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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 November 2003 (27.11.2003)

PCT

(10) International Publication Number
WO 03/097697 A2

(51) International Patent Classification[7]: C07K 16/18, C12N 15/09, A61K 48/00

(21) International Application Number: PCT/EP03/05324

(22) International Filing Date: 21 May 2003 (21.05.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/382,649	22 May 2002 (22.05.2002)	US
60/438,256	3 January 2003 (03.01.2003)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/382,649 (CIP)
Filed on	22 May 2002 (22.05.2002)
US	60/438,256 (CIP)
Filed on	3 January 2003 (03.01.2003)

(71) Applicant *(for all designated States except US)*: **ESBATECH AG** [CH/CH]; Wagistrasse 21, CH-8952 Schlieren (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TISSOT, Kathrin** [DE/CH]; Segantinistrasse 35, CH-8049 Zürich (CH). **EWERT, Stefan** [DE/CH]; Herbstweg 62, CH-8050 Zürich (CH). **AUF DER MAUR, Adrian** [CH/CH]; Herbstweg 110, CH-8050 Zürich (CH). **BARBERIS, Alcide** [CH/CH]; Allenmoosstrasse 64, CH-8057 Zurich (CH). **ESCHER, Dominik** [CH/CH]; Dorfstrasse 6, CH-8911 Rifferswil (CH).

(74) Agent: **E. BLUM & CO.**; Vorderberg 11, CH-8044 Zürich (CH).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

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

Published:

— *without international search report and to be republished upon receipt of that report*

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.

WO 03/097697 A2

(54) Title: IMMUNOGLOBULIN FRAMEWORKS WHICH DEMONSTRATE ENHANCED STABILITY IN THE INTRACELLULAR ENVIRONMENT AND METHODS OF IDENTIFYING SAME

(57) Abstract: Compositions are provided, which can be used as frameworks for the creation of very stable and soluble single-chain Fv antibody fragments. These frameworks have been selected for intracellular performance and are thus ideally suited for the creation of scFv antibody fragments or scFv antibody libraries for applications where stability and solubility are limiting factors for the performance of antibody fragments, such as in the reducing environment of a cell. Such frameworks can also be used to identify highly conserved residues and consensus sequences which demonstrate enhanced solubility and stability.

Immunoglobulin Frameworks which Demonstrate Enhanced Stability in the Intracellular Environment and Methods of Identifying Same

Field of the Invention

The invention relates to protein chemistry, molecular biology, and immunology.

Background of the Related Art

Antibodies can recognize and target almost any molecule with high specificity and affinity. This characteristic has been exploited to turn these natural proteins into powerful tools for diagnostic and therapeutic applications. Advances in recombinant DNA technology have facilitated the manipulation, cloning, and expression of antibody genes in a wide variety of non-lymphoid cells (Skerra, 1988; Martineau, 1998; Verma, 1998). A number of different antibody fragments have been constructed to best suit the various applications. The smallest entity that retains the full antigen-binding capacity of the whole parental immunoglobulin is the single-chain Fv fragment (scFv) (Bird, 1988). This antibody fragment comprises the variable regions of the heavy and the light chains linked by a flexible peptide-linker, which allows the expression of the protein from a single gene.

Antibody fragments have several important advantages in comparison to the entire immunoglobulin molecule. Due to their smaller size, the expression is facilitated and the yield is enhanced in a variety of expression host cells, such as E. coli cells (Plückthun, 1996). Moreover, antibody fragments allow improved tumour penetration in *in vivo* applications (Yokota, 1992) and they can be linked covalently to various effector molecules for therapeutic approaches.

Naturally occurring antibodies, which are secreted by plasma cells, have evolved to function in an extracellular, oxidizing environment. To obtain their functional, folded structure, they generally require the formation of disulfide-bridges within the separate domains, which are crucial for the stability of the immunoglobulin fold. In contrast to full-length antibodies, scFv or Fab antibody fragments can, in principle, be functionally expressed in a reducing environment inside any cell and directed to any compartment to target intracellular proteins and thus evoke specific biological effects (Biocca, 1991). Indeed, some intracellular single chain antibody fragments, which are called intrabodies, have been applied successfully to modulate the function of intracellular target proteins in different biological systems. Thus, resistance against viral infections has been demonstrated in plant biotechnology (Tavladoraki, 1993; Benvenuto, 1995),binding of intrabodies to HIV proteins has been shown(Rondon, 1997), and binding to oncogene products (Biocca, 1993; Cochet, 1998; Lener, 2000) has been described. Moreover, intracellular antibodies promise to be a valuable tool in characterizing the function of a vast number of genes now identified through the sequencing of the human genome (Richardson, 1995; Marasco, 1997). For example, they can be used in a functional genomics approach to block or modulate the activity of newly identified proteins, thereby contributing to the understanding of their functions. Finally, intrabodies have potential diagnostic and therapeutic applications, for example in gene therapy settings.

Despite these great prospects, the generation of functional intrabodies is still limited by their instability and insolubility or propensity to aggregate. The reducing environment of the cytoplasm prevents the formation of the conserved intrachain disulfide bridges, thus rendering a high percentage of antibody fragments unstable and, as a consequence, non-functional inside the

cell (Biocca, 1995; Proba, 1997). Stability and solubility of antibody fragments therefore represents a major obstacle for the application of intrabodies as potential modulators of protein function *in vivo*. So far, no predictions can be made about the sequence requirements that render an antibody fragment functional in an intracellular environment.

There is, therefore, a need for antibody fragments which perform well in a broad range of different cell types and can thus be used as frameworks for diverse binding specificities. Such frameworks can be used to construct libraries for intracellular screening or can serve as an acceptor for the binding portions of an existing antibody.

Besides being uniquely suited for intracellular applications, such antibody fragments or whole antibodies based on very stable variable domain frameworks also have a distinct advantage over other antibodies in numerous extracellular and *in vitro* applications. When such frameworks are produced in an oxidizing environment, their disulfide-bridges can be formed, further enhancing their stability and making them highly resistant towards aggregation and protease degradation. The *in vivo* half-life (and thus the resistance towards aggregation and degradation by serum proteases) is, besides affinity and specificity, the single-most important factor for the success of antibodies in therapeutic or diagnostic applications (Willuda, 1999). The half-life of antibody fragments can further be increased through the covalent attachment of polymer molecules such as poly-ethylene glycol (PEG) (Weir, 2002). Stable molecules of this type represent a significant advance in the use of antibodies, especially, but not exclusively, when the Fc functionality is not desired.

The great practical importance of antibody-fragment libraries has motivated research in this area. Winter (EP 0368684) has provided the initial cloning and

expression of antibody variable region genes. Starting from these genes he has created large antibody libraries having high diversity in both the complementary determining regions (CDRs) as well as in the framework regions. Winter does not disclose, however, the usefulness of different frameworks for library construction.

The teaching of Plückthun (EP 0859841), on the other hand, has tried to improve the library design by limiting the frameworks to a defined number of synthetic consensus sequences. Protein engineering efforts involving introduction of a large amount of rationally designed mutations have previously suggested mutations towards the respective consensus sequence as a suitable means for the improvement of the stability of isolated variable immunoglobulin domains (Ohage 1999; Ohage 1999 and US 5,854,027, hereby incorporated by reference).

Plückthun (EP 0859841) discloses methods for the further optimization of binding affinities based on these consensus sequences. The Plückthun patent also acknowledges the ongoing increase in knowledge concerning antibodies and accordingly aims at including such future findings in the library design. However, no possible further improvements of the synthetic consensus frameworks are suggested.

The teachings of Winter, Plückthun and others (e.g. Soderlind, WO 0175091) have thus tried to create large antibody libraries with a focus on high diversity in the CDRs for selection and application of the selected scFvs under oxidizing conditions. All of these libraries are, however, not optimized for intracellular applications and thus not useful for selection and applications in a reducing environment, or other conditions which set special requirements on stability and solubility of the expressed antibody fragment.

The qualities required for antibody fragments to perform well in a reducing environment, e.g. the cytoplasm of prokaryotic and eukaryotic cells, are not

clear. The application of intracellular antibodies or "intrabodies" is therefore currently limited by their unpredictable behavior under reducing conditions, which can affect their stability and solubility properties (Biocca, 1995; Wörn, 2000). Present patent applications (EP1040201, EP1166121 and WO0200729) and publications (Visintin, 1999) concerning intracellular screening for intrabodies focus on the screening technology but do not disclose specific antibody sequences which are functional in eukaryotic cells, in particular in yeast, and, thus, useful for library construction in this context.

Visintin and Tse have independently described the isolation of a so-called intracellular consensus sequence (ICS) (Visintin, 2002; Tse, 2002). This sequence was derived from a number of sequences that had been isolated from an antigen-antibody-interaction screen in yeast. The input into the intracellular screen was, however, heavily biased due to prior phage-display selection. Thus, all but one of the input-sequences belonged to the VH 3 subgroup in the case of Visintin et al. The published consensus sequence ICS is fully identical to the consensus sequence for the human VH 3 subgroup described by Knappik (2000) and EP0859841. 60 of the 62 amino acids of the ICS are also identical to the general human VH-domain consensus sequence which was proposed by Steipe as a basis for the construction of variable domains with enhanced stability (United States Patent No. 6,262,238, hereby incorporated by reference). These works were, in turn, based on earlier sequence collections (i.e., Kabat, 1991 and definitions of variable domain subgroups and structural determinants (Tomlinson, 1992; Williams, 1996; Chothia, 1989 and Chothia, 1987). However, because the input to the intrabody selection was so heavily biased (i.e., in the case of Visintin et al. all but one of the VH domains was VH3), the isolation of VH3 sequences from intracellular screening is not particularly surprising. Due to the heavy bias of their input

library, the work of Tse et al. and Visintin et al. does not provide a thorough evaluation of the human variable domain repertoire as would be provided by an unbiased inquiry and as is required to identify the useful intrabody frameworks present in the human repertoire.

We have previously described a system, which allows for the selection of stable and soluble intrabodies in yeast, independent of their antigen-binding specificity (Auf der Maur (2001), WO0148017). This approach allows efficient screening of scFv libraries and the isolation of specific frameworks, which are stable and soluble in the reducing environment of the yeast cell. The objective remains to actually isolate framework sequences and use the patterns in a first step to predict what sequence types would be most stable in the reducing environment and in a second step identify by analysis, recombination and further *in vivo* and *in vitro* experiments the optimal sequence.

Brief Summary of the Invention

The present invention fills a missing link in the field of antibody generation. It provides antibody variable domain framework sequences with superior characteristics regarding stability and solubility. These are crucial features for many relevant applications, such as in diagnostics, therapy or research. These frameworks can be used for grafting of existing binding-specificities or for the generation of antibody libraries with high stability and solubility.

ScFv libraries were used for the isolation of frameworks which are stable and soluble in the reducing environment of the yeast cell. The performance of the isolated frameworks has subsequently been characterized in human cell lines and in *in vitro* experiments. The described frameworks can directly serve as acceptor backbones for existing binding specificities or to construct

CDR libraries by randomization of one or more of the hypervariable loops for use in reducing or otherwise challenging environments. The isolated variable domain sequences have further been analyzed by alignment to identify preferred sequence families. From those preferred variable domain sequence families, optimal sequences were chosen based on a structural analysis which excludes sequences containing framework residues which disturb the immunoglobulin fold. The identified variable domain sequence candidates were subsequently recombined in all possible variations and the optimal combinations of variable domains of the light and heavy chain were selected by analysis of their performance in yeast, mammalian cells and *in vitro*.

These optimized scFvs and their constituting variable domain frameworks, as well as other antibody fragments or whole antibodies derived thereof, are ideal as, for example, acceptor backbones for existing binding specificities or for the construction of CDR libraries by randomization of one or more of the hypervariable loops for use in reducing or otherwise challenging environments. Antibodies suitable for intracellular applications are by definition more stable and soluble. Accordingly, their use will also be advantageous in applications outside the intracellular environment.

The invention provides compositions comprising frameworks of antibody variable domains and single-chain Fv antibody (ScFv) fragments which can be incorporated into various antibody fragments or whole antibodies. Classes of antibody variable domains fragments are provided which are the most stable and soluble and thus best suited for intracellular applications. Specific framework sequences of antibody variable domains and scFv antibody fragments which show the highest performance in intracellular assays are also provided. The invention also provides specific framework sequences of antibody variable domains and synthetic combinations of variable do-

mains of the light and heavy chain in scFv fragments which are, for example, optimal for intracellular applications and show an optimal performance in vitro regarding stability and solubility.

The invention provides single-chain framework reagents that have the general structures:

NH_2-VL-linker-VH-COOH or

NH_2-VH-linker-VL-COOH.

In another embodiment of the invention the single-chain framework may be fused to a second protein moiety to yield a fusion construct of the general structure:

NH_2-VL-linker-VH-second protein-COOH

NH_2-second protein-VL-linker-VH-COOH.

The orientation of the VH and VL regions in these fusion constructs may be reversed.

In another embodiment of the invention the variable domains may be incorporated into a Fab fragment, which may additionally be fused to a second protein moiety to yield fusion constructs of the general structure:

NH_2-VH-CH-second protein-COOH and NH_2-VL-CL-COOH

The second protein may be fused to either N- or C-terminus of either the heavy or the light chain.

In a preferred embodiment, the second protein of the single-chain or Fab framework fusion construct is a protein which provides a read-out for intracellular assays, either directly or via transcriptional activation.

Another object of the invention is to provide framework classes of antibody variable domains and sequences of variable domains and scFvs which are suitable for grafting the hypervariable loops from existing antibodies, for example, in order to obtain antibodies which are functional in a reducing or otherwise challenging environment.

Another object of the invention is to provide framework classes of antibody variable domains and sequences of variable domains and scFvs which, for example, through randomization of one or more of the hypervariable loops of such frameworks, are suitable for the creation of libraries for use in a reducing or otherwise challenging environment.

Another object of the invention is the use of the disclosed sequences in the identification of conserved residues and consensus sequences.

The antibodies or antibody fragments resulting from the use of the disclosed frameworks can be used as reagents in target validation and in therapy, prevention and diagnosis of human, animal and plant diseases. The antibodies can be used in the form of protein or DNA encoding such a protein and are not limited to intracellular applications.

Brief Description of the Drawings

Figure 1 shows the result of a typical "quality control" screen in yeast assayed by activation of lacZ expression (see, for example, Example 1). The selected, positive clones (black) were identified in several different screens and the corresponding sequences of the positive clones can be found in Tables 1 and 2. The selected sequences are compared to the positive control, the very stable lambda-graft (dark grey).

Figure 2 shows the performance of the frameworks isolated from a typical "quality control" screen in yeast (black) in the human cell line Hela, assayed by the activation of luciferase expression in comparison to the very stable lambda-graft (dark grey). The positive control Gal4-VP16 (white) gives the maximally possible level of transcriptional activation in the system. Luciferase activity has been corrected for transfection efficiency.

Figure 3 shows the *in vivo* performance of the superior framework combinations assayed in yeast by the activation of lacZ expression. The framework sequences (black) are compared to the positive control (the very stable lambda-graft (dark grey)). The numbering of the frameworks is as described in Table 5.

Figure 4 shows the *in vivo* performance of the superior framework combinations assayed in the human cell line Hela by the activation of luciferase expression and illustrated in comparison to the very stable lambda-graft (dark grey). The positive control, Gal4-VP16 (white) gives the maximal possible level of transcriptional activation in the system. Luciferase activity has been corrected for transfection efficiency.

Figure 5 shows the *in vivo* performance of the superior framework combinations assayed by the amount of soluble protein produced in the cytoplasm of yeast strain *S. cerevisiae* JPY9.

Figure 6 shows the expression behavior of selected framework combinations in the periplasm of *E.coli*. The arrow indicates the location of the band corresponding to the scFv frameworks.

Figure 7 shows the *in vivo* performance of selected superior framework combinations assayed in three human cell lines (Hela, (black), Saos-2 (dark grey) and HEK 293 (white)), by the activation of luciferase expression and illustrated in comparison to the very stable lambda-graft. The positive control Gal4-VP16 gives the maximal possible level of transcriptional activation in the system. Luciferase activity has been corrected for transfection efficiency.

Figure 8 represents the resistance towards aggregation at 37°C of selected framework combinations as quantified by the amount of monomeric protein present before and after incubation as indicated in PBS-buffer. Panel A is representative for frameworks 2.4 and 5.2 and panel B for frameworks 4.4, 6.4 and 7.3.

Figure 9 represents the resistance towards protease degradation aggregation in human serum at 37°C of selected framework combinations, quantified by the amount of soluble full-length protein present before and after prolonged incubation.

Figure 10 shows the in vivo performance of two selected binders on the novel framework 7.3 in the Fab-context, assayed in yeast interaction assay by the activation of lacZ expression. Expression of the Fab-chains is from a bi-directional galactose-inducible promoter, on either an ars/cen or a 2 micron vectors. Expression from the Fab vector yields the antibody light chain and a VH-CH1-Gal4-AD fusion protein. Binders are directed against human Polo-like kinase1 (hPLK1). Binding to the target is compared with the unspecific binding to an unrelated antigen and the binding of the un-randomized framework 7.3. Note that the corresponding scFv that have been included for reference are expressed from an actin promoter (2 micron).

Figure 11 shows the in vivo performance of the scFv frameworks in the Fab-context assayed by the amount of soluble protein produced in the cytoplasm of the yeast strain JPY9. Expression of the Gal4-AD-scFv fusion (actin/2 micron) is compared with the expression of the corresponding Fab-construct, and with the parent framework 7.3 as a Fab, both from two different vectors (Gal-inducible, ars/cen and 2 micron). Expression from the Fab vector yields the antibody light chain and a VH-CH1-Gal4-AD fusion protein, which is detected in this blot.

Table 1 shows an alignment of all VH-domain framework sequences selected from various "quality control" screens in yeast.

Table 2 shows an alignment of all VL-domain framework sequences selected from various "quality control" screens in yeast.

Table 3 shows an alignment of randomly picked sequences from the library.

Table 4 shows a statistical analysis of the sub-class frequency for VH- and VL-domains in the sequences isolated with the "quality control" system. Only those sequences were considered which were subsequently found to be positive in the quantitative yeast assay. The selected sequences are compared with the unselected library as determined from a limited number of random sequences (Table 3).

Table 5 shows the sequences used for further recombination and evaluation of the best combinations in scFvs and their respective abbreviations (abb.), sources and sub-family.

Detailed Description of the Invention

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

As used herein, "identity" refers to the sequence similarity between two polypeptides, molecules or between two nucleic acids. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit (for instance, if a position in each of the two DNA molecules is occupied by adenine, or a position in each of two polypeptides is occupied by

a lysine), then the respective molecules are homologous at that position. The "percentage identity" between two sequences is a function of the number of matching positions shared by the two sequences divided by the number of positions compared x 100. For instance, if 6 of 10 of the positions in two sequences are matched, then the two sequences have 60% identity. By way of example, the DNA sequences CTGACT and CAGGTT share 50% homology (3 of the 6 total positions are matched). Generally, a comparison is made when two sequences are aligned to give maximum homology. Such alignment can be provided using, for instance, the method of Needleman *et al.*, J. Mol Biol. 48: 443-453 (1970), implemented conveniently by computer programs such as the Align program (DNAstar, Inc.).

"Similar" sequences are those which, when aligned, share identical and similar amino acid residues, where similar residues are conservative substitutions for, or "allowed point mutations" of, corresponding amino acid residues in an aligned reference sequence. In this regard, a "conservative substitution" of a residue in a reference sequence is a substitution by a residue that is physically or functionally similar to the corresponding reference residue, *e.g.*, that has a similar size, shape, electric charge, chemical properties, including the ability to form covalent or hydrogen bonds, or the like. Thus, a "conservative substitution modified" sequence is one that differs from a reference sequence or a wild-type sequence in that one or more conservative substitutions or allowed point mutations are present. The "percentage positive" between two sequences is a function of the number of positions that contain matching residues or conservative substitutions shared by the two sequences divided by the number of positions compared x 100. For instance, if 6 of 10 of the positions in two sequences are matched and 2 of 10 positions contain conservative substitutions, then the two sequences have 80% positive homology.

"VH domain" refers to the variable part of the heavy chain of an immunoglobulin molecule.

"VL domain" refers to the variable part of the light chain of an immunoglobulin molecule.

VH or VL "subtype" refers to the subtype defined by the respective consensus sequence as defined in Knappik (2000). The term "subfamily" or " subclass" is used as synonym for "subtype". The term "subtype" as used herein refers to sequences sharing a high degree of identity and similarity with the respective consensus sequence representing their subtype. Whether a certain variable domain sequence belongs to a "subtype" is determined by alignment of the sequence with either all known human germline segments of the respective domain, or the defined consensus sequences and subsequent identification of the greatest homology. Methods for determining homologies and grouping of sequences by using search matrices, such as BLOSUM (Henikoff 1992) are well known to the person skilled in the art.

"Amino acid consensus sequence" as used herein refers to an amino acid sequence, which can be generated using a matrix of at least two or preferably more aligned amino acid sequences, and allowing for gaps in the alignment, it is possible to determine the most frequent amino acid residue at each position. The consensus sequence is that sequence which comprises the amino acids which are most frequently represented at each position. In the event that two or more amino acids are equally represented at a single position, the consensus sequence includes both or all of those amino acids.

The amino acid sequence of a protein can be analyzed at various levels. For example, conservation or variability could be exhibited at the single residue level, multiple residue level, multiple residue with gaps etc. Residues could exhibit conservation of the identical residue or could be conserved at the class level. Examples of amino acid classes include polar but uncharged R

groups (Serine, Threonine, Asparagine and Glutamine); positively charged R groups (Lysine, Arginine, and Histidine); negatively charged R groups (Glutamic acid and Aspartic acid); hydrophobic R groups (Alanine, Isoleucine, Leucine, Methionine, Phenylalanine, Tryptophan, Valine and Tyrosine); and special amino acids (Cysteine, Glycine and Proline). Other classes are known to one of skill in the art and may be defined using structural determinations or other data to assess substitutability. In that sense a substitutable amino acid could refer to any amino acid which could be substituted and maintain functional conservation at that position.

"Polynucleotide consensus sequence" as used herein refers to a nucleotide sequence, which can be generated using a matrix of at least two or preferably more aligned nucleic acid sequences, and allowing for gaps in the alignment, it is possible to determine the most frequent nucleotide at each position. The consensus sequence is that sequence which comprises the nucleotides which are most frequently represented at each position. In the event that two or more nucleotides are equally represented at a single position, the consensus sequence includes both or all of those nucleotides.

"Structural sub-element" as used herein refers to stretches of amino acid residues within a protein or polypeptide that correspond to a defined structural or functional part of the molecule. These can be loops (i.e. CDR loops of an antibody) or any other secondary or functional structure within the protein or polypeptide (i.e., domains, α-helices, β-sheets, framework regions of antibodies, etc.). A structural sub-element can be identified using known structures of similar or homologous polypeptides, or by using the above mentioned matrices of aligned amino acid sequences. Here the variability at each position is the basis for determining stretches of amino acid residues which belong to a structural sub-element (e.g. hypervariable regions of an antibody).

"Sub-sequence" as used herein refers to a genetic module which encodes at least one structural sub-element. It is not necessarily identical to a structural sub-element.

"Antibody CDR" as used herein refers to the complementarity determining regions of the antibody which consist of the antigen binding loops as defined by Kabat et al. (1991). Each of the two variable domains of an antibody Fv fragment contain, for example, three CDRs.

"Antibody" as used herein is a synonym for " immunoglobulin". Antibodies according to the present invention may be whole immunoglobulins or fragments thereof, comprising at least one variable domain of an immunoglobulin, such as single variable domains, Fv (Skerra, 1988), scFv (Bird, 1988; Huston, 1988), Fab , (Fab')2 or other fragments well known to a person skilled in the art.

"Antibody framework" as used herein refers to the part of the variable domain, either VL or VH, which serves as a scaffold for the antigen binding loops of this variable domain (Kabat et al., 1991).

Rationally engineered scFv fragments have demonstrated a clear correlation between the thermodynamic stability of a scFv fragment and its *in vivo* performance (Wörn, 2000; Auf der Maur, 2001). Using a recently developed system named "Quality Control" (Auf der Maur, 2001), specific antibody variable domain framework sequences which are suitable for intracellular applications have been isolated (Table 1 and 2), characterized (Fig. 1 and 2) and further improved (Fig. 3 to 9 and Table 3). As observed in our previous experiments, well performing frameworks selected in the intracellular assay show a high *in vitro* stability as demonstrated by their resistance to aggregation and protease degradation at 37°C (Fig. 8 and 9). Moreover, a pattern emerged which allows a selection of frameworks for intracellular applications on a more general basis, depending on their framework

subfamily (Table 4). Specific antibody variable domain sequences useful for intracellular applications are disclosed here, as well as the general pattern. This allows, on the one hand, the use of these sequences as framework donors in grafting experiments to obtain functional intrabodies which retain the binding specificity of the loop donor. Additionally, antibody libraries can be constructed using the disclosed sequences as frameworks. Such libraries are suitable for intracellular selection systems under reducing conditions, such as those in prokaryotic and eukaryotic cells. Additionally, the disclosed sequences may be used to identify, for example, conserved sequences or residues or motifs. The grafting of structural sub-elements, for example, those of the binding loops of an antibody (e.g. Jung, 1997), as well as the making of libraries of antibodies or fragments thereof (e.g. Vaughan, 1996; Knappik, 2000) has been described in detail and is well known to a person skilled in the art.

Because intracellular applications expose the antibody fragments to very unfavorable conditions (i.e. increased temperatures, reducing environment), the sequences disclosed in the present invention have acquired features that make them resistant to the most adverse conditions. Therefore, when compared to "average" sequences, the disclosed sequences are of outstanding stability and solubility as is demonstrated by their resistance towards aggregation and protease degradation (Fig.8 and 9). These features, together with their excellent expression yield make the disclosed antibody framework sequences uniquely suitable not only for intracellular use, but especially for all therapeutic and diagnostic applications where long half-life, robustness, and ease of production are of great concern.

The present invention enables the design of polypeptide sequences comprising at least the variable part of an antibody that are useful for applications in a

reducing or otherwise challenging environment. In a first embodiment, the invention provides a collection of antibody framework sequences useful for intracellular applications (Table 1 and 2). In a first step, a library of diverse sequences is screened independent of binding affinity using the Quality control system in yeast. The isolated sequences can be evaluated for their intracellular performance in yeast and in mammalian cells (Fig. 1 and 2).

In one embodiment of the invention, the collection of isolated sequences is analyzed by alignment to identify the antibody variable domain sub-classes and consensus sequences that are suitable for intracellular applications.

In a further preferred embodiment of the invention, the collection of antibody framework sequences described above is further analyzed by alignment to each other and grouping into sub-families. All frameworks belonging to one sub-type are compared regarding their intracellular performance in yeast and in mammalian cells (Fig. 1 and 2, as an example) and regarding the occurrence of negative, neutral or positive exchanges in their amino-acid sequence relative to the respective sub-type consensus. A person skilled in the art can distinguish between positive, neutral and negative changes based on the structural environment of the particular exchanged residue in the immunoglobulin domain. Subsequently, framework sequences of variable antibody domains are chosen which show the best intracellular performance and which are devoid of negative exchanges compared to their respective sub-type consensus. Preferably, sequences are selected which further contain amino-acid exchanges which are considered positive.

In a further preferred embodiment, the selected antibody variable domains of the heavy and the light chain are subsequently recombined in all possible combinations into scFv fragments, in order to identify the

combinations with the highest stability and solubility. To this end the novel, recombined scFv fragments are evaluated for their performance under reducing conditions in intracellular interaction assays in yeast (Fig. 3) and in mammalian cell lines (Fig. 4 and 7) and for soluble intracellular expression in yeast (Fig. 5). Promising combinations are further evaluated for their behavior under oxidizing conditions by analyzing the periplasmic expression yield in E.coli (Fig. 6), the resistance to aggregation at elevated temperatures (Fig. 8) and the resistance to aggregation and protease degradation upon prolonged incubation in human serum at 37°C (Fig. 9). These data are used to identify the scFv framework best suitable for any specific application, either intracellular, or under oxidizing conditions.

The selected and optimized framework sequences disclosed herein have a significant advantage not only in intracellular applications, but in all applications which can profit from increased stability and/or solubility of the scFv. Examples are the long-term storage at high concentrations required for diagnostic applications, and prolonged functional half-life in serum at 37°C (as required, for example, in therapeutic applications).

According to one aspect of the present invention, there is provided an intrabody framework comprising a single-chain framework having the general structure:

NH_2-VL-linker-VH-COOH; or

NH_2-VH-linker-VL-COOH

wherein the VH framework is of subtype 1a, 1b or 3.

In another embodiment, the orientation of the VH and VL regions is reversed in the single chain framework described above.

According to one aspect of the present invention, there is provided an intrabody framework comprising a single-chain framework having the general structure:

NH_2-VL-linker-VH-COOH; or

NH_2-VH-linker-VL-COOH

wherein the VH framework is of subtype 1a, 1b or 3 and the VL framework is of subtype λ1, λ3 or κ1.

In another embodiment, the invention provides a single-chain framework fused to a second protein moiety to yield a fusion construct of the general structure:

NH_2-VL-linker-VH-second protein-COOH; or

NH_2-second protein-VL-linker-VH-COOH

wherein the VH framework is of subtype 1a, 1b or 3 and the VL framework is of subtype λ1, λ3 or κ1.

In another embodiment, the orientation of the VH and VL regions in these fusion constructs may be reversed.

In another embodiment, the variable domains may be incorporated into a Fab fragment which may additionally be fused to a second protein moiety to yield fusion constructs of the general structure:

NH_2-VH-CH-second protein-COOH and NH_2-VL-CL-COOH

The second protein may be fused to either N- or C-terminus of either the heavy or the light chain.

As disclosed herein, there is a very strong preference in intracellular applications for VH framework of the subtype 3, but also for 1a and 1b. Regarding the light chain variable domain (VL), there is a clear preference by numbers for frameworks of the kappa 1 type, but lambda 1 and 3 are also enriched. These framework subtypes, i.e. VH 1a, 1b and 3 combined with a kappa 1, lambda 1or 3 VL domain are therefore best suited for intracellular use and other applications which require he folding properties of the scFv. Therefore, in order to reduce the amount of molecules which are not functional in the reducing environment, libraries for intracellular screening systems should preferentially be constructed from a mixture of these framework subtypes.

In a preferred embodiment, the VH domain of the antibody fragments of the invention is of the subgroup 1a, 1b or 3.

In a preferred embodiment, the VL domain of the antibody fragments of the invention is of the subgroup kappa1, lambda 1 or 3.

In a preferred embodiment, antibody fragments used as frameworks are selected from the group consisting of: 1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4, and 6.4 as described in Table 5.

In one embodiment of the invention, at least two and preferably more frameworks are identified and then analyzed. A database of the protein sequences may be established where the protein sequences are aligned with each other. The alignment can then be used to define, for example, residues, sub-elements, sub-sequence or subgroups of framework sequences which show a high degree of similarity in both the sequence and, if that information is available, in the structural arrangement.

The length of the sub-elements is preferably, but not exclusively ranging between 1 amino acid (such as one residue in the active site of an enzyme or a structure-determining residue) and 150 amino acids (for example, whole protein domains). Most preferably, the length ranges between 3 and 25 amino acids, such as most commonly found in CDR loops of antibodies.

In another embodiment, consensus nucleic acid sequences, which are predicted from the analysis are synthesized. This can be achieved by any one of several methods well known to the practitioner skilled in the art, for example, by total gene synthesis or by PCR-based approaches.

In another embodiment, the nucleic acid sequences are cloned into a vector. The vector could be a sequencing vector, an expression vector or a display (e.g. phage display) vector, all which are well known to

those of skill in the art. A vector could comprise one nucleic acid sequence, or two or more nucleic sequences, either in different or the same operon. In the last case, they could either be cloned separately or as contiguous sequences.

In one embodiment, the polypeptides have an amino acid pattern characteristic of a particular species. This can for example be achieved by deducing the consensus sequences from a collection of homologous proteins of just one species, most preferably from a collection of human proteins.

A further embodiment of the present invention relates to fusion proteins by providing for a DNA sequence which encodes both the polypeptide, as described above, as well as an additional moiety.

In further embodiments, the invention provides for nucleic acid sequences, vectors containing the nucleic acid sequences, host cells containing the vectors, and polypeptides obtainable according to the methods described herein.

In a further embodiment, the invention provides for synthesizing or otherwise placing restriction sites at the end of the nucleic acid sequences of the invention allowing them to be cloned into suitable vectors.

In a further preferred embodiment, the invention provides for vector systems being compatible with the nucleic acid sequences encoding the polypeptides. The vectors comprise restriction sites, which would be, for example, unique within the vector system and essentially unique with respect to the restriction sites incorporated into the nucleic acid sequences encoding the polypeptides, except for example the restriction sites necessary for cloning the nucleic acid sequences into the vector.

In another embodiment, the invention provides for a kit, comprising one or more of the list of nucleic acid sequences, recombinant vectors, polypeptides, and

vectors according to the methods described above, and, for example, suitable host cells for producing the polypeptides.

All of the above embodiments of the present invention can be effected using standard techniques of molecular biology known to one skilled in the art.

In another embodiment, the nucleic acid sequence is any sequence capable of encoding the polypeptides of the invention.

In another embodiment, the inventive nucleic acids are used in gene therapy.

In another embodiment, the single chain framework is a variant of any one of sequences 1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4, 6.4 (Table 5), where "variant" as used herein refers to a sequence that exhibits 90% or greater identity, while maintaining enhanced stability.

In another embodiment, the single chain framework is a derivative of any one of sequences 1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4, 6.4 (Table 5) where "derivative" as used herein refers to a sequence that maintains only those amino acids that are critical to the function and stability of the molecule. Isolated neutral or positive exchanges in the framework as described in example 3, are not considered to be relevant change to the antibody frameworks of the present invention.

In a preferred embodiment of the invention, the single chain framework is fused to a second protein, wherein that protein provides a read-out for intracellular assays. The read-out can be either direct, for example in the form of a fusion to a detectable protein, e.g. GFP (green fluorescent protein), enhanced blue fluorescent protein, enhanced yellow fluorescent, protein enhanced cyan fluorescent protein which can be observed by

fluorescence, or other fusion partners with different detection methods. Alternatively, a read-out can be achieved through transcriptional activation of a reporter gene, where the fusion partner in the scFv-fusion protein is either a transcriptional activator, such as the Gal4 activation domain, or a DNA-binding protein, such as the LexA- or Gal4 DNA-binding domain, which activates the transcription of a reporter gene of an enzyme, such as β-galctosidase, luciferase, α-galactosidase, β-glucuronidase, chloramphenicol acetyl transferase and others, which in turn provide a read-out. Fusion proteins, which provide a read out are well known to one of skill in the art.

Another embodiment of the invention is an antibody comprising a framework described herein.

Another embodiment of the invention is the use of the antibody of the instant invention.

A further preferred embodiment of the invention is the use of the described framework classes of antibody variable domains and sequences of variable domains and scFvs for grafting of hypervariable loops from existing antibodies, in order to obtain antibodies which are functional in a reducing or otherwise challenging environment.

Another further preferred embodiment of the invention is the use of the described framework classes of antibody variable domains and sequences of variable domains and scFvs, for example through randomization of one or more of the hypervariable loops of such frameworks, for the creation of libraries for applications in a reducing or otherwise challenging environment.

As would be apparent to one of ordinary skill in the art, the inventive molecules described herein may be used in diagnostic and therapeutic applications, target validation and gene therapy.

The invention may be illustrated by the following examples, which are not intended to limit the scope of the invention in any way.

References

Agatep, R., Kirkpatrick, D. L., Parchaliuk, R. A., Woods and Gietz, R. D. (1998). "Transformation of Saccharomyces cerevisiae by lithium acetate/single-stranded carrier DNA/polyethylene glycol protocol." *Technical Tips Online (http://tto.trends.com)*.

Auf der Maur, A., Escher, D. and Barberis, A. (2001). "Antigen-independent selection of stable intracellular single-chain antibodies." *FEBS Lett* 508: 407-412.

Benvenuto, E. and Tavladoraki, P. (1995). "Immunotherapy of plant viral diseases." *Trends Microbiol* 3(7): 272-5.

Biocca, S. and Cattaneo, A. (1995). "Intracellular immunization: antibody targeting to subcellular compartments." *Trends Cell Biol* 5: 248-252.

Biocca, S., Di Luzio, A., Werge, T. and Cattaneo, A. (1991). "Intracellular immunization: expression of antibody domains in the cytoplasm and in the nucleus of mammalian cells." *Cytotechnology* 5 Suppl 1: 49-50.

Biocca, S., Pierandrei-Amaldi, P. and Cattaneo, A. (1993). "Intracellular expression of anti-p21ras single chain Fv fragments inhibits meiotic maturation of xenopus oocytes." *Biochem Biophys Res Commun* 197(2): 422-7.

Biocca, S., Ruberti, F., Tafani, M., Pierandrei-Amaldi, P. and Cattaneo, A. (1995). "Redox state of single chain Fv fragments targeted to the endoplasmic reticulum, cytosol and mitochondria." *Bio/Technology* 13(10): 1110-5.

Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S. and Whitlow, M. (1988). "Single-chain antigen-binding proteins." *Science* 242(4877): 423-6.

Chothia, C. and Lesk, A. M. (1987). "Canonical structures for the hypervariable regions of immunoglobulins." *J. Mol. Biol.* 196(4): 901-17.

Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, W. R., Spinelli, S., Alzari, P. M. and Poljak, R. J. (1989). "Conformations of immunoglobulin hypervariable regions." *J.Mol. Biol.* 342: 877-883.

Johnson, G., Kabat, E. A. and Wu, T. T. (1996). kabat database of sequences of proteins of immunological interest. WEIR'S Handbook of experimantal Immunology I Immunochemistry and Molecular Immunology. Cambridge, MA, Blackwell Science Inv: 6.1-6.21.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. and Foeller, C. (1991). Sequences of proteins of immunological interest. NIH Publication 91-3242. Wasinghton, DC, US Department of health and human services.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. and Virnekas, B. (2000). "Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides." *J Mol Biol* 296(1): 57-86.

Cochet, O., Kenigsberg, M., Delumeau, I., Virone-Oddos, A., Multon, M. C., Fridman, W. H., Schweighoffer, F., Teillaud, J. L. and Tocque, B. (1998). "Intracellular expression of an antibody fragment-neutralizing p21 ras promotes tumor regression." *Cancer Res* 58(6): 1170-6.

Corpet, F. (1988). "Multalin. Multiple sequence alignment with hierarchical clustering." *Nucleic Acids Res* 16(22): 10881-10890.

Cox, J. P., Tomlinson, I. M. and Winter, G. (1994). "A directory of human germ-line V kappa segments

reveals a strong bias in their usage." *Eur J Immunol* 24: 827-836.

de Haard, H. J., van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderik, P., de Bruine, A. P., Arends, J.-W. and Hoogenboom, H. R. (1999). "a large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies." *J. Biol. Chem.* 274: 18218-18230.

Escher, D., Bodmer-Glavas, M., Barberis, A. and Schaffner, W. (2000). "Conservation of glutamine-rich transactivation function between yeast and humans." *Mol Cell Biol* 20(8): 2774-82.

Escher, D. and Schaffner, W. (1996). "Improved "activator trap" method for the isolation of transcriptional activation domains from random DNA fragments." *Biotechniques* 21(5): 848-54.

Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R. and Plückthun, A. (1997). "Reliable cloning of functional variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system." *J Immunol Methods* 201: 35-55.

Lener, M., Horn, I. R., Cardinale, A., Messina, S., Nielsen, U. B., Rybak, S. M., Hoogenboom, H. R., Cattaneo, A. and Biocca, S. (2000). "Diverting a protein from its cellular location by intracellular antibodies. The case of p21Ras." *Eur J Biochem* 267(4): 1196-205.

Marasco, W. A. (1997). "Intrabodies: turning the humoral immune system outside in for intracellular immunization." *Gene Ther* 4(1): 11-5.

Martineau, P., Jones, P. and Winter, G. (1998). "Expression of an antibody fragment at high levels in the bacterial cytoplasm." *J. Mol. Biol.* 280(1): 117-27.

Ohage, E. and Steipe, B. (1999). "Intrabody construction and expression. I. The critical role of VL domain stability." *J Mol Biol* 291(5): 1119-28.

Ohage, E. C., Wirtz, P., Barnikow, J. and Steipe, B. (1999). "Intrabody construction and expression. II. A synthetic catalytic Fv fragment." *J Mol Biol* 291(5): 1129-34.

Plückthun, A., Krebber, A., Krebber, C., Horn, U., Knüpfer, U., Wenderoth, R., Nieba, L., Proba, K. and Riesenberg, D. (1996). Producing antibodies in Escherichia coli: from PCR to fermentation. Antibody Engineering, a practical approach. Oxford, Oxford University Press.

Proba, K., Honegger, A. and Pluckthun, A. (1997). "A natural antibody missing a cysteine in VH: consequences for thermodynamic stability and folding." *J. Mol. Biol.* 265(2): 161-72.

Richardson, J. H. and Marasco, W. A. (1995). "Intracellular antibodies: development and therapeutic potential." *Trends Biotechnol* 13(8): 306-10.

Rondon, I. J. and Marasco, W. A. (1997). "Intracellular antibodies (intrabodies) for gene therapy of infectious diseases." *Annu Rev Microbiol* 51: 257-83.

Skerra, A. and Plückthun, A. (1988). "Assembly of a functional immunoglobulin Fv fragment in Escherichia coli." *Science* 240(4855): 1038-41.

Tavladoraki, P., Benvenuto, E., Trinca, S., De Martinis, D., Cattaneo, A. and Galeffi, P. (1993). "Transgenic plants expressing a functional single-chain Fv antibody are specifically protected from virus attack." *Nature* 366(6454): 469-72.

Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. and Winter, G. (1992). "The repertoire of human germline VH sequences revealy about 50 groups of VH segments with different hypervariable loops." *J Mol Biol* 227: 776-798.

Tse, E., Lobato, M. N., Forster, A., Tanaka, T., Chung, G. and Rabbitts, T. (2002). "Intracellular antibody capture technology: application to selection of

intracellular antibodies recognising the BCR-ABL oncogenic protein." *J. Mol Biol* 317((1)): 85-94.

Verma, R., Boleti, E. and George, A. J. (1998). "Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems." *J Immunol Methods* 216(1-2): 165-81.

Visintin, M., Settanni, G., Maritan, A., Graziosi, S., Marks, J. D. and Cattaneo, A. (2002). "The intracellular antibody capture technology (IACT): towards a consensus sequence for intracellular antibodies." *J. Mol. Biol.* 317((1)): 73-83.

Visintin, M., Tse, E., Axelson, H., Rabbitts, T. H. and Cattaneo, A. (1999). "Selection of antibodies for intracellular function using a two-hybrid in vivo system." *Proc Natl Acad Sci U S A* 96(21): 11723-8.

Welschhof, M., Ternesss, P., Kolbinger, F., Zewe, M., Dübel, S., Dörsam, H., Hain, C., Finger, M., Jung, M., Moldenhauer, G., Hayashi, N., Little, M. and Opelz, G. (1995). "Amino acid sequence based PCR primers for amplification of rearranged human heavy and light chain immunoglobulin variable region genes." *J Immunol Methods* 179: 203-214.

Williams, S. C., Frippiat, J. P., Tomlinson, I. M., Ignatovic, O., Lefranc, M. P. and Winter, G. (1996). "Sequence and evolution of the human germline V-lambda repertoire." *J Mol Biol* 264: 220-232.

Wörn, A., Auf der Maur, A., Escher, D., Honegger, A., Barberis, A. and Plückthun, A. (2000). "Correlation between in vitro stability and in vivo performance of anti-GCN4 intrabodies as cytoplasmic inhibitors." *J. Biol. Chem.* 275(4): 2795-803.

Yokota, T., Milenic, D. E., Whitlow, M. and Schlom, J. (1992). "Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms." *Cancer Res* 52(12): 3402-8.

Auf der Maur, A., Zahnd, C., Fischer, F., Spinelli, S., Honegger, A., Cambillau, C., Escher, D.,

Plückthun, A. and Barberis, A. (2002). „Direct *in vivo* screening of intrabody libraries constructed on a highly stable single-chain framework." *J Biol Chem* 277(47): 45075-45085.

Gietz, R.D. and Sugino, A. (1988). "New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites" *Gene* 74: 527-534.

Willuda, J., Honegger, A.; Waibel, P., Schubiger, A. Stahel, R., Zangmeister-Wittke, U. and Plückthun, A. (1999). " High thermal stability is essential for tumor targeting of antibody fragments: engineering of a humanized anti-epithelial glycoprotein-2 (epithelial cell adhesion molecule) single-chain Fv fragment" *Cancer Research* 59: 5758-5767.

Wörn, A. and Plückthun, A. (1999). "Different equilibrium stability behavior of scFv fragments: identification, classification, and improvement by protein engineering" *Biochemistry* 38: 8739-8750.

The invention is further illustrated in the following non-limiting examples.

Example 1

Selection of intrabody frameworks through screening of a human library in the "quality control" system in yeast

Screening with the "quality control" system for stable frameworks was essentially performed as described in detail by Auf der Maur (WO0148017, Auf der Maur 2001, each hereby incorporated by reference).

The plasmids for expression of the scFv-fusion constructs for screening in yeast were derived from pESBA-Act (Wörn, 2000). It contains the yeast TRP1 gene for transformation selection in *S. cerevisiae* and

the 2 micron origin of replication to ensure high copy numbers. Moreover it has a constitutive actin promoter for strong expression and the GAL11 transcriptional termination sequence, separated by a multiple cloning site. For handling in bacterial systems, it also has a bacterial origin of replication and the amp resistance gene.

The Gal4 activation domain (AD amino acids 768-881) was originally amplified by PCR using pGAD424 (Clontech) as template with primers including the SV40 T-antigen nuclear localization signal N-terminal to the Gal4-AD. The DNA-fragments encoding amino acids 263-352 of Gal11P were amplified by PCR and cloned in frame, N-terminal to the SV40-NLS-Gal4-AD-construct. The human scFv library, amplified from human spleen-cell cDNA as described elsewhere (Welschhof, 1995; Krebber, 1997; de Haard, 1999), was cloned in frame, N-terminal to this fusion construct via SfiI-sites, and in the orientation V_L-linker-V_H where the linker has the sequence $(GGGS)_4$. Expression thus yields a fusion protein of the general structure scFv-Gal11p-SV40 NLS-Gal4AD.

Screening was carried out in the yeast strain *S. cerevisiae* YDE172 (*MATα ura3-52 leu2Δ1 trp1d63 his3*Δ200 *lys2Δ 385 gal4Δ 11*) (Auf der Maur, 2001), which was derived from the strain JPY9 (Escher, 2000) by integrating the divergently oriented LacZ and HIS3 reporter genes under the control of the natural UAS_G from Gal1-GAL10 regulatory sequences into the *his3Δ200* locus. Transcriptional activation of the reporter system is mediated by the Gal4-AD moiety of the scFv-fusion construct, following the specific interaction of its Gal11P moiety with the Gal4-DNA-binding-domain (DBD, amino acids 1-100). The Gal4-DBD is provided by expression from a second plasmid, pMP83. It contains the yeast LEU2 gene for transformation selection in *S. cerevisiae* and the ARS CEN origin of replication. Moreover, it has a constitutive actin promoter for strong expression and the GAL11 transcriptional termination sequence. For handling in

bacterial systems, it also has a bacterial origin of replication and the amp resistance gene.

For screening, the yeast strain *S. cerevisiae* YDE172 was co-transformed with a scFv-library as fusion construct on the pESBA-Act2 vector while the pMP83-vector provided the Gal4-DBD. A standard lithium acetate transformation protocol was used (Agatep, 1998). Following transformation, the cells were plated on drop-out plates (-Trp/-Leu/-His) containing 80 mM 3-aminotriazole. Colonies were picked after 3 days incubation at 30°C and re-streaked on drop-out plates (-Trp/-Leu/-His) containing 80 mM 3-aminotriazole. Those that re-grew were tested for LacZ-expression by development of blue color in a filter assay on plates containing the substrate X-Gal. Positive clones were taken for further analysis involving isolation of the scFv-carrying plasmid from yeast, transformation into *E.coli* DH5α, isolation of plasmid from single colonies of E.coli and re-transformation into freshly prepared yeast strain *S. cerevisiae* YDE172 for the assay as described below. All methods were performed according to standard procedures, well known to a person of ordinary skill in the art.

In addition, a modified screening procedure was used were the scFv was directly fused to both a DNA-binding domain (LexA amino acids 1-202) and an activation domain (Gal4, amino acids 768-881) to yield a fusion construct of the following structure: scFv-LexA-NLS-Gal4AD. The plasmids for expression of the scFv-fusion constructs for screening in yeast were derived from pESBA-Act2. It contains the yeast TRP1 gene for transformation selection in *S. cerevisiae* and the 2 micron origin of replication to ensure high copy numbers. Moreover, it has a constitutive actin promoter (for strong expression) and the GAL11 transcriptional termination sequence separated by a multiple cloning site. For handling in bacterial systems, it also has a bacterial origin of replication and the amp resistance gene.

Screening was carried out in the yeast strain *S. cerevisiae* ImmunaLHB (*MATα ura3-52 leu2Δ1 trp1d63 his3Δ200 lys2Δ 385*) which was derived from the strain JPY5 by integrating the divergently oriented LacZ and HIS3 reporter genes under the control of a bi-directional promoter with six LexA-binding sites (integrating reporter plasmid pDE200, Escher 2000) into the *his3Δ200* locus and by integrating the LEU2 reporter gene under the control of a promoter with eight LexA-binding sites (derived from EGY48) into the *leu2Δ1* locus. Transcriptional activation of the reporter system is mediated by the Gal4-AD moiety of the scFv-fusion construct. Screening was carried out essentially as described above using drop-out medium (-Trp/-Leu/-His) and 3-aminotriazole concentrations up to 40 mM.

Example 2

Evaluation of *in vivo* performance

a) in yeast

For quantitative analysis of the performance of the selected frameworks in yeast (Fig.1 and 3), *S. cerevisiae* -strain Immuna LHB was transformed with the isolated scFvs as LexA-Gal4-AD-fusion constructs on the pESBA-Act2 vector by following a standard lithium acetate transformation protocol (Agatep, 1998). Following transformation, the cells were plated on drop-out plates (-Trp). 2 ml overnight-cultures in drop-out medium (-Trp) were inoculated in duplicates from streaks containing several colonies and grown at 30°C. Cultures were diluted in 1 ml drop-out medium (-Trp) to an optical density at 600 nm (OD600) of 0.7. They were then grown at 30 °C for 2h. For the assay, 100 µl cell culture were taken, mixed with 900 µl buffer, 45 µl Chloroform and 30 µl 0.1% SDS, vortexed and incubated at room temperature for 5 minutes. The color development was initiated by the addition of 0.2 ml ONPG (4 mg/ml) and stopped with 0.5 ml Na_2CO_3 (1 M). The activity was calculated by taking into

account the OD600 of the assay culture, as well as the incubation time of the color development and the culture volume used

Clones that were at least equal to or better than the positive control (the very stable lambda-graft described before (Wörn, 2000; Auf der Maur, 2001)) were sequenced to identify the framework subtype (framework subtype definitions according to Tomlinson, (1992), Cox, (1994) and Williams, (1996)). Sequencing revealed a striking preference for certain framework subtypes. For the heavy chain variable domain (VH), framework subtypes 2 and 6 were never found and 4 was markedly reduced among the positive clones. Corrected for the performance of the isolated sequences in the yeast intracellular assay, there is a very strong preference for VH framework of the subtype 3, but also for 1a and 1b in intracellular applications. Regarding the light chain variable domain (VL), there is a clear preference for frameworks of the kappa 1, lambda 1 and lambda 3 sub-types (Table 4).

These framework subtypes, i.e. VH 1a, 1b and 3 combined with a kappa 1, lambda 1 and lambda 3 VL domain are therefore best suited for intracellular use and other applications with stringent requirements concerning the folding properties of the scFv. Libraries for intracellular screening systems should, for example, preferentially be constructed from a mixture of these framework subtypes only, to reduce the amount molecules which are not functional in the reducing environment.

b) in mammalian cells

Hela cell line was used for quantitative analysis of the performance of the selected frameworks in human cells (Fig. 2, 4 and 7). The luciferase reporter gene was provided from a co-transfected pGL3 (Promega) reporter plasmid containing the luciferase under the control of the natural Gal4 UAS. The mammalian expression vectors used for transient transfection contains the Gal4

(1-147) fused on the C-terminus to the VP16-AD under the control of a CMV promoter. The isolated scFvs were cloned in frame, C-terminal to a Gal4(1-147)-VP16-fusion to yield a Gal4(1-147)-VP16-scFv-fusion protein upon expression. Cells were cultured in DMEM supplemented with 2.5% FCS and 2 mM l-glutamine. Transient transfections were carried out according to the Polyfect-protocol (Qiagen) in 60 mm tissue culture plates using 0.01-0.1 µg of the vector containing the scFv-construct, 0.5 µg of a CMV promoter-driven Gal4(1-147)-VP16-scFv expression plasmid and 0.5 µg of a LacZ expression vector as reference for transfection efficiency. Cells were harvested 24-48 hours after transfection, resupended in 1000 µl buffer and lysed by three freeze-thaw-cycles. The cell lysate was centrifuged and the supernatant assayed for luciferase activity using luciferase assay solution (Promega) and for LacZ activity according to the standard protocol. The obtained luciferase activity was corrected with the LacZ activity to account for the variation in transfection efficiency.

Example 3

Multiple alignment and analysis of the sequence comparison

To elucidate the general pattern of framework sequences suitable for intracellular applications, all positive clones (i.e. those that grow under selective conditions in the quality control system) were isolated and the part coding for the scFvs was sequenced. Subsequently, the scFv sequences were divided in their light and heavy-chain component to allow alignment of the respective domains (Tables 1 and 2) according to the structural adjusted numbering scheme of immunoglobulin domains by Honegger (2001).

To allow evaluation of the obtained data, an alignment representing the unselected library was generated (Table 3). In order to obtain unselected sequences,

the library was transformed in E.coli cells which do not express the scFv-genes and clones were picked at random for plasmid isolation and sequencing of the scFv-sequence. The library covers the human antibody repertoire as expected and thus has no bias towards specific subgroups, other than expected by the expression pattern generally found in humans.

The VH and VL sequences were grouped according to their subgroup. Changes to the subgroup-specific consensus sequence were highlighted. A person skilled in the art can distinguish between positive, neutral and negative changes based on the structural environment of the particular exchanged residue (e.g. Honegger, 2001). An exchange of a residue belonging to a particular group of amino acids to a residue of the same group is in general validated as a neutral exchange. An exchange of a residue belonging to the group of hydrophobic amino acid pointing into the hydrophobic core of the protein to one amino acid of the group of polar but uncharged or positively or negatively charged amino acids would be highly unfavorable because unsatisfied hydrogen donor/acceptor sites disturb tight packing of the hydrophobic core. Such a change is therefore considered negative. An exchange of a residue belonging to the group of polar but uncharged residues at the surface of the immunoglobulin domain to an amino acid of the group of positively or negatively charged residues is highly favorable as the solubility of the protein is increased. Such a change is therefore validated positively, whereas the exchange from a polar to a hydrophobic residue is highly unfavorable as the solubility of the protein is decreased and is therefore validated negatively. At positions with a conserved positive phi-angle, an exchange of any amino acid to glycine is validated positively whereas an exchange of gylcine to any amino acid is validated negatively because glycine is the only amino acid which is able to form a positive phi-angle. The loss of a conserved salt bridge between posi-

tions 45 - 53, 45 - 100, 77 - 100 and 108 - 137 because of an exchange from an amino acid of the group of positively or negatively charged residues to an uncharged amino acid results in a decreased thermodynamic stability and is therefore considered negative.

Finally, we chose 7 VL domains and 4 VH domains that were preferentially selected during the quality control (i.e. showing the least negative and most positive exchanges from the consensus sequence and cover the subgroups) and that each show high *in vivo* performance in yeast. The sequences are summarized in Table 5 and include two Vκ1 (k I 27 (1.x) and k III 25(2.x)), two Vκ3 (k IV 103 (3.x) and k IV135 (5.x)), one Vλ1 (k IV 107 (4.x)), two Vλ3 (a33 (7.x) and a43 (6.x)), one VH1b (a33 (x.3)) and three VH3 (a fw10 (x.2), a43 (x.4) and a44 (x.1)). These VL and VH domains were shuffled giving 22 novel combinations in the scFv format (1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4, 6.4).

Example 4

Evaluation of *in vivo* performance of shuffled domains

a) Performance in an intracellular assay in yeast and mammalian cells

The 22 combinations were tested for their in vivo performance in yeast and mammalian cells as described in example 2 (Fig. 3 and 4).

b) Expression of soluble protein under reducing conditions in yeast

To compare the yields of soluble protein upon expression under reducing conditions, the selected frameworks were expressed as a fusion to Gal4 AD in the cytoplasm of yeast *S. cerevisiae*. The fusion constructs on the pESBA-Act2 vector had the general structure Gal4 AD-scFv. They were transformed as described above into

the yeast *S. cerevisiae strain* JPY9 and plated on -Trp, drop-out plates.

5 ml overnight-cultures in drop-out medium (-Trp) were inoculated from streaks containing several colonies and grown at 30°C. Cultures were diluted in 50 ml drop-out medium (-Trp) to an optical density at 600 nm (OD600) of 0.5. They were grown at 30 °C for 5h. For the native cell extract, 2.5 ml cell culture normalized to an OD600 of 3 were harvested by centrifugation, frozen in liquid nitrogen and subsequently resuspended in 75µl Y-PER (Pierce) containing protease inhibitor (PMSF). The resuspended cell pellet was vortexed shortly and incubated (slightly shaking) at 20°C for 20 min. Insoluble and aggregated material was pelleted at maximal speed in an eppendorf centrifuge at 4°C for 10 min. The supernatant was mixed with loading dye, heated to 100°C for 5 min. and separated on a 12% SDS-PAGE. The soluble Gal4 AD-scFv fusion constructs were visualized by western blotting via detection of the Gal4-moiety with an anti-Gal4AD monoclonal mouse antibody (Santa Cruz Biotechnology) as a primary antibody and an anti-mouse-peroxidase conjugate (Sigma) as secondary antibody and using a chemoluminescent substrate (Pierce) (Fig. 5). SDS-PAGE and western blotting procedures are well known to a person of ordinary skill in the art.

c) Expression behavior in the periplasm of E.coli

For evaluation of periplasmic expression behavior in *E. coli* (Fig. 6), isolated scFvs-frameworks were cloned in a bacterial vector harbouring the cam resistance gene (*catR*) and the *lacI* repressor gene (Krebber, 1997), with a N-terminal *pelB*-leader sequence and a C-terminal his-tag under the control of the *lac* promoter/operator. Competent E.coli JM83 were transformed with these plasmids. 50 ml dYT-medium containing 35 mg/l chloramphenicol in shaking flasks was inoculated 1:40 with an

over-night culture and incubated at 30°C. Cells were induced at an OD600 of 0.8 with 1mM IPTG and harvested after 3 hours of induction by centrifugation. The pellet was resuspended in 50 mM Tris, pH 7.5, 500 mM NaCl and normalized to an OD600 of 10. Samples of each scFv fragments were analyzed either directly (total extract) or after sonification followed by centrifugation (soluble fraction) by SDS-PAGE. The amount of soluble protein was estimated from the Coomassie-stained gel.

Example 5

Detailed evaluation of 5 combinations with superior properties for extracellular use

Five combinations were chosen as examples which show good performance both in yeast and mammalian intracellular assays, yield soluble protein during expression in yeast and E.coli, and cover the subgroups which were preferentially selected during the quality control (2.4, 4.4, 5.2, 6.4 and 7.3, see Table 5 for details). We analysed these combinations in greater detail to further evaluate their use under reducing, as well as oxidizing conditions.

a) Performance in an intracellular assay in different mammalian cells

The quantitative analysis of the performance of the five combinations in human cells was carried out using Hela cells and in addition using the human osteosarcoma cell line Saos-2 and the human embryonal kidney cell line HEK293 as performed in Example 2 (Fig. 7).

b) Performance *in vitro*

Expression and purification

For evaluation of the *in vitro* performance, the five superior combinations were expressed in the periplasm of E.coli (Fig. 6). The amount of 0.1 l dYT-medium containing 35 mg/l chloramphenicol in shaking

flasks was inoculated 1:40 with an over-night culture and incubated at 30°C. Cells were induced at an OD550 of 1.5 with 1mM IPTG and harvested after 2 hours of induction by centrifugation. For purification of the scFvs, the cell pellet was resuspended and lysed by sonication. Following centrifugation in SS34 at 20krpm, 4°C for 30 minutes, the supernatent was applied to a Ni-MC-affinity column (Hi-Trap™ Chelating HP, 1ml, Amersham Pharmacia) at pH 7.5 and eluted with 200 mM imidazol using an Äkta Basic system from Amersham Pharmacia. The purity of the scFv fragments was greater than 98% as determined by SDS-PAGE (data not shown). The concentration of the purified protein was determined using the calculated extinction coefficient at 280 nm. The yield of soluble purified protein was normalized to a culture volume of 1 l with an OD600 of 10 and varied from 8 to over 55 mg.

Resistance to aggregation

Resistance towards aggregation has been shown to correlate with thermodynamic stability (Wörn, 1999) *in vitro* and the efficiency of tumor localization in a xenografted tumor model in mice (Willuda, 1999). In order to test for the stability, resistance to aggregation and reversibility of unfolding, 200 µl samples of the purified proteins at concentrations of 6 µM in 50 mM Tris, pH 7.5, 100 mM NaCl were either kept 4 days at 4°C or 4 days at 37°C or 3 days at 4°C followed by an incubation of 15 or 60 minutes at 100°C, slow cooling down to room temperature and an overnight incubation at 4°C. The oligomeric state of each sample was subsequently analyzed on a gel filtration column equilibrated with 50 mM Tris, pH 7.5, 100 mM NaCl to estimate the amount of aggregated versus monomeric material (Fig. 8). The proteins were injected on a Superdex-75 column (Amersham Pharmacia) in a volume of 100 µl and a flow-rate of 1 ml/min on a Äkta Basic system (Amersham Pharmacia).

Resistance to protease degradation

To determine the stability of the isolated frameworks towards protease degradation, a parameter that is important for therapeutic applications, we incubated the purified frameworks in human serum at 37°C (Fig. 9).

Purified, his-tagged scFv-protein (see above) at a concentration of 50 µM was diluted tenfold into human serum to a final concentration of 5 µM in 90% serum. The samples were then either incubated at 37°C for either 3 days or 1 day, or taken directly for loading. Before loading insoluble and aggregated material was pelleted at maximal speed in an eppendorf centrifuge at 4°C for 10 min. The supernatant was diluted six-fold with a loading dye to reduce the amount of serum loaded on the gel, heated to 100°C for 5 min. and separated on a 12% SDS-PAGE. The soluble his-tagged scFv fragments were visualized by western blotting via detection of the his-tag with an anti-his monoclonal mouse antibody (Qiagen) as primary and an anti-mouse-peroxidase conjugate (Sigma) as secondary antibody and using a chemoluminescent substrate (Pierce). SDS-PAGE and western blotting procedures are well known to a person of ordinary skill in the art.

Example 6

Selection of antigen binders through screening of a randomized CDR-library on the framework 7.3 in the interaction screening system in yeast

Screening with the interaction system for antigen binders was essentially performed as described in detail before (Auf der Maur, 2002).

The plasmids for expression of the scFv-fusion constructs for screening in yeast were derived from pESBA-Act2. It contains the yeast TRP1 nutritional marker and the 2 micron origin of replication. Moreover it has a constitutive actin promoter for strong expression and the GAL11 transcriptional termination sequence,

separated by a multiple cloning site. For handling in bacterial systems, it also has a bacterial origin of replication and the amp resistance gene.

The Gal4 activation domain (AD amino acids 768-881) was originally amplified by PCR using pGAD424 (Clontech) as template with primers including the SV40 T-antigen nuclear localization signal N-terminal to the Gal4-AD. The scFv library was obtained by PCR-amplification of the scFv-framework 7.3 using primers randomizing 7 amino acids within the CDR3 of VH. The resulting PCR-product was cloned in the framework 7.3, present in the vector in the orientation V_L-linker-V_H, as a C-terminal fusion to Gal4-AD. Expression thus yields a fusion protein of the general structure Gal4-AD-scFv.

Screening was carried out in the yeast strain *S. cerevisiae* Immuna LHB (*MATα ura3-52 leu2Δ1 trp1d63 his3Δ200 lys2Δ 385*). It was derived from the strain JPY5 by integrating the divergently oriented LacZ and HIS3 reporter genes under the control of a bi-directional promoter with six LexA-binding sites (integrating reporter plasmid pDE200, Escher 2000) into the *his3Δ200* locus and by integrating the LEU2 reporter gene under the control of a promoter with eight LexA-binding sites (derived from EGY48) into the *leu2Δ1* locus.

Transcriptional activation of the reporter system is mediated by the Gal4-AD moiety of the scFv-fusion construct, following the specific interaction of its scFv moiety with the antigen-moiety of the bait-fusion protein. The bait-fusion protein consists of the kinase domain of the human polo-like kinase 1 (hPlk1-KD) fused C-terminal to the DNA-binding LexA protein. The kinase domain (amino acid 2-332) was PCR amplified from a hPlk1 cDNA using the upstream primer 5'-tgctctagaagt gctgcagtgactgcag-3' (Seq. Id.No. 12) and downstream primer 5'-ggttgtcgacttacaggctgctgggagcaatcg-3' (Seq. Id. No.13). The resulting PCR product was cloned C-terminal of LexA via *XbaI* and *SalI* into the bait vector. The bait

vector contains the URA3 nutritional marker and an Ars Cen origin of replication. Expression of the bait-fusion protein is driven by a constitutively active actin promoter. Transcription is terminated by the GAL11 termination sequence. The bait vector also carries a bacterial origin of replication and the amp resistance gene for propagation in bacterial systems.

For screening the yeast strain *S. cerevisiae* Immuna LHB was co-transformed with a scFv-library as fusion to Gal4-AD on the pESBA-Act2 vector and the bait-vector providing the LexA-hPLK1-KD fusion by following a standard lithium acetate transformation protocol (Agatep, 1998). Following transformation, the cells were plated on drop-out plates (-Trp/-Leu/-Ura). Colonies were picked after 3 to 5 days incubation at 30°C and restreaked on drop-out plates (-Trp/-Leu/-Ura). Those that re-grew were tested for LacZ expression by development of blue color in a filter assay on plates containing the substrate X-Gal. Positive clones were taken for further analysis involving isolation of the scFv-carrying plasmid from yeast, transformation into *E.coli* DH5α, isolation of plasmid from single colonies of E.coli, sequencing and re-transformation into freshly prepared yeast strain *S. cerevisiae* Immuna LHB for the assay as described below. All methods were performed according to standard procedures, well known to a person of ordinary skill in the art.

Example 7

Evaluation of *in vivo* performance of Fab-constructs derived from novel scFv frameworks

To evaluate the beneficial effect of using stable variable domain frameworks on different antibody formats, Fab expression vector were constructed for use in the yeast interaction screen.

a) Fab constructs for intracellular screening in yeast

Two different expression vectors were constructed to allow different expression levels. The vectors are based on either yEplac 112 (2 micron) or yCplac22 (ars/cen) backbones (Gietz, 1988). Both contain the yeast TRP1 nutritional marker, an inducible, bi-directional Gal1/Gal10 promoter, a bacterial origin of replication and the amp resistance gene for handling in bacterial systems. In one direction, the VH domain of the framework 7.3 was cloned N-terminal to the CH1-domain of IgG1 including the C-terminal cysteine, followed by a linker and the Gal4 activation domain (AD amino acids 768-881) including the SV40 T-antigen. On the other side, the VL domain of the framework 7.3 was cloned N-terminal to the CL (lambda)-domain including the C-terminal cysteine. The terminators are Gal11 terminator on the side of the heavy chain and Cyclin 1 terminator on the side of the light chain.

b) Performance in an intracellular assay in yeast

For quantitative analysis of the performance of the antigen binders in scFv and Fab format in yeast (Fig.1 and 3), *S. cerevisiae* strain Immuna LHB was co-transformed with the isolated scFvs as Gal4-AD-fusion constructs on the pESBA-Act2 vector and the bait vector containing the LexA-hPLK1-KD fusion by following a standard lithium acetate transformation protocol (Agatep, 1998). Following transformation, the cells were plated on drop-out plates (-Trp, -Ura, Glc). 2 ml overnight-cultures in drop-out medium (-Trp, -Ura, Glc) were inoculated in duplicates from streaks containing several colonies and grown at 30°C. Cultures were diluted in 1 ml drop-out medium (-Trp, -Ura, Gal) to an optical density at 600 nm (OD600) of 0.7. They were grown at 30 °C for 5h. The assay was carried out as described above.

c) Expression of soluble protein under reducing conditions in yeast

To compare the yields of soluble protein upon expression under reducing conditions, the scFv and Fab constructs, together with the hPLK1-KD-bait vector, as described above were expressed in the cytoplasm of yeast *S. cerevisiae*. They were transformed as described above into the yeast *strain* YDE173 and plated on -Trp, -Ura, drop-out plates containing glucose.

5 ml overnight-cultures in drop-out medium (-Trp,-Ura, Glc) were inoculated from streaks containing several colonies and grown at 30°C. Cultures were diluted in YPAG to an optical density at 600 nm (OD600) of 0.5. They were grown at 30 °C for 7.5h. For the native cell extract, 2.5 ml cell culture normalized to an OD600 of 3 were harvested by centrifugation, frozen in liquid nitrogen and subsequently resuspended in 75µl Y-PER (Pierce). The resuspended cell pellet was vortexed shortly and incubated slightly shaking at 20°C for 20 min. Subsequently insoluble and aggregated material were pelleted at maximal speed in an eppendorf centrifuge at 4°C for 10 min. The supernatant was mixed with loading dye, heated to 100°C for 5 min and separated on a 12% SDS-PAGE. The soluble Gal4-AD-scFv fusion and the heavy chain part of the Fab fused to the Gal4-AD were visualized by western blotting via detection of the Gal4-moiety with an anti-Gal4-AD monoclonal mouse antibody (Santa Cruz Biotechnology) as primary and an anti-mouse-peroxidase conjugate (Sigma) as secondary antibody and using a chemoluminescent substrate (Pierce) (Fig. 11). SDS-PAGE and western blotting procedures are well known to a person of ordinary skill in the art.

Claims

1. A single-chain framework having the general structure:

NH_2-VL-linker-VH-COOH; or

NH_2-VH-linker-VL-COOH

wherein the VH framework is of subtype 1a, 1b or 3.

2. A single-chain framework fused to a second protein moiety to yield a fusion construct of the general structure:

NH_2-VL-linker-VH-second protein-COOH; or

NH_2-second protein-VL-linker-VH-COOH

wherein the VH framework is of subtype 1a, 1b or 3.

3. The single chain framework of claim 1 or 2, wherein orientation of the VH and VL regions is reversed.

4. The single chain framework according to claim 1 to 3 wherein the VL framework is of the kappa1, lambda 1 or 3 type.

5. The single chain framework according to claim 2, wherein the second protein provides a read-out for intracellular assays.

6. A single chain framework selected from the group consisting of:

AH, BH, CH, DH, EH, FH, GH, AI, BI, CI, DI, EI, FI, GI, AJ, BJ, CJ, DJ, EJ, FJ, GJ, AK, BK, CK, DK, EK, FK, and GK

wherein A is the amino acid sequence (Seq. Id. No. 1)

EIVMTQSPSTLSASVGDRVIITCRASQSISSWLAWYQQKPGKAPKLLIYKASSLESG VPSRFSGSGSGAEFTLTISSLQPDDFATYYCQQYKSYWTFGQGTKLTVLG;

B is the amino acid sequence (Seq. Id. No. 2) EIVLTQSPSSLSASVGDRVTLTCRASQGIRNELAWYQQRPGKAPKRLIYAGSILQSG VPSRFSGSGSGTEFTLTISSLQPEDVAVYYCQQYYSLPYMFGQGTKVDIKR;

C is the amino acid sequence (Seq. Id. No. 3) EIVMTQSPATLSVSPGESAALSCRASQGVSTNVAWYQQKPGQAPR LLIYGATTRASGVPARFSGSGSGTEFTLTINSLQSEDFAAYYCQQYKHWPPWTFGQG TKVEIKR;

D is the amino acid sequence (Seq. Id. No. 4) QSVLTQPPSVSAAPGQKVTISCSGSTSNIGDNYVSWYQQLPGTAPQLLIYDNTKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDSSLSGVVFGGGTKLTVLG;

E is the amino acid sequence (Seq. Id. No. 5) EIVLTQSPATLSLSPGERATLSCRASQTLTHYLAWYQQKPGQAPR LLIYDTSKRATGVPARFSGSGSGTDFTLTISSLEPEDSALYYCQQRNSWPHTFGGGT KLEIKR;

F is the amino acid sequence (Seq. Id. No. 6) SYVLTQPPSVSVAPGQTATVTCGGNNIGSKSVHWYQQKPGQAPVL VVYDDSDRPSGIPERFSGSNSGNTATLTIRRVEAGDEADYYCQVWDSSSDHNVFGSG TKVEIKR;

G is the amino acid sequence (Seq. Id. No. 7) LPVLTQPPSVSVAPGQTARISCGGNNIETISVHWYQQKPGQAPVL VVSDDSVRPSGIPERFSGSNSGNTATLTISRVEAGDEADYYCQVWDSSSDYVVFGGG TKLTVLG;

H is the amino acid sequence (Seq. Id. No. 8) QVQLVQSGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSAISGSGGS TYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAAHVLRFLEWLPDAFDIW GQGTLVTVSS;

I is the amino acid sequence (Seq. Id. No. 9) EIVLTQSPSSLSASLGDRVTITCRASQSISSYLNWYQQKPGKAPK LLIYAASSSQSGVPSRFRGSESGTDFTLTISNLQPEDFATYYCQQSYRTPFTFGPGT KVEIKR;

J is the amino acid sequence (Seq. Id. No. 10) VQLVQSGAEVKKPGASVKVSCTASGYSFTGYFLHWVRQAPGQGLEWMGRINPDSGDT IYAQKFQDRVTLTRDTSIGTVYMELTSLTSDDTAVYYCARVPRGTYLDPWDYFDYWG QGTLVTVSS; and

K is the amino acid sequence (Seq. Id. No. 11)

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGL EWVSAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKDAGI AVAGTGFDYWGQGTLVTVSS.

7. A single chain framework selected from the group consisting of variants of the single chain frameworks according to claim 6.

8. A single chain framework selected from the group consisting of derivatives of the single chain frameworks according to claim 6.

9. A use of the single chain framework, antibody or antibody fragment according to any one of claims 1-7 or 17-18 in target validation, diagnostic applications, library construction or therapeutic applications.

10. The use of at least two framework sequences in the identification of a conserved framework residue class.

11. The use according to claim 10, wherein the conserved framework residue class is selected from the group consisting of:

polar but uncharged R groups;
positively charged R groups;
negatively charged R groups;
hydrophobic R groups; and
special amino acids.

12. The use of at least two framework sequences in the identification of at least one conserved framework sequence.

13. The use according to claim 12, wherein the conserved framework sequence is 2-5 residues.

14. The use according to claim 12, wherein the conserved framework sequence is 5-10 residues.

15. The use according to claim 12, wherein the conserved framework sequence is 10-25 residues.

16. The use according to claims 12-15, wherein the conserved framework sequence has gaps.

17. An antibody comprising the VL or the VH or both from the single chain framework according to any one of claims 1-8.

18. An antibody fragment comprising the VL or the VH or both from the single chain framework according to any one of claims 1-8.

19. A nucleic acid capable of encoding the single chain framework according to any one of claims 1-8.

20. A vector comprising the nucleic acid according to claim 19.

21. A host cell comprising the nucleic acid according to claim 19.

22. The use of the nucleic acid according to claim 19 in gene therapy.

Figure 1: Intracellular performance after Quality control selection in yeast *S. cerevisiae*

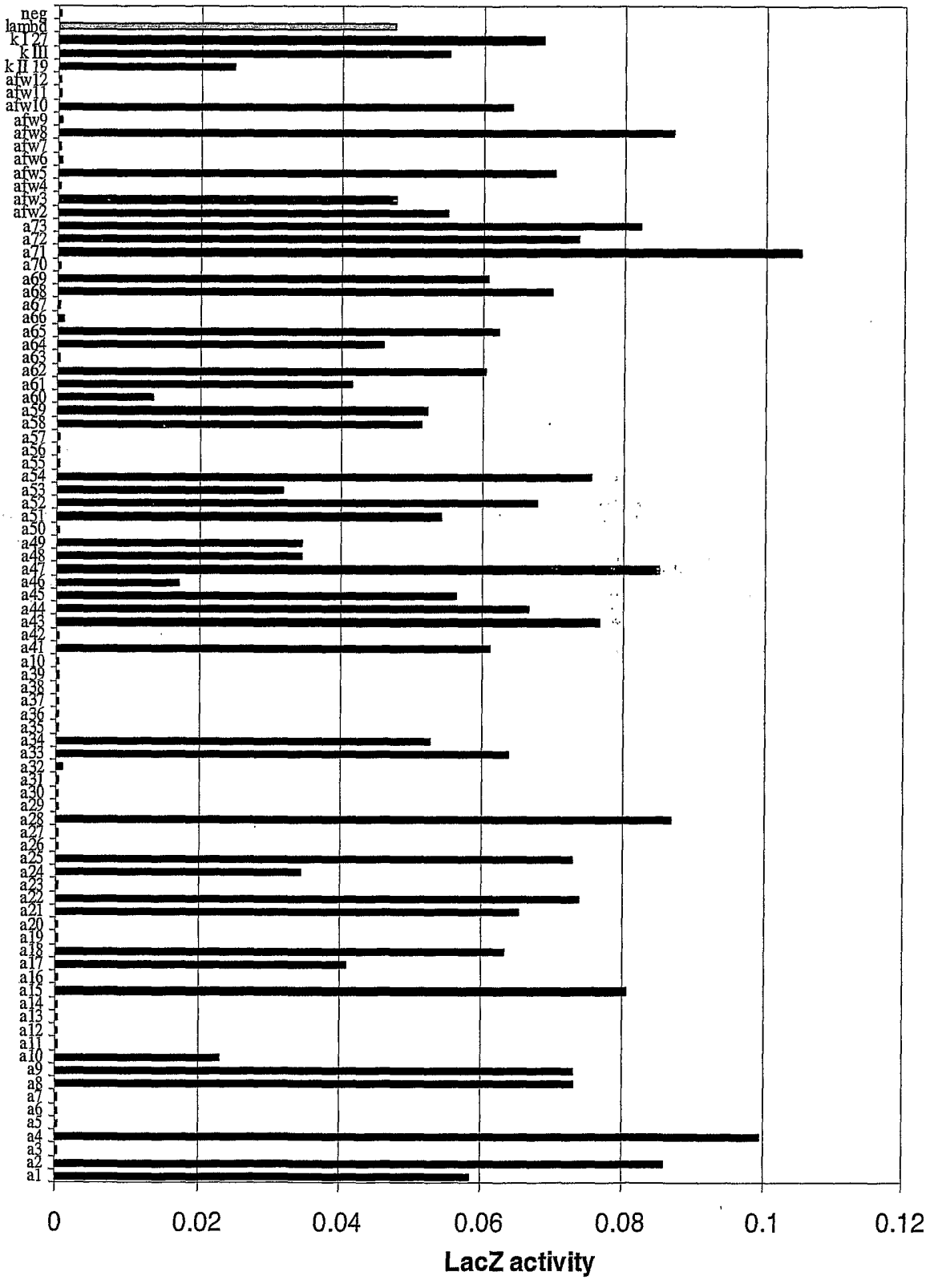


Figure 2: Intracellular performance of selected frameworks in a mammalian Hela cell line

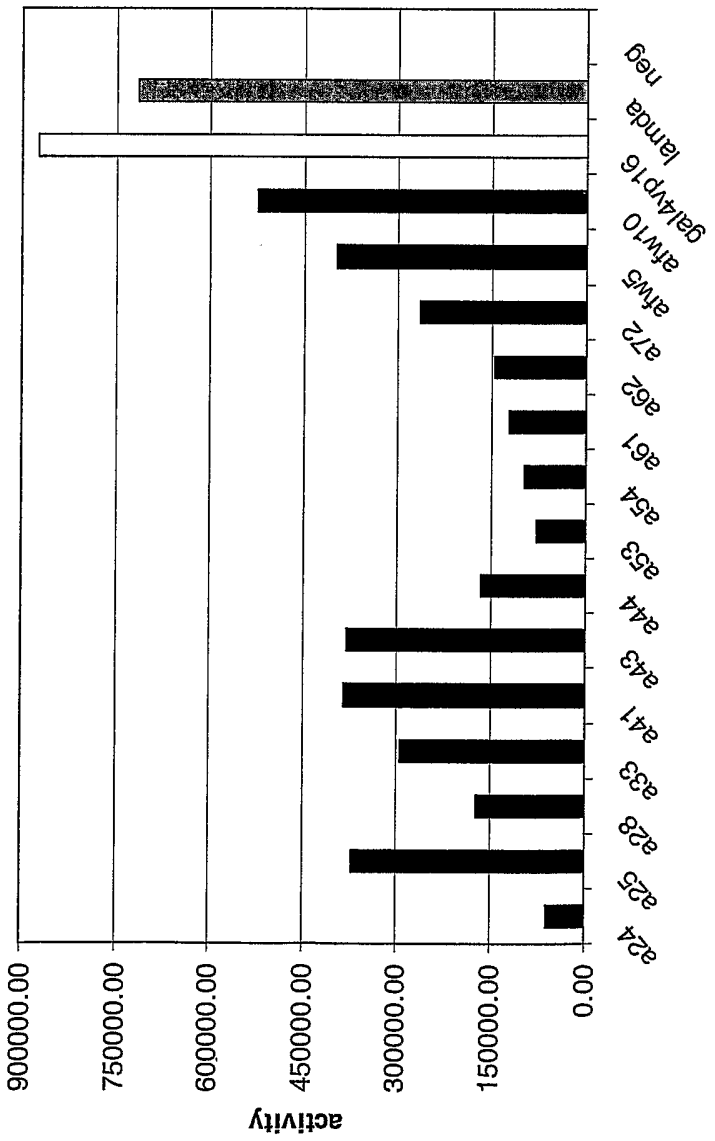


Fig 3:
Intracellular performance of novel frameworks in yeast *S. cerevisiae*

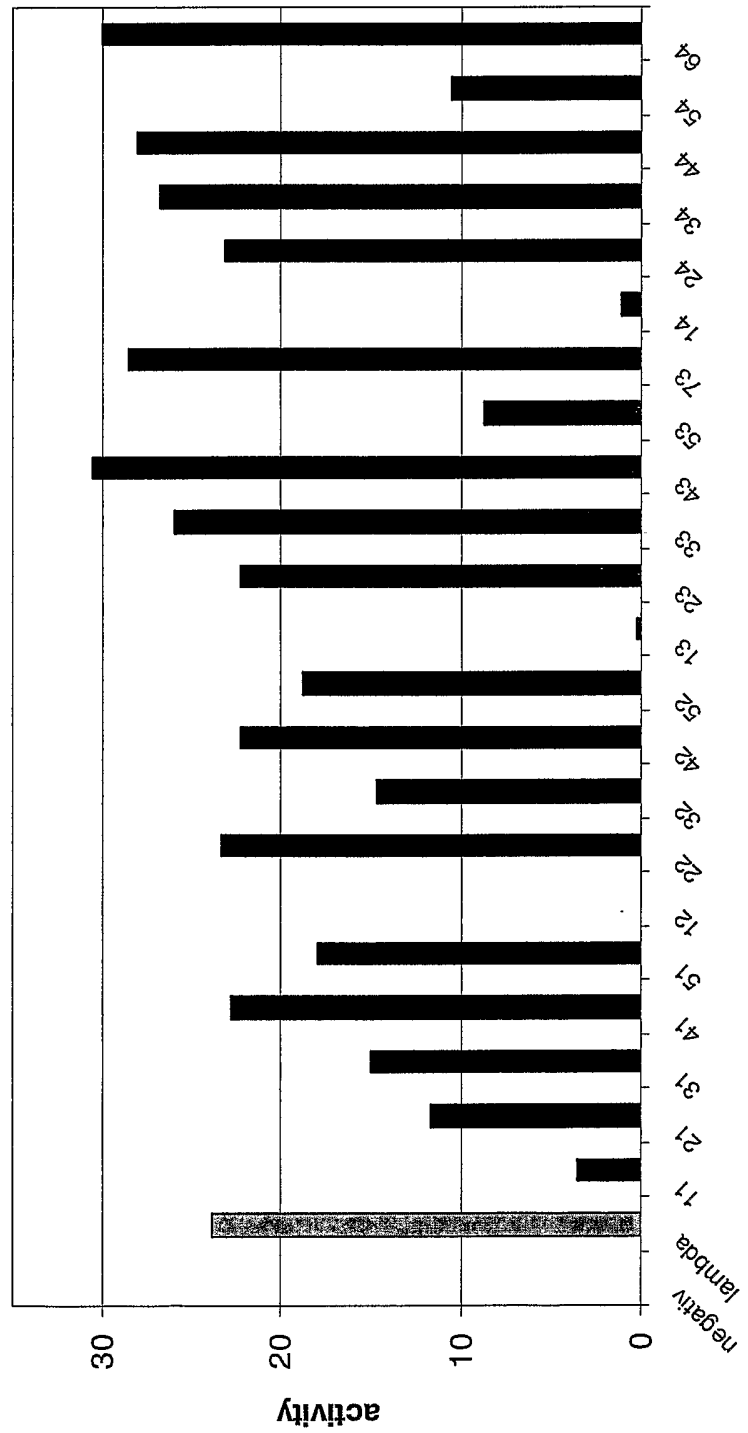
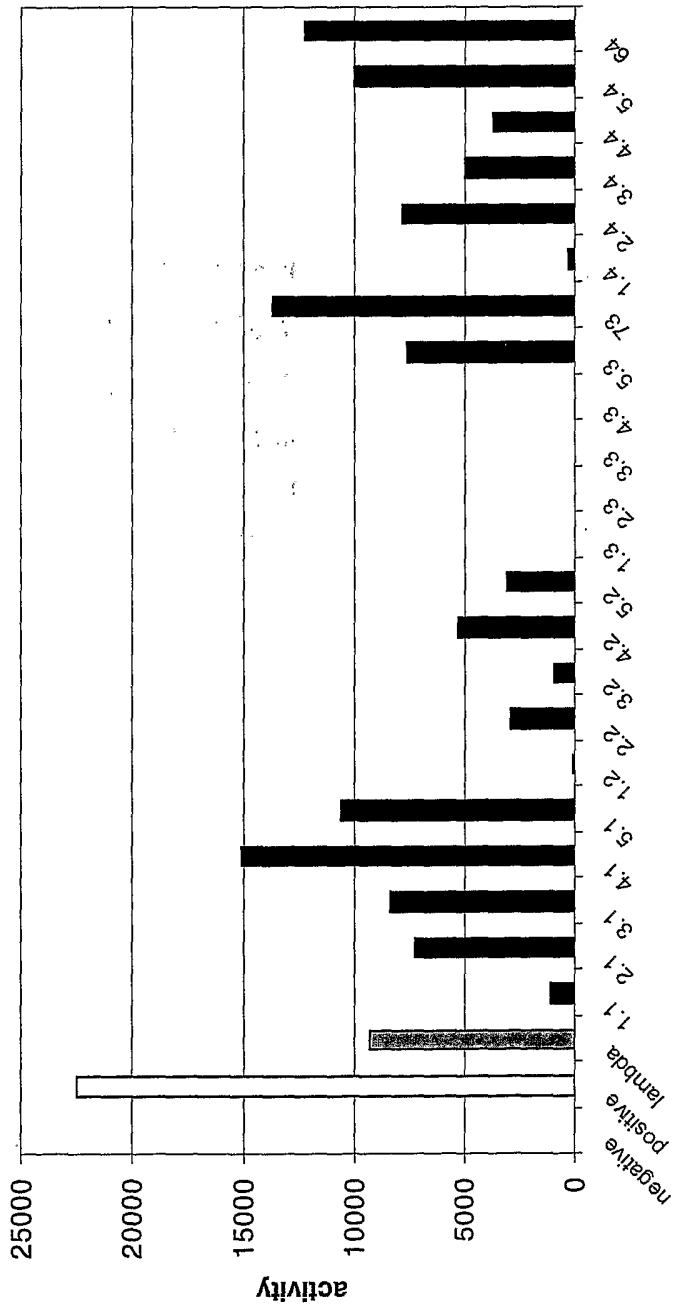


Figure 4:

Intracellular performance of novel frameworks in a mammalian Hela cell line



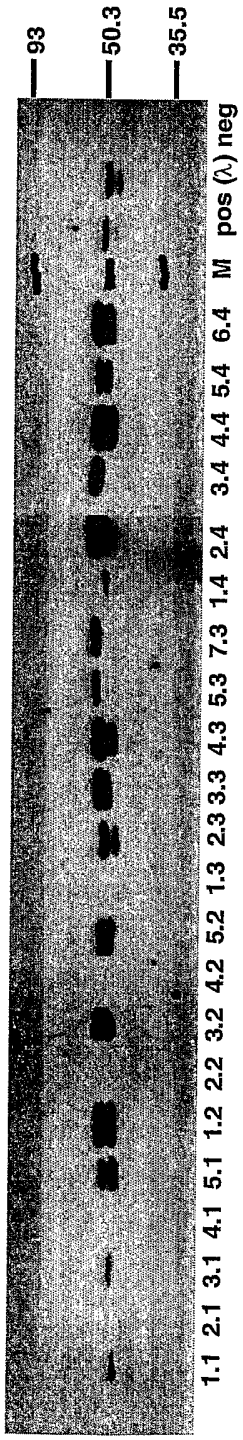


Figure 5: Soluble expression in the cytoplasm of yeast *S. cerevisiae*

Figure 6: Expression behavior in *E.coli*

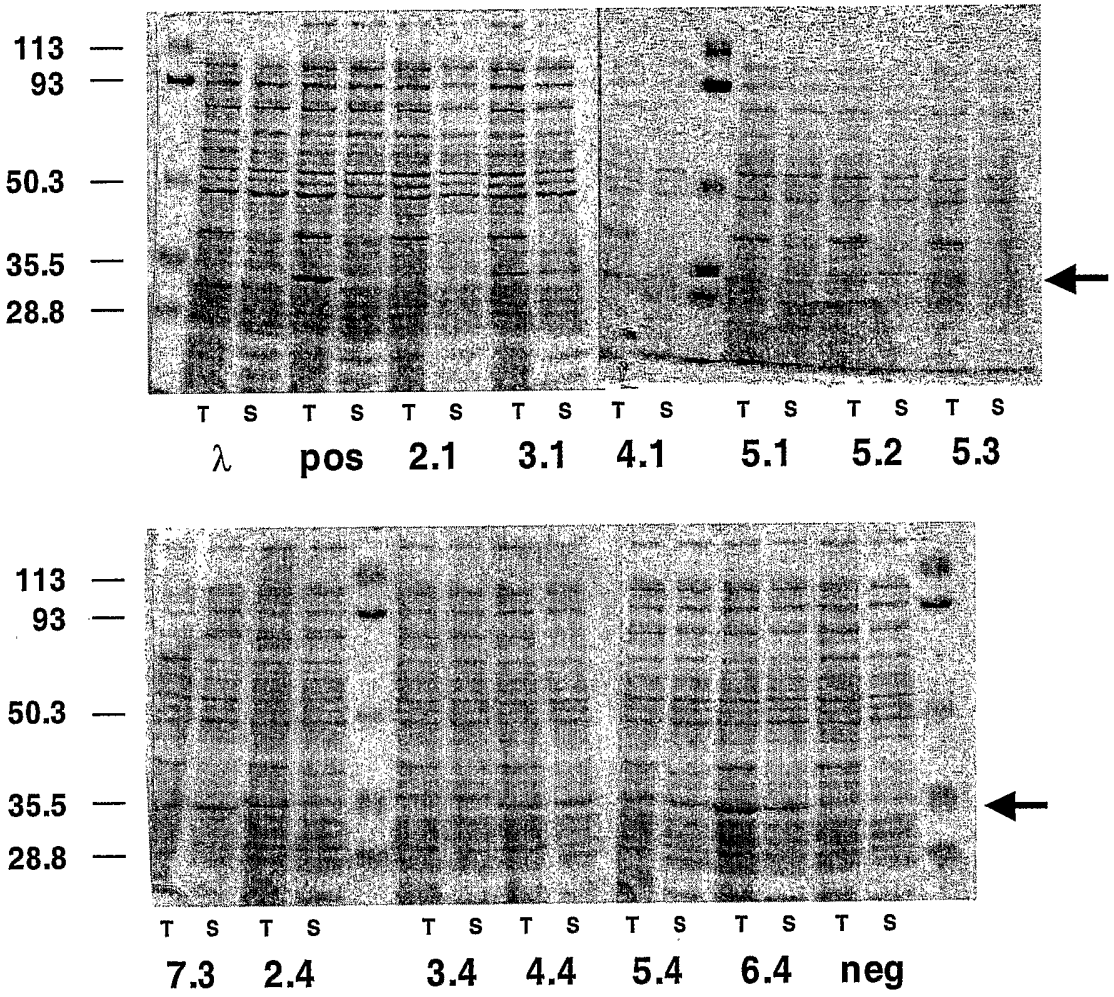


Figure 7:

Intracellular performance of selected novel frameworks in different mammalian cell lines

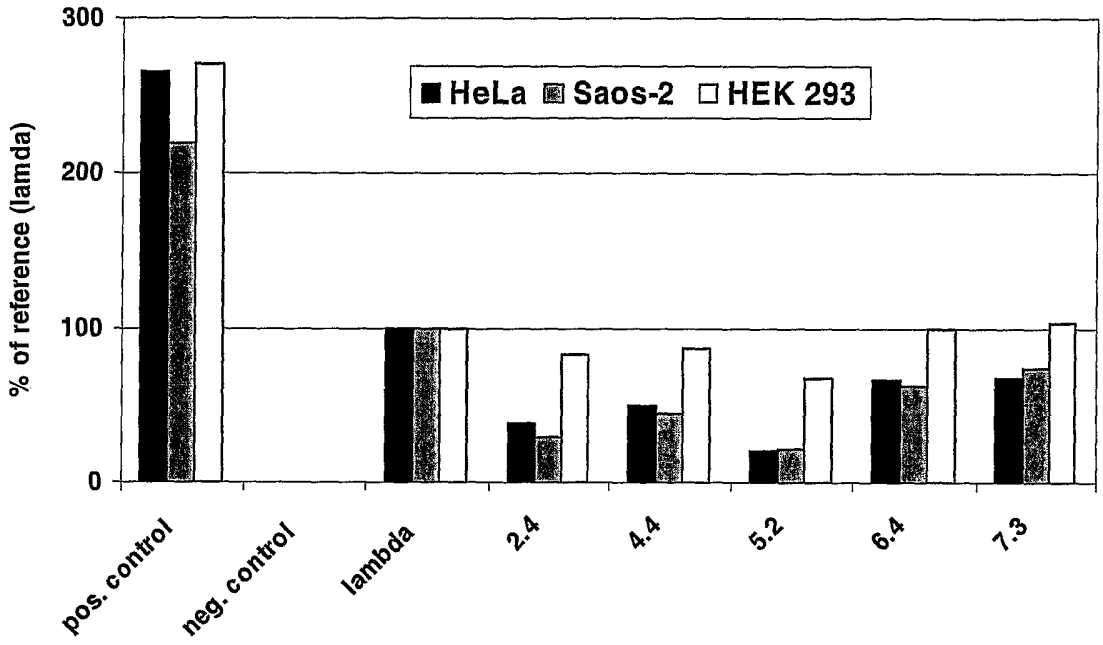


Fig. 8A: Resistance towards aggregation at 37°C

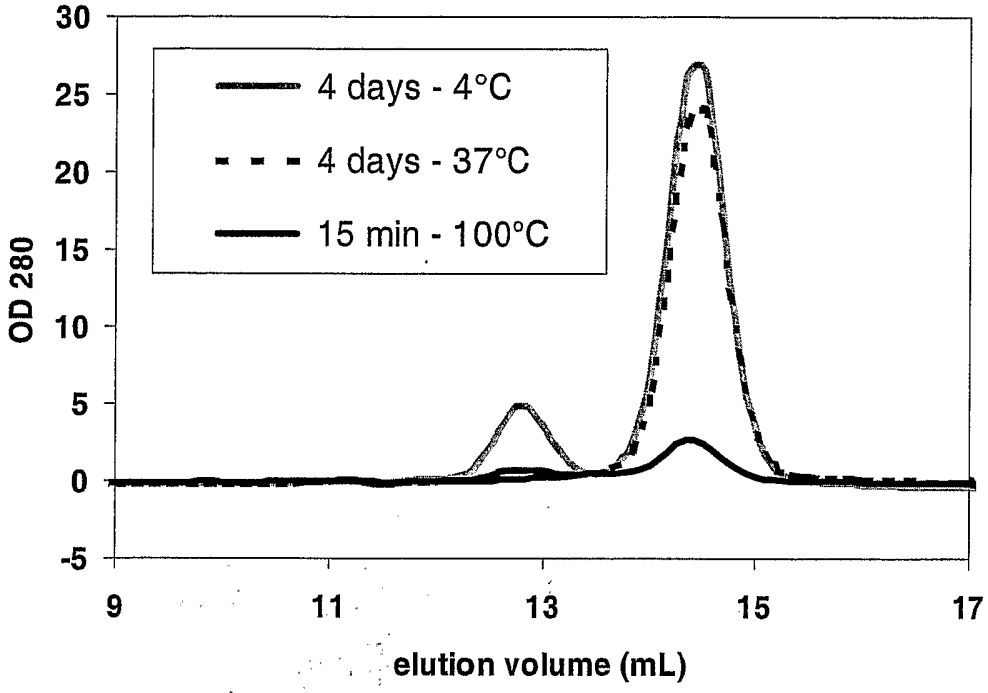
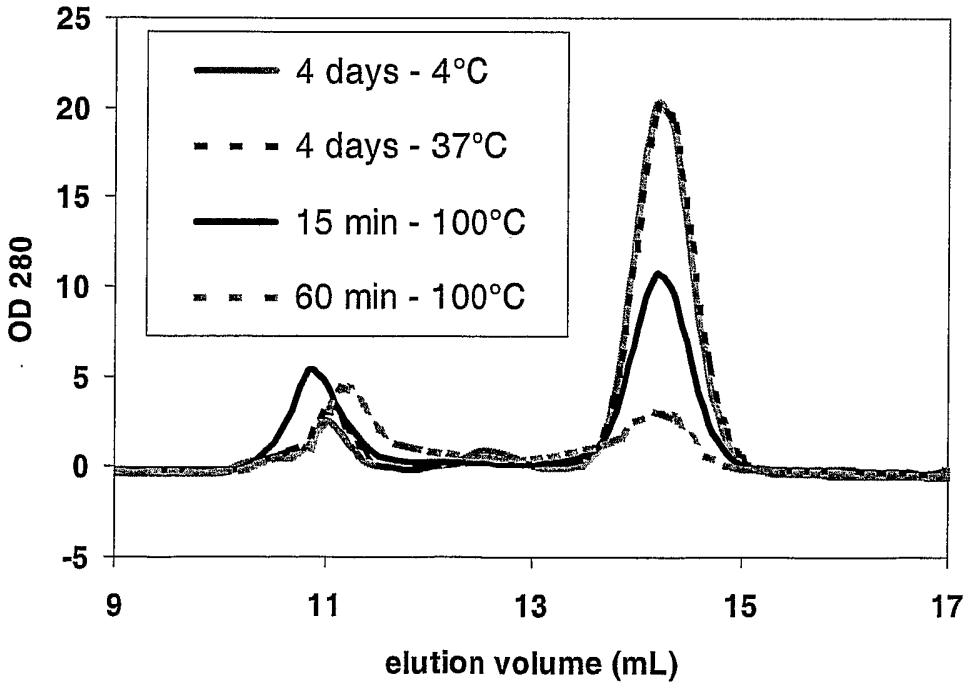


Fig 8B



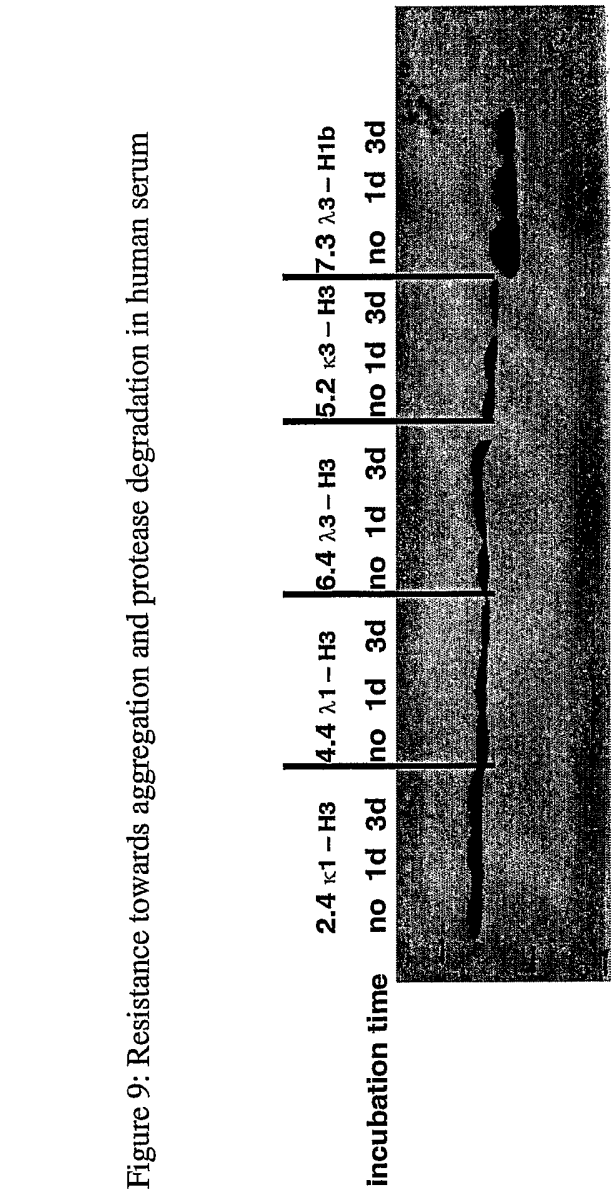


Figure 9: Resistance towards aggregation and protease degradation in human serum

Figure 10

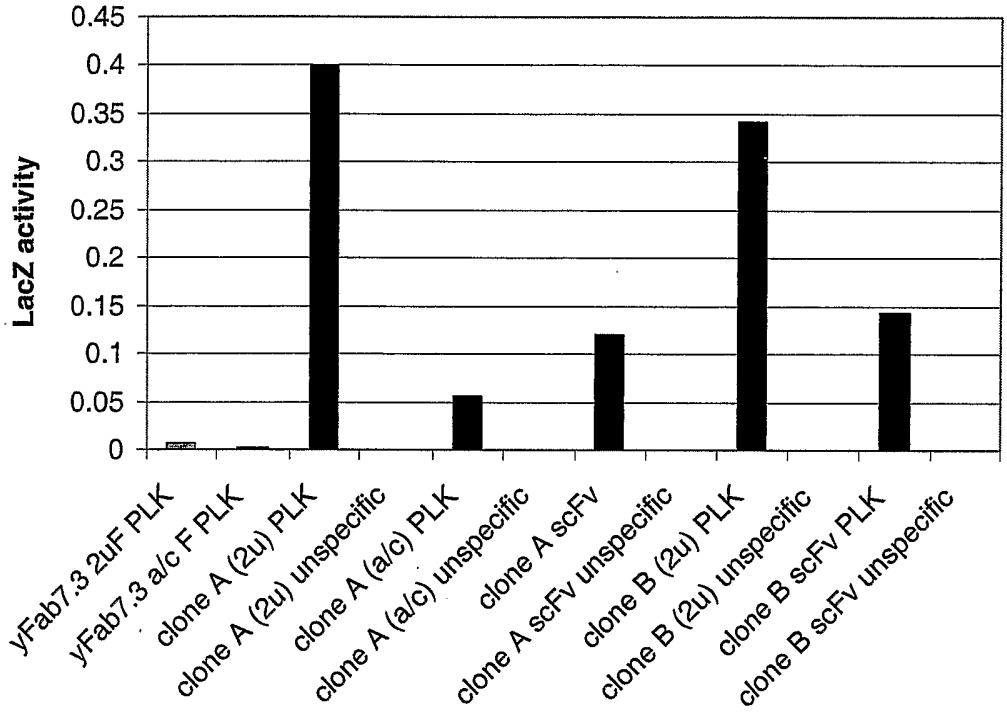

LacZ activity
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
yFab7.3 2uF PLK
yFab7.3 a/c F PLK
clone A (2u) PLK
clone A (2u) unspecific
clone A (a/c) PLK
clone A (a/c) unspecific
clone A scFv
clone A scFv unspecific
clone B (2u) PLK
clone B (2u) unspecific
clone B scFv PLK
clone B scFv unspecific

Figure 11

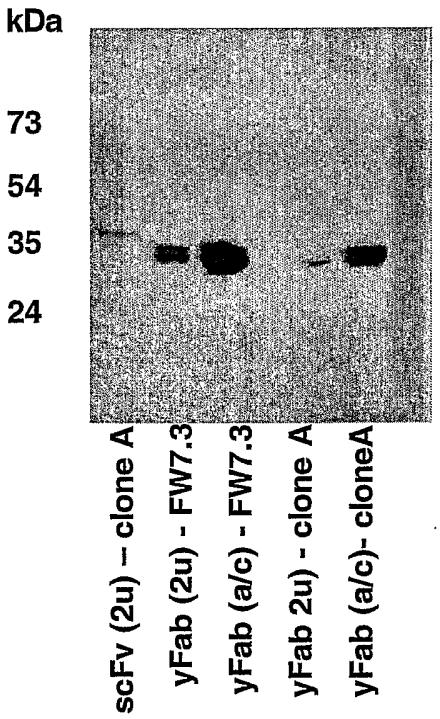

kDa
73
54
35
24
scFv (2u) – clone A
yFab (2u) - FW7.3
yFab (a/c) - FW7.3
yFab 2u) - clone A
yFab (a/c)- cloneA

Table 1: Alignment of selected VH domain sequences

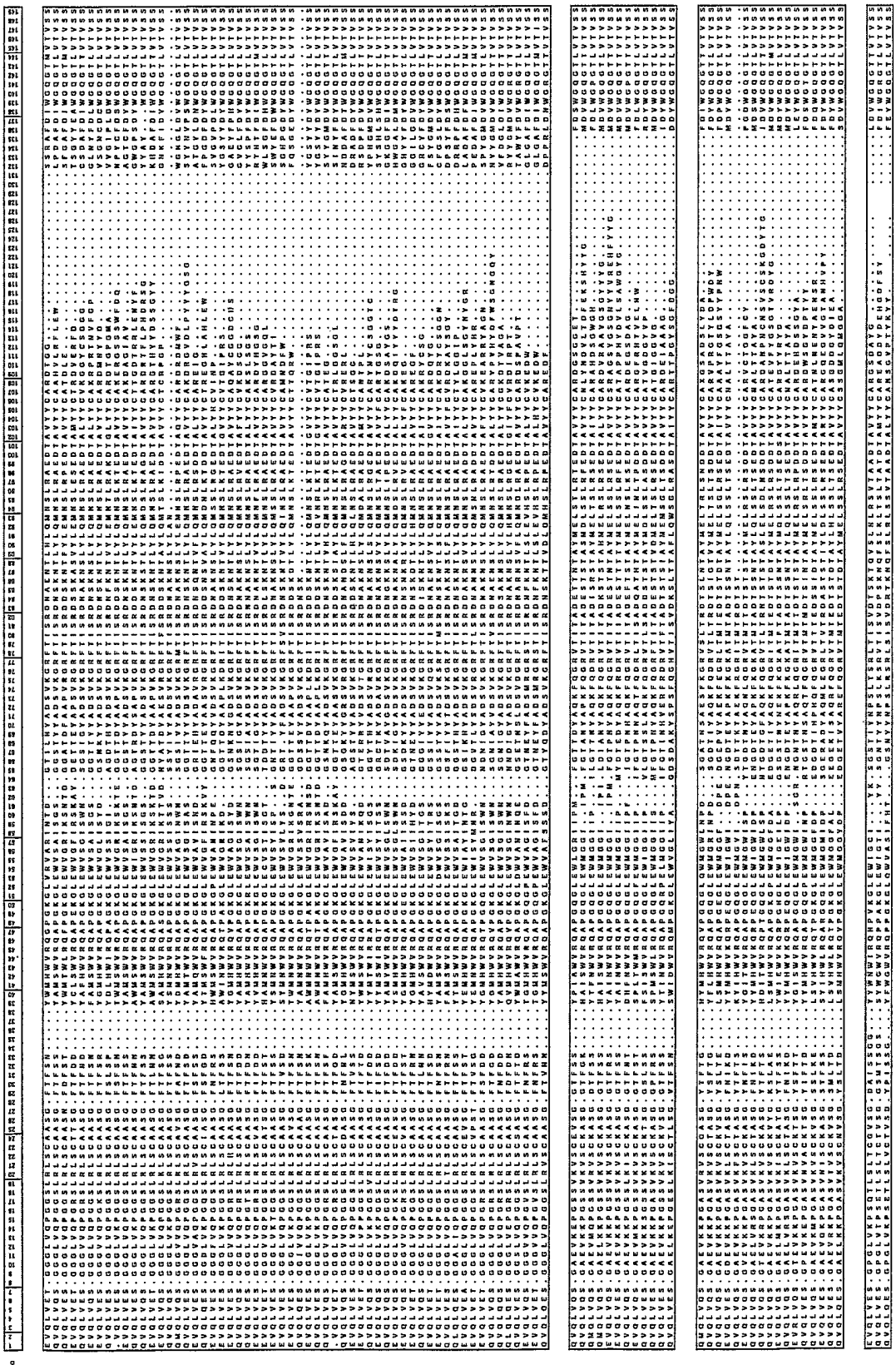


Table 2: Alignment of selected VL domain sequences

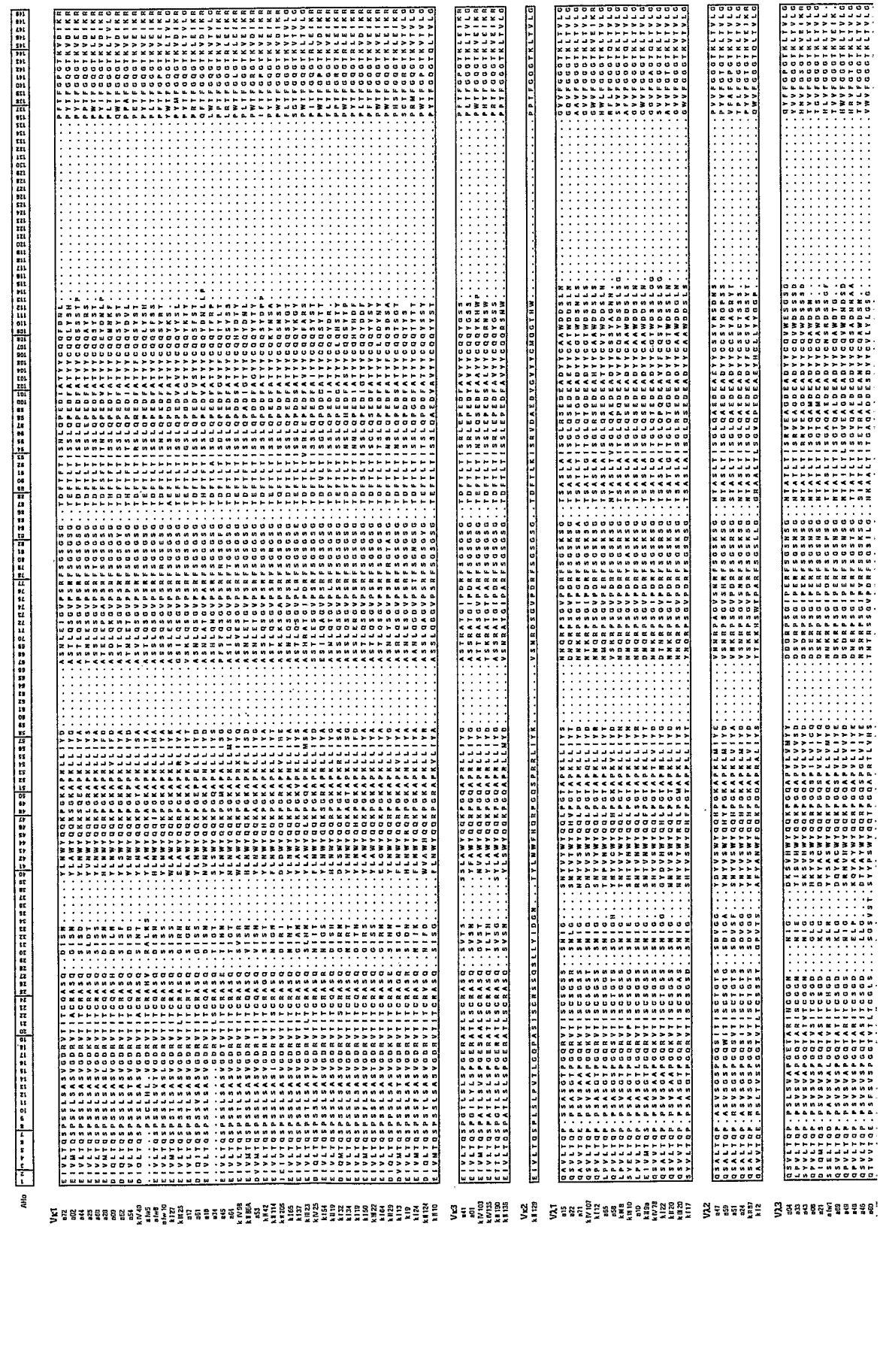


Table 3: Alignment of random library members

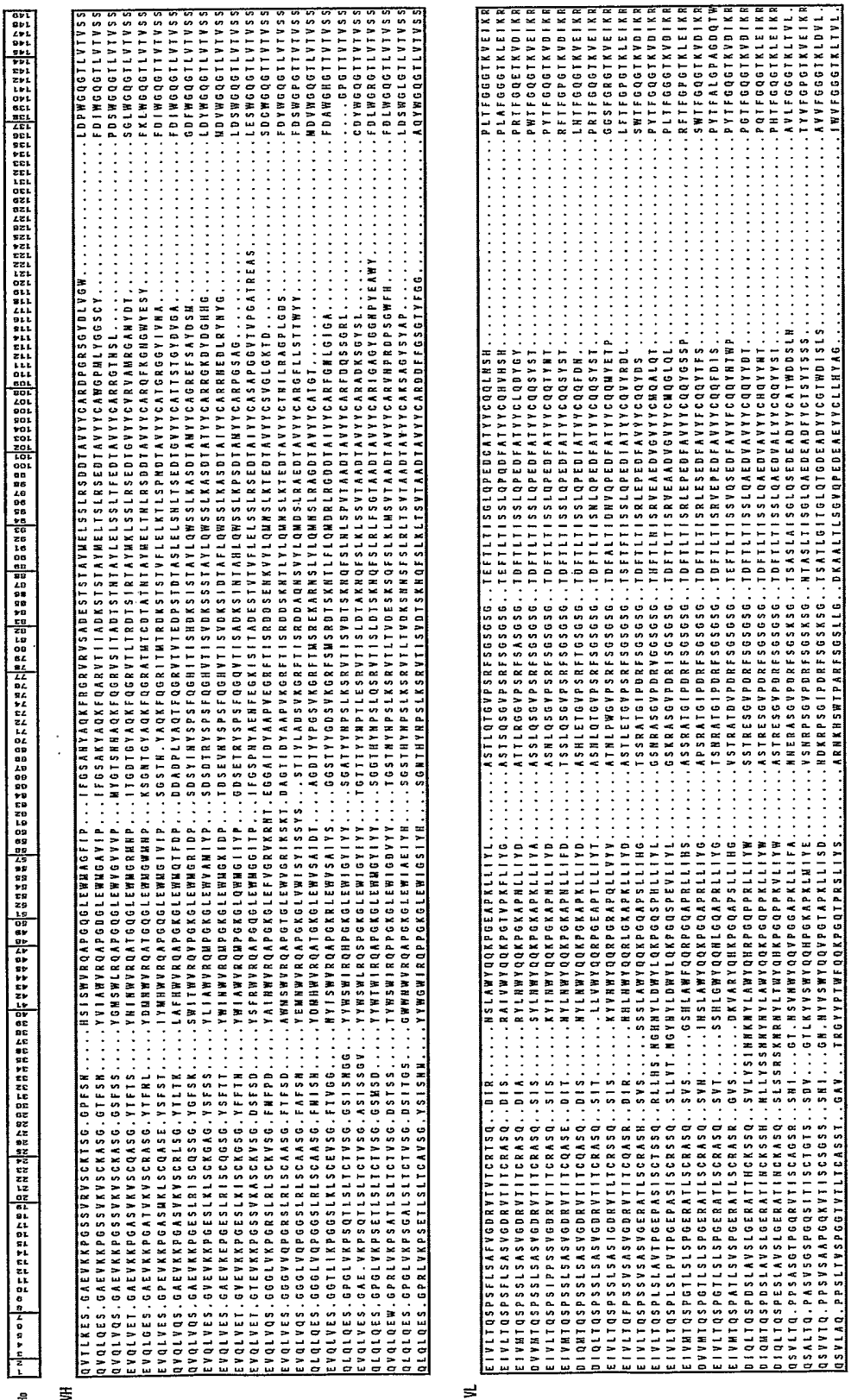


Table 4:

Statistics of subgroup-frequency of well-performing frameworks selected in the quality control system (% of total selected sequences

VLdomain	VL selected	VL random
VL κ1	52%	46%
VL κ2	5%	8%
VL κ3	5%	17%
VL κ4	0	12.5%
VL λ1	19%	4%
VL λ2	5%	4%
VL λ3	14%	8%
VH-domain	VH selected	VH random
VH 1b	19%	39%
VH 1a	9%	13%
VH 2	0	0
VH 3	67%	22%
VH 4	5%	26%
VH 5	0	0%
VH 6	0	0%

Table 5: VH- and VL-domain sequences used for the novel framework combinations			
abb.	origin	sub-group	amino-acid sequence
1.x	K I 27	Vκ1	EIVMTQSPSTLSASVGDRVIITCRASQSISSW LAWYQQKPGKAPKLLIYKASSLESGVPSRFS GSGSGAEFTLTISSLQPDDFATYYCQQYKSY WTFGQGTKLTVLG (Seq. Id. No. 1)
2.x	K III25	Vκ1	EIVLTQSPSSLSASVGDRVTLTCRASQGIRNE LAWYQQRPGKAPKRLIYAGSILQSGVPSRFS GSGSGTEFTLTISSLQPEDVAVYYCQQYYSL PYMFGQGTKVDIKR (Seq. Id. No. 2)
3.x	K IV103	Vκ3	EIVMTQSPATLSVSPGESAALSCRASQGVST NVAWYQQKPGQAPRLLIYGATTRASGVPA RFSGSGSGTEFTLTINSLQSEDFAAYYCQQY KHWPPWTFGQGTKVEIKR (Seq. Id. No. 3)
4.x	K IV107	Vλ1	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGDN YVSWYQQLPGTAPQLLIYDNTKRPSGIPDRF SGSKSGTSATLGITGLQTGDEADYYCGTWD SSLSGVVFGGGTKLTVLG (Seq. Id. No. 4)
5.x	K IV135	Vκ3	EIVLTQSPATLSLSPGERATLSCRASQTLTHY LAWYQQKPGQAPRLLIYDTSKRATGVPARF SGSGSGTDFTLTISSLEPEDSALYYCQQRNS WPHTFGGGTKLEIKR (Seq. Id. No. 5)
6.x	A43	Vλ3	SYVLTQPPSVSVAPGQTATVTCGGNNIGSKS VHWYQQKPGQAPVLVVYDDSDRPSGIPERF SGSNSGNTATLTIRRVEAGDEADYYCQVWD SSSDHNVFGSGTKVEIKR (Seq. Id. No. 6)
7.x	A33	Vλ3	LPVLTQPPSVSVAPGQTARISCGGNNIETISV HWYQQKPGQAPVLVVSDDSVRPSGIPERFS GSNSGNTATLTISRVEAGDEADYYCQVWDS SSDYVVFGGGTKLTVLG (Seq. Id. No. 7)

Table 5 (continued)

x.1	A44	VH3	QVQLVQSGGGLVQPGGSLRLSCAASGFTF SSYAMSWVRQAPGKGLEWVSAISGSGGST YYADSVKGRFTISRDNSKNTLYLQMNSLR AEDTAVYYCAAHVLRFLEWLPDAFDIWG QGTLVTVSS (Seq. Id. No. 8)
x.2	Afw10	VH3	EIVLTQSPSSLSASLGDRVTITCRASQSISSY LNWYQQKPGKAPKLLIYAASSSQSGVPSR FRGSESGTDFTLTISNLQPEDFATYYCQQS YRTPFTFGPGTKVEIKR (Seq. Id. No. 9)
x.3	A33	VH1b	VQLVQSGAEVKKPGASVKVSCTASGYSFT GYFLHWVRQAPGQGLEWMGRINPDSGDTI YAQKFQDRVTLTRDTSIGTVYMELTSLTSD DTAVYYCARVPRGTYLDPWDYFDYWGQ GTLVTVSS (Seq. Id. No. 10)
x.4	A43	VH3	EVQLVESGGGLVQPGGSLRLSCAASGFTFS SYAMSWVRQAPGKGLEWVSAISGSGGST YYADSVKGRFTISRDNSKNTLYLQMNSLR AEDTAVYYCAKDAGIAVAGTGFDYWGQG TLVTVSS (Seq. Id. No. 11)

08420PC.ST25.txt
SEQUENCE LISTING

<110> ESBATECH AG

<120> Immunoglobulin Frameworks which demonstrate enhanced stability in the intracellular environment and methods of identifying same

<130> 08420PC

<150> US 60/382,649
<151> 2002-05-22

<150> US 60/438,246
<151> 2003-01-03

<160> 13

<170> PatentIn version 3.1

<210> 1
<211> 107
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 1

Glu Ile Val Met Thr Gln Ser Pro Ser Thr Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Ile Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Trp
20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Lys Ala Ser Ser Leu Glu Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Ala Glu Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Asp Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Lys Ser Tyr Trp Thr
85 90 95

Phe Gly Gln Gly Thr Lys Leu Thr Val Leu Gly
100 105

<210> 2
<211> 108
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 2

Glu Ile Val Leu Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Leu Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Glu
20 25 30

Leu Ala Trp Tyr Gln Gln Arg Pro Gly Lys Ala Pro Lys Arg Leu Ile
35 40 45

Tyr Ala Gly Ser Ile Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln Tyr Tyr Ser Leu Pro Tyr
85 90 95

Met Phe Gly Gln Gly Thr Lys Val Asp Ile Lys Arg
100 105

<210> 3
<211> 109
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 3

Glu Ile Val Met Thr Gln Ser Pro Ala Thr Leu Ser Val Ser Pro Gly
1 5 10 15

Glu Ser Ala Ala Leu Ser Cys Arg Ala Ser Gln Gly Val Ser Thr Asn
20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
35 40 45

Tyr Gly Ala Thr Thr Arg Ala Ser Gly Val Pro Ala Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Ile Asn Ser Leu Gln Ser
65 70 75 80

Glu Asp Phe Ala Ala Tyr Tyr Cys Gln Gln Tyr Lys His Trp Pro Pro
85 90 95

Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
100 105

<210> 4
<211> 111
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 4

Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Ala Ala Pro Gly Gln
1 5 10 15

Lys Val Thr Ile Ser Cys Ser Gly Ser Thr Ser Asn Ile Gly Asp Asn
20 25 30

Tyr Val Ser Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Gln Leu Leu
35 40 45

Ile Tyr Asp Asn Thr Lys Arg Pro Ser Gly Ile Pro Asp Arg Phe Ser
50 55 60

Gly Ser Lys Ser Gly Thr Ser Ala Thr Leu Gly Ile Thr Gly Leu Gln
65 70 75 80

Thr Gly Asp Glu Ala Asp Tyr Tyr Cys Gly Thr Trp Asp Ser Ser Leu
85 90 95

Ser Gly Val Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
100 105 110

<210> 5
<211> 108
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 5

Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Thr Leu Thr His Tyr
20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
35 40 45

Tyr Asp Thr Ser Lys Arg Ala Thr Gly Val Pro Ala Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro
65 70 75 80

Glu Asp Ser Ala Leu Tyr Tyr Cys Gln Gln Arg Asn Ser Trp Pro His
85 90 95

Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
100 105

<210> 6
<211> 109
<212> PRT
<213> Artificial sequence

<220>
<223> Antibody framework

<400> 6

Ser Tyr Val Leu Thr Gln Pro Pro Ser Val Ser Val Ala Pro Gly Gln
1 5 10 15

Thr Ala Thr Val Thr Cys Gly Gly Asn Asn Ile Gly Ser Lys Ser Val
20 25 30

His Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Val Tyr
35 40 45

Asp Asp Ser Asp Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser
50 55 60

Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Arg Arg Val Glu Ala Gly
65 70 75 80

Asp Glu Ala Asp Tyr Tyr Cys Gln Val Trp Asp Ser Ser Ser Asp His
85 90 95

Asn Val Phe Gly Ser Gly Thr Lys Val Glu Ile Lys Arg
100 105

<210> 7
<211> 109
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 7

Leu Pro Val Leu Thr Gln Pro Pro Ser Val Ser Val Ala Pro Gly Gln
1 5 10 15

Thr Ala Arg Ile Ser Cys Gly Gly Asn Asn Ile Glu Thr Ile Ser Val
20 25 30

His Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Val Ser
35 40 45

Asp Asp Ser Val Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser
50 55 60

Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Ser Arg Val Glu Ala Gly
65 70 75 80

Asp Glu Ala Asp Tyr Tyr Cys Gln Val Trp Asp Ser Ser Ser Asp Tyr
85 90 95

Val Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
100 105

<210> 8
<211> 124
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 8

Gln Val Gln Leu Val Gln Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
20 25 30

Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Ala His Val Leu Arg Phe Leu Glu Trp Leu Pro Asp Ala Phe Asp

100 105 110

Ile Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 9
<211> 108
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 9

Glu Ile Val Leu Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Leu Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Ala Ala Ser Ser Ser Gln Ser Gly Val Pro Ser Arg Phe Arg Gly
50 55 60

Ser Glu Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Asn Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Arg Thr Pro Phe
85 90 95

Thr Phe Gly Pro Gly Thr Lys Val Glu Ile Lys Arg
100 105

<210> 10
<211> 123
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 10

Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala Ser
1 5 10 15

Val Lys Val Ser Cys Thr Ala Ser Gly Tyr Ser Phe Thr Gly Tyr Phe
20 25 30

Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met Gly

35 40 45

Arg Ile Asn Pro Asp Ser Gly Asp Thr Ile Tyr Ala Gln Lys Phe Gln
50 55 60

Asp Arg Val Thr Leu Thr Arg Asp Thr Ser Ile Gly Thr Val Tyr Met
65 70 75 80

Glu Leu Thr Ser Leu Thr Ser Asp Asp Thr Ala Val Tyr Tyr Cys Ala
85 90 95

Arg Val Pro Arg Gly Thr Tyr Leu Asp Pro Trp Asp Tyr Phe Asp Tyr
100 105 110

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 11
<211> 122
<212> PRT
<213> Artificial sequence

<220>
<223> antibody framework

<400> 11

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
20 25 30

Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Asp Ala Gly Ile Ala Val Ala Gly Thr Gly Phe Asp Tyr Trp
100 105 110

Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

08420PC.ST25.txt

<210> 12
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> PCR Primer

<400> 12
tgctctagaa gtgctgcagt gactgcag 28

<210> 13
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> PCR Primer

<400> 13
ggttgtcgac ttacaggctg ctgggagcaa tcg 33