Structure-function studies of an engineered scaffold protein derived from stefin A. I: Development of the SQM variant

Toni Hoffmann[1,5], Lukas Kurt Josef Stadler[1],
Michael Busby[2], Qifeng Song[1], Anthony T.Buxton[1],
Simon D.Wagner[3], Jason J.Davis[2] and Paul Ko
Ferrigno[1,4,5]

[1]Section of Experimental Therapeutics, Leeds Institute of Molecular
Medicine, St James's University Hospital, Beckett St, Leeds LS9 7TF, UK,
[2]Chemical Research Laboratory, University of Oxford, Mansfield Road,
Oxford OX1 3TA, UK, [3]MRC Toxicology Unit and Department of Cancer
Studies and Molecular Medicine Hodgkin Building, University of Leicester,
Lancaster Road, Leicester LE1 9HN, UK and [4]MRC Cancer Cell Unit,
Hutchinson MRC Research Centre, Hills Road, Cambridge CB22 0XZ, UK

[5]To whom correspondence should be addressed.
E-mail: p.koferrigno@leeds.ac.uk (P.K.F.) and t.hoffmann@leeds.ac.uk (T.H.)

**Non-antibody scaffold proteins are used for a range of
applications, especially the assessment of protein–protein
interactions within human cells. The search for a versatile,
robust and biologically neutral scaffold previously led us
to design STM (stefin A triple mutant), a scaffold derived
from the intracellular protease inhibitor stefin A. Here, we
describe five new STM-based scaffold proteins that
contain modifications designed to further improve the ver-
satility of our scaffold. In a step-by-step approach, we
introduced restriction sites in the *STM* open reading
frame that generated new peptide insertion sites in loop 1,
loop 2 and the N-terminus of the scaffold protein. A
second restriction site in 'loop 2' allows substitution of the
native loop 2 sequence with alternative oligopeptides.
None of the amino acid changes interfered significantly
with the folding of the STM variants as assessed by circu-
lar dichroism spectroscopy. Of the five scaffold variants
tested, one (stefin A quadruple mutant, SQM) was chosen
as a versatile, stable scaffold. The insertion of epitope tags
at varying positions showed that inserts into loop 1,
attempted here for the first time, were generally well toler-
ated. However, N-terminal insertions of epitope tags in
SQM had a detrimental effect on protein expression.**
Keywords: circular dichroism spectroscopy/epitope
recognition/peptide aptamer/scaffold protein

Introduction

Although antibodies have proven to be useful tools in a
broad range of applications, they also possess some inherent
limitations. Owing to their complex structure and extensive
glycosylation, antibodies are fragile molecules with low
yields in bacterial expression systems, having a high temp-
erature sensitivity, and expensive production costs. Their
correct folding relies on the formation of intramolecular dis-
ulphide bonds and as a consequence they frequently do not
function in the reducing environment inside eukaryotic cells
(Skerra, 2007). Furthermore, it is often very challenging or
impossible to raise antibodies that are able to distinguish
between different isoforms or mutants of proteins, an often
wished for tool in clinical diagnostics.

It is therefore not surprising that for more than 14 years a
great deal of effort has been invested in the design of non-
antibody scaffold proteins that can present recognition sur-
faces that expand the immune repertoire. This has resulted in
the description of over 40 artificial proteins that have been
engineered to present peptides as recognition moieties
(reviewed in Binz *et al.*, 2005). Major classes of engineered
scaffolds include those based on folds from proteins as
diverse as protein Z (*Affibodies*, Nord *et al.*, 1995), fibronectin
(*Adnexins*, Koide *et al.*, 1998), ankyrin repeat proteins
(*DARPins*, Binz *et al.*, 2003; Kohl *et al.*, 2003), cysteine-knot
miniproteins (*Knottins*, Clark *et al.*, 2006; Kolmar, 2008) or
Armadillo repeat proteins (Parmeggiani *et al.*, 2008), as well
as on full-length proteins such as lipocalins (*Anticalins*,
Schlehuber and Skerra, 2005), ColE7 immunity protein (*Im7*,
Juraja *et al.*, 2006), the green fluorescent protein GFP (Abedi
et al., 1998) and thioredoxin A (LaVallie *et al.*, 1993). A
major rationale of this work has been that conformational con-
straints enforced on inserted peptides by the scaffold protein
will decrease the entropic cost of binding and lead to
increased binding affinities (Ladner, 1995). A second motiv-
ation has been the idea that whereas free peptides are readily
degraded by cellular proteases, full-length folded proteins are
expected to be more stable (Saveanu *et al.*, 2002; Reits *et al.*,
2003). Similarly, antibodies that have evolved to function in
the extracellular environment function poorly within eukary-
otic cells. These considerations led Colas *et al.* to establish a
yeast two-hybrid screening system to identify conformation-
ally constrained peptides that would bind to target proteins
within the intracellular environment of a eukaryotic (yeast)
cell, coining the term 'peptide aptamer' to describe such tools
(Colas *et al.*, 1996; Colas, 2008). In addition to their ability to
function inside cells, peptide aptamers possess several attri-
butes in common with other constrained peptides that dis-
tinguish them from antibodies, including small size (typically
10–15 kDa, compared with 150 kDa for IgG), lack of post-
translational modification (IgG requires extensive glycosyla-
tion of the constant region for stability) and being comprised
of a single polypeptide chain (IgG is comprised of two copies
each of a heavy and a light polypeptide chain). Nonetheless,
peptide aptamers posses binding affinities that are comparable
to most antibodies, in the pM–nM range (Cohen *et al.*, 1998).

A consequence of the simple design is a high protein yield in bacterial expression systems, increased thermal and chemical stability as well as the ability to functionally fuse with reporter proteins (Kolmar and Skerra, 2008).

A major concern with any approach is that the introduction of a scaffold protein itself into cells may lead to phenotypes that may complicate or mask the effects of the displayed peptide. For example, thioredoxin was chosen because the use of a bacterial protein scaffold was thought unlikely to produce interactions with proteins in human cells. However, expression of bacterial thioredoxin protein is anti-apoptotic and is able to decrease ischemia in mouse models (Tao *et al.*, 2004).

Accordingly, we have sought to produce a new scaffold for the presentation of constrained peptides that would be 'biologically neutral', i.e. would not interact with human proteins. After an initial investigation of several candidates, we found that human stefin A (SteA, also called cystatin A) could be engineered to possess all of the desirable characteristics (Woodman *et al.*, 2005). Stefin A is a small (98 amino acid) monomeric protein inhibitor of the cystatin family I (stefins) that inhibits cysteine proteases of the cathepsin family (Turk *et al.*, 1986). It interacts with its partner proteins, cathepsins B, C, H, L and S (Brzin *et al.*, 1984; Green *et al.*, 1984; Abrahamson *et al.*, 1986) using three sites, with key contacts made by glycine at position 4, valine at position 48 and lysine at position 73 (Bode *et al.*, 1988; Stubbs *et al.*, 1990; Martin *et al.*, 1994, 1995; Tate *et al.*, 1995; Pavlova and Björk, 2002; Jenko *et al.*, 2003). Biological neutrality was achieved through a combination of two rational mutations [G4W to abolish interaction with cathepsins (Estrada *et al.*, 1998, 1999) and V48D to abolish interaction with cathepsins and reduce dimer formation through domain swapping (Japelj *et al.*, 2004)] and a third mutation to introduce a unique *Rsr*II restriction site at codons 71–73, which becomes the insertion site for oligonucleotides encoding short peptides. The use of this site allowed us to introduce peptides into the longest loop found in SteA, known as loop 2 (Bode *et al.*, 1988; Stubbs *et al.*, 1990; Martin *et al.*, 1994, 1995; Tate *et al.*, 1995; Pavlova and Björk, 2002). The engineered protein was designated STM, for stefin A triple mutant (Woodman *et al.*, 2005).

However, despite the success of this scaffold in presenting certain peptides for interaction (Woodman *et al.*, 2005; Estrela *et al.*, 2008; Evans *et al.*, 2008; Johnson *et al.*, 2008; Shu *et al.*, 2008; Davis *et al.*, 2009), there was clear room for further improvement. Most importantly in our view, crystal structures of stefin A in complex with its target cathepsin revealed that the parental SteA protein uses three noncontiguous peptide surfaces, comprising the amino terminus, loop 1 and loop 2, to interact with its target (Bode *et al.*, 1988; Stubbs *et al.*, 1990; Martin *et al.*, 1994, 1995; Tate *et al.*, 1995; Pavlova and Björk, 2002; Jenko *et al.*, 2003). We wished to make use of the whole of this 'tripartite wedge' for interaction with targets, so as to maximize both binding affinity and the specificity of interaction.

The goal of the present study is to assess the effect on the SteA protein of introducing step-by-step amino acid changes at specific locations, alone and in combination, that would result from the generation of restriction sites in the open reading frame to be used for oligonucleotide insertion. Having identified a number of potential changes that do not adversely affect the folding of the engineered protein as assessed by circular dichroism (CD) spectroscopy, we asked whether a series of model peptides could be presented by the newly engineered scaffold protein, a stefin A quadruple mutant (SQM). We found (i) that inserts into the amino terminal region that mediates most of the protein contacts in SteA co-crystals frequently disturb the secondary structure of the protein, as analysed by CD spectroscopy and protein expression yields in *E.coli*; (ii) that replacement of loop 2 in SQM may adversely affect protein yields compared with the longer inserts that result from insertion into loop 2 in STM; and most surprisingly (iii) that short inserts into both loops 1 and 2 simultaneously can increase the apparent stability of the protein. Together, these data suggest new ways of working with stefin A derived scaffold proteins and may shed light on determinants of folding and stability of the stefin A protein itself.

Materials and methods

Plasmids and DNA manipulation

pET30a(+), for the expression of His$_6$-tagged proteins in bacteria, was purchased from Novagen (Nottingham, UK). The SteA variant STM in pET30a(+) was used as published (Woodman *et al.*, 2005). DNA manipulations followed standard protocols (Sambrook and Russell, 2001), using enzymes obtained from NEB (MA, USA). Oligonucleotides were from Invitrogen (Paisley, UK) and are listed in Supplementary Table B (Supplementary data are available at *PEDS* online). The genes encoding the STM variants SDM (stefin A double mutant), SQM, SUM (stefin A unique middle), SUN (stefin A unique N-terminus) and SUC (stefin A unique C-terminus) were synthesized and cloned into pET30a(+) by Genscript (Piscataway, NJ, USA). Site-directed mutagenesis was performed according to Fisher and Pei (1997). All DNA manipulations were confirmed by sequencing.

Insertion of peptides into the SteA-based scaffolds

Double-stranded oligonucleotide cassettes flanked by required restriction site overhangs and encoding a peptide tag (Supplementary Table A, Supplementary data are available at *PEDS* online) were made by annealing oligonucleotides (Supplementary Table B, Supplementary data are available at *PEDS* online). Digested dsDNA cassettes were ligated into the appropriate restriction sites of the scaffold-encoding open reading frame in pET30a(+).

Production of STM variant recombinant proteins in E.coli

pET30a(+) STM variants were transformed into the *E.coli* strain BL21 (DE3, Novagen, USA) that provides increased protein stability due to its Ion and OmpT deficiency (Shaw and Ingraham, 1967). The cells were grown to $A_{600} = 0.6$ in Luria–Bertani broth (Sambrook and Russell, 2001) at 37°C, and protein expression was induced with 0.4 mM IPTG. After 3 more hours of growth, the cells were harvested and resuspended in protein lysis buffer [50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, 300 mM NaCl, supplemented with a complete protease inhibitor mix lacking EDTA (Roche)]. The cells were lysed using a cell disrupter (Constant Systems, Northants, UK). The His$_6$-tagged proteins were captured on Ni-NTA columns (Qiagen) and, after extensive washing of

the columns with lysis buffer and lysis buffer supplemented with 30 mM imidazole, eluted with 250 mM imidazole into a 50 mM sodium phosphate, 300 mM NaCl buffer.

Far-UV CD spectroscopy

Purified proteins were transferred into 50 mM sodium phosphate (pH 7.4) using buffer exchange columns (Amicon Ultra 10 kDa, Millipore), and their concentrations determined by measuring the absorbance at 280 nm (NanoDrop ND-8000, Thermo Scientific). CD spectra were recorded on a Jasco J715 spectropolarimeter at $10°C$ using a 300-μl quartz cuvette with $d = 0.1$ cm path length. Folding spectra were collected from 190 to 260 nm. The raw output is given in ellipticity [θ (mdeg)]. The data were normalized by calculating the mean residue ellipticity θ using the following equation:

$$[\theta]_\lambda = \frac{\theta \times 100 \times M_r}{c \times d \times n} \quad (1)$$

where $[\theta]_\lambda$ is the mean residue ellipticity (deg $\times$ cm^2 $\times$ dmol^{-1}), θ the observed ellipticity (in mdeg) at wavelength λ (in nm), M_r the molecular weight of the peptide (in g/mol), c the concentration (in mg/ml), d the path length (in cm) and n the number of residues. Three to eight spectra were taken for each STM variant and averaged, and the average spectrum for the buffer alone was subtracted to produce the final curves. The data were analysed using *Microsoft® Excel®* (version 12.5.1 for Mac OS) and visualized using *Plot* (version 0.997, Michael Wesemann, http://plot.micw.eu/).

Immunoprecipitation

Myc-tagged SQM peptide aptamers were immunoprecipitated using anti-Myc tag antibody-coated agarose resin (Abcam). About 10 μl of the resin was blocked in 250 μl of 50 mM sodium phosphate (pH 7.4) with 4% bovine serum albumin (BSA) and 0.1% Nonidet P-40 (NP-40, Calbiochem) for 1 h at $4°C$. The resin was then washed three times in wash buffer (50 mM sodium phosphate with 0.05% BSA and 0.1% NP-40). Subsequently, the resin was incubated with 2 μg of purified peptide aptamer dissolved in 200 μl of WB for 2 h at $4°C$, followed by seven washes in 200 μl WB. Protein sample buffer was added (Laemmli, 1970), the sample boiled for 5 min and analysed by SDS–polyacrylamide gel electrophoresis followed by western blotting with anti-S-tag monoclonal antibody (Novagen) targeting the S-tag added by the pET30a(+) vector to the amino-terminus of the peptide aptamer.

Microarray experiments

Microarray assays were performed with antibodies (Ab9106 Myc-tag polyclonal, Ab16 918 HA tag monoclonal and Ab24620 AU1 tag monoclonal, from Abcam) labelled 1:1 with Atto dyes (Atto 550-NHS ester and Atto 647N-NHS ester). Labelling was confirmed via a Nanodrop spectrophotometer. All samples have been spotted (BioOdessy, Biorad Corp.) on nickel NTA histag affinity (Xenopore Corp) surfaces using 100 μm capillary pins. Samples were printed at concentrations of 10 μM in print buffer, which comprised of PBST and 10% glycerol. Sample arrays were printed in repeats of four and the entire array was repeated three times across the slide. Spotted volumes were allowed to incubate on the slide for 60 min, prior to a blocking step in 1% BSA

PBST for 60 min. Labelled antibodies were incubated at concentrations of 7–28 nM in 1% BSA PBST for 40 min using 50 μl volume Lifterslips (Thermo Scientific). Slides were washed in PBST twice for 10 min then twice in deionized water for 2 min. The slides were dried under a stream of nitrogen and scanned at a resolution of 5 μm with 543 and 633 nm lasers under the Cy3 and Cy5 detection protocol (Scan Array Express, Perkin Elmer).

Results

Although STM can be used to present peptide aptamers that interfere with biological processes (J.T-H. Yeh, R. Woodman, and P. Ko Ferrigno, unpublished), we found that screens of random peptide libraries created by insertion into loop 2 frequently led to the identification of truncated peptide aptamers. These were produced by oligonucleotides that encoded either in-frame or out-of-frame stop codons, leading to the loss of the C-terminal 23 residues of stefin A (Woodman *et al.*, 2005). We reasoned that the selection of these 'unconstrained' peptides from libraries may result from an inability of the scaffold to tolerate or correctly present a wide range of peptide sequences inserted into loop 2. To circumvent this problem and to increase the available surface area used for interaction with targets, we set out to further engineer the scaffold protein.

Generation of SteA mutants and selection of an improved scaffold protein

We set out to improve the STM scaffold, shown schematically in Fig. 1A. We used a step-by-step approach introducing successive mutations into the *stefin A* gene to generate new insertion sites for oligonucleotides because it was unknown whether any of the planned modifications would be tolerated by the encoded protein. In doing so, we were able to examine the effect of single amino acids changes at different positions, as well as the effects of those changes in combination with others. First, we introduced mutations at unique sites in the wild-type *stefin A* gene.

These engineered proteins were designated SUN (encoded by the *SteA* open reading frame possessing an *Avr*II site at codon 4) and SUM (with an *Nhe*I site at codons 48–50 inclusive). Second, a double mutant was generated, designated SUC (two *Rsr*II sites in codons 71–73 and 82–83 inclusive). Third, a triple mutant designed to allow insertions into loops 1 and 2 was designated SDM (with *Nhe*I at codons 48–50 and two *Rsr*II sites at codons 71–73 and 82–83). Finally, we conceived a quadruple mutant, SQM that combines all of these changes in one open reading frame. The corresponding changes in the amino acid sequences are shown in Fig. 1B. SQM represents a version of the desired final scaffold including all modifications that would allow the insertion of short random peptides into the N-terminus and loop 1 as well as the replacement of loop 2. This combination is designed to provide a high versatility and binding affinity.

We used two methods to determine the effect of the amino acid changes on the scaffold structure. The first method is simply to determine the relative expression levels of the engineered proteins in *E.coli*, with the rationale that most amino acid changes are likely to destabilize the protein and decrease yields of soluble protein. The second method is to
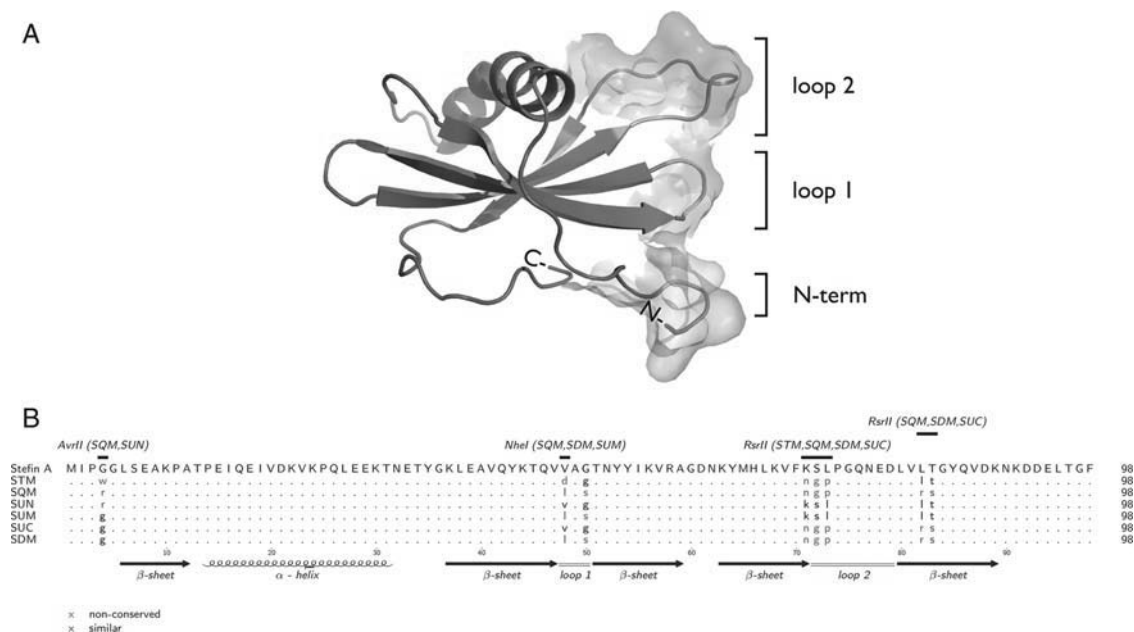
A

B

Fig. 1. Engineering novel scaffolds from stefin A. (**A**) Schematic representation of the structure of the engineered SQM protein. A three-dimensional representation of the SQM protein, using the SWISS-MODEL modelling service (Guex and Peitsch, 1997; Schwede *et al.*, 2003; Arnold *et al.*, 2006) using the stefin A structure as a template (PDB ID: 1nb5k; Jenko *et al.*, 2003). The data were visualized using MacPyMol v0.99 (http://pymol.org/). The α-helix is coloured in red, β-sheets in blue and insertion sites in green. The surface that is potentially covered by inserts is shown in grey. (**B**) Sequence alignment of stefin A-derived scaffold proteins. The amino acids are given in the single-letter code. The positions of the last residue in the alignments are listed in the right-hand column. Secondary structures taken from the model are indicated below the alignment as a coil (α-helix), black arrows (β-sheets) and double lines (loops). The restriction sites (*Avr*II, *Nhe*I, *Rsr*II) we introduced into the open reading frame to allow the insertion of peptide-encoding oligonucleotides are shown on the top and the resulting amino acid changes are shown in lower case. Note that all the stefin A variants used in this work are expressed as HexaHistidine fusion proteins from the pET30a(+) vector (Novagen) and include a 51 residue amino-terminal extension into which we have added a cysteine residue at position 14, seven residues after the His tag, and a StrepII tag (Davis *et al.*, 2007). This extension is not shown in this alignment. The alignment file was produced using ClustalX (Thompson *et al.*, 1997) and visualized by T$_E$Xshade (Beitz, 2000). The colour version of this figure is available at PEDS online.

assess the proportion of secondary structure elements in each protein isoform by CD spectroscopy.

When comparing protein yields from *E.coli*, we measured the amount of purified stefin A variant that could be recovered on Ni-NTA affinity columns of a given volume and standardized this number to the volume of starting culture. We found that each individual change from stefin A either had little effect on normalized yields or occasionally led to increased yield compared with STM. Thus, in five preparations of STM, the average yield was 46 ± 16 mg of purified STM from 1 l of culture, whereas for SUN the average yield was 42 ± 14 mg/l of culture ($n = 3$), SUM 31 ± 8 mg/l of culture ($n = 3$) and SUC 53 ± 9 mg/l of culture ($n = 3$). This was true also when we combined two mutations in one protein, as the yield for SDM was 46 ± 7 mg/l of culture. When we made changes at all three sites simultaneously in SQM, the yield was 49 ± 11 mg of purified protein per litre of culture (average of seven preparations).

CD spectroscopy allows an assessment of secondary structure elements (α-helix and β-sheet) present in a given protein, but can neither readily distinguish between the two nor can CD spectroscopy be used to predict secondary structure *ab initio*. However, we believe CD spectroscopy to be a useful tool for determining whether or not amino acid alterations or peptide insertions have perturbed the secondary structure of the well-characterized stefin A protein as multiple spectra for the parental protein are available (Turk *et al.*, 1983; Zerovnik *et al.*, 1991; Pol *et al.*, 1995; Jerala and Zerovnik, 1999; Kenig *et al.*, 2001). In using CD spectroscopy, we wish simply to ask whether the proportion of

secondary structure remains constant following each engineering step. A change in the shape of the CD spectrum plot would reflect changes in the α-helical and β-sheet content of the folded protein. CD spectra for the different STM variants were obtained between 200 and 260 nm STM (Fig. 2). All STM variants showed similar CD spectra with an inflexion point at around 219 nm. The CD spectrum of STM appears to be slightly shifted towards the far UV compared with other variants. However, we also observed differences in the amplitude of the CD spectra with SQM, SUM and SUN showing a deeper curve compared with STM (Fig. 2; see Discussion).

SQM presents peptides in loop 2 differently than STM

In generating SQM from STM, an additional *Rsr*II site was introduced into the loop 2 encoding region of the STM open reading frame. As a consequence, the leucine at position 82 and threonine at 83 were changed to arginine and serine, respectively. The two *Rsr*II sites allow the substitution of the native 'loop 2' sequence with a random oligopeptide. This was expected to be less disruptive compared with STM where the use of a single insertion site inevitably extends loop 2, but may also affect the way that peptide inserts are presented.

Different oligonucleotides were inserted into each open reading frame so as to present encoded peptides at loop 2 of both the STM and SQM scaffolds. Initially, we inserted oligonucleotides that encode short peptides (A48, A52 and A58, each 10 amino acids long; and A7: 17 amino acid long; Supplementary Table A, Supplementary data are available at
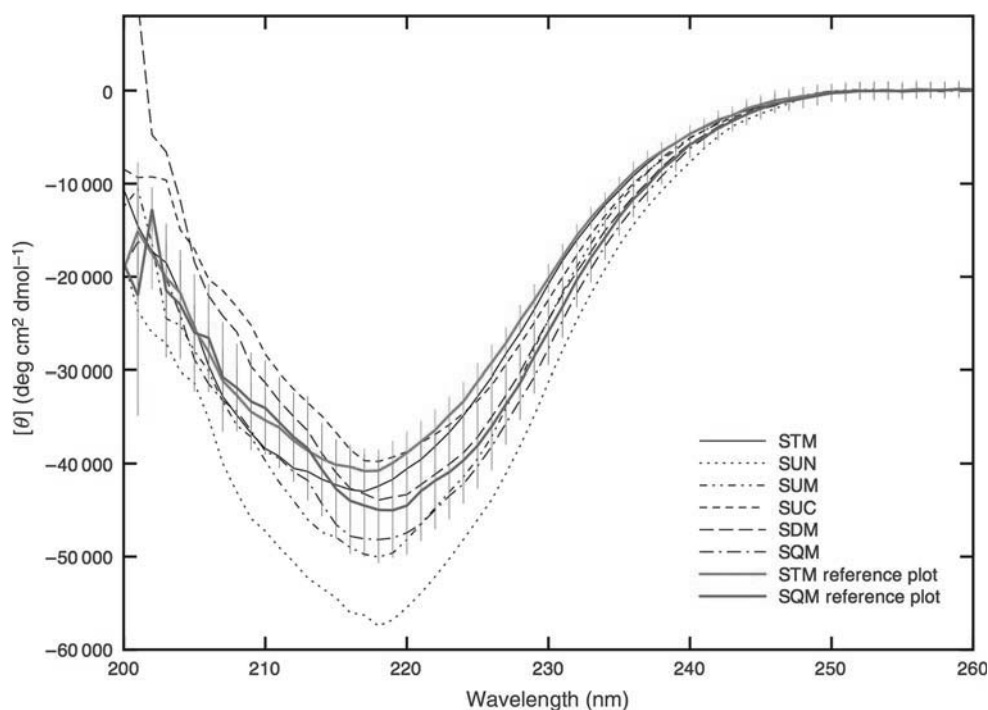
Fig. 2. Far-UV CD spectra of different STM-based scaffold proteins. The far-UV CD spectra of several scaffold proteins that were designed based on the STM scaffold are shown (STM, black line; SUN, dotted line; SUM, streak-double-dotted line; SUC, short streaked line; SDM, long streaked line; SQM, streak-dotted line). The data are expressed as mean residual ellipticity (MRE, $[\theta]$). Vertical error bars indicate the standard deviation of five STM (red) or seven SQM (blue) CD spectra obtained from independent sample preparations. The colour version of this figure is available at PEDS online.

PEDS online) that bind to the POZ domain of BCL-6 when displayed by the thioredoxin A scaffold (Chattopadhyay *et al.*, 2006). In the course of cloning, concatamerization of two oligonucleotides gave rise to a longer peptide, designated A52tandem, comprising 22 amino acids. Expression levels of these peptide aptamers in SQM were comparable to, or slightly exceeded, the expression levels obtained when using STM. Interestingly, one oligopeptide (A7) caused a deviation from the reference CD spectra of 'empty' STM (Fig. 3A) but not of empty SQM (Fig. 3B). The data are consistent with the lower yield of STM-A7 (23 mg protein/l of culture) compared with SQM-A7 (41 mg/l of culture) when expressed in *E.coli* and indicate that, at least in the case of A7, SQM may present peptides in loop 2 differently than STM.

We then turned to a model set of peptide inserts (pep2, pep6, pep10m: 20 amino acids; pep9: 42 amino acids; Supplementary Table A, Supplementary data are available at PEDS online) that were initially identified by Colas *et al.* (1996) as being able to bind to human CDK2 when presented by the thioredoxin A scaffold. To our surprise, the yield of protein from bacterial cultures was now lower when the peptides were in SQM compared with STM (STM and SQM: pep2, 19 and 9 mg/l; pep6, 36 and 2 mg/l; pep10m, 58 and 2 mg/l; pep9, 89 and 2 mg/l). Collected CD spectra for these peptide aptamers also showed a decreased tolerance for insertions in 'loop 2' of SQM than in STM (Fig. 3C and D). These data confirm that peptides in loop 2 of SQM behave differently than peptides inserted in loop 2 of STM.

SQM can present peptide inserts for interaction with target proteins

We wished to make a preliminary assessment of whether SQM would be able to present peptides for interaction.

Accordingly, we turned to our previous strategy (Woodman *et al.*, 2005) of asking whether simple epitope tags would be recognized by their cognate antibodies when presented in the new sites created in SQM. We chose three peptide epitopes (AU1, HA and Myc tags) that differ in both length and physico-chemical characteristics (Supplementary Table A, Supplementary data are available at *PEDS* online). These peptides were inserted singly or in various combinations into the positions we have designed in the SQM scaffold (the N-terminus, loop 1 or loop 2). Initially, we inserted the HA tag into the amino terminal site, the AU1 tag (the shortest peptide) into loop 1 (the shortest loop) and the Myc tag into loop 2. To our surprise, insertions of the HA tag into the amino terminus were only poorly tolerated, with yields of protein decreased more than 2.5-fold compared with the empty scaffold [SQM-HA(N): 19 mg of purified protein/l of culture, compared with SQM: 49 ± 11 mg/l of culture). Similarly, Myc insertion into loop 2 resulted in >5-fold decrease in protein yield in *E.coli*. In contrast, insertion of the AU1 tag into loop 1 did not destabilize SQM and may in fact increase yields (Table I).

These adverse effects of peptide insertion could be due to the insertion of 'any' peptide at each site, or to the insertion of a peptide with a given sequence at that site. Accordingly, we asked whether insertion of the Myc epitope peptide in the amino terminal site had similar effects to the insertion of the HA peptide. Myc insertion into the amino terminus also led to a decrease in yield, especially when comparing SQM with an AU1 tag in loop 1 to SQM with the Myc tag in the N-terminal site combined with the AU1 tag in loop 1 (Table I). In addition, the combination of the N-terminal HA insertion with any other in loop 1 or 2 resulted in lower yields compared with similar peptide aptamers lacking the
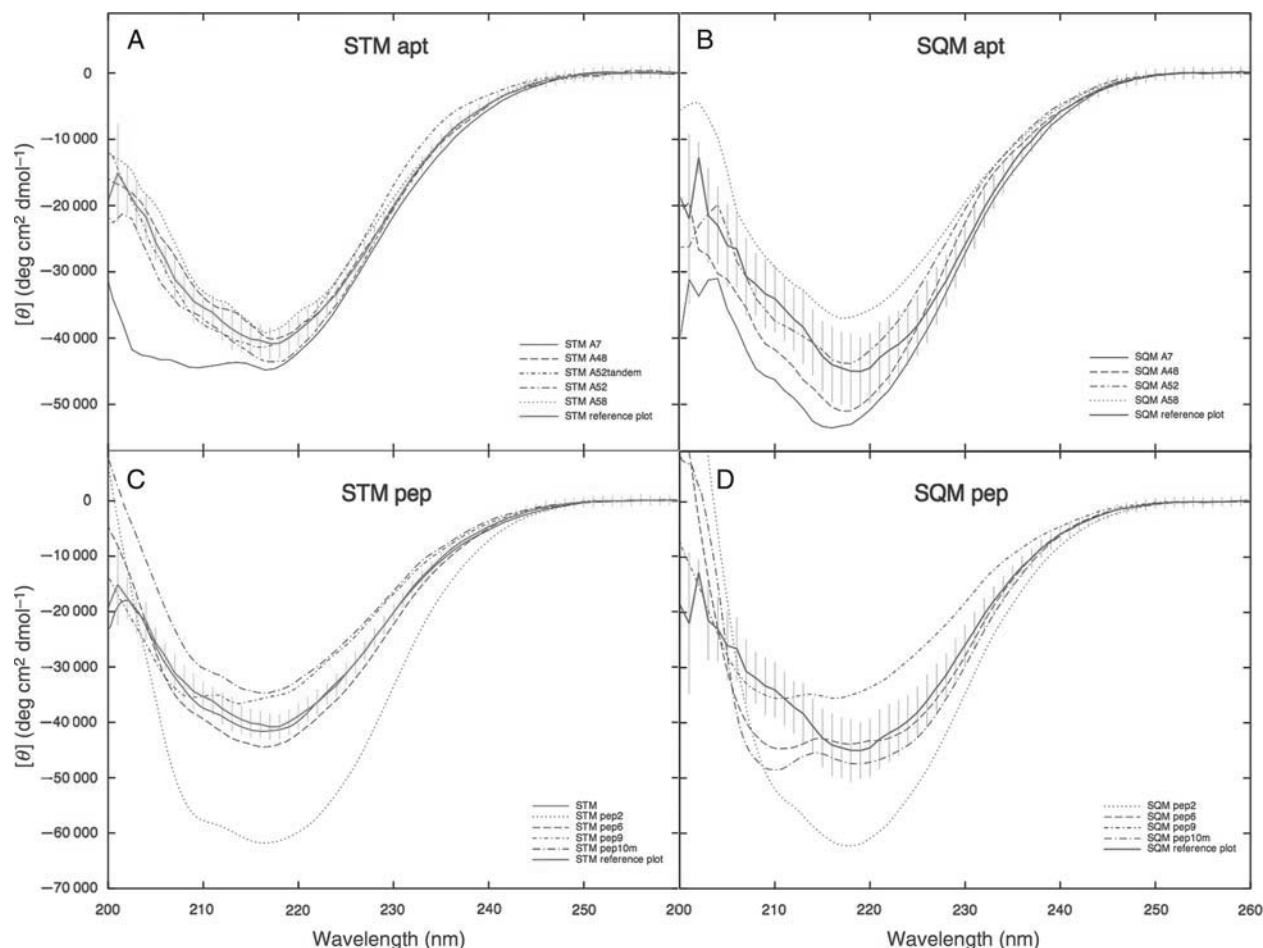
Fig. 3. Far-UV CD spectra of the STM scaffold protein compared with SQM containing short peptide insertions in loop 2. Shown are the far-UV CD spectra of the STM (**A** and **C**) and SQM (**B** and **D**) scaffold proteins containing short peptide inserts [A and B: A7, A48, A52, A52tandem (tandem insertion of A52) and A58; C and D: pep2, pep6, pep9 and pep10m) in loop 2 compared with a reference spectrum of the STM (red) or SQM (blue) scaffold. The data are expressed as mean residual ellipticity (MRE, [θ]). Vertical error bars indicate the standard deviation of five STM (red) or seven SQM (blue) CD spectra obtained from independent sample preparations shown in Fig. 2. The colour version of this figure is available at PEDS online.

HA inserts. Similarly, the Myc insertion at loop 2 was uniformly destabilizing of all other peptide aptamer variants. The highest yield we obtained, over 200 mg of peptide aptamer per litre of culture representing a 4-fold increase in yield over the empty scaffold, involved SQM with insertions of the AU1 tag at loop 1 and the HA tag at loop 2, with no insertion into the amino terminus. We conclude that although some peptides may be tolerated by the amino-terminus of SQM, this will not be a generally useful site.

Analysis of peptide aptamer secondary structure by CD spectroscopy

We next asked whether the insertion of the epitope tags into the three sites as described above (Table I) affected the proportion of secondary structure in the resulting peptide aptamers. To our surprise, even those insertions that decreased protein yields did not appreciably disrupt the secondary structure of the resulting peptide aptamers (Fig. 4A). We did note that the presence of a peptide at the N-terminal site changes the shape of the curve, pushing the inflexion point from 218 towards 209 nm (Fig. 4B SQM-HA). In order to ask whether this may reflect a general effect of insertions at this site on the structure of the scaffold, we analysed the spectra for a range of SQM-derived peptide aptamers with

inserts in loop 1 and/or loop 2, all with an insert at the amino terminus. We consistently found the secondary structure of these proteins to be more disrupted, with inflexion points at 209 nm, than the corresponding proteins lacking an insert at the N-terminal site (Fig. 4B).

Specific recognition of peptide aptamers by antibodies

Having determined that it is possible, within limits, to insert model peptides into each of the three positions, we wished to ask whether the inserted peptides were presented to solvent or were being incorporated into the secondary structure of an aberrantly folded scaffold. In order to do this, we wanted to know whether antibodies specific for each inserted epitope could recognize their target in the scaffold. First, we asked whether the AU1 peptide inserted into 'loop 1' could be recognized by an anti-AU1 antibody. To this end, we asked whether an anti-AU1 antibody could immuno-precipitate model peptide aptamers presenting the AU1 epitope tag in each site in turn. Indeed, the interaction was sufficiently tight that each peptide aptamer could be immuno-precipitated (data not shown). Next, we asked whether other epitope tags would be equally well recognized. To this end, we used scaffold proteins in which the Myc epitope peptide had been inserted in either the amino terminus, loop 1 or loop 2. We

408

Table I. Presentation of a range of peptides by SQM in the N-terminus, loop 1 or loop 2: effects on protein yields

Insert	Single			Double			Triple
	N	L1	L2	L1/L2	N/L1	N/L2	N/L1/L2
HA	19						
AU1		103	1				
AU1AU1		5					
Myc		7	11				
R24		3					
AU1 HA				206			
AU1 Myc				14			
HA AU1					2		
Myc AU1					11		
HA - Myc						4	
Myc AU1 HA							7
HA AU1 Myc							2
HAHA AU1 Myc							1

The numbers given correspond to the normalized amounts, in mg, of purified peptide aptamer recovered from a notional standard litre of bacterial culture.

performed immuno-precipitation experiments with an anti-Myc antibody, to ask whether the peptide was displayed on the surface of the folded protein, and used western blotting to detect the scaffold protein. As before, the epitope tag was equally well recognized regardless of the insertion site and appeared not to be affected by insertions at other sites in the same scaffold (Fig. 5).

Functionality of the SQM derivatives

Because of the combinatorial nature of the SQM scaffold with three possible insertion sites, we have devised a microarray assay that allows us to look at many peptide aptamers simultaneously using cognate antibodies of epitope tags incorporated at several sites independently. Peptide aptamers with various combinations of epitope inserts were spotted and allowed to chelate onto glass slides coated with nickel-NTA in order to induce oriented immobilization by the His tag. Subsequently, surfaces were blocked using 1% BSA PBST buffer and then incubated with the fluorescently labelled antibody at a concentration of 7 nM. After thorough washing, the slides were scanned using a standard DNA microarray scanner (Fig. 6). Tag-less aptamers and print buffer control spots all gave the expected lack of signal. Peptide aptamers containing the Myc and HA tags gave strong signals. Peptide aptamer samples containing the AU1 tag consistently gave much weaker signals compared with the Myc and HA counterparts, and on occasions could not be seen. In order to increase AU1 signal strength, the antibody incubation concentration was increased to 28 nM, which resulted in identifiable spots. Since all our model peptide aptamers were spotted at the same concentration of 10 μM, the possibility of concentration variations correlated with signal variation can be ruled out. Furthermore, internal controls of samples containing all three tags, where by definition the equimolar amounts of all three tags are present, showed the same systematic signal variations. This issue was independent of the insert position and is thus likely due to the short size of the AU1 tag leading to a lower binding affinity by the anti-AU1 antibody (see Discussion).

We have also analysed the expression profiles of 384 random peptides inserted at each of the three sites (Fig. 7). For this experiment, we grew small scale (1 ml) cultures of *E.coli* expressing the random peptide aptamers in 96-well plates, purified the peptide aptamers in high throughput (i.e. without optimizing expression or purification protocols for each well) and spotted an equal volume of each peptide aptamer in duplicate onto a glass microscope slide, creating a small microarray. We then probed the microarray with either of two fluorescently labelled antibodies that recognize the scaffold protein. After washing, the slides were analysed using a standard DNA microarray scanner. The signal intensity at each spot obtained with the antibody is proportional to the amount of peptide aptamer at each feature of the array. Because we do not normalize the peptide aptamers to a standard concentration prior to spotting, the signal intensity should be proportional to the amount of protein expressed in each of the 384 bacterial cultures. We found that insertions at loops 1 and 2 were generally well expressed, whereas only 42% of the peptide aptamers with amino-terminal inserts gave a signal-to-background ratio greater than 3 (Fig. 7 and Table II).

Discussion

In this study, we aimed to improve the stefin A-based STM scaffold and broaden its versatility. Our analysis of six variants of a stefin A-based scaffold indicated structural similarity (Fig. 2). When considering the minimal changes in SUN, SUM, SUC and SDM, we find that the location of the major inflexion point at 219 nm is largely unaffected by each change, indicating that the proportion of secondary structure in stefin A derivatives is largely unchanged by the amino acid alterations. This is true even for SQM, the engineered variant that possesses a combination of all modifications. The slight shift in the CD spectrum of STM towards the far UV compared with all the other variants may be influenced by the STM-specific tryptophan at position 4. We conclude from these data that SQM is likely to be a suitable and versatile scaffold protein.

The insertion of 10 amino acid long peptides into loop 2 was equally well tolerated by SQM and STM. Only a 17 amino acid insert (A7) caused a deviation from the empty scaffold spectra in STM. This indicates that A7 may drive a structural change in the scaffold of STM and that the insertion of inserts longer than 10 amino acids into loop 2 may be better tolerated by SQM than STM. However, 20 amino acid long inserts affected the secondary structure of SQM more than that of STM and decreased the protein yield. This indicates that the two scaffolds are likely to be differentially affected by different peptide inserts, and thus will present the same peptides differently. It might also indicate a length limit of just below 20 amino acids for peptides that are tolerated in loop 2 of SQM. We conclude that SQM is able to present some but not all peptides in loop 2 without adverse effects on its secondary structure.

By asking whether SQM can present peptide inserts for interaction with target proteins, we observed that neither SQM nor STM is able to present thioredoxin-derived BCL-6 POZ domain binders (Chattopadhyay *et al.*, 2006) for positive interaction in yeast two hybrid assays (data not shown). Previous attempts to shuffle peptides between scaffold
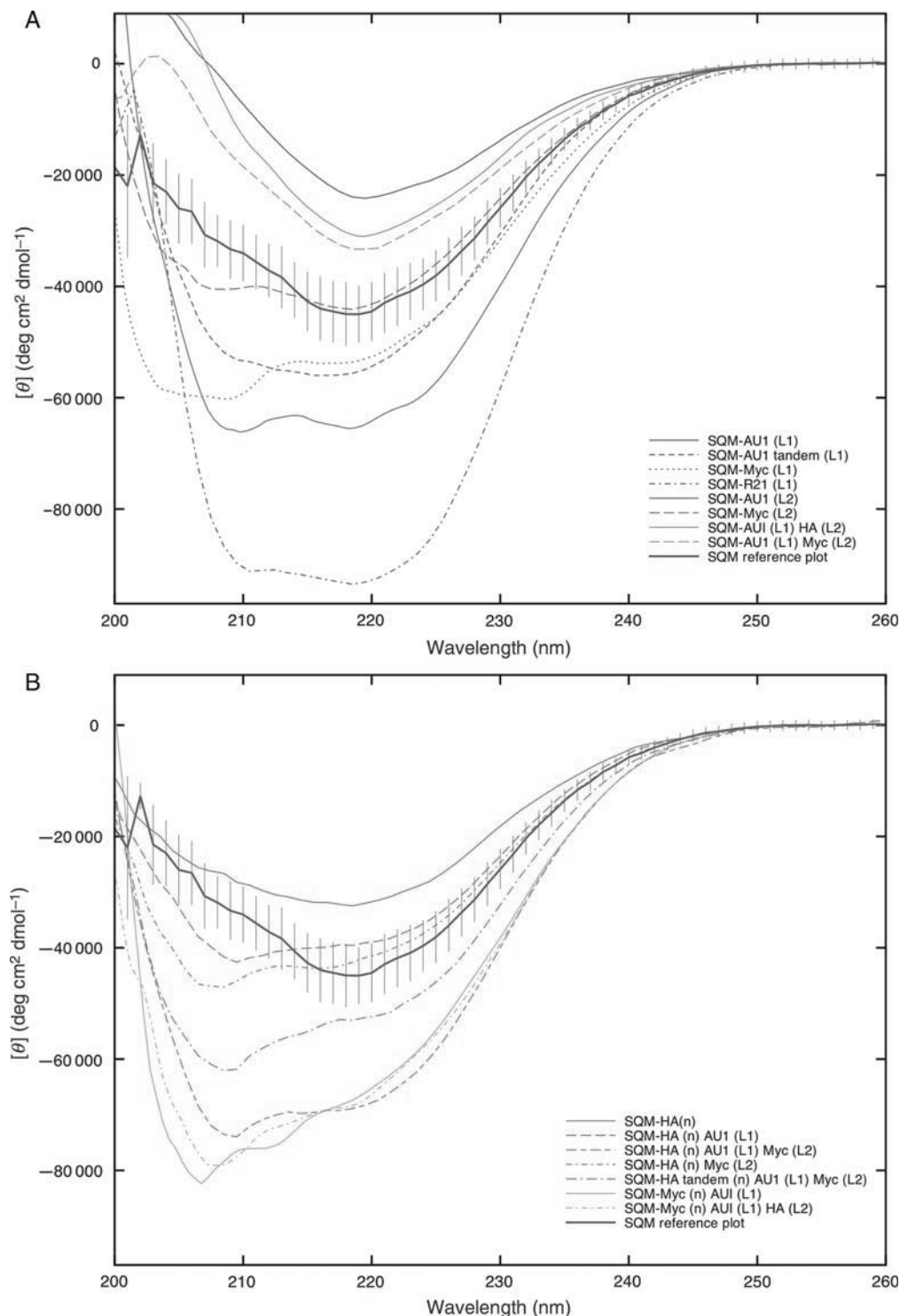
Fig. 4. Far-UV CD spectra of the SQM scaffold presenting epitope tag peptide inserts. Shown are the far-UV CD spectra of SQM scaffold proteins containing short epitope tag encoding inserts (AU1, HA, Myc1) in loop 1 (L1) or loop 2 (L2) compared with a reference spectrum of the SQM scaffold (blue) (**A**), alone or in combination with N-terminal insertions (**B**). The data are expressed as mean residual ellipticity (MRE, [θ]). Vertical error bars indicate the standard deviation of the five STM (red) or seven SQM (blue) CD spectra presented in Fig. 2. The colour version of this figure is available at PEDS online.

proteins as diverse as thioredoxin, staphylococcal nuclease and GFP have suggested that this can be achieved approximately one-third of the time (Klevenz *et al.*, 2002), and we interpret this result to indicate that STM and SQM present peptides differently than thioredoxin.

Insertion of epitope tags into the N-terminus, loop 1 and/or loop 2 had varying effects on protein yield and secondary structure. Most obviously, we observed that N-terminal insertions of short epitope tags were detrimental. Consistent with this, the analysis of a small-scale microarray of 384 spotted random peptide aptamers showed a detrimental effect of insertions at the N-terminus of SQM resulting in lower protein yields. Analysis of the CD spectra broadly confirmed these conclusions, showing that epitope peptide insertions at
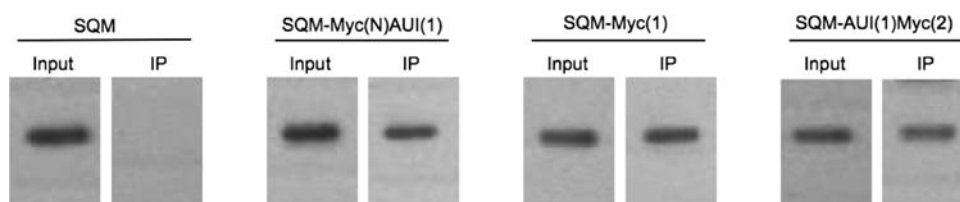
Fig. 5. Immunoprecipitation of SQM presenting the Myc epitope tag. The empty SQM scaffold protein (SQM) or SQM carrying a Myc peptide insert in each of the three available sites (N, loop 1 or loop 2) were expressed in *E.coli*. SQM with Myc inserted at the N-terminus also carried an AU1 tag in loop 1 and is designated SQM-Myc(N)AU1(1). SQM carrying Myc in loop 2 also carries the AU1 epitope insert in loop 1 [SQM-AU1(1)Myc(2)]. The variant carrying the Myc peptide in loop 1 did not have any other peptide inserts and is designated SQM-Myc(1). Each protein sample was incubated with immobilized anti-Myc antibody and proteins eluted from these columns were analysed with an antibody recognizing the S-tag present in the amino terminus of all our model peptide aptamers. Whereas SQM alone was not recruited to the anti-Myc antibody column, all of the other model peptide aptamers were, demonstrating that the Myc tag is available for interaction regardless of which insertion site is used.
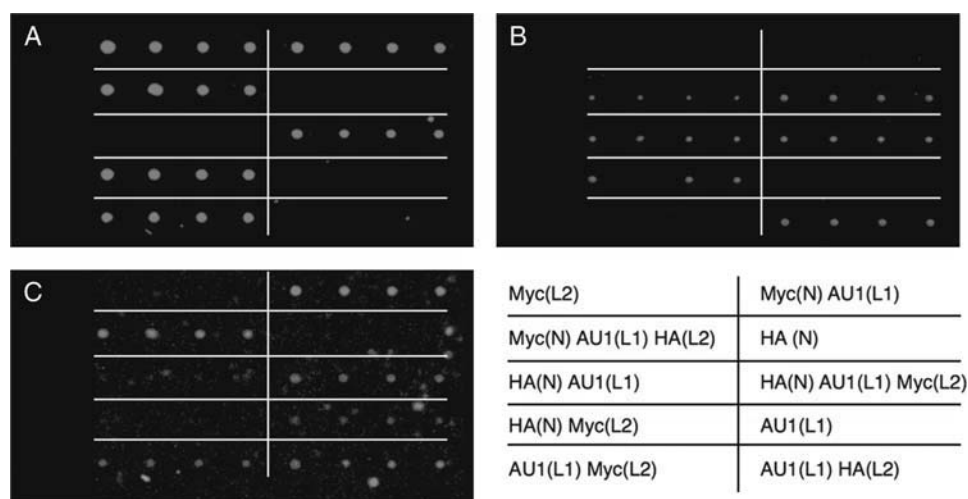

Fig. 6. Fluorescence microarrays of the spotted SQM library on Ni NTA surfaces. Model peptide aptamers with combinations of the three model inserts at each of the three positions were spotted in quadruplicate, with two peptide aptamers spotted per row. Shown are 10 distinct peptide aptamers whose identity is given in (**D**) using the following code: each of the three possible tags (Myc, HA and AU1) is followed by an abbreviation indicating the site at which it is inserted (N: the amino terminal site; L1: loop 1; L2: loop 2). Thus, Myc(L2) refers to a peptide aptamer with a single insertion of the Myc epitope in loop 2, whereas Myc(N)Au1(L1)HA(L2) refers to a peptide aptamer with three inserts: a Myc epitope in the amino-terminus, an AU1 epitope in loop 1 and an HA epitope in loop 2. After spotting, the slides were allowed to dry and were then hybridized with 7 nM anti-Myc-Atto 647N (red, **A**), 7 nM anti-HA-Atto 647N (red, **B**) or with 28 nM anti-AU1-Atto 550 (green, **C**). After washing the images were captured using a microarray scanner. The grid layout is summarized in (D). The colour version of this figure is available at PEDS online.

the N-terminal site caused drastic alterations of the secondary structure. This effect could be due to the insertion at this site, or the nature of the inserted peptide, which was the HA or Myc tag in all cases. We return to this question below. Decreases in protein stability following peptide insertion could be the result of sequence-specific effects of the peptide, or the result of generic effects of peptide insertion on protein stability, or a combination of the two. When considering the expression yields of the scaffold proteins presenting various combinations of epitope tag peptides, e.g. the effects of HA insertion at the amino-terminus, it appears to be a combination of these. Increased protein yields after insertion of the AU1 tag at loop 1 and the HA tag at loop 2 might indicate that double insertions may sometimes be beneficial. However, this effect cannot be generalized as replacement of the HA tag with a Myc tag decreased yield almost 15-fold (Table I).

Taken together, the protein expression data show that SQM is able to present peptides from three sites—the amino terminus, loop 1 and loop 2. Of these, the new loop 1 site appears to be most broadly useful, although loop 2 may be able to present peptides of 10–17 amino acids in length. We

speculate that many (but not all, e.g. A52tandem) longer inserts may affect the formation of the following β-strand and hence destabilize the scaffold. Interestingly, most insertions into the amino terminus greatly affect the secondary structure of the scaffold, rendering use of this site in first-round screens of randomized libraries impractical. We conclude that, although peptides can be presented by insertions at the amino-terminal site, these can also be frequently detrimental to the stability of the scaffold confirming our conclusions from CD spectra of model peptides. Accordingly, we do not propose to screen libraries of random peptides inserted in the amino terminus, but we do expect to be able to make use of this site to improve peptide aptamers that interact with the target protein using loop 1 and/or loop 2 by using a greater surface area and thus create binders with greater affinity and specificity.

Specific recognition of model peptide aptamers by antibodies

Epitope tags that were presented by SQM on a microarray were specifically recognized by antibodies. However, only weak signals were achieved with anti-AU1 antibody. The
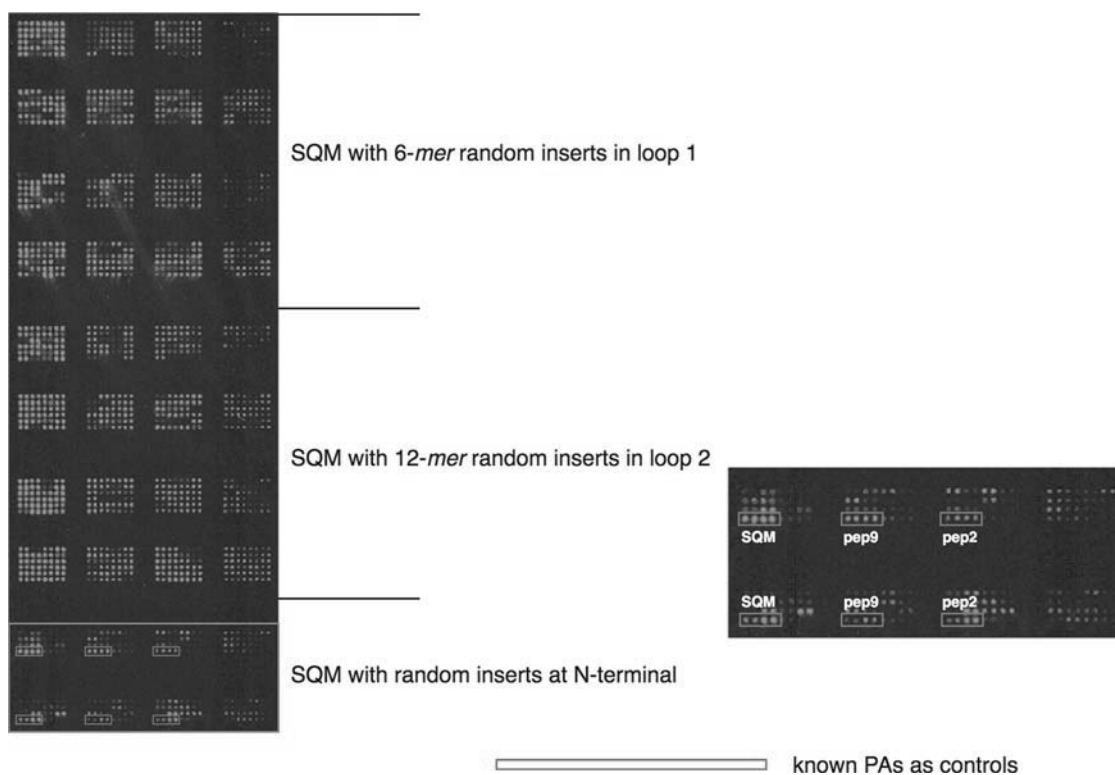
Fig. 7. Microarray comprising 384 random peptide aptamers inserted into loop 1 (top) or loop 2 (centre) or the amino terminus (bottom). Equal volumes of peptide aptamers purified from 1.5 ml bacterial cultures and eluted from equal volumes of Ni-agarose beads were spotted in duplicate onto glass slides. The presence of each peptide aptamer was detected by incubating the whole slide with an antibody that recognizes the scaffold protein (left-hand image). Signal intensity at each spot (feature) of the array correlates with the amount of protein present and was quantified using the scanner software (Table II). Significantly, insertions in the amino terminus and to some extent loop 2 appear less well tolerated than insertions into loop 1. Previously characterized peptide aptamers pep2, pep6, pep9 and pep10m as well as the empty SQM scaffold are given as controls. The colour version of this figure is available at PEDS online.

Table II. Comparison of signal-to-background ratios as an indicator for expression levels of 384 peptide aptamers expressed in duplicates and purified in high throughput format

Signal-to-background ratio	Number of spots	% total spots
SQM with 6-mer random inserts in loop 1		
>3	268	70
2–3	64	17
<2	52	14
SQM with 12-mer random inserts in loop 2		
>3	317	83
2–3	53	14
<2	14	4
SQM with 12mer random inserts in N-terminus		
>3	40	42
2–3	24	25
<2	32	33

lower AU1 signal may be due either to the short length of the AU1 peptide sequence in the epitope loop, with steric hindrance preventing strong association with the antibody, or to the inherently low affinity of the anti-AU1 antibody compared with the anti-Myc and the anti-HA antibodies. Immunoblotting experiments where equal amounts of protein were analysed and the structure of the scaffold is destroyed during SDS–PAGE always gave a stronger signal for the anti-Myc antibody compared with the anti-AU1 antibody (data not shown). This is clear evidence that the affinity of the anti-AU1 antibody is lower than the affinity of the anti-Myc antibody, and this is the most likely explanation for the lower signal for the AU1 tag in the microarray experiments (Fig. 6). On the basis of these experiments, it can be concluded that all epitope tags show selective binding to their corresponding conjugate antibody, though variations in binding affinity exist. Future experiments will determine whether the three insertion surfaces can be simultaneously recognized by three proteins.

In summary, the results described above demonstrate (i) that the stefin A-derived scaffold is amenable to engineering in multiple locations, with each change alone or in combination being well tolerated, (ii) that loop 1 of SQM, utilized here for the first time, appears to be a valid site for peptide insertion, and (iii) that any detrimental effects of the mutations or of insertions are apparently magnified by insertions in the amino terminus. Therefore, the N-terminal insertion site cannot be used routinely. However, some N-terminal inserts are tolerated, which will allow us to use this site to improve the binding affinity and specificity of peptide aptamers in SQM-loop 1 and/or loop 2. SQM appears to be more tolerant to short-sequence insertions into loop 2 (10 amino acids) than STM but longer inserts (>17–42 amino acids) often affect the secondary structure. We conclude that the SQM scaffold is a valid new SteA-based scaffold protein that should allow the selection of peptide aptamers with an increased binding affinity and higher target selectivity than our previous scaffold, STM.

Supplementary data

Supplementary data are available at *PEDS* online.

Acknowledgements

Funding

References

Abedi,M.R., Caponigro,G. and Kamb,A. (1998) *Nucleic Acids Res.*, **26**, 623–630.

Abrahamson,M., Barrett,A.J., Salvesen,G. and Grubb,A. (1986) *J. Biol. Chem.*, **261**, 11282–11289.

Arnold,K., Bordoli,L., Kopp,J. and Schwede,T. (2006) *Bioinformatics*, **22**, 195–201.

Beitz,E. (2000) *Bioinformatics*, **16**, 135–139.

Binz,H.K., Stumpp,M.T., Forrer,P., Amstutz,P. and Pluckthun,A. (2003) *J. Mol. Biol.*, **332**, 489–503.

Binz,H.K., Amstutz,P. and Pluckthun,A. (2005) *Nat. Biotechnol.*, **23**, 1257–1268.

Bode,W., Engh,R., Musil,D., Thiele,U., Huber,R., Karshikov,A., Brzin,J., Kos,J. and Turk,V. (1988) *EMBO J.*, **7**, 2593–2599.

Brzin,J., Popovic,T., Turk,V., Borchart,U. and Machleidt,W. (1984) *Biochem. Biophys. Res. Commun.*, **118**, 103–109.

Chattopadhyay,A., Tate,S.A., Beswick,R.W., Wagner,S.D. and Ko Ferrigno,P. (2006) *Oncogene*, **25**, 2223–2233.

Clark,R.J., Daly,N.L. and Craik,D.J. (2006) *Biochem. J.*, **394**, 85–93.

Cohen,B.A., Colas,P. and Brent,R. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 14272–14277.

Colas,P. (2008) *J. Biol.*, **7**, 2; doi:10.1186/jbiol64.

Colas,P., Cohen,B., Jessen,T., Grishina,I., McCoy,J. and Brent,R. (1996) *Nature*, **380**, 548–550.

Davis,J.J., Tkac,J., Laurenson,S. and Ko Ferrigno,P. (2007) *Anal. Chem.*, **79**, 1089–1096.

Davis,J., Tkac,J., Humphreys,R., Buxton,A., Lee,T. and Ko Ferrigno,P. (2009) *Anal. Chem.*, **81**, 3314–3320.

Estrada,S., Nycander,M., Hill,N.J., Craven,C.J., Waltho,J.P. and Björk,I. (1998) *Biochemistry*, **37**, 7551–7560.

Estrada,S., Pavlova,A. and Björk,I. (1999) *Biochemistry*, **38**, 7339–7345.

Estrela,P., Paul,D., Li,P., Keighley,S.D., Migliorato,P., Laurenson,S. and Ko Ferrigno,P. (2008) *Electrochim. Acta*, **53**, 6489–6496.

Evans,D., Johnson,S., Laurenson,S., Davies,A.G., Ko Ferrigno,P. and Wälti,C. (2008) *J. Biol.*, **7**, 3.

Fisher,C.L. and Pei,G.K. (1997) *Biotechniques*, **23**, 570–571.

Green,G.D., Kembhavi,A.A., Davies,M.E. and Barrett,A.J. (1984) *Biochem. J.*, **218**, 939–946.

Guex,N. and Peitsch,M.C. (1997) *Electrophoresis*, **18**, 2714–2723.

Japelj,B., Waltho,J.P. and Jerala,R. (2004) *Proteins*, **54**, 500–512.

Jenko,S., Dolenc,I., Guncar,G., Dobersek,A., Podobnik,M. and Turk,D. (2003) *J. Mol. Biol.*, **326**, 875–885.

Jerala,R. and Zerovnik,E. (1999) *J. Mol. Biol.*, **291**, 1079–1089.

Johnson,S., Evans,D., Laurenson,S., Paul,D., Davies,A.G., Ko Ferrigno,P. and Wälti,C. (2008) *Anal. Chem.*, **80**, 978–983.

Juraja,S.M., Mulhern,T.D., Hudson,P.J., Hattarki,M.K., Carmichael,J.A. and Nuttall,S.D. (2006) *Protein Eng. Des. Sel.*, **19**, 231–244.

Kenig,M., Jerala,R., Kroon-Zitko,L., Turk,V. and Zerovnik,E. (2001) *Proteins*, **42**, 512–522.

Klevenz,B., Butz,K. and Hoppe-Seyler,F. (2002) *Cell. Mol. Life Sci.*, **59**, 1993–1998.

Kohl,A., Binz,H.K., Forrer,P., Stumpp,M.T., Pluckthun,A. and Grutter,M.G. (2003) *Proc. Natl Acad. Sci. USA*, **100**, 1700–1705.

Koide,A., Bailey,C.W., Huang,X. and Koide,S.J. (1998) *Mol. Biol.*, **284**, 1141–1151.

Kolmar,H. (2008) *FEBS J.*, **275**, 2684–2690.

Kolmar,H. and Skerra,A. (2008) *FEBS J.*, **275**, 2667.

Ladner,R.C. (1995) *Trends Biotechnol.*, **13**, 426–430.

Laemmli,U.K. (1970) *Nature*, **227**, 680–685.

LaVallie,E.R., DiBlasio,E.A., Kovacic,S., Grant,K.L., Schendel,P.F. and McCoy,J.M. (1993) *Biotechnology*, **11**, 187–193.

Martin,J.R., Jerala,R., Kroon-Zitko,L., Zerovnik,E., Turk,V. and Waltho,J.P. (1994) *Eur. J. Biochem.*, **225**, 1181–1194.

Martin,J.R., Craven,C.J., Jerala,R., Kroon-Zitko,L., Zerovnik,E., Turk,V. and Waltho,J.P. (1995) *J. Mol. Biol.*, **246**, 331–343.

Nord,K., Nilsson,J., Nilsson,B., Uhlén,M. and Nygren,P.A. (1995) *Protein Eng.*, **8**, 601–608.

Parmeggiani,F., Pellarin,R., Larsen,A.P., Varadamsetty,G., Stumpp,M.T., Zerbe,O., Caflisch,A. and Pluckthun,A. (2008) *J. Mol. Biol.*, **376**, 1282–1304.

Pavlova,A. and Björk,I. (2002) *Eur. J. Biochem.*, **269**, 5649–5658.

Pol,E., Olsson,S.L., Estrada,S., Prasthofer,T.W. and Björk,I. (1995) *Biochem. J.*, **311**, 275–282.

Reits,E., Griekspoor,A., Neijssen,J., Groothuis,T., Jalink,K., van Veelen,P., Janssen,H., Calafat,J., Drijfhout,J.W. and Neefjes,J. (2003) *Immunity*, **18**, 97–108.

Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning; A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Saveanu,L., Fruci,D. and van Endert,P. (2002) *Mol. Immunol.*, **39**, 203–215.

Schlehuber,S. and Skerra,A. (2005) *Expert. Opin. Biol. Ther.*, **5**, 1453–1462.

Schwede,T., Kopp,J., Guex,N. and Peitsch,M.C. (2003) *Nucleic Acids Res.*, **31**, 3381–3385.

Shaw,M.K. and Ingraham,J.L. (1967) *J. Bacteriol.*, **94**, 157–164.

Shu,W., Laurenson,S., Knowles,T.P.J., Ko Ferrigno,P. and Seshia,A.A. (2008) *Biosens. Bioelectron.*, **24**, 233–237.

Skerra,A. (2007) *Curr. Opin. Biotechnol.*, **18**, 295–304.

Stubbs,M.T., Laber,B., Bode,W., Huber,R., Jerala,R., Lenarcic,B. and Turk,V. (1990) *EMBO J.*, **9**, 1939–1947.

Tao,L., *et al.* (2004) *Proc. Natl Acad. Sci. USA*, **101**, 11471–11476.

Tate,S., Ushioda,T., Utsunomiya-Tate,N., Shibuya,K., Ohyama,Y., Nakano,Y., Kaji,H., Inagaki,F., Samejima,T. and Kainosho,M. (1995) *Biochemistry*, **34**, 14637–14648.

Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) *Nucleic Acids Res.*, **25**, 4876–4882.

Turk,V., Brzin,J., Longer,M., Ritonja,A., Eropkin,M., Borchart,U. and Machleidt,W. (1983) *Hoppe-Seyler's Z. Physiol. Chem.*, **364**, 1487–1496.

Turk,V., Brzin,J., Kotnik,M., Lenarcic,B., Popovic,T., Ritonja,A., Trstenjak,M., Begic-Odobasic,L. and Machleidt,W. (1986) *Biomed. Biochim. Acta*, **45**, 1375–1384.

Woodman,R., Yeh,J.T., Laurenson,S. and Ko Ferrigno,P. (2005) *J. Mol. Biol.*, **352**, 1118–1133.

Zerovnik,E., Lenarcic,B., Jerala,R. and Turk,V. (1991) *Biochim. Biophys. Acta*, **1078**, 313–320.