Template-Assembled Synthetic Proteins with Four-Helix-Bundle Topology. Total Chemical Synthesis and Conformational Studies

Manfred Mutter,*,† Gabriele G. Tuchscherer,‡ Charleen Miller,‡ Karl-Heinz Altmann,†,‖ Robert I. Carey,† Daniel F. Wyss,§ Alexander M. Labhardt,§ and Jean E. Rivier‡

Contribution from the Sêction de Chimie, Université de Lausanne, Rue de la Barre 2, CH-1005 Lausanne, Switzerland, Clayton Foundation Laboratories for Peptide Biology, The Salk Institute, 10010 North Pines Road, La Jolla, California 92037, and Hoffmann-La Roche Ltd., CH-4002 Basel, Switzerland. *Received July 15, 1991*

Abstract: The total chemical synthesis and conformational characterization of three template-assembled synthetic proteins or TASPs [designated T_4-($4\alpha_{11}$), T_4-($4\alpha_{15}$), and T_4-($3\alpha_{16},\alpha_{16}'$)] are described. These TASPs were designed to adopt globular folded structures of four-helix-bundle topology. The four constituent helical building blocks of the putative four-helix bundles were attached to the ϵ-amino groups of a cyclic carrier molecule (template) of the formula Ac-CysLysAlaLysProGlyLysAlaLysCys-NH_2, which serves to induce and direct intramolecular folding of the amphiphilic secondary structure elements. While T_4-($4\alpha_{11}$) and T_4-($4\alpha_{15}$) contain four identical helical building blocks, $11(\alpha_{11})$ and $15(\alpha_{15})$ amino acid residues in length, T_4-($3\alpha_{16},\alpha_{16}'$) incorporates two different sequences comprising 16 amino acid residues each attached to lysines 2, 4, 7, and 9 of the template, respectively. The TASPs were synthesized by solid-phase techniques and extensively purified by a combination of reversed-phase HPLC and high-performance ion-exchange chromatography. TASP molecules of very high purity could be obtained in this way, which demonstrates the capacity of solid-phase peptide synthesis for the rapid preparation of branched macromolecules of this size when used in combination with orthogonal analytical and purification techniques. Comparative conformational studies (isolated helical building blocks vs TASP) by means of CD and NMR spectroscopy in aqueous solution demonstrate that the template exerts a strong secondary structure stabilizing effect and suggest that T_4-($4\alpha_{15}$) as well as T_4-($3\alpha_{16},\alpha_{16}'$) adopts the proposed four-helix-bundle structure. This is further supported by the high conformational stability of T_4-($4\alpha_{15}$) under denaturing conditions, the guanidine hydrochloride induced folding → unfolding transition occurring at denaturant concentrations between 3 and 4 M. In contrast, T_4-($4\alpha_{11}$) exhibits a predominantly unordered conformation, indicating a critical chain length requirement for helix formation of >11 amino acid residues in the template-induced folding process.

Introduction

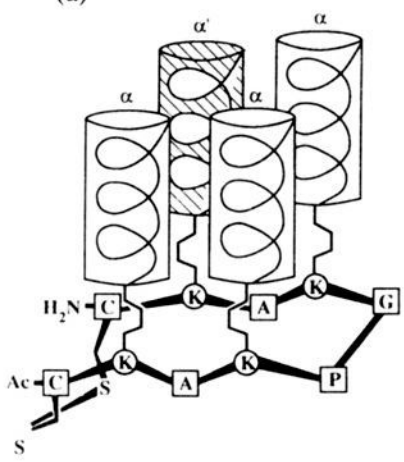
The folding into a well-defined globular structure represents the fundamental prerequisite for any natural protein to fulfill its biological function. Denaturation of a globular protein, i.e., loss of its structural (conformational) integrity, inevitably results in the loss of its biological activity. The development of polypeptide based, protein-like synthetic catalysts or receptors therefore requires, at a first stage, the design of amino acid sequences with the capacity to adopt protein-like globular structures.[1-3] However, at present the de novo design of proteins still poses a formidable challenge to peptide and protein chemists, as our current understanding of protein folding mechanisms does not allow us to reliably predict whether a given amino acid sequence will fold into a well-defined tertiary structure or not.[1-4] Only a few successful attempts on the de novo design of proteins have been reported in the literature, and none of the proposed structures has yet been ultimately confirmed by X-ray or NMR analysis.[5-10] In all of these cases the target molecules were *linear* polypeptide chains incorporating amphiphilic secondary structure blocks that were connected by short loops or turns. However, the design of linear sequences does not address two of the major inherent problems of polypeptide chain folding, i.e., the competition between intermolecular aggregation and intramolecular folding and the high loss in chain entropy associated with the folding process.[1,4] To overcome these limitations of the "linear chain" approach, we have developed a different concept for the de novo design of proteins that is based on the attachment of amphiphilic secondary structure blocks to a topological carrier molecule (or template)[1b,c,11-14] (Figure 1a). The resulting macromolecule is characterized by a branched-chain architecture and is termed "template-assembled synthetic protein" (TASP).[15] While the amphiphilicity of the secondary structure elements still is a necessary requirement for intramolecular folding of TASPs, the template reinforces and

† Université de Lausanne.
‡ The Salk Institute.
§ Hoffmann-La Roche Ltd.
‖ Present address: Ciba-Geigy AG, R-1060.2.14, CH-4002 Basel, Switzerland.

(1) (a) Mutter, M. *Angew. Chem., Int. Ed. Engl.* **1985**, *24*, 639-653. (b) Mutter, M.; Vuilleumier, S. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 535-554. (c) Altmann, K.-H.; Mutter, M. *Int. J. Biochem.* **1990**, *22*, 947-956.
(2) (a) DeGrado, W. F.; Wasserman, Z. A.; Lear, J. D. *Science* **1989**, *243*, 622-628. (b) DeGrado, W. F. *Adv. Protein Chem.* **1988**, *39*, 59-124.
(3) Richardson, J. S.; Richardson, D. C. *Trends Biochem. Sci.* **1989**, *14*, 304-309.
(4) For a comprehensive treatment of protein folding cf., e.g.: (a) Jaenicke, R., Ed. *Protein Folding*; Elsevier: Amsterdam, 1980. (b) Creighton, T. E. *Proteins—Structure and Molecular Properties*; Freeman: New York, 1984. For recent reviews on this subject cf., e.g.: (c) Creighton, T. E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5082-5086. (d) Montelione, G. T.; Scheraga, H. A. *Acc. Chem. Res.* **1989**, *22*, 70-76. (e) King, J. *Chem. Eng. News* **1989**, *67*, 32-54. (f) Baldwin, R. L. *Trends Biochem. Sci.* **1989**, *14*, 291-294. (g) Jaenicke, R. *Biochemistry* **1991**, *30*, 3147-3161.
(5) Gutte, B.; Däumigen, M.; Wittschieber, E. *Nature* **1979**, *281*, 650-655.
(6) Moser, R.; Thomas, R. M.; Gutte, B. *FEBS Lett.* **1983**, *157*, 247-251.
(7) (a) Lau, S. Y.; Taneja, A. K.; Hodges, R. S. *J. Biol. Chem.* **1984**, *259*, 13253-13261. (b) Hodges, R. S.; Semchak, P. D.; Taneja, A. K.; Kay, C. M.; Parker, J. M. R.; Mant, C. T. *Pept. Res.* **1988**, *1*, 19-30.
(8) Mutter, M.; Altmann, K.-H.; Vorherr, T. *Z. Naturforsch.* **1986**, *41B*, 1315-1322.
(9) (a) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976-978. (b) Handel, T.; DeGrado, W. F. *J. Am. Chem. Soc.* **1990**, *112*, 6710-6711.
(10) Hecht, M. H.; Richardson, J. S.; Richardson, D. C.; Ogden, R. C. *Science* **1990**, *249*, 884-891.
(11) Mutter, M. In *Peptides—Chemistry and Biology, Proceedings of the 10th American Peptide Symposium*; Marshall, G. R., Ed.; Escom: Leiden, 1988; pp 349-353.
(12) Mutter, M.; Hersperger, R.; Gubernator, K.; Müller, K. *Proteins* **1989**, *5*, 13-21.
(13) Mutter, M.; Altmann, E.; Altmann, K.-H.; Hersperger, R.; Koziej, P.; Nebel, K.; Tuchscherer, G.; Gremlich, H. U.; Müller, K. *Helv. Chim. Acta* **1988**, *71*, 835-847.
(14) More recently modified forms of this basic approach have also been adopted by others. Cf., e.g.: (a) Sasaki, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1989**, *111*, 380-381. (b) Sasaki, T.; Kaiser, E. T. *Biopolymers* **1990**, *29*, 79-88. (c) Hahn, K. W.; Wieslaw, A. K.; Stewart, J. M. *Science* **1990**, *248*, 1544-1547. (d) Montal, M.; Montal, M. S.; Tomich, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6929-6933.

0002-7863/92/1514-1463$03.00/0

(a)



α = L[1]EALEKAL[8]KEALAKLG[16]
α' = L[1]KALKEAF[8]EKAMAELG[16]

(b)

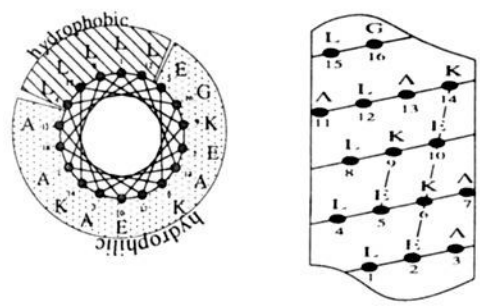


Figure 1. (a) Schematic representation of a four-helix-bundle TASP, T_4-(3α,α'), and amino acid sequences of α_{16} and α_{16}'. The C-terminal glycine is covalently attached via amide bond formation to the side chain of the template lysines. (b) Helical wheel[16] and net[21] representations illustrating amphiphilicity and favorable electrostatic interactions in the designed model helix α_{16}. Ac, acetyl; K, lysine; A, alanine; P, proline; G, glycine; C, cysteine; E, glutamic acid; L, leucine; F, phenylalanine; M, methionine.

directs this process, thus determining the folding topology of the final globular structure. On the basis of the cyclic decapeptide Ac-CysLysAlaLysProGlyLysAlaLysCys-NH_2 as the common template, we have now designed a series of new TASP molecules with the potential to adopt globular structures of four-helix-bundle topology. The choice of the template was suggested by molecular modeling studies which indicated a low-energy conformation for Ac-CysLysAlaLysProGlyLysAlaLysCys-NH_2 with the central (LysAlaLys) tripeptide units adopting an antiparallel β-sheet conformation, i.e., with all of the Lys side chains being located on the same face of the template. The design of the helical building blocks followed established principles for the stabilization of helical conformations, i.e., segregation of hydrophilic and hydrophobic side chains on the helix surface (amphiphilicity) and incorporation of residues with oppositely charged side chains in positions i and $i + 4$ of the sequence to allow for favorable intrachain charge-charge interactions (Figure 1b).[16-20]

The resulting target structures can be represented by the formula

Ac-
Cys-Lys(α1)-Ala-Lys(α1)-Pro-Gly-Lys(α1)-Ala-Lys(α2)-Cys-NH_2

T_4-(4α_{11}): α1 ≡ α2 ≡ α_{11} ≡ H-KALKEALAKLG-

T_4-(4α_{15}): α1 ≡ α2 ≡ α_{15} ≡ X-EALE-α_{11}-

T_4-(4α_{15}-H): X = H

T_4-(4α_{15}-Ac): X = CH_3CO

T_4-(3α_{16},α_{16}'): α1 ≡ α_{16} ≡ CD_3CO-L-α_{15}-
α2 ≡ α_{16}' ≡
CD_3CO-LKALKEAFEKAMAELG-

Due to the identity of the 4 peptides α_{15} of T_4-(4α_{15}) individual ^1H NMR resonance assignments could not be achieved, thus limiting the number of meaningful 2D NOESY experiments. To overcome this problem, one of the helices was modified in T_4-(3α_{16},α_{16}') such as to contain two unique residues (Phe[8] and Met[12]), whose most characteristic resonances (from the phenyl ring and the methyl group, respectively) are readily assignable. In addition, the positions of glutamic acid and lysine residues in α_{16}' are interchanged as compared to α_{16}, allowing for favorable *inter*chain charge–charge interactions between α_{16}' and the two adjacent helices. To further facilitate NMR investigations, all acetyl groups in T_4-(3α_{16},α_{16}') (at the N termini of the helices as well as of the template) were fully deuterated.

Once the design aspects of this research project were solved, we were still faced with the synthetic challenge of assembling the cyclic template and the different blocks with defined secondary structures.[1] Due to their nonnatural branched-chain architecture, TASPs cannot be prepared by recombinant DNA techniques but have to be synthesized chemically. However, the chemical synthesis of branched polypeptides of this size is associated with a variety of special problems.[1,22] While various synthetic strategies for the preparation of TASP molecules are conceivable, we critically evaluate here the potential of stepwise solid-phase peptide synthesis (SPPS) in combination with advanced analytical [reversed-phase (RP)-HPLC, high-performance ion-exchange chromatography (IEC), capillary zone electrophoresis (CZE), mass spectrometry] and purification (preparative RP-HPLC, IEC) techniques to provide TASPs of satisfactory purity. The other part of this work was dedicated to the investigation of the conformational properties of T_4-(4α_{11}), T_4-(4α_{15}), and T_4-(3α_{16},α_{16}') by means of circular dichroism (CD) and NMR spectroscopy. Special emphasis was put on the comparison of the conformational characteristics of TASPs and the respective isolated building blocks, covalently attached to the ε-amino group of Ac-Lys-NH_2, to mimic the situation within the TASP framework. These peptides are denoted Ac-Lys(α_{11})-NH_2, Ac-Lys(α_{15}-H)-NH_2, and Ac-Lys(α_{15}-Ac)-NH_2 and were also obtained by SPPS.

Experimental Section

Materials and Methods. Analytical reversed-phase high-performance *liquid chromatography* (RP-HPLC) was performed on a Hitachi 655A-11 liquid chromatograph equipped with a Hitachi 655A UV variable-wavelength UV monitor and a Hitachi 655A-61 data processor chart recorder using 0.46 × 25 cm stainless steel columns packed with Vydac C_{18} derivatized silica (5 μm, 300 Å) (purchased from the Separations Group, Hesperia, CA). Chromatograms were obtained at flow rates of 2 mL/min with UV detection at 210 nm. Preparative RP-HPLC was carried out on a Prep LC 3000 system from Waters in combination with a Fisher Recordall Series 5000 chart recorder and a Shimadzu SPD-6A UV detector. Semipreparative columns (1 × 25 cm, Vydac C_{18} silica, 5 μm) were purchased from the Separations Group; the preparative

(15) For reasons of simplicity the term "TASP" will be used throughout this paper, in a very liberal sense, for all molecules that incorporate the basic *design* features of the TASP approach. However, this does *not* automatically imply that all of these molecules do actually adopt a well-defined globular structure, i.e., are actual "synthetic proteins".

(16) Schiffer, M.; Edmundson, A. B. *Biophys. J.* **1967**, *7*, 121–135.
(17) Eisenberg, D.; Wilcox, W.; Eshita, S. M.; Pryciak, P. M.; Ho, S. P.; DeGrado, W. F. *Proteins* **1986**, *1*, 16–22.
(18) Ho, S. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751–6758.
(19) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8898–8902.
(20) Merutka, G.; Stellwagen, E. *Biochemistry* **1991**, *30*, 1591–1594.
(21) Dunnil, P. *Biophys. J.* **1968**, *8*, 865–875.
(22) Roth, W.; Heppenheimer, K.; Heidemann, E. R. *Makromol. Chem.* **1979**, *180*, 905–917.

columns were dry packed in our laboratory (Salk Institute) with Vydac C_{18} derivatized silica (15–20 μm). Empty polyethylene cartridges (30 × 5 cm) and frits were obtained from Waters. The following solvent systems were used in the RP-HPLC separations: (I) buffer A, 0.1% triethylammonium phosphate (TEAP) (pH 2.25), buffer B, 60% CH_3CN/40% A; (II) buffer A, 0.1% trifluoroacetic acid (TFA), buffer B, 0.1% TFA in 60% CH_3CN/40% H_2O; (III) buffer A, 0.9% TEAP (pH 2.25), buffer B, 60% CH_3CN/40% A; (IV) buffer A, 0.1% TFA, buffer B, 0.1% TFA in 80% CH_3CN/20% H_2O; (V) buffer A, 0.1% TEAP (pH 2.25), buffer B, 80% CH_3CN/20% A. Ion-exchange chromatography was done on a FPLC system from Pharmacia LKB Biotechnology. Analytical separations were accomplished using a Pharmacia Mono S precision column (cation exchanger) (0.16 × 5.0 cm) at a flow rate of 0.1 mL/min using UV detection at 214 nm. Preparative separations were performed on a Pharmacia Mono S HR 5/5 column (0.5 × 5.0 cm). The solvent system used for separations by IEC (system VI) was as follows: buffer A, 50 mM AcOH in 65% CH_3CN/35% H_2O (pH 5.0); buffer B, 0.7 M NaCl in A.

Capillary zone electrophoresis (CZE) was performed on a P/ACE 2000 capillary electrophoresis system from Beckman Instruments Inc. equipped with a Spectra Physics Chromjet integrator. The capillary cartridges (fused silica 50 cm × 75 μm i.d.) were also from Beckman. Standard conditions were 100 mM phosphate–NaOH buffer, pH 2.5, 30 °C, separation at 20 kV constant, 114 μA, and detection at 214 nm.

Amino acid analyses were performed after 24 h of hydrolysis in 4 N methanesulfonic acid on a Perkin-Elmer high-pressure liquid chromatograph using *o*-phthalaldehyde postcolumn derivatization in combination with fluorescence detection.

LSI mass spectra were measured with a JEOL JMS-HX 110 double-focusing mass spectrometer (JEOL, Tokyo) fitted with a Cs^+ gun. An accelerating voltage of 10 kV and a Cs^+ gun voltage of 25 kV were employed. The sample was added directly to a glycerol/3-nitrobenzyl alcohol (1:1) matrix.

Circular dichroism (CD) spectra were recorded on a JASCO 500 circular dichrometer using quartz cells of 0.05 and 0.1 cm path lengths. Ellipticities are expressed as mean residue ellipticities. If not noted otherwise, all spectra were taken at 25 °C. In the case of T_4-(4α_{15}-Ac) the material obtained after preparative RP-HPLC purification was used for the CD studies.

NMR spectra of aqueous solutions were recorded at 500.13 MHz (^1H) on a Bruker AM-500 spectrometer using a dedicated probehead optimized for spectroscopy in H_2O. The water resonance was suppressed by selective presaturation. All samples were dissolved in 90% H_2O/10% D_2O. No salt was added, and the pH was titrated to 3 by adding aliquots of concentrated HCl. Concentrations were in the range 1–5 mM. The low pH value was chosen to prevent the rapid exchange of labile amide protons,[23] which would result in the bleaching of the corresponding resonances. The TASPs T_4-(4α_{15}-Ac) and T_4-(3α_{16},α_{16}') exhibit nearly invariant α-helical-type CD spectra between 20 and 80 °C (data not shown). Two-dimensional NOESY spectra were acquired on nonspinning samples; 2048 real time data points in quadrature detection were collected in the phase-sensitive mode (TPPI)[24] for 512 t_1 increments. The mixing time was set to 150 ms. The water resonance was suppressed by continuous irradiation during the relaxation delay but not during the mixing time.

Solid-Phase Peptide Synthesis (SPPS). All protected amino acid derivatives were purchased from Bachem (Torrence, CA) and were directly used without further purification. Solvents were obtained from Aldrich. All syntheses were carried out manually starting from a methylbenzhydrylamine–1% cross-linked polystyrene (MBHA) resin prepared according to a modified literature procedure[25] (using *p*-toluoyl chloride instead of benzoyl chloride). With one exception (cf. synthesis of T_4-(3α_{16},α_{16}') N^{α}-BOC protection was used throughout and the general synthesis protocol followed established SPPS methodology:[26] BOC removal with 80% TFA and 1% ethanedithiol (EDT) in CH_2Cl_2 (25 min); washes with 1% EDT in 2-propanol, MeOH, and CH_2Cl_2; neutralization with 10% Et_3N/CH_2Cl_2; coupling with a 2-fold excess of activated amino acid derivative using either diisopropyl carbodiimide (DIC) [with or without 1-hydroxybenzotriazole (HOBt) as additive] or benzotriazol-1-yl-oxytris(dimethylamino)phosphonium hexafluorophosphate (BOP) as activating agents. Depending on the solubility of the BOC amino acid, couplings were performed in CH_2Cl_2 or CH_2Cl_2/DMF mixtures. Each coupling step was monitored by the qualitative ninhydrin test, and recouplings were performed as necessary. N-Terminal acetylations were performed with 10% Ac_2O in CH_2Cl_2 or in 1% pyridine/CH_2Cl_2. Cleavage of the peptides from the resin was accomplished with anhydrous HF in the presence of 5% anisole and 5% ethyl methyl sulfide using a Kel-F line (0 °C, 60–75 min) designed in-house (Salk Institute). After evaporation of HF, the peptides were precipitated with ether, collected by filtration, and redissolved in 20% acetic acid.

T_4-(4α_{11}) and T_4-(4α_{15}). The subsitution level of the starting MBHA resin was 0.33 mmol of NH_2/g. The linear protected precursor of the cyclic decapeptide template was assembled according to the general protocol except that the Pro-Gly unit was introduced by BOP-mediated coupling of the dipeptide BOC-Pro-Gly-OH. Cys sulfhydryls were protected by the *p*-methoxybenzyl (Mob) group, while the Lys side chains were protected by the 9-flourenylmethoxycarbonyl (Fmoc) group. Following removal of the Fmoc groups by treatment with 20% piperidine/DMF for 10 min, the synthesis was continued with the parallel assemblage of the four identical helical building blocks according to the standard protocol [Lys side chains protected by the 2-chlorobenzyloxycarbonyl (2-ClZ) group, Glu side chains as cyclohexyl (cHx) esters]. At the appropriate stages resin samples were removed and treated with anhydrous HF, and the crude products thus obtained were cyclized as will be described below for T_4-(3α_{16},α_{16}').

T_4-(4α_{11}-H) was purified by semipreparative RP-HPLC [(1) linear gradient from 15% B to 55% B in 80 min (system I) at a flow rate of 3.0 mL/min; (2) linear gradient from 10% B to 70% B in 20 min at a flow rate of 4.0 mL/min (system II, desalting)]. As judged by CZE and analytical RP-HPLC, the purity of T_4-(4α_{11}-H) was >90%.

T_4-(4α_{15}-H) was also purified by semipreparative RP-HPLC [(1) linear gradient from 40% B to 80% B in 80 min at 3.0 mL/min (system I); (2) linear gradient from 30% B to 90% B in 25 min at 3.0 mL/min (system II)]. As judged by CZE and analytical RP-HPLC, the purity of T_4-(4α_{15}-H) was >80%.

T_4-(4α_{15}-Ac) was purified by preparative RP-HPLC [(1) 55% B to 85% B in 60 min at 100 mL/min (system III); (2) 60% B to 90% B in 40 min at 100 mL/min (system II)]. The material thus obtained (estimated purity >90%) was further purified by cation-exchange chromatography on a Mono S HR 5/5 column [0% B to 50% B in 30 min (system VI)] followed by desalting on a preparative RP-HPLC column [40% B to 95% B in 20 min at 100 mL/min (system II)]. As judged by CZE and analytical RP-HPLC, this material was >95% pure.

T_4-(3α_{16},α_{16}'). The substitution level of the starting MBHA resin was 0.45 mmol of NH_2/g. After coupling of BOC-Cys(Mob) and Fmoc-Lys(BOC), the N^{ϵ}-BOC protecting group was removed and α_{16}' was assembled according to the general protocol. Following removal of the N^{α}-Fmoc protecting group (20% piperidine/CH_2Cl_2; 5 and 10 min) the template was completed using BOC-Lys(Fmoc) for the incorporation of the remaining lysines. The Fmoc groups were then removed and helices α_{16} were synthesized in parallel. Side chain protecting groups in the synthesis of helical blocks were 2-ClZ for Lys and cHex for Glu. All acetylations were carried out with fully deuterated acetic anhydride. After HF cleavage (1.5 g of peptide resin) and precipitation with ether, the peptide material was redissolved in 60 mL of 20% acetic acid. This solution was diluted with 3 L of water and then added dropwise to a vigorously stirred solution of $K_3[Fe(CN)_6]$ (200 mg) and NH_4OAc (10 g) in 1 L of water (total addition time ~4 h). During this period the pH was kept between 6.8 and 6.9 by simultaneous addition of 10% aqueous NH_4OH. The solution was stirred at 4 °C overnight, the pH was adjusted to 5 by addition of AcOH, and a slurry (~10 mL) of Bio-Rad AG3-X4A anion-exchange resin (analytical grade, 100–200 mesh, Cl^- form) was added. The mixture was stirred for 10 min and filtered, and the filtrate was applied to an ion-exchange column (15 mL gel bed volume) filled with AG3-X4A. The eluant was subsequently loaded on a cation-exchange column (30 mL gel bed volume) filled with Bio-Rex 70 resin (analytical grade, 100–200 mesh, H^+ form) where the peptide is initially retained and then eluted with 50% AcOH (~400 mL). After lyophilization, the crude T_4-(3α_{16},α_{16}') was purified by preparative HPLC [(1) 50% B to 90% B in 50 min at 105 mL/min (system V); (2) 50% B to 90% B in 20 min at 105 mL/min (system IV)] to yield 34 mg of lyophilized material; 21 mg of this material was further purified by high-performance IEC [0% B to 80% B in 20 min at 1.0 mL/min (system VI)] and then desalted by RP-HPLC [40% B to 95% B in 15 min at 1 mL/min (system IV)]. Total yield from seven runs was 9.5 mg. For all TASP syntheses mass spectral analysis gave the correct molecular ion peak and amino acid analyses gave the expected amino acid ratios. Sequence analysis of a T_4-(4α_{18}) obtained upon further addition of three amino acids to the T_4-(4α_{15}) was carried out on an ABI pulsed liquid-

(23) Molday, R. S.; Englander, S. W.; Kallen, R. G. *Biochemistry* **1972**, *11*, 150–158.

(24) Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967–974.

(25) Penke, B.; Rivier, J. *J. Org. Chem.* **1987**, *52*, 1197–1200.

(26) (a) Barany, G.; Merrifield, R. B. *The Peptides—Analysis, Synthesis, Biology*; Gross, E., Meienhofer, J., Eds.; Academic Press: New York, 1980; Vol. 2, pp 1–284. (b) Stewart, J. M.; Young, J. D. *Solid Phase Peptide Synthesis*, 2nd ed.; Pierce Chemical: Rockford, IL, 1984.

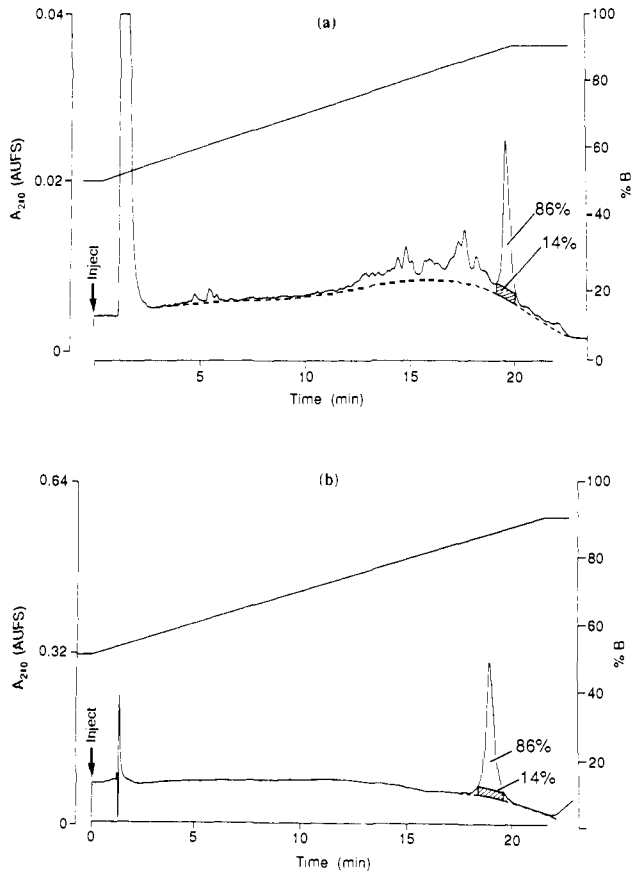


Figure 2. (a) Analytical RP-HPLC of crude cyclized T_4-($3\alpha_{16},\alpha_{16}'$). (b) Analytical RP-HPLC after preparative RP-HPLC purification. Gradient: 50% B to 90% B over 40 min (system IV).

phase protein sequencer column. No amino acid preview was observed, suggesting the absence of significant amounts of deletion sequences.

Ac-Lys(α_{11}-H)-NH_2, Ac-Lys(α_{15}-H)-NH_2, and Ac-Lys(α_{15}-Ac)-NH_2 were synthesized according to standard solid-phase procedures and purified by RP-HPLC. As demonstrated by analytical RP-HPLC, they were all >98% pure.

Results

Peptide Synthesis. As detailed under Experimental Section, all TASPs were synthesized by N^{α}-BOC-based SPPS on a MBHA resin.[26,27] Two slightly different strategies had to be followed for TASPs containing four identical helices [T_4-($4\alpha_{11}$) and T_4-($4\alpha_{15}$)] and T_4-($3\alpha_{16},\alpha_{16}'$), respectively. While the former could be prepared by parallel assemblage of all of the helical building blocks subsequent to the synthesis of the full-length (linear) template and removal of the N^{ϵ}-Fmoc protecting groups from lysines, the latter was synthesized by first assembling the unique helix α_{16}' [attached to the ϵ-amino group of resin-bound Fmoc-Lys-Cys(Mob)] followed by completion of the template and subsequent parallel synthesis of helices α_{16}. The completed peptides were simultaneously deprotected and cleaved from the resin by treatment with liquid HF, and the crude material was directly cyclized with $K_3[Fe(CN)_6]$ under conditions of infinite dilution. Figure 2a shows an analytical RP-HPLC trace of crude T_4-($3\alpha_{16},\alpha_{16}'$) after cyclization and subsequent concentration on a cation-exchange column. Similar chromatograms were also obtained in the syntheses of T_4-($4\alpha_{11}$), T_4-($4\alpha_{15}$-H), and T_4-($4\alpha_{15}$-Ac). The major peak was isolated by preparative RP-HPLC using a TEAP buffer/CH_3CN system and then repurified on the same column but applying a 0.1% TFA buffer/CH_3CN system. The analytical RP-HPLC trace of this material is depicted in Figure 2b and demonstrates the dramatic improvement in the purity of T_4-($3\alpha_{16},\alpha_{16}'$) that could be achieved in the RP-HPLC purification steps. However, on the basis of the appearance of the original chromatogram of the crude peptide (Figure 2a) it seemed conceivable that up to 14% of the material still had to be accounted for by various coeluting impurities. This contention was confirmed by CZE (Figure 3a) and high-performance IEC (chromatogram not shown) analysis, revealing the presence of nonnegligible amounts of impurities that had not been resolved by RP-HPLC. Further purification by high-performance IEC yielded a preparation of T_4-($3\alpha_{16},\alpha_{16}'$) which we conservatively estimate to be >90% pure (Figure 3b). This was additionally confirmed by amino acid analysis and LSI mass spectrometry (Figure 4). TASP T_4-($4\alpha_{15}$-Ac) was taken through an analogous purification procedure, resulting in a material with an estimated purity of >95%. TASPs T_4-($4\alpha_{11}$) and T_4-($4\alpha_{15}$-H), which were only purified by preparative RP-HPLC, were obtained in purities of >90% and >80%, respectively.

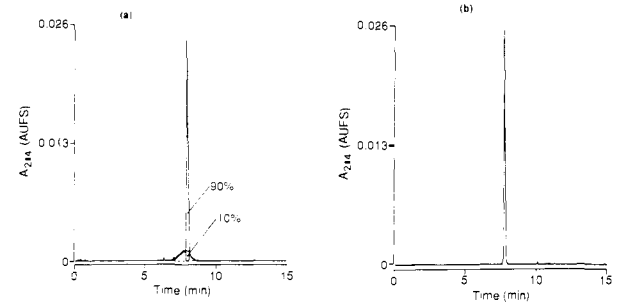


Figure 3. Electropherogram of T_4-($3\alpha_{16},\alpha_{16}'$) after preparative RP-HPLC purification and before (a) and after (b) further purification by high-performance IEC.

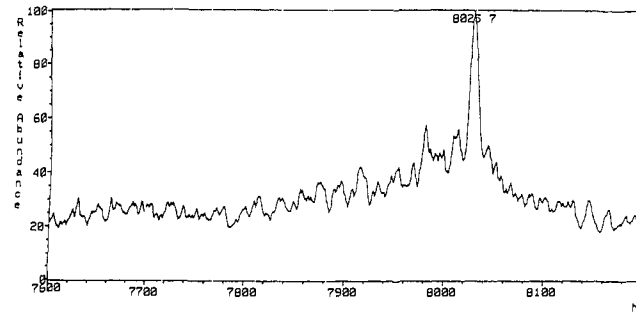


Figure 4. LSI mass spectrum of T_4-($3\alpha_{16},\alpha_{16}'$) after purification by RP-HPLC and ion-exchange chromatography. Accelerating voltage, 10 kV; CsI ion gun voltage, 25 kV; postaccelerating voltage, 20 kV.

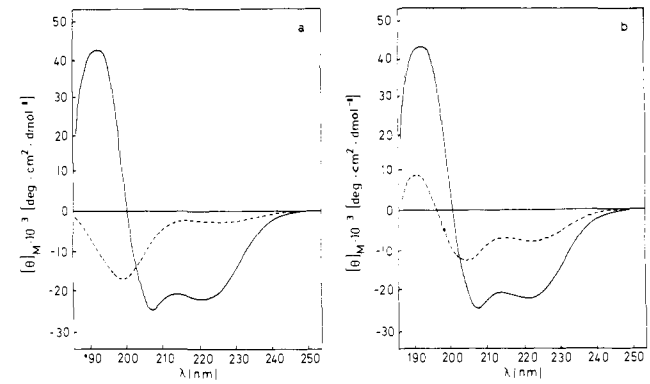


Figure 5. CD spectra in aqueous phosphate buffer (50 mM phosphate, pH 7.0; from Merck): (a) T_4-($4\alpha_{11}$) (---) and T_4-($4\alpha_{15}$-H) (—), $c = 10^{-4}$ M; (b) Ac-Lys(α_{15}-H)-NH_2 (---), $c = 4 \times 10^{-4}$ M, and T_4-($4\alpha_{15}$-H) (—), $c = 10^{-4}$ M.

Circular Dichroism. The CD spectra of TASPs T_4-($4\alpha_{11}$) and T_4-($4\alpha_{15}$-H) in aqueous solution at pH 7 are shown in Figure 5a. The spectrum of T_4-($4\alpha_{11}$) is characterized by a strong negative Cotton effect centered around 198 nm, which indicates that the peptide exists in a predominantly unordered conformation under

(27) (a) Barany, G.; Kneib-Cordonier, N.; Mullen, D. G. *Int. J. Pept. Protein Res.* **1987**, *30*, 705–739. (b) Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161–214.

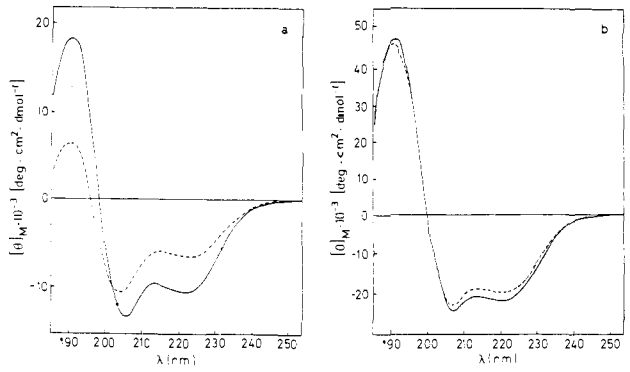


Figure 6. CD spectra in aqueous phosphate buffer (50 mM phosphate, pH 7): (a) Ac-Lys(α_{15}-H)-NH_2 (---) and Ac-Lys(α_{15}-Ac)-NH_2 (—), $c = 4 \times 10^{-4}$ M; (b) T_4-($4\alpha_{15}$-H) (—) and T_4-($4\alpha_{15}$-Ac) (---), $c = 10^{-4}$ M.

these conditions. A similar spectrum was obtained for Ac-Lys-(α_{11})-NH_2 (data not shown). In contrast, the CD curve of T_4-($4\alpha_{15}$-H) displays the typical features of peptides in a α-helical conformation, i.e., strong negative Cotton effects at 222 ($[\theta]_M$ = −22 000 deg cm² dmol⁻¹) and 207 nm ($[\theta]_M$ = −25 000 deg cm² dmol⁻¹), a zero-crossover at 202 nm, and a strong positive Cotton effect at 193 nm. The spectrum is unaffected by temperature changes between 0 and 25 °C. Figure 5b compares the CD spectra of T_4-($4\alpha_{15}$-H) and its constituent helical building block attached to the ϵ-amino group of Ac-Lys-NH_2. While the latter peptide definitely exhibits some degree of helicity in water at pH 7, the attachment of α_{15}-H to the cyclic decapeptide template results in a dramatic increase in secondary structure content [$[\theta]_{222}$ = −6400 deg cm² dmol⁻¹ for Ac-Lys(α_{15}-H)-NH_2 vs −22 000 for T_4-($4\alpha_{15}$-H)], indicating a strong secondary structure inducing effect of the template. As opposed to T_4-($4\alpha_{15}$-H) the CD spectrum of Ac-Lys(α_{15}-H)-NH_2 proved to be strongly concentration dependent, increasing peptide concentration leading to an increased degree of helicity. Striking differences in the conformational behavior of T_4-($4\alpha_{15}$-H) and α_{15} are also observed upon N-terminal acetylation of the helical building blocks in T_4-($4\alpha_{15}$-H) and Ac-Lys(α_{15}-H)-NH_2, respectively. As can be seen from Figure 6a, acetylation of the isolated building block significantly enhances helix stability [$[\theta]_{222}$ = −10 300 deg cm² dmol⁻¹ for Ac-Lys-(α_{15}-Ac)-NH_2 vs −6400 deg cm² dmol⁻¹ for Ac-Lys(α_{15}-H)-NH_2]. It is important to note that in 10^{-4}–10^{-3} M solutions no further increase in the helical content of the single helix Ac-Lys(α_{15}-Ac)-NH_2 could be observed; consequently, the lower degree of helicity (even at higher concentrations of the single helix) compared to that of the TASP molecule clearly points to a stabilization of the helical bundle structure by the template. On the other hand, acetylation of the helix N termini in T_4-($4\alpha_{15}$-H) causes only marginal changes in the CD spectrum, which is a rather surprising finding in view of the expected helix destabilization by a positively charged N terminus and the results obtained for Ac-Lys(α_{15}-Ac)-NH_2 vs Ac-Lys(α_{15}-H)-NH_2. The conclusions that may be derived from these findings will be discussed below.

NMR Spectroscopy. Figure 7 shows the amide proton region of the 1D NMR spectra of the cyclic decapeptide Ac-CysLysAlaLysProGlyLysAlaLysCys-NH_2 (Figure 7a), of Ac-Lys(α_{15}-Ac)-NH_2 (Figure 7b), and of T_4-($4\alpha_{15}$-Ac) (Figure 7d) in aqueous solution at pH 3 and 47 °C. The resonances labeled "NH_2" in Figure 7a (and also Figure 7d,e) can be assigned to the protons of the C-terminal primary amide group of Cys^{10} of the template. In Figure 7a resonances of ϵ-amino protons of lysines are not visible due to rapid exchange with the solvent. The labels "NH_2" and "ϵ-NH Lys" in Figure 7b [Ac-Lys(α_{15}-Ac)-NH_2] denote the primary amide proton resonances and the resonances of the ϵ-NH group, respectively, of the C-terminal Lys. Included in Figure 7 (Figure 7c) is also the superposition in a 4:1 ratio of spectra b and a, respectively, of Figure 7, thus creating an artificial spectrum for a TASP molecule lacking any interactions between the template-attached peptide blocks [except for additional resonances arising from the primary amide function and the α-NH proton of the C-terminal Lys in Ac-Lys(α_{15}-Ac)-NH_2]. It is important to note that the actual NMR spectrum of T_4-($4\alpha_{15}$-Ac) is distinctly different from the spectrum depicted in Figure 7c and thus does not represent a simple superposition of the spectra of its constituent components. This suggests, even in the absence of detailed resonance assignments, the occurrence of significant conformational changes upon covalent attachment of α_{15}-Ac to the template. The 1D NMR spectrum of T_4-($3\alpha_{16},\alpha_{16}'$) in water at pH 3 and 47 °C is shown in Figure 7e. The absence of any aromatic residues other than Phe^8 in α_{16}' allows straightforward assignment of the aromatic proton resonances of this residue (labeled "Phe 8"). Except for these additional signals, the 1D spectrum of T_4-($3\alpha_{16},\alpha_{16}'$) in the region between 7.0 and 8.5 ppm has the same appearance as that of T_4-($4\alpha_{15}$-Ac). An excerpt from the 2D NOESY spectrum of T_4-($3\alpha_{16},\alpha_{16}'$) is depicted in Figure 8. Three crosspeaks are observed between an aromatic proton resonance at 7.14 ppm and three different resonances arising from γ-methyl groups of Leu residues; it follows that *at least* two different leucines, possibly more, must be involved in close contacts with Phe^8 of α_{16}'. The implications of this finding for the conformational characterization of T_4-($3\alpha_{16},\alpha_{16}'$) will be evaluated under Discussion.

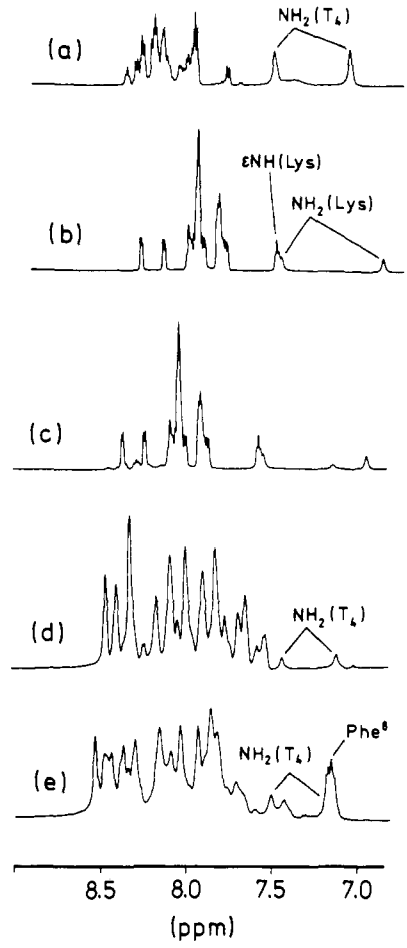


Figure 7. Amide region of the 1D ^1H NMR spectra in water, pH 3.0, at 47 °C. (a) Ac-CysLysAlaLysProGlyLysAlaLysCys-NH_2 (template); (b) Ac-Lys(α_{15}-Ac)-NH_2; (c) superposition in a 4:1 ratio of spectra b and a; (d) T_4-($4\alpha_{15}$-Ac); (e) T_4-($3\alpha_{16},\alpha_{16}'$). Resonances labeled "$NH_2(T_4)$" pertain to the C-terminal primary amide protons of the template and those labeled "ϵNH(Lys)" and "NH_2(Lys)" to the ϵ-NH proton and the primary amide protons, respectively, of the C-terminal Lys in Ac-Lys-(α_{15}-Ac)-NH_2. "Phe^8" denotes the aromatic proton resonances in α_{16}' of T_4-($3\alpha_{16},\alpha_{16}'$).

Denaturation Experiments. Denaturation studies on T_4-($4\alpha_{15}$-H) were performed by following the ellipticity changes in the CD spectrum at 222 nm as a function of guanidine hydrochloride (Gdn-HCl) concentration. To elucidate the effect of the template,

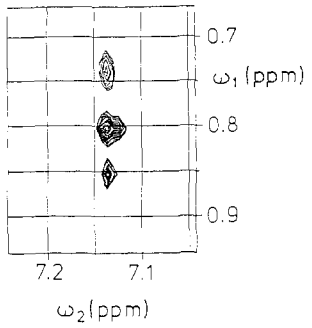


Figure 8. Aromatic (ω_2) vs methyl (ω_1) region of a NOESY spectrum of T_4-($3\alpha_{16},\alpha_{16}'$) recorded in H_2O, pH 3.0, at 47 °C with a mixing time of 150 ms.

identical peptide concentrations (1 mg of peptide/mL) were used. As can be seen in Figure 9, the Gdn-HCl-induced folding → unfolding transition occurs at ~3 M denaturant concentration. No complete transition curve could be obtained for Ac-Lys-(α_{15}-H)-NH_2) due to its low helix content even at zero denaturant concentration.

Discussion

One of the principal objectives of this work was to critically evaluate the potential of stepwise SPPS for the synthesis of TASP molecules of high chemical purity. The synthesis of TASPs by SPPS poses a number of problems that are not normally encountered in the stepwise solid-phase synthesis of linear peptides, e.g., the requirement for N^{ϵ}-protecting groups of template lysines exhibiting orthogonal stability with respect to the N^{α}-protecting groups *and* the anchoring group between peptide and resin as well as the parallel synthesis of several peptide chains attached to the side chains of residues that are part of the same template peptide. Based on the current status of solid-phase methodology[26,27] the most reasonable option for the synthesis of TASPs T_4-($4\alpha_{11}$), T_4-($4\alpha_{15}$), and T_4-($3\alpha_{16},\alpha_{16}'$) seemed to consist in a synthetic scheme employing an HF cleavable linker group (MBHA), a TFA-labile N^{α}-protecting group (BOC), HF-cleavable protecting groups for permanent side chain protection, and a base-labile, but completely TFA-stable, protecting group (Fmoc) for temporary purposes [on either the ϵ- or the α-amino group of template Lys residues (cf. Results)]. As demonstrated by the RP-HPLC trace depicted in Figure 2a for T_4-($3\alpha_{16},\alpha_{16}'$), the crude mixture of reaction products after cyclization clearly contains a readily identifiable major component rather than a number of principal products having formed in comparable amounts. This is in agreement with the fact that no special problems were encountered in the course of the parallel assemblage of helices α_{16} [and also α_{15} and α_{11} in the case of T_4-($4\alpha_{15}$) and T_4-($4\alpha_{11}$), respectively], allowing the completion of TASP syntheses within a matter of days using standard solid-phase methodology. The initial purification of crude TASPs by RP-HPLC proved fairly straightforward, and the purified material eluted as a single peak from an analytical RP-HPLC column showed the expected amino acid composition and gave the correct molecular ion peak upon LSI mass spectral analysis. However, reanalysis of peptides purified by RP-HPLC using the same chromatographic technique exclusively, although often practiced, does not represent an entirely reliable tool for the assessment of the quality of the purified material (as indicated in Figure 2 by the shaded areas at the base of the major peak). Further analysis of RP-HPLC-purified TASPs by techniques involving separation mechanisms distinctly different from those of RP-HPLC (high-performance IEC and CZE) indeed revealed a certain heterogeneity of these TASP preparations that had not been evident in the RP-HPLC analysis (Figure 3). Clearly, this finding not only bears important implications for the purification of TASP molecules but touches on the question of peptide purity in general, if long sequences (>30–40 amino acid residues) are purified and reanalyzed by RP-HPLC exclusively. The purity of TASPs T_4-($4\alpha_{15}$-Ac) and T_4-($3\alpha_{16},\alpha_{16}'$) could be markedly improved by a second purification step employing preparative high-performance IEC (cf. Figure 3), resulting in

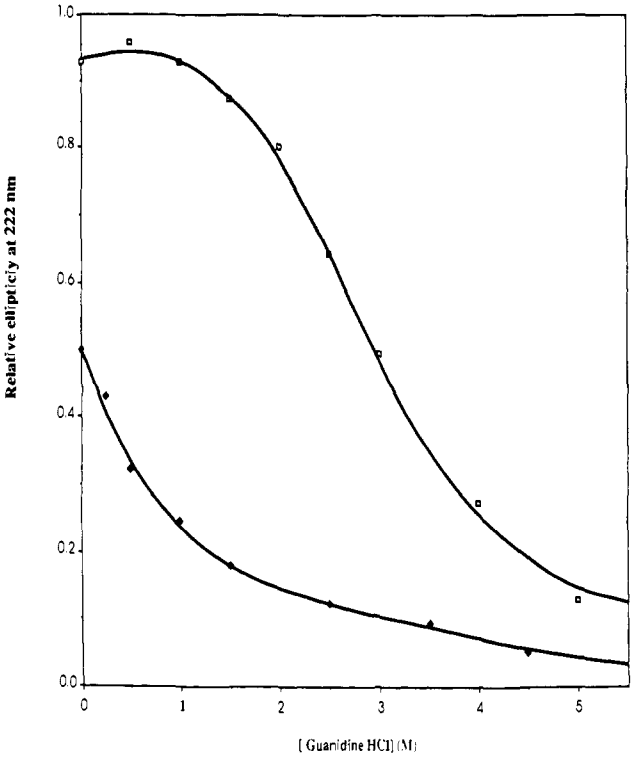


Figure 9. Denaturation of T_4-($4\alpha_{15}$-H) (□) (concentration = 1 mg of peptide/mL) by increasing concentrations of guanidine hydrochloride (guanidine HCl). The solutions were prepared in 50 mM phosphate and adjusted to pH 7 by addition of 4 N NaOH. The curve for the single helix block, Ac-Lys(α_{15}-H)-NH_2 (c = 1 mg of peptide/mL) is shown for comparison (◆).

preparations of very conservatively estimated purities of >95% and >90%, respectively. No comparable levels of purity have been achieved in previous TASP syntheses by stepwise SPPS[11-13] that did not incorporate purification schemes based on *two-dimensional* chromatographic techniques.[28] A few examples have been reported in the more recent literature[14] of the synthesis and partial purification of TASP-like molecules having molecular weights similar to that of T_4-($4\alpha_{15}$) and T_4-($3\alpha_{16},\alpha_{16}'$). However, from the paucity of analytical data provided in those papers it appears at least questionable whether the purity of the synthetic material was comparable to that of our preparations of T_4-($4\alpha_{15}$-Ac) and T_4-($3\alpha_{16},\alpha_{16}'$).

The second major objective of this study was the detailed investigation of the TASP structural properties to further evaluate the general applicability of the TASP concept in the area of protein de novo design. The TASPs investigated in this study differ from previously designed structures[11-13] in several important respects: (i) While all previous TASPs were based on *linear* peptide templates containing appropriately located putative β-turn motifs, T_4-($4\alpha_{11}$), T_4-($4\alpha_{15}$), and T_4-($3\alpha_{16},\alpha_{16}'$) incorporate a *cyclic* peptide template that is expected to enhance the thermodynamic stability of the designed four-helix-bundle structures. (ii) TASPs T_4-($4\alpha_{15}$-Ac) and T_4-($3\alpha_{16},\alpha_{16}'$) are of unprecedented chemical purity, which has allowed us for the first time to conduct reliable structural studies by 1D and 2D ^1H NMR techniques. This also required (iii) the design of an appropriate four-helix-bundle TASP [T_4-($3\alpha_{16},\alpha_{16}'$)] incorporating helical building blocks of different amino acid sequences as the four-fold symmetry of T_4-($4\alpha_{15}$-Ac) essentially precludes unambiguous resonance assignments and thus meaningful 2D NOESY experiments.

The results of the conformational studies by means of CD spectroscopy clearly establish a strong helix stabilizing effect of

(28) We have, however, recently succeeded in the synthesis of highly pure 4- and 6-helix-bundle TASPs employing a fragment condensation strategy.[29,30]

(29) Ernest, I.; Vuilleumier, S.; Fritz, H.; Mutter, M. *Tetrahedron Lett.* **1990**, *31*, 4015–4018.

(30) Steiner, V.; Mutter, M. Unpublished results.

the cyclic template. This is convincingly demonstrated by the dramatic difference in helicity between T_4-($4\alpha_{15}$-H) and Ac-Lys(α_{15}-H)-NH_2 [the respective helix content being ~66% and ~20%, using helical poly(Lys) as a reference[31,32]] (Figure 5b) that can be best rationalized by *intra*molecular helix stabilization within the TASP framework, especially since no concentration or temperature dependence of the CD spectrum of T_4-($4\alpha_{15}$-H) was observed. Most notably, the helix content of the single helical block could not be increased by increasing the concentration ($[\theta]_{222}$ being constant in the range between 10^{-4} and 10^{-3} M solutions); thus, the difference in the helix content, as well as in the thermodynamic stability (Figure 9), between T_4-($4\alpha_{15}$-H) and the self-associating single helices α_{15}-H must originate in a helix stabilizing template effect in the TASP molecule. This is further supported by the distinctly different effects caused by N-terminal acetylation of Ac-Lys(α_{15}-H)-NH_2 and the template-attached helices in T_4-($4\alpha_{15}$-H), respectively (Figure 6). While the increase in helicity of Ac-Lys(α_{15}-H)-NH_2 upon acetylation can be readily explained by the elimination of unfavorable electrostatic interactions between a positively charged N terminus and the helix dipole,[33-35] this argument does not hold for T_4-($4\alpha_{15}$-H) as no significant differences were found between the CD spectra of T_4-($4\alpha_{15}$-H) and T_4-($4\alpha_{15}$-Ac). The latter finding implies that the conformational equilibrium between helix and coil forms of the template-attached peptide blocks in T_4-($4\alpha_{15}$-H) is already shifted to the helical side to such a degree that no sizable further increase in helix content can be induced by N-terminal acetylation, at least not under folding conditions. The secondary structure inducing effect of the template thus outweighs unfavorable interactions of a positively charged helix N terminus with the helix dipole[35] and even the electrostatic repulsion between charged N termini of helices associated in a parallel fashion. The conclusions derived from CD studies are corroborated by the results of 1D and 2D ^{1}H NMR experiments on T_4-($4\alpha_{15}$-Ac) and T_4-($3\alpha_{16},\alpha_{16}'$). Although no sequence-specific resonance assignments were achieved for T_4-($4\alpha_{15}$-Ac), thus limiting the applicability of 2D NOESY experiments, some useful information can be extracted from the 1D NMR spectrum when analyzed in combination with the spectra of the isolated cyclic decapeptide template and Ac-Lys(α_{15}-Ac)-NH_2. The linear combination of the latter in a 1:4 ratio results in a theoretical spectrum of T_4-($4\alpha_{15}$-Ac) with noninteracting building blocks (Figure 7c) that is typical of peptides without a preferred conformation in solution. In contrast, the recorded spectrum of T_4-($4\alpha_{15}$-Ac) (Figure 7d) is strongly reminiscent of spectra obtained for globular proteins [despite the fact that T_4-($4\alpha_{15}$-Ac) contains only five different amino acids (Leu, Ala, Glu, Lys, Gly)], indicating significant intramolecular interactions among the template-attached helical building blocks. More direct evidence for intramolecular folding could be obtained for T_4-($3\alpha_{16},\alpha_{16}'$) by 2D NOESY experiments, due to straightforward assignment of aromatic proton resonances of the unique Phe^8 residue in α_{16}'. Three crosspeaks were observed in the NOESY spectrum between the aromatic proton resonance at 7.14 ppm and three different resonances corresponding to Leu γ-methyl protons (Figure 8). None of the spectra (e.g., ROESY spectra measured in identical conditions; data not shown) display exchange crosspeaks or give any other evidence for multiple conformations of T_4-($3\alpha_{16},\alpha_{16}'$) in slow exchange. Hence, the most likely interpretation for the three NOESY cross-peaks in Figure 8 is that they have to be accounted for by close spatial proximity of Phe^8 and *at least* two different Leu residues. Assuming a helical conformation of α_{16}', the only conceivable *intra*chain contact of Phe^8 (with a Leu residue) is with Leu^4 of α_{16}', which leads to the conclusion that Phe^8 must also be in close spatial proximity to Leu residues of α_{16} chains. This is as expected for a four-helix-bundle structure of T_4-($3\alpha_{16},\alpha_{16}'$), where Phe^8 of α_{16}' would be located directly adjacent to three Leu residues 8 of the α_{16} chains. Due to the high degree of sequence similarity between T_4-($3\alpha_{16},\alpha_{16}'$) and T_4-($4\alpha_{15}$-Ac), their similar CD spectra [data not shown for T_4-($3\alpha_{16},\alpha_{16}'$)],[36] and the similar appearance of the amide region of their 1D ^{1}H NMR spectra, the conclusions derived for T_4-($3\alpha_{16},\alpha_{16}'$) are likely to apply also to T_4-($4\alpha_{15}$-Ac).

While all of the above data demonstrate the strong helix promoting effect of the template or explicitly establish close *inter*chain residue contacts in TASP molecules, they do not strictly *prove* the formation of a globular structure of four-helix-bundle topology. However, given the topology of the template (with a low-energy conformation having the side chains of all of the template lysines pointing in the same direction), the high frequency of occurrence of the four-helix-bundle motif in proteins.[37,38] and the (concentration-dependent) tendency even of isolated amphiphilic peptides to form tetrameric aggregates,[17,18] it seems likely that the observed template-induced stabilization of helices indeed occurs via the formation of monomeric four-helix bundles. Interestingly, it was recently demonstrated in a theoretical study[39] that the formation (and also the unfolding) of isolated four-helix-bundle structures is a highly cooperative process, with no intermediates incorporating two or three interacting helices being detected in these calculations. This points to the marginal stability of intramolecular helix dimers and trimers as compared to that of four-helix-bundle structures and agrees well with our experimental finding of a cooperative Gdn-HCl-induced folding $\rightleftharpoons$ unfolding transition for T_4-($4\alpha_{15}$-H). With a denaturant concentration at the midpoint of the transition of ~3 M the stability of T_4-($4\alpha_{15}$-H) is roughly comparable to that of globular proteins of similar molecular weight.[2a,14b]

In contrast to T_4-($4\alpha_{15}$) and T_4-($3\alpha_{16},\alpha_{16}'$), T_4-($4\alpha_{11}$) definitely exists in a predominantly unordered conformation in aqueous solution (Figure 3a). This finding has important implications for the future design of TASP molecules, as it demonstrates that even for template-attached peptide sequences a chain length of more than 11 amino acid residues is required to form stable helical structures.

Conclusions

We have accomplished the total chemical synthesis of TASP molecules of unprecedented purity by means of stepwise SPPS. This demonstrates that SPPS can definitely serve as a suitable tool for the preparation of TASPs incorporating up to almost 80 amino acid residues provided that rigorous analytical and purification protocols are applied. The use of *orthogonal* preparative as well as analytical techniques (RP-HPLC vs CZE and IEC, respectively) proved to be of decisive importance for the successful purification and the proper assessment of the purity of our target peptides. Conformational studies conducted on these molecules by CD as well as NMR spectroscopy clearly suggest that they adopt the designed folded structure of four-helix-bundle topology, if the length of the helical building blocks is in the range of about 15 amino acid residues. This reaffirms the potential of the TASP concept for the construction of artificial tertiary structures, the use of a cyclic peptide template resulting in TASPs of thermodynamic stabilities that are comparable to those of globular proteins of similar molecular weight. The molecules presented in this study may thus serve as a starting point for the future design of artificial proteins that not only structurally but also functionally resemble their natural counterparts. Studies along these lines are presently under way in our laboratories.

Acknowledgment. We gratefully acknowledge financial support by the Swiss National Science Foundation. Parts of this research were also supported by NIH Grants HD-13527 and DK-26741

(31) Greenfield, N.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108-4116.
(32) Due to uncertainties with respect to the CD contribution of the template and in view of the lack of any significant increase in helicity upon acetylation of the helix termini of T_4-($4\alpha_{15}$-H), these numbers are very likely to represent only very rough estimates.
(33) Hol, W. G. J.; van Duijnen, P. T.; Berendson, H. J. C. *Nature* **1978**, *273*, 443-446.
(34) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Nature* **1987**, *326*, 563-567.
(35) Chou, K.-C.; Maggiora, G. M., Nēmethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4295-4299.
(36) Toniolo, C.; Mutter, M.; et al. Unpublished data.
(37) Richardson, J. S. *Adv. Prot. Chem.* **1981**, *34*, 167-339.
(38) Presnell, S. R.; Cohen, F. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6592-6596.
(39) Sikorski, A.; Skolnick, J. *Biopolymers* **1989**, *28*, 1097-1113.

and the Hearst Foundation. R.I.C. gratefully acknowledges Postdoctoral Training Grant GM 13678-01 from the National Institutes of Health. MS measurements were carried out in part at the Rockefeller University Biotechnology Research Resource supported by the division of Research Resources, NIH Grant RR 862. We are indebted to Drs. E. Gassmann and F. Raschdorf, Ciba-Geigy AG, Basel, Switzerland, for initial CZE and MS analyses, respectively. Finally, we thank Dr. A. Craig for taking the LSI mass spectrum and John Dyckert for the sequence analysis.

Design of a Nonnatural Deoxyribonucleoside for Recognition of GC Base Pairs by Oligonucleotide-Directed Triple Helix Formation

Jong Sung Koh and Peter B. Dervan*

Contribution from the Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology, Pasadena, California 91125. *Received August 8, 1991*

Abstract: The deoxyribonucleoside 1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]pyrimidin-7-one (**P1**) was designed such that two specific hydrogen bonds would form with guanine (G) of a Watson–Crick guanine–cytosine (GC) base pair in the major groove of double-helical DNA. One edge of the **P1** heterocycle mimics N3-protonated cytosine, which would circumvent the pH dependence observed for the formation of triple helices containing C + GC base triplets. **P1** was synthesized in five steps and incorporated by automated methods in pyrimidine oligodeoxyribonucleotides. From affinity cleaving analyses, the stabilities of base triplets decrease in the order P1·GC ≫ P1·CG ≫ P1·AT ~ P1·TA (pH 7.4, 35 °C). **P1** binds GC base pairs within a pyrimidine triple-helix motif as selectively and strongly as C but over an extended pH range. Oligodeoxyribonucleotides containing **P1** residues were shown to bind within plasmid DNA a single 15 base pair site containing five GC base pairs at pH 7.8 and a single 16 base pair site containing six *contiguous* GC base pairs at pH 7.4.

Introduction

Oligodeoxyribonucleotide-directed triple-helix formation offers a powerful chemical approach for the sequence-specific recognition of double-helical DNA.[1,2] Pyrimidine oligodeoxyribonucleotides bind specifically to purine sequences in double-helical DNA to form a local triple-helical structure.[1,2] These oligonucleotides bind in the major groove of DNA parallel to the purine Watson–Crick (W–C) strand through the formation of specific (Hoogsteen) hydrogen bonds.[1-4] Specificity is derived from thymine (T) recognition of adenine–thymine (AT) base pairs (T·AT triplets) and protonated cytosine (C+) recognition of guanine–cytosine (GC) base pairs (C + GC triplets). Complexes of triple-helical nucleic acids containing cytosine (C) and thymine (T) on the Hoogsteen strand are stable in acidic to neutral solutions but dissociate with increasing pH.[1-4] An important factor is the required protonation of the cytosine N3 in the third strand to enable the formation of two hydrogen bonds.[3] Moreover, the stability of triple-helical complexes depends on the number and positions of C + GC triplets. In particular, we have observed that the stability of complexes containing *contiguous* C + GC triplets decreases markedly with increasing pH.[5] Oligodeoxyribonucleotides which contain 5-methylcytosine (mC) bind duplex DNA with greater affinities and over an extended pH range.[1c,2e,6] However, protonation of N3 is still required (mC + GC triplets). For the binding of *contiguous* tracts of GC base pairs on double-helical DNA by the pyrimidine motif[1,2] over a wide range of pH, even oligodeoxyribonucleotides containing 5-methylcytosine may be inadequate.

We report the design of nonnatural bases that bind GC base pairs without protonation in a pyrimidine-motif triple-helical complex.[7,8] Our rationale was to use a molecular frame that would place hydrogen bond donors and acceptors in the appropriate positions to form two specific hydrogen bonds to GC base pairs while maintaining a phosphate–deoxyribose backbone geometry compatible with the pyrimidine triple-helix motif. From mod-

(1) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645. (b) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927. (c) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059. (d) Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 7286. (e) Maher, L. J., III; Wold, B. J.; Dervan, P. B. *Science* **1989**, *245*, 725. (f) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967. (g) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435. (h) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73. (i) Maher, L. J., III; Dervan, P. B.; Wold, B. J. *Biochemistry* **1990**, *29*, 8820. (j) Plum, G. E.; Park, Y. W.; Singleton, S.; Dervan, P. B.; Breslauer, K. T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436. (k) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172.

(2) (a) Praseuth, D.; Perrouault, L.; Le Doan, T.; Chassignol, M.; Thuong, N. T.; Lhomme, J.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349. (b) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D.; Cantor, C. R. *Nucleic Acids Res.* **1988**, *16*, 2165. (c) Francois, J. C.; Saison-Behmoaras, T.; Thuong, N. T.; Helene, C. *Biochemistry* **1989**, *28*, 9617. (d) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Nucleic Acids Res.* **1990**, *18*, 157. (e) Collier, D. A.; Thuong, N. T.; Helene, C. *J. Am. Chem. Soc.* **1991**, *113*, 1457.

(3) (a) Rajagopal, P.; Feigon, J. *Nature* **1989**, *239*, 637. (b) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859. (c) de los Santos, C.; Rosen, M.; Patel, D. *Biochemistry* **1989**, *28*, 7282.

(4) For early studies on triple helical nucleic acid polymers, see: (a) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023. (b) Hoogsteen, K. *Acta Crystallogr.* **1959**, *12*, 822. (c) Lipsett, M. N. *Biochem. Biophys. Res. Commun.* **1963**, *11*, 224. (d) Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256. (e) Howard, F. B.; Frazier, J.; Lipsetts, M. N.; Miles, H. T. *Biochem. Biophys. Res. Commun.* **1964**, *17*, 93. (f) Miller, J. H.; Sobell, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1201. (g) Michelson, A. M.; Massoulie, J.; Guschlbauer, W. *Prog. Nucleic Acid Res. Mol. Biol.* **1967**, *6*, 83. (h) Felsenfeld, G.; Miles, H. T. *Annu. Rev. Biochem.* **1967**, *36*, 407. (i) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63. (j) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucleic Acids Res.* **1979**, *6*, 3073.

(5) Kiessling, L. L.; Dervan, P. B. Unpublished observations.

(6) Polynucleotides which contain 5-methylcytosine (mC) form triplexes (mC·G·mC) up to pH 8. Lee, J. S.; Woodsworth, M. L.; Latimer, L. J. P.; Morgan, A. R. *Nucleic Acids Res.* **1984**, *12*, 6603.

(7) Purine oligodeoxyribonucleotides bind purine tracts of double helical-DNA at extended pH (GGC and AAT or TAT triplets); see: (a) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456. (b) Beal, P. B.; Dervan, P. B. *Science* **1991**, *251*, 1360. (c) Pilch, D. S.; Levenson, D.; Shafer, R. H. *Biochemistry* **1991**, *30*, 6081.

(8) 2'-Methylpseudoisocytidine will substitute for 2'-deoxycyctidine in triplex formation at neutral pH. The sequence specificity and thermodynamic stability for binding contigious GC base pairs at extended pH is not yet known. See: Ono, A.; Tso, P. O. P.; Kan, L. *J. Am. Chem. Soc.* **1991**, *113*, 4032.

0002-7863/92/1514-1470$03.00/0