FEBS Letters 377 (1995) 92–96 FEBS 16384

An antibody VH domain with a *lox*-Cre site integrated into its coding region: bacterial recombination within a single polypeptide chain

Julian Davies, Lutz Riechmann*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Received 24 October 1995

Abstract **Bacterial *lox*-Cre recombination within a single antibody VH domain was achieved through integration of a *loxP* site into its coding sequence. The 5′ half of the VH gene, in which the H2 loop was replaced by a mutant *loxP* site, was fused to geneIII in an 'acceptor' fd-phage vector containing also a wild type *loxP* site. With a 'donor' plasmid vector harbouring the 3′ *half of the VH gene flanked by the same, differing loxP* sites it recombined into a full-length VH with the *loxP* site-H2 loop. This VH was purified from bacterial periplasm, where it folded into a typical immunoglobulin domain. The system allows the generation of large VH repertoires using *lox*-Cre recombination.**

Key words: Cre; Recombination; Antibody; VH; Circular dichroism; Thermodenaturation

1. Introduction

The Cre recombinase of bacteriophage P1 recognises a 34 base pair DNA sequence (*loxP* site), which consists of two inverted repeats seperated by an 8 base pair spacer [1,2]. If two *loxP* sites of identical sequence are present on a DNA fragment, Cre recombinase catalyses deletion of the sequence between the two sites. A single point mutation in the spacer sequence between the inverted repeats in only one of two *loxP* sites on a single piece of DNA can prevent deletion through Cre recombination [2]. However, combinatorial exchange with a second DNA fragment takes place, if this is flanked by the same differing *loxP* sites present on the first DNA fragment.

Using such a system, it was possible to recombine two seperate repertoires of genes from different replicons to generate and display a large library of antibody Fab fragments on phage [3,4]. The resulting repertoire with more than 10^{10} members allowed the antigen driven selection of antibody Fab fragments with dissociation constants for their ligands in the nanomolar range [4].

The same strategy can easily be adopted for bacterial recombination of other (poly)peptide based recognition units as long as they are encoded by at least two seperate genes like the light and heavy chain genes in the case of Fab fragments. Then the two *loxP* sites, which remain after Cre recombination, are easily accomodated: one is placed outside of the genes to be recombined, while the other is positioned between them.

To achieve *lox*Cre recombination of a single polypeptide chain in bacteria, the placing of the *loxP* site in between the two DNA fragments to be recombined becomes more difficult. One of the remaining *loxP* site must either be spliced out on the RNA level or it has to be integrated into the coding sequence of the gene itself. To ultimately prepare a large repertoire of single antibody domains we describe here the integration of a *loxP* site into the middle of the coding sequence of a camelised human VH domain to achieve recombination of two halves of the VH gene while located on seperate replicons.

2. Experimental

2.1. Vectors and recombination

The vector fd-loxA (Fig. 1) was prepared by cloning the 5′ half of the camelised, human antibody VH domain VH-P37 (identical to VH-P1 [5] except for residue 37 (V37F) mutated for a different purpose) into the vector fd-2loxvkdel [4]. The fd vector fd-VH harbouring the gene for VH-P37 between the leader and the remainder of the fd-geneIII [6] was amplified using PCR with oligonucleotides fdPCRBAK [7] and fdLOXFOR (5′-TTT GCA TGC CTG CAG **ATA ACT TCG TAT AGT ATA CAT TAT ACG AAG TTA** TAG CTGA GAC CCC CTC-3′). Following restriction with *Pst*I and *Apa*LI the PCR product containing the *loxP511* site was cloned into fd-2loxvkdel providing a wild type *loxP* site. The *PstI*/*Not*I fragment between the *loxP* sites was then replaced by stuffer DNA restoring the reading frame throughout the geneIII-VH fusion (Fig. 1).

The vector pUC-loxB (Fig. 1) was constructed through cloning of the 3′ portion of VH-P37 from fd-VH after amplification by PCR with oligonucleotides fdSEQ1 [7] and pUCLOXBAK (5′-GAG TGG GTC AAG CTT **ATA ACT TCG TAT AAT GTA TAC TAT ACG AAG** TTA TAC TAC GCA GAC TCC GTG-3′) and restriction with *Hin*dIII and *Not*I into pUC19–2loxVHdel [4]. pUCLOXBACK introduced the *loxP511* site, while the 3′ wild type *loxP* site was provided by pUC19–2loxVHdel.

Recombination of the vectors fd-loxA and pUC-loxB was principally performed as described [3] except for the use of a chloramphenicol resistance confering plasmid (S. Williams, A. Nissim, O. Hartley, J.L. Harrison, A. Griffiths and G. Winter, in preparation) providing arabinose inducable Cre expression instead of phage P1Cm*c*1.100. Tetracycline resistant colonies obtained after recombination were tested by PCR screening with the primers fdPCRBACK and fdSEQ1 for the presence of fd-loxR.

2.2. Protein preparation and analysis

To express soluble VH-lox protein the *Xho*I/*Not*I fragment of fd-loxR was subcloned into the vector pUC-VH47G [6] containing a *Xho*I site at the beginning of CDR1. For expression bacterial cultures were grown overnight at 37°C in rich medium containing ampicillin (100 mg/l) and 0.1% glucose followed by a three hour induction period at 30°C with 1 mM isopropyl β-D-thiogalactopyranoside. VH-lox and VH-P37 proteins were purified from the periplasm [8] by IMAC [9,10] according to the manufacturer's instructions (Diagen).

Purified VH was analysed by SDS-polyacrylamide gel electrophore-

*Corresponding author. Fax: (44) (1223) 41-2178.
E-mail: lutz@mrc-lmb.cam.ac.uk

Abbreviations: CD, circular dichroism; CDR1, CDR2, CDR3, complementarity determining regions 1, 2 and 3; Fab, antigen binding antibody fragment; Fv, heterodimer of VH and VL; H1, H2, H3, hypervariable loops 1, 2 and 3; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; SDS, sodium dodecyl sulphate; VH, heavy chain variable domain; VL, light chain variable domain. Amino acids are abbreviated using the one letter code; mutations are denoted by the wild type amino acid followed by the residue number and the new amino acid.

SSDI 0014-5793(95)01313-X

sis [11] in the presence and absence of β-mercaptoethanol. The Ellman assay [12,13] was used to detect possible free cysteine residues. Circular dichroism analysis was performed at a VH concentration of 10 μg/ml in phospate buffered saline solution, pH 7, using a Jasco J720 spectropolorimeter connected to a Vigelin III/33 computer and a Neslab RTE-110 waterbath through an RS232 interface. Wavelength scans were recorded at 20°C. Thermodenaturation was measured by following the elipticity at 235 nm in steps of 0.2°C at a rate of 50°/h.

3. Results and discussion

3.1. VH domains

The Fv fragment of an antibody comprises a pair of variable domains from heavy (VH) and light chain (VL). It is usually viewed as the smallest functional antibody fragment with respect to antigen binding. However, some early experimental evidence indicated that VH domains alone can retain some of the antigen affinity [14]. More recently, antigen binding was also detected in expression libraries of mouse VH domains [15]. For camels even antibodies were discovered, which lack a natural light chain partner [16]. Furthermore, antigen binding VH domains [6] or size reduced fragments thereof [17] were selected *in vitro* from synthetic phage-display libraries.

For an efficient antigen driven selection of phage displayed, human VH domains non-specific binding of these VH had to be reduced through three mutations in the former light chain interface (G44E, L45R and W47G [5,6]). These mutations were

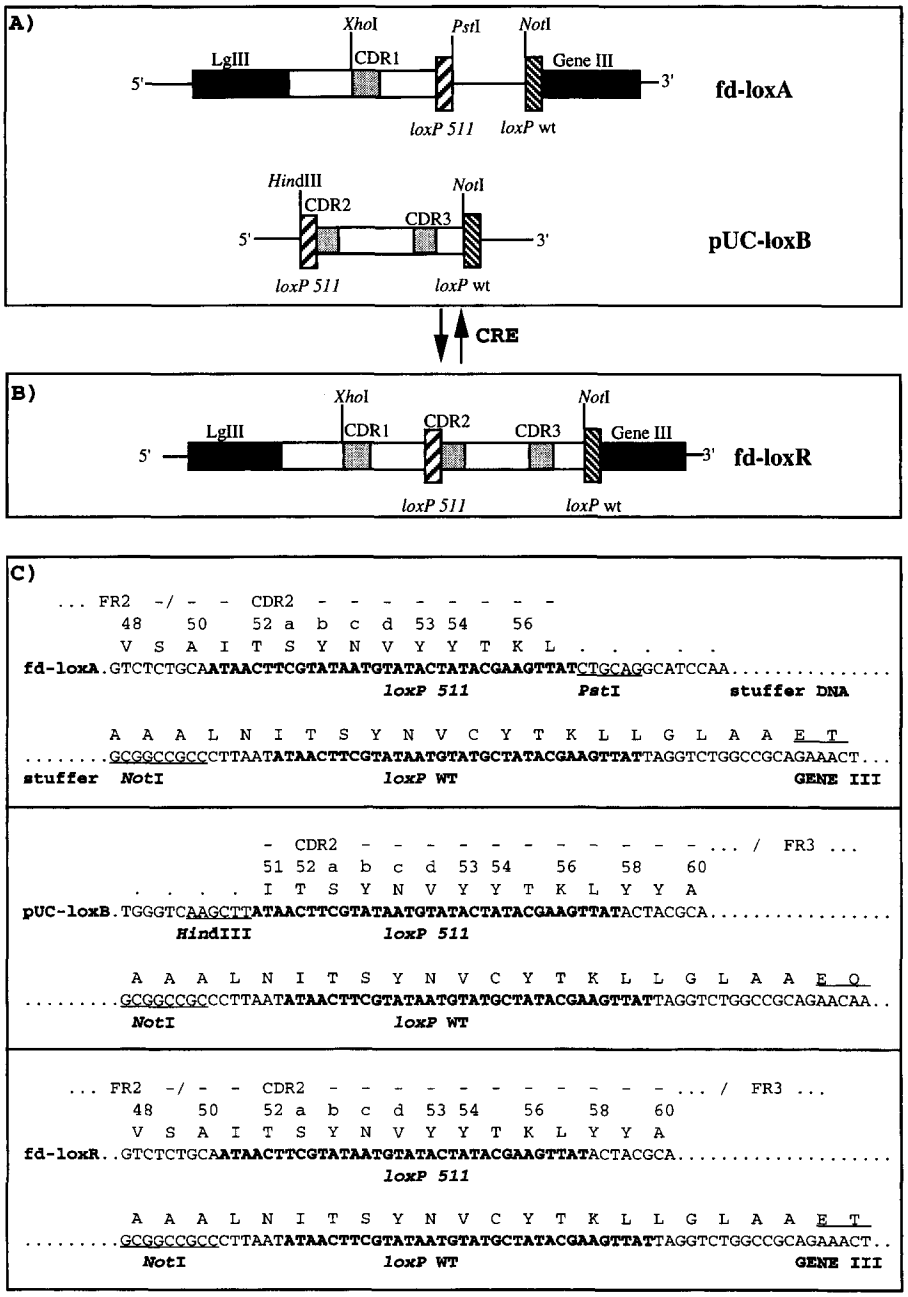

Fig. 1. *Lox*-Cre recombination 'acceptor' vector fd-loxA (A), 'donor' vector pUC-loxB (B), recombined vector fd-loxR (B) and DNA sequences close to the *loxP* sites in all three vectors (C).

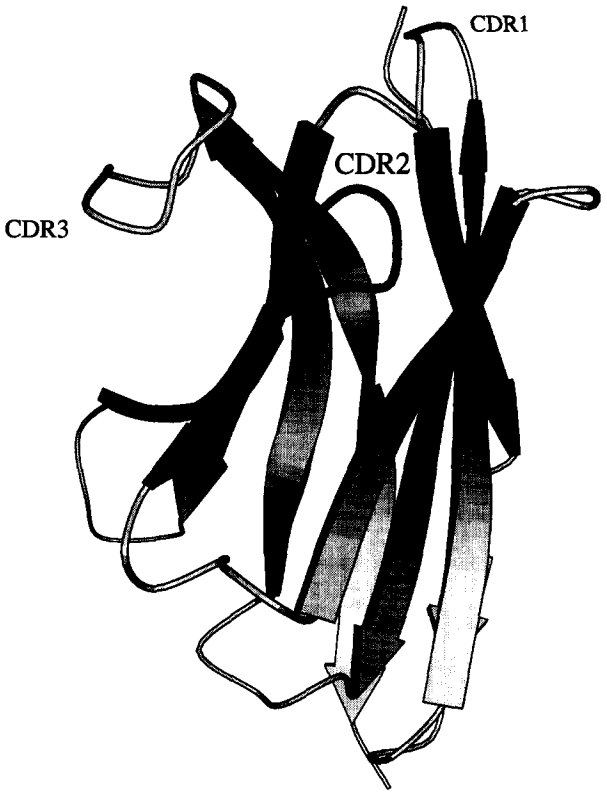

Fig. 2. MolScript [31] ribbon diagram of a human antibody VH domain taken from the crystal structure of the Fv Pot [32] with the H2 loop in CDR2 shown in black.

introduced to mimic the camelid heavy chains, which occur naturally without light chain [16,18]. Then highly specific VH domains could be isolated after selection on protein or hapten–protein conjugates [6]. However, as found for phage-display repertoires of Fv [19] or Fab [20] fragments affinities of isolated antibody fragments for antigen were moderate as long as the size of the libraries was not bigger than about 10^8. Dramatic increases in affinity with dissociation constants in the nanomolar range of selected Fab fragments and their ligands were only observed when considerably larger libraries ($>10^{10}$) were used [4].

3.2. LoxP *site integration*

Libraries of such size cannot be obtained by simple transformation of bacteria. In the case of the large Fab library a combinatorial approach was used, in which a light chain library displayed on phage was recombined with a heavy chain library through infection of bacteria harbouring a plasmid based heavy chain repertoire [4]. Site-specific recombination was facilitated through the *lox*-Cre system, for which the heavy chain on the 'donor' plasmid was flanked by the same two differing *loxP* sites present 3′ of the light chain gene on the 'acceptor' phage. The two *loxP* sites on each vector had to be different in sequence to prevent deletion within a vector. Recombination of *loxP* site flanked DNA between the two vectors however took place in the bacteria when Cre recombinase was present and created a repertoire, which in practice was limited only by the number of infective events.

To employ the same system to produce large phage-display libraries of VH domains 'acceptor' phage and 'donor' plasmid vectors (Fig. 1) were designed analogous to those used for the recombination of the Fab library. The 'acceptor' fd phage vector fd-loxA (confering tetracycline resistance) contains the 5′ half (from the N-terminus up to the H2 loop) of the camelised VH-P37 fused at its N-terminus to the geneIII leader peptide. At the C-terminus an engineered H2 loop based on the mutant *loxP511* DNA sequence was integrated into the coding sequence. This was necessary as in the case of bacterial expression of a single polypeptide like the VH the recombined gene will normally not contain any noncoding sequences between N and C-terminus, which would allow accomodation of the 5′ of the two *loxP* sites outside of coding DNA. In fd-loxA a second, but now wild type *loxP* site is located between the following stuffer DNA and the remainder of geneIII. Thus resulting phage particles are highly infectious as they contain an intact proteinIII as part of a fusion protein in their coat.

The 'donor' plasmid vector pUC-loxB (confering ampicillin resistance) contains the 3′ half of the VH (from the H2 loop to the C-terminus). This is flanked by the same *loxP* sites present on the 'acceptor' vector and should therefore allow recombination of this region into the 'acceptor' phage when Cre recombinase is present. Neither phage nor plasmid DNA should recombine within itself as their two *loxP* sites were not identical.

Recombination between the two replicons should create a VH gene, in which the H2 loop residues of the 'wild type' VH-P37 (V51-S52-G52a-S53-G54-G55-S56-T57) are replaced by I51-T52-S52a-Y52b-N52c-Y53-Y54-T55-K56-L57 (numbering according [21]) resulting in a lengthening of the hypervariable loop in CDR2 by two residues. We opted for the H2 loop to locate the *loxP511* sequence, because VH domains must be able to accomodate a variety of sequences in this region, which is by definition hypervariable. Use of CDR2 will also allow later a seperated randomisation of CDR1 on the 'acceptor' phage and CDR3 on the 'donor' plasmid. These two CDRs are together with CDR2 the most variable regions of VH genes in natural repertoires [21] and their variation will create a repertoire of binding sites around a scaffold formed by the translated *loxP* site in the H2 loop.

H2 loops identical in length to that in the newly created VH were also found among other members of the originating germline VH3 gene family [22]. It is worth noting that among

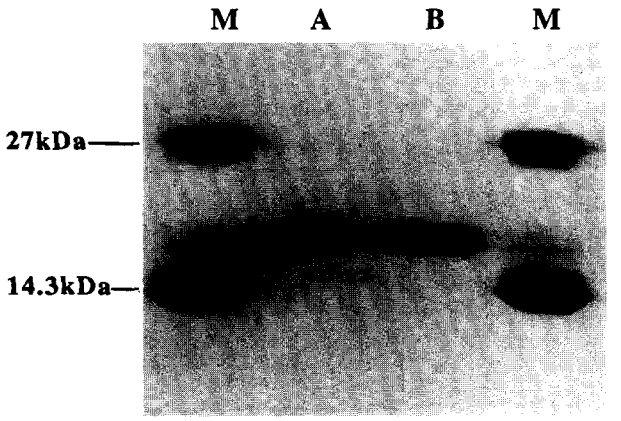

Fig. 3. Non-reducing SDS polyacrylamide gel of purified VH-lox (A), VH-P37 (B) and molecular weight markers (M).

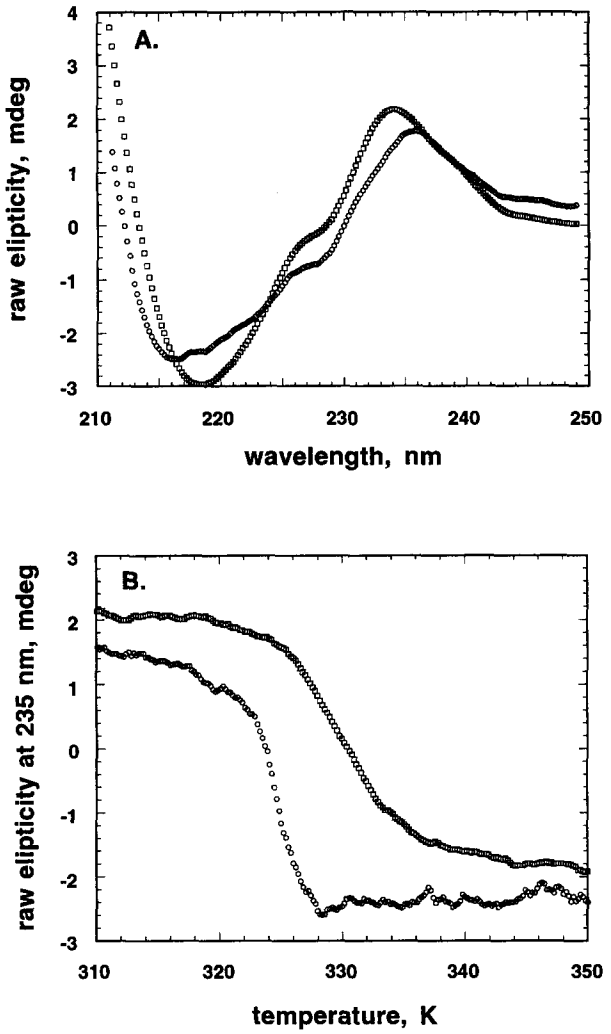

Fig. 4. Circular dichroism spectra (A) and thermodenaturation curves (B) of VH-lox (circles) and VH-P37 (squares).

the residues replaced only those at position 51, 52 and 57 are part of regular β-strand in the VH while residues 52a–56 form the turn of the H2 loop [23].

The two vectors fd-loxA and pUC-loxB were used to perform a Cre recombination. Bacteria harbouring the 'donor' plasmid in addition to a second plasmid confering inducable Cre recombinase expression were infected with 'acceptor' phage. 30% of resulting tetracycline resistant colonies contained a recombined VH gene as part of fd-loxR phage DNA (Fig. 1). The recombination result was similar to that observed for recombination in case of the Fab phage libraries. The recombination frequency should be sufficient to guarantee expression of a recombined VH–geneIII fusion protein from each infected bacterium as there are multiple copies of plasmid and phage replicons in each cell, which will produce at least 60 phage particles during an overnight growth period [4].

3.3. VH-lox protein

We then investigated if the resulting VH domain (termed VH-lox) still formed the same general protein fold as the VH with the wild type CDR2. The VH-lox gene was subcloned from fd-loxR into a soluble expression vector allowing correctly folded protein to be secreted into the periplasm via the N-terminal pectate lyase leader peptide, which is cleaved off during the membrane transport [24]. This system is being widely used for soluble expression of antibody fragments [25] and was also used for soluble expression of VH-P37, which is identical to VH-lox except for its CDR2. Homonuclear NMR spectra of VH-P37 (L.R., unpublished) were very similar to those of the closely related VH-P8 [23], from which it differs only at postion 37 (V37F) and 47 (G47I). The solution structure of VH-P8 has been solved through NMR spectroscopy [23 and L.R., in preparation] and is principally identical of that of VH domains in intact antibodies (Fig. 2).

VH-lox was secreted from bacteria just as VH-P37. The purification yield for VH-lox was 1.3 mg/l bacterial culture compared to 3 mg obtained with VH-P37. Both proteins tested negative for free SH-groups and both run as monomers of expected size on reducing and non-reducing SDS-polyacrylamide gels (Fig. 3). Therefore the VH typical intramolecular disulphide bond between the cysteines at position 22 and 92 had to be correctly formed in VH-lox, because there were neither free SH groups present nor multimeric complexes formed through formation of intermolecular disulphide bridges.

Circular dichroism analysis (Fig. 4) indicated anti-parallel β-sheets as the predominant secondary structure element in both VH domains, which was expected for immunoglobulin domain and was also determined from nuclear Overhauser enhancement NMR analysis of the closely related VH-P8 [23]. The CD curves had a negative band near 220 nm typical for anti-parallel β-sheets [26]. The maximum at about 235 nm was indicative for aromatic or disulphide chromophores [26] and could be expected for a folded protein containing eight tyrosines, five phenylalanines and two tryptophans in combination with an intramolecular disulphide bond among altogther 121 (VH-P37) or 124 (VH-lox) VH domain residues.

We also investigated whether the *loxP* site based H2 loop in VH-lox compromised structural stability. Melting points of VH-P37 and VH-lox were determined in thermodenaturation experiments by following their elipticity at 235 nm. Both VH showed a single denaturation transition between 20°C and 90°C (Fig. 4). VH-P37 had a melting point (T_m) of 56.9°C, whereas VH-lox had a T_m of 51.7°C. Thus VH-lox was somewhat less stable than VH-P37. However, VH-lox was still considerably more stable than for example the VH domain of the mouse antibody D1.3, whose T_m was determined as 43°C [27]

The difference in folding stability was approximated in terms of energy from the denaturation curves and difference in melting points. Van't Hoff's enthalpy ΔH_{vh} of unfolding was estimated from the denaturation curve of VH-P37 as 87.5 kcal/mol (compare equation 22 in [28]) and the enthalpy ΔS (0.265 kcal/mol·K) calculated as $\Delta H_{vh}/T_m$. The difference in free energy change of unfolding by thermodenaturation between VH-P37 and VH-lox was then approximated from $\Delta\Delta G = \Delta S_{VH\text{-}P37} \times \Delta T_m$ as 1.38 kcal/mol. This effect on the overall stability was comparable to that observed on average for the loss of a mythelene group in a fully buried residue of barnase (1.5 kcal/mol [29]). Thus replacement of the H2 loop in VH-P37 with the *loxP* site sequence did neither altered the principle structure of the VH nor did it substantially affect the overall folding stability.

4. Outlook

These results demonstrated that integration of the *loxP* sequence into the CDR2 coding sequence of a human VH domain caused no structural or microbiological problems. The VH-lox protein was created through an in vivo recombination experiment, which exhibited the same recombination frequency as one performed with the corresponding *loxP* site integrated into non-coding sequences. The VH-lox protein was purified from the bacterial periplasm in amounts comparable to the VH domain with the 'wild type' CDR2, which in itself suggested correct folding. CD spectra recorded with VH-lox showed that its general fold was identical to that of a typical immunoglobulin VH domain consisting of nine anti-parallel orientated β-strands. Furthermore, the *loxP*-CDR2 had only a small effect on the overall stability of the VH domain resulting in a 5°C decrease of its melting point. It should be possible to increase this stability later through suitable modifications like the mutation of residue 47 from glycine to isoleucine, which increased the stability of the camelised VH-P8 [5] by more than 10°C (L.R., unpublished).

It needs now to be investigated how well the *loxP*-CDR2 performs as a scaffold, around which an antigen binding sites is formed in a VH domain. Assuming that VH domains utilise the three CDRs for their antigen binding sites like a VH in combination with a VL [30], a library of 5′ VH regions containing randomised CDR1 residues in the vector fd-loxA can be recombined with a second library of 3′ VH portions containing randomised CDR3 residues in the vector pUC-loxB. Phage selection of the recombined VH-lox library will then reveal if the *loxP*-CDR2 is suitable as a building block for the formation of high affinity ligand binding sites on VH domains. If so, large repertoires of VH-lox libraries could also be used for a secondary recombination with light chain repertoires thereby creating even more diverse synthetic repertoires of antibody Fv or Fab fragments.

Acknowledgements: J.D. was recipient of an MRC studentship. We thank Andrew Griffiths for the generous supply of vectors and the IRC for Protein Engineering, Cambridge, for use of its CD facilities.

References

[1] Sternberg, N. and Hamilton, D. (1981) J. Mol. Biol. 150, 467–486.
[2] Hoess, R.H., Wierzbicki, A. and Abremski, K. (1986) Nucleic Acids Res. 14, 2287–2300.
[3] Waterhouse, P., Griffiths, A.D., Johnson, K.S. and Winter, G. (1993) Nucleic Acids Res. 21, 2265–2266.
[4] Griffiths, A.D., Williams, S.C., Hartley, O., Tomlinsom, I.M., Waterhouse, P., Crosby, W.L., Kontermann, R.E., Jones, P.T., Low, N.M., Allison, T.J., Prospero, T.D., Hoogenboom, H.R., Nissim, A., Cox, J.P.L., Harrison, L., Zaccolo, M., Gherardi, E. and Winter, G. (1994) EMBO J. 13, 32345–3260.
[5] Davies, J. and Riechmann, L. (1994) FEBS Lett. 339, 285–290.
[6] Davies, J. and Riechmann, L. (1995) Bio/Technology 13, 475–479.
[7] Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P. and Winter, G. (1991) Nucleic Acids Res. 19, 4133–4137.
[8] Breitling, F., Dübel, S., Seehaus, T., Klewinghaus, I. and Little, M. (1991) Gene 104, 147–153.
[9] Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. and Struber, D. (1988) Bio/Technology 6, 1321–1325.
[10] Hoffmann, A. and Roeder, R.G. (1991) Nucleic Acids Res. 19, 6337–6338.
[11] Laemmli, U.K. (1970) Nature 227, 680–685.
[12] Ellman, G.L. (1959) Arch. Biochem. Biophys. 82, 70–77.
[13] Riddles, P.W., Blakeley, R.L. and Zerner, B. (1983) Methods Enzymol. 91, 49–60.
[14] Utsumi, S. and Karush, F. (1964) Biochemistry 3, 1329–1338.
[15] Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T. and Winter, G. (1989) Nature 341, 544–546.
[16] Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Bajyana Songa, E., Bendahman, N. and Hamers, R. (1993) Nature 363, 446–448.
[17] Martin, F., Toniatti, C., Salvati, A.L., Venturini, S., Ciliberto, G., Cortese, R. and Sollazzo, M. (1994) EMBO J. 13, 5303–5309.
[18] Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J.A.R.G. and Hamers, R. (1994) Protein Eng. 7, 1129–1135.
[19] Hoogenboom, H.R. and Winter, G. (1992) J. Mol. Biol. 227, 381–388.
[20] Barbas, C.F., Amberg, W., Simoncsits, A., Jones, T.M. and Lerner, R.A. (1993) Gene 137, 57–62.
[21] Kabat, E.A., Wu, T.T., Perry, H.M., Gottesmann, K.S. and Foeller, C. (1991) Sequences of immunological interest, 5th edn., US Department of Health and Human Services, Washington, DC.
[22] Tomlinson, I.M., Walter, G., Marks., J.D., Llewelyn, M.B. and Winter, G. (1992) J. Mol. Biol. 227, 776–798.
[23] Riechmann, L. and Davies, J. (1995) J. Biomol. NMR 6, 141–152.
[24] Lei, S.-P., Lin, H.-C., Wang, S.-S., Callaway, J. and Wilcox, G. (1987) J. Bacteriol. 169, 4379–4383.
[25] Better, M., Chang, C.P., Robinson, R.R and Horwitz, A.H (1988) Science 240, 1041–1043.
[26] Perczel, A., Park, K. and Fasman, G.D. (1992) Proteins: Structure, Function and Genetics 13, 57–69.
[27] Yasui, H., Ito, W. and Kurosawa, Y. (1994) FEBS Lett. 353, 143–146.
[28] Agashe, V.R. and Udgaonkar, J.B. (1995) Biochemistry 34, 3286–3299.
[29] Serrano, L., Kellis, J.T., Cann, P., Matouschek, A. and Fersht, A. (1992) J. Mol. Biol. 224, 783–804.
[30] Davies, D.R., Padlan, E.A. and Sheriff, S. (1990) Annu. Rev. Biochem. 59, 439–473.
[31] Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950.
[32] Fan, Z.-C., Shan, L., Guddat, L.W., He, X.-M. and Gray, W.R. (1992) J. Mol. Biol. 228, 188–207.