Protein Engineering vol.14 no.11 pp.897–901, 2001

Four-helix bundle topology re-engineered: monomeric Rop protein variants with different loop arrangements

Hans P.Kresse[1,2], Martin Czubayko[3], Gerald Nyakatura[4], Gerrit Vriend[5], Chris Sander[6] and Helmut Bloecker[3]

[1]Experimentelle Kinderkardiologie, Deutsches Herzzentrum, Lazarettstrasse 36, D-80636 Munich, [3]GBF, 38124 Braunschweig, [4]MWG Biotech AG, 85560 Ebersberg, [5]EMBL, 69117 Heidelberg, Germany and [6]EBI, Hinxton, UK

[2]To whom correspondence should be addressed.
E-mail: kresse@dhm.mhn.de

We converted the small homodimeric four-helix bundle repressor of primer protein (Rop) into a monomeric four-helix bundle by introduction of connecting loops. Both left- and right-handed four-helix bundles were produced. The left-handed bundles were more stable and were used to introduce biologically interesting peptides in one of the loops.

Keywords: four-helix bundle/loop engineering/peptide presentation/topology

Introduction

The repressor of primer protein (Rop) is a small, dimeric molecule consisting of two identical chains of 63 amino acids. Each monomer consists of two helices connected by a short turn and a seven-residue C-terminal tail. The two monomers pack together as a fully antiparallel four-helix bundle. The structure of Rop has been elucidated by X-ray (Banner *et al.*, 1987) and NMR (Eberle *et al.*, 1990) techniques. Two helix–turn–helix hairpins in a head-to-tail arrangement form a left-handed, antiparallel four-helix bundle with ~20° interhelical angles and a right-handed superhelical topology. The Rop protein is a constituent of the *Escherichia coli* ColE1 plasmid copy number control mechanism (Cesareni *et al.*, 1991).

Here we present a protein engineering study in which we created a monomeric four-helix bundle starting from the Rop dimer. Our long-term goal is to create a small and stable molecule that can be used for *in vitro* and *in vivo* presentation studies of biologically active peptides.

Several successful topological reorganizations have been realized in different structural classes of proteins. They include circular permutations (Buchwalder *et al.*, 1992; Zhang *et al.*, 1993; Ay *et al.*, 1998) and the transfer of loop modules (Hynes, 1989). MacBeath *et al.*, for example, redesigned the topology of chorismate mutase by directed evolution (MacBeath *et al.*,1998). Finkelstein's group has *ab initio* designed a beta barrel protein with unusual topology, i.e. the strands are connected by loops in a way not seen before in proteins with known structure (Abdullaev *et al.*, 1997). These are all examples of more or less successful engineering and design projects. However, total *ab initio* protein design is still an interesting challenge (Bryson *et al.*, 1995; Walsh *et al.*, 1999).

Sander's group designed a monomeric four-helix bundle Rop (Sander, 1994). This bundle was designed to be left-handed and the success of the designed was confirmed by NMR (W.Eberle and C.Sander, personal communication). Regan's group (Predki and Regan, 1995; Nagi and Regan, 1997) have made an interesting right-handed monomer Rop design similar to the right-handed constructs that we present here.

Principally there exist four different ways in which a fully antiparallel monomeric Rop four-helix bundle topology can be organized when the three loops connect helices that are directly adjacent to each other. Figure 1 shows the loop engineering that is required to arrive at these four possibilities starting from the Rop dimer. It can be seen that three of the four possibilities require the design of only two new loops while one of the right-handed bundles can only be obtained by designing three new loops. Two of these bundles have a left-handed and two a right-handed super helical topology. The Rop wild-type four-helix bundle is left-handed.

The studies by the Sander and Regan groups have revealed the possibility of converting the Rop dimer into a monomeric four-helix bundle with different topologies. These studies concentrated mainly on the feasibility of loop insertions and reorganizations. Regan and co-workers' studies showed that the loop length of simple polyglycine loops is a crucial determinant for the stability of the bundle. Their findings are in good agreement with the theoretical predictions by Thomas, who studied the relation between loop length and protein stability from a polymer physics point of view (Thomas, 1990). Other studies (e.g. Castagnoli *et al.*, 1989), however, have indicated that the physico-chemical nature of individual residues can be equally important. In the present study we combined these ideas and introduced loops of various lengths. We attempted to overcome the inevitable loss in stability by careful selection of the mending residues.

We constructed four left-handed and three right-handed Rop monomers. One of the right-handed constructs could not be isolated. The six successful constructs could be produced in milligram quantities and have been extensively characterized. As expected, our left-handed constructs are more stable than the right-handed ones. We therefore used the left-handed construct for the introduction of biologically interesting loops. In this stage it was our main goal to determine the boundaries within which we can move in this engineering project and we do not yet concentrate on specific applications. The results indicate that the Rop molecule is indeed a useful vehicle for the presentation of biologically interesting peptides. Preliminary studies show the applicability of our methodology to, for example, the presentation of inhibitor peptides to the HIV-1 proteinase.

Materials and methods

Monomer gene construction

The gene for pro-LM-Rop was kindly provided by S.C.E.Emery (Emery, 1990). We introduced the Pro59Asn mutation into LM-Rop in order to increase the stability and solubility. This mutation was introduced by a gapped-duplex single-strand mutagenesis method (Stanssens *et al.*, 1989) and the resultant

Table I. Sequences of wild-type Rop and the monomeric constructs[a]

WILD TYPE
--HHHHHHHHHHHHHHHHHHHHHHHHHHH---HHHHHHHHHHHHHHHHHHHHHHHH-
MTKQEKTALNMARFIRSQTLTLLEKLNELDADEQADICESLHDHADELYRSCLASFgddgenl

LM-Rop
--HHHHHHHHHHHHHHHHHHHHHHHHHHH---HHHHHHHHHHHHHHHHHHHHHHHH-
MTKQEKTALNMARFIRSQTLTLLEKLNELDADEQADICESLHDHADELYRSCLASF

------HHHHHHHHHHHHHHHHHHHHHHHHHH----HHHHHHHHHHHHHHHHHHHHHHHHH-
KKNGQIDEQADICESLHDHADELYRSCLASF**GGS**KQEKTALNMARFIRSQTLTLLEKLNEL**akg**

RM-Rop1
--HHHHHHHHHHHHHHHHHHHHHHHHHHH-----HHHHHHHHHHHHHHHHHHHHHHHHH-
MTKQEKTALNMARFIRSQTLTLLEKLNEL**ESKAG**QEKTALNMARFIRSQTLTLLEKLNEL

--HHHHHHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHHHHHHH-
DADEQADICESLHDHADELYRSCLASF**KKNGQI**DEQADICESLHDHADELYRSCLASFgddgenl

HGHR-LM-Rop
HH----------------HHHH
ASFKGCRSPERETFSCGQIDEQA

ACHR-LM-Rop
HH--------------------HHHH
ASFKGCEIIVTHFPFDEQNCGQIDEQA

GLOOP-LM-Rop
HH--------------------HHHH
ASFKGNDGRTPGSRNLANIPGQIDEQA

RM-Rop2
HHH------HHHHH
LNELSQSNGSQEKT

RM-Rop3
HHH-------HHHHH
LNELAGGDATKQEKT

[a]From top to bottom are listed the full sequences of wild-type Rop, one LM-Rop variant and one RM-Rop variant. The newly inserted residues are in bold. The loops that were different in further variants are underlined. For three more LM-Rop variants and two more RM-Rop variants only the varied loop regions are shown with the residues that are different underlined. Wild-type and all RM-Rop variants contain a C-terminal extension (GDDGENL), the LM-Rop variants have an introduced C-terminal (AKG); these extensions are given in lower-case characters. LM-Rop consists of A1–A2(KKNGQI)B2(GGS)B1(AKG). RM-Rop1 consists of A1(ESKAG)B1–B2(KKNGQI)A2(GDDGENL).

pro-LM-Rop-Pro59Asn will be called LM-Rop throughout this study. For this mutagenesis we used the phagemid pMa/c 5–8 system (kindly provided by H.-J.Fritz) and *Pst*I/*Eco*RI sites for insertion of the gene into this vector. The insertion variants were generated by ligation of synthetic *Nhe*I/*Bcl*I fragments into the *Nhe*I/*Bcl*I digested LM-ROP gene (all amino acid sequences are listed in Table I). All oligonucleotide syntheses were performed on a Pharmacia gene assembler.

All genes were ligated into a pEX 43-related plasmid (pEX70, wtRop gene deleted) via the *Eco*RI and a new *Pst*I endonuclease cloning sites (Castagnoli *et al.*, 1989). The RM-Rop1 gene was synthesized completely, ligated and purified according to the method of Frank *et al.* (Frank *et al.*, 1987). A synthetic gene coding for wtRop was constructed and ligated into our pEX 70 plasmid via the *Pst*I/*Eco*RI sites. Three synthetic DNA fragments, each coding for the RM-ROP helices 2 and 3 (B1–B2) were inserted into the wtRop gene via *Sac*I and *Nde*I restriction endonuclease sites and ligated. The constructs then were transformed into the *E.coli* host strain 71/72 (Castagnoli *et al.*, 1989). All cloned genes were analysed using a Pharmacia ALF sequencer.

Protein expression and purification

The monomer Rop genes were cloned into the pMAL-c fusion vector (New England Biolabs) via a PCR technique with a 21 bp non-mutagenic ATG$(N)_{18}$ primer and the pMAL-c reverse sequencing primer for generation of a semi-blunt end/*Eco*RI fragment. The constructs were subcloned, sequenced and transformed. The proteins were expressed in the host strain TB 1 on a 2 l scale as C-terminal fusions of the maltose-binding protein (MBP-Rop) using a commercially available system (New England Biolabs) and following the manufacturer's protocols. After cell lysis in a 20 mM Tris buffer, pH 8.0 with a French Press and sonification (2×1 min, 50 mW, 4°C), the samples were centrifuged using a Sorvall SS-34 rotor (15 min at 4°C and 12 000 *g*). The crude cell extracts were subjected to anion-exchange chromatography (FPLC) at 4°C on a 40 ml Q-Sepharose FF column (Pharmacia) in 20 mM Tris–HCl buffer, pH 8.0, and eluted with the same buffer with a 0–1 M NaCl gradient. The fusion protein was cleaved overnight at 20°C by 80 units of factor Xa protease (New England Biolabs) in elution buffer with 2 mM $CaCl_2$. After cleavage, the NaCl concentration in the protein solution was reduced by ultrafiltration to 150 mM with a YM-10 membrane (Amicon) and addition of Tris–HCl buffer at 4°C. For the cation-exchange chromatography of LM-Rop, the buffer was exchanged completely. LM-Rop eluate was loaded on a 20 ml Mono S column (Pharmacia) in 40 mM phosphate buffer, pH 6.4, containing 50 mM NaCl and eluted with this buffer with a 0–1 M NaCl gradient. The right-handed monomers and insertion variants were subjected, after the ultrafiltration, to anion-exchange chromatography on an 8 ml Mono Q column (Pharmacia) and eluted with a 0–1 M NaCl gradient in Tris–HCl buffer. The monomer Rop fractions were combined and concentrated to 10 ml by ultrafiltration using a YM-5 membrane. In a final gel filtration step, the proteins were separated on a Superdex G 75 (26/60) column (Pharmacia) in 40 mM

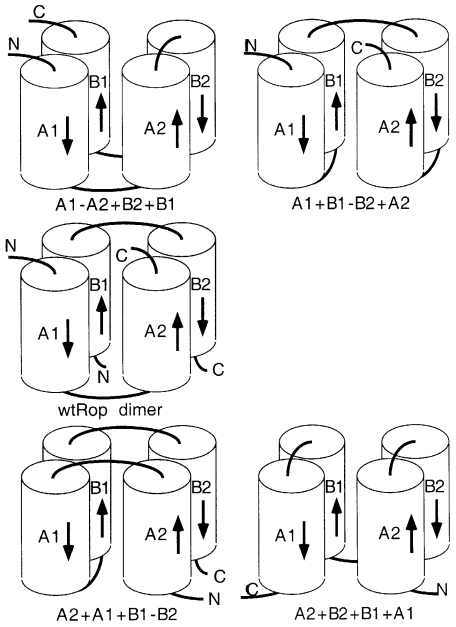

Fig. 1. Schematic representation of the four possible topologies for monomeric Rop variants. Wild-type Rop (wtRop) is shown in the middle for reference. The top two drawings are schematic representations of LM-Rop and RM-Rop. The two left-handed monomers are shown at the left and the right-handed monomers at the right. The helices are shown as cylinders and loop regions as solid lines. The termini are indicated by the letters N and C, respectively. The resulting order of the helices is indicated under the drawings; + signs indicate newly constructed loops and – signs indicate loops that are the same as in wtRop.

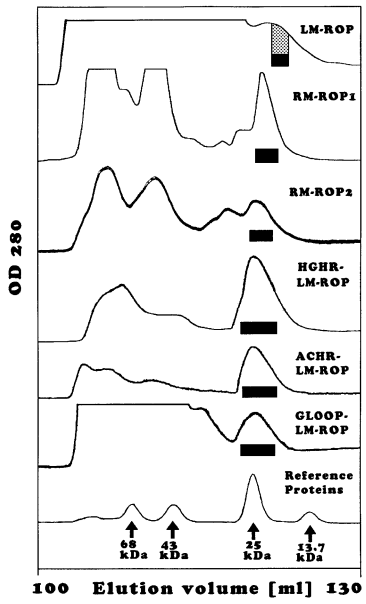

Fig. 2. Elution diagrams from gel filtration chomatography of ROP variant on a Superdex G-75 HR 26/60 column (Pharmacia). Rop protein was found in fractions marked with a black bar. For LM-Rop we used intermediately purified material and detected the elution peak immunologically.

phosphate buffer containing 200 mM NaCl at a flow-rate of 0.5 ml/min. The protein fractions were analysed by 12% SDS–PAGE according to the method of Schägger and von Jagow (1987) and stained with Coomassie Brilliant Blue. Eluate fractions were concentrated by ultrafiltration as described above. The identity of the Rop monomers was confirmed by Western blotting.

Analytical gel filtrations of partially enriched protein fractions were performed on a Superdex G 75 26/60 column under the buffer conditions described above at a flow-rate of 0.5 ml/min. A Pharmacia standard marker set and 0.5 mg each of BSA (67 kDa), ovalbumin (43 kDa), chymotrypsinogen (25 kDa) and RNase A (13.7 kDa) was used for molecular size determination.

Circular dichroism (CD) measurements in the 200–260 nm range were routinely done with a JASCO J 600 spectropolarimeter, in a 0.5 or 1.0 mm quartz cuvette at a protein concentration of 0.1 mg/ml [except for HGHR-ROP (0.25 mg/ml) and ACHR-LM-Rop (0.71 mg/ml)] in 40 mM phosphate buffer, pH 6.4 (LM-Rop) or pH 7.6 (other LM-Rop and RM-Rop variants), containing 50 mM NaCl. The secondary structure contents were calculated using the CONTIN software (Provencher, 1982).

Loops were extracted from PDB files (Bernstein *et al.*, 1977) using the WHAT IF loop search algorithm (Vriend, 1990). This algorithm searches for loops that show a good overlap with the framework of the protein for three residues at either side of the insertion (i.e. the r.m.s. misfit for the six α-carbons is <1.0 Å). We introduced the additional constraint that the loop should be short and have the potential for tight packing between the residues in the loop and the helical framework of Rop. Obviously, the chance of finding the perfect loop by database searching only is very small and some adjustments were needed in all cases. The mutations required to convert the loops that were extracted from the database into the loops that were inserted in the Rop variants were predicted using a combination of WHAT IF's mutation prediction software and visual inspection. These adaptations were kept as conservative as possible.

Results

Seven fully antiparallel, short-loop, four-helix bundles with a left- or a right-handed topology were engineered. The four left-handed (LM-Rop) and three right-handed constructs (RM-Rop) had a helix sequence A1 – A2 + B2 + B1 and A1 + B1 – B2 + A2, respectively, in which the + sign indicates a newly introduced loop. Table I shows the sequences of the wild-type and the seven monomer constructs. The Brookhaven PDB file of the wild-type Rop protein X-ray structure (1ROP; Banner *et al.*, 1987) served as the starting point for our studies.

The LM-Rop variant was created by connecting the loops in the order A1 – A2 + B2 + B1. The two new loops (indicated by a +; bold face in Table I) were taken from the proLM-Rop variant designed by Sander's group (Emery, 1990). Their design had a three-residue C-terminal extension AKG which they introduced to cap properly the fourth helix of the bundle. We kept this extension where appropriate.

The RM-Rop1 variant was created by connecting the loops in the order A1 + B1 – B2 + A2. Visual inspection of the three-dimensional models indicated very little interaction between the A1 to B1 and B2 to A2 loops. Therefore, the A2 to B2 loop from the LM-Rop variant could also be used for the B2 to A2 connection in RM-Rop1 (remember that A2 and B2 have identical sequences so that the A2–B2 and B2–A2 connections are identical). The A1 to B1 loop was obtained using the WHAT IF loop search algorithm as described in the Materials and methods section. A well-fitting loop that would insert the five residues ESKRF was found in the PDB file 1RHD (Ploegman *et al.*, 1978).

Previous studies indicated that the A2 to B2 loop in LM-

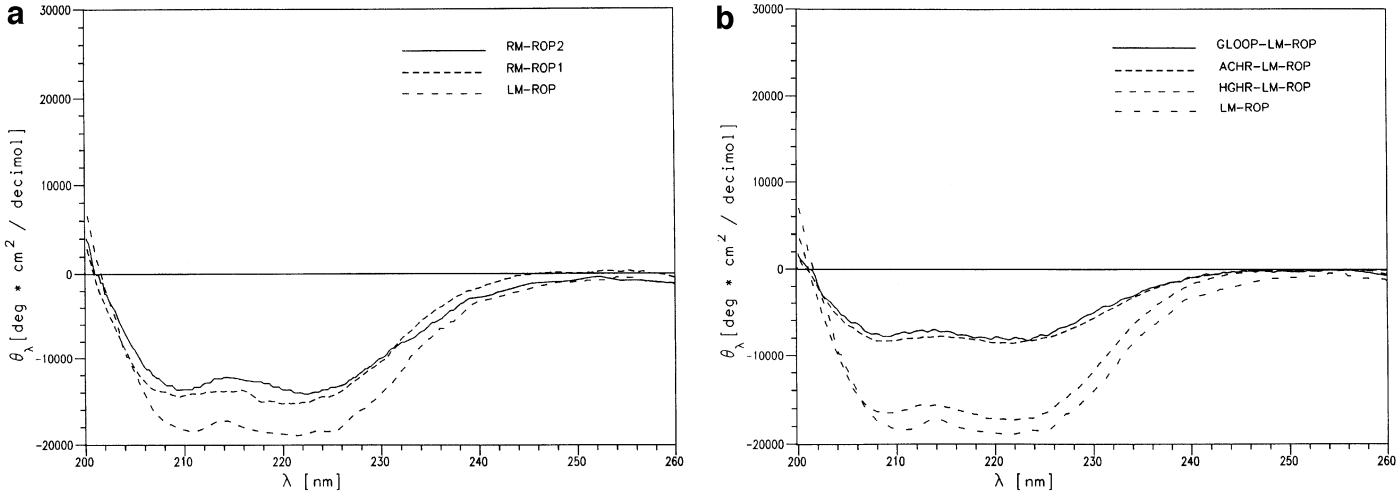

Fig. 3. Circular dichroism (CD) spectra of ROP monomers.

Table II. Results of Rop monomer gel filtration and circular dichroism analysis

Rop protein	wtRop dimer	LM-ROP	RM-ROP1	RM-ROP2	HGHR-LM-ROP	ACHR-LM-ROP	GLOOP-LM-ROP
Nominal M_r	14.3	13.7	14.2	14.3	14.8	15.0	15.3
M_r (gel filt.)	17–18[a]	19.7	22.6	23.5	24.5	24.5	24.5
Helicity (CD)	90[b]	74	68	54	71	~70[c]	N.D.
Helicity (calc.)[d]	79	83	80	79	78	75	75

[a]According to others (Kokkinidis *et al.*, 1983; Lacatena *et al.*, 1984).
[b]Value calculated with a different method as measured by Peters *et al.* (Peters *et al.*, 1997).
[c]Value determined at a different concentration. At this concentration the spectra of HGHR-LM-Rop and ACHR-RM-Rop are virtually indistinguishable.
[d]Calculation based on the helical content as indicated in Table I.

Rop can easily be modified without great loss of stability. We exploited this fact by making three insertions in this six-residue loop leading to the variants HGHR-LM-Rop (16 residues; from the human growth hormone receptor), ACHR-LM-Rop (20 residues; from the nicotinic acetylcholine receptor) and GLOOP-LM-Rop (20 residues; the immunogenic G loop from lysozyme) (Table I). These loops were chosen solely because they were biologically 'interesting' and their N- and C-terminal residues must be close to each other because of cysteine bridges that connect residues at or near these loop termini.

We decided not to vary the B2 to B1 loop because this loop connects two helix ends that are terminal in the wild-type variant. We did not mutate the A1 to A2 and B1 to B2 loops because they are native loops and we wanted to minimize the differences between the monomeric constructs and the native dimer. The B2 to A2 loop in RM-Rop1 was not varied because this loop is identical with the A2 to B2 loop in LM-Rop and variation of this loop would not provide new information.

The RM-Rop2 and RM-Rop3 variants were constructed using the loop VESNGT from 1OVO (Papamokos *et al.*, 1982) and the loop AGGDATE from 2B5C (Mathews *et al.*, 1972), respectively. Both loops were adapted manually by residue exchanges to their new environment (Table I).

The maltose binding protein fusion protein expression yielded stable Rop monomers for all constructs except GLOOP-LM-Rop, which is stable but nicked, and RM-Rop3, which could not be isolated. The cleavage in the A2 to B2 loop of GLOOP-LM-Rop could be monitored during the purification. The typical yields from a 2 l fermentation were 2–5 mg. SDS–PAGE analysis confirmed the homogeneity of the final products. All stable monomers were soluble at concentrations of $\geqslant$5 mg/ml.

Figure 2 shows the results of the Superdex gel filtrations of the Rop monomers. Table II shows the apparent molecular masses calculated from the elution volumes. Each Rop protein eluted as a single, well-defined peak, and LM-Rop was determined immunologically in the fractions. The GLOOP-LM-Rop also eluted as a single peak, but a subsequent SDS–PAGE indicated that cleavage had occurred in the A2 to B2 loop. The wild-type Rop variant (nominal molecular mass of the dimer: 14.3 kDa) displayed an apparent molecular mass of 17–18 kDa in the gel filtrations of earlier workers (Lacatena *et al.*, 1984; Kokkinidis *et al.*, 1993). Obviously, the rod-shaped molecular structure of Rop leads to unusual chromatographic behaviour. The LM-Rop molecule shows a slightly (~2 kDa) higher apparent molecular mass compared with wtRop. The apparent molecular masss of our right-handed monomers and insertion variants are increased by several kDa (Figure 2, Table II).

The purified Rop monomers were characterized by CD to determine the influence of the different loop rearrangements on their secondary structure (Figure 3, Table II). All variants showed a typical highly α-helical spectrum similar to wtRop. The LM-Rop variant had the highest α-helix content, followed by the insertion variant HGHR-LM-Rop. ACHR-LM-Rop was measured at a high concentration, where LMR-HGHR-Rop had a virtually identical CD spectrum. The cleaved GLOOP-LM-Rop had a reduced secondary structure content. RM-Rop1 has a similar and RM-Rop2 a significantly reduced α-helix content compared with the left-handed variants.

Discussion

Using the results from early Rop studies, Sander proposed the 'core hypothesis'. 'According to this 'core hypothesis', the precise sequence in the loop regions and on the protein surface is of secondary importance, provided that it is consistent with loop formation and surface solvation' (Sander, 1994). Other studies by Regan's group (Regan and DeGrado, 1988; Nagi and Regan, 1997) indicate that the loop length in helix bundles is more important for the stability than is suggested by the core hypothesis. We conclude from our Rop monomer experiments that loop length and residue composition are important for the A1 to B1 loop in RM-Rop variants, whereas the structural integrity seems rather independent of loop length in the A2 to B2 loop in LM-Rop variants. For the A1 to B1 loop variants of RM-Rop a loop length effect with another optimum than that observed by Regan's group (Predki and Regan, 1995) could be supplemented by strong influences of our three individual loop sequences. Consequently, we found different optimal A1–B1 and B2–A2 loop lengths than did Regan's group. Obviously the actual sequence that is introduced plays a crucial role. Introduction of a proline residue into the wtRop loop destabilized the whole molecule (Peters *et al.*, 1997). Castagnoli *et al.* found that a reorganization of one helix end could compensate for the deletion of five residues (Asp30–Gln34) in the wtRop loop region (Castagnoli *et al.*, 1989).

Our studies on the LM-Rop variants suggest that if the residues directly adjacent to the helix are carefully selected, the loop can tolerate a wide variety of loop insertions.

Predki and Regan designed a right-handed Rop monomer variant like we did, but they had to introduce a stabilizing Asp30Gly mutation, leading to an increase in the T_m of ~13°C (Predki and Regan, 1995). Despite this stabilization, their initially designed protein showed strong aggregation behaviour. Consequently, Regan's group varied the lengths of the two newly inserted loops. For the A1–B1 loop they found that the six-residue GGGGTK loop (using GGGGTK where we use ESKAG; see Table I) yielded a monomer Rop of higher stability (they measured the melting temperatures, helix content and molecular volumes) than A1–B1 loops with five or seven residues. However, we found that the five-residue ESKAG A1–B1 loop of RM-Rop1 yielded more compact structures with higher helix content than the six-residue SQSNGS A1–B1 loop of RM-Rop2 or the seven-residue AGGDATK construct in the presumably highly unstable RM-Rop3. For the B2–A2 loop, Predki and Regan found that the five-residue loop GGGGA (using GGGGA where we use KKNGQI with six residues; see Table I) yielded a structure of higher stability than loops with more or fewer glycines. Our B2–KKNGQI–A2 helix–loop–helix structure module in RM-Rop has the same peptide sequence as the A2–KKNGQI–B2 structure in LM-Rop where we used this peptide and longer structures for loop insertion.

These differences once again confirm the importance of the actually used amino acids. Loop length surely plays an important role, but the loop length effects are supplemented by the interactions made by the loop residues.

It can be concluded that the influence of the loops on the Rop monomer secondary structure and the apparent molecular volume is small provided that the new loops are designed carefully. Further studies have to reveal if the observations made in this study and the studies by Regan's group regarding loop lengths are general rules or whether those loops for which these observations were made will also tolerate many more modifications once we have found the ideal residues. In summary, the Rop monomer can serve well as a vehicle for the presentation of bioactive peptides in an *in vivo* system. Preliminary studies, in which HIV-1 proteinase-inhibiting peptides derived from proteinase recognition sequences were inserted into the A2–B2 loop context of an HGHR-LMR variant, seem to confirm this conclusion. We constructed seven new monomers by introduction of heptamer peptides into the loop with flexible three-residue linkers on each side. The linker sequences and the overall loop lengths were homologous with the HGHR-LMR. Upon expression in *E.coli*, five of the seven monomers were biologically stable.

Acknowledgements

The authors gratefully acknowledge the work of Stephen C.Emery who constructed the gene of the proLM-Rop variant, which we mutagenized to the LM-Rop. H.P.K. acknowledges financial support from the EC.

References

Abdullaev,Z.K., Latypov,R.F., Badretdinov,A.Y., Dolgikh,D.A., Finkelstein,A.V., Uversky, V.N. and Kirpichnikov, M.P. (1997) *FEBS Lett.*, **414**, 243–246.
Ay,J., Hahn,M., Decanniere,K., Piotukh,K., Borriss.,R. and Heinemann,U. (1998) *Proteins*, **30**, 155–167.
Banner,D.W., Kokkinidis,M. and Tsernoglou,M. (1987) *J. Mol. Biol.*, **196**, 657–675.
Bernstein,F.C., Koetzle,T.F., Williams,C.J.B., Meyer,E.F.,Jr., Brice, M.D., Rodgers, J.R., Kennard,O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.*, **112**, 535–542.
Bryson,J.W., Betz,S.F., Lu,H.S., Suich,D.J., Zhou,H.X., O'Neill,K.T. and DeGrado,W.F. (1995) *Science*, **270**, 935–941.
Buchwalder,A., Szadkowski,H. and Kirschner,K. (1992) *Biochemistry*, **31**, 1621–1630.
Castagnoli,L., Scarpa,M., Kokkinidis,M., Banner,D.W., Tsernoglou,D. and Cesareni,G. (1989) *EMBO J.*, **8**, 621–629.
Cesareni,G. Helmer Citterich,M. and Castagnoli,L. (1991) *Trends Genet.*, **7**, 230–235.
Eberle,W., Klaus,W., Cesareni,G., Sander,C. and Rosch,P. (1990) *Biochemistry*, **29**, 7402–7407.
Emery,S.C.E. (1990) *PhD Thesis*, Sheffield.
Frank,R., Meyerhans,A., Schwellnus,K. and Blöcker,H. (1987) *Methods Enzymol.*, **154**, 221–249.
Hynes,T.R. (1989) *Nature*, **339**, 73–76.
Kokkinidis,M., Vlassi,M., Papanikolaou,Y., Kotsifaki,D., Kingswell,A., Tsernoglou,D. and Hinz,H.-J. (1983) *J. Mol. Biol.*, **170**, 1059–1060.
Lacatena,M., Banner,D.W., Castagnoli,L. and Cesareni,G. (1984) *Cell*, **37**, 1009–1014.
Lacroix,E., Viguera,A.R. and Serrano,L., (1998) *Fold. Des.*, **3**, 79–85.
MacBeath,G., Kast,P. and Hilvert,D. (1998) *Science*, **279**, 1958–1961.
Mathews,F.S., Argos,P. and Levine,M. (1972) *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 387–395.
Nagi,A.D. and Regan,L. (1997) *Fold. Des.*, **2**, 67–75.
Olszewski,K.A., Kolinski,A. and Skolnick,J. (1996) *Protein Eng.*, **9**, 5–14.
Papamokos,E., Weber,E., Bode,W., Huber,R., Empie,M.W. and Kato,I. (1982) *J. Mol. Biol.*, **158**, 515–537.
Peters,K., Hinz,H.-J. and Cesareni,G. (1997) *Biol. Chem.*, **378**, 1141–1152.
Ploegman,J.H., Drent,G., Kalk,K.H. and Hol,W.G.J. (1978) *J. Mol. Biol.*, **123**, 557–594.
Predki,P.F. and Regan,L. (1995) *Biochemistry*, **34**, 9834–9839.
Provencher,S.W. (1982) *Comput. Phys. Commun.*, **27**, 229–242.
Regan,L. and DeGrado,W.F. (1988) *Science*, **241**, 976–978.
Sander,C. (1994) *Trends Biotechnol. Sci.*, **12**, 163–167.
Schägger,H. and von Jagow,G. (1987) *Anal. Biochem.*, **166**, 368–379.
Stanssens,P., Opsomer,C., McKeown,Y.M., Kramer,W., Zabeau,M. and Fritz,H.-J. (1989) *Nucleic Acids Res.*, **17**, 4441–4454.
Thomas,D. (1990) *J. Mol. Biol.*, **216**, 459–465.
Vriend,G. (1990) *J. Mol. Graphics*, **8**, 52–56.
Walsh,ST, Cheng H., Bryson J.W., Roder,H. and DeGrado, W.F. (1999) *Proc. Natl Acad. Sci.*, **96**, 5486–5491.
Zhang,T., Bertelsen.,E, Benvegnu,D. and Alber,T., (1993) *Biochemistry*, **32**, 12311–12318.

Received September 7, 2000; revised May 31, 2001; accepted June 18, 2001