Pergamon

Bioorganic & Medicinal Chemistry 7 (1999) 23–27

BIOORGANIC & MEDICINAL CHEMISTRY

Design of Proteins Using Rigid Organic Macrocycles as Scaffolds

Ashley S. Causton and John C. Sherman*

Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, BC, V6T 1Z1 Canada

Received 4 May 1998; accepted 4 September 1998

Abstract—We have designed and synthesized new three-helix template-assembled synthetic proteins (TASPs) **1a–c**. The template was the rigid cyclotribenzylene (CTB) macrocycle **2**, which has C_3 symmetry. Thiol moieties on the CTB template were used to link cysteine-containing peptide strands **3a–c** via disulfide bonds. With designed peptide strands of 15 and 18 residues in length, the structure of TASPs **1a–c** were determined to be helical in water according to circular dichroism (CD) spectroscopy. The helicities of TASPs **1a–c** were unchanged over large ranges of pH (2–12) and salt concentrations (0–2 M KCl). TASPs **1a–c** were also extremely resistant to chemical denaturants: it requires a guanidine hydrochloride (GnHCl) concentration of 7.4 M for TASPs **1a–c** to lose 50% of their helicity. The major force for stabilization of TASPs **1a–c** is the hydrophobic bundling of the helices.

Introduction

The discovery of the reversible folding/unfolding of proteins led to the realization that all the information necessary for a protein to fold into its native structure is contained in its primary amino acid sequence.[1] The specific three-dimensional structure of a protein is the result of a multitude of noncovalent interactions involving both the polypeptide backbone and its side chains (which may be polar, nonpolar, or charged).[2] Determining how a tertiary structure of a protein arises from its primary amino acid sequence has become known as the protein-folding problem. The ultimate goal for researchers in this area is the prediction of protein structure from an amino acid sequence alone. De novo design of proteins that manifest unique properties is another challenging goal in this area. Investigations into simple de novo proteins provide an iterative process where the protein design increases in sophistication while clues to solving the folding problem are obtained.

Our research group is interested in studying the noncovalent interactions responsible for protein tertiary structure. To this end, we have created a family of water-soluble template-assembled synthetic proteins (TASPs). The concept of TASP was developed by Mutter, who put forward the idea of overcoming the potentially complex folding patterns available to a polypeptide by linking peptides with predetermined secondary structure to a template.[3] This concept has been used with organic macrocycles as the template to model four helix bundles using porphyrins,[4] and by our group using a cavitand macrocycle.[5] We chose cavitand macrocycles as our template because they have a rigid, well defined bowl-like structure with an enforced cavity that has the potential for recognition of apolar moieties.[6] The other advantage of using different cavitand macrocycles is that we can model three or four helical bundles in a parallel or antiparallel fashion, depending on the design. Here we report the design, synthesis, and preliminary characterization of a new type of water-soluble three-helix bundle, TASPs **1a–c** (Fig. 1).

Design

Careful design of the TASPs is crucial to their success. Our strategy for the synthesis of TASPs **1a–c** was to covalently link unprotected peptide strands to cyclotribenzylene (CTB) template **2** via disulfide bonds. We chose CTB **2** as the template for manifold reasons: CTB **2** is a known compound,[7] and it is a rigid macrocycle, which is an important characteristic for preorganization of the peptide strands. The three-thiol groups on the rim of the macrocycle can be used as a synthetic handle for the attachment of peptides. In addition, the three-thiol moieties are 8–10 Å apart, which will result in the appropriate inter-helical distance found for three-helix bundles in nature.[8] CTB **2** has an enforced cavity with potential for binding apolar substrates. This may be important if these TASPs are to be used as substrate binders, catalysts, or drug delivery devices. Finally, CTB **2** is a mixture of two enantiomers (designated as the M and P isomers using IUPAC nomenclature),[9] which upon addition of three chiral peptide strands would form two diastereomers, which are, in principle,

Key words: Helical bundles; TASP; de novo protein; cyclotribenzylene.

*Corresponding author. E-mail: sherman@chem.ubc.ca

0968-0896/99/$ - see front matter
PII: S0968-0896(98)00224-7

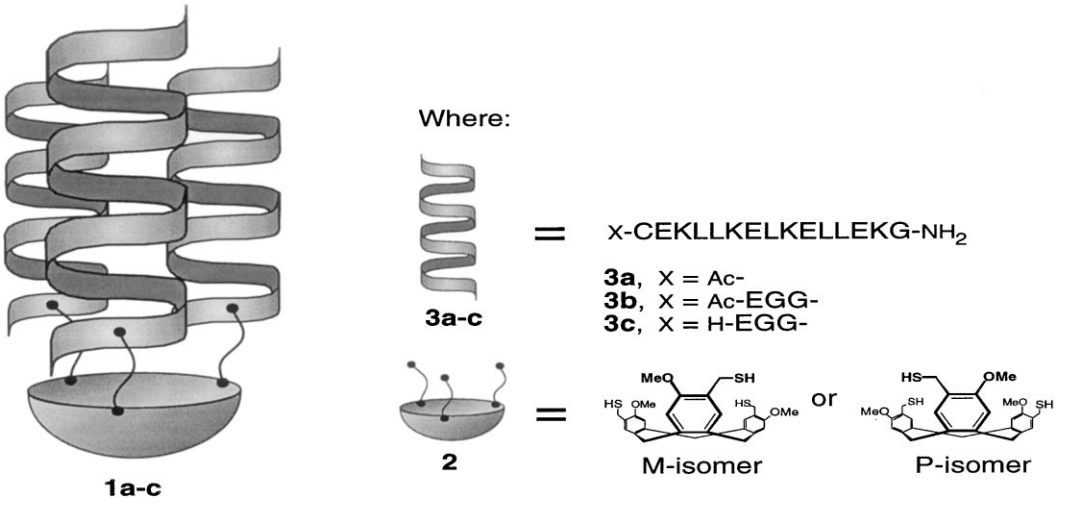


Figure 1. Schematic diagram of the TASPs **1a–c**. Template **2** is linked to peptide strands **3a–c** via disulfide bonds. C = cysteine, E = glutamic acid, K = lysine, L = leucine, G = glycine, -NH_2 = amidated C-terminus, Ac- = acylated N-terminus, and H- = free N-terminus.

separable and may manifest different physical properties. Indeed, subtle interactions between helices could potentially be probed via analysis of such diastereomers. Before we can obtain a crystal structure of our TASP, it is essential to first separate these diastereomers.

Peptides **3a–c** were designed to fold into an amphiphilic α-helix;[10] the template-enhanced bundling of three of these helices leads to the desired folding motif (see helical wheel diagram in Figure 2). The peptide contains nonpolar leucine residues occupying one face of the helices to drive inter-helical bundling via the hydrophobic effect in water.[2] Hydrophilic lysine and glutamic acid residues are incorporated to reside on the exterior of the three-helix bundle, thus imparting water solubility. The *i, i+4* spacing of these oppositely charged residues is optimal for intra-helical stabilization by salt bridges.[11] The charge distribution within the helix helps to stabilize the helix macrodipole,[2] e.g. the negatively charged glutamic acid residues (when compared to the positively charged lysine residues) tend to be towards the electropositive N-terminus of the peptide strand.

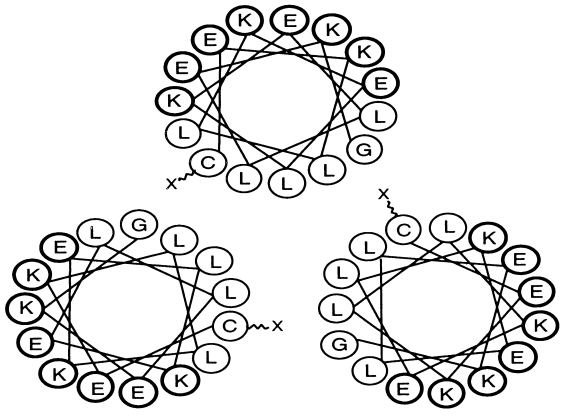


Figure 2. A helical wheel diagram of TASPs **1a–c** showing the proposed bundling. The helical portion of the peptide strands is viewed from the N-terminus, looking down the helical axis. For **3a** X = Ac-, **3b** X = Ac-EGG-, **3c** X = H-EGG-. Dark circles represent hydrophilic residues.

Favorable inter-helical salt bridges such as those designed into coiled coils[12] are yet to be included in our model. A glycine residue is included as a C-cap for unattached backbone hydrogen-bonding moieties to diminish fraying of the helices.[13] A cysteine residue (currently towards the N-terminal end of the peptide) is the synthetic handle for disulfide formation. The post-cysteine N-terminal three residue sequences of peptides **3b** and **3c** aid in separation of the diasteomeric TASPs. All peptides were amidated at their C-terminus and peptides **3a** and **3b** were acylated at their N-terminus to prevent the termini from being charged and therefore interact unfavorably with the helix macrodipole.[14] The N-terminus of peptide **3c** was left as the free amine.

The essential step in the synthesis of TASPs **1a–c** involves linking the unprotected peptide strands to the template. Previously this has been achieved in our group using peptides that have been bromo- or chloro-acetylated at their N-terminus and linked to phenolic or thiophenolic cavitand moieties.[5,15] Others have linked unprotected peptide strands to various templates using thioethers,[4b,16] thioesters,[17] oximes,[18] and maleimides.[19] A more versatile peptide-template linker would be to form disulfide bonds using peptides containing a cysteine residue, which could conceivably be placed anywhere in the sequence. Protein crosslinking reagents that rely on disulfide bonds, formed using a thiol-containing protein or crosslinker that has been activated with 2,2′-dipyridyl disulfide (DPDS) have been known for many years.[20] More recently, DPDS was employed for making disulfide linked peptide heterodimers[21] and in the assembly of multiple peptide units.[22] Thus, we explored the formation of disulfides, via DPDS, to covalently link unprotected peptide strands to our thiol-containing template.

Results

TASPs **1a–c** were prepared by first activating the cysteine residues of peptides **3a–c** using five equivalents

of 2,2′-dipyridyl disulfide (DPDS) in ethanol (Scheme 1). After HPLC purification, six equivalents of the relevant activated peptide **4a–c** were reacted with CTB **2** in dimethylformamide (DMF) with diisopropylethylamine (DIPEA) as base to produce TASPs **1a–c**. The reaction proceeded very cleanly in less than an hour, with the bonus that the starting activated peptides **4a–c** could be recovered during HPLC purification. Only TASP **1c** was separable into its constituent diastereomers (given the codes $\mathbf{1c_f}$ and $\mathbf{1c_s}$—depending whether it was the fast or slow eluting diastereomer by HPLC). The properties of TASPs **1a–c** were compared to a single-stranded control peptide **4c**.

TASPs **1a–c** and control peptide **4c** show CD spectra consistent with an α-helix (Figure 3 shows $\mathbf{1c_f}$, $\mathbf{1c_s}$ and **4c**). By monitoring the CD signal of the α-helix at the 222 nm minimum (i.e. $[\theta]_{222}$),[23] the helicity of TASPs **1a–c** were found to be concentration independent within the range studied, which suggests that they are likely to be monomeric in water (Figure 4 shows $\mathbf{1c_f}$ and $\mathbf{1c_s}$). In contrast, single-stranded control peptide **4c** shows concentration-dependent CD spectra with higher helicity at increasing concentrations (Fig. 4), which is consistent with aggregation of the amphiphilic peptide strands.[24]

Titration with trifluoroethanol (TFE), which is known to induce helicity in non-helical peptides,[25] shows a minor (>10%) increase in the helicities of **1a–c** with TFE solutions of over 20% (data not shown).

The GnHCl induced denaturation curves of $\mathbf{1c_f}$, $\mathbf{1c_s}$ and control peptide **4c** are shown in Figure 5. Peptide **4c** (180 μM) is 50% unfolded at 1.6 M GnHCl,[24] whereas for TASPs $\mathbf{1c_f}$ and $\mathbf{1c_s}$, 50% unfolding is estimated at 7.4 M for both diastereomers. TASPs **1a–b** exhibited almost identical denaturation curves as $\mathbf{1c_f}$ and $\mathbf{1c_s}$ (data not shown). The shapes of the denaturation curves are consistent with a cooperative unfolding of the TASPs.

The thermal stability of TASPs **1a–c** was investigated by heating them in 6 M GnHCl. Loss of helicity as linear with increasing temperature. At 70 °C TASPs **1a–c** had lost only about 20% of their helicity (data not shown).

High (2 M KCl) and no salt, and a pH range of 2–12 have no effect on the helicity of TASPs **1a–c** (data not shown).

Discussion and Conclusions

We have successfully designed and synthesized a new family of three-helix bundle TASPs. TASPs **1a–c** are remarkably stable to the chemical denaturant GnHCl, and insensitive to changes in salt concentration and extremes of pH. Their stability appears to come from inter-strand hydrophobic effects that are enhanced due to the peptide preorganization by the template. TFE titration resulted in a modest increase in the helicity of TASPs **1a–c**: this implies the TASPs are close to their maximum helicity in water. Intra- and inter-helix electrostatic interactions do not appear to be crucial in TASPs **1a–c**, as shown by the negligible effect of high salt and changes in pH.

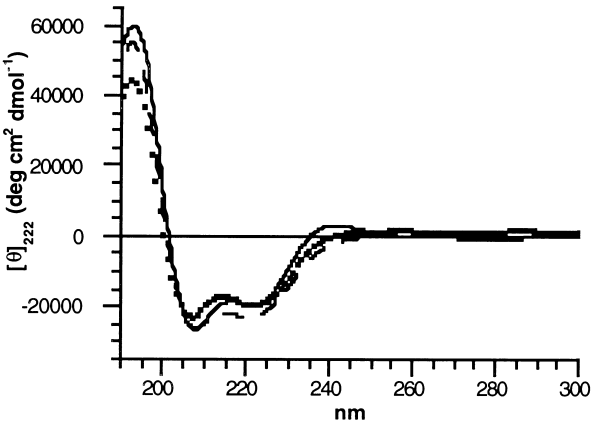
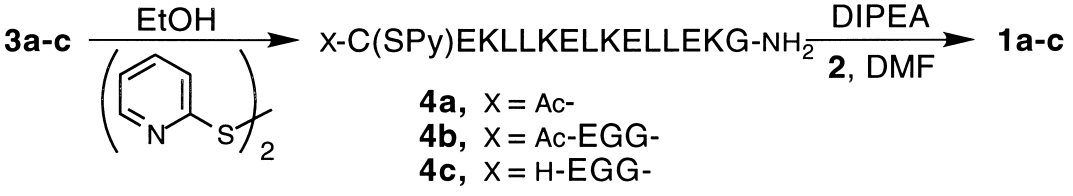


Figure 3. CD spectra of TASPs $\mathbf{1c_f}$ (solid line) (62 μM), $\mathbf{1c_s}$ (dotted line) (64 μM), and control peptide **4c** (dashed line) (180 μM) at pH 7.5 (10 mM phosphate buffer) and 25 °C. Concentrations of stock solutions were determined by amino acid analysis.

The absolute configurations of the CTB enantiomers associated with diastereomers $\mathbf{1c_f}$ and $\mathbf{1c_s}$ were not assigned. We plan to crystallize the individual diastereomers and then identify them.

TASPs **1a–c** did not show any detectable difference in their stability to GnHCl denaturation; therefore, in order to better study the noncovalent forces involved in our TASPs, we plan to make less stable models. Initially we plan to reduce the number of amino acids in the peptide sequence; a shorter peptide would possess less stabilizing interactions, thus reducing the overall stability of the TASP and make it more accessible for comparative studies. For example, we could then better investigate subtle differences between diastereomeric TASPs since their melt temperature and GnHCl would not be so nearly off scale.

Other plans for CTB **2** are to link peptides by their C-terminus, and compare them to the N-terminally linked peptides. We also plan to extend the use of CTB **2** to



Scheme 1. Synthesis of S-pyridyl activated peptides **4a–c**, and their subsequent reaction with CTB **2** to form TASPs **1a–c**. EtOH = ethanol, DIPEA = diisopropylethylamine, DMF = dimethylformamide, Py = pyridyl.

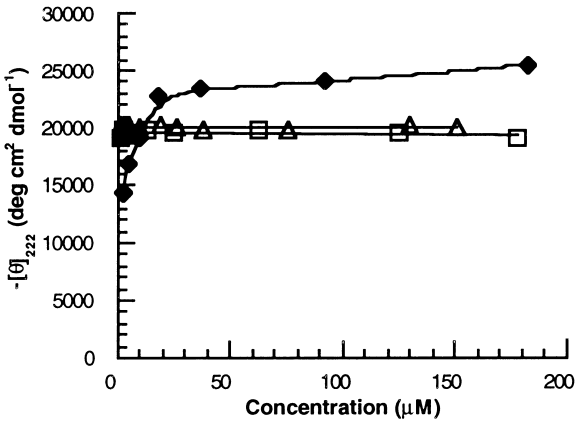


Figure 4. $[\theta]_{222}$ versus concentration for TASPs $\mathbf{1c_f}$, (□), $\mathbf{1c_s}$ (Δ) and control peptide **4c** (◆) at pH 7.5 (10 mM phosphate buffer) and 25 °C.

model coiled-coils, anti-parallel three-helix bundles, β-α-β motifs and β-sheets. Thus, coupled with the other macrocycle containing TASPs in our research group, we hope to first gain an insight into the features that are important in our family of TASPs, and then to use them to help solve the protein folding problem.

Experimental

General. Chemicals (Aldrich or BDH) used for the synthesis were reagent grade except dimethyl formamide that was dried over 4 Å molecular sieves and degassed by bubbling dry N_2 through it for 5 min before use. CTB **2** was prepared as reported.[26] Peptides were prepared on a 0.25 mmol scale using an Applied Biosystems 431A peptide synthesizer using FastMoc™ chemistry with peptide synthesis grade (Advanced Chemtech, Aldrich, or Richelieu Biotechnologies) reagents. The peptides and TASPs were purified using preparative scale C18 reverse phase HPLC (Perkin–Elmer), compound separation was effected by running gradients using water (containing 0.1% TFE) and HPLC grade acetonitrile (containing 0.05% TFA); 229 nm was the wavelength monitored for product elution. Peptides were characterized by their correct mass using LSIMS performed on a Kratos Concept IIH32. TASPs were characterized using electrospray MS run on a Perkin–Elmer SCIEX API 300 LC/MS/MS system. CD's were run on a Jasco J-710 spectropolarimeter using quartz cuvettes of 1 mm and 1 cm path length. The temperature was kept at 25 °C using a Haake FX 10 circulating bath. Samples were allowed 5 min to equilibrate to the desired temperature. Blanks were subtracted from the CD spectra, which were the average of three scans. Test samples for CD were prepared in duplicate or triplicate from stock solutions of known concentrations from amino acid analysis. The pH was measured using a Fisher Scientific Accumet® pH meter model 915.

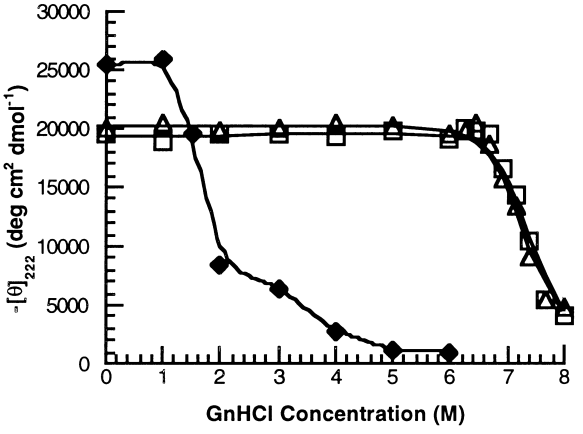


Figure 5. Guanidine hydrochloride denaturation curves for TASPs $\mathbf{1c_f}$ (□) (62 μM), $\mathbf{1c_s}$ (Δ) (64 μM), and control peptide **4c** (◆) (180 μM) at 25 °C.

Peptides 3a–c were synthesized on Rink resin (which results in a C-terminal amide upon cleavage) using standard FastMoc® procedures.[27] The N-termini of resin bound peptides **1a** and **1b** were acylated (after first deprotecting the N-terminus) using a standard DMF/acetic anhydride procedure.[27] The peptides were deprotected/cleaved from the resin using a trifluoroacetic acid/water/ethane dithiol mixture and then isolated by filtration.[27] The crude peptides **3a–c** were then purified by HPLC and lyophilized. LSIMS; **3a**; *m/z* 1814 $(M+H)^+$, **3b**; *m/z* 2057 $(M+H)^+$, **3c**, *m/z* 2015 $(M+H)^+$.

Activated peptide 4a was prepared by the addition of peptide **3a** (20 mg, 11 μmol) in 3 mL ethanol to a stirred solution of dipyridyl disulfide (12 mg, 55 μmol) in 2 mL ethanol. The reaction was stirred at room temperature for 1 h. The ethanol was removed in vacuo and the residue redissolved in water (1.5 mL) containing 0.1% TFA, filtered (0.45 μm nylon filter), and purified by HPLC. Lyophilization resulted in 18 mg (85%) of **4a**. LSIMS; *m/z* 1923 $(M+H)^+$.

Activated peptides 4b and **4c** were prepared from **3b** and **3c** in a similar manner to **4a**, resulting in **4b** (65%) (LSIMS; *m/z* 2166 $(M+H)^+$), and **4c** (58%) (LSIMS; *m/z* 2124 $(M+H)^+$).

TASP 1a was prepared by adding DIPEA (24 mg, 190 μmol) to a mixture of activated peptide **4a** (36 mg, 19 μmol) and CTB **2** (1.6 mg, 3.2 μmol) in DMF under N_2. The reaction was stirred at room temperature for 1 h. The solvent was removed in vacuo, keeping the temperature of the water bath below 40 °C. The residue was re-dissolved in water (1.5 mL), filtered (0.45 μm nylon filter), and injected directly onto a preparative HPLC column. Lyophilization resulted in 9 mg (45%) of **1a** (electrospray MS; 5934 $(M+H)$). In addition, 5 mg of **4a** was recovered.

TASPs 1b and **1c** were prepared as described for **1a**, but using activated peptide **4b** or **4c** instead of **4a**. **1b** was prepared in 36% yield (electrospray MS; 6663 $(M+H)$). For compound **1c**, it was possible to separate the two different diastereomers (see text) by HPLC. The faster eluting diastereomer $\mathbf{1c_f}$, (electrospray MS; 6537

(M + H)) was recovered in 31% yield, and the slower moving diastereomer **1c$_s$**, (electrospray MS; 6537 (M + H)) was recovered in 23% yield (therefore a 54% overall yield).

Acknowledgements

We thank the Natural Sciences and Engineering Research Council of Canada and the donors of the Petroleum Research Fund, administered by the ACS, for financial support. We also thank S. He for running the electrospray mass spectrometer and M. Fryzuk for the use of his HPLC.

References and Notes

1. Anfinsen, C. B. *Science* **1973**, *181*, 223.
2. Creighton, T. E. In *Proteins: Structures and Molecular Properties*, 2nd ed.; W. H. Freeman and Co.: New York, 1993.
3. (a) Mutter, M.; Tuchscherer, G. G.; Miller, C.; Altmann, K.-H.; Carey, R. L.; Wyss, D. F.; Labhardt, A. M.; Rivier, J. E. *J. Am. Chem. Soc.* **1992**, *114*, 1463. (b) Mutter, M.; Tuchscherer, G. *Makromol. Chem., Rapid Commun.* **1988**, *9*, 437. (c) Mutter, M.; Altmann, E.; Altmann, K.-H.; Hersperger, R.; Koziej, P.; Nebel, K.; Tuchscherer, G.; Vuilleumer, S.; Gremlich, H.-U.; Müller, K. *Helv. Chim. Acta* **1988**, *71*, 835.
4. (a) Arai, T.; Kobata, K.; Mihara, H.; Fujimoto, T.; Nishino, N. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 1989. (b) Choma, C. T.; Kaestle, K.; Åkerfeldt, K. S.; Kim, R. M.; Groves, J. T.; DeGrado, W. F. *Tetrahedron Lett.* **1994**, *35*, 6191. (c) Åkerfeldt, K. S.; Kim, R. M.; Camac, D.; Groves, J. T.; Lear, J. D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 9656. (d) Sasaki, T.; Kaiser, E. T. *Biopolymers* **1990**, *29*, 79. (e) Sasaki, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1989**, *111*, 380.
5. Gibb, B. C.; Mezo, A. R.; Sherman, J. C. *Tetrahedron Lett.* **1995**, *36*, 7587.
6. Cram, D. J.; Cram, J. M. *Container Molecules and Their Guests*; The Royal Society of Chemistry: Cambridge, 1994.
7. Collet, A. *Tetrahedron* **1987**, *43*, 5725.
8. Lovejoy B.; Choe, S.; Cascio, D.; McRorie, D. K.; DeGrado, W. F.; Eisenberg, D. *Science* **1993**, *259*, 1288.
9. Collet, A.; Gabard, J.; Jacques, J.; Cesario, M.; Guilhem, J.; Pascard, C. *J. Chem. Soc., Perkin Trans. 1* **1981**, 1630.
10. (a) Kamtekar, S.; Hecht, M. H. *FASEB J.* **1995**, *9*, 1013. (b) DeGrado, W. F. *Adv. in Protein Chem.* **1988**, *39*, 51.
11. Stellwagen, E.; Park, S.-H.; Shalongo, W.; Jain, A. *Biopolymers* **1992**, *32*, 1193.
12. Cohen, C.; Parry, D. A. D. *Proteins: Structure, Function and Genetics* **1990**, *7*, 1.
13. Aurora, R.; Srinivasan, R.; Rose, G. D. *Science* **1994**, *264*, 1126.
14. Shoemaker, K. R.; Kim, P. S.; Brems, D. N.; Marqusee, S.; York, E. J.; Chaiken, I. M.; Stewart, J. M.; Baldwin, R. L. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 2349.
15. Gibb, B. C.; Mezo, A. R.; Causton, A. S.; Fraser, J. R.; Tsai, F. T. S.; Sherman, J. C. *Tetrahedron* **1995**, *51*, 8719.
16. Geier, G. R.; Sasaki, T. *Tetrahedron Lett.* **1997**, *38*, 3821.
17. (a) Futaki, S.; Sogawa, K.; Maruyama, J.; Asahara, T.; Niwa, M. *Tetrahedron Lett.* **1997**, *38*, 6237. (b) Dawson, P. E.; Kent, S. B. H. *J. Am. Chem. Soc.* **1993**, *115*, 7263.
18. Tuchscherer, G. *Tetrahedron Lett.* **1993**, *34*, 8419.
19. Nefzi, A.; Sun, X.; Mutter, M. *Tetrahedron Lett.* **1995**, *36*, 229.
20. (a) Frisch, B.; Boeckler, C.; Schuber, F. *Bioconjugate Chem.* **1996**, *7*, 180. (b) Haselgrübler, T.; Amerstorfer, A.; Schindler, H.; Gruber, H. J. *Bioconjugate Chem.* **1995**, *6*, 242. (c) Carlsson, J.; Drevin, H.; Axén, R. *Biochem. J.* **1978**, *173*, 723.
21. Semchuk, P. D.; Monera, L. H.; Kondejewski, L. H.; Gannon, C.; Daniels, L.; Wilson, I.; Hodges, R. S. *Peptides: Chemistry, Structure and Biology* **1996**, 73.
22. Futaki, S.; Kitagawa, K. *Tetrahedron* **1997**, 7479.
23. The maximum calculated helicities for peptides of 18 and 15 residues in length are −31,111 and −29,333 deg cm^2 dmol^{-1} respectively.[28] The $[\theta]_{222}$ values (and percent of the maximum calculated helicity) of the compounds studied were **1a** −23,600 (80%), **1b** −23,800 (77%), **1c$_f$** −19,500 (63%) and **1c$_s$** −20,100 (65%) deg cm^2 dmol^{-1}. At 275 μM, control peptide **4c** has a $[\theta]_{222}$ value of −26,600 (86%). The $[\theta]_{222}$ values also include the contribution from the CTB moiety. The three N-terminal residues of the peptides in TASPs **1b–c** are not expected to be part of the helix, nor are the glycine C-cap.
24. The properties of peptide **4c** were expected to vary according to its aggregation state (in terms of both number and orientation of the helices). In relation to this investigation, peptide **4c** serves to demonstrate that the properties of the TASP are not intrinsic to the peptide sequence alone, but due to the template-enhanced bundling. Peptides akin to **4c** have been investigated by other researchers.[29]
25. Nelson, J. W.; Kallenbach, N. R. *Proteins: Structure, Function and Genetics* **1986**, *1*, 211.
26. Keipert, S. J., Ph.D. thesis, University of California, Los Angeles. **1985**.
27. In Users Manual for *Applied Biosystems 431A Peptide Synthesizer* **1992**.
28. Lyu, P. C.; Sherman, J. C.; Chen, A.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 5317.
29. (a) Chmielewski, J.; Lipton, M. *Int. J. Peptide Protein Res.* **1994**, *44*, 152. (b) Osterhout, J. J.; Handel, T.; Na, G.; Toumadje, A.; Long, R. C.; Connolly, P. J.; Hoch, J. C.; Johnson, Jr. W. C.; Live, D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 331. (c) Ciesla, D. J.; Gilbert, D. E.; Feignon, J. *J. Am. Chem. Soc.* **1991**, *113*, 3957. (d) Hill, C. P.; Anderson, D. H.; Wesson, L.; DeGrado, W. F.; Eisenberg, D. *Science* **1990**, *249*, 543. (e) Ho, S. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751.