(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets

(11) **EP 2 134 841 B9**

(12) **CORRECTED EUROPEAN PATENT SPECIFICATION**

(15) Correction information:
Corrected version no 1 (W1 B1)
Corrections, see
Claims EN 1

(48) Corrigendum issued on:
11.06.2014 Bulletin 2014/24

(45) Date of publication and mention
of the grant of the patent:
22.02.2012 Bulletin 2012/08

(21) Application number: **08719320.7**

(22) Date of filing: **17.03.2008**

(51) Int Cl.:
C12N 15/10 (2006.01)   *C07K 16/00* (2006.01)

(86) International application number:
PCT/IB2008/000628

(87) International publication number:
WO 2008/110914 (18.09.2008 Gazette 2008/38)

(54) **METHODS FOR PRODUCING ACTIVE SCFV ANTIBODIES AND LIBRARIES THEREFOR**

VERFAHREN ZUR HERSTELLUNG AKTIVER SCFV-ANTIKÖRPER UND BIBLIOTHEKEN DAFÜR

PROCÉDÉS DE PRODUCTION D'ANTICORPS SCFV ACTIFS ET BIBLIOTHÈQUES DE CEUX-CI

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MT NL NO PL PT RO SE SI SK TR

(30) Priority: **15.03.2007 US 894947 P**

(43) Date of publication of application:
23.12.2009 Bulletin 2009/52

(73) Proprietors:
- **INSERM**
  75654 Paris Cédex 13 (FR)
- **UNIVERSITE LOUIS PASTEUR DE STRASBOURG**
  F-67084 Strasbourg Cédex (FR)
- **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)**
  75016 Paris (FR)

(72) Inventors:
- **MARTINEAU, Pierre, Emile, Ulysse**
  34980 Saint Gely du Fesc (FR)
- **WEISS, Etienne**
  67640 Fegersheim (FR)

(74) Representative: **Jacobson, Claude et al**
Cabinet Lavoix
2, Place d'Estienne d'Orves
75441 Paris Cedex 09 (FR)

(56) References cited:
WO-A-00/54057   WO-A-01/48017
WO-A-99/45959

- LADEN JEAN-CHARLES ET AL: "Expression and folding of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm." RESEARCH IN MICROBIOLOGY SEP 2002, vol. 153, no. 7, September 2002 (2002-09), pages 469-474, XP002509241 ISSN: 0923-2508
- JUNG S ET AL: "IMPROVING IN VIVO FOLDING AND STABILITY OF A SINGLE-CHAIN FV ANTIBODY FRAGMENT BY LOOP GRAFTING" PROTEIN ENGINEERING, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 10, no. 8, 1 January 1997 (1997-01-01), pages 959-966, XP000971779 ISSN: 0269-2139
- AUF DER MAUR ADRIAN ET AL: "Antigen-independent selection of intracellular stable antibody frameworks." METHODS (SAN DIEGO, CALIF.) OCT 2004, vol. 34, no. 2, October 2004 (2004-10), pages 215-224, XP004526807 ISSN: 1046-2023
- DER MAUR ADRIAN AUF ET AL: "Direct in vivo screening of intrabody libraries constructed on a highly stable single-chain framework." THE JOURNAL OF BIOLOGICAL CHEMISTRY 22 NOV 2002, vol. 277, no. 47, 22 November 2002 (2002-11-22), pages 45075-45085, XP002284890 ISSN: 0021-9258

EP 2 134 841 B9

- MARTINEAU P ET AL: "IN VITRO FOLDING AND THERMODYNAMIC STABILITY OF AN ANTIBODY FRAGMENT SELECTED IN VIVO FOR HIGH EXPRESSION LEVELS IN ESCHERICHIACOLI CYTOPLASM" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 292, 1 October 1999 (1999-10-01), pages 921-929, XP004462583 ISSN: 0022-2836 cited in the application
- MARTINEAU P ET AL: "Expression of an antibody fragment at high levels in the bacterial cytoplasm" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 280, no. 1, 3 July 1998 (1998-07-03), pages 117-127, XP004453942 ISSN: 0022-2836

- PHILIBERT PASCAL ET AL: "A focused antibody library for selecting scFvs expressed at high levels in the cytoplasm." BMC BIOTECHNOLOGY 2007, vol. 7, 22 November 2007 (2007-11-22), page 81, XP002509301 ISSN: 1472-6750

Description

FIELD

[0001]   This disclosure relates to recombinant single-chain antibodies and methods of producing and using such antibodies.

BACKGROUND

[0002]   Genetic engineering approaches have allowed the production of recombinant antibodies having specific binding specificities, specific domain structures, and other desirable properties. One type of genetically engineered antibody is the single chain Fv fragment (scFv). Single chain Fv fragments are genetically engineered polypeptides that contain a heavy chain variable region (VH) linked to a light chain variable region (VL) via a flexible peptide linker. Each VH and VL domain contains three complementarity determining regions (CDRs). CDRs are short amino acid sequences that vary greatly among antibody molecules, and thus, are responsible for generating the great diversity of antibody binding specificity. The combination of the CDRs of the VH plus the CDRs of the VL determines the binding specificity of any given antibody.

[0003]   Single chain Fv fragments display the binding specificity and monovalent binding affinity of full-size antibodies and provide the added benefit of relative ease of genetic manipulation and expression (because scFvs are encoded by and expressed from a single coding sequence, rather than from separate coding sequences, as are full-size antibodies). Single chain Fv fragments and other recombinant antibodies are used in a broad variety of applications, for example, in medical diagnostic tests, in basic research; and as therapeutic antibody treatments for various diseases.

[0004]   Intrabodies are genetically-engineered antibody molecules that are ectopically expressed within cells. Intrabodies can be used to visualize or to modulate the function of a target antigen within living cells. For example, the use of intrabodies can induce a phenotypic knockout either by directly inhibiting the function of the targeted antigen or by diverting the targeted antigen from its normal intracellular location (e.g., an intrabody can redirect its target antigen to the degradation machinery). Intrabodies can also enhance or change the function of their target antigens. For protein targets, intrabodies can be targeted to a specific post-translational modification or to a specific antigen conformation. Moreover, an intrabody-induced phenotypic knockout can be confined to a specific cell compartment by targeting an intrabody to the specific subcellular compartment using an addressing signal (e.g., a nuclear localization signal, a mitochondrial localization signal, or an endoplasmic reticulum retention signal). Intrabodies can also modulate target function by modifying the oligomeric structure of the target.

[0005]   Because intrabody phenotypic knockout relies only on the binding capacity of the antibody molecule to its target, it is not necessary to express within the cell a complete antibody molecule but only its binding site, which is entirely located within the variable region (Fv). Given their advantages of small size and antigen specificity encompassed within a single polypeptide chain, scFvs are the most common type of recombinant antibody fragment used for intracellular antibody expression.

[0006]   One serious limitation to the use of intrabodies is that most scFvs are not able to fold under the reducing conditions of the cell cytosol and nucleus. Under such conditions the two conserved disulfide bridges of scFvs are reduced, thereby destabilizing and inactivating the binding activity of many scFvs. *In vitro,* most scFvs cannot be renatured under reducing conditions. Statistical analyses of scFv sequences have shown that fewer than 1% of the scFvs are stable enough to be expressed and active in absence of disulfide bond formation. In addition, even if a scFv protein is indeed stable enough in its reduced form to be expressed and active *in vivo,* other parameters such as protease susceptibility or folding kinetics may also influence the final *in vivo* fate of the intrabody and thus are critical for ultimate intrabody expression and activity.

[0007]   To obtain an active intrabody, current approaches often involve two successive steps. First, a panel of scFv or Fab antibodies that specifically bind an antigen of interest are identified (for example, by screening a phage display library). Second, the specifically-binding antibodies are tested for their ability to bind and/or inhibit the target antigen *in vivo.* Because fewer than 1% of scFvs are potentially useful as intrabodies (because they are not expressed and/or cannot properly fold under the reducing conditions that exist within a cell), identification of a single scFv that can be used as an intrabody requires the isolation of more than 100 scFv clones, a number that is unlikely to be obtained in most cases.

[0008]   Application WO 00/54057 discloses a method for determining the ability of an immunoglobulin to bind to a target antigen in an intracellular environment. The method disclosed consists of providing i) an intracellular immunoglobulin associated with the a first molecule, and ii) an intracellular target antigen associated with a second molecule, then in assessing the intracellular interaction between immunoglobulin to the target in intracellular, knowing that when the first and second molecules are brought into stable interaction by binding of immunoglobulin to the target in intracellular environment, a signal is generated. To evaluate the possibility to conduct this method in mammalian cells, a test has been carried out with a first fusion protein comprising beta-galactosidase, and a second fusion protein the anti-beta-

galactosidase scFvR4.

[0009] Application WO 01/48017 discloses a method for isolating an scFv with a framework which is stable and soluble in a reducing environment. The method disclosed is performed by using a "two-hybrid assay" comprising the following steps:

a) a scFv library with varied frameworks and constant CDRs is generated by mutation of at least one framework encoding region of DNA sequence of a scFv to a known antigen and by introduction of such mutations into suitable expression vectors,

b) host cells able to express a specific known antigen and only surviving in the presence of antigen-scFv-interaction are transformed with said scFv library,

c) the thus transformed host cells are cultivated under conditions suitable to express the antigen and the scFv and allowing cell survival only in the presence of antigen-scFv-interaction,

d) the scFv expressed in surviving cells and having a defined framework that is stable and soluble in reducing environment is isolated.

[0010] Der Maur Adrian et al. (Methods, 34: 215-224, 2004) and Der Maur Adrian et al. (Journal of Biological Chemistry, 277(47): 45075-45085, 2002) disclosed methods for isolating an scFv with a framework which is stable and soluble in a reducing environment.

[0011] Der Maur Adrian et al. (Methods, 34 : 215-224, 2004) teaches in Table 1 that of a total of 10^7 screened clones, only 0.1% allow cell growth under selective conditions, and only 60% of these clones which allow cell growth under selective conditions show scFv-dependent lacZ expression.

[0012] Laden et al. (Research in Microbiology, 153: 469-474, 2002) disclosed a mutant of a human anti-beta-galactosidase single chain antibody fragment, called scFv13R4, and shown that scFv13R4 is able to fold properly in Escherichia coli cytoplasm and is active *in vivo* in yeast since it is able to interact with beta-galactosidase in a two-hybrid assay.

[0013] Jung and plückthun (protein Engineering, 10(8): 959-966, 1997) discloses a method for improving in vivo folding and stability of a single-chain Fv antibody fragment by using CDR grafting. The method is performed with the framework of humanized 4D5 antibody.

[0014] Application WO 99/45959 discloses a humanized antibody framework motif wherein said framework has been selected from a human library based upon comparison to a murine antibody, and the heavy chain is encoded by the VH gene of K5B8, and the light chain is encoded by the VL gene of TR1.6.

[0015] However, none of the prior art documents recited above discloses an antibody library comprising at least 10^6 unique scFv antibody clone wherein each clones encodes a framework sequence identical to a framework encoded by antibody clone scFv13R4 (SEQ ID NO: 32).

[0016] In view of the foregoing difficulties in producing and identifying antibodies that can be used as intrabodies for use in medical and research applications, what is needed are more efficient methods of producing and selecting antibodies that can be used as intrabodies.

SUMMARY

[0017] The inventors have described herein an antibody library that includes at least about 10^6 unique scFv clones, wherein at least about 20% of the scFv clones encode an antibody that can detectably specifically bind a target antigen within a cell when the antibody encoded by the scFv clone is expressed within the cell.

[0018] The inventors have also described herein an antibody library, wherein the antibody library includes at least about 10^6 unique scFv antibody clones, wherein at least about 20% of the scFv antibody clones can be expressed within an *E. coli* cell to produce soluble antibody at a level of at least about 5 mgs per liter of *E. coli* cells, wherein the *E. coli* cells have been grown to an OD_{600nm} of about 5.

[0019] In an aspect, of the invention described herein is an antibody library including at least about 10^6 unique scFv antibody clones, wherein each unique scFv antibody clone encodes a unique scFv antibody comprising at least one of a unique CDR3 VH sequence and a unique CDR3 VL sequence, and wherein the unique scFv antibody clones encode a framework sequence substantially identical to a framework sequence encoded by scFv13R4 (SEQ ID NO: 32).

[0020] In the above antibody library of the invention, the unique scFv antibody clones can encode scFv antibodies including a unique CDR3 VH sequence.

[0021] In the above antibody library of the invention, the unique scfv antibody clones can encode scFv antibodies including a unique CDR3 VH, sequence.

[0022] In the above antibody library of the invention, the unique scFv antibody clones can encode scFv antibodies including a unique CDR3 VL sequence.

[0023] In the above antibody library of the invention the unique scFv antibody clones can encode scFv antibodies including a unique CDR3 VH sequence and a unique CDR3 VL sequence.

[0024] The inventors have, described herein an scFv antibody that can be expressed as substantially soluble protein under reducing conditions, wherein the scFv antibody is isolated from the library described above in the third aspect. The scFv antibody can specifically bind to a target antigen under reducing conditions.

[0025] The inventors have described herein a method of producing an scFv antibody, including expressing the scFv antibody described above in the fourth aspect within a cell, thereby producing the scFv antibody. The method can include further purifying the scFv antibody from the cell.

[0026] In a second aspect of the invention, described herein is a method for preparing an scFv antibody library enriched for scFv antibody clones that can be expressed within a cell, including: a) providing a first collection of scFv antibody clones, wherein the first collection comprises clones comprising a unique sequence within a CDR3 loop of VH, wherein the first collection has been enriched for scFv antibody clones that can be detectably expressed when introduced into a cell; b) providing a second collection of scFv antibody clones, wherein the second collection comprises clones comprising a unique sequence within a CDR3 loop of VL, wherein the second collection has been enriched for scFv antibody clones that can be detectably expressed when introduced into a cell; c) joining VH domains from scFv antibody clones of the first collection with VL domains from scFv antibody clones of the second collection to obtain a third collection of scFv antibody clones, wherein the third collection contains scFv antibody clones comprising a unique sequence within the CDR3 loop of VH and a unique sequence within the CDR3 loop of VL, thereby preparing the scFv antibody library enriched for scFv antibody clones that can be expressed within a cell.

[0027] In the above method for preparing an scFv antibody library, the first collection includes scFv antibody clones that contain a VL sequence identical to an scfv13R4 (SEQ ID NO:33) VL sequence and the second collection can include scFv antibody clones that contain a VH sequence identical to an scFv13R4 (SEQ ID NO:33) VH sequence.

[0028] In the above method for preparing an scFv antibody library, the first collection can include scFv antibody clones that include identical CDR1 and CDR2 sequences in the VH domain and the second collection can include scFv antibody clones that include identical CDR1 and CDR2 sequences in the VL domain.

[0029] In a third aspect of the invention, described herein is an antibody library produced by the method described above in the sixth aspect.

[0030] The inventors also described herein an antibody selected from an antibody library produced by the method described above in the second aspect.

[0031] In a ninth aspect, the invention features a method for constructing an antibody library including: a) selecting an scFv antibody framework; b) introducing sequence diversity into a VH CDR3 region of the scFv antibody framework to generate a first library including scFv antibody clones including a unique VH CDR3 region; c) introducing sequence diversity into a VL CDR3 region of the scFv antibody framework to generate a second library including scFv antibody clones including a unique VL CDR3 region; d) removing, from the first library, clones that do not detectably express scFv antibody; e) removing, from the second library, clones that do not detectably express scFv antibody; and f) recombining the first and second libraries to generate a final library comprising scFv antibody clones comprising a unique VH CDR3 region and a unique VL CDR3 region, thereby constructing the antibody library.

[0032] In the above method, the scFv is scFv13R4.

[0033] It is an object of the present invention to provide a novel antibody library for the isolation of scFvs expressed in the cytoplasm that may be used as intrabodies.

[0034] It is another object of the present invention to provide a novel antibody library based on a single framework and optimized for intracellular expression.

[0035] A further object of the present invention is to provide novel methods of constructing and validating a novel antibody library for the isolation of scFvs expressed in the cytoplasm that may be used as intrabodies.

[0036] Another object of the present invention is to provide novel methods of constructing and validating a novel antibody library based on a single framework and optimized for intracellular expression.

[0037] Still another object of the present invention is to provide novel methods of using an antibody library in order to produce highly expressed scFvs that may be used as intrabodies.

[0038] Yet another object of the present invention is to provide novel methods of using an antibody library in order to produce scFvs based on a single framework and optimized for intracellular expression.

[0039] These and other objects, features, and advantages of the present invention will become apparent after review of the following detailed description of the disclosed embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0040]

Figure 1. (a) Schematic outline of the steps followed during library construction. The critical steps are: introduction of tailored CDR3 loops in an unique human scFv framework; removal of non-expressed clones by fusion with the CAT enzyme and selection on CAM plates; recombination of the 13 VH and 5 VL libraries, and display on phage.

(b) Summary of the CDR3 loops collected in the database. (c) Distribution of the amino acids at each position of the 5 amino acid long VH CDR3s from 55 rearranged human antibodies.

Figure 2. Schematic representation of CDR3 length distribution. Distribution of the CDR3 lengths in the database and in 118 sequenced clones from the library .

Figure 3. Western blot analysis of twenty clones from the library that were cloned in a cytoplasmic expression vector and expressed in *E. coli* under the control of the T7 promoter. Twenty clones from the library were cloned in a cytoplasmic expression vector and expressed in *E. coli* under the control of the T7 promoter. Soluble extracts were prepared, separated by SDS-PAGE, and analyzed by Coomassie staining (a) or Western-blot (b) using 9E10 and an alkaline phosphatase conjugated anti-mouse IgG antibody (substrate BCIP/NBT). Each lane corresponded to 2 x 10^7 (2 X 10 to the 7th) cells. The arrow on the left indicates the position of the scFv.

Figure 4. Selected fluorescent micrographs of expression of randomly picked scFvs in HeLa cells. Cells were transfected with scFv-EGFP constructs as indicated. 13R4 and 1F4 represent the positive and negative controls, respectively. At 24 h post-transfection, cells were fixed and visualized under a fluorescent microscope with the fluorescein isothiocyanate filter set. The micrographs represent typical fields containing a similar number of cells in each case. Magnification: x 400.

Figure 5. Selection of binders against five purified proteins. (a) 2.5×10^{10} phages from each round of selection were tested by ELISA against their respective antigen and revealed using an anti-M13 HRP conjugated monoclonal antibody (Pharmacia). R0 is the non selected library and Rn the stock obtained after the nth round of selection. Specificity was tested on BSA for the 3rd rounds of selection (A450 nm ~0.1 - 0.3). (b) Monoclonal phages from each round of selection against GST:Syk protein were tested by ELISA as in (a). Percentage of soluble active scFvs (Absorbance > 0.1) selected after three rounds against the indicated protein and expressed either in the periplasm (gray) or in the cytoplasm (black).

Figure 6. Photomicrographs of the expression of anti-histones scFvs fused to EGFP in HeLa cells. The cells were transfected and treated as indicated in the legend of figure 4. The pictures represent typical cells transfected with scFv13R4 and three representative anti-histones clones (2, 5 and 10). D, DAPI staining (blue) merged with the GFP signal.

Figure 7. Graphical representation of the expression of the activity of selected scFvs under reducing conditions. Extracts of cells expressing the scFvs in the cytoplasm were prepared as in figure 3 in presence or absence of 10 mM DTT. ELISA was performed in presence or absence of DTT, accordingly. x-axis: amount of extract per well. y-axis: ELISA signal at 450 nm.

Figure 8. Photomicrographs of double staining of HeLa cells using an intrabody and by immunofluorescence. HeLa cells were transfected with anti-histones clone 5 (Figure 6) fused to the dsRed-monomer GFP. Microtubules were revealed in the transfected cells by IF using the anti-tubulin scFv 2F12C (Table 3). Cells were observed at the appropriate wavelength to visualize: (A) 2F12C scFv (microtubules alone); (B) clone 5 intrabody (histones alone); (C) 2F12C and clone 5 (both microtubules and histones); (D) as C plus nucleus staining with DAPI.

DETAILED DESCRIPTION OF THE INVENTION

[0041] Provided herein are methods for constructing an scFv library enriched for scFvs that can be expressed within a prokaryotic or eukaryotic cell. The scFvs maintain their structure under the reducing conditions present within a cell, retain their ability to specifically bind a target antigen within a cell, and thus can be used as intrabodies in various therapeutic and research approaches. Because the methods described herein can be used to produce libraries of scFvs that are stable under reducing conditions, these libraries are also useful for producing scFvs that can be expressed by and purified from prokaryotic or eukaryotic cells (e.g., by lysing the cells and purifying the desired scFv by well-known techniques). Such purified scFvs are useful as antibodies for use in research, diagnostic tests, and for disease therapies.

[0042] An improved strategy for stabilizing scFvs to be used *in vivo* is to construct a scFv library tailored for intracellular expression. Ideally, such a library should only contain scFvs able to fold under reducing conditions such as those found in the cytoplasm of a cell. Another strategy is to construct an scFv library based on a single optimized antibody framework for intrabody selection.

[0043] Described herein are methods for constructing a novel antibody library that is based on a single optimized antibody framework and tailored for intracellular expression. Through molecular evolution, we obtained a human scFv called scFv13R4, which is expressed at high levels in *E. coli* cytoplasm. In addition, this scFv is highly expressed, soluble, and displays specific binding activity to a target antigen in yeast and mammalian cells. This scFv is very stable *in vitro* and can be renatured in presence of a reducing agent. In addition, analysis of its folding kinetics showed that it folds faster than the parent scFv and aggregates more slowly *in vitro.*

[0044] The human scFv library based on the framework of the optimized scFv13R4 contains more than a billion clones, which is larger than previous libraries of 10^7 to 10^8. The diversity of the present human scFv library is much greater than that of previous libraries because we designed the present library to encode VH and VL CDR3 loops of various different

lengths. In addition, we used a biased mix of degenerate oligonucleotide sequences to encode CDR3 loops that mimic human CDR3. Using optimized CDR3 loops and filtering steps to eliminate the non-expressed clones, we purged the library of non-expressed scFvs and retained the cytoplasmically expressed scFvs without compromising the diversity of the clones, as confirmed through testing with several proteins used as the antigens. Contrary to previously described scFv libraries, most of the scFvs in the library are well expressed in *E. coli* and in mammalian cytoplasm.

[0045] This new approach to building scFv libraries allows facile and large-scale selection of functional intrabodies. For example, several strong binders against different proteins, including the Syk and Auroroa-A protein kinases, the $\alpha\beta$ tubulin dimer, the papillomavirus E6 proteins, the core histones, gankyrin, and MAPK11-14, have been isolated from the library. Some of the selected scFvs are expressed at an exceptionally high level in the bacterial cytoplasm, allowing the purification of 1 mg or more of active scFv from only 20 ml of culture. Moreover, after three rounds of selection against core histones as a target antigen, more than half of the selected scFvs were active when expressed in vivo in human cells and were essentially localized in the nucleus. This new type of library, methods of creating and using such libraries, and antibodies isolated from such libraries, are useful not only for the simple and large-scale selection of functional intrabodies but also for the expression and purification of highly expressed scFvs that can be used in numerous biotechnological, diagnostic, and therapeutic applications.

Intrabodies

[0046] Intrabodies are genetically-engineered antibody molecules that are ectopically expressed within cells. Intrabodies can be used to visualize or to inhibit a targeted antigen in living cells, and thus find use in various research and medical (e.g., diagnostic and therapeutic) applications. However, intrabody technology has been limited by the fact that fewer than 1% scFvs in a typical antibody expression library are stable enough to be expressed and/or active in vivo, because the intracellular environment reduces the two conserved disulfide bridges that the vast majority of scFvs require for stability. Described herein are methods of producing libraries of scFvs that are greatly enriched for scFvs that are stable under reducing conditions and thus are suitable for use as intrabodies. The intrabodies can be used for various research, diagnostic, and therapeutic approaches that employ intrabodies.

[0047] In most cases, obtaining efficient intrabodies currently requires two successive steps. First, a panel of antibodies against the target antigen must be isolated. Due to the availability of very high quality naive antibody libraries displayed on phage, this step is now easily accomplished by phage-display and can be automated in order to isolate binders against several proteins in parallel. In a second step, these antibody fragments (scFv or Fab) must be tested *in vivo* for their ability to inhibit their target. However, most scFvs do not fold properly under the reducing conditions found in the cytosol and the nucleus of the cell where most of the interesting targets are located. This can result in aggregated and inactive scFvs, which are unable to interact with their target. Since less than 1% of scFvs are efficient as intrabodies, getting a single binder against a protein requires the isolation of 100 different scFvs, a number which is unlikely to be obtained in most cases. This makes the process of identifying intrabodies from regular scFv libraries a difficult procedure even when the screening is done *in vivo* using two-hybrid or equivalent systems. In addition, this low proportion of active scFvs in the current libraries results in a 100-fold decrease in the potential repertoire screened, making the isolation of intrabodies against different epitopes of the same protein unlikely.

Libraries

[0048] Described herein are novel phage-display libraries of scFvs optimized for intracellular expression and novel methods of constructing and using the library. The library is constructed on a single antibody framework of a parental scFv which was selected because of its improved activity inside the cytoplasm. The parental scFv is very stable, has favorable folding and aggregation kinetics and is expressed at very high levels in all tested cell types. Having a single framework for the construction of a library should allow more comparable expression levels between clones since most of their sequences are conserved.

[0049] Because CDR sequences also play a role in scFv folding and expression, however, we believed that the expression level of the clones would still exhibit some variability. To minimize these differences, we introduced variability only within the CDR3 loops because these loops are the most variable in antibodies and are thus more likely to be highly tolerant to sequence and length variations. Introduction of variability in the CDR3 is sufficient to generate antibodies against most proteins. In addition, we carefully biased the frequencies of the amino acids in these loops so as to recover the distribution observed among natural human sequences. When the expression levels of randomly selected clones were compared in the cytoplasm, despite some clear differences, a high proportion of them were correctly expressed both in *E. coli* and in mammalian cells. This proportion of intracellularly expressed scFvs is much higher than in previously described libraries.

[0050] We were also able to select binders (i.e., antibodies that specifically bind to a selected target antigen) against five different proteins (Aurora-A, GST:Syk, tubulin, histones, and E6 protein from papillomavirus HPV16). In subsequent

studies we have also isolated binders against new targets including gankyrin and MAPK11-14. For MAPK11-14, the four proteins involved are the four isoforms of the p38 MAP kinase. These proteins are very homologous (~60-74% identity). In all cases we were able to isolate scFvs specific for the isoform used for the selection. This underlines the specificity of the scFvs that can be isolated from this library.

[0051] A frequent concern when constructing scFv libraries is the simultaneous optimization of the library's diversity and size. Generally, the size of such a library is limited by the transformation efficiency to about 10^{10} clones. Given this limited number of clones, it is thus of premium importance to avoid non-expressed scFvs or duplicates. To solve this problem we first selected an antibody framework that was already optimized for intracellular expression, and then used a two-step procedure to further optimize the library.

[0052] First, we constructed 18 "small" libraries for each CDR3 length (13 VH CDR3 lengths and 5 VL CDR3 lengths) to be used for constructing the scFv library (see Example 1, herein below). Each of these 18 libraries was made by replacing the corresponding CDR3 of the parental scFv13R4 by the randomized CDR3. The resulting library with one and only one randomized CDR3 was then fused to the gene encoding chloramphenicol acetyltransferase (CAT), which we used as a selectable marker. After transformation of each CDR3 library into *E. coli*, the libraries were plated on chloramphenicol (CAM)-containing medium to select for those CDR3-CAT fusions that were expressed in the *E. coli* cytoplasm. This step reduced the diversity of each library by about 10-30%.

[0053] Although the CAT gene is used in the examples provided herein, one of ordinary skill in the art will understand that nucleic acids encoding any appropriate selectable marker can be fused with CDR3-encoding-nucleic acids and used in the methods provided herein to enrich scFv libraries for scFvs that are expressed in prokaryotic or eukaryotic cells. For example, suitable selectable markers for use in bacterial cells include, but are not limited to, the kanamycin resistance gene. Examples of selectable markers for use in mammalian cells include but are not limited to, e.g., the neomycin resistance gene, the puromycin resistance gene, and the hygromycin resistance gene. Examples of selectable markers for yeast cells include URA3, HIS3, and purE. In addition, markers such as green fluorescent protein (GFP) and derivatives thereof, beta-galactosidase, luciferase, or other luminescent, fluorometric, and/or colorimetric markers can be used, in all types of cells, for example, with fluorescence-activated cell sorting (FACS), to enrich scFv libraries for clones that encode scFvs that are capable of being expressed. One of ordinary skill in the art will understand how to use the teachings herein, together with what is known in the art, to select an appropriate selectable marker and an appropriate selection scheme to construct the scFv libraries described herein.

[0054] In a second step, we used PCR to randomly assemble the selected VH and VL libraries to generate the final diversity, based on the hypothesis that if a newly-generated scFv clone containing a new CDR3 sequence in its VH region plus the original VL of scFv13R4 was expressed, and if a newly-generated scFv clone containing a new CDR3 sequence in its VL region plus the original VH region of scFV 13R4 was expressed, then a recombined scFv clone containing the new CDR3 sequence in its VH region plus the new CDR3 sequence in its VL region would also be expressed, thereby resulting in a library containing only expressed scFvs. This was the case since 19 out of 20 clones selected at random were expressed at least partially in a soluble form in *E. coli* cytoplasm. Importantly, since this selection step was done early during the library construction, the diversity of the final library was only limited by the final transformation. This final recombination step, by generating a high diversity, ensured that all the clones were unique in the final library. Altogether, this approach resulted in a library of 1.5 billion expressed and different scFvs.

[0055] Successful use of scFvs as intrabodies on a large scale requires several essential points to be fulfilled by the library. First, the scFv must be easy to isolate. This is the case for the presently-described methods, since not only we were able to isolate binders against all the tested proteins, but also a single cycle of selection was enough to get close to 100% of binders. This very high enrichment rate is presumably due both to the high quality of the biased library which contains only well expressed scFvs and to the use of the trypsin-sensitive helper phage KM13. This is of premium importance since it may be possible to use a single panning cycle before *in vivo* testing of the scFvs as intrabodies, allowing a better diversity of the targeted epitopes. Also, this high enrichment rate also reduces by a huge amount the quantity of purified antigen needed for the panning steps. We were able to select on microtiter plates with as little as 1 μg of protein per well. Since further selection and confirmation of the binding activity by ELISA is not necessary because of the very high proportion of binders, it is now possible to isolate good intrabodies with a very small amount of antigen. Second, the scFv should be able to fold in all the cell compartments, particularly in the reducing ones such as the cytoplasm and the nucleus. Again, this is the case for the scFv library described herein, since more than 80% of the tested clones are at least partially soluble in the cell. In addition, we have shown that good cytoplasmic binders can be obtained from the phage selected scFvs in *E. coli* and in eukaryotic cells.

[0056] Provided herein are antibody libraries that comprises at least about: 10^6, 5×10^6, 10^7, 5×10^7, 10^8, 5×10^8, 10^9, 1.5×10^9, 5×10^9, 10^{10}, 5×10^{10}, 10^{11}, $5 \times 10"$, 10^{12}, 10^{13}, 10^{14}, 10^{15}, or 10^{16} unique scFv clones, wherein at least about: 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, or more of the scFv clones encode an antibody that can detectably specifically bind a target antigen when the antibody encoded by the scFv clone is expressed within a cell (e.g., as an intrabody). The scFvs described herein can be expressed to detectably specifically bind to a target antigen in any prokaryotic or eukaryotic cell in which it would be desirable to detectably specifically bind a target antigen (e.g.,

but not limited to, a bacterial cell, a yeast cell, an insect cell, an amphibian cell, an avian cell, a mammalian cell, and the like).

[0057] By "specifically binds" is meant that an scFv antibody preferentially binds to its target antigen rather than to another antigen. By "detectably specifically binds" is meant that specific binding of scFv antibody to its target antigen can be observed, e.g., but not limited to a phenotypic change in the cell in which the scFv antibody binds to its target antigen, or detection of the interaction between the scFv antibody and its target antigen (e.g., co-localization of the scFv antibody and its target antigen within a cell).

Exemplary Uses

[0058] Though the libraries described herein have been optimized for the isolation of intrabodies, the libraries can also be used to select scFvs for diagnostic and therapeutic applications. Furthermore, because we designed the CDR3 diversity using expressed human sequences, the scFvs present in the library are fully human and should not induce an anti-scFv antibody response in patients.

[0059] The libraries described herein and the antibodies produced by it are useful not only for identifying functional intrabodies but also for the isolation of highly expressed scFvs that could be used in numerous biotechnological and therapeutic applications. For example, the library and the antibodies produced by it may be have uses related to or including, but not limited to, gene delivery strategies therapeutic agents, drug discovery tools, counteracting agglutination of unwanted target molecules, combating disease states relating to misfolded protein disorders, and binding, neutralizing, or modifying the function of a cancer-related target.

[0060] For example, the intrabodies described herein can be used to modulate the activity of syk tyrosine kinase and other proteins implicated in allergic disorders (see, e.g., WO2005106481; see also Ulanova et al., Expert Opin Ther Targets 2005, 9:901-921. MAP kinases (MAPK) are key mediators of cell proliferation and are often targeted for inhibition in cancer therapy (see, e.g., Sebolt-Leopold JS and Herrera, R, Nat Rev Cancer, 4:937-947 (2004)) Other interesting targets are microtubules and associated proteins (see Jordan MA and Wilson, L, Nat Rev Cancer, 4:253-265 (2004)).

[0061] Immunobodies can be used to treat or prevent diseases in commercially valuable plants, such as crops, e.g., using the methods described in Villani, ME et al. Immunomodulation of cucumber mosaic virus infection by antibodies selected in vitro from a stable single-framework phage display library, Plant Molecular Biology 58(3):305 (2005)).

[0062] The intrabodies described herein can be used to treat, or prevent infections in human or animal cells. For example, intrabodies can be used to treat, ameliorate, or prevent infection of cells by the human immunodeficiency virus, using methods as described, for example, in Swan, CH et al, T-cell protection and enrichment through lentiviral CCR5 gene delivery, Gene Ther. 13:1480-1492.

[0063] The intrabodies described herein can be used to target proteins involved in neurological disorders, e.g., as described, e.g., in Miller, TW et al (A human single-chain Fv intrabody preferentially targets amino-terminal Huntingtin's fragments in striatal models of Huntington's disease, Neurobiol dis. 19:47-56 (2005)) and Miller and Messer (Intrabody applications in neurological disorders: progress and future prospects, Mol. Ther. 12:394-401 (2005).

[0064] The intrabodies described herein can be used for cancer therapy by targeting a protein involved in cancer like oncoproteins, as described, e.g., in Williams, BR and Zhu, Z (Intrabody-based approaches to cancer therapy: status and prospects, Current med Chem 13:1473-1480 (2006) and Doorbar J. and Griffin H. (2007) Intrabody strategies for the treatment of human papillomavirus-associated disease. Expert Opin. Biol. Ther. 7(5), 677-689.

[0065] The intrabodies described herein can be used to treat infections (e.g. but not limited to Epstein-Barr Virus) by targeting a protein expressed by an infectious agent, e.g., as described in Fang CY et al (Modulation of Epstein0-Barr Virus Latent Membrane Protein 1 Activity by Intrabodies, Intervirology 50:254-263 (2007)

[0066] As described herein, an intrabody can be administered to a cell by administering to the cell an expression vector encoding the intrabody of interest. Expression vectors that are suitable for expression of intrabodies are well-known in the art. Administration of expression vectors that encode the intrabodies described herein, can be achieved by any one of numerous, well-known approaches, for example, but not limited to, direct transfer of the nucleic acids, in a plasmid or viral expression vector, alone or in combination with carriers such as cationic liposomes. Such expression vectors (which contain promoter and enhancer sequences suitable for expressing an operably-linked coding sequence when the expression vector is introduced into a cell) and methods for making, using, and delivering such vectors to cells are well known in the art and readily adaptable for use for administering intrabodies to cells.

[0067] Vectors can be any nucleotide construction used to deliver genes into cells, e.g., a plasmid or viral vector, such as a retroviral vector which can package a recombinant retroviral genome (see e.g., Pastan et al., Proc. Natl. Acad. Sci. U.S.A. 85:4486, 1988; Miller et al., Mol. Cell. Biol. 6:2895, 1986). The recombinant retrovirus can then be used to infect and thereby deliver a nucleic acid of the invention to the infected cells. The exact method of introducing the altered nucleic acid into mammalian cells is, of course, not limited to the use of retroviral vectors. Other techniques are widely available for this procedure including the use of adenoviral vectors (Mitani et al., Hum. Gene Ther. 5:941-948, 1994), adeno-associated viral (AAV) vectors (Goodman et al., Blood 84:1492-1500, 1994), lentiviral vectors (Naidini et al., Science 272:263-267, 1996), pseudotyped retroviral vectors (Agrawal et al., Exper. Hematol. 24:738-747, 1996). Physical

transduction techniques can also be used, such as liposome delivery and receptor-mediated and other endocytosis mechanisms (see, for example, Schwartzenberger et al., Blood 87:472-478, 1996). This invention can be used in conjunction with any of these or other commonly used gene transfer methods. Appropriate means for transfection, including viral vectors, chemical transfectants, or physico-mechanical methods such as electroporation and direct diffusion of DNA, are described by, for example, Wolff, J. A., et al., Science, 247, 1465-1468, (1990); and Wolff, J. A. Nature, 352, 815-818, (1991).

[0068]   Cell-permeable intrabodies (transbodies) can be administered to cells by fusing an scFv antibody with a protein transduction domain (PTD) that allows the cell-permeable intrabody to cross the cell membrane and enter the cell, for example, according to the methods described in Heng, BC and Cao, T (Making cell-permeable antibodies (Transbody) through fusion of protein transduction domains (PTD) with single chain variable fragment (scFv) antibodies: potential advantages over antibodies expressed within the intracellular environment (Intrabody), Med Hypotheses 64:1105-1108 (2005)). Alternatively, the scFv could be mixed with cationic lipids and delivered efficiently into cells, as shown with complete antibodies (Courtête J., Sibler A.P., Zeder-Lutz G., Dalkara D., Zuber G. & Weiss E. (2007) Suppression of cervical carcinoma cell growth by intracytoplasmic co-deiivery of anti-oncoprotein E6 antibody and siRNA. Mol. Cancer Ther. 6, 1728-36). Such cell-permeable intrabodies can be used in cell culture (e.g., for research purposes) and for diagnostic purposes (e.g., to detect a virus or microorganism in sample of cells from a human, animal, or plant suspected of harboring an infectious agent. Such cell-permeable intrabodies can be administered to research animals (e.g. but not limited to systemic administration, e.g., by intravenously administering an intrabody to a research animal, or by topical administration, for example) to modulate the activity of a particular target antigen and thereby alter a phenotype in the animal. Such cell-permeable intrabodies can also be administered to human or non-human animal patients to treat or prevent disease or infection as described above. For example, cell-permeable intrabodies can be administered intravenously, topically, orally, or by other well-known methods, as will be appreciated by one of ordinary skill in the art.

Use of the present scFv antibodies for large-scale antibody production

[0069]   Because the scFv antibody clones encode antibodies that fold under reducing conditions, the libraries described herein are also useful for identifying scFv antibodies that can be produced in large quantities by expressing the scFv antibodies in cells (e.g., prokaryotic cells such as bacterial cells, or eukaryotic cells such as yeast cells, insect cells, mammalian cells, or the like) and isolating the scFv antibodies from the cells. Single chain scFv antibodies for which it would be desirable to purify large quantities of antibody include but are not limited to, for example, scFv antibodies that are useful in laboratory research, for medical diagnostic tests, for commercial diagnostic tests or other types of diagnostic tests (e.g., to detect contaminating microorganisms in drinking water or food), and for antibody therapeutics (e.g., to treat cancer or to treat infectious or other types of diseases in which an antibody can be used to treat the disease).

[0070]   Provided herein are antibody libraries for the isolation of expressed scFv antibodies, wherein the antibody library comprises at least about: 10^6, 5×10^6, 10^7, 5×10^7, 10^8, 5×10^8, 10^9, 1.5×10^9, 5×10^9, 10^{10}, 5×10^{10}, 10^{11}, 5×10^{11}, 10^{12}, 10^{13}, 10^{14}, 10^{15}, or 10^{16} unique scFv antibody clones, wherein at least about: 20%, 25%, 30%, 35%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, or more than 80% of the scFv antibody clones can be expressed within a cell to produce soluble antibody at a level of at least about: 5 mgs, 8 mgs, 10 mgs, 15 ms, 20 mgs, 25 mgs, 30 mgs, 35 mgs, 40 mgs, 45 mgs, 50 mgs, or more than 50 mgs per liter of cells in a flask grown to an OD_{600nm} of about 5.

[0071]   There are many well-accepted approaches for expressing and purifying proteins such as expressed scFv antibodies from cells and one of ordinary skill in the art will understand how to select the most appropriate type of cell expression system from which to express and purify scFv antibodies. Many manuals that describe methods for protein expression are known in the art. See, e.g., Current Protocols in Molecular Biology, John Wiley and Sons, 2006.

Definitions

[0072]   By "unique sequence" is meant that, within a collection of scFv antibody clones, there are at least 10^6 clones that contain a CDR3 sequence that is different from the CDR3 sequences of other clones within the collection of scFv antibody clones. With reference to an scFv antibody clone or a library of such clones, a "unique sequence" is different from the sequence present at the corresponding position within scFv13R4.

[0073]   The terms "antibody framework" and "framework sequence" refer to any non-unique portion of the scFv antibody clones, e.g., any portion of the scFv antibody that is not a unique CDR3 VH region and/or unique CDR3 VL region. As referred to herein, the antibody framework or framework sequence is that of scFv13R4, the parental scFv clone upon which the present libraries are based.

[0074]   By "substantially identical" is meant that two or more amino acid sequences being compared are at least 98%, 98.5%, 99%, or 99.5% the same, or that two or more nucleic acid sequences being compared encode amino acid sequences that are at least 98%, 98.5%, 99%, or 99.5% the same.

[0075]   By "common sequence" is meant a nucleotide or amino acid sequence that is shared among scFv clones.

[0076] By "scFv antibody clone" is meant a nucleic acid molecule that encodes an individual species of scFv antibody that comprises a unique sequence within the VH CDR3 domain, the VL CDR3 domain, or within both VH and VL CDR3 domains.

[0077] By "target antigen" is meant an antigen that is preferentially bound by a particular antibody, compared to the binding of that particular antibody to another antigen that is not a target antigen.

[0078] By "specifically binds" binds is meant that an antibody strongly and preferentially binds to a particular target antigen, compared to the binding of the antibody to other antigens.

[0079] By "ectopically expressed" is meant that expression of an antibody of interest is not naturally expressed within a particular type of cell in which the antibody is being expressed, i.e., the antibody is expressed within the cell because an expression construct encoding the antibody has been introduced into the cell.

[0080] By "isolated" or "purified" is meant that an scFv antibody has been substantially separated, produced apart from, or purified away from other biological components in the cell in which it has been produced, that is, substantially separated away from other cellular components, such as other cellular proteins, DNA, RNA, lipids, and the like. The term "isolated" or "purified" does not require absolute purity; rather, it is intended as a relative term. Preferably, an scFv antibody is purified or isolated away from other cellular components such that the scFv antibody represents at least: 25%, 30%, 40%, 50%, 60%, 70% 80%, 90%, 95%, or greater, of the total content of the scFv antibody preparation.

[0081] By "soluble" antibody means that an antibody molecule is not in aggregate form. A soluble antibody has the ability to specifically bind its target antigen.

[0082] By "substantially soluble" is mean that at least 20%, e.g., at least 25%, 30%, 40%, 50%, 60%, 70%, 80%, or more of an scFv antibody as described herein is properly folded and thus can specifically bind its target antigen.

Example 1: Construction of an scFv library optimized for intracellular expression

MATERIALS

Bacterial strains, chemicals and enzymes

[0083] LB and 2xYT media were previously described (Miller, J.H. A Short Course in Bacterial Genetics: a Laboratory Manual and Handbook for Escherichia coli and Related Bacteria. Cold Spring Harbor Laboratory Press; 1992). Strain C-Max5F' (Bio-rad laboratories) is *E. coli* K-12, [F' *lacI*q *Tn*10] φ80d*lacZ*ΔM15 Δ*(lacZYA-argF)*U169 *recA*1 *endA*1 *hsdR*17(r$_k^-$ m$_k^+$) *phoA supE*44 λ- *thi*-1 *gyrA*96 *relA*1. MC1061 (ATCC #53338) is *E. coli* K-12, F-λ- *hsdR*2 *hsdM*+ *hsdS*+ *mcrA mcrB*1 *araD*139 Δ*(ara-leu)*7696 Δ*(lacIPOZY)X*74 *galE*15 *galUgalK*16 *rpsL thi.* Non-suppressor strain HB2151 is *E. coli* K-12 [F'*proA+B+ lacI*q *lacZ*ΔM15] *ara* Δ*(lac-pro) thi.* Chemicals were purchased from Sigma. Restriction enzymes and cloned *Taq* polymerase were from Fermentas. ProofStart and *Pfu* DNA polymerases were respectively purchased from Qiagen and Promega. Plasmid DNA, PCR and agarose-separated DNA were purified using Macherey-Nagel Nucleospin kits.

Oligonucleotides

[0084] Eighteen spiked oligonucleotides used to introduce degenerate CDR3 loops were synthesized and purified using high-pressure liquid chromatography (HPLC) by IBA GmbH (Goettingen, Germany). The sequences of the 18 spiked oligonucleotides is as follows.

[0085] H3_n = n amino acid long VH CDR3 loop; K3_n = n amino acid long VL kappa (κ) CDR3 loop; L3_n = n amino acid long VL lambda (λ) CDR3 loop. For the degenerate positions, the percentages of the 4 bases are given as N(A/C/G/T). The proportions of each nucleotide used at each spiked (degenerate) position is shown in Table I below.

H3_5:

[0086]

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 1);

H3_6:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNTCTCACACAGTAATAAACAGC

CG (SEQ ID NO: 2);

H3_7:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNTCTCACACAGTAATAAAC AGCCG (SEQ ID NO: 3);

H3_8:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNTCTCACACAGTAAT AAACAGCCG (SEQ ID NO: 4);

H3_9:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNTCTCACACAGT AATAAACAGCCG (SEQ ID NO: 5);

H3_10:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNTCTCACAC AGTAATAAACAGCCG (SEQ ID NO: 6);

H3_11:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNTCTCA CACAGTAATAAACAGCCG (SEQ ID NO: 7);

H3_12:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 8);

H3_13:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 9);

H3_14:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 10);

H3_15:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 11);

H3_16:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 12);

H3_17:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 13);

K3_9:

GGACGAGGCTGATTATTACTGCNNNNNNNNNNNNNNNNNNNNNNNNNTTCGGC
GGAGGGACCAAG (SEQ ID NO: 14);

K3_10:

GGACGAGGCTGATTATTACTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTC
GGCGGAGGGACCAAG (SEQ ID NO: 15);

L3_9:

GGACGAGGCTGATTATTACTGCNNNNNNNNNNNNNNNNNNNNNNNNNTTCGGC
GGAGGGACCAAG (SEQ ID NO: 16);

L3_10:

GGACGAGGCTGATTATTACTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTC
GGCGGAGGGACCAAG (SEQ ID NO: 17);

L3_11:

GGACGAGGCTGATTATTACTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
TTCGGCGGAGGGACCAAG (SEQ ID NO: 18).

Table 1.

*For the degenerated positions, the percentages of the 4 bases are given as N(A/C/G/T)

H3_5

AGGGTGCCTCTGCCCCA N(40/5/50/5) N(5/5/5/85) N(60/5/25/10) N(40/10/45/5)
N(15/30/5/50) N(5/75/10/10) N(30/35/30/5) N(50/20/10/20)N(60/20/5/15)
N(30/25/40/5) N(15/20/25/40) N(35/40/5/20) N(25/50/20/5) N(10/50/10/30)
N(5/55/35/5) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 1)

H3_6

AGGGTGCCTCTGCCCCA N(55/5/35/5) N(10/5/5/80) N(75/15/5/5) N(45/5/45/5)
N(5/10/10/75) N(5/85/5/5) N(45/35/15/5) N(55/20/10/15)N(55/20/10/15)
N(40/30/25/5) N(10/40/15/35) N(25/45/5/25) N(45/30/20/5) N(15/40/20/25)
N(10/35/30/25) N(20/15/60/5) N(10/60/10/20) N(5/70/10/15)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 2)

H3_7

AGGGTGCCTCTGCCCCA N(35/5/55/5) N(15/5/15/65) N(65/15/10/10)
N(55/5/35/5) N(5/5/5/85) N(5/85/5/5) N(50/5/40/5) N(75/5/10/10) N(60/10/20/10)
N(10/45/40/5) N(15/35/25/25) N(25/55/5/15) N(35/35/25/5) N(10/45/15/30)
N(25/45/15/15) N(25/25/25/25) N(20/40/15/25)N(10/35/25/30) N(40/15/40/5)
N(5/50/15/30) N(5/80/10/5) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 3)

H3_8

AGGGTGCCTCTGCCCCA N(55/5/35/5) N(10/5/10/75) N(75/10/5/10) N(50/5/40/5)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(75/5/10/10) N(65/5/20/10) N(25/45/25/5)
N(10/25/40/25) N(30/45/5/20) N(55/10/30/5) N(5/55/15/25) N(5/55/15/25)
N(35/35/25/5) N(15/45/20/20) N(15/40/10/35) N(30/40/25/5) N(20/35/20/25)
N(15/40/20/25) N(65/5/20/10) N(15/45/5/35) N(5/70/15/10)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 4)

H3_9

AGGGTGCCTCTGCCCCA N(40/5/50/5) N(10/5/10/75) N(75/5/15/5) N(40/5/50/5)

N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(70/5/15/10) N(55/25/15/5)
N(10/25/35/30) N(35/45/15/5) N(10/25/60/5) N(10/35/20/35) N(20/45/15/20)
N(30/30/35/5) N(15/40/20/25) N(20/40/15/25) N(65/20/10/5) N(20/40/20/20)
N(5/45/15/35) N(15/30/50/5) N(20/45/20/15) N(5/40/25/30) N(30/25/40/5)
N(10/40/10/40) N(5/60/20/15)TCTCACACAGTAATAAACAGCCG (SEQ ID NO:
5)

H3_10

AGGGTGCCTCTGCCCCA N(50/5/40/5) N(20/5/10/65) N(70/10/15/5) N(60/5/30/5)
N(5/5/5/85) N(5/85/5/5) N(35/5/55/5) N(85/5/5/5) N(75/5/15/5) N(30/35/30/5)
N(10/20/45/25) N(30/45/15/10) N(30/35/30/5) N(5/30/20/45) N(30/40/10/20)
N(40/35/20/5) N(15/35/25/25) N(25/35/15/25) N(50/30/15/5) N(10/45/20/25)
N(20/50/10/20) N(25/20/50/5) N(20/40/20/20) N(5/50/15/30) N(20/20/40/20)
N(20/35/20/25) N(5/35/30/30)N(5/30/60/5) N(15/35/10/40) N(5/60/15/20)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 6)

H3_11

AGGGTGCCTCTGCCCCA N(55/5/35/5) N(10/5/10/75) N(80/5/10/5) N(60/5/30/5)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(75/10/10/5) N(40/30/25/5)
N(10/20/35/35) N(35/35/20/10) N(40/30/25/5) N(10/25/20/45) N(35/35/15/15)
N(35/30/30/5) N(15/30/25/30) N(35/30/15/20) N(20/35/40/5) N(20/30/25/25)
N(30/40/10/20) N(5/35/50/10) N(20/30/30/20) N(30/40/15/15) N(35/35/25/5)
N(20/35/20/25) N(20/40/15/25) N(45/25/10/20) N(20/35/20/25) N(20/30/25/25)
N(45/25/25/5) N(15/35/10/40) N(5/60/20/15) TCTCACACAGTAATAAACAGCCG
(SEQ ID NO: 7)

H3_12

AGGGTGCCTCTGCCCCA N(45/5/45/5) N(15/5/15/65) N(70/5/15/10) N(40/5/50/5)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(75/5/15/5) N(10/30/55/5)
N(10/30/25/35) N(40/40/15/5) N(25/25/45/5) N(10/25/20/45) N(35/35/10/20)
N(40/40/15/5) N(20/20/35/25) N(40/30/15/15) N(10/40/45/5) N(20/25/30/25)
N(35/35/15/15) N(30/30/35/5) N(15/40/20/25) N(20/45/5/30) N(15/25/40/20)
N(20/30/25/25) N(20/40/15/25)N(15/25/40/20) N(20/35/20/25) N(25/40/15/20)
N(50/30/15/5) N(20/45/20/15) N(5/35/35/25) N(45/25/25/5) N(15/35/10/40)
N(5/65/15/15) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 8)

H3_13

AGGGTGCCTCTGCCCCA N(45/10/40/5) N(15/5/15/65) N(75/5/15/5) N(35/5/55/5)

N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(80/5/10/5) N(35/25/35/5)
N(10/30/25/35) N(40/45/10/5) N(15/25/55/5) N(10/25/10/55) N(35/35/10/20)
N(55/30/10/5) N(15/20/25/40) N(45/25/10/20) N(25/30/40/5) N(20/30/20/30)
N(30/35/15/20) N(15/25/55/5) N(15/35/20/30) N(25/45/10/20) N(10/45/40/5)
N(20/25/25/30) N(30/40/10/20)N(30/25/30/15) N(20/25/25/30) N(35/35/15/15)
N(25/30/30/15) N(20/30/20/30) N(25/35/20/20) N(40/30/25/5) N(20/45/20/15)
N(5/35/35/25) N(45/25/25/5) N(10/35/15/40) N(5/65/20/10)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 9)

H3_14

AGGGTGCCTCTGCCCCA N(50/5/40/5) N(20/5/15/60) N(60/10/25/5) N(35/5/55/5)
N(5/5/5/85) N(5/85/5/5) N(30/20/45/5) N(85/5/5/5) N(80/5/5/10) N(30/35/30/5)
N(10/35/20/35) N(40/40/15/5) N(50/25/20/5) N(10/25/10/55) N(40/30/10/20)
N(20/20/55/5) N(15/20/20/45) N(45/25/15/15) N(15/25/55/5) N(15/30/25/30)
N(35/30/15/20) N(40/30/25/5) N(15/25/25/35) N(35/35/10/20) N(20/30/40/10)
N(15/30/25/30) N(25/40/10/25)N(50/25/15/10) N(15/35/25/25) N(35/35/10/20)
N(15/25/45/15) N(15/30/25/30) N(35/35/10/20) N(15/30/40/15) N(20/30/20/30)
N(30/40/10/20) N(30/30/15/25) N(20/35/25/20) N(20/30/30/20) N(30/25/40/5)
N(10/30/15/45) N(5/70/15/10) TCTCACACAGTAATAAACAGCCG (SEQ ID NO:
10)

H3_15

AGGGTGCCTCTGCCCCA N(30/5/45/20) N(45/5/5/45) N(45/35/5/15) N(45/5/45/5)
N(5/5/5/85) N(5/85/5/5) N(45/20/30/5) N(85/5/5/5) N(85/5/5/5) N(20/35/40/5)
N(5/40/25/30) N(35/50/10/5) N(15/5/75/5) N(10/10/20/60) N(50/30/10/10)
N(25/10/55/10) N(15/20/20/45) N(50/25/15/10) N(35/30/30/5) N(15/25/25/35)
N(50/20/15/15) N(35/5/35/25) N(10/30/30/30) N(35/35/10/20) N(20/40/35/5)
N(20/20/35/25) N(35/35/15/15)N(45/35/15/5) N(15/30/25/30) N(35/35/10/20)
N(35/35/25/5) N(20/30/20/30) N(30/35/10/25) N(30/25/40/5) N(15/30/20/35)
N(35/30/15/20) N(35/25/35/5) N(15/35/20/30) N(25/35/25/15) N(5/30/45/20)
N(20/35/20/25) N(15/30/30/25) N(45/25/25/5) N(15/35/10/40) N(5/70/15/10)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 11)

H3_16

AGGGTGCCTCTGCCCCA N(35/10/30/25) N(65/5/5/25) N(25/55/5/15)
N(45/5/45/5) N(5/5/5/85) N(5/85/5/5) N(35/30/30/5) N(85/5/5/5) N(75/5/5/15)
N(15/35/45/5) N(5/45/20/30) N(35/55/5/5) N(50/5/40/5) N(10/5/15/70)
N(60/20/10/10) N(65/5/25/5) N(10/10/25/55) N(55/15/15/15) N(30/5/60/5)
N(10/15/30/45) N(55/15/15/15) N(25/25/45/5) N(20/30/15/35) N(40/35/10/15)
N(40/20/30/10) N(20/30/25/25) N(35/30/10/25)N(45/30/15/10) N(15/35/25/25)

N(30/35/10/25) N(35/30/25/10) N(20/25/30/25) N(35/35/10/20) N(30/20/30/20)
N(20/35/20/25) N(30/40/10/20) N(45/10/25/20) N(20/25/20/35) N(30/35/15/20)
N(45/30/20/5) N(15/35/25/25) N(25/35/20/20) N(40/15/5/40) N(20/35/20/25)
N(15/30/30/25) N(30/30/35/5) N(15/30/15/40) N(5/75/15/5)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 12)

H3_17

AGGGTGCCTCTGCCCCA N(45/10/20/25) N(70/5/5/20) N(15/60/5/20)
N(40/5/50/5) N(5/5/5/85) N(5/85/5/5) N(25/20/50/5) N(85/5/5/5) N(80/5/5/10)
N(10/45/40/5) N(5/50/20/25) N(30/60/5/5) N(5/5/85/5) N(5/5/15/75) N(65/20/5/10)
N(45/5/45/5) N(10/10/20/60) N(55/10/20/15) N(45/5/45/5) N(15/15/20/50)
N(60/15/15/10) N(60/20/15/5) N(10/30/20/40) N(45/25/15/15) N(15/35/45/5)
N(15/15/35/35) N(40/30/15/15) N(30/25/40/5)N(20/30/25/25) N(30/35/10/25)
N(20/30/40/10) N(15/35/25/25) N(30/35/10/25) N(40/30/25/5) N(20/30/25/25)
N(40/35/10/15) N(35/20/40/5) N(20/25/20/35) N(40/35/10/15) N(35/20/40/5)
N(20/20/20/40) N(35/35/15/15) N(30/20/30/20) N(20/30/20/30) N(25/30/25/20)
N(35/35/25/5) N(25/35/20/20) N(5/35/40/20) N(20/30/45/5) N(15/30/10/45)
N(5/70/15/10) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 13)

K3_9

GGACGAGGCTGATTATTACTGC N(5/85/5/5) N(85/5/5/5) N(5/5/85/5)
N(5/85/5/5) N(85/5/5/5) N(40/5/50/5) N(5/5/10/80) N(65/25/5/5) N(5/50/5/40)
N(30/5/25/40) N(60/10/20/10) N(5/40/10/45) N(80/5/10/5) N(30/15/50/5)
N(5/60/5/30) N(35/5/10/50) N(30/40/10/20) N(5/15/45/35) N(5/85/5/5) N(5/85/5/5)
N(5/5/85/5) N(5/40/5/50) N(30/10/25/35) N(5/35/50/10) N(85/5/5/5) N(5/85/5/5)
N(5/50/30/15) TTCGGCGGAGGGACCAAG (SEQ ID NO: 14)

K3_10

GGACGAGGCTGATTATTACTGC N(5/85/5/5) N(85/5/5/5) N(5/5/85/5)
N(5/85/5/5) N(85/5/5/5) N(40/5/50/5) N(5/5/5/85) N(80/10/5/5) N(25/20/5/50)
N(35/5/40/20) N(45/5/45/5) N(5/30/15/50) N(75/10/10/5) N(40/5/50/5) N(5/45/5/45)
N(20/5/5/70) N(10/45/40/5) N(5/5/85/5) N(5/85/5/5) N(5/85/5/5) N(5/5/85/5)
N(5/85/5/5) N(5/65/15/15) N(5/5/85/5) N(15/5/10/70) N(20/5/30/45) N(5/30/50/15)
N(85/5/5/5) N(5/85/5/5) N(5/40/5/50)TTCGGCGGAGGGACCAAG (SEQ ID NO:
15)

L3_9

GGACGAGGCTGATTATTACTGC N(5/85/5/5) N(80/5/5/10) N(5/5/85/5)
N(20/5/30/45) N(5/75/5/15) N(85/5/5/5) N(5/5/5/85) N(35/5/55/5) N(5/10/80/5)
N(5/10/80/5) N(80/10/5/5) N(5/60/5/30) N(65/5/25/5) N(15/10/70/5) N(5/65/5/25)
N(70/5/20/5) N(20/10/65/5) N(5/60/5/30) N(65/10/20/5) N(15/45/20/20)
N(5/40/10/45) N(15/5/50/30) N(10/15/25/50) N(5/20/70/5) N(10/5/80/5) N(5/5/5/85)
N(5/40/10/45)TTCGGCGGAGGGACCAAG (SEQ ID NO: 16)

L3_10

GGACGAGGCTGATTATTACTGC N(45/5/10/40) N(15/15/60/10) N(10/50/5/35)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(5/5/5/85) N(85/5/5/5) N(5/50/5/40)
N(25/5/65/5) N(45/45/5/5) N(5/45/5/45) N(55/5/35/5) N(10/10/75/5) N(5/40/5/50)
N(85/5/5/5) N(15/10/70/5) N(5/50/5/40) N(70/5/20/5) N(30/20/45/5) N(5/75/5/15)
N(70/15/10/5) N(25/45/15/15) N(5/35/5/55) N(5/15/35/45) N(15/5/30/50)
N(5/15/75/5) N(10/5/80/5) N(5/5/5/85) N(5/50/10/35) TTCGGCGGAGGGACCAAG
(SEQ ID NO: 17)

L3_11

GGACGAGGCTGATTATTACTGC N(10/55/25/10) N(65/20/10/5) N(5/10/80/5)
N(20/5/20/55) N(5/85/5/5) N(5/5/85/5) N(5/5/5/85) N(30/5/60/5) N(5/5/85/5)
N(5/5/85/5) N(85/5/5/5) N(5/45/5/45) N(75/5/15/5) N(20/10/65/5) N(5/60/5/30)
N(85/5/5/5) N(5/5/85/5) N(5/35/5/55) N(5/5/5/85) N(5/30/10/55) N(85/5/5/5)
N(75/5/15/5) N(45/20/30/5) N(5/20/5/70) N(5/30/50/15) N(25/25/35/15)
N(5/45/15/35) N(10/5/50/35) N(20/5/25/50) N(5/20/60/15) N(5/5/85/5) N(5/5/5/85)
N(70/5/20/5)TTCGGCGGAGGGACCAAG (SEQ ID NO: 18)

[0087]   Other oligonucleotides use to construct the libraries were synthesized by MWG and have the following sequences:

T7.back CCGGATATAGTTCCTCCTTT (SEQ ID NO: 19);
T7.for CTGCTAACCAGTAAGGCAAC (SEQ ID NO: 20);
M13rev-49 GAGCGGATAACAATTTCACACAGG (SEQ ID NO: 21);
M13uni-43 AGGGTTTTCCCAGTCACGACGTT (SEQ ID NO: 22);
scFvCAT.rev AACGGTGGTATATCCAGTGA (SEQ ID NO: 23);
scFvCAT2.rev CGGTGGTATATCCAGTGATTTT (SEQ ID NO: 24);
PliaisonH3 TGGGGCAGAGGCACCCT (SEQ ID NO: 25);
PliaisonH3.back AGGGTGCCTCTGCCCCA (SEQ ID NO: 26);
PliaisonL3 GCAGTAATAATCAGCCTCGTCC (SEQ ID NO: 27).

Plasmids

[0088]   Phagemid vector pCANTAB6 (McCafferty J et al., Appl Biochem Biotechnol 1994, 47:157-171) was used for N-terminal fusion of *Ncol/Notl-scFv* fragments to the minor coat protein pIII of filamentous phage M13. This phagemid is derived from pUC 119 and contains the following sequences in the following order: a *lac* promoter, the *pelB* leader sequence, *Ncol* and *Notl* sites for scFv cloning, a His6 and a c-myc tag recognized by the 9E10 monoclonal antibody, an amber codon and the pIII gene sequence.

[0089]   For cytoplasmic expression of the scFvs in *E. coli* we used plasmid pET23NN. This plasmid is derived from pET23d(+) (Novagen) and contains a T7 promoter, followed by a *Ncol* site containing the ATG initiator, a *Notl* site followed by a *c*-myc and a His6 tag.

[0090]   Plasmid pscFv CAT is derived from pTrc99A and contains a *tac* promoter, followed by a *Ncol* site containing

the ATG initiator of an out-of-frame scFv, a *Not*I site followed by the CAT gene. When a scFv is inserted within the *Nco*I-*Not*I sites, the scFv is expressed as a fusion with the CAT protein. The construction was done as follows: first, the unique *BstE*II site of pTrc99A (A13038) was removed by digestion followed by 5' overhang fill-in to form blunt ends, and ligation. The resulting plasmid was digested with *Nco*I and *Not*I, and the 4210bp fragment purified (fragment I). Second, the unique *Nco*I site of plasmid pACYCI 84 (X06403) located within the CAT gene was removed by site directed mutagenesis by changing the Thr 172 codon from ACC to ACG. Then the CAT gene was amplified by PCR using CAT-NotI.for (TAAG<u>GCGGCCGCA</u>***ATGGAGAAAAAAATCACTG***; SEQ ID NO: 28) and CAT-HindIII.back (ACTGCCTTAAA<u>AAGC**T-TACGCC**</u>; SEQ ID NO: 29). In the oligonucleotide sequences, the introduced restriction sites are underlined and the beginning and the end of the CAT gene are in bold-italic. The 660 bp PCR fragment was digested by *Not*I and *Hind*III, and purified (fragment II). Third, a 750 bp *Nco*I-*Not*I scFv13E6 fragment (a grafted version of the scFv13R4 containing the CDR loops of an anti-E6 monoclonal antibody. Philibert *et al.,* to be published) was purified (fragment III). Fourth, the three fragments I, II and III were ligated to give plasmid pscFvCAT. Finally, an internal deletion of 165 bp was introduced in the scFv by removing the fragment. between the two *Pst*I sites of the gene. The resulting plasmid, called pscFvCAT, is Amp^R and CAM^s since the deletion of the *Pst*I fragment resulted in a frameshift in the scFv.

[0091] Plasmid p513-EGFP is a derivative of pSG5 (Green, S., et al. Nucleic Acids Res 1988, 16:369) and harbors the EGFP coding region (Clontech, Inc.) under the control of the SV40 early promoter. The p513-scFv-EGFP constructs correspond to in frame fusions of the scFv's and the EGFP coding region with a linker of 10 residues.
[0092] The scFv coding regions were amplified with oligonucleotide primers

5'-ACTGATAAGCTTGCCACCATGGCCGAGGTGC (SEQ ID NO: 30)
and 5'-TTGATTACTAGTGAGTTTTTGTTCTGCGGCC (SEQ ID NO: 31)

and inserted into the *HindIII-SpeI* digested p51 3-EGFP vector.

Optimized Antibody Framework

[0093] To maximize the effectiveness of the scFv library, construction of the library begain with the selection of a single optimized antibody framework for intrabody selection. Through molecular evolution, a human scFv called scFv13R4 was obtained, which is expressed at high levels in *E. coli* cytoplasm. This scFv is also expressed and has a soluble and active conformation in both yeast and mammalian cells. This scFv is very stable *in vitro* and can be renatured in presence of a reducing agent. In addition, analysis of its folding kinetics showed that it folds faster than the parent scFv and aggregates more slowly *in vitro.* The mutations isolated are mainly located in the VH domain and seem to be highly specific to this particular scFv since they cannot be transferred to a very homologous VH domain. The nucleotide and amino acid sequences of scFv13R4 are shown below.

```
Nucleotide and Amino Acid Sequence of scFv13R4 (from nucleotides 1 to 819)

      ---------|---------|---------|---------|---------|---------|
    1 atggccgaggtgcagctggtggagtctggggggaagcctggtcaagcctggggggtccctg 60
    1 M   A   E   V   Q   L   V   E   S   G   G   S   L   V   K   P   G   G   S   L   20

      ---------|---------|---------|---------|---------|---------|
   61 agactctcctgtgcagcctctggattcaccttcagtaactatagcatgaactgggtccgc 120
   21 R   L   S   C   A   A   S   G   F   T   F   S   N   Y   S   M   N   W   V   R   40
```

```
         ---------|---------|---------|---------|---------|---------|
     121 caggctccagggaaggggctggagtggatctcatccattagtggtagtagtagatacata 180
      41 Q   A   P   G   K   G   L   E   W   I   S   S   I   S   G   S   S   R   Y   I   60

         ---------|---------|---------|---------|---------|---------|
     181 tactacgcagacttcgtgaagggccgattcaccatctccagagacaacgccacgaactca 240
      61 Y   Y   A   D   F   V   K   G   R   F   T   I   S   R   D   N   A   T   N   S   80

         ---------|---------|---------|---------|---------|---------|
     241 ctgtacctgcaaatgaacagcctgagagccgaggacacggctgtttattactgtgtgaga 300
      81 L   Y   L   Q   M   N   S   L   R   A   E   D   T   A   V   Y   Y   C   V   R   100

         ---------|---------|---------|---------|---------|---------|
     301 tccagtattacgatttttggtggcggtatggacgtctggggcagaggcaccctggtcacc 360
     101 S   S   I   T   I   F   G   G   G   M   D   V   W   G   R   G   T   L   V   T   120

         ---------|---------|---------|---------|---------|---------|
     361 gtctcctcaggtggaggcggttcaggcggaggtggcagcggcggtggcggatcgcagtct 420
     121 V   S   S   G   G   G   G   S   G   G   G   G   S   G   G   G   G   S   Q   S   140

         ---------|---------|---------|---------|---------|---------|
     421 gtgctgactcagcctgcctccgtgtctgggtctcctggacagtcgatcaccatctcctgc 480
     141 V   L   T   Q   P   A   S   V   S   G   S   P   G   Q   S   I   T   I   S   C   160

         ---------|---------|---------|---------|---------|---------|
     481 gctggaaccagcagtgacgttggtggttataactatgtctcctggtaccaacaacaccca 540
     161 A   G   T   S   S   D   V   G   G   Y   N   Y   V   S   W   Y   Q   Q   H   P   180

         ---------|---------|---------|---------|---------|---------|
     541 ggcaaagccccccaaactcatgatttatgaggacagtaagcggccctcaggggtttctaat 600
     181 G   K   A   P   K   L   M   I   Y   E   D   S   K   R   P   S   G   V   S   N   200

         ---------|---------|---------|---------|---------|---------|
     601 cgcttctctggctccaagtctggcaacacggcctccctgacaatctctgggctccaggct 660
     201 R   F   S   G   S   K   S   G   N   T   A   S   L   T   I   S   G   L   Q   A   220

         ---------|---------|---------|---------|---------|---------|
     661 gaggacgaggctgattattactgcagctcatatacaaccaggagcactcgagttttcggc 720
     221 E   D   E   A   D   Y   Y   C   S   S   Y   T   T   R   S   T   R   V   F   G   240

         ---------|---------|---------|---------|---------|---------|
     721 ggagggaccaagctggccgtcctaggtgcggccgcagaacaaaaactcatctcagaagag 780
     241 G   G   T   K   L   A   V   L   G   A   A   A   E   Q   K   L   I   S   E   E   260

         ---------|---------|---------|---------
     781 gatctgaatggggccgcacatcaccatcatcaccattaa 819   (SEQ ID NO: 32)
     261 D   L   N   G   A   A   H   H   H   H   H   *   272   (SEQ ID NO: 33)
```

METHODS

Database of CDR3 sequences

[0094]   Release 5 (August, 1992) of the Kabat database was used (Johnson, G., and Wu, T.T.: Kabat Database and its applications: 30 years after the first variability plot. Nucleic Acids Res 2000, 28:214-218). This dataset contained 44990 sequences. First, 4643 human VH sequences which were not a pseudogene and were not humanized were extracted. H3 sequences were then extracted from this dataset, first taking into account the nucleotide sequence when present, then the amino acid sequence. Finally, the 3469 complete H3 sequences that contained only the 20 regular amino acids were kept, among which 2703 were unique. The same procedure was followed for λ and κ light chains, respectively, resulting in 1044 and 1291 sequences from which 775 and 828 were unique.

[0095]   CDR3 sequences from the IMGT/LIGM-DB database as it existed on 27 November 2003 were also extracted (Giudicelli, V., et al. IMGT/LIGM-DB, the IMGT(R) comprehensive database of immunoglobulin and T cell receptor

nucleotide sequences. Nucleic Acids Res 2006, 34:D781-784). Only the "productive/regular/human/cDNA+RNA/rear-ranged" genes were considered. 5179 H3, 1432 K3, and 1131 L3 sequences were obtained, of which 4323 H3, 974 K3, and 812 L3 sequences were unique.

[0096] 127 additional human antibody sequences were also collected from the Protein data bank (Berman, H.M., et al. The Protein Data Bank. Nucleic Acids Res 2000, 28:235-242.). For this we used the file of 510 sequences already compiled by Andrew Martin on August 19, 2003 (Allcorn, L.C., and Martin, A.C.R. SACS-self-maintaining database of antibody crystal structure information. Bioinformatics 2002, 18:175-181).

Spiked oligonucleotide design

[0097] In biasing the representations of the amino acids, optimized mixtures of the nucleotides at each of the three codon positions were calculated as described previously (Wang, W., and Saven, J.G. Designing gene libraries from protein profiles for combinatorial protein experiments. Nucleic Acids Res 2002, 30:e120; Park, S., et al. Progress in the development and application of computational methods for probabilistic protein design. Comput Chem Eng 2005, 29:407-421.). Premature termination of protein sequences was limited by imposing an upper bound of 0.05 on the probability of realizing a stop codon. For the 34 positions which did not satisfactorily recover the desired probabilities of the amino acids, a second optimization was done with the same method, but with no constraint on the stop codon frequency. For oligonucleotide synthesis, the calculated frequencies were rounded in increments of 5% as follows: all the frequencies between 0% and 5% were rounded to 5%; other frequencies were rounded to the nearest 5%; if the resulting sum was higher than 100%, 5% was removed from the rounded amino acid frequency larger than 5% for which the difference between the rounded and the target frequency was maximal and the process iterated until the sum was 100%; if the sum was lower than 100%, 5% was added to the rounded frequency lower than 95%, for which the difference between the rounded and the target frequency was maximal and the process iterated until the sum was 100%.

Construction of VH and VL libraries

[0098] Variable CDR3 sequences were introduced in scFv13R4 by PCR assembly using a hot-start proofreading polymerase (ProofStart, Qiagen) using as a template plasmid pAB 1-scFv13R4p (Martineau, P., and Betton, J.M. In vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm. J Mol Biol 1999, 292:921-929.) To introduce random H3 loops, the 5' of the gene with the random H3 sequence was obtained with oligonucleotides M13rev-49 and one of the 13 degenerate oligonucleotides, and the 3' was obtained with PliaisonH3 and M13uni-43 (both for 20 cycles at 55°C). The two purified bands were thus assembled by PCR (30 cycles, 55°C) using M13rev-49 and M 13uni-43. The resulting PCR was purified using a commercial kit (Nucleospin, Macherey-Nagel), digested for 16 hours at 37°C with *Nco*I and *Not*I enzymes, and then purified on a gel. The same procedure was followed to introduce random L3 and K3 loops except that the pairs of primers used were M13rev-49/PliaisonL3 for the 5' and one of the degenerate oligonucleotides encoding the L3/K3 loop (K3_n or L3_n) with M13uni-43 for the 3' part of the gene.

[0099] Each digested band was ligated for 16 hours at 16°C with 1 µg of *Nco*I-*Not*I, digested, and purified pscFvΔCAT in 100 µl using 10 Weiss units of T4 DNA ligase. The ligation was heat inactivated and purified using a commercial kit (Nucleospin). The ligation was then electroporated in 300 µl of MC 1061 competent cells (Sidhu, S.S., et al. Phage display for selection of novel binding peptides. Methods Enzymol 2000, 328:333-363) and plated on a 600 cm^2 square plate of LB containing 100 µg/ml of ampicillin, then incubated for 16 hours at 37°C. The 18 libraries (13 VH and 5 VL) were scrapped in 10 ml of LB with 10% glycerol, and 10^9 bacteria were immediately plated on a 600 cm^2 square plate of LB containing 100 µg/ml of ampicillin, 1 mM IPTG and 30 µg/ml of CAM, then incubated for 16 hours at 37°C. The 18 libraries were then scrapped in 10 ml of LB with 10% glycerol and frozen at -80°C. An aliquot was used to prepare DNA for the library assembly.

Library assembly

[0100] The 13 VH libraries were amplified using primers M13rev-49/PliaisonH3.back using Pfu polymerase, and the 5 VL libraries were amplified using scFvCAT.rev/H3_Liaison (30 cycles, 55°C). The 18 PCR bands were first purified, then carefully quantified on gel using ImageJ software. The 13 VH bands were pooled in amounts proportional to their frequency in human H3. This mix was called VHpool. The 2 VL κ bands were pooled in order to obtain 75% of 9 amino acid-length loops and 25% of 10 amino acid-length loops. The VL λ bands were pooled to obtain 30% of the 9, 30% of the 10, and 40% of the 11 amino acid-length loops. Finally, the κ and λ mixes were pooled in order to get 50% of each class in the final mix called VLpool.

[0101] VHpool and VLpool were assembled by PCR using *Taq* DNA polymerase and primers M13rev-49/scFvCAT2.rev in 500 µl (30 cycles, 55°C). The PCR was successively digested with 20 units of *Nco*I and *Not*I for at least 6 h each,

purified, and then quantified on gel. 50 μg of vector pCANTAB6 was successively digested with 80 units of *Nco*I and *Not*I for at least 6 h each, purified then quantified on gel. 5 μg of linearized pCANTAB6 was ligated with an equal molar amount of insert (0.84 μg) in 500 μl at 16°C using 50 Weiss units of T4 DNA ligase. The ligation was heat inactivated and purified using a commercial kit (Nucleospin). The purified ligation was then electroporated in 10×300 μl of C-Max5αF' competent cells (Sidhu, S.S., et al. Phage display for selection of novel binding peptides. Methods Enzymol 2000, 328:333-363), and plated on ten 600 cm^2 square plate of LB containing 1% of glucose and 100 μg/ml of ampicillin. After incubation for 16 h at 37°C, cells were scrapped in 2xYT containing 10% of glycerol and kept frozen at -80°C in aliquots corresponding to twenty times the diversity.

Antigens

[0102] Aurora-A is an His-tagged protein and was produced in *E.coli*. GST:Syk was expressed in *E.coli*. (Dauvillier, S., et al. Intracellular single-chain variable fragments directed to the Src homology 2 domains of Syk partially inhibit Fc epsilon RI signaling in the RBL-2H3 cell line. J Immunol 2002, 169:2274-2283). E6 protein from papillomavirus HPV16 was expressed in cyanobacterium Anabaena (Desplancq *et al.,* to be published). Histones (a mix of H2a, H2b, H3 and H4) were purchased from Sigma (type II-AS. #H7755). Tubulin was purified from pig brain (Williams, R.C.J., and Lee, J.C. Preparation of tubulin from brain. Methods Enzymol 1982, 85 (Pt B):376-385).

Library rescue and selection

[0103] Library rescue was done essentially as previously described using a trypsin-sensitive helper phage (Kristensen, P., and Winter, G. Proteolytic selection for protein folding using filamentous bacteriophages. Fold Des 1998, 3:321-328). Briefly, an aliquot of the library corresponding to a 10 to 20-fold excess over the diversity (2-3 x 10^{10} bacteria) was inoculated in 1000 ml of 2xYT containing 100 μg/ml ampicillin and 1% glucose and grown with shaking at 37°C until OD$_{600nm}$ was 0.7. 200 ml ($\sim 3 \times 10^{10}$ cells) were infected with 5×10^{11} helper phage KM 13 (Kristensen, P., and Winter, G. Proteolytic selection for protein folding using filamentous bacteriophages. Fold Des 1998, 3:321-328) and incubated without shaking for 30 min at 37°C. Cells were pelleted, resuspended in 1000 ml of 2xYT containing 100 μg/ml ampicillin and 25 μg/ml kanamycin and incubated overnight with vigorous shaking at 30°C. The supernatant containing phages was precipitated twice by adding 1/5th of the volume of PEG-8000 20%, NaCl 2.5 M, and resuspended in PBS supplemented with 15% of glycerol. Aliquots containing 10^{11}-10^{12} phages were stored at -80°C.

[0104] To select for binders, 100 μl of purified antigens were coated in a Nunc Maxisorp 96-well plate. For the first round, an antigen concentration of 10-100 μg/ml was used. For subsequent rounds, an antigen concentration of 1-10 μg/ml was used. The plate was washed 3 times with PBS containing 0.1% of Tween20 (PBST) and saturated for 2 hours at room temperature with PBS containing 2% non-fat milk (MBPS). 10^{11}-10^{12} phages were added per well in 2% MPBS and incubated for 2 hours at room temperature. The plate was washed 20 times (first round) or 40 times (2nd and 3rd rounds) with PBST, and then washed 3 times with PBS. Excess PBS was removed, and the phages were eluted by adding 100 μl of 100 mM triethylamine for 10 minutes at room temperature. The eluted phage suspension was neutralized with 50 μl of 1 M Tris-HCl pH 7.4, then digested 15 minutes at room temperature with trypsin by adding 1.5 μl of 0.1 M CaCl$_2$ and 15 μl of 10 mg/ml TPCK-treated trypsin (Sigma). 1 ml of a 37°C exponentially growing Cmax5aF' strain in 2xYT was infected with 40 μl of trypsin-treated phages, incubated 30 min at 37°C without shaking, then plated on a 15 cm round Petri dish (LB, 100 μg/ml ampicillin, 1% glucose). After overnight incubation at 37°C, bacteria were recovered from the plate and used to prepare a new stock of phages using KM13 helper phage. 10^{11}-10^{12} phages of this stock were used for the next round of selection.

Periplasmic and cytoplasmic screening

[0105] For periplasmic screening, phages from round 3 were used to infect the non-suppressive strain HB2151. Individual clones were tested for scFv expression by ELISA on antigen-coated 96-well microtiter plates as described (Harrison, J.L., et al. Screening of phage antibody libraries. Methods Enzymol 1996, 267:83-109.) For cytoplasmic screening, plasmid was prepared from the pool of bacteria of the 2nd or 3rd selection round, digested with *Nco*I and *Not*I enzymes, and the 750 bp band was cloned in *Nco*I-*Not*I digested and dephosphorylated plasmid pET23NN. Ligation was transformed in C-Max5αF', and the cells were plated on LB with 100 μg/ml ampicillin and incubated for 16 hours at 37°C. Cells were scrapped, and the plasmid DNA was prepared and used to transform chemically competent BL21(DE3) pLysS. Individual clones were grown in a 96-well microtiter plate containing 100 μl of 2xYT, 100 μg/ml of ampicillin with vigorous shaking at 37°C until OD$_{600nm}$ reached 0.6. IPTG was added to 0.4 mM final and the microtiter plate was incubated for 16 hours at 24-30°C with vigorous shaking in a humidified atmosphere. After centrifugation, cells were resuspended in 100 μl of 50 mM Tris-HCl pH7.5, 5 mM EDTA, freeze/thawed, and incubated 1 hour on ice. MgCl$_2$ was added up to 10 mM and the DNA was digested with 10 μg/ml of DNaseI. 5-20 μl were used in an ELISA on an antigen-

coated 96-well microtiter plate (Nunc Maxisorp). Revelation was done using 9E10 monoclonal antibody followed by an HRP conjugated anti-mouse IgG antibody.

Purification of scFv

[0106] scFvs cloned in plasmid pET23NN were purified from the cytoplasm of BL21(DE3) pLysS and purified on a Ni-NTA column as described for the parental scFv13R4 (Martineau, P., et al. Expression of an antibody fragment at high levels in the bacterial cytoplasm. J Mol Biol 1998,280:117-127.).

Cell transfection and immunofluorescence

[0107] HeLa cells were maintained in Dulbecco's modified Eagle's tissue culture medium (DMEM; Invitrogen) supplemented with L-glutamine (2 mM), penicillin (100 IU/ml), streptomycin (25 μg/ml) and 10% heat-inactivated fetal calf serum at 37°C in a humidified 5% CO_2 atmosphere. Transient transfection was carried out with the TransFectin lipid reagent (Bio-Rad, Hercules, CA, USA) according to the manufacturer's instructions. Cells were seeded on coverslips in 6-well plates at 2.5×10^5 cells/well the day before transfection. 1 μg DNA and 2 μl of reagent diluted in 100 μl of DMEM were mixed and left at room temperature for 20 minutes. Cells were grown at 37°C for 24 hours after addition of the mixture. The expressed GFP-tagged proteins were visualized after fixation of the transfected cells with 4% paraformaldehyde in PBS for 45 minutes at room temperature. After extensive washing with PBS, cells were dried and mounted with Fluoromount-G (SouthernBiotech, Birmingham, UK). The processed cells were examined with a Zeiss Axioplan fluorescence microscope equipped with an Olympus DP50 camera. Images were collected with a Zeiss 40× planneofluar objective and processed using Adobe Photoshop 5.5. For Figure 8, HeLa were transfected with anti-histones clone 5 fused to the dsRed-monomer GFP, fixed as above and permeabilized with Triton ×-100 (0.2%, 5 min). The microtubule network was revealed with the 2F12C scFv (Table 3) at 3 μg/ml using the 9E10 anti-myc and an Alexa Fluor 488 anti-mouse IgG antibody. Cells were observed by confocal microscopy (×63).

RESULTS

Step 1: Selection of Antibody Framework

[0108] In order to maximize the effectiveness of the scFv library, construction of the library begain with the selection of a single optimized antibody framework for intrabody selection. Through molecular evolution, a human scFv called scFv13R4 was obtained, which is expressed at high levels in *E. coli* cytoplasm. This scFv is also expressed and has a soluble and active conformation in both yeast and mammalian cells. This scFv is very stable *in vitro* and can be renatured in presence of a reducing agent. In addition, analysis of its folding kinetics showed that it folds faster than the parent scFv and aggregates more slowly *in vitro.* The mutations isolated are mainly located in the VH domain and seem to be highly specific to this particular scFv since they cannot be transferred to a very homologous VH domain. The nucleotide and amino acid sequences of scFv13R4 are shown below.

Step 2: Introduction of Diversity into CDR3 Sequences

a)Database of human CDR3 sequences

[0109] Human CDR3 sequences were compiled from three main sources: the Kabat database (Johnson, G. and Wu, T.T. Kabat Database and its applications: 30 years after the first variability plot. Nucleic Acids Res 2000, 28:214-218), the IMGT database (Giudicelli, V., et al. IMGT/LIGM-DB, the IMGT(R) comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. Nucleic Acids Res 2006, 34:D781-784), and the RCSB PDB (Berman, H.M., et al. The Protein Data Bank. Nucleic Acids Res 2000, 28:235-242). After removing the duplicates, the database contained 5179 H3, 1432 K3, and 1131 L3 CDR3 unique sequences. It can be noted that most of the H3 sequences were unique since, for instance, in the Kabat database, 2368 H3 sequences (88%) were found only once among the 2703 complete H3 sequences. The result was comparable in the case of L3 and K3 since, respectively, 87% and 82% of the sequences were unique in the Kabat database. This underlines the very high variability of the human CDR3 sequences.
[0110] This variability, however, is not evenly distributed in the loop, and the frequency of each amino acid varies from one position to another and for each loop length. In addition, the amino acid distribution depends on the origin of the antibody sequence. This bias can be due to a structural constraint as for instance in the case of the antepenultimate residue which is frequently an aspartate (D101 using Kabat numbering scheme) and plays an important role in the switch between the extended and the kinked conformation of the H3. In other cases, this bias may only be due to the limited number of sequences available for the D and J segments, and amino acids other than those found in natural antibodies

may be tolerated.

[0111] For the construction of the library it was decided to make CDR3s that mimic the natural distribution for two main reasons: i) one goal was to general scFvs that would be as human as possible for possible use in human therapy; and ii) maintaining the natural amino acid distribution will be more likely result in functional antibodies.

[0112] CDR3 sequences from the database were aligned by length, and the frequency for each of the 20 possible amino acids at each position and for each loop length were calculated. In the case of the light chain CDR3s, sequences were analyzed independently for each class. In the case of the H3 sequences, this resulted in 35 tables, one for each H3 length between 1 and 34 amino acids. For a loop of length n, this table contained 20n frequencies.

b) Oligonucleotide design for encoding CDR3 loops

[0113] Eighteen oligonucleotides were designed to follow the amino acid distribution found in the compiled CDR3 database. One-hundred-ninety-two optimized mixes of the four nucleotides at each position of the codon were used, to match as well as possible with the desired amino acid distribution. The main advantage of this approach is that it only requires classical oligonucleotide synthesis resulting in better oligonucleotide quality. Due to the restrictions of the genetic code, however, it is not possible to follow precisely an arbitrary amino acid distribution. In addition, if the library does not strictly follow the natural amino acid distribution, this may introduce interesting non-natural diversity in the CDR3 loops.

[0114] Optimized mixes for the 249 variable positions were first calculated to match the distribution with a minimal frequency of stop codons, then 34 positions which were too distant from the target distribution were further optimized by relaxing this last constraint. Due to the constraint of the genetic code, some positions were not perfectly optimized. For instance, at position 3, alanine and glycine were under-represented in our mix and it was necessary to introduce a substantial amount of some non-naturally found amino acids like cysteine in order to match other amino acid frequencies. There was, however, a good overall agreement between the database and the oligonucleotide-encoded frequencies since the most frequently found amino acids were represented at the highest rates in the library and the rare amino acids were usually present at a low frequency. The sequences of the 18 degenerate oligonucleotides used to construct the CDR3 loops are provided above.

c) Construction of CDR3 loop libraries for VH and VL chains

[0115] Independent libraries for each CDR3 loop length were constructed. This was done independently for each of the heavy and the two classes of light chains. For each library, random CDR3 loops were introduced by PCR and the resulting library was then cloned back in the scFv 13R4 gene, which was fused to the CAT gene in vector pscFvCAT. This resulted in libraries of scFv13R4 clones with one and only one randomized CDR3 loop.

[0116] The 5 amino acid long H3 loop library was constructed first. Forty-three randomly chosen clones were sequenced. Some positions diverged from the expected values of frequencies for the 20 amino acids, but on average, the distribution of amino acids in the library matched the expected distribution. This showed that the quality of the oligonucleotides was good and that the resulting library followed the natural distribution of the amino acid in human H3 loops.

Table 2

Diversity of the CDR3 libraries

		CAM phenotype[b]			
	Initial diversity[a]	++	+	-	final diversity[c]
H3-5	1.4e8	15/20	3/20	2/20	1.1e8
H3-6	2.7e7	20/20	0/20	0/20	2.7e7
H3-7	9.2e7	16/20	3/20	1/20	7.4e7
H3-8	5.0e6	14/20	4/20	2/20	3.5e6
H3-9	2.4e8	16/20	1/20	3/20	1.9e8
H3-10	1.0e7	10/20	6/20	4/20	5.0e6
H3-11	2.2e7	9/12	1/12	2/12	1.7e7
H3-12	1.0e7	15/20	2/20	3/20	7.5e6
H3-13	2.3e7	8/12	3/12	1/12	1.5e7
H3-14	1.0e7	12/20	4/20	4/20	6.0e6
H3-15	2.1e7	8/12	3/12	1/12	1.4e7
H3-16	1.0e7	12/20	2/20	6/20	6.0e6
H3-17	1.1e7	12/20	4/20	4/20	6.6e6
K3-9	3.6e6	11/16	1/16	4/16	2.5e6

(continued)

Diversity of the CDR3 libraries

	Initial diversity[a]	CAM phenotype[b]			final diversity[c]
		++	+	-	
K3-10	4.4e6	10/16	1/16	5/16	2.8e6
L3-9	1.7e7	11/20	6/20	3/20	9.4e6
L3-10	1.5e7	10/20	9/20	1/20	7.5e6
L3-11	1.8e7	15/20	3/20	2/20	1.4e7

[a] Initial diversity of the library cloned as fusion with CAT and selected on ampicillin. This is the number of clones obtained after transformation.

[b] Between 12 and 20 clones from the transformation were checked on CAM plates. Plates were incubated for 16 h at 37°C and colony size estimated. Columns ++,+ and - give, respectively, the fraction of clones that grew normally, gave tiny colonies, and did not grow at all.

[c] Diversity of the libraries selected on CAM. The diversity is estimated from column "Initial Diversity" and "CAM phenotype" by assuming that the final diversity is close to (Initial Diversity) $\times$ (CAM phenotype ++). The real diversity may be a higher since some of the clones noted + in column "CAM phenotype" may be present in low abundance.

Step 3: *Removal of Non-Expressed Sequences*

[0117] As expected, not all of the clones formed from the construction of the VH and VL libraries were functional for three main reasons: i) oligonucleotides used to introduce diversity may contain stop codons; ii) stop codons or frameshifts may be introduced by the PCR and the cloning steps; or iii) the scFvs are poorly expressed in the cytoplasm. To remove these non-functional scFv clones, expressed clones were selected by fusing the scFv gene and the chloramphenicol acethyl transferase (CAT) enzyme using the method of Maxwell KL et al. (A simple in vivo assay for increased protein solubility, Protein Sci 1999; 8:1908-1911) as described below.

Creation of scFV-CAT fusion proteins

[0118] Briefly, the scFv libraries were independently cloned between the *NcoI* and *NotI* sites of vector pscFvCAT under the control of the *tac* promoter and in frame with a downstream CAT gene. The scFv-CAT fusion protein was thus expressed in the cytoplasm. If a scFv was not properly expressed because of the inclusion of a stop codon or frameshift mutation, or if it was unable to fold in the cytoplasm, the resulting scFv-CATprotein would be either not expressed or not active, resulting in a chloramphenicol sensitive (CAM[s]) phenotype. On the other hand, if the scFv was properly expressed, the resulting scFv-CAT protein would be well expressed in the cytoplasm, and the bacterium would be chloramphenicol resistant (CAM[R]). By adjusting the chloramphenicol (CAM) concentration, one can even select for expression of different solubility levels of the scFv-CAT protein.

[0119] Different CAM concentrations were tested for this selection step ranging from 15 to 200 µg/ml. At the highest concentration of CAM, the library was enriched in well-expressed scFvs, but also in clones containing recombined plasmids harboring partial or complete deletions of the scFv gene. Next, the libraries were plated on a medium CAM concentration (30 µg/ml). This concentration was high enough to remove all the non-expressed or strongly aggregating scFvs, but did not result in a detectable amount of plasmids harboring scFv deletion.

[0120] To estimate the final size of the libraries, at least 12 clones were isolated from each initial library, which was selected on ampicillin and on CAM to determine the fraction of CAM[R] clones. Some clones grew quickly and formed colonies on ampicillin/CAM/IPTG plates (Table 2, column ++), some grew slowly (column +), and some did not grow at all (column -). The size of the libraries of expressed scFv (selected on Amp+CAM+IPTG) was thus estimated as the product of the original library size (selected on ampicillin) by the frequency of the CAM[R] clones. The sizes of the 18 libraries are given in the last column of Table 2 and ranged from 2.5×10^6 to 1.9×10^8.

Step 4: Assembly of CAM-selected library

[0121] The final library was assembled by recombining the 13 CAM-selected VH libraries with the 5 CAM-selected VL libraries. The theoretical possible diversity is about 10^{15} (~13 VH x 10^7 x 5 VL x 10^6). This is much larger than a library

that could be obtained by electroporation. It is thus very unlikely to obtain twice the same clone in the final library.

[0122] The 13 VH and 5 VL libraries were amplified by PCR with an overlapping sequence of 17 nucleotides, then purified and quantified on agarose gel. The VH and VL purified fragments were then pooled in amounts proportional to the natural distribution of the CDR3 loop lengths in human antibodies (Figure 3a). Finally, the VH and VL mixes were assembled by PCR, digested and cloned in a vector suitable for phage display. The library was electroporated in strain Cmax5αF', resulting in a library of 1.5x10^9 clones containing a scFv insert, as checked by PCR on 100 randomly picked colonies. The 18 CAM-selected libraries were assembled in amounts proportional to the natural distribution of CDR3 loops lengths in human antibodies to form a final library of more than a billion clones.

[0123] One hundred and eighteen clones were sequenced to determine loop lengths and sequences. Almost all loop lengths were found in the library. 11 and 16 amino acid long loops were also under-represented in the library. This is presumably due to the poor quality of these oligonucleotides as shown by their profile on an Agilent Bioanalizer. Loop lengths ranging from 7 to 12 were over-represented in the library but only by a two-fold factor. The loop lengths between 8 and 17 amino acids, which are the most frequently found in human antibodies, were all present in the library. The number of sequenced clones was too small to analyze the frequency of the amino acids found in the CDR3 loops. Except for some contamination with the original scFv 13R4 sequence, no CDR3 sequence was found twice in the library.

Expression of scFvs in the cell

[0124] Because of the novel use of the CAM selection step, the VH and VL libraries were independently optimized for expression in the cell. Because of this optimization of the VH and VL libraries the result should be only expressed scFv proteins. Furthermore, since only the CDR3 loops are modified between the original scFv13R4 antibody framework and the resulting scFv libraries, most of the interface residues between the two domains are conserved between clones. It is thus likely that any VH will assemble correctly with any VL and that the expression level of the resulting scFv will be close to that of both clones from the VH and VL libraries selected on CAM. In other words, if a VH, with a modified H3 loop, is well expressed when fused to the VL13R4, it will be also well expressed when coupled with a VL with a modified CDR3 loop and selected as a fusion with the VH1 3R4. This hypothesis was tested by picking random clones from the final library and expressing them in *E. coli* cytoplasm and in mammalian cytosol.

[0125] DNA was prepared from the final library and the scFv genes cloned in a plasmid for cytoplasmic expression under the control of the strong T7 promoter. It must be noted that the very high expression levels obtained under such a strong promoter favor aggregation over soluble expression because of the high kinetics order of the aggregation process. The stringency of this test is thus high and it could be possible to increase the soluble versus insoluble ratio by using a weaker promoter. Twenty clones were tested in *E. coli* and 19 of them showed at least some soluble expression in the cytoplasm (Figure 3). One-fourth of the clones (5/20; clones 3, 10, 11, 16, 19) were expressed at very high levels since the scFvs were clearly visible on a Coomassie stained gel. To obtain a more global view of the soluble expression levels in *E. coli,* the library was cloned in front of the GFPuv gene under the control of the T7 promoter. If the scFv is soluble and expressed in the cytoplasm, this should result in green fluorescent protein (GFP) activity that can be directly monitored on an UV transilluminator. About 1000 clones were tested for the presence of detectable GFP activity and approximately 60% exhibited a GFP+ phenotype, which again indicated that most of the scFv clones from the final library were correctly expressed in *E. coli* cytoplasm. These two tests demonstrated that the novel method of constructing an scFv library as described above was very successful in generating cytoplasmically expressed scFvs in *E. coli.*

[0126] Next, the expression of the library in mammalian cells was tested. Fifteen scFvs were cloned in a mammalian expression vector as fusions with the EGFP gene and under the control of the SV40 early promoter, then transfected in HeLa cells. Typical results are shown in Figure 4. Three clones were expressed at a high soluble level, comparable to that of the parental scFv13R4 (clone 15), 10 scFvs were found to be mainly soluble but some aggregated material was still present in the cell (clones 33 and 36), and 2 clones accumulated essentially as cytoplasmic aggregates (clone 24), as observed with the hybridoma-derived anti-oncoprotein E6 scFv 1F4 (Figure 4). In conclusion, thirteen out of the fifteen scFvs tested were expressed as soluble proteins that could be easily detected in the cytoplasm and in the nucleus of the transfected cells.

[0127] Together, these results showed that more than 85% of the clones from the final library expressed soluble scFv in *E. coli* (16/20) and mammalian cytoplasm (13/15), while about 20% of them expressed scFv at a very high level (5/20 in *E. coli* and 3/15 in eukaryotic cells). Overall, most of the clones were well expressed under the reducing conditions of the bacterial and eukaryotic cytoplasm. This is a great improvement over results previously obtained with non-optimized scFv libraries.

Selection of binders

[0128] As shown above, the library contains a very high proportion of expressed clones. The next step was selecting antibodies from the library against particular proteins. Thus, the phage display library was used to select for binders

against five different antigens using purified proteins adsorbed on microtiter plates. Three rounds of selection were performed, and the eluted phages were tested by ELISA against the immobilized antigens. In all the cases a positive signal was obtained after a single round of selection. This signal increased strongly after two rounds and did not increase further during the third round of selection. This very fast selection process was presumably due both to the focused library itself, which contains only expressed scFvs resulting in a low background, and to the use of a trypsin sensitive helper phage that further decreased the background level.

[0129] To characterize the selection process, 60 clones were selected from the three selection rounds against GST:Syk fusion. These clones were used to prepare monoclonal phages, which were then tested for binding to the antigen by ELISA. Figure 7 shows the distribution of the ELISA values obtained for each selection round. The distribution was normal with a strong homogeneity of the signal in each round of selection since more than half of the clones showed an ELISA signal within 0.1 of the peak value. During the selection process the peak signal increased from 0.4-0.5 after a single round to 0.9 after round 2 and 1.0 after round 3, in good correlation with the results obtained with the polyclonal phages. Moreover, after a single round of selection nearly 100% of the clones already recognized the antigen. This showed that using the optimized library in combination with a trypsin-sensitive helper phage results in almost a total absence of background during the selection process.

[0130] Next tested was whether the library contained clones expressing soluble scFvs in the periplasm. The non-suppressive HB2151 strain was infected with the phages eluted after the third round of selection against tubulin, GST:Syk and the core histones. Periplasmic extracts were prepared and tested for binding activity by ELISA. In the three cases, 12-20% of the clones gave a strong signal with absorbance values higher than 0.5 (10 times the background), and about 30% were clearly positive with an absorbance value higher than 0.1 (twice the background). These results compared favorably with those reported with other scfv libraries, underlining again the high proportion of well-expressed clones present in the library. In addition, this showed that the CAM-selection approach selected efficiently for constructs without stop codons present in the oligonucleotides. This is indeed of premium importance to isolate soluble scFvs from phage-displayed libraries since amber stop codons in CDRs are frequently selected during panning of synthetic and semi-synthetic libraries.

[0131] In both the previous characterizations, the scFvs were expressed under oxidizing conditions in *E. coli* periplasm, either as scFv-pIII fusion or as soluble protein. In addition, panning was done on phage, again under oxidizing conditions. To test whether the scFvs were indeed also expressed in the cytoplasm, the same pool of clones (Round 3) was subcloned in a cytoplasmic expression vector under the control of the strong T7 promoter. For each antigen, ninety-five clones were tested by ELISA for binding to their respective antigen. In each case, the number of positive clones was comparable or even better than in the periplasmic screen. For instance, in the case of GST: Syk, 80% of the tested clones were positive after three rounds of selection. This demonstrated that the periplasmic selection step did not decrease the proportion of soluble scFvs in the cytoplasm. Furthermore, when using the optimized library as described above it is not necessary to directly select within the cytoplasm to avoid introducing a bias during the selection.

[0132] Individual clones from the 2nd and the 3rd round of selection against tubulin were sequenced. Sequences are shown in Table 3.

Table 3

Sequences of some anti-tubulin scfvs

Name	VH CDR3 Sequence	length	VL CDR3 Sequence	length	frequency[a]	yield[b]	WB[c]	IF[d]
Round 2								
C12C	SSITIFGGGMDV (SEQ ID NO: 34)	12	HSREVHRTF (SEQ ID NO: 35)	9	1/5	19		
E12C	SGGNTFDY (SEQ ID NO: 36)	8	QQYYRKPWT (SEQ ID NO: 37)	9	1/5	53		
F1C	GNADGGENWELFDK (SEQ ID NO: 38)	14	QLYQNTLWT (SEQ ID NO: 39)	9	2/5	52		
H6C	SSITIFGGGMDV (SEQ ID NO: 40)	12	QQNWTSPLS (SEQ ID NO: 41)	9	1/5	nd		
Round 3								
2C1C	RGRDY (SEQ ID NO: 42)	5	QQYNTSPFS (SEQ ID NO: 43)	9	1/6	8.6	+	-
2E11C	GRNVLNY (SEQ ID NO: 44)	7	QQNSSSPRFT (SEQ ID NO: 45)	10	2/6	8.7	+	-

(continued)

Round 3								
2F12C	GRRALGN (SEQ ID NO: 46)	7	QQYNTSPFS (SEQ ID NO: 47)	9	1/6	45	+	+
2G4C	GRRALGN (SEQ ID NO: 48)	7	LTWSMRSAI (SEQ ID NO: 49)	9	1/6	15	+	+
2G9C	GRRALGN (SEQ ID NO: 50)	7	LTTENSVYRLV (SEQ ID NO: 51)	11	1/6	50	+	-

Sequences of the CDR3s of the best positive clones in an ELISA using cytoplasmically expressed scFv from the **2nd** (5 clones) and the **3rd** (6 clones) round of selection (Table 3). [a] Frequency of apparition of the scFv among sequenced clones of the same round. [b]Yield: mg of scFv purified from 1 liter of cells grown in a flask (OD_{600} = 5). [c]WB: detection of tubulin in brain extracts by Western blot. [d] IF: + means that the scFv is able to reveal microtubule network by Immunofluorescence (Fig. 8). The sequences of the clones 2C1C, 2E11C, 2F12C, 2G4C and 2G9C have been submitted to the EMBL database and their accession numbers are respectively AM886280, AM886281, AM886282, AM886283 and AM886284.

[0133]   In all cases, the clones sequenced were those giving the best signal in the ELISA performed with cytoplasmically expressed scFvs. Most of the clones were different since only one clone from the 2nd round and one from the 3rd round were found twice. This demonstrated that a high diversity is still present after 3 rounds of selection. Eight of the anti-tubulin scFvs were purified by affinity chromatography from the cytoplasm. In all cases, more than 8 mg of scFv was purified from one liter of cells grown in a flask (OD_{600} = 5), and some scFvs were expressed at a level per cell comparable to the exceptionally high expression level reported for an anti-HER2 in *E. coli* periplasm.

Functionality of scFvs as intrabodies

[0134]   To determine if the isolated scFvs were able to bind to their target *in vivo,* the anti-histone scFvs expressed in human cells was characterized. The third round of selection was cloned in vector p513-EGFP and ten randomly chosen clones were transfected in HeLa cells. Typical results of the cells expressing the scFv-EGFP fusions and observed by fluorescence microscopy are shown in Figure 6. One scFv was expressed as cytoplasmic aggregates. Four scFvs were expressed as soluble cytoplasmic proteins, as judged by the homogeneous staining of the cells, at a level comparable to that of the scFv13R4. Finally, the expression of three scFvs gave rise to a stronger staining of the nucleus (Figure 6, clone 2) and two scFvs were exclusively localized in the nucleus (Figure 6, clones 5 and 10). Since these scFvs-EGFP fusion proteins were expressed in the cytoplasm of the cell and did not contain a nuclear localization signal, this suggested that they were able to interact *in vivo* with the histones and were thus active inside the cell. This analysis showed that about half of the clones present after the third round of selection against core histones were able to bind to their nuclear target *in vivo.* This was confirmed in vitro by western and dot blot with purified scFv. In addition, sequencing of these clones showed that they contained different heavy and light chain CDR3 regions.

In vitro characteriza*tion of anti-tubulin scFvs

[0135]   To demonstrate the activity of the anti-tubulin scFv under the reducing conditions in the cell cytoplasm, the scFvs were extracted in the presence of a reducing agent and compared the ELISA signal with that obtained with the scFvs extracted under oxidizing conditions. As shown in Figure 7, the five scFvs tested gave the same ELISA signal under both conditions, demonstrating that the scFVs retain full activity under reducing conditions even in the absence of disulfide bond formation. The five scFvs were able to recognize unfolded tubulin by western blot in brain extracts and the native protein in a competition ELISA. The ability of the five scFvs to interact with microtubules in cells was tested by IF. Clones 2F12C and 2G4C revealed the microtubule network in cells.

[0136]   Figure 8 illustrates the utility of the library as a source for both in vitro and in vivo proteomic studies: HeLa cells were transfected with the anti-histones clone 5 fused to a Red-GFP, and he microtubule network was revealed by IF using the 2F12C scFv.

[0137]   Altogether our results show that the library described in this report is highly diverse and functional and allows fast and easy isolation *of in vivo* active fully human intrabodies.

Claims

1. An antibody library comprising at least about 10^6 unique scFv antibody clones, wherein each unique scFv antibody clone encodes a unique scFv antibody comprising at least one of a unique CDR3 VH sequence and a unique CDR3 VL sequence, and wherein the unique scFv antibody clones encode a framework sequence identical to a framework sequence encoded by an scFv13R4 antibody clone of SEQ ID NO: 32.

2. The antibody library of claim 1, wherein the unique scFv antibody clones encode scFv antibodies comprising a unique CDR3 VH sequence.

3. The antibody library of claim 1, wherein the unique scFv antibody clones encode scFv antibodies comprising a unique CDR3 VL sequence.

4. The antibody library of claim 1, wherein the unique scFv antibody clones encode scFv antibodies comprising a unique CDR3 VH sequence and a unique CDR3 VL sequence.

5. A method for preparing an scFv antibody library enriched for scFv antibody clones that can be expressed within a cell, comprising:

   a) providing a first collection of scFv antibody clones, wherein the first collection comprises clones comprising a unique sequence within a CDR3 loop of VH, wherein the first collection has been enriched for scFv antibody clones that contain a VL sequence identical to the VL sequence of an scFv13R4 antibody clone of SEQ ID IVO: 33;
   b) providing a second collection of scFv antibody clones, wherein the second collection comprises clones comprising a unique sequence within a CDR3 loop of VL, wherein the second collection has been enriched for scFv antibody clones that contain a VH sequence identical to the VH sequence of an scFv13R4 antibody clone of SEQ ID NO: 33;
   c) joining VH domains from scFv antibody clones of the first collection with VL domains from scFv antibody clones of the second collection to obtain a third collection of scFv antibody clones, wherein the third collection contains scFv antibody clones comprising a unique sequence within the CDR3 loop of VH and a unique sequence within the CDR3 loop of VL,

   thereby preparing the scFv antibody library enriched for scFv antibody clones that can be expressed within a cell.

6. The method of claim 5, wherein the first collection comprises scFv antibody clones that comprise identical CDR1 and CDR2 sequences in the VH domain and wherein the second collection comprises scFv antibody clones that comprise identical CDR1 and CDR2 sequences in the VL domain.

7. An antibody library produced by the method of claim 5.

8. A method for constructing an antibody library comprising:

   a) selecting the framework of an scFv13R4 antibody clone of SEQ ID NO: 33 as an scFv antibody framework;
   b) introducing sequence diversity into a VH CDR3 region of the scFv antibody framework to generate a first library comprising scFv antibody clones comprising a unique VH CDR3 region;
   c) introducing sequence diversity into a VL CDR3 region of the scFv antibody framework to generate a second library comprising scFv antibody clones comprising a unique VL CDR3 region;
   d) removing, from the first library, clones that do not detectably express scFv antibody;
   e) removing, from the second library, clones that do not detectably express scFv antibody; and
   f) recombining the first and second libraries to generate a final library comprising scFv antibody clones comprising a unique VH CDR3 region and a unique VL CDR3 region;

   thereby constructing the antibody library.

Patentansprüche

1. Antikörperbank, umfassend mindestens etwa 10^6 nur einmal vorkommende scFv-Antikörperklone, wobei jeder nur einmal vorkommende scFv-Antikörperklon einen nur einmal vorkommenden scFv-Antikörper kodiert, der mindestens

eine aus einer nur einmal vorkommenden CDR3-VH-Sequenz und einer nur einmal vorkommenden CDR3-VL-Sequenz umfasst, und wobei die nur einmal vorkommenden scFv-Antikörperklone eine Rahmensequenz kodieren, die mit einer durch einen scFv13R4-Antikörperklon der SEQ ID NO: 32 kodierten Rahmensequenz identisch ist.

2. Antikörperbank nach Anspruch 1, wobei die nur einmal vorkommenden scFv-Antikörperklone scFv-Antikörper kodieren, die eine nur einmal vorkommende CDR3-VH-Sequenz umfassen.

3. Antikörperbank nach Anspruch 1, wobei die nur einmal vorkommenden scFv-Antikörperklone scFv-Antikörper kodieren, die eine nur einmal vorkommende CDR3-VL-Sequenz umfassen.

4. Antikörperbank nach Anspruch 1, wobei die nur einmal vorkommenden scFv-Antikörperklone scFv-Antikörper kodieren, die eine nur einmal vorkommende CDR3-VH-Sequenz und eine nur einmal vorkommende CDR3-VL-Sequenz umfassen.

5. Verfahren zum Herstellen einer für scFv-Antikörperklone, die in einer Zelle exprimiert werden können, angereicherten scFv-Antikörperbank, umfassend:

   (a) Bereitstellen einer ersten Sammlung von scFv-Antikörperklonen, wobei die erste Sammlung Klone umfasst, die eine nur einmal vorkommende Sequenz in einem CDR3-Loop von VH umfassen, wobei die erste Sammlung für scFv-Antikörperklone, die eine mit der VL-Sequenz eines scFv13R4-Antikörperklons der SEQ ID NO: 33 identische VL-Sequenz enthalten, angereichert worden ist;
   (b) Bereitstellen einer zweiten Sammlung von scFv-Antikörperklonen, wobei die zweite Sammlung Klone umfasst, die eine nur einmal vorkommende Sequenz in einem CDR3-Loop von VL umfassen, wobei die zweite Sammlung für scFv-Antikörperklone, die eine mit der VH-Sequenz eines scFv13R4-Antikörperklons der SEQ ID NO: 33 identische VH-Sequenz enthalten, angereichert worden ist;
   (c) Verknüpfen von VH-Domänen von scFv-Antikörperklonen der ersten Sammlung mit VL-Domänen von scFv-Antikörperklonen der zweiten Sammlung, um eine dritte Sammlung von scFv-Antikörperklonen zu erhalten, wobei die dritte Sammlung scFv-Antikörperklone enthält, die eine nur einmal vorkommende Sequenz im CDR3-Loop von VH und eine nur einmal vorkommende Sequenz im CDR3-Loop von VL umfassen,

   wodurch die für scFv-Antikörperklone, die in einer Zelle exprimiert werden können, angereicherte scFv-Antikörperbank hergestellt wird.

6. Verfahren nach Anspruch 5, wobei die erste Sammlung scFv-Antikörperklone umfasst, die identische CDR1- und CDR2-Sequenzen in der VH-Domäne umfassen, und wobei die zweite Sammlung scFv-Antikörperklone umfasst, die identische CDR1- und CDR2-Sequenzen in der VL-Domäne umfassen.

7. Antikörperbank, hergestellt durch das Verfahren nach Anspruch 5.

8. Verfahren zum Konstruieren einer Antikörperbank, umfassend:

   (a) Auswählen des Rahmens eines scFv13R4-Antikörperklons der SEQ ID NO: 33 als scFv-Antikörper-Rahmen;
   (b) Einführen von Sequenzdiversität in eine VH-CDR3-Region des scFv-Antikörper-Rahmens, um eine erste Bank zu erzeugen, die scFv-Antikörperklone, umfassend eine nur einmal vorkommende VH-CDR3-Region, umfasst;
   (c) Einführen von Sequenzdiversität in eine VL-CDR3-Region des scFv-Antikörper-Rahmens, um eine zweite Bank zu erzeugen, die scFv-Antikörperklone, umfassend eine nur einmal vorkommende VL-CDR3-Region, umfasst;
   (d) Entfernen von Klonen aus der ersten Bank, die einen scFv-Antikörper nicht nachweisbar exprimieren;
   (e) Entfernen von Klonen aus der zweiten Bank, die einen scFv-Antikörper nicht nachweisbar exprimieren, und
   (f) Rekombinieren der ersten und zweiten Bank, um eine endgültige Bank zu erzeugen, die scFv-Antikörperklone, umfassend eine nur einmal vorkommende VH-CDR3-Region und eine nur einmal vorkommende VL-CDR3-Region, umfasst;

wodurch die Antikörperbank konstruiert wird.

Revendications

1. Bibliothèque d'anticorps comprenant au moins environ 10^6 clones d'anticorps scFv unique, chaque clone d'anticorps scFv unique codant un anticorps scFv unique comprenant au moins une séquence parmi une séquence VH CDR3 unique et une séquence VL CDR3.unique, et dans lequel les clones d'anticorps scFv unique codent une séquence d'ossature identique à une séquence d'ossature codée par un clone d'anticorps scFv13R4 de SEQ ID NO : 32.

2. Bibliothèque d'anticorps selon la revendication 1, dans laquelle les clones d'anticorps scFv unique comprennent une séquence VH CDR3 unique.

3. Bibliothèque d'anticorps selon la revendication 1, dans laquelle les clones d'anticorps scFv unique comprennent une séquence VL CDR3 unique.

4. Bibliothèque d' anticorps selon la revendication 1, dans laquelle les clones d'anticorps scFv unique codent des anticorps scFv comprenant une séquence VH CDR3 unique et une séquence VL CDR3 unique.

5. Procédé de préparation d'une bibliothèque d'anticorps scFv enrichis en clones d'anticorps scFv qui peuvent exprimer dans une cellule, consistant à :

   a) fournir une première collection de clones d'anticorps scFv, la première collection comprenant des clones comprenant une séquence unique dans une boucle de VH de CDR3, la première collection ayant été enrichie en clones d'anticorps scFv qui contiennent une séquence VL identique à la séquence V1 d'un clone d' anticorps scFv13R4 de SEQ ID NO : 33 ;
   b) fournir une deuxième collection de clones d'anticorps scFv, la seconde collection comprenant des clones comprenant une séquence unique dans une boucle de VL de CDR3, la deuxième collection ayant été enrichie en clones d'anticorps scFv qui contiennent une séquence VH identique à la séquence VH d'un clone d'anticorps scFv13R4 de SEQ ID NO : 33 ;
   c) relier les domaines VH. provenant des clones d'anticorps scFv de la première collection à des domaines VL provenant de clones d'anticorps scFv de la seconde collection pour obtenir une troisième collection de clones d'anticorps scFv, la troisième collection contenant des clones anticorps scFv comprenant une séquence unique dans la boucle de VH de CDR3 et une séquence unique dans la boucle de VL de CDR3, en préparant ainsi la bibliothèque d'anticorps scFv enrichie en clones d'anticorps scFv qui peuvent exprimer dans une cellule.

6. Procédé selon la revendication 5, dans lequel la première collection comprend des clones d'anticorps scFv qui comprennent des séquences CDR1 et CDR2 identiques dans le domaine VH, et dans lequel la deuxième collection comprend des clones d' anticorps scFv qui comprennent des séquences CDR1 et CDR2 identiques dans le domaine VL.

7. Bibliothèque d'anticorps produite par le procédé selon la revendication 5.

8. Procédé de construction d'une bibliothèque d'anticorps, consistant à :

   a) sélectionner l'ossature d'un clone d'anticorps scFv13R4 de SEQ ID NO : 33 en tant que'ossature d'anticorps scFv;
   b) introduire une diversité de séquences dans une région VH CDR3 de l'ossature d'anticorps scFv pour produire une première bibliothèque comprenant des clones d'anticorps scFv comprenant une région VH CDR3 unique ;
   c) introduire une diversité de séquences dans une région VL CDR3 de l'ossature d'anticorps scFv pour produire une deuxième bibliothèque comprenant des clones d'anticorps scFv comprenant une région VL CDR3 unique;
   d) enlever, à partir de la première bibliothèque, les clones qui n'ont pas exprimé de manière détectable un anticorps scFv ;
   e) enlever, de la deuxième bibliothèque, des clones qui n'ont pas exprimé de manière détectable un anticorps scFv; et
   f) recombiner les premières et deuxième bibliothèques pour produire une bibliothèque finale comprenant des clones d'anticorps scFv comprenant une région VH CDR3 unique et une région VL CDR3 unique,

   construisant ainsi la bibliothèque d'anticorps.
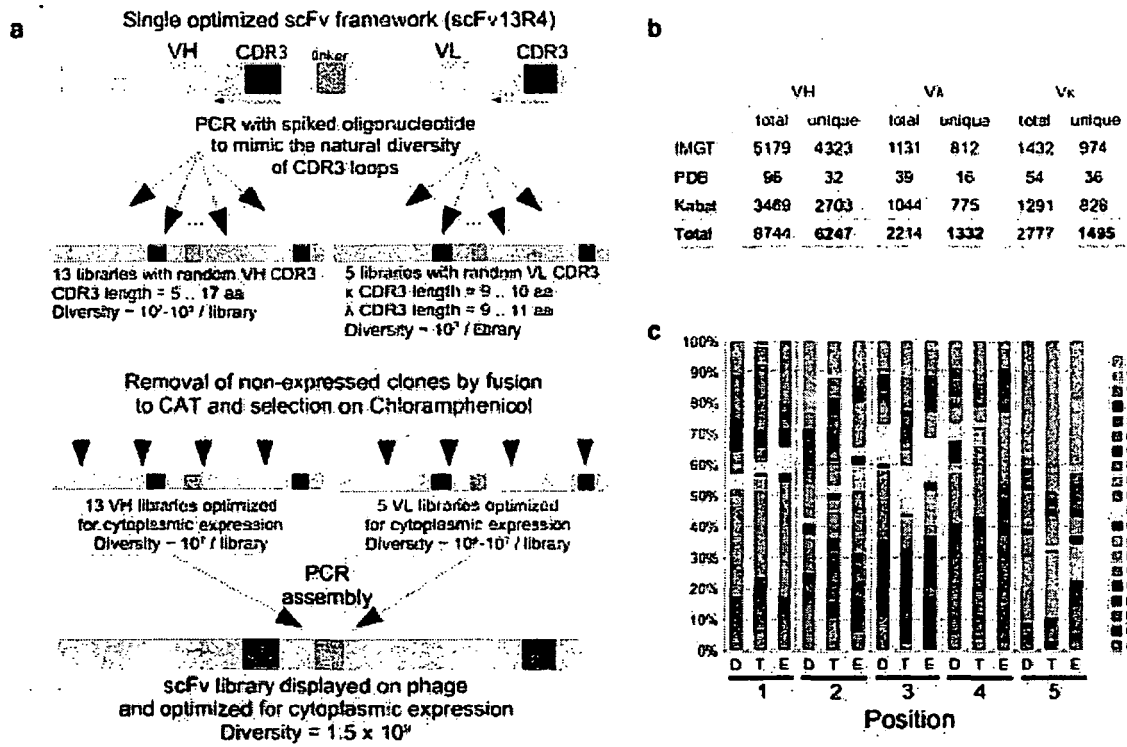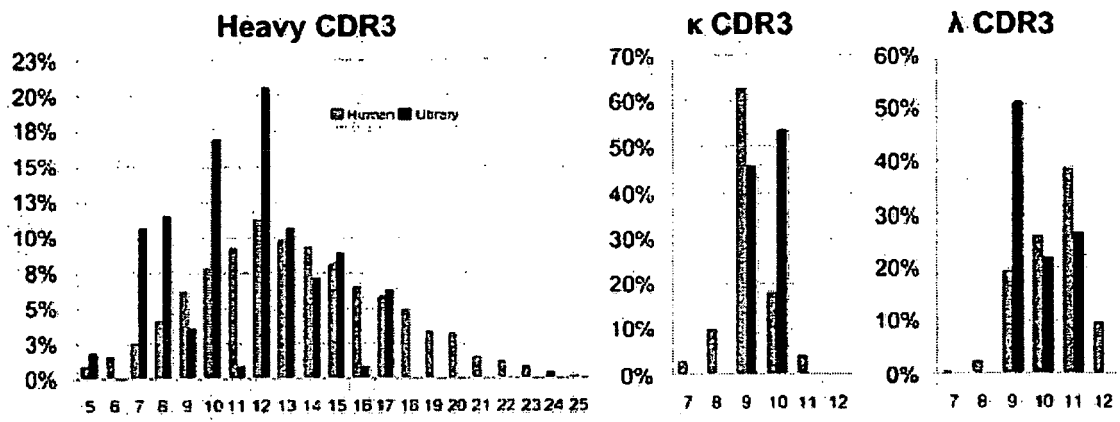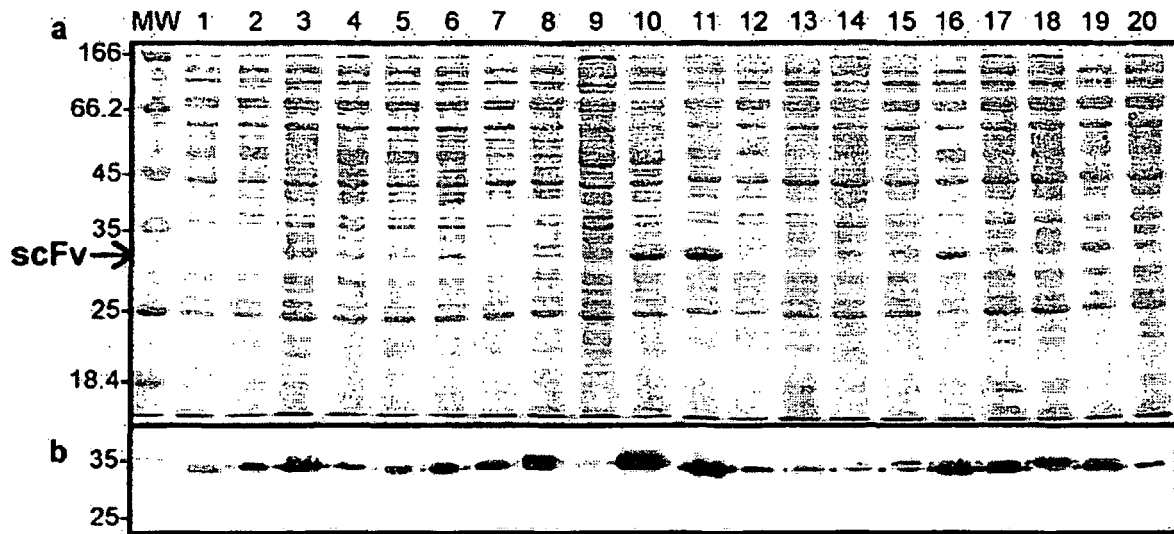
Figure 1

Heavy CDR3

κ CDR3

λ CDR3

Figure 2

Figure 3

Figure 4

a

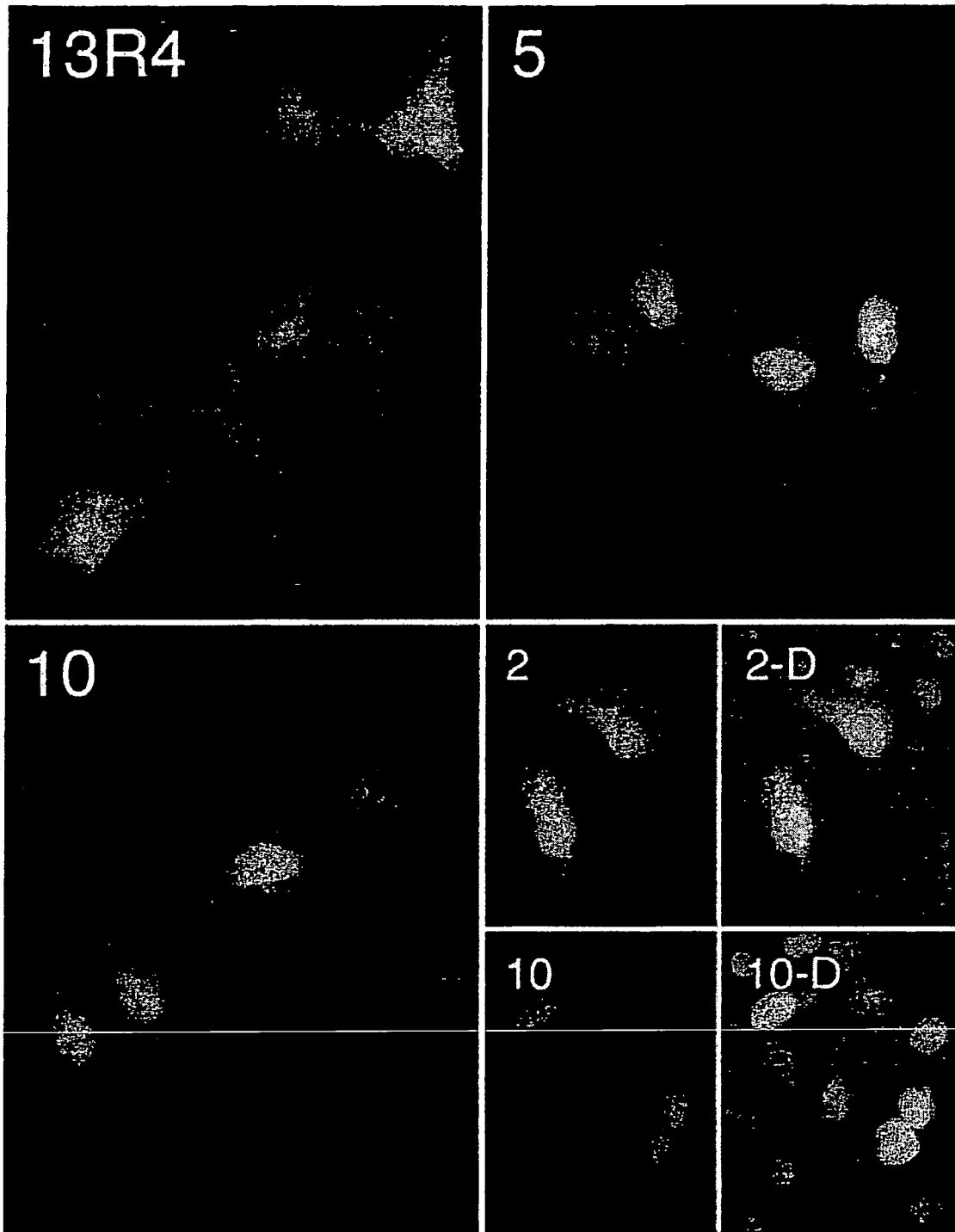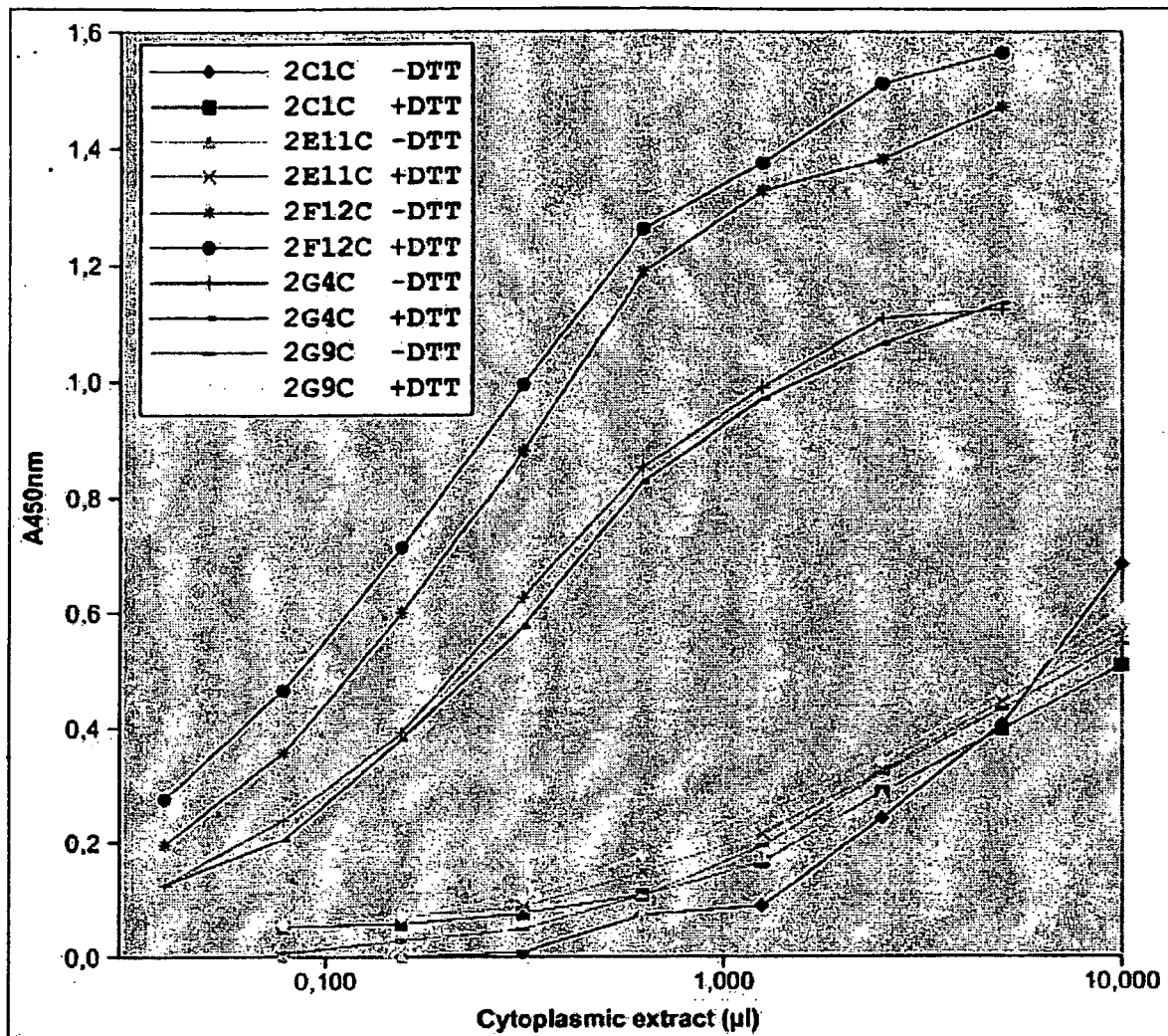**b**
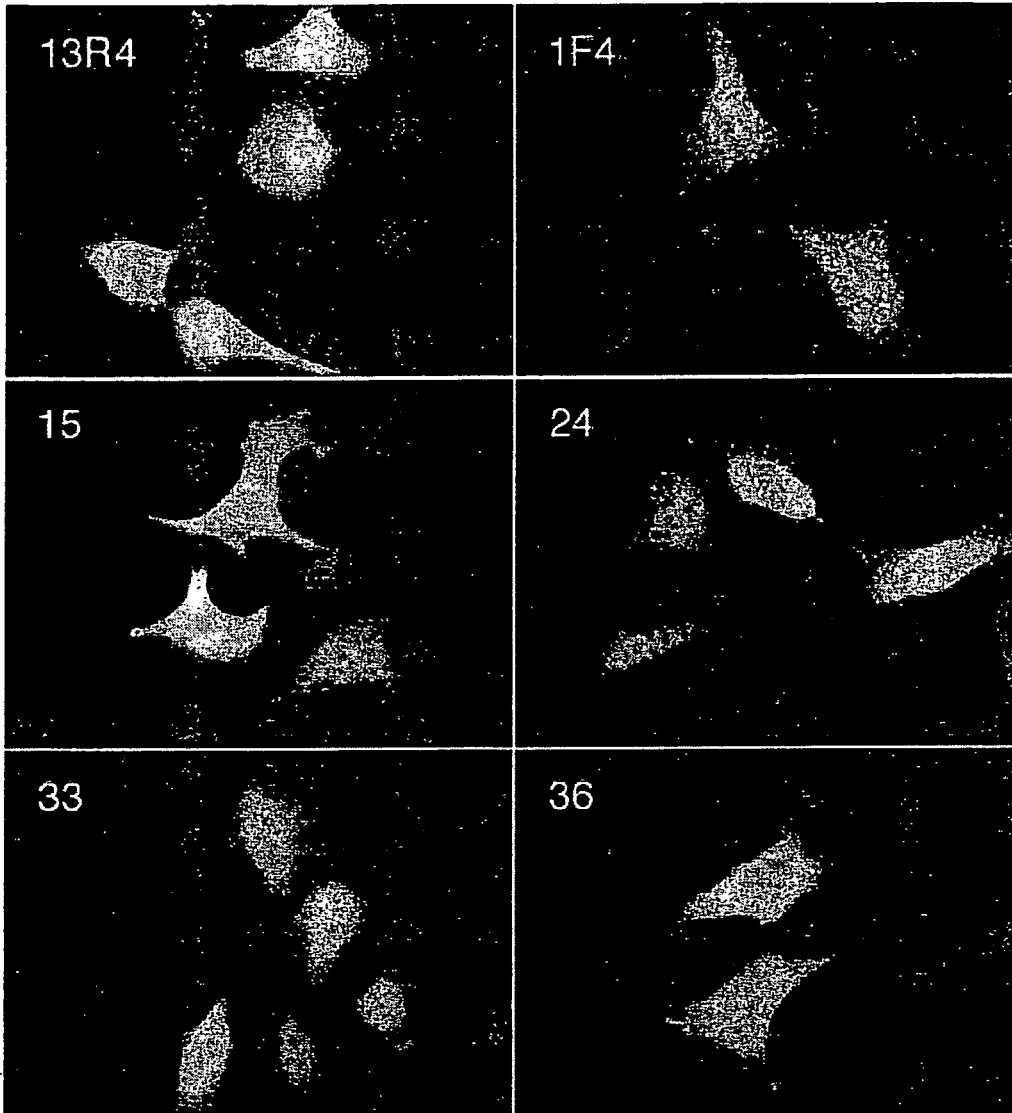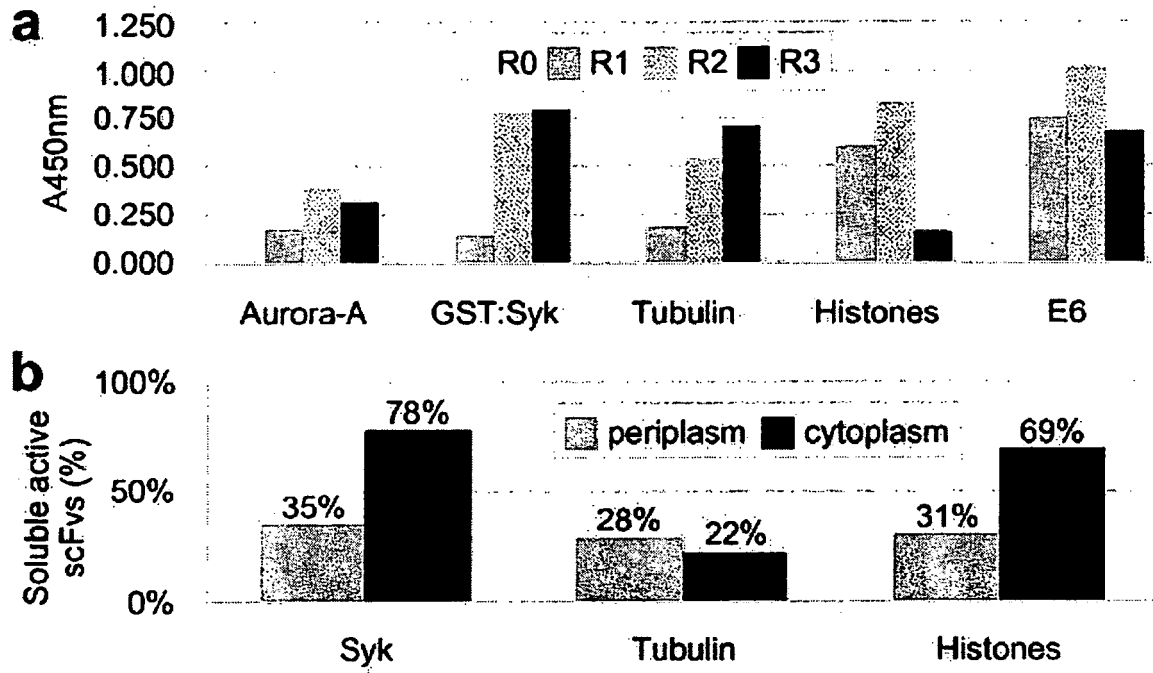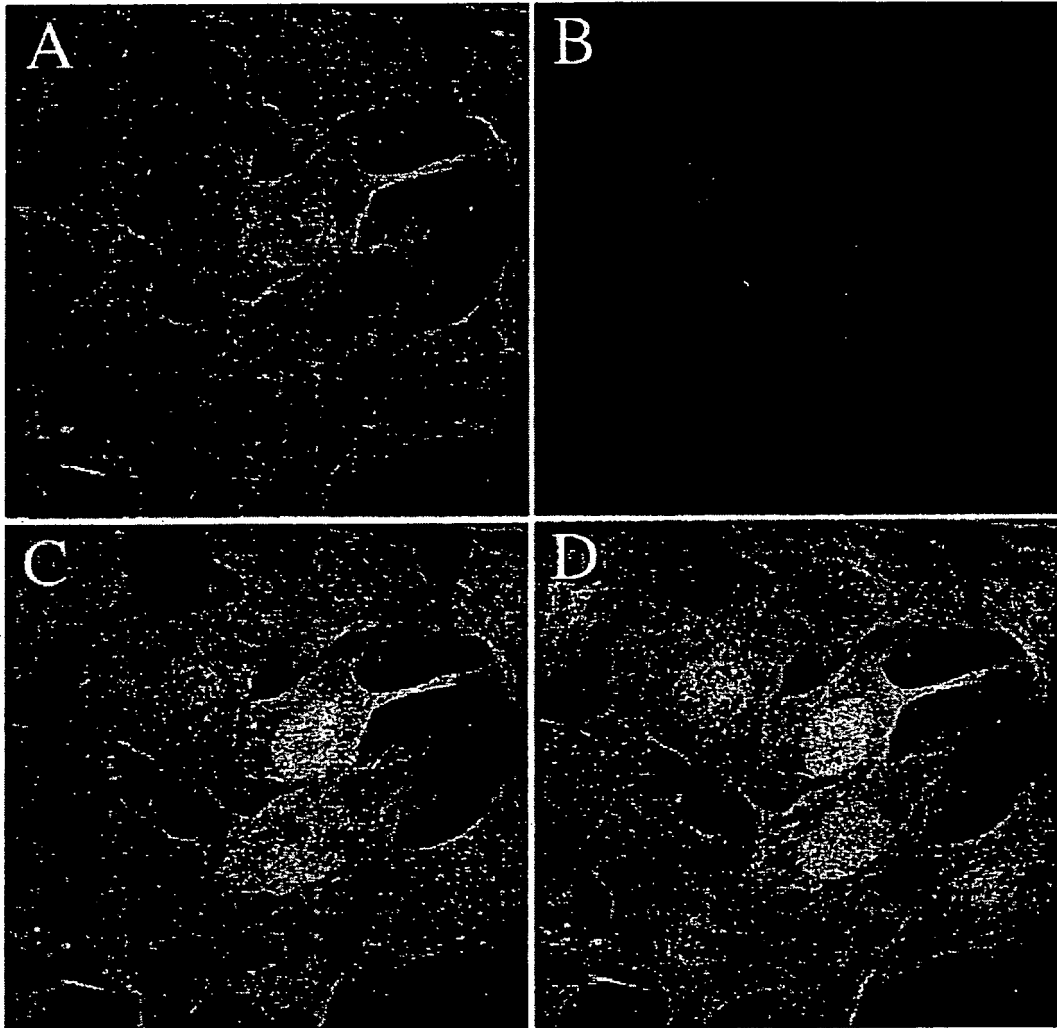

Figure 5

Figure 6

Figure 7

Figure 8

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 0054057 A **[0008]**
- WO 0148017 A **[0009]**
- WO 9945959 A **[0014]**
- WO 2005106481 A **[0060]**

Non-patent literature cited in the description

- **DER MAUR ADRIAN et al.** *Methods,* 2004, vol. 34, 215-224 **[0010]**
- **DER MAUR ADRIAN et al.** *Journal of Biological Chemistry,* 2002, vol. 277 (47), 45075-45085 **[0010]**
- **MAUR ADRIAN et al.** *Methods,* 2004, vol. 34, 215-224 **[0011]**
- **LADEN et al.** *Research in Microbiology,* 2002, vol. 153, 469-474 **[0012]**
- **JUNG.** *protein Engineering,* 1997, vol. 10 (8), 959-966 **[0013]**
- **ULANOVA et al.** *Expert Opin Ther Targets,* 2005, vol. 9, 901-921 **[0060]**
- **SEBOLT-LEOPOLD JS ; HERRERA, R.** *Nat Rev Cancer,* 2004, vol. 4, 937-947 **[0060]**
- **JORDAN MA ; WILSON, L.** *Nat Rev Cancer,* 2004, vol. 4, 253-265 **[0060]**
- **VILLANI, ME et al.** Immunomodulation of cucumber mosaic virus infection by antibodies selected in vitro from a stable single-framework phage display library. *Plant Molecular Biology,* 2005, vol. 58 (3), 305 **[0061]**
- **SWAN, CH et al.** T-cell protection and enrichment through lentiviral CCR5 gene delivery. *Gene Ther,* vol. 13, 1480-1492 **[0062]**
- **MILLER, TW et al.** A human single-chain Fv intrabody preferentially targets amino-terminal Huntingtin's fragments in striatal models of Huntington's disease. *Neurobiol dis.,* 2005, vol. 19, 47-56 **[0063]**
- **MILLER ; MESSER.** Intrabody applications in neurological disorders: progress and future prospects. *Mol. Ther.,* 2005, vol. 12, 394-401 **[0063]**
- **WILLIAMS, BR ; ZHU, Z.** Intrabody-based approaches to cancer therapy: status and prospects. *Current med Chem,* 2006, vol. 13, 1473-1480 **[0064]**
- **DOORBAR J. ; GRIFFIN H.** Intrabody strategies for the treatment of human papillomavirus-associated disease. *Expert Opin. Biol. Ther.,* 2007, vol. 7 (5), 677-689 **[0064]**
- **FANG CY et al.** Modulation of Epstein0-Barr Virus Latent Membrane Protein 1 Activity by Intrabodies. *Intervirology,* 2007, vol. 50, 254-263 **[0065]**
- **PASTAN et al.** *Proc. Natl. Acad. Sci. U.S.A.,* 1988, vol. 85, 4486 **[0067]**
- **MILLER et al.** *Mol. Cell. Biol.,* 1986, vol. 6, 2895 **[0067]**
- **MITANI et al.** *Hum. Gene Ther.,* 1994, vol. 5, 941-948 **[0067]**
- **GOODMAN et al.** *Blood,* 1994, vol. 84, 1492-1500 **[0067]**
- **NAIDINI et al.** *Science,* 1996, vol. 272, 263-267 **[0067]**
- **AGRAWAL et al.** *Exper. Hematol.,* 1996, vol. 24, 738-747 **[0067]**
- **SCHWARTZENBERGER et al.** *Blood,* 1996, vol. 87, 472-478 **[0067]**
- **WOLFF, J. A. et al.** *Science,* vol. 247, 1465-1468 **[0067]**
- **WOLFF, J. A.** *Nature,* 1991, vol. 352, 815-818 **[0067]**
- *Med Hypotheses,* 2005, vol. 64, 1105-1108 **[0068]**
- **COURTÊTE J. ; SIBLER A.P. ; ZEDER-LUTZ G. ; DALKARA D. ; ZUBER G. ; WEISS E.** Suppression of cervical carcinoma cell growth by intracytoplasmic co-deiivery of anti-oncoprotein E6 antibody and siRNA. *Mol. Cancer Ther.,* 2007, vol. 6, 1728-36 **[0068]**
- Current Protocols in Molecular Biology. John Wiley and Sons, 2006 **[0071]**
- **MILLER, J.H.** A Short Course in Bacterial Genetics: a Laboratory Manual and Handbook for Escherichia coli and Related Bacteria. Cold Spring Harbor Laboratory Press, 1992 **[0083]**
- **MCCAFFERTY J et al.** *Appl Biochem Biotechnol,* 1994, vol. 47, 157-171 **[0088]**
- **GREEN, S. et al.** *Nucleic Acids Res,* 1988, vol. 16, 369 **[0091]**
- **JOHNSON, G. ; WU, T.T.** Kabat Database and its applications: 30 years after the first variability plot. *Nucleic Acids Res,* 2000, vol. 28, 214-218 **[0094] [0109]**
- **GIUDICELLI, V. et al.** IMGT/LIGM-DB, the IMGT(R) comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Res,* 2006, vol. 34, D781-784 **[0095] [0109]**
- **BERMAN, H.M. et al.** The Protein Data Bank. *Nucleic Acids Res,* 2000, vol. 28, 235-242 **[0096] [0109]**

- **ALLCORN, L.C. ; MARTIN, A.C.** R. SACS-self-maintaining database of antibody crystal structure information. *Bioinformatics,* 2002, vol. 18, 175-181 **[0096]**
- **WANG, W. ; SAVEN, J.G.** Designing gene libraries from protein profiles for combinatorial protein experiments. *Nucleic Acids Res,* 2002, vol. 30, 120 **[0097]**
- **PARK, S. et al.** Progress in the development and application of computational methods for probabilistic protein design. *Comput Chem Eng,* 2005, vol. 29, 407-421 **[0097]**
- **MARTINEAU, P. ; BETTON, J.M.** In vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm. *J Mol Biol,* 1999, vol. 292, 921-929 **[0098]**
- **SIDHU, S.S. et al.** Phage display for selection of novel binding peptides. *Methods Enzymol,* 2000, vol. 328, 333-363 **[0099] [0101]**
- **DAUVILLIER, S. et al.** Intracellular single-chain variable fragments directed to the Src homology 2 domains of Syk partially inhibit Fc epsilon RI signaling in the RBL-2H3 cell line. *J Immunol,* 2002, vol. 169, 2274-2283 **[0102]**
- **WILLIAMS, R.C.J. ; LEE, J.C.** Preparation of tubulin from brain. *Methods Enzymol,* 1982, vol. 85, 376-385 **[0102]**
- **KRISTENSEN, P. ; WINTER, G.** Proteolytic selection for protein folding using filamentous bacteriophages. *Fold Des,* 1998, vol. 3, 321-328 **[0103]**
- **HARRISON, J.L. et al.** Screening of phage antibody libraries. *Methods Enzymol,* 1996, vol. 267, 83-109 **[0105]**
- **MARTINEAU, P. et al.** Expression of an antibody fragment at high levels in the bacterial cytoplasm. *J Mol Biol,* 1998, vol. 280, 117-127 **[0106]**
- **MAXWELL KL et al.** A simple in vivo assay for increased protein solubility. *Protein Sci,* 1999, vol. 8, 1908-1911 **[0117]**