Protein Science (2000), *9*:942–955. Cambridge University Press. Printed in the USA.

Synthesis and NMR solution structure of an α-helical hairpin stapled with two disulfide bridges

PHILIPPE BARTHE,[1] SANDRINE ROCHETTE,[2] CLAUDIO VITA,[2] AND CHRISTIAN ROUMESTAND[1]

[1]Centre de Biochimie Structurale, CNRS-UMR 9955, INSERM-U414, Université de Montpellier I, Faculté de Pharmacie, 15 Avenue Charles Flahault, 34060 Montpellier Cedex, France
[2]Commissariat à l'Energie Atomique, Département d'Ingénierie et d'Etude des Protéines, C.E. Saclay, 91191 Gif-sur-Yvette, France

(RECEIVED January 3, 2000; FINAL REVISION March 24, 2000; ACCEPTED March 30, 2000)

Abstract

Helical coiled-coils and bundles are some of the most common structural motifs found in proteins. Design and synthesis of α-helical motifs may provide interesting scaffolds that can be useful as host structures to display functional sites, thus allowing the engineering of novel functional miniproteins. We have synthesized a 38-amino acid peptide, α_2p8, encompassing the α-helical hairpin present in the structure of p8^{MTCP1}, as an α-helical scaffold particularly promising for its stability and permissiveness of sequence mutations. The three-dimensional structure of this peptide has been solved using homonuclear two-dimensional NMR techniques at 600 MHz. After sequence specific assignment, a total of 285 distance and 29 dihedral restraints were collected. The solution structure of α_2p8 is presented as a set of 30 DIANA structures, further refined by restrained molecular dynamics, using simulated annealing protocol with the AMBER force field. The RMSD values for the backbone and all heavy atoms are 0.65 ± 0.25 and 1.51 ± 0.21 Å, respectively. Excised from its protein context, the α-hairpin keeps its native structure: an α-helical coiled-coil, similar to that found in superhelical structures, with two helices spanning residues 4–16 and 25–36, and linked by a short loop. This motif is stabilized by two interhelical disulfide bridges and several hydrophobic interactions at the helix interface, leaving most of its solvent-exposed surface available for mutation. This α-helical hairpin, easily amenable to synthetic chemistry and biological expression system, may represent a stable and versatile scaffold to display new functional sites and peptide libraries.

Keywords: conformational peptide library; disulfide bridges; helical coiled-coils; NMR spectroscopy; solution structure; structural scaffolds

Engineering small and simplified structural motifs is of primary interest in protein folding, because it may provide models of early intermediates in protein folding as well as precious information on the forces stabilizing protein structure. In addition, newly designed structures may be very useful to develop "maquettes" of much larger and complex biological systems (Johnsson et al., 1993; Robertson et al., 1994; Gibney et al., 1996, 1997) to test reaction mechanisms, or to obtain small-size mimics of biological relevant proteins, representing potential leads in drug design (Cunningham & Wells, 1997).

α-Helical bundles and coiled coil have been the first de novo designed motifs (DeGrado et al., 1989; Hecht et al., 1990; Kamtekar et al., 1993), because the structural determinants governing α-helical structures are understood to a satisfactory level (Schafmeister & Stroud, 1998). However, early de novo designed α-helical structures lacked the well-packed hydrophobic cores, the thermodynamic stability, and cooperative unfolding transition, which are characteristic of natural proteins. These constructions adopted a more dynamic structure, characteristic of a molten globule conformation (Handel et al., 1993; Bryson et al., 1995). In the past decade, however, the engineering of independently folded motifs has reached a remarkable level of complexity and sophistication. The 3D structures of several de novo proteins are known (Raleigh et al., 1995; Struthers et al., 1996; Schafmeister et al., 1997; Dahiyat & Mayo, 1997; Harbury et al., 1998; Walsh et al., 1999), and fully support the success of these new attempts in the design of α-helical bundles. Also, structures all-β, difficult to obtain in early

Reprint requests to: Christian Roumestand, Centre de Biochimie Structurale, CNRS-UMR 9955, INSERM-U414, Université de Montpellier I, Faculté de Pharmacie, 15 Avenue Charles Flahault, 34060 Montpellier Cedex, France; e-mail: roume@cbs.univ-montp1.fr.

Abbreviations: 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; (^{2}H)TSP, 3-methylsilyl-[2,2,3,3-$^2H_4$]propionate; CD, circular dichroism; DQF-COSY, double-quantum-filtered scalar-correlated spectroscopy; Fmoc, fluorenylmethyloxycarbonyl; HPLC, high-performance liquid chromatography; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; RMSD, root-mean-square deviation; *z*-TOCSY, *z*-filtered total correlation spectroscopy; TFE, 2,2,2-trifluoroethanol.

designs because of solubility difficulties (Quinn et al., 1994; Bianchi et al., 1994), have been finally successfully engineered (Ilyina et al., 1997; Kortemme et al., 1998; De Alba et al., 1999). Motifs containing combined α/β elements have been also obtained by iterative design and analysis procedure (Struthers et al., 1996) or by an automated approach utilizing a computational design algorithm, based on physical chemical potential functions and stereochemical constraints (Dahiyat & Mayo, 1997).

An alternative approach to de novo design, which is particularly suited for the production of novel proteins expressing a specific biological activity, is based on the utilization of stable natural structures as scaffolds, thus taking advantage of millions years of evolution aimed, among other things, to the selection of stable scaffolds. New functions are then incorporated into these frameworks by (1) the transfer of an exogenous functional site to scaffold regions structurally compatible with the selected site (Vita, 1997); (2) random substitution of residues at the molecular surface combined with a screening assay, leading to selection of variants with a specific biological function (Nygren & Uhlén, 1997); and (3) a combination strategy where a natural structural motif is minimized and optimized for both structural stability and biological function by phage display methodology (Cunningham & Wells, 1997).

The choice of the scaffold relies on several parameters: as a general rule, these frameworks should be preferably small, soluble, robust, easily engineered, and effectively produced in large scale using either chemical synthesis or low-cost expression systems. α-Helical coiled-coils or bundles represent what is probably the most widespread structural motif found in proteins, and support extremely diverse functions, from DNA binding to muscle regulation and dimer assembly (Gentz et al., 1989; Landschultz et al., 1989; Cohen & Parry, 1990; O'Shea et al., 1991). Coiled-coils are made of two, three, or four right-handed amphipathic α-helices that wrap around each other in a left-handed supercoil with a crossing angle of $\sim$20° between helices, such that their hydrophobic surfaces are in continuous contact to form dimeric, trimeric, or tetrameric coiled-coils (Crick, 1953; Cohen & Parry, 1990). Their formation is dependent primarily on the presence of heptad repeat sequence, denoted *a b c d e f g* (McLachlan & Stewart, 1975), where positions *a* and *d* are characteristically occupied by hydrophobic residues, thus forming the core of the coiled-coil, while the *e* and *g* positions, often occupied by charged residues, flank the hydrophobic interface and pack against the residues of the hydrophobic core, and *b*, *c*, *f* are generally hydrophilic and exposed to the solvent. Numerous studies have been devoted to the design of small stable coiled-coil structures that either have been obtained by de novo design (DeGrado et al., 1989; Zhou et al., 1992a, 1992b; Kuroda et al., 1994; Fezoui et al., 1994; Myszka & Chaiken, 1994; Walsh et al., 1999) or by recruiting naturally existing coiled-coil motifs in proteins or domains for further engineering (de Wolf et al., 1996; Predki et al., 1996; Braisted & Wells, 1996; Starovasnik et al., 1997; Domingues et al., 1999).

The structure of p8^{MTCP1} could provide an interesting "natural" α-hairpin scaffold, potentially useful in the engineering of α-helical protein mimics. P8^{MTCP1} (Soulier et al., 1994) is a small 8 kDa protein coded by the human oncogene *MTCP1* (also called *c6.1B*) (Fisch et al., 1993; Stern et al., 1993), a gene unequivocally identified in the heterogeneous group of uncommon leukemias having a mature phenotype. Although the participation of p8^{MTCP1} to the oncogenesis cannot be discarded, this mitochondrial protein is expressed at low levels in most human tissues, suggesting that p8^{MTCP1} may be associated with a function common to many cell types. Its solution structure has been recently solved (Barthe et al., 1997) and reveals a novel scaffold consisting of three α-helices, associated with an unusual cysteine motif. Two of the helices are paired covalently by two disulfide bridges and form an α-hairpin, which exhibits numerous structural similarities with a classical antiparallel helical coiled-coil. The third helix is orientated roughly parallel to the plane defined by the antiparallel α-helical motif, and its axis forms an angle of about 60° with the main axis of this motif. A disulfide bonds between Cys39 and Cys50, and only few hydrophobic contacts link this helix to the α-hairpin. The present study reports on the synthesis and on the 3D solution structure determination of a p8^{MTCP1} peptide analogue, named α_2p8, that comprises the antiparallel α-helical hairpin. This peptide, 38-amino acid long, stabilized by two internal disulfides, may represent a potential, stable, and minimal mimic of a two-stranded α-helical scaffold, useful to display peptide libraries or exogenous functional sites.

Results

Design and synthesis

Analysis of p8^{MTCP1} structure (Barthe et al., 1997) suggested that its three-helix system could be simplified into an antiparallel two-helix motif (Fig. 1). In fact, helix III is well exposed to the solvent and is linked to helix II by a flexible loop, a disulfide bond and few hydrophobic contacts only. Furthermore, helix III is less well defined than the two other helices, and its backbone exhibits a strongly different dynamic behavior, as demonstrated by a ^{15}N relaxation study (Barthe et al., 1999), suggesting a reduced contribution of this helix to the stability of the antiparallel α-helical motif. Thus, a new sequence encompassing the α-hairpin motif of p8^{MTCP1} was designed as follow: (1) helix III was deleted starting from residue 43; (2) the four N-terminal residues, disordered in the native pro-

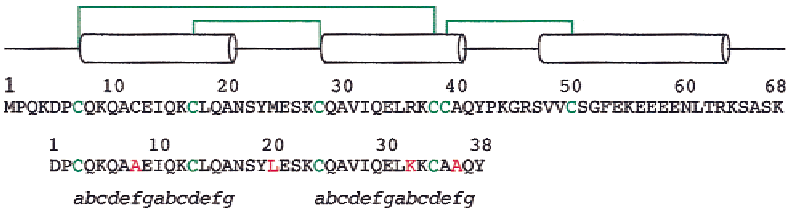


Fig. 1. Alignment of the α_2p8 sequence (below) and wild-type p8^{MTCP1} (above). Cysteine residues involved in disulfide bonds are in green, mutated residues of α_2p8 in red. The secondary structure is schematized at the top, with the positioning of the disulfide bridges in green lines. Italic letters below the α_2p8 sequence indicate the heptad repeats.

tein, were deleted, leaving an Asp residue at the N-terminus with a potential N-capping role; (3) Cys39 of helix II was mutated to alanine, to eliminate potential undesired disulfide bonds; (4) unpaired Cys12 in helix I was mutated to alanine (in a previous work (Barthe et al., 1999), we have demonstrated that mutation Cys12Ala does not alter the geometry of the native protein structure); (5) Met24 (in the loop connecting the two helices) and Arg36 (in position e' were mutated into Leu and Lys, respectively, to facilitate the synthesis. The 5–42 α-hairpin of $p8^{MTCP1}$, in this way modified, was renumbered and dubbed α_2p8 (Fig. 1). The resulting 38 amino acid peptide was synthesized by solid-phase peptide synthesis in a fully automated peptide synthesizer, by using Fmoc chemistry. Monitoring of deprotection steps during synthesis indicates that the designed sequence did not present synthetic difficulties (not shown). The crude product, analyzed in reversed-phase HPLC, presented a major peak representing about 85% of the total eluted material (Fig. 2A). Even if $p8^{MTCP1}$ forms the native disulfide bonds spontaneously and in high yields (Barthe et al., 1997), the newly designed α_2p8 sequence might have lost the original folding efficiency, as a consequence of sequence modifications and deletions. Thus, Cys13–Cys24 and Cys3–Cys34 disulfide bonds were formed, one by one, by using during synthesis two different and orthogonal protecting groups for cysteine residues, i.e., trityl for Cys13, Cys24, and acetamidomethyl for Cys3, Cys34; subsequent cysteine bonding was performed by using specific chemical reactions (see Materials and methods). Purity and identity of the purified peptide were verified by analytical HPLC (Fig. 2B), amino acid analysis (not shown), and electrospray mass spectrometry (4,268.6 Da, determined mass; 4,268.9 Da theoretical mass). The oxidized peptide was obtained in 28% overall yields.

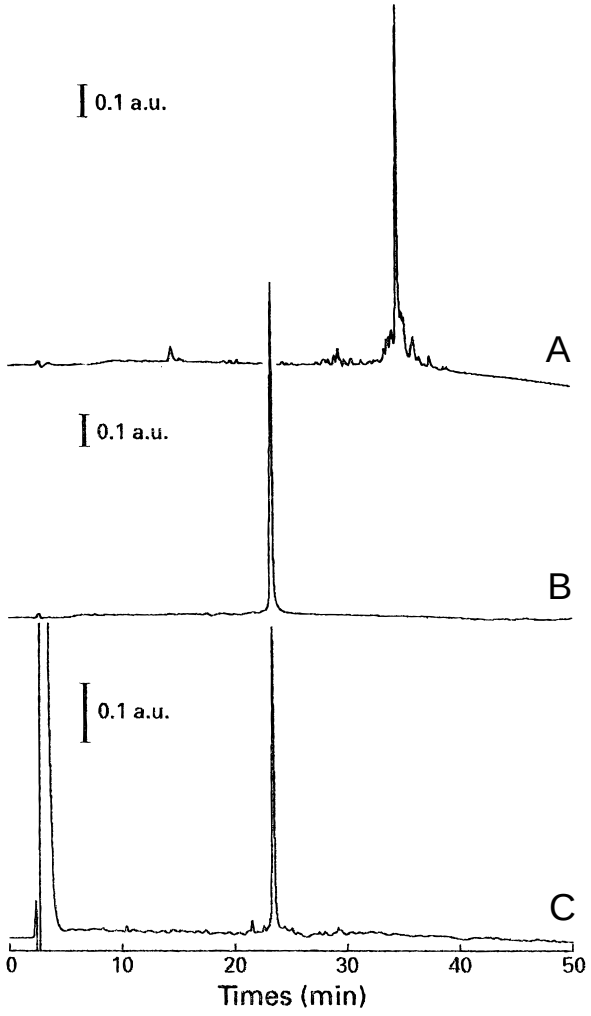


Fig. 2. Analytical reversed-phase HPLC profiles of (**A**) crude, (**B**) oxidized and purified, and (**C**) glutathione refolded α_2p8 peptide. A Vydac C18 column (0.46 × 15 cm) was used, equilibrated in 0.1% TFA and eluted with a linear 0–80% acetonitrile gradient over 50 min at 1.0 mL/min flow rate. Detection was at 215 nm.

The designed α_2p8 sequence was also tested for its ability to form the native disulfide bonds spontaneously. Thus, purified reduced α_2p8 was incubated in a redox buffer, containing reduced and oxidized glutathione (Saxena & Wetlaufer, 1970) (see Materials and methods). HPLC analysis clearly indicates that the peptide is readily oxidized in the presence of glutathione, because it is eluted at the same retention time than the species possessing the native disulfides (Fig. 2C). Thus, the designed α_2p8 sequence maintains the capability to form the native disulfide bonds efficiently. The conservation of this property in the newly designed sequence is expected to increase purification yields significantly in future chemical syntheses (as a consequence of the utilization of only one protecting group for all four cysteine residues) and in biological expression systems.

Conformational characterization

The CD spectra of α_2p8 and of the peptide (denominated CM-α_2p8) presenting the four cysteines carboxamidomethylated (see Materials and methods) are shown in Figure 3B. The molar ellipticities were measured in mild conditions (5 mM phosphate buffer, pH 7.5) and in the presence of increasing concentrations of TFE. TFE is considered to be a noninteracting (inert) solvent that induces helicity in a single-chain polypeptide, potentially α-helical (Nelson & Kallenbach, 1986). α_2p8 exhibits a spectrum characteristic of an α-helix conformation with double minima at 222 and 208 nm as well as a maximum at 192 nm, and the value of $[\theta]_{222}$ indicates a helical content of ≈65% (Chen et al., 1974) in mild condition. This content increases until ≈80% at high TFE concentration. This result strongly suggests that α_2p8 is predominantly α-helical in water, and that this geometry is further stabilized by TFE addition. On the other hand, CM-α_2p8 (the peptide devoid of disulfide bridges) appears less ordered in mild conditions, with less than 15% of helical content as deduced from its CD spectrum. However, its helicity increases considerably upon addition of TFE: 20% TFE is able to induce a CD spectrum similar to that of α_2p8 and, at higher concentration of TFE, the carboxamidomethylated peptide increases its ordered α-helical conformation up to 90% helicity. This result indicates a high propensity of this peptide to adopt a helical conformation. Interestingly, at high TFE concentration CM-α_2p8 acquires a proportion of α-helical ordered structure that is higher than that of α_2p8, suggesting that the structure of the two peptides are different in those conditions. The existence of ordered structure for α_2p8 in water is further demonstrated by the ^{1}H-NMR spectrum, which, especially in the amide proton region, is much better resolved in the case of α_2p8 compared to CM-α_2p8 (Fig. 3A). In the following study, we focus our analysis only on α_2p8. From NMR diffusion experiments, we obtained a value of $133 \pm 1\ \mu m^2\,s^{-1}$ for the self-diffusion coefficient (D_s) of α_2p8, corresponding to a molecular weight of 4.4 ± 0.1 kDa (see

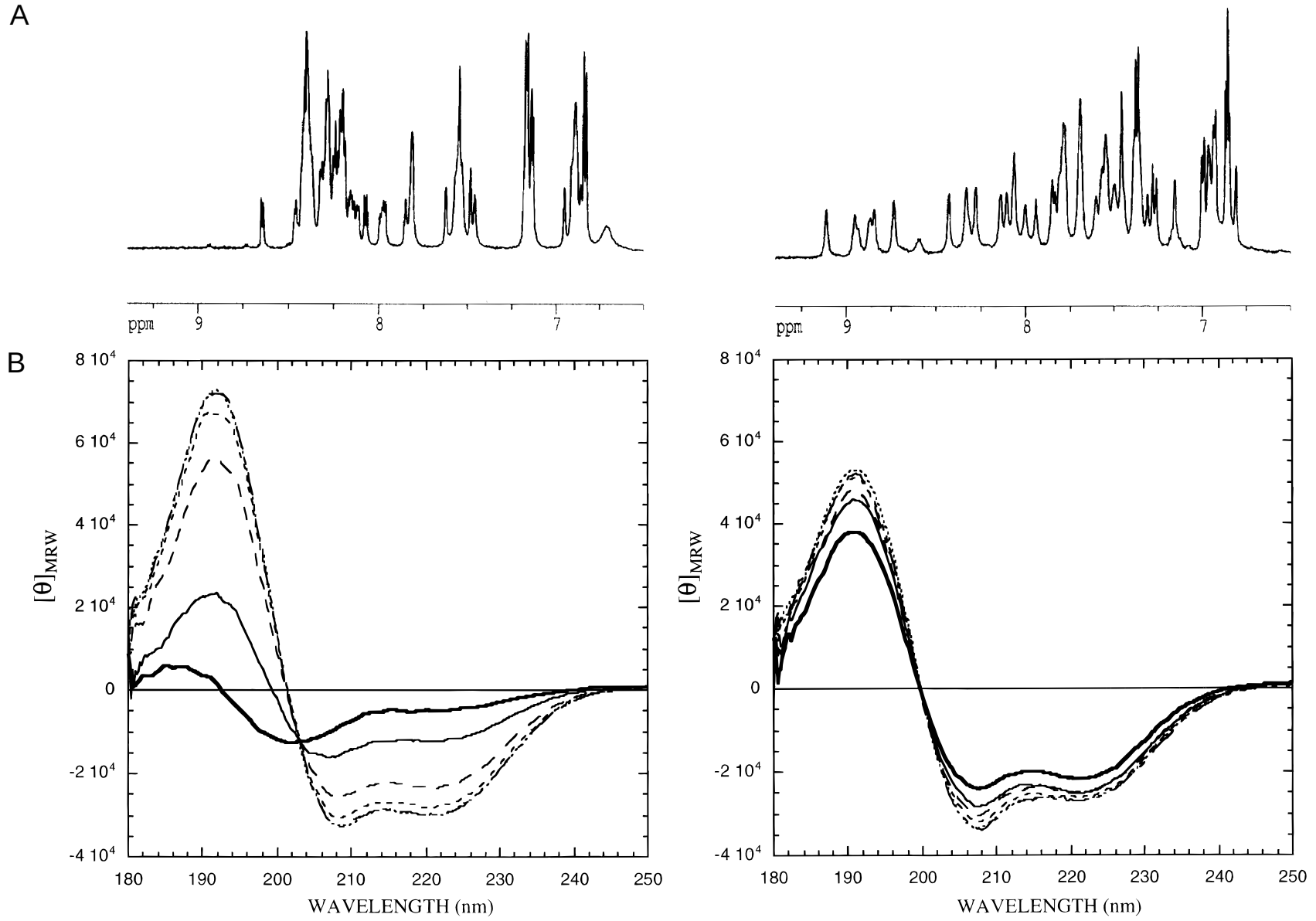


Fig. 3. **A:** Amide region of the 600 MHz 1D spectra of (right) α_2p8 and (left) CM-α_2p8 recorded on samples dissolved in H_2O, pH 5.5. **B:** CD spectra of (right) α_2p8 and (left) CM-α_2p8 in mild conditions (5 mM sodium phosphate, pH 7.5) (bold line) and in the presence of increasing concentrations of TFE (thin line) added at the following percentage: 10, 20, 40, 60, and 80.

Materials and methods): this indicates that the peptide is monomeric in solution under the experimental conditions used for the study.

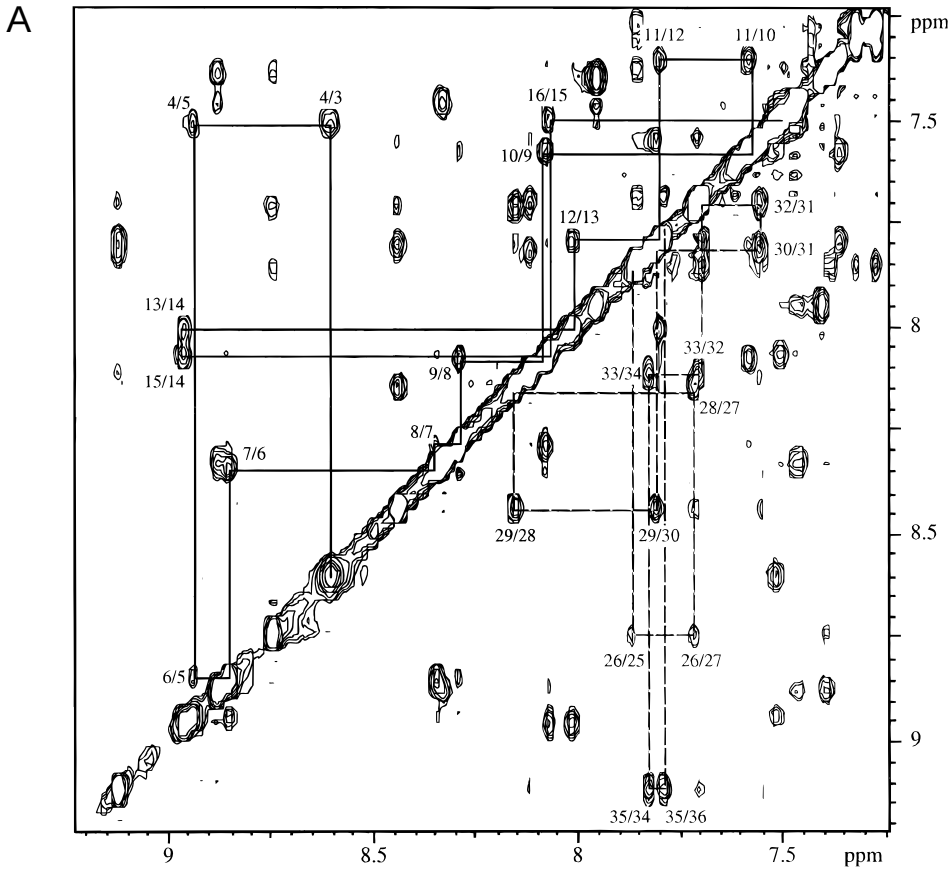
Structure resolution

Sequence-specific assignment of the proton resonances was performed according to the method of Wüthrich (1986). The procedure involves: (1) identification of the amino acid spin systems on the basis of 2D TOCSY spectra and (2) differentiation of identical spin systems as well as sequence-specific assignment using the sequential $d_{\alpha N}$ and d_{NN} cross peaks in the 2D NOESY spectra. At this step, the assignment of the backbone resonances was considerably facilitated by the generally small difference in chemical shift found between the protons of corresponding residues in α_2p8 and in native p8^{MTCP1}. Significant differences were only found near mutated residues and at the N- and C-termini, probably due either to the presence of additional residues or to the vicinity of helix III in p8^{MTCP1}. This already suggests a high similarity between the structure of α_2p8 and of the α-hairpin motif of p8^{MTCP1}.

The chemical shifts of the ^{1}H resonances identified for the individual amino acid residues are tabulated in Table 1. A total of 29 residues out of the 38 have their C$^\alpha$H proton resonance shifted upfield from the random-coil values, indicating a 3D fold rich in α-structures (Wishart et al., 1992). This was further supported by the presence of two large stretches of sequencial d_{NN} and medium-range NOE contacts ($\leq i, i + 4$) (Fig. 4A), as well as by the generally weak $^3J_{NH\text{-}H\alpha}$ coupling constants measured on these peptide segments (Table 1). This indicates that the main secondary elements are α-helices encompassing residues 4–16 (helix I) and 25–36 (helix II). The presence of $d_{\alpha N}(i, i + 2)$ connectivities at the extremities of helices I and II, with the concomitant disappearance of $d_{\alpha N}(i, i + 4)$ connectivities, indicates that initial and final turns are distorted toward a 3_{10} conformation. The disposition of nu-

Table 1. *Proton chemical shifts and $^3J_{NH\text{-}H\alpha}$ coupling constant values measured for α_2p8 relative to (^{2}H)TSP in water (H_2O/D_2O, 90/10) at pH 6.5 and 298 K*

Residue	NH	$^3J_{HN\text{-}H\alpha}$	C$^\alpha$H	C$^\beta$H	C$^\gamma$H	Others
Asp1			4.42	2.79, 2.74		
Pro2	—		4.47	1.97, 1.60	1.74	C$^\delta$H$_2$ 3.74, 3.66
Cys3	8.60	(8.00)	5.22	3.36, 2.91		
Gln4	7.51	(6.00)	4.01	2.27, 2.15	2.47	
Lys5	8.94	(6.50)	4.06	1.86, 1.72	1.40	C$^\delta$H$_2$ 1.55; C$^\epsilon$H$_2$ 2.99
Gln6	8.85	(7.00)	3.92	2.41, 1.84	2.59	N$^\epsilon$H$_2$ 7.70, 6.97
Ala7	8.34	(4.50)	4.15	1.56		
Ala8	8.29	(5.50)	1.59			
Glu9	8.08	(7.50)	4.08	2.16,1.90	2.47	
Ile10	7.58	(7.00)	3.70	2.06	1.81, 0.91	C$^{\gamma'}$H$_3$ 0.86; C$^\delta$H$_3$ 1.10
Gln11	7.36	(7.00)	3.95	2.26, 2.19	2.50, 2.47	
Lys12	7.80	(6.00)	3.99	1.86, 1.58	1.40	C$^\delta$H$_2$ 1.70; C$^\epsilon$H$_2$ 2.94
Cys13	8.01	(6.00)	4.23	3.42, 3.18		
Leu14	8.96	(7.50)	3.60	1.85, 1.79	0.59	C$^\delta$H$_3$ 0.96, 0.92
Gln15	8.07	(6.00)	3.50	2.15, 2.06	2.58, 2.55	
Ala16	7.50	(7.50)	4.37	1.50		
Asn17	7.46	(11.00)	5.10	2.95, 2.47		N$^\delta$H$_2$ 7.95, 7.40
Ser18	8.33	(7.50)	4.08	4.03, 4.00		
Tyr19	8.88	(7.00)	3.35	3.27, 3.11		C(2,6)H 7.01; C(3,5)H 6.86
Leu20	7.39	(7.50)	4.37	1.82	1.34	C$^\delta$H$_3$ 1.49, 0.92
Glu21	8.75	(6.00)	3.59	1.98, 1.95	2.34, 2.17	
Ser22	8.32		4.19	4.01, 3.98		
Lys23	7.55	(14.50)	4.48	2.00	1.59, 1.43	C$^\delta$H$_2$ 1.92; C$^\epsilon$H$_2$ 3.05
Cys24	7.83	(10.00)	5.33	3.43, 2.39		
Gln25	7.88	(5.50)	3.97	2.15	2.52, 2.46	
Ala26	8.75	(5.50)	4.20	1.51		
Val27	7.71	(8.50)	4.10	2.31	1.21, 1.10	
Ile28	8.15	(5.50)	3.79	1.94	1.86, 1.09	C$^{\gamma'}$H$_3$ 1.01; C$^\delta$H$_3$ 0.84
Gln29	8.44	(6.00)	4.04	2.29, 2.20	2.56, 2.53	
Glu30	7.80	(6.00)	4.14	2.27, 2.17	2.49	
Leu31	7.55	(5.50)	4.18	2.26, 1.40	1.79	C$^\delta$H$_3$ 1.11, 1.02
Lys32	7.70	(6.50)	3.86	1.90, 1.76	1.37	C$^\delta$H$_2$ 1.64; C$^\epsilon$H$_2$ 2.95
Lys33	8.11	(7.50)	4.04	1.90	1.59, 1.414	C$^\delta$H$_2$ 1.70; C$^\epsilon$H$_2$ 2.94
Cys34	7.82	(4.50)	4.36	3.36, 3.32		
Ala35	9.12	(7.00)	4.14	1.49		
Ala36	7.79	(6.00)	4.22	1.51		
Gln37	7.69	(6.00)	4.05	1.97, 1.83	2.16, 2.13	N$^\epsilon$H$_2$ 7.32, 7.27
Tyr38	7.85	(11.00)	4.57	3.30, 2.79		C(2,6)H 7.39; C(3,5)H 6.88



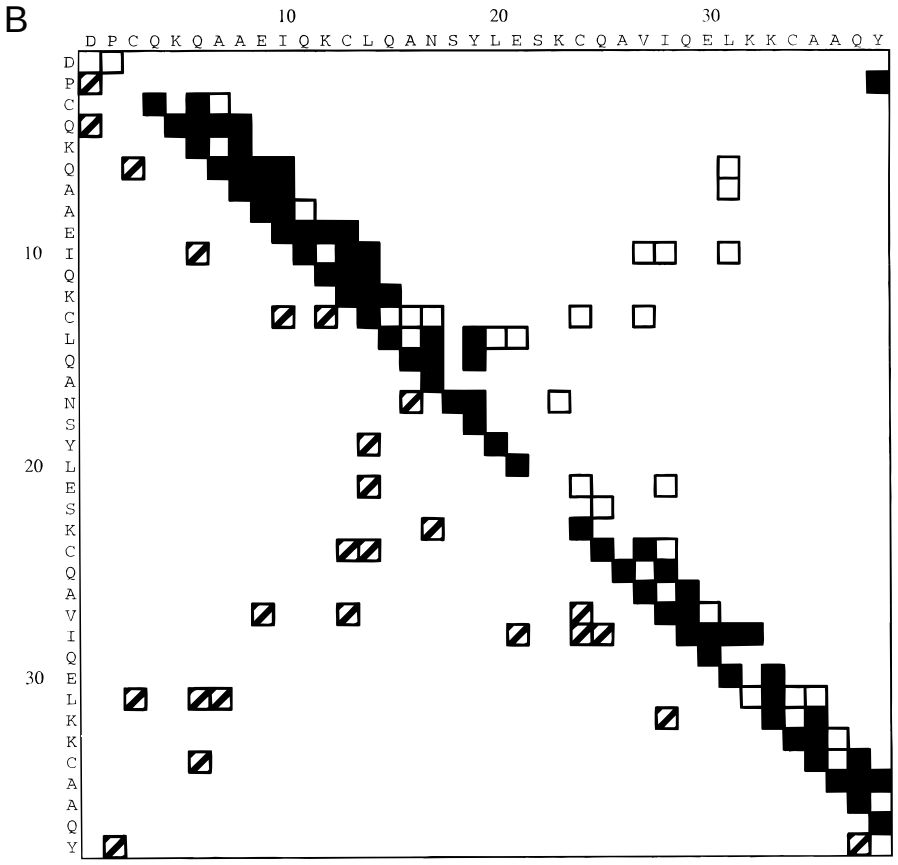


Fig. 4. A: NOESY spectrum (amide proton region) of α_2p8 (600 MHz, 200 ms mixing time) recorded in H_2O at pH 5.5 and 293 K. The sequential d_{NN} connectivities shown for different regions [residues 3 to 16 (above the diagonal) and residues 25–36 (below the diagonal)] indicate the presence of two helices in the secondary structure of α_2p8. **B:** Diagonal plot of the NOE contacts observed for α_2p8. Above the diagonal: filled squares indicate backbone–backbone NOE, open squares backbone–side-chain NOE. Below the diagonal, dashed squares indicate side-chain–side-chain NOE.

merous long-range ($>i, i + 4$) NOE contacts in the diagonal plot perpendicular to helix I and helix II diagonal segments (Fig. 4B) indicates that these two helices are orientated in an antiparallel fashion.

A total of 314 structural (distance and dihedral; Table 2) constraints were collected and taken into account for structure calculations using DIANA (see Material and methods). The 30 structures with the best target value (<0.37 Å^2) were further refined by simulated annealing and restrained energy minimization using the AMBER force field. The average global RMSD calculated for backbone and heavy atoms goes from 0.98 ± 0.20 Å and 1.90 ± 0.22 Å to 0.65 ± 0.25 Å and 1.51 ± 0.21 Å, respectively, after refinement. The survey of the structural statistics and of the residual violations of experimental constraints for these 30 conformers is shown in Table 2. The high number of NOEs per residue (Fig. 5A), in conjunction with the low RMSD values, indicates a good sampling of the conformational space. Analysis of the local backbone displacements (Wagner et al., 1987) (Fig. 5B) for the 30 energy-refined conformers indicates that the two helices of α_2p8, as well as the two interlocking turns joining helix I and helix II, are well defined. In addition, a value very close to 1 is generally observed for the order parameter S (Hyberts et al., 1992) for most φ and ψ angles in the two helices, reflecting their high definition (data not shown), even though no dihedral constraints are available for all φ

Table 2. *Experimental constraints and refinement statistics of the 30 conformers representing the solution structure of α_2p8 before and after restrained energy minimization*

Distance constraints			
Intraresidue	26		
Sequential	100		
Medium range	105		
Long range	54		
Disulfide bonds	2		
Constraints per residue	14		
Dihedral constraints			
Φ	17		
χ_1	12		
Parameter		DIANA	DIANA + AMBER
Target function (Å^2)		0.27 ± 0.06	
Upper limit violations			
Number > 0.2 Å		0	0
Sum of violations (Å)		1.62 ± 0.26	1.23 ± 0.16
Maximum violation (Å)		0.16 ± 0.06	0.21 ± 0.05
Dihedral angle violations			
Number > 5°		0	0
Sum of violations (°)		2.14 ± 1.71	3.47 ± 4.39
Maximum violation (°)		1.17 ± 1.09	2.74 ± 2.84
van der Waals violations			
Number > 0.2 Å		0	
Sum of violations (Å)		1.05 ± 0.25	
Maximum violation (Å)		0.11 ± 0.03	
AMBER energies (kcal mol^{-1})			
Bond energy			11.3 ± 0.6
Valence angle energy			49.7 ± 2.4
van der Waals energy			−238.5 ± 5.6
Electrostatic energy			−1,123.6 ± 15.7
Constraint energy			7.8 ± 1.1
Total nonbonding energy			−858.0 ± 11.7
Total energy			−612.1 ± 9.7
RMSD values (Å)			
Residues			
1–38 (all)	BA[a]/HA[b]	0.98 ± 0.20/1.90 ± 0.22	0.65 ± 0.25/1.51 ± 0.21
4–36 (structural part)	BA[a]/HA[b]	0.84 ± 0.17/1.79 ± 0.20	0.43 ± 0.16/1.40 ± 0.18
4–16,25–36 (helices)	BA[a]/HA[b]	0.71 ± 0.16/1.65 ± 0.18	0.39 ± 0.18/1.44 ± 0.20
4–16 (helix I)	BA[a]/HA[b]	0.58 ± 0.17/1.63 ± 0.24	0.21 ± 0.06/1.44 ± 0.24
25–36 (helix II)	BA[a]/HA[b]	0.52 ± 0.15/1.40 ± 0.20	0.26 ± 0.19/1.23 ± 0.24

[a] Backbone atoms.
[b] All heavy atoms.

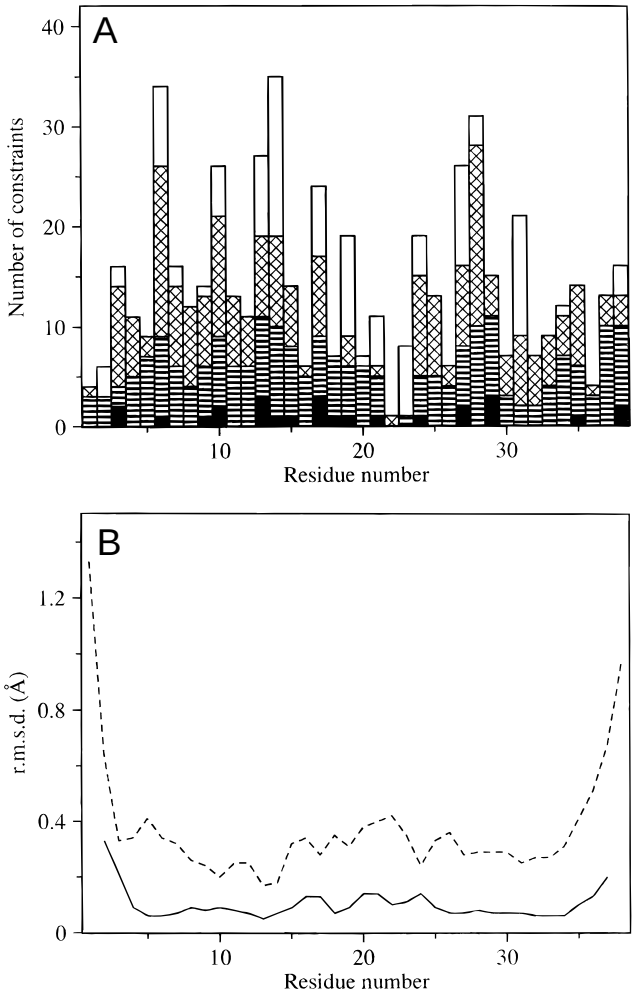


Fig. 5. A: Plot of the number of NOE constraints used in the final structure calculation of α_2p8 as a function of the amino acid sequence. NOE categories are shown as follows: intraresidue (filled bars); sequential (dashed bars); medium-range (cross-hatched bars); long-range (open bars). **B:** Plots, as a function of the amino acid sequence, of the mean of the global RMSD for the backbone atoms superimposed over the whole structure (residues 1–38) (broken line), and the mean local RMSD for the backbone superposition of all tripeptide segments along the sequence (solid line). The RMSD values for the tripeptide segments are plotted at the position of the central residue.

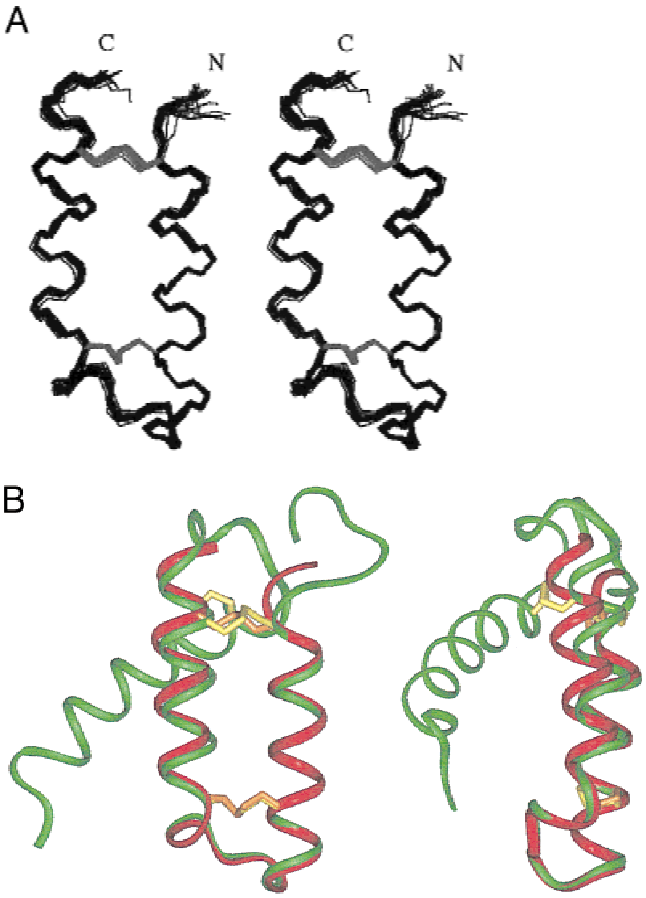


Fig. 6. A: Stereoview of the polypeptide backbone of the 30 conformers representing the structure of α_2p8 in solution. The side-chain heavy atoms of the four cysteine residues are reported (gray line). Conformers were superimposed for minimum pairwise RMSD of the backbone atoms of residues 3 to 34. **B:** Ribbon diagram of the average NMR structure of α_2p8 (red) superimposed to the average NMR structure of p8^{MTCP1} (green), from residue 3 to residue 34. Two views are shown, rotated by a 90° about the vertical axis. The heavy atoms of the disulfide bridges are reported in orange for α_2p8 and in yellow for p8^{MTCP1}. The coordinates of the average structure of α_2p8 have been deposited to the Protein Data Bank (code access: 1EI0; code access for the coordinates of the average structure of p8^{MTCP1}: 1hp8).

angles. Finally, inspection of the Ramachandran plot (ProCheck; Laskowski et al., 1993) built from the energy-minimized (Pearlman et al., 1995) average structure calculated from these 30 structures shows that 29 residues (i.e., 82.9%)—out of 35 "meaningful" residues—fall in the most favored regions, six residues (i.e., 17.1%) are in the additional allowed regions, and no residue falls in the generously allowed or disallowed regions, supporting the good geometric features of the structure. The coordinates of the structure of α_2p8 are being deposited to the Protein Data Bank (Brookhaven National Laboratory).

The peptide α_2p8 folds into an α-helical hairpin consisting of two amphipathic antiparallel helices spanning residues 4 to 16 (helix I) and 25 to 36 (helix II), connected by a short loop (residues 17–24), and stapled in an antiparallel orientation by the two disulfides 3–34 and 13–24 (Fig. 6A). As already observed for p8^{MTCP1}, the structure deviates from a canonical coiled-coil (Seo & Cohen, 1993), showing an interhelical crossing angle of about 10°, which is significantly smaller than the 20° value usually measured for a typical supercoil. As indicated by χ_1 angular order parameter values close to 1 (Fig. 7), all the side chains in the hydrophobic interface between the two helices adopt a single predominant rotameric state: the two disulfide bridges have the common left-handed spiral conformation (Richardson, 1981) in all the calculated structures—as already observed in the 3D structure of p8^{MTCP1}—and the other hydrophobic side chains display typical "knobs-into-holes" packing, where the hydrophobic *a* and *d* chains (knobs) fit into the spaces (holes) between four residues on the neighboring helices. Interestingly, stereospecific assignment of the $C^{\beta}H$ protons was not available for all these residues, suggesting that their well-defined χ_1 is a consequence of the protein packing rather than the direct experimental data. Inspection of the hydrogen bonds reveals the presence of N- and C-caps stabilizing the ex-

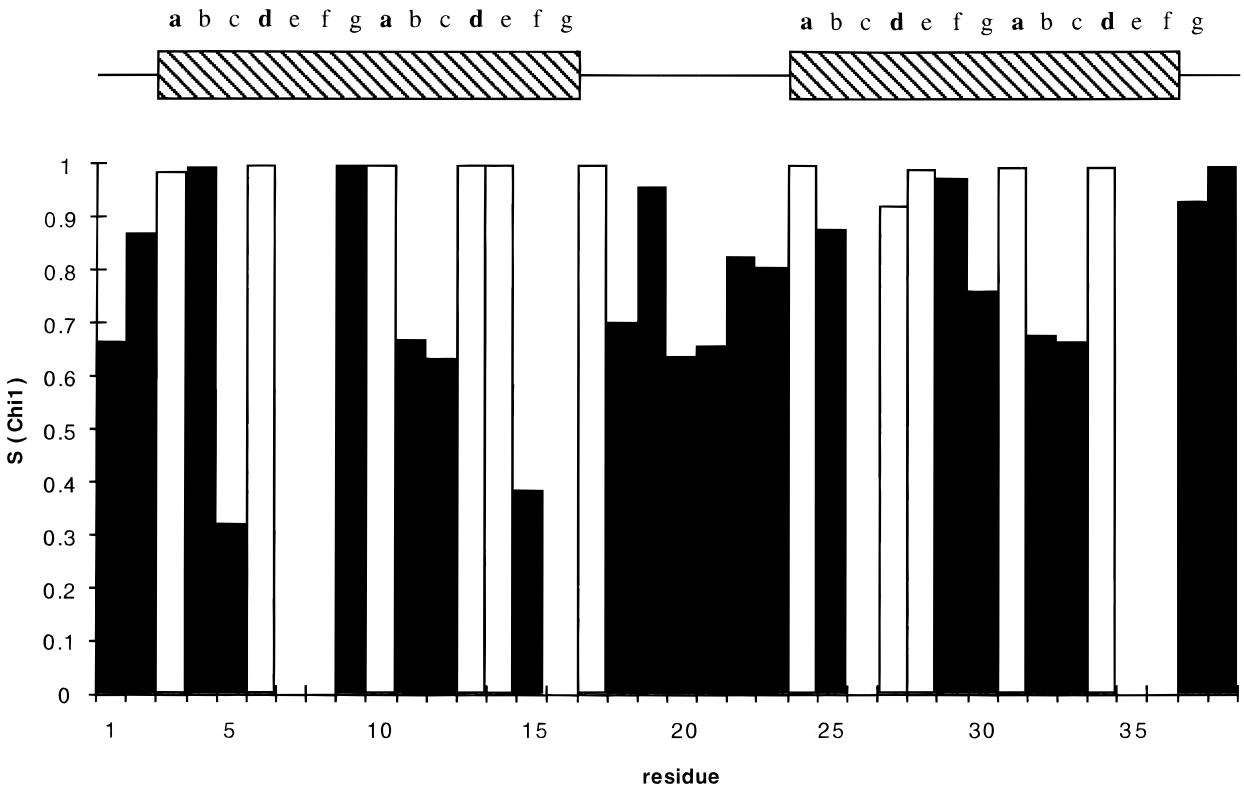


Fig. 7. Plot of the angular order parameters S calculated for the χ_1 angles from the 30 conformers of α_2p8, as a function of the amino acid sequence. Hydrophobic residues at the interface of the two helices are labeled as open bars. The secondary structure and the heptad repeats positioning are reported on top of the figure.

tremities of helix I. At the N-terminal end, the oxygens of the side-chain carboxyl group of Asp1 H-bonds the backbone amide proton of Gln4. In the 3D structure of p8^{MTCP1}, a similar capping of the N-terminus of helix I was observed: one carboxyl oxygen of the corresponding residue (Asp5) was engaged in a hydrogen bond with the backbone amide proton of Gln8, but the other carboxyl oxygen H-bonded the backbone amide proton of Val48, thus capping the N-terminus of helix III. At the C-terminal end of helix I in the peptide, the side-chain amino group of Asn17 H-bonds the backbone carbonyl group of Cys13. In addition, noncanonical H-bonds involve the backbone carbonyl group of Leu14 with the backbone amide groups of Asn17 and Tyr19, further stabilizing the C-terminal end of helix I. Similar interactions were already present in the 3D structure of p8^{MTCP1}.

As shown by structure superimposition in Figure 6B, the α-helical hairpin found in α_2p8 is highly similar to that observed in p8^{MTCP1}. When enforcing the alignment of the two average NMR structures on helices I and II only (residues 4–16 and 25–36 in α_2p8, and residues 8–20 and 29–40 in p8^{MTCP1}), the RMSD measured in these helical segments for the backbone heavy atoms is below 0.65 Å. Nevertheless, the crossing angle between the two helices of the α-hairpin is significantly larger in α_2p8 ($\approx$10° against $\approx$5°), probably due to the lack of the third helix packed against the α-hairpin in the structure of p8^{MTCP1}.

Discussion

The present report demonstrates that it is possible to minimize the structure of p8^{MTCP1} into a 38-residue α-helical structure, without altering the native geometry or compromising the native stability of the α-helical hairpin. Deletion of the third helix of p8^{MTCP1} and minimal sequence modifications to cope with the changes introduced are well accepted, and produce a new antiparallel α-helical motif, presenting a well-defined 3D structure.

Previous studies have revealed that hydrophobic packing within the dimer interface of a helical coiled-coil is the major force contributing to the overall structure stability (Zhou et al., 1992a, 1992b; Zhu et al., 1993). This imposes a minimal length to the helix to gain sufficient stabilization. It has been shown that three heptads or six helix turns correspond to the minimum length for a peptide to adopt the coiled-coil conformation in aqueous media (Su et al., 1994). α_2p8 provides an original alternative to the stabilization of two shorter antiparallel α-helices (three helical turns per helices) through interhelical disulfide bridging. Disulfide linkages between α-helical structures (Pullman & Pullman, 1974) are scarcely observed and may involve unusual stereochemical constraints. α_2p8 together with p8^{MTCP1} represent the first examples of a "natural" motif containing an antiparallel α-helical hairpin stabilized by two disulfide bridges without any distortion of either the helices or the disulfide geometries. However, the steric constraints induced by the two-disulfide bridges impose an angle of about 10° between the helix axes (5° in p8^{MTCP1}), instead of the typical value of 20° observed in a helical supercoil.

In α_2p8, the half-cystine residues Cys3, Cys13, and Cys24, Cys34 are located at position *a*, *d*, and *a′*, *d′* of helix I and helix II, respectively. This fixes a regular spacing of half-cystine residues, -Cys-X$_9$-Cys-X$_m$-Cys-X$_9$-Cys-, which is characteristic of p8^{MTCP1} (Barthe et al., 1997) and leads to the formation of *a*–*d′* and *a′*–*d*

disulfide bridges, which appear in a favorable geometry for structure stabilization, the disulfides being buried in the hydrophobic interface. This geometry may probably explain the capability of the reduced peptide to form the native disulfides and may also be advantageous to confer to the motif an enhanced stability in applications involving exposure to reducing environment. In fact, incubation of α_2p8 with 5 mM reduced glutathione for 2 h was not sufficient to reduce disulfide bridges, and reduction occurred only in the presence of 6 M guanidine hydrochloride (data not shown). Neurotoxin B-IV from the marine worm *Cerebratulus lacteus* (Barnham et al., 1997) that also presents a α-hairpin motif with a correct helical geometry possesses a different regular half-cystine spacing, Cys-X_7-Cys-X_m-Cys-X_7-Cys. However, with this spacing, the half-cystine residues, delimiting the regular helical region of the hairpin (Cys16–Cys23 and Cys41–Cys48) of this neurotoxin, are located at *d*, *d′* positions of the heptad and are involved in *d–d′* disulfide bridges. Therefore, the disulfide bonds lay on the same face of the hairpin and are exposed to the solvent. Intrahelical *i* to *i* + 4 disulfide (Jackson et al., 1991) or lactam bridges (Houston et al., 1996b), have also been proposed to stabilize helical structure and to enhance dimerization between single strand α-helices. NMR studies have demonstrated that residues located within such bridges are strongly constrained into a helical conformation (Houston et al., 1995). These bridges act as constrained C- or N-terminal capping boxes (Harper & Rose, 1993; Zhou et al., 1994) and inhibit fraying of the N- and C-termini, which in turn enhance hydrophobic interactions required for coiled-coil stabilization. Nevertheless, it has been demonstrated that the introduction of a disulfide cross-link between the two helices, *via* an appropriate linker, is often necessary to further stabilize newly designed coiled-coil structures (Houston et al., 1996a). In the case of α_2p8, the stabilization of the coiled-coil conformation is obtained with a thermodynamically favorable canonical geometry both for the disulfides and the helices, as well as through proper packing of the hydrophobic side chains at the helix interface.

As a matter of fact, even if the two-disulfide bridges are likely to play an important role in conferring conformational stability to the antiparallel α-helical motif of α_2p8, as assessed by the reduced helical content of the derivative devoid of disulfides, the contribution of interhelix hydrophobic interactions cannot be neglected. As observed in typical α-helical coiled-coils, helix I and helix II are markedly amphipathic, and regular heptad motifs are observed in helix I and II (Fig. 8). The residues at position *a* and *d*, which typically form the hydrophobic core in α-helical coiled-coils, are also hydrophobic residues or half-cystines in α_2p8. Thus, the sulfur–sulfur covalent bonds of disulfide bridges take the place of non-covalent hydrophobic interactions to stabilize the α-helical assembly. Nevertheless, the presence of Gln6 at a *d* position is an indication of the stabilization brought to the α-hairpin by the two disulfides, because the polar side chain of this residue can fit in the hydrophobic core without any distortion or destabilizing effects. As usually found in helical coiled-coils, the highly solvent exposed positions *b*, *c*, and *f* are also occupied by hydrophilic residues in α_2p8. Charged residues, especially glutamate, arginine and lysine, are generally located at the *e* and *g* positions of the heptad repeat. The apolar portion of these side chains shields the hydrophobic core, whereas their polar moieties are engaged in weakly attractive electrostatic and hydrogen-bonding interhelical *e–e′* or *g–g′* interactions, thereby increasing the stability of the helical motif. This is not the case for α_2p8. At positions *g* and *g′*, which are located in the same solvent-exposed face of the α-hairpin, residues with the same charge, Glu9 and Glu30, are found so that no favorable interaction can occur. On the opposite face of the antiparallel α-helical motif, positions *e* and *e′* are occupied by hydrophobic residues, Ala7, Leu14, and Ile28, Ala35, respectively. The equivalent residues in $p8^{MTCP1}$ are involved in hydrophobic contacts with helix III. Presumably, in α_2p8, due to the presence of the two disulfide bridges, additional interhelical *e–e′* or *g–g′* contacts are not needed to stabilize the α-helical assembly.

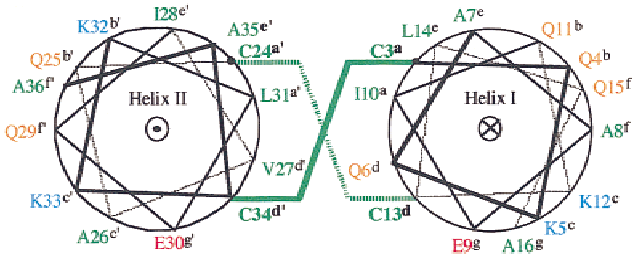


Fig. 8. "Wheel diagram" of the two α_2p8 helices. For the sake of clarity, the two helices are represented as regular α-helices, whereas the pitch of helix I and II has an intermediate value between α-helix and coiled-coil. Hydrophobic residues and disulfide bridges are labeled in green, polar residues in orange, positively charged residues in blue, and negatively charged residues in red. Letters at exponent index indicate the heptad repeats.

In conclusion, in α_2p8 the major contribution to structure formation and stabilization appears to come from residues at positions *a*, *a′* and *d*, *d′*; the backbone geometry appears rather insensitive to the nature of the (solvent-exposed) side chains at all other positions *b*, *c*, *e*, *f*, *g* and *b′*, *c′*, *e′*, *f′*, *g′* of the heptads: this leaves a large molecular surface of this motif available for mutations. Indeed, the presence of hydrophobic residues at the solvent-exposed position *e* in the heptads of α_2p8 demonstrates that this highly constrained scaffold has the potential to effectively present hydrophobic residues to solvent. Thus, α_2p8 appears as a stable and permissive α-helical scaffold, particularly promising for the display of new functional sites and peptide libraries, including hydrophobic residues that, usually, when exposed to solvent, dramatically destabilize the structure.

It has been predicted that the formation of a coiled-coil is primarily dependent on the presence of heptad repeat sequences, and that discontinuities in heptad periodicity induce changes in the direction of the polypeptide chain in coiled-coils and four-α-helix bundles, including formation of α-hairpin bends (Cohen & Parry, 1990; Banner et al., 1987). In a previous work, it has been speculated that the regular heptad repeats of $p8^{MTCP1}$ probably participate in the earlier events of protein folding, contributing to the correct positioning of helices I and II prior to disulfide bridge formation (Barthe et al., 1997). Supporting this assumption, the disulfide bridges form spontaneously during protein purification and only one protein species is obtained from the *Escherichia coli* culture. Interestingly, the correct disulfide bonds of α_2p8 also form spontaneously from the reduced peptide in a redox buffer, indicating that excising the α-hairpin from its proteic context does not alter its folding capabilities. In addition, to facilitate peptide synthesis, this suggests the possibility to obtain α_2p8 derivatives using low-cost expression systems and phage library (Dunn, 1996).

Conclusions

The selection of α-helical domains or motifs from natural protein has experienced exciting successes with the redesign of scaffolds

able to incorporate novel functions (Nord et al., 1995; Braisted & Wells, 1996; Starovasnik et al., 1997; Domingues et al., 1999). In the present paper, we have demonstrated that the α-hairpin present in $p8^{MTCP1}$ can be excised from its natural context without altering its geometry. It is obtainable in high yields by chemical synthesis and may thus incorporate nonnatural substitutions. The geometry of the peptide is structurally homologous to other motifs found in natural helical coiled-coils. This scaffold is stabilized by two interhelical disulfide bridges and should be more permissive for mutations than earlier proposed oligomeric multistranded coiled-coils. The large solvent-exposed surface of the helix coiled-coil is ideally suited to create a protein recognition surface either by the transfer of suitable functional sites (Vita, 1997) or by sequence randomization and functional selection (Nygren & Uhlén, 1997). Because local interactions in the interhelical loops do not exert a dominant influence on the structure of helical proteins (Brunet et al., 1993), in addition to the coiled-coil region, the interhelical loop could be also utilized, thus enlarging the capabilities of this scaffold in the design of conformationally homogeneous peptide libraries. Hence, this new scaffold should provide an excellent framework for the engineering of new functions within a well-defined α-helical miniprotein system.

Materials and methods

Materials

Fluorenylmethyloxycarbonyl (Fmoc)-amino acid derivatives and 4-(2′,4′-dimethoxyphenyl)hydroxymethylphenoxy polystyrene resin (Rink amide resin, 0.47 mmol/g) were obtained from Novabiochem (Laufelfingen, Switzerland). Solvents and chemicals used in synthesis and purification were from SDS (Peypin, France). The 1-hydroxybenzotriazole (HOBt), (2-(1-*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium) hexafluorophosphate (HBTU) were from Applied Biosystems (Paris, France). Chemicals used in the deprotection were from Aldrich (Milwaukee, Wisconsin). 2,2,2-Trifluoroethanol was from Fluka (Buchs, Switzerland). Analytical reversed-phase HPLC was carried out on a Spectra-Physics System (TSP, Les Ulis, France), consisting of a P2000 pump and a UV3000 detector, driven by a PC operating with a PC1000 software, using a Vydac C18 column (0.46 × 15 cm i.d.; 5-μm silica particles; 300 Å pore size); preparative reversed-phase HPLC was carried out on a Jasco system (JASCO France, Nantes, France), consisting of two PU986 pumps, a UV975 detector, and a Merck-Hitachi D7500 recorder, using a Vydac C18 column (1 × 25 cm; 5 μm silica particles, 300 Å pore size).

Peptide synthesis

Synthesis was carried out by the stepwise solid-phase approach on an Applied Biosystems Mod. 433 Peptide Synthesizer. The polypeptide chain was assembled by HOBt/HBTU-mediated single coupling of 1 mmol of the Fmoc-protected amino acids to 0.1 mmol Rink-amide resin (0.43 mmol/g charge), by using FastMoc Applied Biosystems protocol (Fields et al., 1991) and the manufacturer-suggested amino acid protections. Trityl and acetamidomethyl groups were used as orthogonal protection of cysteine residues. Ultraviolet (UV) monitoring of Fmoc deprotection was used to optimize the deprotection and coupling time. Final side-chain deprotection and cleavage of the polypeptide chain from the resin was carried out by treatment with Reagent K (King et al., 1990) (82.5% trifluoroacetic acid, 5% water, 5% phenol, 5% thioanisole, and 2.5% ethanedithiol) for 2 h at room temperature. By this procedure all protecting groups were removed, except the acetamidomethyl group. Deprotected peptide was precipitated by t-butylmethylether, washed three times with ether, dissolved in 30% acetic acid, and lyophilized.

Disulfide bond formation and peptide purification

The first disulfide bond (Cys13–Cys24) was formed by air oxidation, by dissolving the reduced crude peptide at 0.2 mg/mL in 20 mM ammonium bicarbonate, pH 8.5, and stirring the solution for 48 h. The product was then lyophilized and purified on reversed-phase HPLC. The second disulfide bond, Cys3–Cys34, was formed by iodine oxidation of the one-disulfide purified product: 50 equivalents of iodine were added to the peptide dissolved (2.0 mg/mL) in 50% acetic acid solution, containing 1 N HCl (0.1 mL/mg peptide), and reaction allowed to occur at room temperature for 1 h. Excess iodine was then extracted four times with carbon tetrachloride, and the aqueous solution lyophilized. The peptide was lyophilized and finally purified by reversed-phase HPLC.

Disulfide bonds were also formed directly from reduced purified peptide by incubating the peptide (0.1 mg/mL) in degassed 20 mM Tris·HCl buffer, containing 0.1 M NaCl, 5 mM oxidized, and 0.5 mM reduced glutathione, pH 7.8, for 1 h.

Preparation of reduced and carboxamidomethylated peptide (CM-α2p8)

Oxidized and purified peptide was dissolved (0.6 mg/mL) in degassed Tris·HCl buffer, pH 8.0, containing 6 M guanidine hydrochloride, and treated with 40 equivalents of tris-[2-carboxyethyl]phosphine (TCEP) for 1 h at room temperature. A part of this material was saved for folding experiments, the remainder was then treated with iodoacetamide (80 equivalents) for 1 h in the dark; then, it was acidified with 50% acetic acid and the peptide purified by reversed-phase HPLC.

CD spectroscopy

CD spectra were recorded on a Jobin Yvon CD6 dichrograph, equipped with a thermostatically controlled cell holder and an IBM PC operating with a CD6 data acquisition and manipulation program. Spectra were recorded at 20 °C in 5 mM phosphate buffer, pH 7.5, or in a mixture of buffer-TFE, by accumulating four scans obtained with an integration time of 0.5 s every 0.2 nm. In the far-UV region (180–250 nm), the peptide sample was at 2.0×10^{-5} M in a 0.1 cm pathlength quartz cell. Spectra are presented as mean residue ellipticity ($[\theta]_{MRW}$ in deg cm^2 dmol^{-1}).

NMR measurements

All NMR experiments, i.e., DQF-COSY (Rance et al., 1983), z-TOCSY (Braunschweiler & Ernst, 1983; Davis & Bax, 1985; Rance, 1987), NOESY (Jeener et al., 1979; Kumar et al., 1980) were carried out at 600 MHz, on a Bruker AMX600 spectrometer equipped with a z-gradient ^{1}H-^{13}C-^{15}N triple resonance probe. All samples were 2 mM in peptide, and the pH (pD) was adjusted to 5.5. The pD values are uncorrected for isotopic effects. Samples dissolved in H_2O contained 10% (v/v) ^2H_2O for the lock. Data

were acquired on samples generally maintained at 20 °C, and chemical shifts are reported relative to (^{2}H)TSP. Deuterated reagents ^2H_2O, ^{2}HCl, NaO^{2}H, (^{2}H)DTT, (^{2}H)TSP were from EURISOTOP.

For NOESY experiments, two mixing times of 80 and 200 ms were used. *z*-TOCSY experiments were carried out using the TOWNY isotropic transfer sequence (Kadkhodaei et al., 1993) and a mixing time of 60 ms. In all 2D experiments, quadrature detection in the indirectly observed dimension was obtained with States-TPPI (Marion et al., 1989b). Solvent suppression in *z*-TOCSY and NOESY experiments was carried out using the WATERGATE method (Piotto et al., 1992) in association with water-flip-back pulses (Lippens et al., 1995; Dalhuin et al., 1996). For DQF-COSY, water suppression was obtained with a low-power irradiation of the solvent signal during the relaxation delay. In addition, a solvent-suppression filter was applied to the time-domain data (Marion et al., 1989a). The spectral width used in both dimension was 7,800 Hz; 4 K data points in t_2 and 512 experiments in t_1 were usually acquired, except for DQF-COSY where 1,024 experiments were acquired in t_1. Typically, 128 scans (NOESY, DQF-COSY) or 96 scans (*z*-TOCSY) were acquired per increment. The data sets were processed using GIFA (Pons et al., 1996). Prior to Fourier transformation, the raw data were multiplied by a squared cosine window function in t_2 and a shifted (p/4) sine bell window function in t_1 and were zero-filled to the next power of 2 leading to a resolution of 1.9 Hz in the f_2 dimension and 7.62 Hz (NOESY, z-TOCSY) or 3.8 Hz (DQF-COSY) in the f_1 dimension. Residual baseline distorsions in f_2 were removed with a fifth-order polynomial baseline correction, and a linear baseline correction was applied in the f_1 dimension.

NH-C$^\alpha$H and C$^\alpha$H-C$^\beta$H coupling constants were measured in the DQF-COSY spectra (2 Hz of digital resolution). To compensate for the overestimation of the $^3J_{NH\text{-}H\alpha}$ due to the broad linewidth of the antiphase multiplets, we used the method previously described by Ludvigsen et al. (1991). Spin systems determination and sequential assignment of the ^{1}H resonances were computer assisted using the in-house CINDY software, operating on a Silicon Graphics O_2 workstation.

Diffusion experiments were recorded using a modified LED sequence using bipolar gradients (Wider et al., 1994) to prevent any parasitic effect arising from eddy currents, thus allowing short recovery delays. The classical triple resonance probe and Bruker gradient linear amplifier were used for these measurements. Special care was taken for solvent signal suppression: the WATERGATE sequence was used in association with low-power irradiation of the water signal during both the relaxation delay (1 s) and the constant diffusion delay (200 ms). The temperature was maintained at 20 °C, closed to the room temperature, to minimize spurious effects from convection. To determine the self-diffusion constant, 36 1D experiments with an identical diffusion delay were performed with pulse field gradients of 1.8 ms duration and variable gradient strength ranging from 1 to 47.5 G/cm (2 to 95% of the maximum output power of the amplifier). Each individual experiment has a time domain of 2 K complex points and was acquired with 32 transients (minimum phase cycling). The different values of the gradient strength were determined from an exponential sampling function. The diffusion coefficient was obtained from the intensity decays of selected nonexchangeable peaks when increasing the gradient intensity by Maximum Entropy processing with an inverse Laplace transform computed on 128 points. The corresponding molecular weight was deduced from a standard curve previously established in the same experimental conditions from a series of proteins of known molecular weights ranging from 4 to 50 kD: charybdotoxin (4.1 kD), BPTI (6.5 kD), γ-cardiotoxin (6.8 kD), and α-neurotoxin (6.8 kD) from *Naja nigricollis*, ubiquitin (8.6 kD), thioredoxin (11.6 kD), cytochrome *c* (12.4 kD), ribonuclease A (13.7 kD), lysozyme (14.6 kD), myoglobin (16.9 kD), β-lactoglobulin (17.7 kD), chymotrypsynogen C (25 kD), ovalbumin (43 kD), GST (2 × 26 kD), all sample concentrations being 3 mg/mL (to be published elsewhere).

Molecular modeling calculations

All calculations were carried out on an Helwett-Packard HP735 workstation, following our standard protocol previously described for p8^{MTCP1} (Barthe et al., 1997). Briefly, 285 distance restraints (26 intraresidues, 100 sequential, 105 medium-range ($i - j < 5$), and 54 long-range upper bound restraints, 255 lower bound restraints) were obtained from the volume of cross peaks measured on 2D NOESY spectra and 17 angular restraints on the ϕ dihedral angles were obtained from the coupling constants $^3J_{NH\text{-}H\alpha}$ using the empirically calibrated Karplus relation (Karplus, 1963; Pardi et al., 1984). In addition, 12 χ_1 angles were obtained from the analysis of the $^3J_{H\alpha\text{-}H\beta}$ coupling constants and intraresidue NOEs (Hyberts et al., 1987), and the usual distance constraints were added to enforce the two disulfide bridges, i.e., ranges of 2.0–2.1 Å for $d(S^\gamma, S^\gamma)$ and 3.0–3.1 Å for $d(C^\beta, S^\gamma)$ and 3.5–4.5 Å for $d(C^\beta, C^\beta)$. From these restraints, a set of 50 structures was generated with the variable target function program (Güntert & Wüthrich, 1991) DIANA. Each run started from 999 randomized conformers. When no stereospecific assignment was possible, pseudo-atoms were defined and corrections added as described by Wüthrich et al. (1983). Among the 50 preliminary structures generated by DIANA, 30 presented a value of the target function smaller than 0.37 Å^2, and no distance violation larger than 0.35 Å. The φ and χ_1 angles showed no violation greater than 4.1°. The low value of the target function obtained from a high number of constraints indicates the consistency of the data set.

The 30 best structures (based on the final target penalty function values) were further submitted to molecular mechanics energy refinement with the SANDER module of AMBER 4.1 (Pearlman et al., 1995) using the 1994 force field (Cornell et al., 1995) as previously described for p8^{MTCP1}. Briefly, 5,000 cycles of restrained energy minimization were first carried out followed by a 50 ps long simulated annealing procedure in which the temperature was raised to 900 K for 30 ps then gradually lowered to 300 K. During the molecular dynamics runs, the covalent bond lengths were kept constant by applying the SHAKE algorithm (van Gunsteren & Berendsen, 1977), allowing a 2 fs time step to be used. During this stage, the force constants for the NMR distance constraints and for the angular constraints were gradually increased from 3.2 to 32 kcal mol^{-1} Å^{-2} and 0.5 to 50 kcal mol^{-1} rad^{-2}, respectively. A final restrained minimization led to the refined structures discussed below.

The structures were displayed and analyzed on a Silicon Graphics O_2 station using the INSIGHT program (version 98.0, Biosym Technologies, San Diego).

Acknowledgments

This work was supported by grants from the Ligue Nationale contre le Cancer and from the Association pour la Recherche sur le Cancer.

References

Banner DW, Kokkinidis M, Tsernoglou D. 1987. Structure of the ColE1 Rop protein at 1.7 Å resolution. *J Mol Biol 196*:657–675.

Barnham KJ, Dyke TR, Kem WR, Norton SN. 1997. Structure of neurotoxin B-IV from the marine worm *Cerebratulus lacteus*: A helical hairpin cross-linked by disulphide bonding. *J Mol Biol 268*:886–902.

Barthe P, Chiche L, Declerck N, Delsuc MA, Guignard L, Lefèvre JF, Malliavin T, Mispelter J, Stern MH, Lhoste JM, Roumestand C. 1999. Refined solution structure and backbone dynamics of ^{15}N-labeled C12A-p8^{MTCP1} studied by NMR relaxation. *J Biomol NMR 15*:271–288.

Barthe P, Yang YS, Chiche L, Hoh F, Strub MP, Guignard L, Soulier J, Stern MH, van Tilbeurgh H, Lhoste JM, Roumestand C. 1997. Solution structure of human p8^{MTCP1}, a cysteine-rich protein encoded by the MTCP1 oncogene, reveals a new α-helical assembly motif. *J Mol Biol 274*:801–815.

Bianchi E, Venturini S, Pessi A, Tramontano A, Solazzo M. 1994. High level expression and rational mutagenesis of a designed protein, the minibody. From an insoluble to a soluble molecule. *J Mol Biol 236*:649–659.

Braisted AC, Wells JA. 1996. Minimizing a binding domain from protein A. *Proc Natl Acad Sci USA 93*:5688–5692.

Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J Magn Reson 53*:521–528.

Brunet AP, Huang ES, Huffine ME, Loeb JE, Weltman RJ, Hecht MH. 1993. The role of turns in the structure of an α-helical protein. *Nature 364*:355–358.

Bryson JW, Betz SF, Lu HS, Suich DJ, Zhou HX, O'Neil KT, DeGrado WF. 1995. Protein design: A hierarchic approach. *Science 270*:935–941.

Chen YH, Yang JT, Chau KH. 1974. Determination of the helix and beta-form of proteins in aqueous solution by circular dichroism. *Biochemistry 13*: 3350–3359.

Cohen C, Parry DAD. 1990. α-Helical coiled-coils and bundles: How to design an α-helical protein. *Proteins 7*:1–15.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. 1995. A second generation force field for the simulation of proteins and nucleic acids. *J Am Chem Soc 117*:5179–5197.

Crick FHC. 1953. The packing of α-helices: Simple coiled-coils. *Acta Crystallogr 6*:689–697.

Cunningham BC, Wells JA. 1997. Minimized proteins. *Curr Opin Struct Biol 7*:457–462.

Dahiyat BI, Mayo SL. 1997. De novo protein design: Fully automated sequence selection. *Science 278*:82–87.

Dalhuin C, Wieruszeski JM, Lippens G. 1996. An improved homonuclear TOCSY experiment with minimal water saturation. *J Magn Reson B 111*:168–170.

Davis DG, Bax A. 1985. Assignment of complex ^{1}H NMR spectra with two-dimensional homonuclear Hartmann-Hahn spectroscopy. *J Am Chem Soc 107*:2820–2821.

De Alba E, Santoro J, Rico M, Jimenez MA. 1999. De novo design of a monomeric three-stranded antiparallel β-sheet. *Protein Sci 8*:854–865.

DeGrado WF, Wasserman ZR, Lear JD. 1989. Protein design, a minimalistic approach. *Science 243*:622–628.

de Wolf E, Gill R, Geddes S, Pitts J, Wollmer A, Grötzinger J. 1996. Solution structure of a mini IGF-1. *Protein Sci 5*:2193–2202.

Domingues H, Cregut D, Sebald W, Oschkinat H, Serrano L. 1999. Rational design of a GCN4-derived mimetic of interneukin-4. *Nat Struct Biol 6*:652–656.

Dunn IS. 1996. Phage display on proteins. *Curr Opin Technol 7*:547–553.

Fezoui Y, Weaver DL, Osterhout JJ. 1994. De novo design and structural characterization of an α-helical hairpin peptide: A model system for the study of protein folding intermediates. *Proc Natl Acad Sci USA 91*:3675–3679.

Fields CG, Lloyd DH, Macdonald RL, Otteson KM, Noble RL. 1991. HBTU activation for automated Fmoc solid-phase peptide synthesis. *Peptide Res 4*:95–101.

Fisch P, Forster A, Sherrington PD, Dyer MJ, Rabbit TH. 1993. The chromosomal translocation t(X;14)(q28;q11) in T-cell pro-lymphocytic leukaemia breaks within one gene and activates another. *Oncogene 8*:3271–3276.

Gentz R, Rauscher FJ III, Abate C, Curran T. 1989. Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains. *Science 243*: 1695–1699.

Gibney BR, Mulholland SE, Rabanal F, Dutton PL. 1996. Ferredoxin and ferredoxin-heme maquettes. *Proc Natl Acad Sci USA 93*:15041–15046.

Gibney BR, Rabanal F, Dutton PL. 1997. Synthesis of novel proteins. *Curr Opin Chem Biol 1*:537–542.

Güntert P, Wüthrich K. 1991. Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J Biomol NMR 1*:447–456.

Handel TM, Williams SA, DeGrado WF. 1993. Metal binding modulation of dynamics of a designed protein. *Science 261*:879–885.

Harbury PB, Plecs JJ, Tidor B, Alber T, Kim PS. 1998. High-resolution protein design with backbone freedom. *Science 282*:1462–1467.

Harper ET, Rose GD. 1993. Helix stop signals in proteins and peptides: The capping box. *Biochemistry 32*:7605–7609.

Hecht MH, Richardson JS, Richardson DC, Ogden RC. 1990. De novo design, expression, and characterization of Felix: A four-helix bundle protein of native-like sequence. *Science 249*:884–891.

Houston ME Jr, Campbell AP, Lix B, Kay CM, Sykes BD, Hodges RS. 1996a. Lactam bridge stabilization of α-helices; the role of hydrophobicity in controlling dimeric versus monomeric α-helices. *Biochemistry 35*:10041–10050.

Houston ME Jr, Gannon CL, Kay CM, Hodges RS. 1995. Lactam bridge stabilization of α-helical peptides: Ring size, orientation and positional effects. *J Peptide Sci 1*:274–282.

Houston ME Jr, Wallace A, Bianchi E, Pessi A, Hodges RS. 1996b. Use of a conformationally restricted secondary element to display peptide libraries: A two-stranded α-helical coiled-coil stabilized by lactam bridges. *J Mol Biol 262*:270–282.

Hyberts SG, Goldberg MS, Havel TF, Wagner G. 1992. The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci 1*:736–751.

Hyberts SG, Märki W, Wagner G. 1987. Stereospecific assignments of side-chain protons and characterization of torsion angles in Eglin c. *Eur J Biochem 164*:625–635.

Ilyina E, Roontga V, Mayo KH. 1997. NMR structure of a *de novo* designed peptide 33mer with two distinct compact beta-sheet folds. *Biochemistry 36*:5245–5250.

Jackson DY, King DS, Chmielewski J, Singh S, Schultz PG. 1991. General approach to the synthesis of short α-helical peptides. *J Am Chem Soc 113*:9391–9392.

Jeener J, Meier BH, Bachman P, Ernst RR. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Chem Phys 71*:4546–4553.

Johnsson K, Allemann RK, Widmer H, Benner SA. 1993. Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides. *Nature 365*:530–532.

Kadkhodaei M, Hwang TL, Tang J, Shaka AJ. 1993. A simple windowless mixing sequence to suppress cross relaxation in TOCSY experiments. *J Magn Reson A 105*:393–399.

Kamtekar S, Schiffer JM, Xiong H, Babik JM, Hecht MH. 1993. Protein design by binary patterning of polar and nonpolar amino acids. *Science 262*:1680–1685.

Karplus M. 1963. Vicinal proton coupling in nuclear magnetic resonance. *J Am Chem Soc 85*:2870–2871.

King DS, Fields CG, Fields GB. 1990. A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. *Int J Peptide Protein Res 36*:255–266.

Kortemme T, Ramirez-Alarado M, Serrano L. 1998. Design of a 20-amino acid, three-stranded beta-sheet protein. *Science 281*:253–256.

Kumar A, Ernst RR, Wüthrich K. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Commun 95*:1–6.

Kuroda Y, Nakai T, Ohkubo T. 1994. Solution structure of a de novo helical protein by 2D-NMR spectroscopy. *J Mol Biol 236*:862–868.

Landschultz WH, Johnson PF, McKnight SL. 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. *Science 243*:1681–1688.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Lippens G, Dalhuin C, Wieruszeski JM. 1995. Use of a water flip-back pulse in the homonuclear NOESY experiment. *J Biomol NMR 5*:327–331.

Ludvigsen S, Andersen KV, Poulsen FM. 1991. Accurate measurements of coupling constants from two dimensional nuclear magnetic resonance spectra of proteins and determination of φ angles. *J Mol Biol 217*:731–736.

Marion D, Ikura M, Bax A. 1989a. Improved solvent suppression in one- and two-dimensional NMR spectra by convolution of time-domain data. *J Magn Reson 84*:425–430.

Marion D, Ikura M, Tschudin R, Bax A. 1989b. Rapid recording of 2D NMR data without phase cycling. Application to the study of hydrogen exchange in proteins. *J Magn Reson 85*:393–399.

McLachlan AD, Stewart M. 1975. Tropomyosin coiled-coil interactions: Evidence for an unstaggered structure. *J Mol Biol 98*:293–304.

Myszka DF, Chaiken IM. 1994. Design and characterization of an intramolecular antiparallel coiled coil peptide. *Biochemistry 33*:2363–2372.

Nelson, JW, Kallenbach NR. 1986. Stabilization of the ribonuclease S-peptide alpha-helix by trifluoroethanol. *Proteins 1*:211–217.

Nord K, Nilsson J, Nilsson B, Uhlén M, Nygren PÅ. 1995. A combinatorial library of an α-helical bacterial receptor domain. *Protein Eng 8*:601–608.

Nygren PÅ, Uhlén M. 1997. Scaffolds for engineering novel binding sites in proteins. *Curr Opin Struct Biol 7*:463–469.

O'Shea EK, Klemm JD, Kim PS, Alber T. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science 254*:539–544.

Pardi A, Billeter M, Wüthrich K. 1984. Calibration of the angular dependence of the amide proton-Cα proton coupling constants, $^3J_{HN\alpha}$, in a globular protein. *J Mol Biol 180*:741–751.

Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE III, Ferguson DM, Seibel GL, Chandra Singh U, Weiner PK, Kollman PA. 1995. AMBER 4.1. San Francisco: University of California.

Piotto M, Saudek V, Sklenar V. 1992. Gradient-tailored excitation for single quantum spectroscopy of aqueous solutions. *J Biomol NMR 2*:661–665.

Pons JL, Malliavin TE, Delsuc MA. 1996. Gifa V.4: A complete package for NMR data set processing. *J Biomol NMR 8*:445–452.

Predki PF, Agrawal F, Brünger AT, Regan L. 1996. Amino acid substitutions in a surface turn modulatye protein stability. *Nat Struct Biol 3*:54–58.

Pullman B, Pullman A. 1974. Molecular orbital calculations of the conformation of amino acid residues of proteins. *Adv Protein Chem 28*:347–526.

Quinn TP, Tweedy NB, Williams RW, Richardson JS, Richardson DC. 1994. Betadoublet: De novo design, synthesis and characterization of a beta-sandwich protein. *Proc Natl Acad Sci USA 91*:8747–8751.

Raleigh DP, Betz SF, DeGrado WF. 1995. A de novo designed protein mimics the native state of natural proteins. *J Am Chem Soc 117*:7558–7559.

Rance M. 1987. Improved techniques for homonuclear rotating-frame and isotropic mixing experiments. *J Magn Reson 74*:557–564.

Rance M, Sorensen OW, Bodenhausen G, Ernst RR, Wüthrich K. 1983. Improved spectral resolution in COSY ^{1}H NMR spectra of proteins via double quantum filtering. *Biochem Biophys Res Commun 117*:479–495.

Richardson JS. 1981. The anatomy and taxonomy of protein structure. *Adv Protein Chem 34*:167–330.

Robertson DE, Farid RS, Moser CC, Urbauer JL, Mulholland SE, Pidikiti R, Lear JD, Wand AJ, DeGrado WF, Dutton PL. 1994. Design and synthesis of multi-haem proteins. *Nature 368*:425–432.

Saxena VP, Wetlaufer DB. 1970. Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme. *Biochemistry 9*:5015–5023.

Schafmeister CE, LaPorte SL, Miercke JW, Stroud RM. 1997. A designed four helix bundle protein with native-like structure. *Nat Struct Biol 4*:1039–1046.

Schafmeister CE, Stroud RM. 1998. Helical protein design. *Curr Opin Biotechnol 4*:350–353.

Seo J, Cohen C. 1993. Pitch diversity in alpha-helical coiled coils. *Proteins Struct Funct Genet 15*:223–234.

Soulier J, Madani A, Cacheux V, Rosenzwajg M, Sigaux F, Stern MH. 1994. The MTCP-1/c6.1B gene encodes for a cytoplasmic 8 kD protein overexpressed in T cell leukemia bearing a t(X;14) translocation. *Oncogene 9*:3565–3570.

Starovasnik MA, Braisted AC, Wells JA. 1997. Structural mimicry of a native protein by a minimal binding domain. *Proc Natl Acad Sci USA 94*:10080–10085.

Stern MH, Soulier J, Rosenzwajg M, Nakahara K, Canki-Klain N, Aurias A, Sigaux F, Kirsch IR. 1993. A novel gene of the human chromosome Xq28 translocated to the T cell receptor .alpha/.delta. locus in mature T Cell proliferations. *Oncogene 8*:2475–2483.

Struthers MD, Cheng RP, Imperiali B. 1996. Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science 271*:342–345.

Su JY, Hodges RS, Kay CM. 1994. Effect of chain length on the formation and stability of synthetic α-helical coiled-coils. *Biochemistry 33*:15501–15510.

van Gunsteren WF, Berendsen HJC. 1977. Algorithms for macromolecular dynamics and constraints dynamics. *Mol Phys 34*:1311–1327.

Vita C. 1997. Engineering novel proteins by transfer of active sites to natural scaffolds. *Curr Opin Struct Biol 5*:443–449.

Wagner G, Braun W, Havel TF, Schaumann T, Go N, Wüthrich K. 1987. Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J Mol Biol 196*:611–639.

Walsh STR, Cheng H, Bryson JW, Roder H, DeGrado WF. 1999. Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proc Natl Acad Sci USA 96*:5486–5491.

Wider G, Dötsch V, Wüthrich K. 1994. Self-compensating pulsed magnetic-field gradients for short recovery times. *J Magn Reson A108*:255–258.

Wishart DS, Sykes BD, Richards FM. 1992. The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry 31*:1647–1651.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.

Wüthrich K, Billeter M, Braun W. 1983. Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton–proton distance constraints with nuclear magnetic resonance. *J Mol Biol 169*:949–961.

Zhou HX, Lyu P, Wemmer DE, Kallenbach NR. 1994. Alpha helix capping in synthetic model peptides by reciprocal side chain–main chain interactions: Evidence for an N terminal "capping box." *Proteins Struct Funct Genet 18*:1–7.

Zhou NE, Kay CM, Hodges RS. 1992a. Synthetic model proteins. Positional effects of leucine residues at the non-equivalent positions of the 3-4 hydrophobic repeat to the stability of the two stranded α-helical coiled-coil. *J Biol Chem 267*:2664–2670.

Zhou NE, Kay CM, Hodges RS. 1992b. Synthetic model proteins: The relative contribution of leucine residues at the non-equivalent positions of the 3-4 hydrophobic repeat to the stability of the two stranded α-helical coiled-coil. *Biochemistry 31*:5739–5746.

Zhu BY, Zhou NE, Kay CM, Hodges RS. 1993. Packing and hydrophobicity effects on protein folding and stability: Effects of β-branched amino acids, valine and isoleucine on the formation and stability of twostranded α-helical coiled-coils/leucine zippers. *Protein Sci 2*:383–394.