AU9515487

(12) PATENT ABRIDGMENT (11) Document No. AU-B-15487/95
(19) AUSTRALIAN PATENT OFFICE (10) Acceptance No. 696186

(54) Title
BACTERIAL RECEPTOR STRUCTURES

International Patent Classification(s)
(51)[6] **C07K 014/705** **C12N 015/00** **C07K 014/735**

(21) Application No. : **15487/95** (22) Application Date : **16.01.95**

(87) PCT Publication Number : **WO95/19374**

(30) Priority Data

(31) Number	(32) Date	(33) Country
9400088	**14.01.94**	**SE SWEDEN**

(43) Publication Date : **01.08.95**

(44) Publication Date of Accepted Application : **03.09.98**

(71) Applicant(s)
PHARMACIA & UPJOHN AKTIEBOLAG

(72) Inventor(s)
BJORN NILSSON; PER-AKE NYGREN; MATHIAS UHLEN

(74) Attorney or Agent
PHILLIPS ORMONDE & FITZPATRICK , 367 Collins Street, MELBOURNE VIC 3000

(56) Prior Art Documents
CEDERGREW ET AL PROTEIN ENGINEERING JUNE 1993 6(4) 441-8 (ABSTRACT)

(57) Claim

1. A modified binding protein derived from a native (wild-type) bacterial receptor having an $\alpha\alpha\alpha$ helical domain structure, wherein said modified binding protein derived from said bacterial receptor comprises the B domain of staphylococcal protein A, and wherein said modified binding protein is produced by mutagenesis of amino acid residues 9, 11, 14, 27 and 35 of said B domain structure, which results in a modified binding protein which specifically binds to a different binding partner than the unmodified native (wild-type) bacterial receptor, and which substantially retains the basic structure and stability of the native (wild-type) bacterial receptor.

OPI DATE 01/08/95 APPLN. ID 15487/95
AOJP DATE 14/09/95 PCT NUMBER PCT/SE95/00034
AU9515487

INT

(51) International Patent Classification 6 : C07K 14/705, 14/735, C12N 15/00

A1

(11) International Publication Number: WO 95/19374

(43) International Publication Date: 20 July 1995 (20.07.95)

(21) International Application Number: PCT/SE95/00034

(22) International Filing Date: 16 January 1995 (16.01.95)

(30) Priority Data:
9400088-2 14 January 1994 (14.01.94) SE

(71) Applicant *(for all designated States except US)*: ~~PHARMACIA AB [SE/SE]; S-171 97 Stockholm (SE).~~

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: NILSSON, Björn [SE/SE]; Torkils väg 9, S-191 73 Sollentuna (SE). NYGREN, Per-Åke [SE/SE]; Pilotgatan 22, S-128 32 Skarpnäck (SE). UHLÉN, Mathias [SE/SE]; Kvarnbogatan 30, S-752 39 Uppsala (SE).

(74) Agent: AWAPATENT AB; P.O. Box 45086, S-104 30 Stockholm (SE).

(71) Pharmacia & Upjohn Aktiebolag
S-112 87 Stockholm Sweden

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.

AUSTRALIAN PATENT OFFICE SEC 104

(54) Title: BACTERIAL RECEPTOR STRUCTURES

(57) Abstract

Novel proteins obtainable by mutagenesis of surface-exposed amino acids of domains of natural bacterial receptors, said proteins being obtained without substantial loss of basic structure and stability of said natural bacterial receptors; proteins which have been selected from a protein library embodying a repertoire of said novel proteins; and methods for the manufacture of artificial bacterial receptor structures.

Bacterial receptor structures

The present invention relates to new bacterial receptor structures originating from natural bacterial receptor structures which have been modified in regard to amino acid residues involved in the original interaction function, whereby said original interaction function has been substantially inhibited and replaced by a modified interaction function directed to a desired interaction partner.

Several bacteria known to invade mammals have evolved surface proteins capable of binding to a variety of substances including host specific carbohydrates and proteins. Several such receptors from Gram-positive bacterial pathogens have been isolated and characterized in detail as will be shown below. Most well-characterized are the Fc receptors, named after the capability of binding to the constant Fc part of IgG. Based on binding experiments to IgG from various mammalian sources, and subclasses thereof, Fc receptors have been divided into six types I-VI. The receptor from *S.aureus*, protein A [SPA], defining the type I receptor, has been the subject of immense studies.

SPA binds IgG from most mammalian species, including man. Of the four subclasses of human IgG, SPA binds to IgG1, and IgG4 but shows very weak or no interaction with IgG3 [Eliasson, M. et al, 1989 J.Biol.Chem. 9:4323-4327]. This pseudoimmune reaction has been used for more than 20 years for the purification and detection of antibodies in diagnostic, research and therapeutic applications. Cloning, sequencing and *Escherichia coli* expression of defined fragments of the SPA gene revealed a highly repetitive organization, with five IgG binding domains [E-D-A-B-C], a cell wall spanning region and membrane anchoring sequence [XM] [Uhlén, M. et al, 1984 J.Biol.Chem. 259:1695-1702; Moks, T. et al, 1986 Eur.J.Biochem. 156:637-643]. A vast number of plasmid vectors have been constructed, allowing gene fusions to different fragments of the gene for the

purpose of fusion protein production in different hosts [Nilsson B. and Abrahmsén, L. 1990 Meth.Enz. 185:144-161] (Fig. 2a).

The structure for a complex between human Fc [IgG1] and a single domain [B] of SPA has been determined by X-ray crystallography at a resolution of 2.8 Å [Deisenhofer, J. et al 1981 Biochemistry 20:2361-2370]. Based on this structure and additional information from NMR experiments, the B domain can be viewed as a compact structure consisting of three anti-parallel α-helices connected with loops. In the Fc binding, which is of both electrostatic and hydrophobic nature, only side chains of residues from helices 1 and 2 are involved, whilst the third helix is not participating in the binding. Based on this domain B, a synthetic IgG-binding domain [Z] [Nilsson, B. et al 1987 Prot.Eng. 1:107-113] has been constructed, suitable as fusion partner for the production of recombinant proteins which allows purification by IgG affinity chromatography. The high solubility and the stable structure of the Z domain has been utilized for production, purification and renaturation of a large number of recombinant proteins. [Josephsson, S. and Bishop, R. Trends Biotechnol. 6:218-224; Samuelsson, E. et al 1991 Bio.Technol. 9:363-366]

Streptococcal strains of serological groups C and G display a binding repertoire for mammalian IgGs, including human IgG3, which is even broader than for the type I receptor. The name protein G was suggested for this type III receptor from group G streptococci. In 1986 Olsson and coworkers reported on the cloning and sequencing of the gene from the serological group G streptococci [G148] [Guss, B. et al, 1987 EMBO J. 5:1567-1575; Olsson, A. et al, 1987 Eur.J.Biochem. 168:319-324]. In analogy with SPA is SPG a repetitively arranged molecule, comprising an IgG-binding region of three homologous domains [C1,C2,C3], spaced by smaller D-regions (Fig. 2A). Compared to SPA, SPG displays a different binding spectra for immunoglobulins from different species and subclasses thereof. The IgG binding

domains of protein G are now widely used as an immunological tool, i.e. in the affinity purification of monoclonal antibodies. Production of subfragments constructed by DNA-technology, have shown that an individual C-region is sufficient for full IgG-binding. Recently, the structure for a complex between the C1-domain from SPG and human Fc was determined with X-ray crystallography (Fig. 2B). This shows that SPG binds to the CH2-CH3 interface but at a different site compared to SPA. The binding is mainly of electrostatic nature which is in contrast to the large contribution of hydrophobic forces seen for the SPA-Fc interaction. Moreover, the 3-D structure of C1 differs from the X structure in that it is built up by two β-sheets connected by an α-helix [ββ-α-ββ]. The residues of C1 which according to the structure are involved in the binding, corresponds to the α-helix, the loop and the following β-sheet.

An additional activity of SPG is the capability to bind serum albumin. The binding strength is species dependent, and among the tested samples, SPG binds strongest to serum albumin from rat, man and mouse. Production and binding studies of subfragments of SPG shows that the two binding activities are structurally separated and that the serum albumin binding function is located at the repetitive A-B region [Nygren et al 1990 Eur.J.Biochem. 193:143-148]. This region has been used for several biotechnological purposes. Recombinant proteins have been produced as fusions to the region which enables the purification by affinity chromatography, where human serum albumin most frequently has been used as immobilized ligand. Proteins found to be proteolytically sensitive have been produced as "dual affinity fusions" in which they are flanked by two different affinity tails derived from SPA and SPG, respectively. Purification schemes employing both the N- and C-terminal are thus possible which ensures the recovery of an intact target protein [Hammarberg et al 1989 Proc.Natl.Acad.Sciences USA 86:4367-4371]. The strong and

specific binding to serum albumin has also been used for the *in vivo* stabilization of therapeutic proteins.

Through complex formation with the very long-lived serum albumin, the receptor is carried in the circulation (macaque apes) with a half-life which is close to the half-life for serum albumin itself. Studies in mice with the for HIV/AIDS therapy interesting, but rapidly cleared T-cell receptor CD4, showed that it was substantially stabilized when fused to the serum albumin binding region, when compared with an unfused control protein [Nygren et al 1991 Vaccines 91 Cold Spring Harbor Press 363-368]. The slow clearance can probably be explained by the complex formation with serum albumin which circumvents elimination by the liver and excretion in the kidney.

In order to determine the minimal extension required for maintained binding to serum albumin, several smaller fragments of the A-B region have been produced and analyzed. The smallest fragment so far with serum albumin binding activity is a 46 residue fragment ["B2A3"] comprising region A3 flanked by 13 and 9 residues, respectively, from regions B2 and S.

Based on homology and binding studies of other partial fragments SPG is regarded to be trivalent with regard to binding to serum albumin. Similar to the monovalent IgG-binding domains Z and C1 B2A3 is relatively small and shows high solubility and stability and is therefore a suitable candidate for modification.

Summary of the invention

The present invention has for its main purpose to provide new bacterial receptor structures by modifying natural bacterial receptors in regard to their original interaction functions to result in new structures having modified interaction functions.

Another object of the invention is to provide artificial bacterial receptor structures which are stable and more resistant to various conditions, such as increased

temperatures.

Yet another object of the invention is to provide artificial bacterial receptor structures, the interaction functions of which have been modified to direct same to other desired interaction partners.

With these and other objects that will be clear from the following disclosure in mind the invention provides for novel proteins obtainable by mutagenesis of surface-exposed amino acids of domains of natural bacterial receptors said proteins being obtained without substantial loss of basic structure and stability of said natural bacterial receptors. Said proteins have preferably been selected from a protein library embodying a repertoire of said novel proteins. In such novel bacterial receptor structures, at least one amino acid residue involved in the interaction fuction of the original bacterial receptor has been made subject to substitution by another amino acid residue so as to result in substantial loss of the original interaction capacity with a modified interaction capacity being created, said substitution being made without substantial loss of basic structure and stability of the original bacterial receptor.

It is preferred that said bacterial structures originate from Gram-positive bacteria. Among such bacteria there may be mentioned *Staphylococcus aureus*, *Streptococcus pyogenes* [group A], *Streptococcus* group C,G,L, bovine group G streptococci, *Streptococcus zooepidemicus* [group C], *Streptococcus zooepidemicus* S212, *Streptococcus pyogenes* [group A], streptococci groups A,C,G, *Peptostreptococcus magnus*, *Streptococcus agalactiae* [group B].

Of special interest are thermophilic bacteria evolved to persist in environments of elevated temperatures. Receptors from species like e.g. *Bacillus stearothermophilus*, *Thermus aquaticus*, *Thermococcus litoralis* and *Pyrococcus* have the potential of being naturally exceptionally stable, thus suitable for providing structural

frameworks for protein engineering according to the invention.

It is particularly preferred to use as a starting material for the modification of the interaction function bacterial receptor structures originating from staphylococcal protein A or streptococcal protein G.

Among preferred receptors there may be mentioned bacterial receptors originating from Fc[IgG]receptor type I, type II, type III, type IV, type V and type VI, fibronectin receptor, M protein, plasmin receptor, collagen receptor, fibrinogen receptor or protein L [K light chains], protein H [human IgG], protein B [human IgA,A1], protein Arp [human IgA].

Particularly preferred bacterial receptors originate from the Fc[IgG]receptor type I of staphylococcal protein A or the serum albumin receptor of streptococcal protein G.

In order to maintain stability and the properties of the original bacterial receptor structure it is preferred in accordance with the present invention that the substitution involving amino acid residues taking part in the interaction function of the original bacterial receptor does not involve more than about 50% of the amino acid residues of the original bacterial receptor. It is particularly preferred that not more than about 25% of the amino acid residues of the original bacterial receptor are made subject to substitution.

In regard to the original bacterial receptor structures selected for modification of their interaction functions it is particularly preferred to use receptors originating from the IgG-binding domains Z, C1, and the serum albumin binding domains B2A3.

In order to maintain as far as possible the stability and properties of the original receptor structure subject to modification in accordance with the present invention it is preferred that substitution thereof involves not more than substantially all of the amino acid residues

taking part in the interaction function of the original bacterial receptor.

In order to obtain favourable properties concerning stability and resistance to various conditions it is preferred that the bacterial receptor according to the present invention is comprised of not more than about 100 amino acid residues. It is known from scientific reports that proteins of a relatively small size are fairly resistant to increased temperatures and also to low pH and certain chemicals. For details concerning temperature resistance c.f. the article by Alexander et al. in Biochemistry 1992, 31, pp 3597-3603.

With regard to the modification of the natural bacterial receptor structure it is preferred to perform the substitution thereof by resorting to a genetic engineering, such as site-directed mutagenesis.

With regard to the interaction partner of the modified natural bacterial receptor a multitude of substances are conceivable, such as proteins, lipids, carbohydrates and inorganic substances. Among carbohydrates examples are blood group determinants and pathogen specific oligosaccharides.

In regard to proteins conceivable interaction partners are IGF-I, IGF-II, hGH, Factor VIII, insulin and apolipoprotein and their respective receptors as interaction partners. Furthermore, by selecting new receptor variants with specificity for different folding forms of proteins, affinity resins or analytical tools to facilitate the isolation of correctly folded molecules can be produced. Further examples are viral coat proteins, bacterial antigens, biotin and cell markers, such as CD 34 and CD 4.

Although the present invention is applicable to a variety of natural bacterial receptors the following illustration of the invention more in detail will be directed to the use of the IgG-binding domains Z, C1 and B2A3. The concept of the present invention residing in the

use of artificial bacterial receptors based on the natural structures of naturally occurring bacterial receptors is associated with several advantages. Thus, the invention makes it possible to use robust and stable, highly soluble and secretion competent receptors. This is in contrast to previous techniques based on the use of polyclonals and monoclonals, such as for diagnostic purposes, which are not very stable in connection with storage, varying conditions, such as varying temperatures etc. Furthermore, the invention makes it possible to modify natural bacterial receptors to obtain desired interaction capacities for specific purposes.

For the selection of such functional variants in a large repertoire, a powerful selection system must be employed. Recent developments in this field offer alternative methods. One of the most important tools for protein engineering that has emerged during the last years is the phage display of proteins. By recombinant DNA techniques, single phage particles can be prepared which on their surface carries a protein fused to a phage-coat protein. By panning from a large pool of phages bearing different proteins, or variants of a specific protein, specific phage clones can be sorted out, which displays a certain binding characteristic [WO92/20791 to Winter et al]. Since the phage particle contains packed DNA encoding the phage protein components, a coupling between the specific variant of the displayed protein and the corresponding genetic information is obtained. Using this technique, typically 10^9 phage clones can simultaneously be generated and subjected to panning for screening of desired characteristics. The phage display technique can be used for selection of both small peptides as well as more complicated proteins such as antibodies, receptors and hormones. For selection of proteins which cannot be secreted, which is a prerequisite for phage display, intracellular systems have been developed in which the library of proteins are fused to a repressor protein with affinity

for a specific plasmid-borne operator region resulting in a coupling between the specific protein variant and the plasmid that encoded it. An alternative to the phages as bearer of protein libraries would be to use bacterial cells. Recently, display of recombinant proteins on the surface of *Staphylococcus xylosus* based on fusions to the cell-wall anchoring domain was demonstrated, which opens the possibility of display also of repertoires of proteins for affinity selection of specific variants [Hansson, M. et al 1992 J.Bacteriology 174:4239-4245]. Furthermore, by structure modelling using computer graphic simulations, predictions of the binding and function of altered variants of a protein can theoretically be done before the construction of the gene encoding the protein.

As indicated above the present invention describes the construction of novel proteins based on the mutagenesis of surface exposed amino acids of domains derived from bacterial receptors. These artificial bacterial receptors can be selected for different applications using a phage display system. The benefits from using bacterial receptors as structural frameworks are several. They have evolved to express a binding function without disturbing the overall structure. They are naturally highly soluble, robust to unphysiological conditions, such as pH and heat, folding efficient and are in addition secretion competent.

The invention finds use in several different areas.

The introductory part of the above-identified patent specification WO92/20791 gives an excellent survey on antibodies and their structure. Reference is particularly made to page 1 thereof.

The bacterial receptors SPA and SPG have been widely used in antibody technology for detection and purification of antibodies from e.g. hybridom supernatants and ascites fluids. However, not all antibodies are recognized by these receptors, depending on species and subclass. For the smaller subfragments of antibodies (Fig. 4), SPA and SPG show a limited binding, and efficient tools for gene-

ral purification schemes are lacking. However, from a repertoire of mutant receptors including SPA and SPG, forms displaying a broader affinity for antibodies and subfragments thereof can potentially be selected.

The complex structural organization of antibodies has a number of consequences for their use in different applications as well for the production of recombinant derivatives. For use in immunosorbents, the arrangement of subunits connected by disulphide bonds can lead to a leakage of released heavy and light chains from columns. The requirement of successful docking of the two subunits contributing to the antigen binding site makes the production in bacteria of small subfragments with a low association difficult. The folding of the antibody is dependent on the formation of intra- and inter chain disulphidebonds, which are not able to form in the intracellular environment of bacterial cells. High-level intracellular expression systems for recombinant antibodies leads to the formation of inclusion bodies, which have to be renatured to obtain biological activity. These limitations make it worthwhile to search for alternatives for use as protein domains capable of specific binding, to replace antibodies in a vast number of applications.

The CDR regions forming the antigen bidning part of an antibody forms a total surface available for the antigen of approximately 800 Å^2, with typical 10-20 residues from the antibody involved in the binding. Using the structure of the complex determined by X-ray crystallography between an individual domain B of SPA and human fc[IgGI] as a starting point about 15 amino acids of the said domain involved in this binding can be determined or postulated. The binding surface of about 600 Å^2 is of the same order of magnitude as between an antibody and its antigen. By arbitrary *in vitro* mutagenesis of these positions simultaneously there is obtained a large library of Z variants with modified functional properties. In view of the fact that the regions of the Z domain constituting the

very stabile so called three-helix bundle is maintained in its native form spectra of proteins are generated which could be considered as "artificial antibodies" and which have the expected high solubility and excellent folding properties capable of binding to a large number of new ligands. Fusions of these artificial receptors to constant regions can be constructed to recruite effector functions, such as complement binding or triggering of ADCC (antibody dependent cellular cytotoxicity).

There are several potential advantages of utilizing the SPA structure [Z] as a starting point for such "artificial antibodies" or artificial bacterial receptors. For a period of about 10 years a large number of proteins have been produced as fusions to SPA, where one has utilized the unique properties of the fusion partner in expression, refolding and purification. In these applications the Z domain has been found to be extremely soluble, stable against proteases, easy to produce in large amounts and foldable to a correct structure also intracellularly in *Escherichia coli* (no cysteins). Immunoglobulins (Ig:s) are substantially tetramers built up from so called β-sheet structures which stabilize the orientation of the antigen-binding loops which in turn consist of continuous peptide sequences. This is to be compared to the monomeric Z domain built up from so called three-helix bundle consisting of three closely packed α-helix structures, where the Fc-binding amino acids are found discontinuously in the sequence but in the folded protein are positioned on one and the same binding surface. This difference with regard to the structural elements contributing to the formation of the binding surface enables new possible conformations which cannot be obtained in natural antibodies. The ability of Z to be folded to the native structure also under conditions prevailing in the site of cytoplasma opens the possibility of using also derivatives thereof clinically. Genes coding for artificial antibodies with for example virus-neutralizing capacity can be distributed to cells

through so called gene therapy resulting in interrupting the infection at an early stage.

From structure data for the complex between an individual Ig-binding domain [C1] of SPG and human Fc the binding surface can be studied. The binding which is essentially of an electrostatic nature involves side chains from amino acids from the α-helix as well as from the subsequent β-sheet [#3]. These differences in structure compared to the Z domain makes it useful also to create a library of C1 variants to investigate whether differences in binding patterns for artificial antibodies can be observed depending on the different conditions as regards the topology of the binding surface. Repertoires based on the structures of these and other receptors therefore offer different possibilities in the creation of artificial forms with novel functions.

When producing recombinant proteins the purification of the product is frequently a major problem. By expressing the target protein as a fusion to a so called affinity tail the hybrid product can effectively and selectively be recovered from the cell lysate or in certain cases from the culture medium by passage through a column containing an immobilized ligand. Several such gene fusion systems have been described which are based on the interaction of a certain protein with a ligand. For industrial applications it is often desirable to clean effectively the columns between the runs to satisfy purity requirements by authorities. Depending on the nature of proteins the relatively harsh conditions (NaOH, acids, heat) often used for organic or physical matrises, for example in ion exchange chromatography and gel filtration, can normally not be used. Here the use of new ligands based on stable structures originating from bacterial receptors are of great importance. In this connection the Z domain from SPA is an excellent example since said domain can be subjected to such difficult conditions as a pH of 1 or heating to 80°C without denaturating non-reversibly (see Example 2

below). From the library of for example Z variants interesting protein products can be selected for use immobilized on a solid phase for affinity chromatography. These protein ligands are resistant to effective purification conditions and are therefore useful repetitively on a large scale. In traditional immuno affinity chromatography where immobilized monoclonal antibodies are used for the selective purification of a certain product there are problems with leakage from the column of subunits (heavy and light chains) of the antibody since it consists of four polypeptide chains linked by cystein bridges. Since the artificial bacterial receptors consist only of one polypeptide chain this problem will be avoided. One particular area of interest is selection for binding to carbohydrates. Lectins, nature's binders to this large and important group of biomolecules, have been found to be difficult to purify and have limited stability. Since the generation of antibodies against carbohydrates has been found to be quite complicated a selection of new artificial lectins will be of great importance to research, diagnostics and therapy.

In the production of recombinant proteins in bacterial hosts precipitates of the gene product are frequently formed, so called inclusion bodies. In order to obtain a native structure of the protein this must be subjected to refolding *in vitro*. A limitation in such process one is often confronted with is the fact that a great part of the material precipitates in the procedure which results in low yields. By producing the protein with an extension in the form of either a short hydrofilic peptide or an easily soluble complete domain [Samuelsson, E. et al 1991 Bio/Technol. 9:363-366] substantially higher concentrations of the protein will be obtained without precipitation taking place during renaturation. For example the high solubility of the said domain enables the use of increased solubility of proteins in either refolding from inclusion bodies or in so called reshuffling of disulphide

bridges. From libraries of artificial receptors new forms can be selected having improved properties to facilitate and even make refolding of recombinant proteins possible (cis-acting chaperones).

Recently a new unit operation for the purification of recombinant proteins based on ion exchange chromatography in so called expanded bed has been described [Hansson, M. et al 1994 Bio/Technol. in press]. In this connection there is used a difference in isoelectric point between the target protein and the proteins of the host cell for selective enrichment on a positively charged ion exchange matrix. By fusion to the acid Z domain (pI 4.7) the ion exchange step can take place at a pH, at which the majority of the contaminants were of the opposite charge compared to the fusion protein. By constructing libraries of bacterial receptors where a selection of amino acids have been replaced by the acid amino acids aspartate and glutamate also very acid and solubility increasing domains can be produced for use as fusion partners in the production of recombinant proteins.

As previously described affinity systems based on protein ligands are not totally suitable for industrial purposes in view of the harsh conditions required in the cleaning of columns. Therefore, there is a need for fusion partners having specific affinity towards simple and cheap organic ligands. Panning of phage display libraries of different bacterial receptors against such ligands provide novel affinity tails suitable for the use as fusion partners for the production purification of recombinant proteins.

The present invention provides means for producing and selecting proteins with novel functions. According to the invention this is achieved by extensive mutation of defined residues of stable domains of bacterial receptors. Due to the novel functions of the artificial bacterial receptors, these can be used as specific binders for therapeutic, diagnostic, biotechnology or in research.

The present invention will now be described more in detail by specific examples with reference to the appended drawings. In the drawings:

Figure 1. A. Schematic representation of staphylococcal protein A showing the signal peptide (S), five IgG-binding regions [E-D-A-B-C], followed by cell wall anchoring region [X-M].
B. Computer graphic representation of the complex between domain B from SPA and human Fc1 determined by X-ray crystallography. Note that the third helix of SPA is not seen in this figure.

Figure 2. A. Schematic representation of streptococcal protein G from the strain G148 showing the signal peptide (Ss), region E (E), the repetitive serum albumin binding A-B region, the spacer region (S), followed by the IgG binding domains C1 through C3, spaced by the D regions and finally the cell wall anchoring region W.
B. Computer graphic representation of the complex between domain C1 of SPG and human Fc1 determined by X-ray crystallography.

Figure 3. Schematic representation of the three helix bundle structure of the 58 residue SPA analogue Z. Indicated are some of the side chains proposed to be involved in the binding to Fc with the exception of F30, which stabilizes the helix-helix packing.

Figure 4. IgG antibody structure, showing the different subfragments Fab,Fd,Fc and the scFv composed of the VH and VL connected by a short (ca 15 aa) linker.

Figure 5. A. General concept for the gene assembly strategy used for the generation of the Z gene libraries. For the construction of the library of acid Z derivatives, only residues 9, 11, 14, 27 and 35 were altered using the degenerated oligonucleotides ACID-1, ACID-2. The PCR primers used for the amplification of the assembled library were ZLIB-3 (PCR primer 5') and ZLIB-5 (PCR primer 3').
B. The PCR products from the amplification of the assembled library encoding 46 of the 58 residues of the Z-domain can be cloned into phagemid DNA harboring the remaining C-terminal part of Z. This gene is fused in frame with the gene for protein III of the M13 family of *E.coli* bacteriophages. This enables the display on the phage surface of the repertoire of acid Z-variants.

Figure 6. Oligonucleotides used for the construction of Z-libraries. For the library of acid Z-variants described in Example 2, only oligonucleotides ZLIB-1, 2, 3, 4, 5, LONGBRIDGE, ACID-1 and ACID-2 were used.

Figure 7. DNA sequences of clones derived from the acid Z protein library. Bold figures indicate amino acid positions in the Z-domain. For clarity the positions of the restriction sites Acc I and Nhe I are shown.

Figure 8. Result from analysis of the temperature stability of an individual Z domain at pH 2.9. The content of α-helicity in the sample was monitored by measuring the ellipticity at s222 nm during a temperature scan.

Figure 9. *Phagemid vector pKN1*. The library PCR products encoding the variegated helices 1 and 2 (both the acid and the extensive library) was subcloned into the phagemid vector, pKN1, containing the gene for residues 44-58 of the wild type Z domain (essentially helix 3), followed by the gene for a 46 residues serum albumin binding region (ABP) derived from streptococcal protein G linked in frame with a truncated version of the M13 phage coat protein 3 gene. The phagemid contains the origin of replication derived from plasmid pBR322 as well as the intergenic region (f1 ori) required for packing into the phage particles.

Figure 10. **SDS-PAGE.** HSA-affinity purified proteins from the periplasm of *Escherichia coli* cells producing the wild type Z domain and two different acid Z-variants as ABP fusion proteins encoded from their respective phagemid vectors were analyzed by SDS/PAGE. M, molecular weight marker; lane 1, wild type Z domain; lane 2, clone 10; lane 3, clone 12.

Figure 11. **CD-data.** Overlay plot of CD spectra obtained for the wild type Z domain and two variants of the Z-protein library. The signals of the proteins were obtained after subtraction of the CD signal contribution of the ABP tail, present during the analysis.

Figure 12. **Ion exchange chromatography.** The two acid Z-variant proteins no. 10 and no. 12 together with the wild type Z-domain (produced as ABP fusion proteins) were each subjected to analysis at pH 5.5, employing an anion exchange chromatography column. Elution of the proteins

from the column was obtained by a NaCl gradient. Top: acid Z-variant no. 12; middle, acid Z-variant no. 10; bottom, Z (wild type). Note that the wild type Z protein was not retarded on the column at this pH.

Figure 13. **Z-domain structure.** A main-chain trace representation of the model of the structure of the native Z-domain. The structure of helices one and two are from the co-crystal structure between domain B of SPA and Fc (Deisenhofer, (1981) Biochemistry, 20, 2361-2370). The third helix was built based on the secondary structure assignments from NMR spectroscopy (Gouda et al., (1992) Biochemistry, 31, 9665-9672). Non-hydrogen atoms of side-chains of residues that were mutated in the construction of the combinatorial library are displayed as ball-and-stick models. The display was generated by the program MOLSCRIPT (Kraulis (1991) J.Appl. Cryst., 24, 946-950).

Figure 14. **Amino acid sequences.** Result from DNA-sequencing of 31 randomly chosen Z-variants from the library. The residues subjected to the mutagenesis are boxed. Horisontal lines indicate nucleotide identity with the wild type Z sequence listed at the top. Indicated are the clones that were expressed and characterized as fusion proteins to the ABP-tail.

Figure 15. **Aminoacid distribution.** Result from the statistical analysis of the deduced amino acids at the mutated positions. In total, 13 residues from 31 clones (403 codons) were included in the calculation. The ratios between observed and expected frequencies are shown for all 20

amino acids as well for the only termination signal (TAG) included in the NNG/T degeneracy profile.

Figure 16. **SDS-PAGE analysis.** HSA-affinity purified proteins from the periplasm of E.coli cells producing the wild type Z domain and four different Z-variants as ABP fusion proteins encoded from their respective phagemid vectors were analyzed by SDS/PAGE. Lanes 1-5: Reduced conditions. Lanes 6 and 7: Non-reduced conditions. Lane 1, wild type Z domain; lane 2, clone 16; lane 3, clone 21; lane 4, clone 22; lane 5, clone 24; M, molecular weight marker; lane 6, clone 16 and lane 7, clone 22.

Figure 17. **CD-data.** Overlay plot of CD spectra obtained for the wild type Z domain and four variants of the α-helical protein surface library. The signals of the variants were obtained after subtraction of the CD signal contribution of the ABP tail, present during the analysis.

Figure 18. **Biosensor assay.** An overlay plot of sensorgrams obtained from the BIA-core™ analysis of the wild type Z domain and four different variants (no. 16,21,22,24; Figure 4) fused to the ABP tail. The IgG-binding activities of the different proteins were analyzed using a sensor chip coated with approx. 5000 RU human polyclonal IgG and injections of 45 µl pulses at 2 µl/min of 1500 nM solutions of the different proteins. Note that the differences in plateau values of signals during the injections of the variants no. 16,21,22 and 24 is due to divergent dilutions into the driving buffer.

All reagents and DNA constructions are available at the department for Biochemistry and Biotechnology, Royal Institute of Technology, Stockholm, Sweden.

Material

The oligonucleotides (Fig. 6) were purchased from Scandinavian Gene Synthesis (Sweden), and phosphorylated where indicated according to [Maniatis et al (1988) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press]. ZLIB-1 was biotinylated in the 5'-end enabling immobilization on paramagnetic beads M-280 Streptavidin purchased from Dynal A/S (Norway). Washing/binding buffer was 1 M NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA (ethylenediamine tetraacetic acid). The annealing/ligation buffer was 30 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.2 mM ATP, 1 mM 1.4 dithiothreitol (DTT). DNA ligase were from Boehringer Mannheim, Germany. 10 x PCR buffer contained 20 mM $MgCl_2$, 2 mM dNTPs, 100 mM Tris-HCl, pH 8.3, 50 mM KCl, 1% Tween 20. *Taq* DNA polymerase was from Cetus Inc., USA. The thermal cycler was a Perkin-Elmer 9600. For the temperature/stability scanning a J-720 spectropolarimeter (JASCO, Japan) was used. *Escherichia coli* strain RR1ΔM15 [Rüther, U. (1982) Nucl.Acids Res. 10:5765-5772] prepared for competence [Maniatis et al (1988) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press] was used as host for the transformation. Agar plates contained 100 µg/ml of ampicillin.

EXAMPLE 1

Construction of an acid Z-library

The synthetic 58 residue SPA analogue Z [Nilsson et al, Prot.Eng] was subjected to a mutagenesis approach to construct new variants with an altered pI, in order to produce fusion partners for recombinant proteins to be purified by ion-exchange chromatography. Based on the crystal structure of the complex between the B-domain of SPA and human Fc1 [Deisenhofer, J. et al 1981, Biochemistry

20:2361-2370], five residues from the B-domain participating in the binding were chosen as targets for mutagenesis. These five codons corresponding to the Z-residues No. 9, 11, 14, 27 and 35 positioned in helices 1 and 2 were altered simultaneously using degenerate oligonucleotides with the triplet sequence G(C/A)(C/A) at these positions resulting in the codons for the amino acids alanine (50%), aspartic acid (25%) and glutamic acid (25%), respectively. Using a solid phase gene assembly strategy [Ståhl et al, Biotechniques 14:424-434] a library of genes encoding 3^5 (243) acidic variants of the synthetic IgG-binding Z-domain was created (Fig. 5). Twenty microlitres (200 µg) of paramagnetic streptavidin-coated beads were washed with washing/binding buffer and incubated with 15 pmole of pre-hybridized oligonucleotides ZLIB-1 (biotinylated) and ZLIB-2, for 15 min at RT at a final volume of washing/-binding buffer of 40 µl. After ligation and washing, approximately 15 pmole each of the oligonucleotides ACID-1 (degenerated), LONGBRIDGE, and ACID-2 (degenerated) and the preannealed linker pair ZLIB-4/ZLIB-5 were added in a repetitive manner, with intervening washing steps according to Ståhl et al [Biotechniques 14:424-434]. After completed assembly, the different fragments were ligated for 15 min at 37°C. To amplify the amount of DNA coding for the Z(Acid)-library still immobilized onto the beads, a fraction was withdrawn and subjected to PCR. The PCR mix (50 µl) contained one pmole each of PCR primers ZLIB-3 and ZLIB-5, 5µl each of the ligation mix, 10 x PCR buffer and 10 x CHASE, 1 unit of *Taq* polymerase and sterile water to 50 µl. The temperature cycling programme was: 96°C, 1 min, 60°C, 1 min and 72°C, 2 min, repeated for 35 cycles. Analysis by 1% agarose gel electrophoresis showed a band of the expected size of 179 bp, showing the feasibility of the assembly concept. The 179 bp band from the PCR of the Z(Acid)-library, was cut out from the gel and purified (Geneclean™, Bio 101, Inc. USA) prior to insertion in a plasmid vector (TA-cloning™ kit, Invitrogen, Inc. USA)

suitable for solid phase DNA sequencing [Hultman et al, 1988]. After transformation and spreading on ampicillin containing agarplates two colonies were chosen for analysis of the obtained sequences. The results (Fig. 6) show that the expected degeneracy was found at the desired positions.

EXAMPLE 2

Measurement of the temperature stability of the Z conformation

The temperature stability of the Z conformation was determined by following the ellipticity at 222 nm by circular dichroism (CD) spectroscopy through a temperature scan. This wavelength is used to monitor the presence of α-helicity of Z [Cedergren et al. 1993 Prot. Eng. 6:441-448]. The experiment was performed at a rather low pH (approximately 2.9) in order to destabilize the molecule since the mid-point of temperature denaturation (Tm) is ~95°C at neutral pH (data not shown), which is outside the range that can be determined by a complete scan through the transition under normal atmospheric pressure. The experiment shows (Fig. 4) that the Tm (as defined by the inflexion point of the temperature scan) of the Z domain is as high as 71°C at pH 2.9. This demonstrates the extreme temperature stability of the α-helices of the Z molecule.

The experiment was performed in a J-720 spectropolarimeter (JASCO, Japan) and the temperature was controlled by circulating water through the cuvette holder from a NESLAB water bath. The temperature was monitored in the cuvette through a micro sensor device (JASCO, Japan). The buffer was 50 mM acetic acid, pH 2.9. The protein was domain Z [Cedergren et al 1993 Prot. Eng. 6:441-448] at a protein concentration of 50 µg/mL and the cuvette cell path length was 1 cm. The temperature scan speed in the experiment was 50°C/h.

Example 3. Characterization of proteins derived from the acid Z-library.

Two protein variants derived from the acid Z-library were expressed in *Escherichia coli*, purified and characterized using SDS-PAGE, circular dichroism and ion exchange chromatography. The PCR products from a solid phase gene assembly (see example 1) were restricted with 45 U *Esp* 3I (Labassco AB, Sweden) and 50 U *Nhe* I (Pharmacia, Sweden) in 200 µl buffer (33 mM Tris-acetate, pH 7.9, 10 mM Mg-acetate, 66 mM potassium-acetate, 0.5 mM DTT and 0.1 mg/ml BSA). The mix was overlaid with mineral oil and incubated at 37°C over night. The restricted fragments (approximately 5 µg) were purified by phenol/chloroform/isoamylalcohol extraction followed by additional washing with chloroform and later ethanol precipitated before ligation at 15°C over night to *Mlu* I-*Nhe* I cleaved pKN1 vector (1 µg) (see below) using 13.5 Weiss units of T4 DNA ligase. The ligation mixture was heat-treated at 70°C for 20 min, extracted with phenol/chloroform/isoamylalcohol followed by washing with chloroform, ethanol precipitated and redissolved in 20 µl of sterile water.

The phagemid vector pKN1 (figure 9) was constructed in several steps as follows. A double stranded linker encoding the invariant residues 44-58 of the Z-domain was formed from oligonucleotides ZLIB-6 and ZLIB-7 and cloned as a *Mlu* I-*Xho* I fragment into phagemid pKP986 (A kind gift from Dr. Lars Abrahmsén, Pharmacia BioScience Center, Sweden), resulting in pKN. Plasmid pKP986 encodes the *E. coli* Omp A leader peptide followed by residues 249-406 of fd filamentous phage coat protein 3 (Lowman *et al.* (1991) *Biochemistry*, 30, 10832-10844.) under the control of a *lac* promoter. A gene fragment encoding a monovalent serum albumin binding region derived from streptococcal protein G was amplified by PCR from the plasmid pB2T (Eliasson *et al.*, *Molecular Immunol.*, 28, 1055-1061), using primers ABP-1 and ABP-2 (which contain *Xho* I and *Sal* I recognition sites, respectively) and cloned into *Xho* I restricted plasmid pKN, yielding pKN1. This phagemid vector thus encodes for the Omp A signal peptide, the third helix of the wild type Z domain followed by a 46 residue albumin binding protein (ABP) linked to residues 249-406 of fd phage protein III and is adapted for insertion of *Esp* 3I/*Nhe* I-digested PCR products encoding variegated helices one and two of the Z domain.

Freeze competent *E. coli* RR1ΔM15 (*supE44 lacY1 lacZ ara-14 galK2 xyl-5 mtl-1 leuB6 proA2 Δ(mrcC-mrr) recA⁺ rpsL20 thi-1 lambda⁻* F'[*lacI*q *lacZ*ΔM15]) (Rüther, (1982) *Nucleic Acids Research*, 10, 5765-5772) cells were transformed with the ligation mixture according to Maniatis and coworkers (Maniatis *et al.* (1982)

Molecular cloning: A Laboratory Manual, Cold Spring Harbor, Cold Spring Harbor Laboratory Press) and plated on agar plates containing 100 µg/ml ampicillin (Sigma, USA) and 1% glucose. Small amount of cells from randomly picked colonies were separately subjected to two-step PCR amplifications (30 cycles: 96°C, 15 s; 72°C, 2 min) on a GeneAmp PCR System 9600 (Perkin Elmer, USA), using 5 pmoles of primers RIT-27 and NOKA-2 (biotinylated) in 20 mM TAPS (pH 9.3), 2 mM $MgCl_2$, 50 mM KCl, 0.1% Tween 20, 0.2 mM deoxyribonucleoside triphosphates (dNTPs) and 1.0 U of *Taq* DNA polymerase (Perkin-Elmer). The solid-phase DNA sequencing of the PCR products was performed employing the FITC labeled sequencing primers NOKA-3 (for the immobilized strand) and ABP-2 (for the eluted strand) on a robotic workstation (Biomek™ 1000, Beckman Instruments, Fullerton, CA) and an Automated Laser Fluorescent (A.L.F.) DNA Sequencer™ (Pharmacia Biotech, Sweden) as described by Hultman and coworkers (Hultman *et al.*, (1989) *Nucleic acids Research*, **17**, 4937-4946).

Two clones with the different encoded acid aminoacid substitutions (bold face) at the positions 9, 11, 14, 27 and 35 in the Z-domian according to table 1 were selected for further analysis. The wild type Z domain and the two different acid Z-variant proteins (clones no. 10 and 12) were expressed from their respective phagemid vectors as fusions to the serum albumin binding tail (ABP) and purified by human serum albumin-affinity chromatography.

Table 1. Amino acid substitutions for selected clones in the acid Z-library [a].

Clone no.	*Encoded aminio acid at position no.* 9	11	14	27	35
w.t.	Gln	Asn	Tyr	Arg	Lys
10	**Glu**	**Asp**	**Asp**	Ala	**Glu**
12	**Glu**	**Asp**	**Asp**	Ala	Ala

a. Letters in bold face indicate acid aminoacids

Colonies of *E. coli* RR1ΔM15 cells harbouring the corresponding phagemid vectors were used to inoculate 100 ml of Tryptic Soy Broth (Difco), supplemented with ampicillin (100 µg/ml). The cultures were grown at 37°C to an OD_{600nm}=1, followed by induction with a final concentration of 1 mM IPTG and incubation at 30°C over night. The cells were harvested by centrifugation at approximately 5000 *g* for 10 min and periplasmic proteins released by osmotic shock. The periplasmic content from the cells was subjected to affinity chromatography on HSA-Sepharose as described by Nygren and coworkers (Nygren *et al.*, (1988) *J. Mol. Recognit.*, 1, 69-74) and analyzed by SDS/PAGE on a homogeneous 12% slab gel (BioRad Inc., USA), which was stained with Coomassie Brilliant Blue R-250. For all proteins appr. 1.5-2.5 mg/L culture could be recovered, indicating similar production and secretion efficiencies for the variants and the wild type domain. In addition, the results from the SDS-PAGE analysis (Figure 10) of purified proteins suggest that the acid Z variants analyzed are stably expressed in *E. coli*.

To investigate if the secondary structure content of the derivatives was preserved after the surface mutagenesis, a subtractive circular dichroism analysis was performed. IgG- or HSA-affinity chromatography purified proteins Z, Z-ABP, the acid derivatives no. 10 and 12 fused to the ABP tail as well as the ABP-tail itself were subjected to a 250 to 184 nm (far UV) circular dichroism analysis at room temperature using a J-720 spectropolarimeter instrument (JASCO, Japan). The scanning speed was 10 nm/min. The cell pathlength was 1 mm. Solutions (approximately 0.1 mg/ml) of the different proteins were prepared in 20 mM phosphate buffer pH 6.5, supplemented with 0.05 % Tween 20 (Kebo AB, Sweden). Accurate protein concentrations were determined by amino acid analysis on a Beckman 6300 amino acid analyzer equipped with System Gold data handling system. CD signals for the derivatives were obtained by subtracting the signal obtained for the ABP tail, after adjustments for differences in protein concentrations, followed by normalization for amino acid contents.

A comparison of signals obtained from 250 to 184 nm for the wild type Z domain and the acid variants fused to the ABP-tail was performed after subtraction of the contribution from the ABP tail itself. The result shows that for the two acid Z-derivatives, spectra similar to the wild type Z domain were obtained with a characteristic minimum at 208 nm and an inflexion point at 222 nm (Johnson, 1990) (Figure 11). This suggests that the three helix bundle framework is preserved in these mutants.

The two Z-variants, no. 10 and 12, contain four and three introduced acid aminoacids, respectively, compared to the native Z domain. In order to investigate if the introduced acidity was reflected as differences in their isoelectric points, they were subjected to a gradient elution from an anion exchange column. The proteins Z(wild type) and the acid variants no. 10 and no. 12 (all produced as ABP fusion proteins) were each (5 µg) dissolved in 300 µl of 20 mM Piperazine buffer (pH 5.5) and separately applied at 100 µl/min on a MonoQ, PC 1.6/5 column (Pharmacia, Sweden). Elution of the proteins were performed by applying a NaCl gradient in Piperazine buffer (pH 5.5) (Sigma, USA) ranging from 0-50 % NaCl in 20 min. The results from the analysis (figure 12) shows that the two acid Z.-variant proteins were eluted at different NaCl concentrations suggesting clear differences in isoelectric points. In contrast, at the pH chosen during the experiments, the wild type Z-domain did not interact with the resin, and was therefore seen in the flow-through.

Thus, the series of experiments performed on the two acid Z-variant proteins shows that the expression behaviour, proteolytic stability and secondary structure content of the variants were unchanged when compared to the native Z-domain. Furthermore, a novel functions were introduced into the two Z-variants by the substitution of surface located positions with acid amino acids. The two acid variants can be used e.g. as fusion partners to facilitate purification of recombinant proteins by ion exchange chromatography at low pH. Thus, it is showed that among the members of the acid Z-library, variants with novel functions can be isolated.

Example 4. Construction and characterization of a combinatorial library of Z-variants.

A library of Z-variants was assembled using a solid-phase gene assembly strategy (see example 1). Most of the amino acid residues suggested to take part in the binding to Fc (Deisenhofer, (1981) *Biochemistry*, 20, 2361-2370) were found to be on the molecule surface (Q9, Q10, N11, F13, Y14, L17, N28, Q32 and K35), and therefore included in the mutagenesis. In addition, based on their surfacial location, other residues (H18, E24, E25 and R27) were also decided to be included. In total, 13 residues in the Z scaffold where thus chosen for simultaneous and random mutagenesis. A set of oligonucleotides (Figure 6) were synthesized for construction of the library of surface mutants of the 58-residues monovalent IgG-binding domain denoted Z. In this library, the codons for Q9, Q10, N11, F13, Y14, L17 and H18 located in the first α-helix and E24, E25, R27, N28, Q32 and K35 in the second α-

helix of the Z domain (Figure 13) were substituted for degenerate NNK (K=G or T) codons using a solid phase strategy utilizing the single stranded degenerate oligonucleotides for the assembly. The chosen NNK degeneracy includes 32 codons covering all 20 amino acids, including the TAG (amber) termination signal.

Oligonucleotide ZLIB-1 was synthesized with a 5' biotin group to enable robust anchoring onto streptavidin-coated paramagnetic beads used as solid support during the gene assembly. This ZLIB-1 oligonucleotide, together with its complementary sequence (ZLIB-2) encodes residues 1-8 of the Z domain, preceeded by the first six residues of region E of protein A which were included to facilitate the *E. coli* secretion of the Z variants (Abrahmsén *et al.*, (1986) *EMBO J.*, 4, 3901-3906). The oligonucleotides DEGEN-1 and DEGEN-2 (Table I) encode the two mutated helices of the Z domain, respectively, normally involved in Fc-binding. Theoretically, full and simultaneous NNK degeneracy at the 13 selected positions would yield a combinatorial library of appr. $8 \cdot 10^{16}$ protein variants encoded by $3.7 \cdot 10^{19}$ different DNA sequences. However, here the assembly of the library was initiated by the immobilization of appr. 15 pmole of prehybridized oligonucleotides ZLIB-1 and ZLIB-2 (Figure 6), which limits the theoretical size of the Z-library to appr. $0.9 \cdot 10^{13}$ different DNA sequences encoding appr. $2 \cdot 10^{10}$ Z variants. The assembly was continued by the addition and ligation of a preformed construct, obtained after ligation of equimolar amounts of oligonucleotides DEGEN-1 and DEGEN-2, facilitated by the bridging oligonucleotide BRIDGE (Figure 6).

To complete the assembly, a fragment consisting of the prehybridized oligonucleotides ZLIB-4 and ZLIB-5 was added to the beads for ligation. This fragment encodes the second loop and the first six residues of the unaltered third helix of the Z domain. After completed assembly, oligonucleotides ZLIB-3 and ZLIB-5, containing the recognition sequences for the endonucleases *Esp* 3 I and *Nhe* I respectively, were used as primers for PCR amplification of the assembled constructs using one tenth of the bead-immobilized ssDNA as template (theoretically corresponding to $2 \cdot 10^9$ protein variants). To avoid unwanted interference during the amplification, oligonucleotides ZLIB-2, BRIDGE and ZLIB-5 were first eluted with alkali. The resulting PCR product was analysed by agarose gel electrophoresis and found to be homogenous and of the expected size, 179 bp.

The PCR product was subcloned into the pKN1 phagemid vector containing the gene for residues 44-58 of the wild type Z domain in frame with a truncated version of the fd phage coat protein 3 gene for surface display on phage particles upon helper phage

superinfection of phagemid transformed *E. coli* cells (Lowman *et al.*, (1991) *Biochemistry*, 30, 10832-10844) (Figure 9). In addition, the phagemid vector contains an interspaced in-frame cassette encoding a 5 kDa (46 aa) serum albumin binding region (denoted ABP) derived from streptococcal protein G (Nygren *et al.*, (1988) *J. Mol. Recognit.*, 1, 69-74; Nilsson *et al.*, (1994) *Eur. J. Biochem.*, 224, 103-108), enabling efficient affinity purification of produced Z variants devoid of their native Fc-binding activity. Furthermore, the serum albumin binding activity can potentially be used for pre-selection of phage particles carrying recombinant molecules, prior to the panning for Z variants with new binding functions, to decrease the background originating from unspecifically bound non-recombinant phage particles.

After transformation, PCR screening (using the oligonucleotides RIT-27 and NOKA-2) of 25 clones showed that over 95% (24/25) of the clones contained an insert of the expected length, suggesting that the gene assembly procedure was carried out with high efficiency. Fortyfive transformants were randomly selected and subjected to direct solid phase DNA sequencing (see Example 3) in order to further analyze the quality and heterogeneity of the library. Approximately 69% of the clones were correct, containing wild type and degenerate codons at expected positions. The remaining clones had spurious discrepancies which in part can be attributed to the oligonucleotide synthesis or errors introduced during PCR. The correct clones (31 clones) (Figure 14) were further analyzed for codon representation at the 13 degenerate positions. The distribution of the total 403 resulting deduced amino acids among the 32 codons included in the NNK degeneracy profile shows a close correlation with the expected frequencies for these yet unselected clones (Figure 15). To investigate the expression and stability of the Z-variants, four clones (no. 16, 21, 22, 24; Figure 14) with different degrees of substitution as well as the wild type Z domain were produced as ABP fusions encoded from their respective phagemid vectors. Soluble proteins from the periplasm of IPTG-induced cultures were subjected to HSA-affinity chromatography employing the ABP-tail for general and efficient recovery (Nygren *et al.*, (1988) *J. Mol. Recognit.*, 1, 69-74). For all proteins appr. 1.5-2.5 mg/L culture could be recovered, indicating similar production and secretion efficiencies for the variants and the wild type domain. The results from a SDS-PAGE analysis (Figure 16) of purified proteins suggest that the four Z variants analyzed are stably expressed in *E. coli*. The smaller band with HSA-binding activity, seen with different intensities most probably corresponds to the ABP-tail itself (5 kDa), resulting from proteolytic cleavage between the Z variant and the ABP tail. Interestingly, both Z-variants (no. 16 and 22) with introduced cysteine residues

formed dimers, which could be observed under non-reducing conditions during SDS-PAGE (Figure 13; lanes 6 and 7).

To investigate if the secondary structure content of the derivatives was preserved after the extensive surface mutagenesis, a subtractive circular dichroism analysis was performed (see example 3). A comparison of signals obtained from 250 to 184 nm for the wild type Z domain and the four variants fused to the ABP-tail was performed after subtraction of the contribution from the ABP tail itself. The result showed that for three of the four derivatives spectra similar to the wild type Z domain were obtained, with a characteristic minimum at 208 nm and an inflexion point at 222 nm (Johnson, (1990) *Prot. Struct. Funct.. Genet.*, 7, 205-224) (Figure 17). This suggests that the three helix bundle framework probably is preserved in these mutants. However, for the fourth derivative (no. 24), a spectrum was obtained which resembles spectra seen for random coils, indicating a low content of secondary structure elements (Johnson, 1990). This derivative contains a glutamine to proline substitution at position 32 in helix 2, suggesting a destabilization leading to a collapse of the helix bundle framework.

In order to further investigate the four Z-variants, the interaction with polyclonal human IgG (hIgG) (Pharmacia AB) for wild type Z and four different Z variant clones (no. 16, 21, 22, 24; Figure 14) fused to the ABP tail were compared using biosensor technology (BIAcore™, Pharmacia Biosensor AB, Sweden). The carboxylated dextran layer of a CM-5 sensor chip was activated using *N*-hydroxysuccinimide (NHS) and *N*-ethyl-*N*'-[3-diethylaminopropyl]-carbodiimide (EDC) chemistry according to the manufacturers' recommendations. For immobilization of hIgG, 20 µl of a 500 nM hIgG solution in 50 mM acetate, pH 4 was injected at a flow rate of 5 µl/min over the activated surface, resulting in the immobilization of approximately 5000 resonance units (RU). Fortyfive-microlitre samples of the five fusion proteins, dissolved to approximate concentrations of 1500 nM in NaCl/Hepes (10 mM Hepes, pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.5% surfactant P-20), were injected in separate experiments at a flow rate of 2 µl/min. After each sample injection, the hIgG surface was regenerated with 20 mM HCl. As expected, only the wild type Z-domain showed any detectable Fc-binding activity (Figure 18).

In conclusion, the results show that a library of SPA variants with a substituted surface made up from 13 residues located in the α-helices can be constructed. The high degree of conservation of the overall framework of the native Z-domain suggests that derivatives with novel functions grafted onto a stable and soluble scaffold could be isolated for use as artificial antibodies in biochemistry, immunology and biotechnology.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A modified binding protein derived from a native (wild-type) bacterial receptor having an ααα helical domain structure, wherein said modified binding protein derived from said bacterial receptor comprises the B domain of staphylococcal protein A, and wherein said modified binding protein is produced by mutagenesis of amino acid residues 9, 11, 14, 27 and 35 of said B domain structure, which results in a modified binding protein which specifically binds to a different binding partner than the unmodified native (wild-type) bacterial receptor, and which substantially retains the basic structure and stability of the native (wild-type) bacterial receptor.

2. The modified binding protein of claim 1, wherein said modifications result in a glutamic acid residue at position 9, an aspartic acid at position 11, an aspartic acid at position 14, an alanine at position 27, and either a glutamic acid or alanine at position 35 of said B domain.

3. A modified binding protein derived from a native (wild-type) bacterial receptor having an ααα 3-helix domain structure, wherein said modified binding protein derived from a bacterial receptor comprises the B domain of staphylococcal protein A, and wherein said modified binding protein is produced by mutagenesis of at least two of residues 9, 10, 11, 13, 14, 17, 28, 32, 35, 18, 24, 25 and 27 in the Z Scaffold of said B domain structure, and wherein said modifications result in a modified binding protein that specifically binds to a different binding partner than the unmodified native (wild-type) bacterial receptor, and which substantially retains the basic structure and stability of the native (wild-type) bacterial receptor.

4. A modified binding protein according to any one of the preceding claims substantially as hereinbefore described with reference to any of the figures and/or examples.

DATED: 22 July 1998

PHARMACIA & UPJOHN AB

By: PHILLIPS ORMONDE & FITZPATRICK

Patent Attorneys per:

AUSTRALIAN PATENT OFFICE

S	E	D	A	B	C	X	M

Fig. 1A

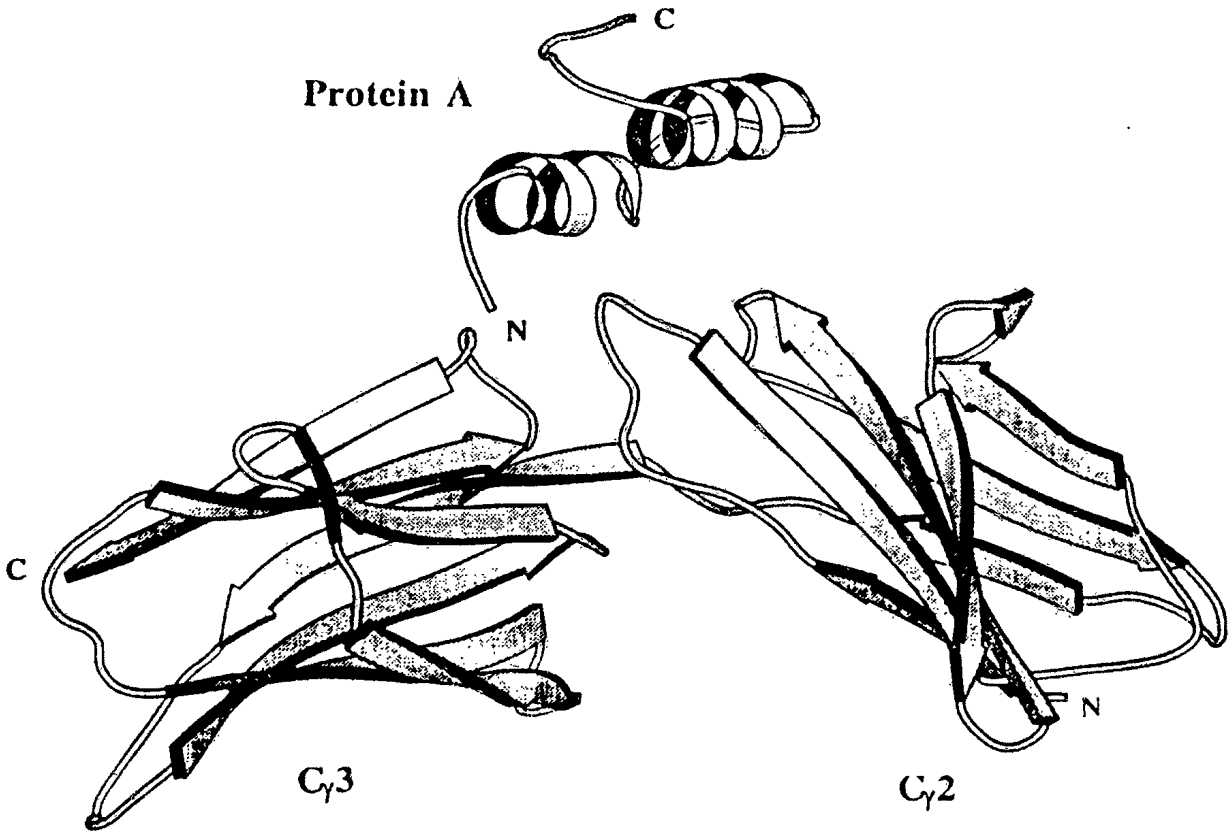


Fig. 1B

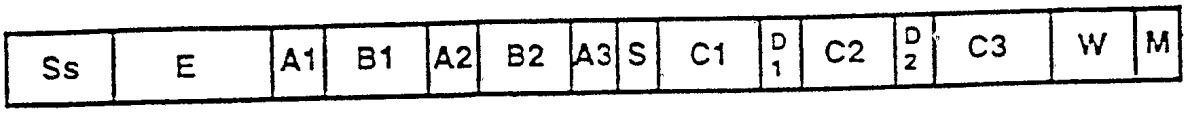

Ss
E
A1
B1
A2
B2
A3
S
C1
D 1
C2
D 2
C3
W
M

Fig. 2A

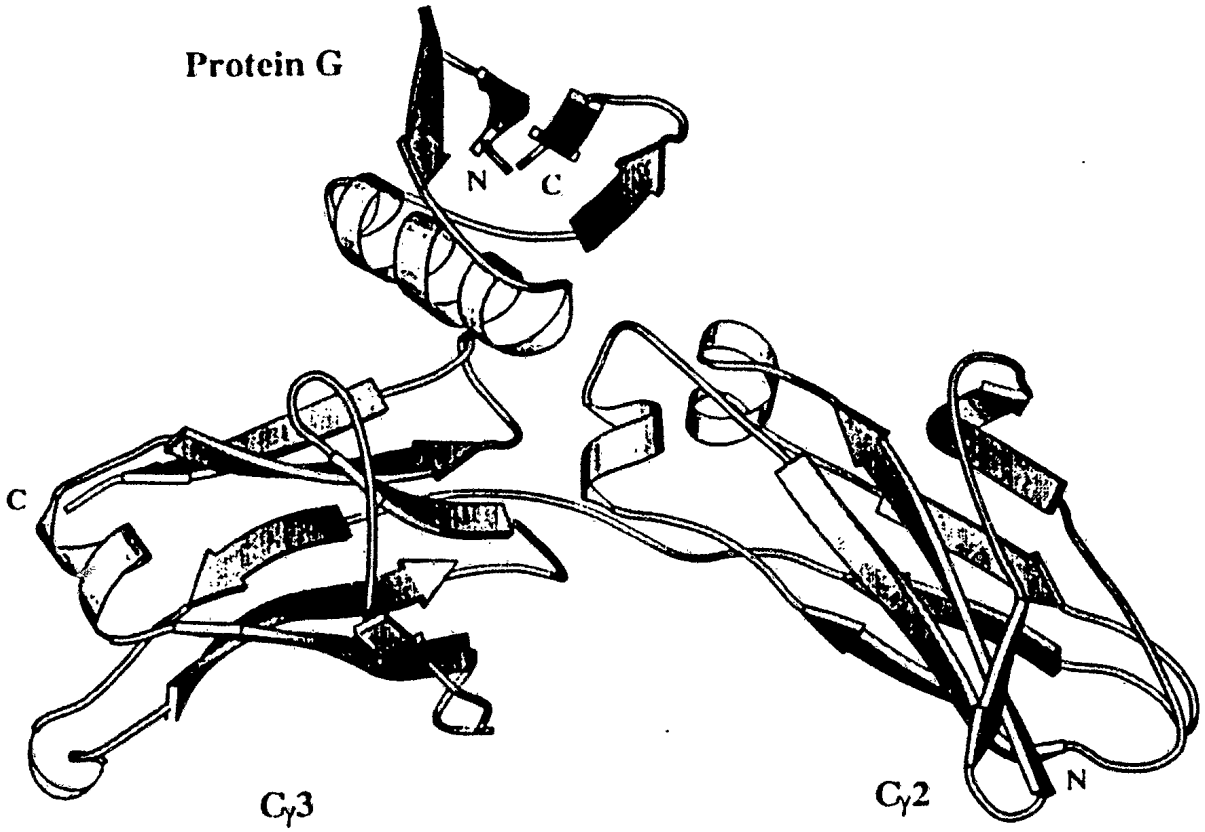

Protein G
N
C
C
$C_\gamma 3$
$C_\gamma 2$
N

Fig. 2B

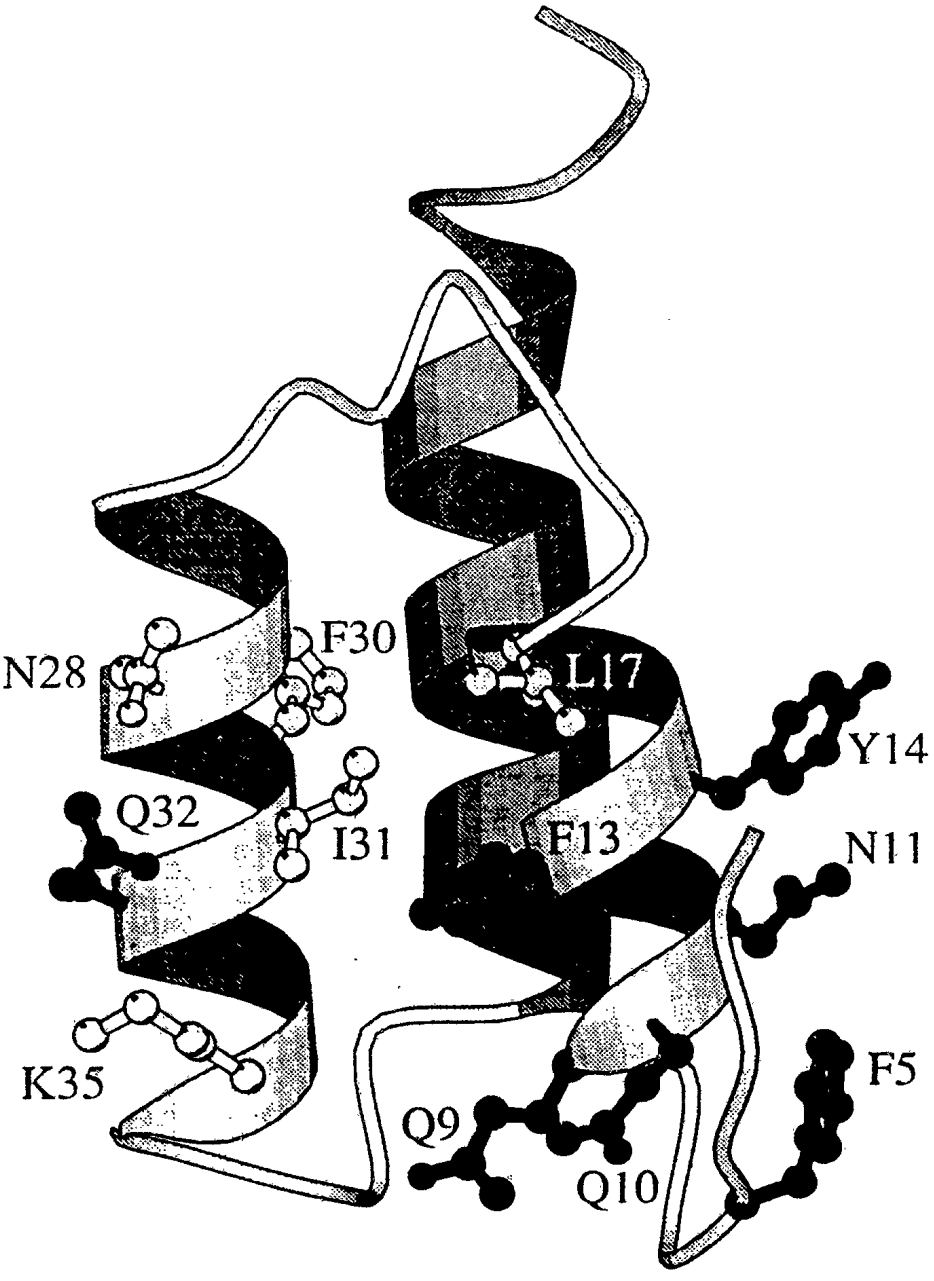


Fig. 3

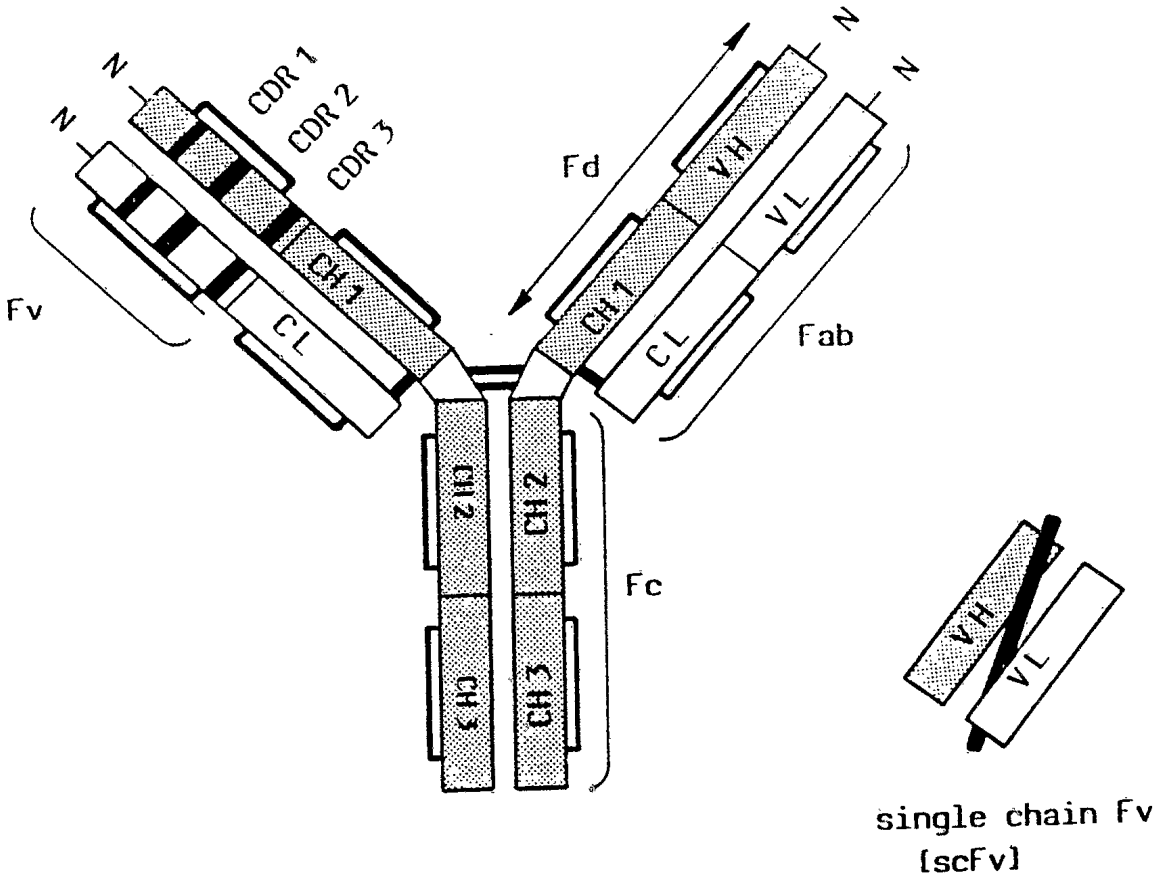

N
N
N
CDR 1
CDR 2
CDR 3
Fd
VH
VL
CH 1
CL
Fv
CH 1
CL
Fab
CH 2
CH 2
Fc
CH 3
CH 3
VH
VL
single chain Fv
(scFv)

Fig. 4

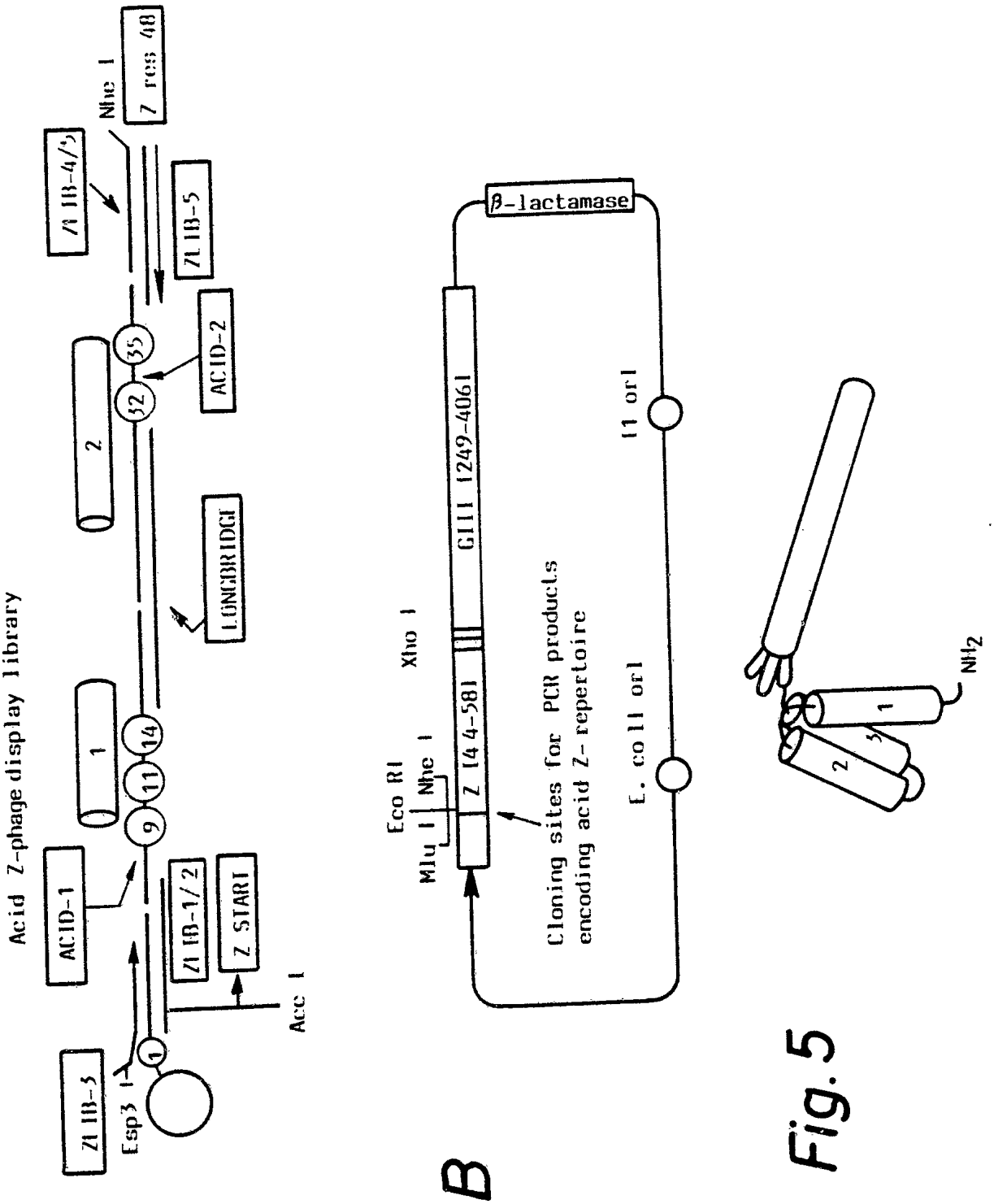

A
Acid Z-phage display library
ZHB-3
Esp3 I
ACID-1
ZHB-1/2
Z START
Acc I
1
9
11
14
1
2
LONGBRIDGE
12
35
ACID-2
ZHB-5
ZHB-4/3
Nhe I
Z res 48
B
β-lactamase
G111 1249-4061
Z 144-581
Eco RI
Nhe I
Mlu I
Xho I
Cloning sites for PCR products
encoding acid Z-repertoire
E. coli ori
11 ori
1
2
3
NH2

Fig.5

SUBSTITUTE SHEET

ZLIB-1	Biotin - 5' - GCG CAA CAC GAT GAA GCC GTA GAC AAC AAA TTC AA - 3'
ZLIB-2	5' - TTC TT GTT GAA TTT GTT GTC TAC GGC TTC ATC GTG TTG CGC - 3'
ZLIB-3	5' - CG CGC GCG TCT CAC GCG GCG CAA CAC GAT GAA GCC GTA - 3'
ZLIB-4	5' - A AGC CAA AGC GCT AAC TTG CTA GCA GGG - 3'
ZLIB-5	5' - CCC CCC TGC TAG CAA GTT AGC GCT TTG GCT TGG GTC ATC - 3'
ZLIB-6	5' - C GCG TGA ATT CTG CTA GCA GAA GCT AAA AAG CTA AAT GAT CGT CAG GCG CCG AAA AGC - 3'
ZLIB-7	5' - T CGA GCT TTT CGG CGC CTG AGC ATC ATT TAG CTT TTT AGC TTC TGC TAG CAG AAT TCA - 3'
LONG BRIDGE	5' - TTG TTC TTC GTT TAA GTT AGG TAA ATG TAA GAT CTC - 3'
ACID-1	5' - C AAA GAA G^{CA}_{AC} CAA G^{CA}_{AC} GCG TTC G^{CA}_{AC} GAG ATC TTA CAT TTA CCT A - 3'
ACID-2	5' - AC TTA AAC GAA GAA CAA G^{CA}_{AC} AAC GCC TTC ATC CAA AGT TTA G^{CA}_{AC} GAT GAC CC - 3'
DEGEN-2	5' - AC TTA AAC NN^{G}_{T} NN^{G}_{T} CAA NN^{G}_{T} NN^{G}_{T} GCC TTC ATC NN^{G}_{T} AGT TTA NN^{G}_{T} GAT GAC CC - 3'
DEGEN-1	5' - C AAA GAA NN^{G}_{T} NN^{G}_{T} NN^{G}_{T} GCG NN^{G}_{T} NN^{G}_{T} GAG ATC NN^{G}_{T} NN^{G}_{T} TTA CCT A - 3'
BRIDGE	5' - GTT TAA GTT AGG TAA - 3'

Fig. 6

```
1                                   9       11          14
GTA GAC AAC AAA TTC AAC AAA GAA GAC CAA GCA GCG TTC GAC NAG ATC TTA CAT TTA CCT AAC TTA AAC
GTA GAC AAC AAA TTC AAC AAA GAA GAC CAA GAC GCG TTC GAC GAG ATC TTA CAT TTA CCT AAC TTA AAC
Acc I

            27                                  35
GAA GAA CAA GAC AAC GCC TTC ATC CAA AGT TTA GCA GAT GAC CCA AGC CAA AGC GCT AAC TTG CTA GC
GAA GAA CAA GAA AAC GCC TTC ATC CAA AGT TTA GCA GAT GAC CCA AGC CAA AGC GCT AAC TTG CTA GC
                                                                                Nhe I
```

Fig. 7

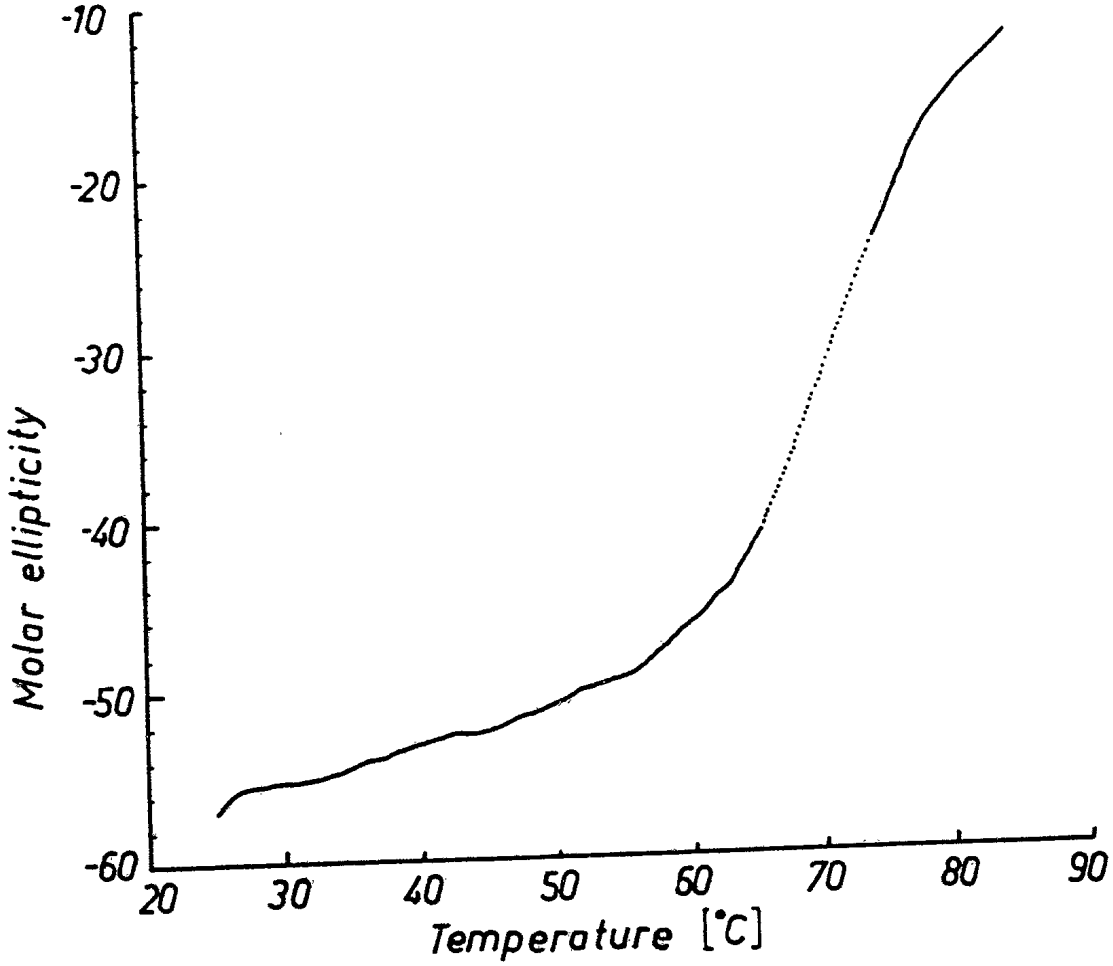

-10
-20
-30
-40
-50
-60
Molar ellipticity
20
30
40
50
60
70
80
90
Temperature [°C]

Fig. 8

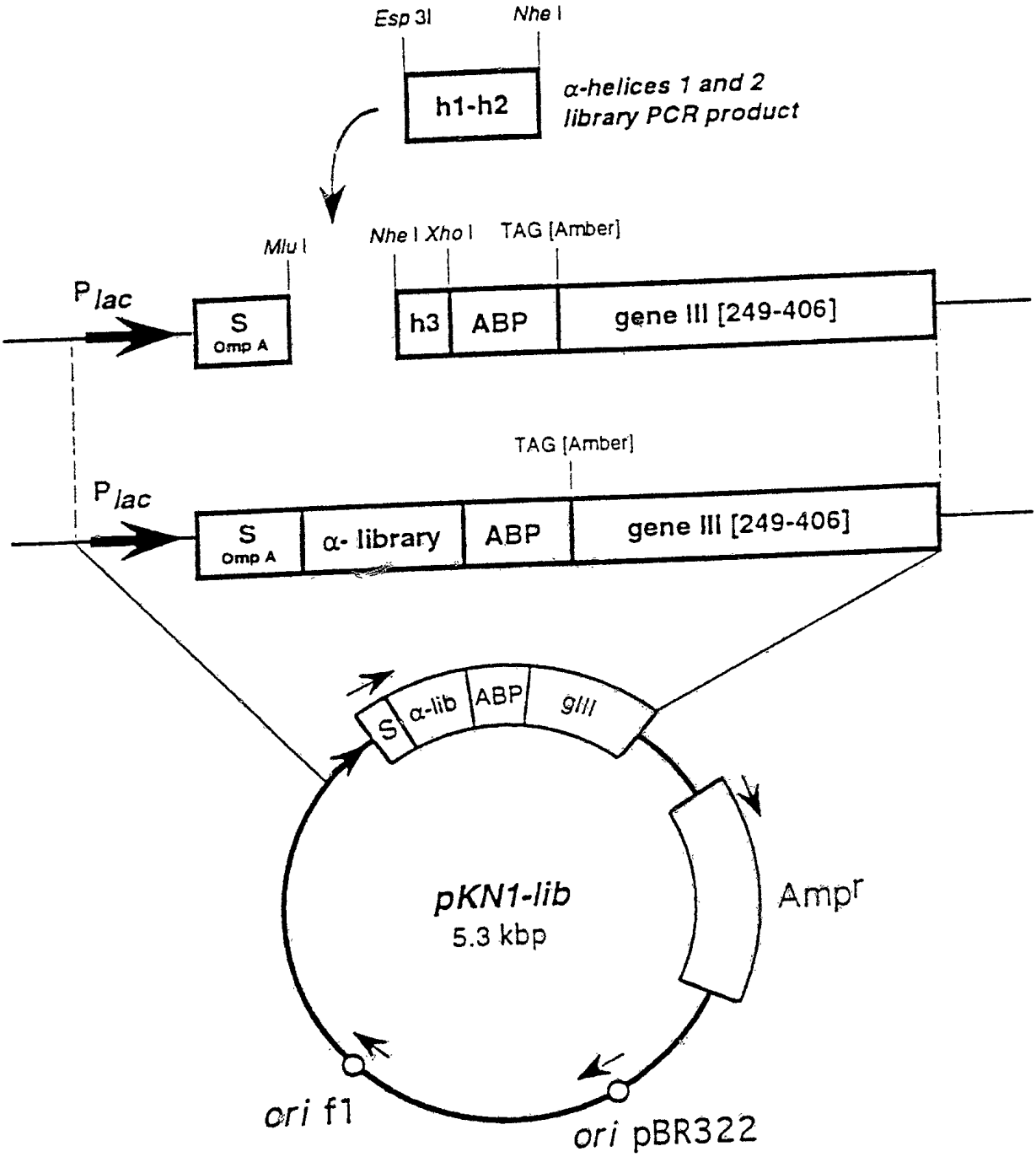

Esp 3I
Nhe I
h1-h2
α-helices 1 and 2
library PCR product
Mlu I
Nhe I Xho I
TAG [Amber]
P lac
S
Omp A
h3
ABP
gene III [249-406]
TAG [Amber]
P lac
S
Omp A
α- library
ABP
gene III [249-406]
S
α-lib
ABP
gIII
pKN1-lib
5.3 kbp
Ampr
ori f1
ori pBR322

Fig. 9

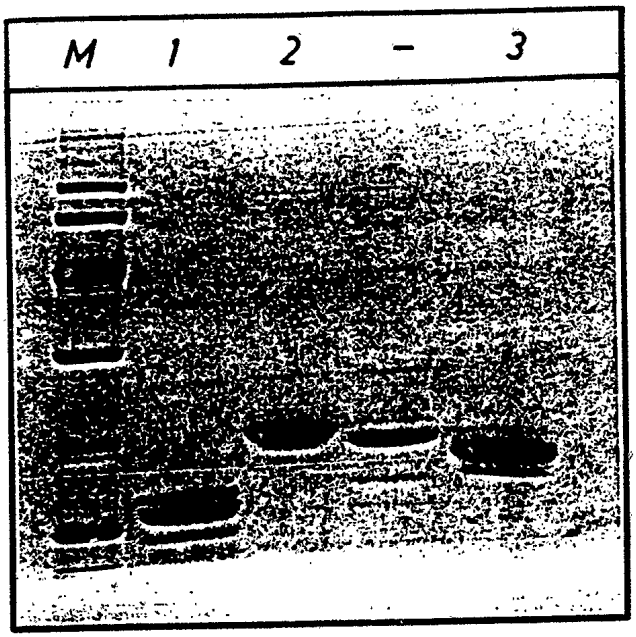


Fig. 10

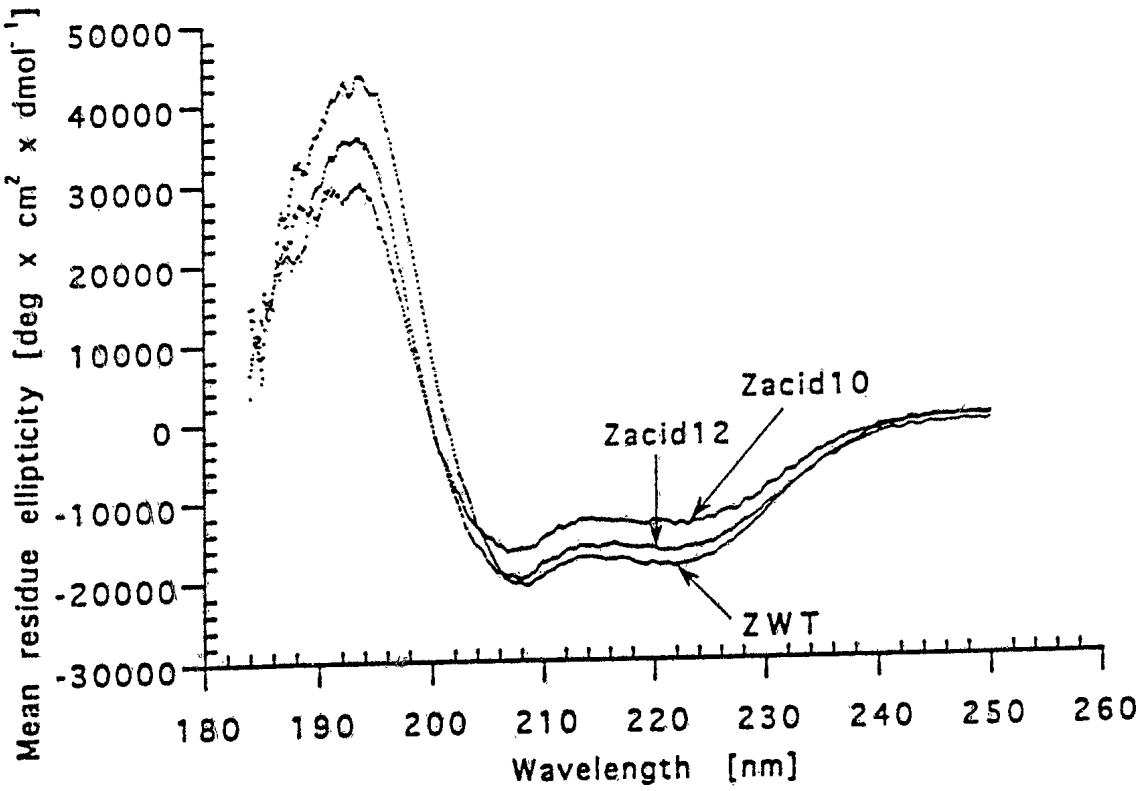

Mean residue ellipticity [deg x cm^2 x dmol^{-1}]
50000
40000
30000
20000
10000
0
-10000
-20000
-30000
180
190
200
210
220
230
240
250
260
Wavelength [nm]
Zacid10
Zacid12
ZWT

Fig. 11

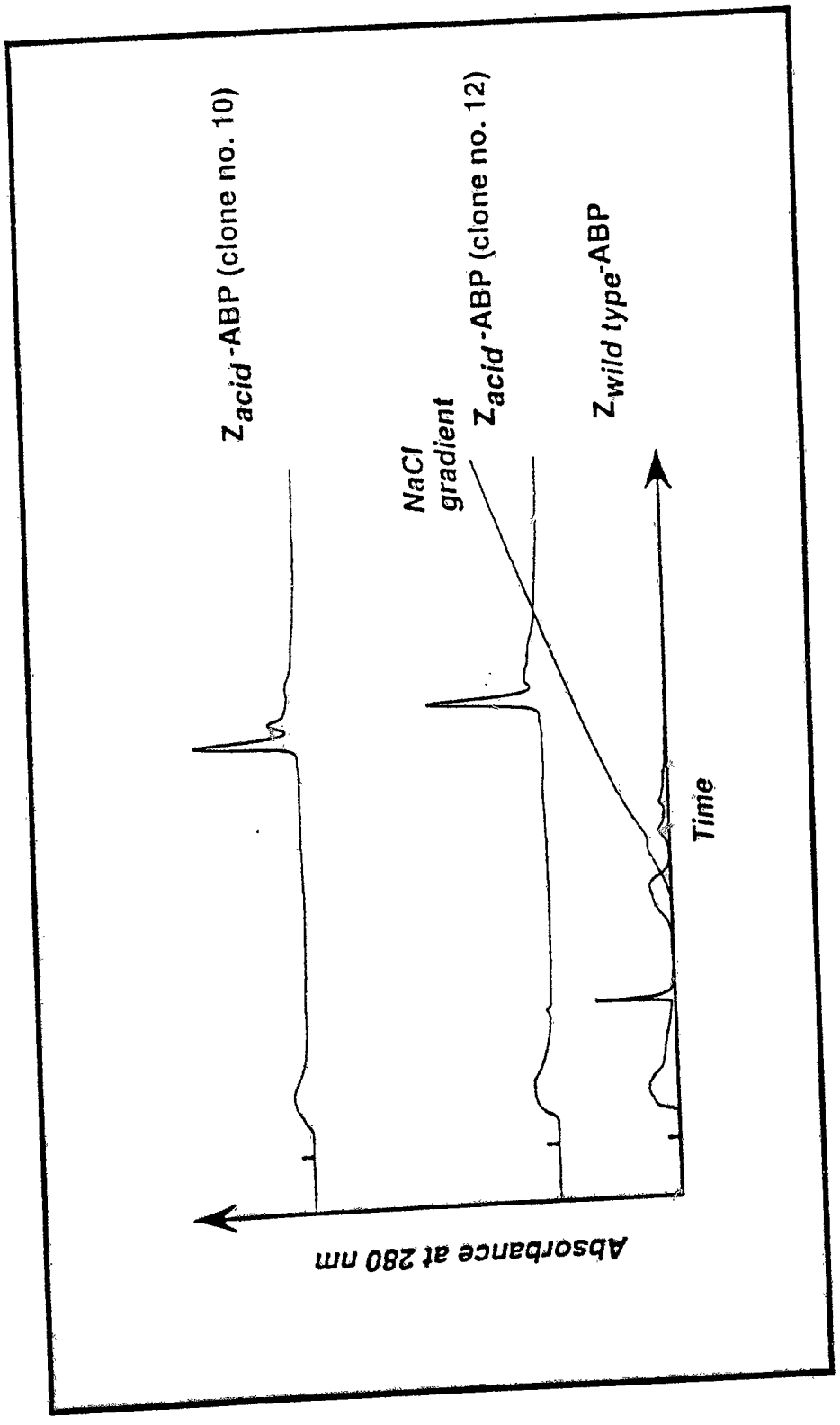

Z_{acid}-ABP (clone no. 10)
NaCl gradient
Z_{acid}-ABP (clone no. 12)
$Z_{wild\ type}$-ABP
Time
Absorbance at 280 nm

Fig. 12

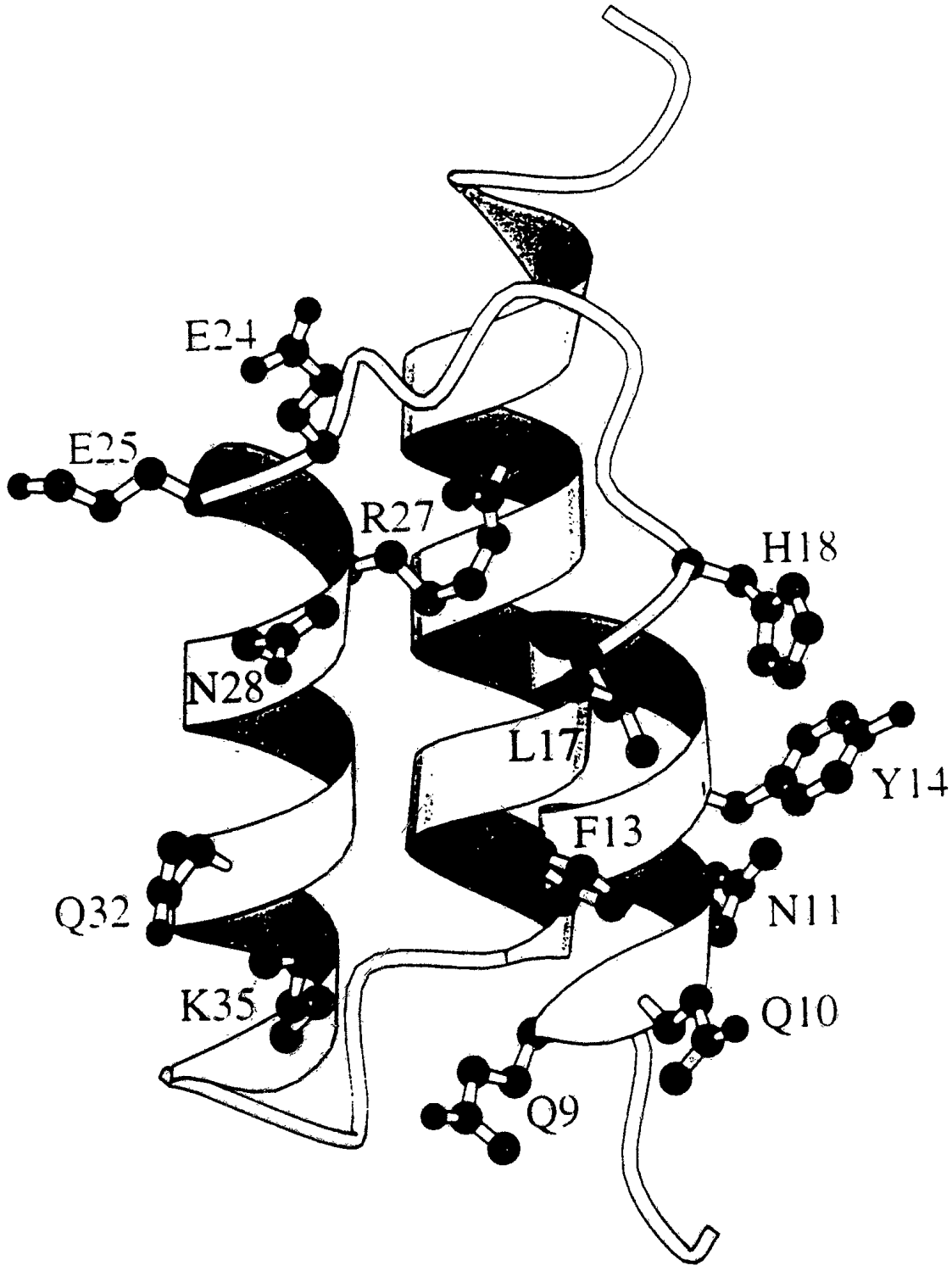
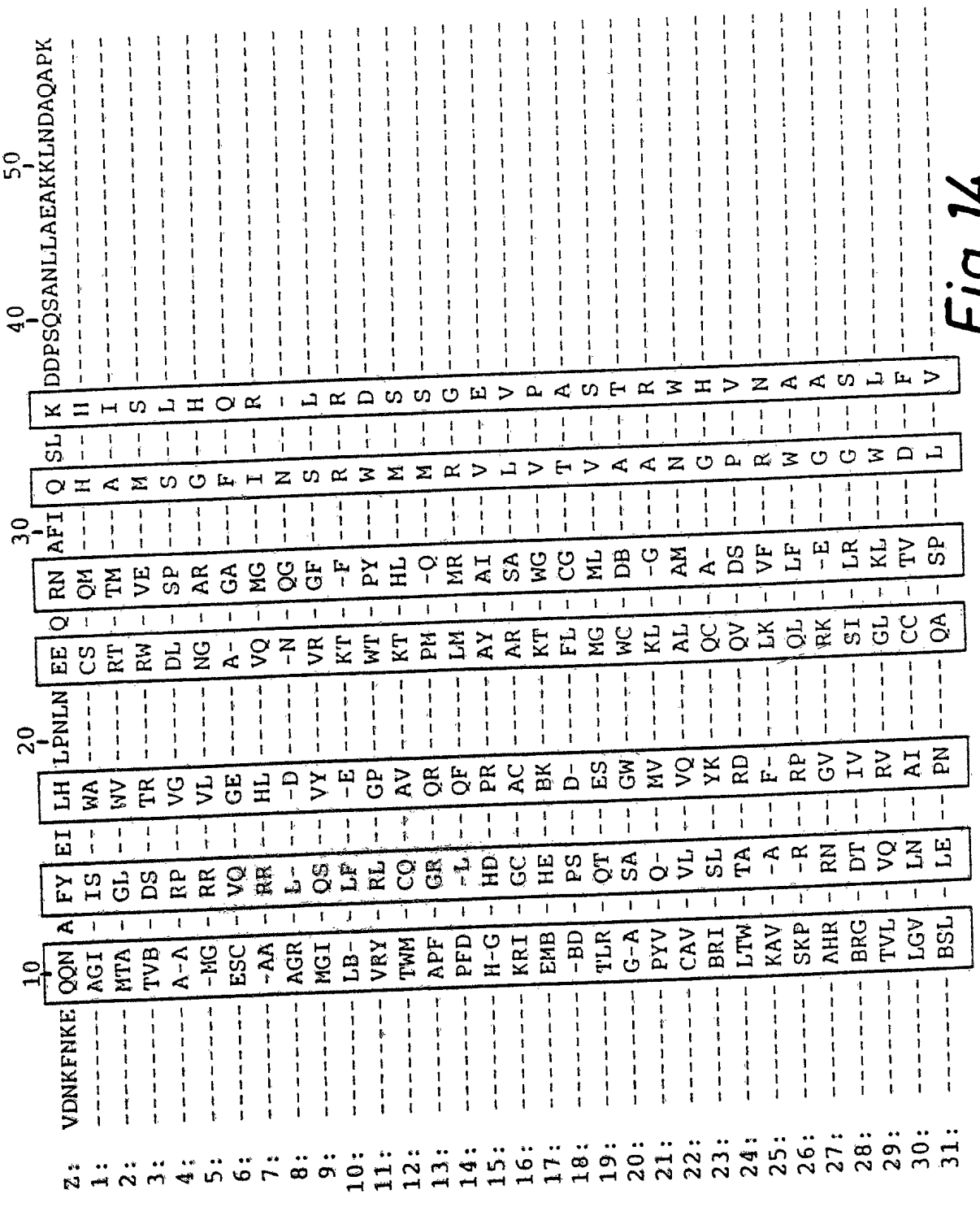


Fig. 13



		10					20				30				40		50	
Z:	VDNKFNKE	QQN	A	FY	EI	LH	LPNLN	EE	Q	RN	AFI	Q	SL	K	DDPSQSANLLAEAKKLNDAQAPK			
1:	--------	AGI	-	IS	--	WA	-----	CS	-	QM	---	H	--	H	----------			
2:	--------	MTA	-	GL	--	WV	-----	RT	-	TM	---	A	--	I	----------			
3:	--------	TVB	-	DS	--	TR	-----	RW	-	VE	---	M	--	S	----------			
4:	--------	A-A	-	RP	--	VG	-----	DL	-	SP	---	S	--	L	----------			
5:	--------	-MG	-	RR	--	VL	-----	NG	-	AR	---	G	--	H	----------			
6:	--------	ESC	-	VQ	--	GE	-----	A-	-	GA	---	F	--	Q	----------			
7:	--------	-AA	-	RR	--	HL	-----	VQ	-	MG	---	I	--	R	----------			
8:	--------	AGR	-	L-	--	-D	-----	-N	-	QG	---	N	--	-	----------			
9:	--------	MGI	-	QS	--	VY	-----	VR	-	GF	---	S	--	L	----------			
10:	--------	LB-	-	LF	--	-E	-----	KT	-	-F	---	R	--	R	----------			
11:	--------	VRY	-	RL	--	GP	-----	WT	-	PY	---	W	--	D	----------			
12:	--------	TWM	-	CQ	--	AV	-----	KT	-	HL	---	M	--	S	----------			
13:	--------	APF	-	GR	--	QR	-----	PM	-	-Q	---	M	--	S	----------			
14:	--------	PFD	-	-L	--	QF	-----	LM	-	MR	---	R	--	G	----------			
15:	--------	H-G	-	HD	--	PR	-----	AY	-	AI	---	V	--	E	----------			
16:	--------	KRI	-	GC	--	AC	-----	AR	-	SA	---	L	--	V	----------			
17:	--------	EMB	-	HE	--	BK	-----	KT	-	WG	---	V	--	P	----------			
18:	--------	-BD	-	PS	--	D-	-----	FL	-	CG	---	T	--	A	----------			
19:	--------	TLR	-	QT	--	ES	-----	MG	-	ML	---	V	--	S	----------			
20:	--------	G-A	-	SA	--	GW	-----	WC	-	DB	---	A	--	T	----------			
21:	--------	PYV	-	Q-	--	MV	-----	KL	-	-G	---	A	--	R	----------			
22:	--------	CAV	-	VL	--	VQ	-----	AL	-	AM	---	N	--	W	----------			
23:	--------	BRI	-	SL	--	YK	-----	QC	-	A-	---	G	--	H	----------			
24:	--------	LTW	-	TA	--	RD	-----	QV	-	DS	---	P	--	V	----------			
25:	--------	KAV	-	-A	--	F-	-----	LK	-	VF	---	R	--	N	----------			
26:	--------	SKP	-	-R	--	RP	-----	QL	-	LF	---	W	--	A	----------			
27:	--------	AHR	-	RN	--	GV	-----	RK	-	-E	---	G	--	A	----------			
28:	--------	BRG	-	DT	--	IV	-----	SI	-	LR	---	G	--	S	----------			
29:	--------	TVL	-	VQ	--	RV	-----	GL	-	KL	---	W	--	L	----------			
30:	--------	LGV	-	LN	--	AI	-----	CC	-	TV	---	D	--	F	----------			
31:	--------	BSL	-	LE	--	PN	-----	QA	-	SP	---	L	--	V	----------			

Fig. 14

Fig. 15

SUBSTITUTE SHEET

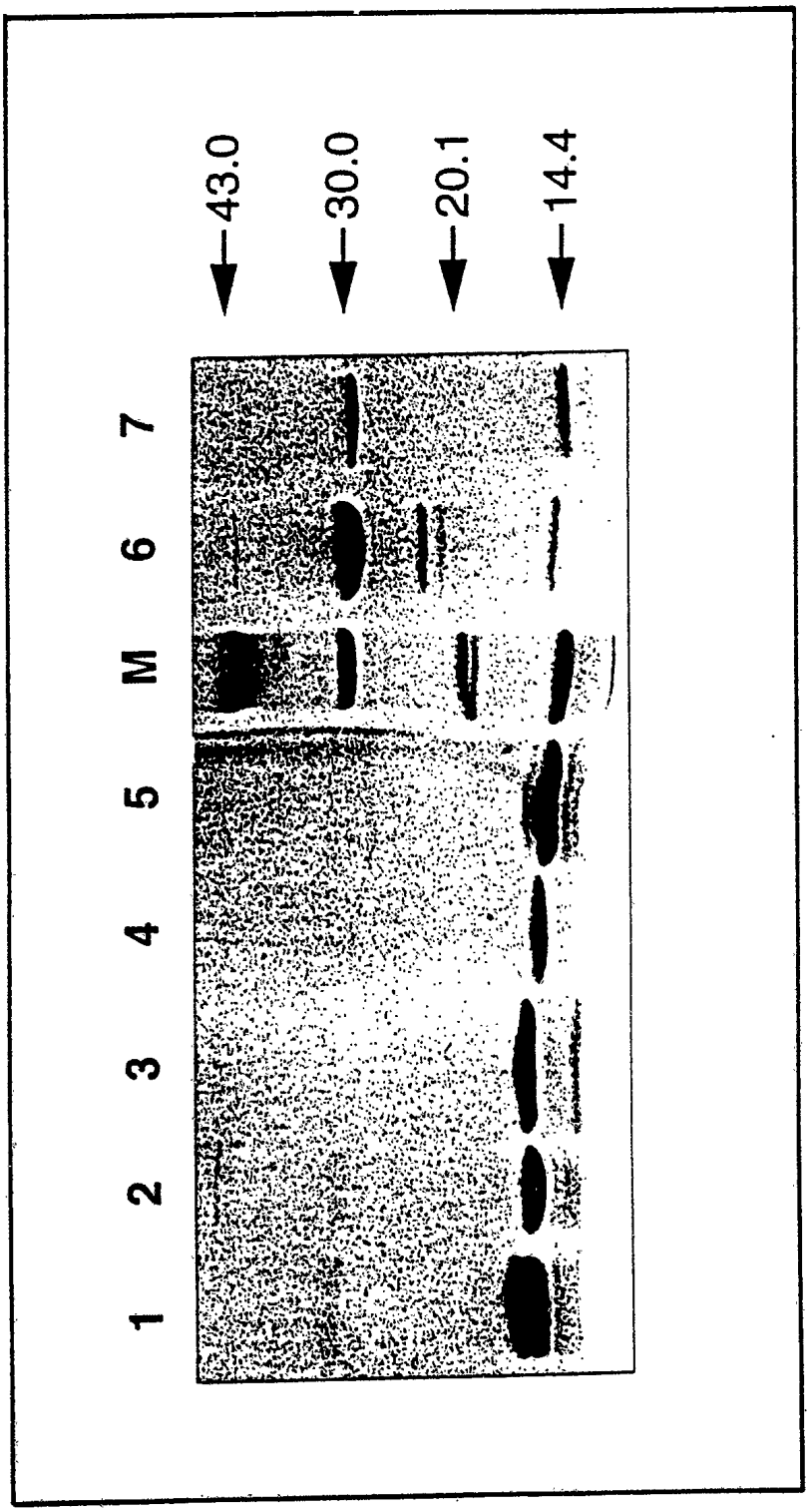


Fig. 16

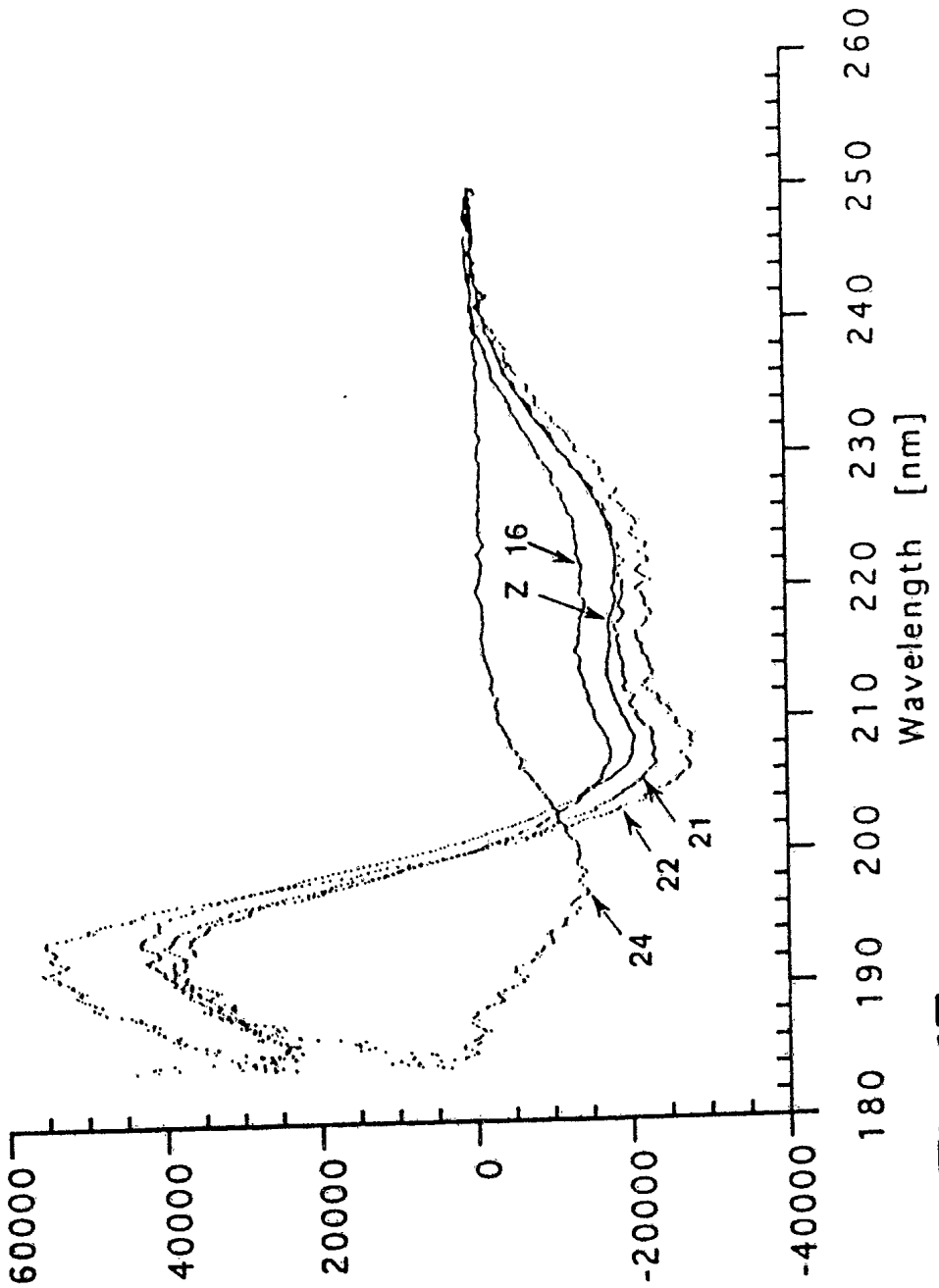

60000
40000
20000
0
-20000
-40000
180
190
200
210
220
230
240
250
260
Wavelength [nm]
Z 16
21
22
24
Mean residue ellipticity [deg x cm^2 x dmol^{-1}]

Fig. 17

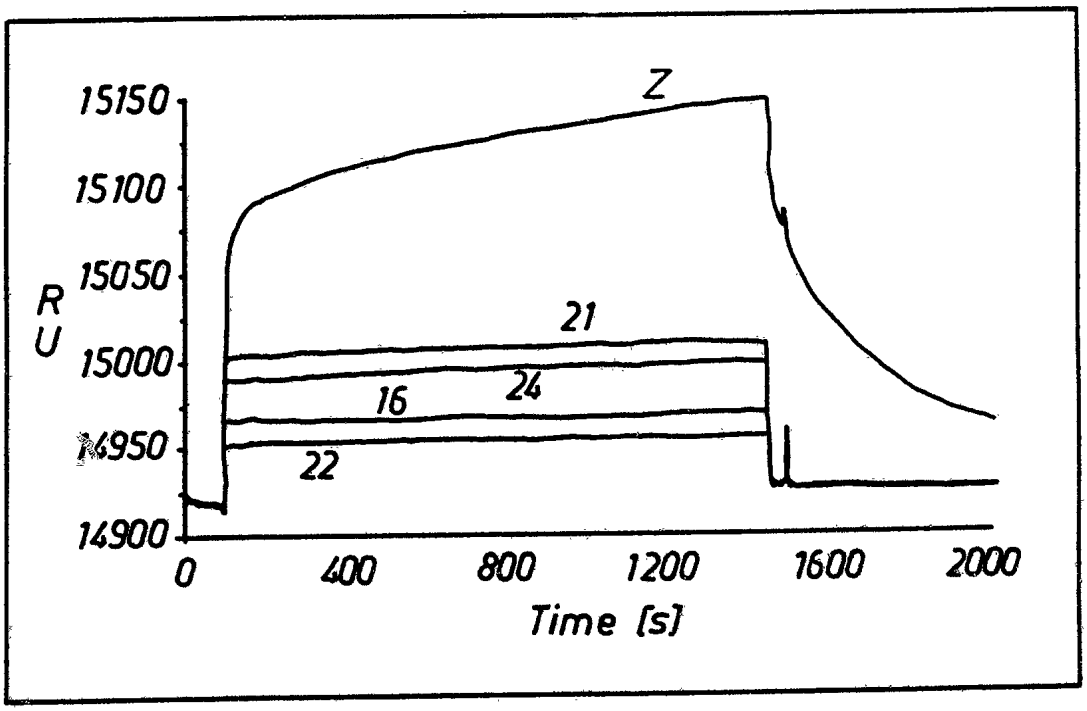

15150
15100
15050
15000
14950
14900
RU
Z
21
16
24
22
0
400
800
1200
1600
2000
Time [s]

Fig. 18

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 95/00034

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C07K 14/705, C07K 14/735, C12N 15/00
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C07K

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

SE,DK,FI,NO classes as above

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

WPIDS, CLAIMS, CA, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US, A, 5084559 (ALBERT T. PROFY), 28 January 1992 (28.01.92), see column 10 --	1-7,9-10, 12-19
X	US, A, 4954618 (STEPHEN R. FAHNESTOCK), 4 Sept 1990 (04.09.90), column 7 - column 8 --	1-7,9-10, 12-22
X	Protein Engineering, Volume 6, No 4, 1993, Lena Cedergren et al, "Mutational analysis of the interaction between staphylococcal protein A and human IgG1" page 441 - page 448 --	1-7,9-10, 12-19

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" erlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 April 1995	03 -05- 1995
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Carolina Gómez Lagerlöf Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 95/00034

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	The Journal of Biological Chemistry, Volume 263, No 9, March 1988, Margareta Eliasson et al, "Chimeric IgG-binding Receptors Engineered from Staphylococcal Protein A and Streptococcal Protein G" page 4323 - page 4327 -- --------	1-22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 95/00034

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 1-3
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

The formulation of claim 1 is so vague and broad in scope, that the claim and consequently claims 2-3 do not define the matter for which protection is sought. The claims are not clear and concise. (See Art 6)

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.