg ccctttgacg	3120
ttggagtcca cgttctttaa tagtggaactc ttgttcaaaa ctggaacaac actcaaccct	3180
atctcgggtct attcctttga tttataaggg attttgccga tttcggccta ttggttaaaa	3240
aatgagctga tttaacaaaa atttaacgcg aattttaaca aaatattaac gtttacaatt	3300
tcagggtggc cttttcgggg aaatgtgcgc ggaaccccta tttgtttatt tttctaaata	3360
cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca ataatttga	3420
aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt ttttgcgga	3480
ttttgccttc ctgtttttgc tcaccagaa acgctggtga aagtaaaaga tgcgaagat	3540

Sequence Listing

cagttgggtg cacgagtggg ttacatcgaa ctggatctca acagcggtaa gatccttgag	3600
agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagttct gctatgtggc	3660
gcggtattat cccgtattga cgccgggcaa gagcaactcg gtcgccgcat aactattct	3720
cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga tggcatgaca	3780
gtaagagaat tatgcagtgc tgccataagc atgagtata aactgctggc caacttactt	3840
ctgacaacga tcggaggacc gaaggagcta accgcttttt ttcacaacat gggggatcat	3900
gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt	3960
gacaccacga tgctgtagc aatggcaaca acgttgcgca aactattaac tggcgaacta	4020
cttactctag cttcccgga acaattaata gactggatgg aggcggataa agttgcagga	4080
ccacttctgc gctcgccct tccggtggc tggtttattg ctgataaatc tggagccggt	4140
gagcgtgggt ctgcgggtat cattgcagca ctggggccag atggtaagcc ctccgtatc	4200
gtagttatct acacgacggg cagtcaggca actatggatg aacgaaatag acagatcgt	4260
gagataggtg cctcactgat taagcattgg taactgtcag accaagttta ctcatatata	4320
ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa gatccttttt	4380
gataatctca tgacaaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc	4440
gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg	4500
caaacaaaaa aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact	4560
ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt ctttctagt	4620
tagccgtagt taggccacca cttaagaac tctgtagcac gcctacata cctgctctg	4680
ctaatoctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttgac	4740

Sequence Listing

tcaagacgat agttaccgga taaggcgag cggtcgggct gaacgggggg ttcgtgcaca	4800
cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg tgagcattga	4860
gaaagcgcca cgcttcccga agggagaaaag gcgacaggt atccggtaag cggcagggtc	4920
ggaacaggag agcgcacgag ggagcttcca ggggggaacg cctggatatct ttatagtcct	4980
gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggccg	5040
agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct	5100
tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc	5160
tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgacgca gtcagtgagc	5220
gaggaagcgg aagagcgct gatgcggtat tttctcctta cgcactctgt cggtatattca	5280
caccgcatat atggtgcact ctgagtacaa tctgctctga tgccgcatag ttaagccagt	5340
atacactccg ctatcgctac gtgactgcaa ggagatggcg cccaacagtc ccccgccac	5400
ggggcctgcc accataccca cgccgaaaca agcgctcatg agccgaagt ggcgagcccg	5460
atcttcccca tcggtgatgt cggcgatata ggcgccagca accgcacctg tggcgccggt	5520
gatgccggcc acgatgcgtc cggcgtagag gatcttgaga tctcgatccg cgaaat	5576

<210> 12
 <211> 4263
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> pKL4 plasmid full sequence

<400> 12	
taatacgact cactataggg agaccacaac ggtttcctc tagaaataat ttgtttaac	60

Sequence Listing

tttaagaagg agatatacat atgcatcacc atcaccatca cgatgtgaag ctggtggaat	120
ctggaggagg cttagtgca cctggagggt ccctgaaact ctctgtgca acctctggat	180
tcactttcag tgactattac atgtattggg ttcgccagac tccagagaag aggttgaggt	240
gggtcgcata cattagtaat gatgatagtt cgcgcgtta ttcagacact gtaaagggcc	300
ggttcacat ctccagagac aatgccagga acaccctcta cctgcaaag agccgtctga	360
agtctgagga cacagccata tattctgtg caagaggact ggcctgggga gcctggtttg	420
cttactgggg ccaagggact ctggtcactg tctctgcagc caaaacgaca ccccatctg	480
tctatccact ggcccctgga tctgtgccc aaactaactc catggtgacc ctgggatgcc	540
tggtcaaggg ctatttcct gagccagtga cagtgcctg gaactctgga tccctgtcca	600
gcggtgtgca caccttccca gctgtcctgc agtctgacct ctacactctg agcagctcag	660
tgactgtccc ctccagcacc tggcccagcg agaccgtcac ctgcaacgtt gccacccgg	720
ccagcagcac caaggtggac aagaaaattg tgcccaggga ttgtggtgct aagccttgca	780
tagctacaca agcttcgggt ggtggcgat ctggagggtg cggaagcgga ggtcccagg	840
tgacaggggg aatggcaagc aagtgggatc agaagggat ggacattgcc tatgaggagg	900
cggccttagg ttacaaagag ggtggtgttc ctattggcgg atgtcttacc aataacaaag	960
acggaagtgt tctcggtcgt ggtcacaaca tgagatttca aaagggatcc gccacactac	1020
atggtgagat ctccacttg gaaaactgtg ggagattaga gggcaaagtg taaaagata	1080
ccactttgta tacgacgtg tctccatgcg acatgtgtac aggtgccatc atcatgtatg	1140
gtattccacg ctgtgtgtgc ggtgagaacg ttaatttcaa aagtaagggc gagaaatatt	1200
tacaactag aggtcacgag gttgtgtgtg ttgacgatga gaggtgtaa aagatcatga	1260

Sequence Listing

aacaatttat cgatgaaaga cctcaggatt ggtttgaaga tattggtgag taggaattcg	1320
gctgctaaca aagcccgaaa ggaagctgag ttggctgctg ccaccgctga gcaataacta	1380
gcataacccc ttgggcctct aaacgggtct tgaggggttt tttgctgaaa ggaggaacta	1440
tatccggatc ggagatcaat tctggcgtaa tagcgaagag gcccgaccg atcgcccttc	1500
ccaacagttg cgtagcctga atggcgaatg ggacgcgccc tgtagcggcg cattaagcgc	1560
ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgcgc tagcgcgcgc	1620
tcctttcgct ttcttccctt cctttctcgc cagttcgcgc ggctttcccc gtcaagctct	1680
aaatcggggg ctcccttttag ggttccgatt tagtgcttta cggcacctcg acccaaaaa	1740
acttgattag ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc	1800
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact	1860
caaccctatc tcggtctatt cttttgattht ataagggtt ttgccgattt cggcctattg	1920
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat ttttaaaaaa tattaacgtt	1980
tacaatttca ggtggcactt ttccgggaaa tgtgcgcgga acccctattt gtttattttt	2040
ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata	2100
atattgaaaa aggaagagta tgagtattca acatttcctg gtcgccotta ttcccttttt	2160
tgcggcattt tgccttcctg tttttgctca ccagaaaacg ctggtgaaag taaaagatgc	2220
tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtgaagat	2280
ccttgagagt tttcgccccg aagaacgttt tccaatgatg agcactttta aagtcttgct	2340
atgtggcgcg gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca	2400
ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg	2460

Sequence Listing

catgacagta agagaattat gcagtgtctgc cataagcatg agtgataaca ctgcggccaa	2520
cttacttctg acaacgatcg gaggaccgaa ggagctaacc gctttttttc acaacatggg	2580
ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga	2640
cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg	2700
cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt	2760
tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg	2820
agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc	2880
ccgtatcgta gttatctaca cgacgggcag tcaggcaact atggatgaac gaaatagaca	2940
gatcgtgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc	3000
atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat	3060
cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc	3120
agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg	3180
ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct	3240
accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct	3300
tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct	3360
cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg	3420
gttggaactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggggtc	3480
gtgcacacag ccagccttg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga	3540
gcattgagaa agcggcacgc ttcccgaagg gagaaaggcg gacaggatc cggtaaagcg	3600
cagggtcgga acaggagagc gcacgaggga gcttccaggg gggaacgcct ggtatcttta	3660

Sequence Listing

tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gtcggtcagg	3720
ggggccgagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg	3780
ctggcctttt gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat	3840
taccgccttt gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc	3900
agtgagcgag gaagcggaag agcgctgat gcggtatfff ctccttacgc atctgtgcgg	3960
tatttcacac cgcatatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta	4020
agccagtata cactccgcta tcgctacgtg actgcaagga gatggcgccc aacagtcccc	4080
cggccacggg gcctgccacc ataccacgc cgaaacaagc gctcatgagc ccgaagtggc	4140
gagcccgatc ttccccatcg gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg	4200
cgccggtgat gccggccacg atgcgtccgg cgtagaggat cttgagatct cgatccgcga	4260
.aat	4263

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2004/001595

A. CLASSIFICATION OF SUBJECT MATTER**IPC7 C07K 19/00, C07K 16/00, C07K 14/21**

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, C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Patents and Applications for Inventions since 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PubMed, Delphion, CA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	Park, J.H., et al. "A divalent recombinant immunotoxin formed by a disulfide bond between the extension peptide chains" Mole. Cells, 2001, Vol. 12(3): pages 398-402, See entire document.	1-18, 20 19
X Y	US 4956453 (Cetus Corporation), Sep. 11, 1990, See entire document.	1-18, 20 19
A	Bera, T.K., et al. "Bivalent disulfide-stabilized fragment variable immunotoxin directed against mesotheliomas and ovarian cancer" Mole. Cancer Ther., 2001, Vol. 1: pages 79-84, See entire document.	1-20
A	Reiter, Y., et al. "Antitumor activity and pharmacokinetics in mice of a recombinant immunotoxin containing a disulfide-stabilized Fv fragment" Cancer Res., 1994, Vol. 54(10): 2714-2718, See entire document.	1-20
A	Zhu, Z., et al., "Remodeling domain interfaces to enhance heterodimer formation" Protein Sci., 1997, Vol. 6(4): 781-788, See entire document.	1-20

☐ Further documents are listed in the continuation of Box C.☒ 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 application or patent 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 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

22 OCTOBER 2004 (22.10.2004)

Date of mailing of the international search report

25 OCTOBER 2004 (25.10.2004)

Name and mailing address of the ISA/KR



Korean Intellectual Property Office
920 Dunsan-dong, Seo-gu, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

PARK, JEONG UNG

Telephone No. 82-42-481-8159



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2004/001595

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of :

a. type of material



a sequence listing



table(s) related to the sequence listing

b. format of material



in written format



in computer readable form

c. time of filing/furnishing



contained in the international application as filed



filed together with the international application in computer readable form



furnished subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2004/001595

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 4956453 (Cetus Corporation)	Sep. 11, 1990	EP 0226418B1 JP 62209098A2	May 27, 1992 Sep. 14, 1987