Non-repetitive single-chain Fv linkers selected by selectively infective phage (SIP) technology

**Frank Hennecke, Claus Krebber[1] and
Andreas Plückthun[2]**

Biochemisches Institut der Universität Zürich, Winterthurerstrasse 190,
CH-8057 Zürich, Switzerland

[1]Present address: Maxygen, 3410 Central Expressway, Santa Clara,
CA 95051, USA

[2]To whom correspondence should be addressed

By using the selectively infective phage (SIP) technology, we selected non-repetitive linkers for a single-chain Fv fragment to have genes more robust against deletions in PCR-based gene assembly and directed evolution experiments than is the case for the classical $(Gly_4Ser)_3$ linker. We designed linkers encoding turns at both ends and random positions in the middle where glycines and polar and charged residues were allowed to occur. After only a single round of SIP, all clones obtained were fully functional. Properties such as antigen binding constants, urea denaturation curves and expression of soluble scFv fragments were identical with those of the parental fragment with the $(Gly_4Ser)_3$ linker. This demonstrates that SIP is a very fast and powerful technique to remove rapidly sequences of poor functionality, exclusively yielding sequences of the desired overall property in a single round.

Keywords: filamentous phage/fluorescein/library/scFv fragment/selection

Introduction

Immunoglobulin fragments constitute a very popular subject for protein engineering. This is due to a large body of knowledge that has accumulated owing to extensive analyses of structural (Bork *et al.*, 1994; Padlan, 1994) and physico-chemical properties (Plückthun, 1992) as well as progress in accessing antibody genes from various libraries (Winter *et al.*, 1994). The enormous potential for diagnostic and therapeutic applications (George *et al.*, 1994, 1996) keeps driving the technology forward.

The binding site of an antibody is made up from the V_H and V_L domains, which associate only poorly to form an Fv fragment, however. Thus, in general, they have to be linked to ensure a stable V_H–V_L complex in solution (Glockshuber *et al.*, 1990). This is commonly achieved by introducing a peptidic linker (Bird *et al.*, 1988; Huston *et al.*, 1988) and the sequence $(Gly_4Ser)_3$ connecting the domains has been most popular (Huston *et al.*, 1991). The repetitive sequence of this linker, however, frequently causes problems during PCR-based assembly of amplified V_H and V_L domains (Winter *et al.*, 1994) or during the introduction of random mutations by mutagenic DNA shuffling (Stemmer, 1994). In these PCR-based technologies, the repetitive character of the scFv linker frequently leads to irregular annealing of homologous sequences and hence to undesired length variants.

To make scFv fragments more amenable to PCR-based gene

assembly and mutagenic DNA shuffling, we designed and selected non-repetitive linker sequences suited to connect covalently the domains of an Fv fragment without interfering with the desired properties, such as proper folding, solubility and ligand binding. Using the fluorescein binding scFv fragment FITC-E2 (Vaughan *et al.*, 1996) as a model system, we constructed an scFv library with several positions of the linker region fixed to encode particular structural features and with others randomized, but restricted to particular amino acid types.

We used the selectively infective phage (SIP) technology (Spada and Plückthun, 1997; Spada *et al.*, 1997) for identifying the most desirable sequences. In SIP, the scFv fragment is displayed fused to g3p, in all copies of the protein. In the present experiments, the phage is missing the N-terminal domain N1 and hence completely non-infective. Upon being provided with the soluble N1 domains, coupled to the antigen fluorescein, it becomes infective again (Figure 1). Already after one round of SIP, all of the selected clones bound to fluorescein at least as efficiently as the original FITC-E2. Three of the selected clones were characterized in more detail and turned out to be fully comparable to the original scFv fragment containing the $(Gly_4Ser)_3$ linker in terms of binding, folding, expression and solubility.

Materials and methods

Construction of linker library and SIP

The region coding for the randomized scFv linker was assembled by PCR from oligonucleotides LiLib (5′-CT-GGGATCCMGRAGCAGAACYAGTNTBMGRAGCAGAA-CYACYNTBGCTCGCGCCGTTAGG-3′) and LiFill (5′-CT-GGTCACCGTCTCGAGTCCTAACGGCGCGAGC-3′) (M = A or C; R = A or G; Y = C or T; N = A, T, C or G; B = C, G or T). With the resulting cassette, the $(Gly_4Ser)_3$ scFv linker of the scFlu-'medium' phage (Krebber *et al.*, 1997b) was replaced in several cloning steps, resulting in the randomized linker shown in Figure 2A. All constructions were performed using standard methods of DNA manipulation (Sambrook *et al.*, 1989). With the resulting scFlu-'medium' phage vectors, *Escherichia coli* XL1-Blue cells (Stratagene) were electroporated and grown overnight at 37°C on 2YTG plates (Krebber *et al.*, 1997b) containing 25 µg/ml chloramphenicol (cam), yielding 4×10^6 colonies, which were scraped
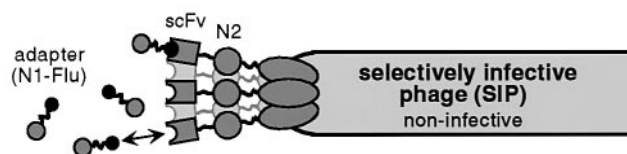

Fig. 1. Schematic representation of the SIP experiment. The N1 domain of the phage gene 3 protein was replaced by the scFv FITC-E2 containing the randomized linker region. The resulting scFv-displaying phages are non-infective owing to the lack of the N1 domain. Infectivity is restored by binding of adapter molecules (N1-Flu) consisting of fluorescein (black circle) covalently coupled to the missing N1 domain.

A

| V$_H$ | | | | | | LINKER | | | | | | | | | | | | | V$_L$ |

```
         V_H                                               LINKER                                                          V_L

     AGT CCT AAC GGC GCG AGC VAN RGT RGT TCT GCT YCK VAN ACT RGT TCT GCT YCK GGA TCC CAG
      S   P   N   G   A   S  E(2)S(8)S(6) S   A  S(6)E(0) T S(9) S   A  S(7) G   S   Q
                                D(1)G(5)G(7)        P(8)D(0)   G(4)        P(7)
                                K(4)                    K(5)
                                N(3)                    N(3)
                                Q(2)                    Q(3)
                                H(1)                    H(2)
```

B

LLA1
```
AGT CCT AAC GGC GCG AGC CAA AGT AGT TCT GCT TCG CAC ACT GGT TCT GCT CCG GGA TCC CAG
 S   P   N   G   A   S   Q   S   S   S   A   S   H   T   G   S   A   P   G   S   Q
```

LLA3
```
AGT CCT AAC GGC GCG AGC AAT AGT GGT TCT GCT CCT GAC ACT AGT TCT GCT CCG GGA TCC CAG
 S   P   N   G   A   S   N   S   G   S   A   P   D   T   S   S   A   P   G   S   Q
```

LLA4
```
AGT CCT AAC AGC GCG AGC CAC AGT GGT TCT GCT CCG AAT ACT AGT TCT GCT CCG GGA TCC CAG
 S   P   N   S   A   S   H   S   G   S   A   P   N   T   S   S   A   P   G   S   Q
```

LLB13
```
AGT CCT AAC GGC GCG AGC GAA AGT GGT TCT GCT TCG AAG ACT AGT TCT GCT TCT GGA TCC CAG
 S   P   N   G   A   S   E   S   G   S   A   S   K   T   S   S   A   S   G   S   Q
```

LLB14
```
AGT CCT AAC GGC GCG AGC AAC AGT GGT TCT GCT CCT AAA ACT GGT TCT GCT TCG GGA TCC CAG
 S   P   N   G   A   S   N   S   G   S   A   P   K   T   G   S   A   S   G   S   Q
```

LLB17
```
AGT CCT AAC GGC GCG AGC AAA AGT GGT TCT GCT TCG CAA ACT AGT TCT GCT CCG GGA TCC CAG
 S   P   N   G   A   S   K   S   G   S   A   S   Q   T   S   S   A   P   G   S   Q
```

LLB18
```
AGT CCT AAC GGC GCG AGC CAC AGT AGT TCT GCT TCG CAG ACT GGT TCT GCT TCG GGA TCC CAG
 S   P   N   G   A   S   H   S   S   S   A   S   Q   T   G   S   A   S   G   S   Q
```

LLB24
```
AGT CCT AAC GGC GCG AGC AAA AGG ... ... ... ... ... ... ... ... TCT GCT CCG GGA TCC CAG
 S   P   N   G   A   S   K   R                                   S   A   P   G   S   Q
```

LLB25
```
AGT CCT AAC GGC GCG AGC CAT AGT GGT TCT GCT CCT CAC ACT AGT TCT GCT TCG GGA TCC CAG
 S   P   N   G   A   S   H   S   G   S   A   P   H   T   S   S   A   S   G   S   Q
```

Fig. 2. Sequence of the randomized linker region. (**A**) The sequence of the partially randomized linker region (bold letters) including the last codon of the V$_H$ domain and the first codon of the V$_L$ domain (non-bold letters) is shown. Randomized positions (using the IUPAC nomenclature of mixed bases: V = A, C or G; N = A, C, G or T; R = A or G; Y = C or T; K = G or T) are underlined. Below the nucleotide sequence, the amino acid residues which can be encoded are shown. Numbers in parentheses correspond to the frequencies of the respective residue types found in 14 clones of the unselected library. Where the sum of observed residues is 13 instead of 14, a different residue type was found. (**B**) Nucleotide and deduced amino acid sequences of nine random clones after one round of SIP.

from the plates and resuspended in 50 ml of fresh medium. A 50 ml volume of 2YTG containing 25 µg/ml cam was inoculated with 1 ml of the resulting suspension and incubated overnight at 37°C with shaking. Phage particles were precipitated from the culture supernatant by addition of 0.25 vol. 17% PEG 6000, 3.3 M NaCl, 1 mM EDTA, incubation at 4°C for 30 min and centrifugation at 4°C and 4500 *g* for 30 min. The pellet was resuspended in 1.5 ml of TBS (25 mM Tris–HCl, 140 mM NaCl, 2.7 mM KCl, pH 7.4), passed through a sterile filter (0.22 µm), PEG precipitated again and resuspended in a final volume of 100 µl of TBS. For selection of fluorescein binders, 1 µl of the resulting phage suspension was incubated with 813 µM N1-Flu (prepared as described by Krebber *et al.*, 1997b) in a total volume of 6 µl in TBS. After overnight incubation at 4°C, 1 µl of this mixture was incubated with 100 µl of an exponentially growing XL1-Blue culture at 37°C for 1 h and plated on 2YTG agar containing cam (25 µg/ml) and tetracycline (15 µg/ml).

Phage ELISA

For monovalent display of selected clones, the scFv encoding *Sfi*I fragments were cloned into pAK100 and phage particles were prepared as described (Krebber *et al.*, 1997a). For phage ELISA, a 96-well microtiter plate was coated with fluorescein isothiocyanate-conjugated bovine serum albumin (FITC–BSA) overnight at 4°C and subsequently blocked for 1 h with 4% skim milk in TBS. Phagemid particles (10^8 per well) were incubated for 1 h in 100 µl TBS + 2% skim milk in the presence or absence of soluble fluorescein (10^{-7} M). After incubation with HRP–anti-M13 conjugate (Pharmacia) for 1 h, relative amounts of bound phages were determined colorimetrically with soluble BM Blue POD substrate (Boehringer). For each phagemid clone, three independent measurements were performed in parallel.

Expression and purification of scFv fragments

For soluble expression of the scFv fragments, the corresponding *Sfi*I fragments were cloned into pAK400 (Krebber *et al.*, 1997a) and the resulting vectors were transformed into *E.coli* SB536 (Bass *et al.*, 1996). The cells (1.5 l) were grown at 25°C, induced at an OD$_{550}$ of 0.5, grown for another 3.5 h, centrifuged and resuspended in 15–20 ml of IMAC buffer (20 mM HEPES, 150 mM NaCl, pH 7.0) on ice and passed through a French press three times. The suspension was centrifuged (48 000 *g*, 4°C) and the supernatant was passed through a 0.22 µm sterile filter. The protein was purified on an Ni-NTA column (Qiagen) and eluted with a step gradient at 100 mM imidazole. The scFv containing fractions were

then concentrated, with a buffer change to 20 mM MES, pH 6.0, 50 mM NaCl. The remaining impurities were removed on a Sepharose-SP column in this buffer, using an NaCl step gradient, where the scFv samples eluted at 130–175 mM NaCl. The final protein concentration was determined from the OD_{280} (Pace *et al.*, 1995).

Determination of dissociation constants

The dissociation constant of fluorescein was determined by fluorescence titration as described previously (Pedrazzi *et al.*, 1997).

Reversible, urea-induced unfolding

For determining equilibrium denaturation curves, equal amounts of scFv fragment (final concentration 0.2 μM) and various concentrations of urea were incubated in 1.8 ml of 20 mM HEPES, 150 mM NaCl, pH 6.8, overnight at 10°C. The exact urea concentration in each sample was determined from its refractive index (Pace and Scholtz, 1997). Upon equilibration at 20°C for at least 1 h, unfolding was monitored by fluorescence spectroscopy (λ_{ex} = 280 nm, λ_{em} = 320–365 nm). From the wavelength shift of maximum fluorescence, the fraction of scFv fragment unfolded at each distinct urea concentration was determined according to Pace and Scholtz (1997).

Detection of scFv variants in soluble and insoluble fractions of cell lysate

To compare the soluble expression of the scFv variants, the partitioning into soluble and insoluble fractions of crude *E.coli* lysate was determined by quantitative Western blotting. *E.coli* SB536 (Bass *et al.*, 1996), harboring the pAK400-based secretion vectors (Krebber *et al.*, 1997a), were grown at 25°C in 20 ml of 2YT medium (Sambrook *et al.*, 1989) to an OD_{550} of 0.5, induced by addition of 1 mM IPTG and grown for a further 3.5 h. Upon centrifugation, resuspended cells were disrupted by three passages through a French press and recentrifuged (16 000 *g*, 15 min, 4°C). Upon removal of the supernatant containing the soluble fraction of total cellular protein, insoluble proteins were solubilized from the pellet in the presence of 2% SDS. Aliquots of soluble and insoluble fractions that corresponded to 25 μl of bacterial suspension with an OD_{550} of 32 were subjected to SDS–PAGE and blotted onto a nitrocellulose membrane. ScFv fragments were detected with an anti-His tag scFv–alkaline phosphatase fusion protein (Lindner *et al.*, 1997).

Gel filtration

For analytical gel filtration, 50 μg of the purified scFv fragments were run on a Superose 12 PC 3.2/30 column on a SMART System (Pharmacia) in 20 mM HEPES, 150 mM NaCl, pH 6.8, at a flow rate of 60 μl/min. The eluted proteins were monitored by measuring the absorbance at 280 nm.

Results

Construction of an scFv library with partially randomized linker region

As a model system for introducing a non-repetitive linker, we used the fluorescein binding antibody scFv fragment FITC-E2 (Vaughan *et al.*, 1996), which was in the orientation V_H–$(Gly_4Ser)_3$–V_L. The linker was replaced by a randomized oligonucleotide, designed according to the following principles. In modeling the scFv fragment it becomes obvious that the protein chain must make a turn in going from the end of V_H

into the linker. In a natural antibody, the chain continues in the same direction at the end of V_H, creates the elbow bend and leads into C_H1. In the common scFv fragments with $(Gly_4Ser)_3$ linker, the linker would be of sufficient flexibility to allow the turn required for an scFv fragment, but it would have no structural preference for it. Consequently, diabodies can result, in which two scFv fragments associate, if the V_L domain does not associate with its covalently attached V_H (requiring the linker to make a turn), but with V_H of another molecule (requiring no turn) (Perisic *et al.*, 1994). We therefore designed the sequence Ser-Pro-Asn-Gly at the beginning of the linker region that encodes a typical type I β-turn (Wilmot and Thornton, 1988). At the end, the linker again has to change its direction when leading into V_L. Here we designed the sequence (Pro/Ser)-Gly-Ser-Gln, allowing selection either for a proline-mediated kink or for the sequence Ser-Gly-Ser-Gln encoding a type II′ β-turn (Mattos *et al.*, 1994).

We chose the length of the linker to be 19 residues, which should be sufficient to drive the monomer–diabody equilibrium completely to the side of the monomers. Increasing the length of the linker favors monomers (Desplancq *et al.*, 1994) and in this model system already the $(Gly_4Ser)_3$ linker gave exclusively monomers (see Figure 6). In the region between the two turns, we wished to allow charged and polar residues in addition to a number of glycines, such that the conformation could adapt to the requirements and the preferred ionic composition would be selected. For instance, two glycine-rich spacers separate the domains of the minor coat protein g3p (Beck and Zink, 1981) and these contain a number of acidic residues.

We therefore designed the oligonucleotide shown in Figure 2A, where three different codon mixtures were specified, allowing for the amino acids given below the nucleotide sequence to occur with equal probability. To verify randomization of the respective codons, 20 unselected clones of the library that were correct on the basis of restriction analysis were sequenced. Of these, six had an insertion or deletion in the linker region. Among the remaining 14 clones, the frequencies of different residues observed at each randomized position were almost as expected (Figure 2A).

SIP

The constructed scFv library was displayed on the scFv-'medium' SIP phage (Krebber *et al.*, 1997b) and selected for fluorescein binding clones by one round of SIP using N1-Flu (Krebber *et al.*, 1997b) as adapter (Figure 1). Fluorescein was chemically coupled to the N1 domain of g3p to give the N1-Flu adapter. This adapter has the advantage over N1-N2-Flu that it does not inhibit infection at high concentrations (Krebber *et al.*, 1997b) and therefore its concentration is not critical as long as it is high enough. The use of 'medium' phages (containing the N2 domain) results in higher infectivity than that of 'short' phages (not containing the N2 domain) (Krebber *et al.*, 1997b). Growth of the phage library at 30°C instead of 37°C led to a larger number of clones obtained after SIP (80 000 vs. 2500), some of which, however, were only functional when regrown at 30°C but not when regrown at 37°C (data not shown).

Phage ELISA

Of the colonies obtained after SIP, 22 were picked randomly, grown in liquid culture and after overnight growth the culture supernatant was screened by ELISA for fluorescein binding
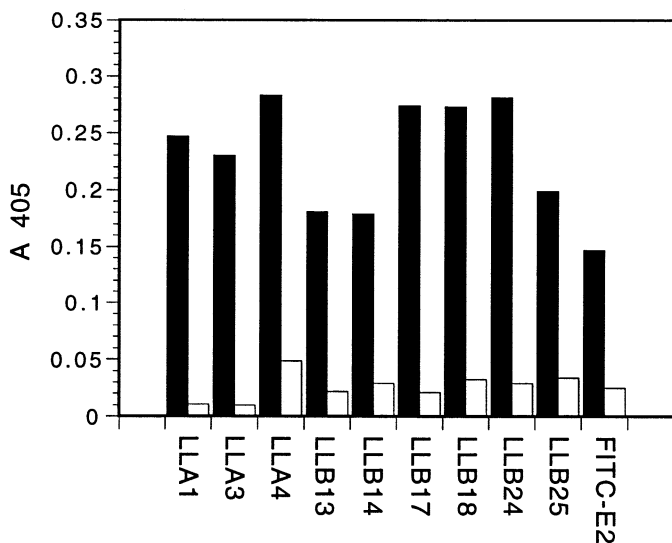
Fig. 3. ELISA of scFv displaying phagemid particles on FITC–BSA-coated plates in the absence (filled bars) and presence (white bars) of 10^{-7} M fluorescein as competitive inhibitor. Nine clones selected for fluorescein binding by one round of SIP as well as the original clone FITC-E2, containing the $(Gly_4Ser)_3$ linker, are shown.

phage particles. For all clones, binding to the FITC–BSA coated plate could be detected, whereas before SIP selection, very few (1 out of 23) clones gave rise to an ELISA signal (data not shown). To confirm these results with monovalent display—the SIP phages lead to multivalent display—nine of the selected scFv variants were cloned to the phagemid vector pAK100 (Krebber *et al.*, 1997a) and the antigen-binding ELISA was repeated with the resulting phagemid particles (Figure 3). Clone FITC-E2, containing the original $(Gly_4Ser)_3$ linker, also in the pAK100 format, was used as a control. All SIP selected variants investigated gave rise to an ELISA signal at least as high as that of clone FITC-E2 and all were efficiently inhibited in the presence of 10^{-7} M fluorescein.

Sequences

To determine whether selection for fluorescein binding resulted in selection for distinct sequences, the clones shown in Figure 3 were sequenced (Figure 2B). Except for the second randomized position, where exclusively serine residues were found, no obvious bias towards any particular sequence was observed. Owing to a deletion, clone LLB24 had a severely shortened linker of 12 amino acid residues that sterically should not allow intramolecular association of V_H and V_L domains. The most plausible explanation of why this clone gives rise to an antigen-binding ELISA signal is that it contains an scFv diabody on the phage (McGuinness *et al.*, 1996; Schier *et al.*, 1996), with one monomer covalently fused to g3p, the other one being soluble owing to proteolytic cleavage of the scFv–g3p linker.

Determination of dissociation constants

The intensities of phage ELISA signals may be influenced by a number of parameters such as the antigen binding constant and the display frequency on the phage, which in turn is dependent on various properties such as folding efficiency, thermodynamic stability, proteolytic stability and solubility of the displayed scFv fragments. In order to determine which parameters to select for, three of the selected clones (LLA1, LLA4, LLB17) were subjected to a more detailed analysis of physicochemical properties. For this purpose, the respective
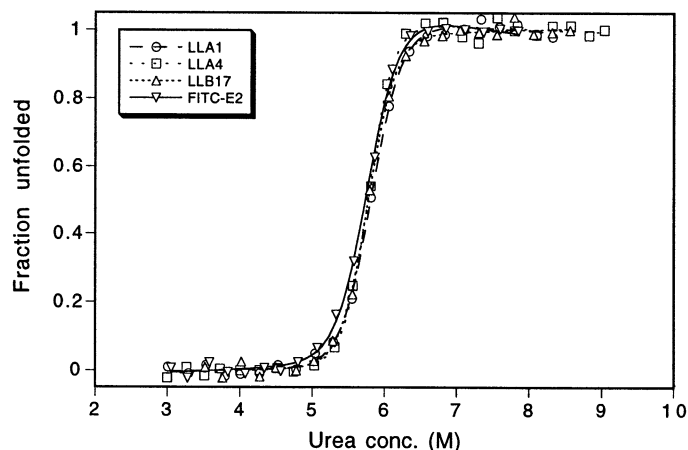
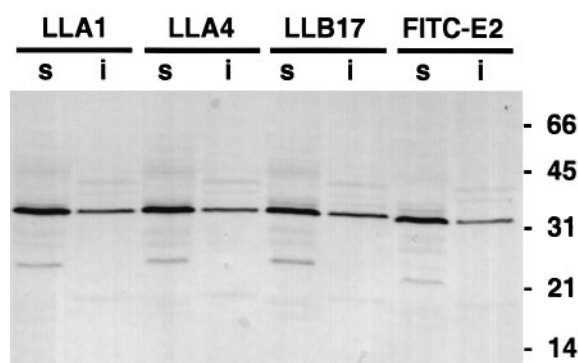**Fig. 4.** Urea-induced unfolding of scFv variants.

Fig. 5. Quantitative Western blot analysis of the partitioning of scFv variants into soluble (s) and insoluble (i) fractions of crude *E.coli* lysate. Protein samples loaded per lane correspond to equal amounts of cells. The electrophoretic mobility of marker proteins ($M_r \times 10^{-3}$) is indicated beside the blot.

Table I. Dissociation constants determined by fluorescence quench titration

scFv fragment	K_D (nM)
LLA1	0.56 (± 0.02)
LLA4	0.53 (± 0.05)
LLB17	0.72 (± 0.02)
FITC-E2	0.73 (± 0.09)

genes were cloned into the expression vector pAK400 (Krebber *et al.*, 1997a) and soluble scFv protein was purified. Determination of affinities to fluorescein in solution by fluorescence quench titration revealed dissociation constants of ~0.5–0.7 nM for all variants (Table I). The selected linker variants thus exhibited high affinities to fluorescein identical with the original FITC-E2, within experimental error.

Urea unfolding

In order to investigate the influence of the variant linker regions on conformational stability, reversible unfolding of the scFv fragments in the presence of increasing concentrations of urea was monitored (Figure 4). As is evident from the unfolding curves, all variants seem to be fairly stable, the midpoint of transition being ~5.8 M urea in each case. Furthermore, all variants can be closely superimposed on the wild-type with only subtle deviations.
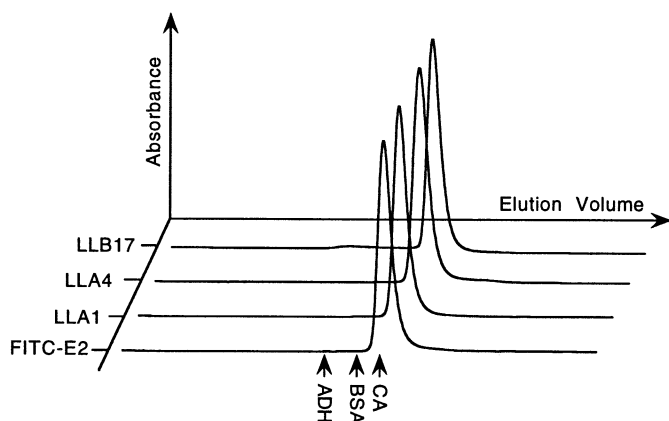
Fig. 6. Analytical gel filtration. The elution profiles of the scFv fragments LLA1, LLA4, LLB17 and FITC-E2 are shown. The elution volumes of calibration proteins alcohol dehydrogenase (ADH) (M_r = 150 000), bovine serum albumin (BSA) (M_r = 66 000) and carbonic anhydrase (CA) (M_r = 29 000) are indicated by arrows.

Soluble expression of scFv fragments

To compare the soluble expression of the variants, the partitioning of the respective scFv proteins into soluble and insoluble fractions of crude *E.coli* lysates was determined by quantitative Western blot analysis (Figure 5). As is evident from the blot, the total expression level, proteolytic stability and the ratio of soluble and insoluble scFv fragments accumulated inside the cells are about the same for all variants and do not deviate significantly from the wild-type with the (Gly$_4$Ser)$_3$ linker.

Analytical gel filtration

To investigate whether the selected scFv linkers favor intramolecular V_H/V_L association over the formation of diabodies, the purified scFv fragments were subjected to analytical gel filtration (Figure 6). Only a single molecular species could be detected in all cases. From the elution volumes, apparent molecular masses of 28 800 (LLA1), 25 700 (LLA4), 28 300 (LLB17) and 28 100 (FITC-E2) were calculated, which are in good agreement with the theoretical value of 30 000. Hence all variants form exclusively monomers in solution.

Discussion

We have selected, by using SIP technology, a series of scFv linkers from partially designed and partially randomized non-repetitive linker sequences. The selected linkers give rise to scFv fragments which are indistinguishable in properties from that containing the original (Gly$_4$Ser)$_3$ linker, but are highly advantageous in that they do not contain internal repeats that may lead to recombination and shortening of the linker during PCR. This feature will simplify the assembly of scFv fragments from amplified V_L and V_H by PCR. In addition, non-repetitiveness was previously shown to be an important prerequisite for the generation of scFv fragments by mutagenic DNA shuffling (Stemmer, 1994; Proba *et al.*, 1998). Even though the three Gly$_4$Ser repeats have often been encoded by different codons to minimize this recombination (Ge *et al.*, 1995; Krebber *et al.*, 1997a), it cannot be completely eliminated (Proba *et al.*, 1998). For example, in phage panning, shortened variants are frequently selected, presumably forming diabodies on the phage (Schier *et al.*, 1996). Furthermore, a sequence stretch homologous to the linker exists at the beginning of

V_H, leading again to problems in PCR-based mutagenesis (Proba *et al.*, 1998).

A number of studies have used linkers other than (Gly$_4$Ser)$_n$ (summarized by Huston *et al.*, 1991). However, they have rarely been compared and have not been optimized in any way. Three previous approaches have used selection of random linkers (Stemmer *et al.*, 1993; Tang *et al.*, 1996; Turner *et al.*, 1997). As in our case, several different sequences were found by Tang *et al.* (1996) and Turner *et al.* (1997) by phage panning, indicating that there are many possibilities of linking V_H and V_L. It appeared that the only feature selected for in the study by Tang *et al.* (1996) was a proline at the beginning, which we therefore encoded as a turn in the whole linker library. Interestingly, a slight bias of basic residues was observed, which is also consistent with our experiments, even though the trend is very weak. In contrast to Tang *et al.* (1996) and Turner *et al.* (1997), we observed no influence of the linker on expression. However, these authors started with poorly expressed scFv fragments with the aim of increasing the yield of functional scFv protein. In contrast, the aim of our study was to generate highly functional scFv fragments containing non-repetitive linkers. Therefore, we used as a model system an scFv fragment with favorable overall properties and compared the selected mutants with it. It seems that, in our study, any poorly expressed mutants which might conceivably have been generated by randomizing the linker sequence have been efficiently counterselected.

The new linkers appear to perform identically with the original (Gly$_4$Ser)$_3$ variant. This shows that, as expected, the folding and expression properties of the scFv fragments are governed by the antibody domains and not by the linker. The linker thus appears to be a passive connector, consistent with the observation in NMR (Freund *et al.*, 1993) that there are no cross peaks between linker and antibody domains and hence no stable contacts. Furthermore, both NOE and T_2 values indicate an extremely high mobility (Freund *et al.*, 1993), which is consistent with the fact that the linker has been invisible in X-ray structures of scFv fragments (Kortt *et al.*, 1994; Zdanov *et al.*, 1994). This is also borne out by molecular dynamics simulations (A.Caflisch, unpublished work), which indicate very high flexibility and almost no contact between the linker and the antibody domains. Nevertheless, we cannot conclude that the set of linkers selected in this study will be generally favorable independent of the context of the scFv sequence. However, here we have presented a reliable procedure for easily adapting the sequences to any desired scFv fragment.

SIP technology was shown to lead to the selection of useful clones in a single round with immediate elimination of all non-functional linker sequences and most of the clones containing insertions or deletions. Although one could conceivably select for the best variants among the selected clones by further rounds of SIP, this would limit the variability to the given library diversity. Instead, we have now constructed a robust system which can be randomized by any number of PCR-based technologies and selected for the best variants including changes all over the protein.

Beyond the mere selection of functional non-repetitive scFv linkers, we have demonstrated with this study the application of SIP technology to the screening of a randomized immunoglobulin library. It stringently selects not only for high binding constants (already present in our starting molecule) but also for satisfactory overall properties of the displayed molecule,

including expression and folding. We believe that this technology is a powerful tool for the molecular optimization of a wide variety of ligand binding proteins.

Acknowledgements

References

Bass,S., Gu,Q. and Christen,A. (1996) *J. Bacteriol.*, **178**, 1154–1161.

Beck,E. and Zink, B. (1981) *Gene*, **16**, 35–58.

Bird,R.E. *et al.* (1988) *Science*, **242**, 423–426.

Bork,P., Holm,L. and Sander,C. (1994) *J. Mol. Biol.*, **242**, 309–320.

Desplancq,D., King,D.J., Lawson,A.D.G. and Mountain,A. (1994) *Protein Engng*, **7**, 1027–1033.

Freund,C., Ross,A., Guth,B., Plückthun,A. and Holak,T.A. (1993) *FEBS Lett.*, **320**, 97–100.

Ge,L., Knappik,A., Pack,P., Freund,C. and Plückthun,A. (1995) In Borrebaeck, C.A.K. (ed.), *Antibody Engineering*. Oxford University Press, Oxford, pp. 229–266.

George,A.J.T., Spooner,R.A. and Epenetos,A.A. (1994) *Immunol. Today*, **15**, 559–561.

George,A., Huston,J.S. and Haber,E. (1996) *Nature Biotechnol.*, **14**, 584.

Glockshuber,R., Malia,M., Pfitzinger,I. and Plückthun,A. (1990) *Biochemistry*, **29**, 1362–1367.

Huston,J.S. *et al.* (1988) *Proc. Natl Acad. Sci. USA*, **85**, 5879–5883.

Huston,J.S., Mudgett-Hunter,M., Tai,M.-S., McCartney,J., Warren,F., Haber,E. and Oppermann,H. (1991) *Methods Enzymol.*, **203**, 46–88.

Kortt,A.A. *et al.* (1994) *Eur. J. Biochem.*, **221**, 151–157.

Krebber,A., Bornhauser,S., Burmester,J., Honegger,A., Willuda,J., Bosshard,H.R. and Plückthun,A. (1997a) *J. Immunol. Methods*, **201**, 35–55.

Krebber,C., Spada,S., Desplancq,D., Krebber,A., Ge,L. and Plückthun,A. (1997b) *J. Mol. Biol.*, **268**, 607–618.

Lindner,P. *et al.* (1997) *BioTechniques*, **22**, 140–149.

Mattos,C., Petsko,G.A. and Karplus,M. (1994) *J. Mol. Biol.*, **238**, 733–747.

McGuinness,B.T., Walter,G., FitzGerald,K., Schuler,P., Mahoney,W., Duncan,A.R. and Hoogenboom, H.R. (1996) *Nature Biotechnol.*, **14**, 1149–1154.

Pace,C.N. and Scholtz,J.M (1997) In Creighton,T.E. (ed.), *Protein Structure: a Practical Approach*. IRL Press, Oxford, pp. 299–321.

Pace,C.N., Vajdos,F., Fee,L., Grimsley,G. and Gray,T. (1995) *Protein Sci.*, **4**, 2411–2423.

Padlan,E.A. (1994) *Mol. Immunol.*, **31**, 169–217.

Pedrazzi,G., Schwesinger,F., Honegger,A., Krebber,C. and Plückthun,A. (1997) *FEBS Lett.*, **415**, 289–293.

Perisic,O., Webb,P.A., Holliger,P., Winter,G. and Williams,R.L. (1994) *Structure*, **2**, 1217–1226.

Plückthun,A. (1992) *Immunol. Rev.*, **130**, 151–188.

Proba,K., Wörn,A., Honegger,A. and Plückthun,A. (1998) *J. Mol. Biol.*, **275**, 245–253.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schier,R., Bye,J., Apell,G., McCall,A., Adams,G.P., Malmqvist,M., Weiner,L.M. and Marks,J.D. (1996) *J. Mol. Biol.*, **255**, 28–43.

Spada,S. and Plückthun,A. (1997) *Nature Med.*, **3**, 694–696.

Spada,S., Krebber,C. and Plückthun,A. (1997) *Biol. Chem.*, **378**, 445–456.

Stemmer,W.P.C. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 10747–10751.

Stemmer,W.P.C., Morris,S.K. and Wilson,B.S. (1993) *BioTechniques*, **14**, 256–265.

Tang,Y., Jiang,N., Parakh,C. and Hilvert,D. (1996) *J. Biol. Chem.*, **271**, 15682–15686.

Turner,D.J., Ritter,M.A. and George,A.J.T. (1997) *J. Immunol. Methods*, **205**, 43–54.

Vaughan,T.J. *et al.* (1996) *Nature Biotechnol.*, **14**, 309–314.

Wilmot,C.M. and Thornton,J.M. (1988) *J. Mol. Biol.*, **203**, 221–232.

Winter,G., Griffiths,A.D., Hawkins,R.E. and Hoogenboom,H.R. (1994) *Annu. Rev. Immunol.*, **12**, 433–455.

Zdanov,A., Li,Y., Bundle,D.R., Deng,S.-J., MacKenzie,C.R., Narang,S.A., Young,N.M. and Cygler,M. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 6423–6427.