Structure of synthetic peptide analogues of an eggshell protein of *Schistosoma mansoni*

C. RUSSELL MIDDAGH,¹ JOHN A. THOMSON,^{4,5} CARL J. BURKE,¹ HENRYK MACH,¹
ADEL M. NAYLOR,² MICHAEL J. BOGUSKY,³ JAMES A. RYAN,¹
STEVEN M. PITZENBERGER,³ HANLEE JI,⁴ AND JOHN S. CORDINGLEY⁴

Departments of ¹ Pharmaceutical Research, ² Molecular Systems, and ³ Medicinal Chemistry,
Merck Research Laboratories, West Point, Pennsylvania 19486

⁴ Department of Molecular Biology, University of Wyoming, 3944 University Station, Laramie, Wyoming 82071

(RECEIVED December 22, 1992; REVISED MANUSCRIPT RECEIVED February 26, 1993)

Abstract

The peptide (Gly-L-Tyr-L-Asp-L-Lys-L-Tyr)₆, referred to as F4-6, was synthesized as a model for a schistosome eggshell protein in which the Gly-Tyr-Asp-Lys-Tyr consensus sequence is repeated over 40 times. Analysis by CD, Fourier transform infrared spectroscopy, potentiometric and spectrophotometric titrations, NMR, and molecular modeling suggests that F4-6 forms some type of left-handed structure. A likely possibility appears to be a left-handed α -helix stabilized by Lys_i-Asp_{i+4} salt bridges and possibly Asp_i-Tyr_{i+4} hydrogen bonding and Tyr-Tyr interactions. Spectroscopic studies of a number of F4-6 analogues support this conclusion. For example, substitution of D-Ala for Gly produces a peptide with enhanced left-handed helical spectral characteristics, whereas an L-Ala substitution results in a peptide with minimal structure. These studies suggest that the F4 protein from *Schistosoma mansoni* may be the first example of a naturally occurring protein devoid of proline and carbohydrate that forms a left-handed helix composed of L-amino acids, although alternative forms of other left-handed structures have yet to be rigorously excluded.

Keywords: F4-6 peptide; *Schistosoma mansoni*; schistosome eggshell protein

The pathology of schistosomiasis is due to the inflammatory response to the eggs of the parasitic trematode, *Schistosoma mansoni* (or other closely related schistosomes) lodged in the hosts' body tissues. Granulomas form around the eggs in response to soluble carbohydrate and protein antigens secreted by the eggs. When an egg is deposited by the female schistosome, it contains a fertilized ovum and 30–40 vitelline cells. Development of the miracidium larva occurs inside the eggshell lodged in the host tissues. The eggshell is formed shortly before egg laying from protein precursors secreted by the vitelline cells that subsequently become surrounded by the eggshell. These vitelline cells serve as a food supply for the developing miracidium and disappear before hatching. The secreted

eggshell proteins are rapidly and extensively cross-linked, and the resulting eggshell is very resistant to protease digestion (see Wells & Cordingley [1991] for more details on eggshell formation).

Two eggshell precursor proteins have been reported (Bobek et al., 1986; Johnson et al., 1987). One is Gly rich (Bobek et al., 1986), while the other, known as F4, has a high Tyr content (Johnson et al., 1987; Chen et al., 1992). The F4 protein may be involved in the hardening of the eggshell through cross-linking of Tyr residues (Johnson et al., 1987). Both show homology to silk moth chorion proteins (Bobek et al., 1986; Johnson et al., 1987). The majority of the predicted amino acid sequence of F4 consists of a repeating pentapeptide with the consensus sequence Gly-Tyr-Asp-Lys-Tyr (Johnson et al., 1987). Characterization of a genomic clone of the single copy of the gene present in our homozygous laboratory strain of *S. mansoni* shows that there are over 40 repeats, 23 exact and the remainder differing by one conservative substitution. Examination of the DNA sequences shows that silent substitutions are frequent and strongly suggests that selective pressure has been maintaining the observed

Reprint requests to: C. Russell Middaugh, Department of Pharmaceutical Research, WP26-331, Merck Research Laboratories, West Point, Pennsylvania 19486.

⁵ Present address: Vertex Pharmaceuticals Incorporated, 40 Allston Street, Cambridge, Massachusetts 02139-4211.

Abbreviations: α_L , left-handed α -helical; α_R , right-handed α -helical; F4-6, (Gly-L-Tyr-L-Asp-L-Lys-L-Tyr)₆; FTIR, Fourier transform infrared spectroscopy; HPLC, high-performance liquid chromatography; MD, molecular dynamics; TFE, trifluoroethanol.

amino acid sequence for a long period (Johnson et al., 1987). This simple repetitive pattern suggested to us that some type of regular secondary structure is present in the F4 protein. The predominant pentapeptide sequence suggests that a short synthetic peptide can be used as a model for the native protein. In this work, a variety of approaches have been used to examine a hexameric repeat of the consensus pentapeptide sequence (henceforth referred to as F4-6) and analogue peptides, which indicate that the most likely conformation of a large portion of this eggshell protein is in some type of left-handed structure.

Results

Circular dichroism spectroscopy

The most prominent feature of the CD spectrum of F4-6 (Fig. 1A) is the intense negative peak at 188 nm. Less intense positive peaks are present at 201 and 227 nm, with a shoulder near 215 nm. The peak at 188 nm was used to monitor changes induced by temperature, pH, and ionic strength. The negative ellipticity decreases in intensity above 25 °C in 20 mM phosphate buffer and also in the presence of greater than 1 M LiClO₄ at 20 °C (Fig. 1B). The cause of the second thermal transition is not clear but may be an artifact of surface adsorption. At 10 °C, a LiClO₄ concentration above 2 M is required to attenuate the 188-nm signal (data not shown). The CD signal is also influenced by changes in pH (Fig. 1C). The $[\theta]_{188}$ decreases in intensity below pH 4.5 and above pH 8.5, where the Asp side chains and the Lys and Tyr side chains are titrated, respectively.

In methanol or TFE (Fig. 2), the CD spectrum differs markedly from that seen in aqueous solution. These spectra have negative ellipticity in the 210–220-nm range, and an intense positive peak is seen at 196 nm and 200 nm in TFE and methanol, respectively. Although they are not exact mirror images of the spectrum in phosphate buffer, many features are inverted and resemble the distinctive CD of a right-handed α -helix. Figure 3A shows the spectrum of Tyr at the concentration that Tyr residues are present in the F4-6 peptide. Spectral features like those seen in F4-6 are the positive peaks at 201 and 225 nm, and negative ellipticity at 190 nm.

To further assess the contribution of the Tyr rings to the CD spectrum of F4-6, synthetic polymers were studied. The random copolymer poly-(Tyr₃₅,Ala₃₁,Glu₃₄), which is reported to be partially α -helical (Ramachandran et al., 1971), approximates the 40% Tyr content found in F4-6. The CD of poly-(Tyr₃₅,Ala₃₁,Glu₃₄) shows a marked helix-indicating shoulder near 220 nm (Fig. 3B). Nevertheless, significant random structure is clearly present as indicated by the strong negative ellipticity near 200 nm. Because this helical structure is thought to be stabilized by Glu_i-Tyr_{i+4} interactions (Ramachandran et al., 1971), an increase in

Fig. 1. A: CD of F4-6 under various conditions; pH 6.5 at 10 °C (solid line) and 70 °C (dashed line); pH 2.7 at 10 °C (dotted line). Spectra were recorded at a peptide concentration of 15 μ M employing 2-mm path-length cells. B: Alteration of F4-6 structure (open circles) and versus LiClO₄; ellipticity at 188 nm versus temperature (open circles) and versus LiClO₄ concentration at 20 °C (solid circles). C: F4-6 ellipticity at 10 °C and 190 nm as a function of pH. The data were collected as 60-s averages using a bandwidth of 1.5 nm.

pH would be expected to disrupt the conformation. At pH 11.4, the CD spectrum has a stronger peak at 200 nm and little intensity near 220 nm. This change indicates a transition to a more disordered form upon deprotonation of the Tyr side chain.

Fig. 2. CD spectra of F4-6 in methanol (solid line) and TFE (dashed line). The data were collected at 10 °C using a cuvette with a pathlength of 2 mm.

In addition to the complication of potential Tyr contributions, association of peptides may give rise to altered CD signals. No concentration dependence, however, was observed in the range that the CD spectra have been obtained (2–260 μ M), suggesting that the aggregation state is not changing. Gel filtration and equilibrium sedimentation experiments indicate that F4-6 as well as its analogues exist in a monomeric state (data not shown). Dityrosine formation has been observed in proteins after exposure to UV radiation (Lehrer & Fasman, 1967) and occurs in F4-6. The CD intensities, however, are not affected (data not shown).

FTIR spectroscopy

The low solubility of F4-6 prevented recording infrared spectra in aqueous solution. In films cast from water solution, the FTIR spectrum of F4-6 is characterized by an amide I peak near 1,668 cm^{-1} with weaker peaks at 1,696, 1,682, 1,655, 1,642, and 1,629 cm^{-1} (Fig. 4A). The peaks at 1,696 and 1,629 cm^{-1} can be assigned to the Asp and Lys side chains, respectively (Venjaminov & Kalnin, 1990a). A peak at 1,668 cm^{-1} is not consistent with the usually encountered secondary structures (Susi & Byler, 1986) although turns and left-handed helical β -aspartyl esters have shown absorbance at this position (Bradbury et al., 1962; Hashimoto & Arakawa, 1967). The sharp peaks at 1,515 and 1,600 cm^{-1} are assigned to the unionized and ionized Tyr residues, respectively (Venjaminov & Kalnin, 1990a). The signals at 1,642 and 1,655 cm^{-1} presumably arise from different forms of disordered structure. The sharp peak near 1,614 cm^{-1} may reflect the presence of some intermolecular β -structure in the film (Jackson & Mantsch, 1991). Interestingly, even in the spectrum (Fig. 4B) of the intact eggshell, an amide I band near 1,664 cm^{-1} , and an amide II signal at 1,536 cm^{-1} , also seen in β -aspartyl esters can be distinguished.

Fig. 3. CD spectra of L-Tyr and L-Tyr-containing polymers. **A:** Spectrum of the estimated contribution of Tyr to the F4-6 CD spectrum (dashed curve) and the spectrum of F4-6 itself (solid curve). The former spectrum was calculated by assuming additive contributions from all Tyr residues in F4-6. **B:** Poly-(Tyr₃₅,Ala₃₁,Glu₃₄) in 20 mM phosphate buffer under different tyrosyl hydroxyl protonation states; at pH 6.5 (protonated, solid curve); at pH 11.4 (ionized, dashed curve). **C:** The spectrum of poly-(Tyr,Glu) in 20 mM phosphate at pH 6.5. The polymers are randomly ordered with respect to the primary structure. All spectra were recorded at 10 °C as described in Figure 1.

Potentiometric and spectrophotometric titrations

In the pH region from 2 to 6.5, the aspartic acid side chains and the C-terminal carboxyl group of F4-6 can be titrated. Figure 5A shows a potentiometric titration of F4-6 in this range, with a broad transition centered at pH 4.0. The pK_a of an Asp side chain is typically near 4.5 (Creighton, 1984).

Fig. 4. FTIR spectra of the F4-6 peptide and the *S. mansoni* eggshell. Spectrum A is of a film cast from water solution. Spectrum B is of an area from an individual eggshell obtained through an FTIR microscope. Both spectra have been partially deconvoluted by the method of Kaupinen et al. (1981).

A spectrophotometric titration of F4-6 was also performed to probe the environments of the F4-6 Tyr residues. The complex nature of the titration curve (Fig. 5B) indicates the presence of multiple Tyr environments.

NMR spectroscopy

The proton NMR spectrum of the F4-6 peptide in water shows extensive peak overlap as a consequence of the highly repetitive nature of the peptide. Therefore, detailed resonance assignments are not feasible. Nevertheless, changes in both the chemical shifts and linewidths of some resonances can still be used to follow secondary structure perturbations induced by changes in temperature or pH. At a pH of 2.1, the spectrum is indicative of an unstructured peptide with narrow resonance linewidths and limited chemical shift dispersion. As the pH is raised to less extreme values between ca. 3.4–9.2, several changes in the spectrum are evident (Fig. 6). Most notable is the increase in chemical shift dispersion of resonances, particularly in the aliphatic region of the spectrum (1–5 ppm) with numerous resonances appearing at positions well

shifted from values characteristic of unstructured peptides (Bundi & Wüthrich, 1979). These changes are also evident for the Tyr aromatic protons near 7 ppm (Fig. 6). Furthermore, most of the narrow components that were evident in the spectrum at pH 2.1 have now disappeared, suggesting formation of stable secondary structure. As the pH is raised above 11, the chemical shift dispersion of the resonances is greatly reduced again implying an unstructured conformation.

At pH 5.5, similar spectral changes are evident as the temperature is increased. At low temperatures (5–15 °C), the spectra are indicative of a structured peptide. As the temperature is raised above 25 °C, linewidths begin to sharpen and chemical shift dispersion is reduced, indicating the disruption of ordered secondary structure, which is complete at 35 °C (data not shown).

Molecular modeling

α -Helices were found to be the lowest energy conformations of the starting structures examined. Therefore, right- and left-handed forms were minimized in the presence of 200 water molecules using the potential of Hagler et al. (1974). The final energy values for the minimized α_R - and α_L -helices are 10 and 1.7×10^3 kcal/mol, respectively. The minimized structure of the α_R -helix showed little interaction among the side chains, whereas Lys_{*i*}–Asp_{*i*+4} salt bridges are apparent in the α_L -helix (Fig. 7). Consistent with the observation of an elevated pK_a , Tyr residues are positioned to form a hydrogen bond with the aspartyl carboxyl group forming a Lys_{*i*}–Asp_{*i*+4}–Tyr_{*i*+8} triad that repeats along the helix (Kinemage 1). The proximity of many phenolic rings to one another also suggests possible Tyr–Tyr interactions as well.

To further investigate possible side-chain interactions and their role in stabilizing the structure of F4-6, more detailed MD simulations were performed on α_R - and α_L -helices as well as structures with multiple type I and type I' β -turns.

For each of the solvated, fully movable production MD simulations, the total energy of the simulation was well conserved and did not require rescaling of velocities. Likewise, the simulation temperature remained stable at 300 ± 15 K, and the total potential energy was well maintained with fluctuations within ± 30 kcal.

The components of the potential energy for the four secondary structural types investigated were all examined separately. For the two-turn conformations and the α_R -helical structure, the partitioned energies showed slight fluctuations but overall were quite well behaved. The average peptide self energies for each of these three structures were comparable. Similarly, the average peptide–solvent interaction energies and the solvent self energies for each form were not drastically different (Table 1). Because all structures were hydrated by a sheath of water, the number of water molecules is not the same for each

Fig. 5. Potentiometric (A) and spectrophotometric (B) titration of F4-6 at 20 °C. The titration curve of *N*-acetyl tyrosine amide is represented by open circles in B. Experimental details are described in the text.

simulation. As a result, the small differences in the total energies of the different structure forms are probably not significant. Also, the compensatory oscillations observed during the simulation (described below) cause the relative variation of the overall energy to be small relative to the component energies.

This type of energetic analysis for the α_L -helical conformation, however, was substantially different. Whereas the total potential energy of the α_L -helical system re-

mained relatively consistent, the component energy plots clearly showed oscillatory behavior (Fig. 8). Two forms of peptide conformational stability are apparent corresponding to the lower (type A) and upper (type B) segments of the energy diagram. From the simulations, it appears that the α_L -helical peptide adopts each conformational type for approximately half of the time. Throughout the type A periods of the simulation (60–80 ps and 105–125 ps), the peptide self energy of the α_L -helical form is lower than that of the α_R -helical and the two-turn structures. During the type B behavior, the α_L -helical form adopts a different conformation, which possesses a peptide self energy comparable to those found for the α_R -helical and turn secondary structures. In the type B case, both the peptide and solvent self energies are higher than they are for type A yet the peptide–solvent interaction energy has dropped relative to the type A value. Within experimental error, the total energy for all four of the structures is consistent with the low conformational stability of any one particular form of the peptide.

Examination of the three-dimensional structure of the peptide at various times during the simulation reveals that the oscillating pattern in the α_L -helical energetic trajectory profiles can be attributed to the making and breaking of favorable salt bridges between the residues Lys_i and Asp_{i+4} . During the MD simulation, only the electrostatic component of the peptide self energy shows marked fluctuations (not illustrated), and these changes in electrostatic energy correlate with those seen in the peptide self energy while other components of the total energy remain relatively constant. An examination of the separation distances between $\text{Lys}_i \text{N}_\epsilon$ and $\text{Asp}_{i+4} \text{O}_{\delta 1}$ or $\text{O}_{\delta 2}$ during the simulation offers a structural explanation for this phenomenon. For type A structure, five salt bridges exist in the peptide conformation allowing each ionic residue to form a stabilizing electrostatic interaction (see

Fig. 6. ^1H -NMR spectra at 500 MHz of F4-6 in D_2O at various pH values. Spectra were collected at 13 °C with a peptide concentration of 150 μM .

Fig. 7. Stereo pair of the energy-minimized structure of the α_L -helical conformation of F4-6. The minimization was performed using DISCOVER (Biosym Inc.) with a 3-Å hydration layer (not shown) to a maximum derivative of 0.01 kcal/Å. The display depicts (Gly/white, Tyr/blue, Asp/red, Lys/green, Tyr/magenta)₆ with the backbone and carbonyl groups in yellow. The final structure suggests Lys-Asp-Tyr and Tyr-Tyr interactions.

Fig. 9). In the type B regions, two of the five possible salt bridges are disturbed, and water solvates each ionic side chain.

Analogue peptides

To further explore the conformation of the F4-model peptide, analogues were synthesized in which potentially important structural residues were altered. The GYDKY consensus sequence (Johnson et al., 1987; Chen et al., 1992) shows substitutions in the repeat at various points

along the sequence with Ser and Glu residues frequently found in the place of Tyr and Asp side chains, respectively. Thus, a peptide (F4-Ser) based on one of these natural sequences was prepared (Table 2). The far-UV CD spectrum of this peptide possesses much weaker negative and positive ellipticity at 190 and 228 nm, respectively, as well as reduction of the 200-nm shoulder compared to F4-6 (Fig. 10A). The near-UV CD of F4-Ser is also reduced in intensity and blue shifted approximately 7 nm relative to F4-6 (inset, Fig. 10A). The amide I FTIR spectrum of F4-Ser films manifests a significant decrease in intensity

Table 1. Molecular dynamics energies for F4-6 during 60.5–160 ps^a

	α_L -Helix	α_R -Helix	Type I β -turn	Type I' β -turn
Peptide self	$-1,036 \pm 52$	-948 ± 34	-965 ± 34	-913 ± 38
Solvent self	$-1,741 \pm 60$	$-1,570 \pm 47$	$-1,786 \pm 40$	$-1,570 \pm 37$
Peptide-solvent	$-1,688 \pm 107$	$-1,986 \pm 69$	$-1,833 \pm 55$	$-1,946 \pm 55$
Kinetic	$1,247 \pm 21$	$1,195 \pm 22$	$1,226 \pm 20$	$1,196 \pm 21$
Total	$-3,218 \pm 1$	$-3,308 \pm 1$	$-3,319 \pm 1$	$-3,233 \pm 1$

^a Units are CHARMM energy units, errors are ± 1 standard deviation from the mean.

Fig. 8. Energy profiles of F4-6 in α -helical and β -turn conformations during MD simulations. **A, B:** Peptide self energies. **C, D:** Peptide-solvent interaction energies. In **A** and **C**, the solid trace represents the α_L -helix energy, and the dashed line is the α_R -helix energy. In **B** and **D**, the solid and dashed lines represent type I and I' β -turns, respectively. Energy profiles from the final phases of the simulation are displayed.

and shift of the F4-6 $1,667\text{ cm}^{-1}$ band to $1,670\text{ cm}^{-1}$ concomitant with increases in peaks at $1,655$ and $1,641\text{ cm}^{-1}$, which can be assigned to disordered structure as indicated above. Very little change in the CD spectrum of F4-Ser is observed between 5 and 90°C . Overall, the spectral characteristics of the F4-Ser peptide resemble those of structurally disrupted F4-6.

Because of the possible importance of salt bridges in stabilizing F4-6, a second analogue was created in which each Asp residue was replaced with an Asn. The CD spectrum of the Asp $\rightarrow$ Asn analog (F4-Asn) displays a 3-nm red shift in the intense 190-nm negative peak with a small increase in the magnitude of the signal when compared to the parent F4-6 peptide (Fig. 10A). In F4-Asn, the positive signal near 200 nm is reduced to a shoulder, but comparable positive ellipticity is maintained between 208 and 240 nm. The near-UV CD spectrum of F4-Asn (inset, Fig. 10A) displays a slightly greater peak intensity than that observed in F4-6. The amide I FTIR spectrum of F4-Asn is very similar to that of F4-6 (Figs. 4A, 11B). Although F4-Asn has CD spectral features similar to F4-6, the dependence of the spectrum on the presence of salts differs significantly. The CD spectrum of F4-6 remains unchanged in the presence of up to 1 M LiClO_4 , whereas the addition of less than 100 mM NaCl or LiClO_4 (in addition to the 20-mM sodium phosphate buffer) decreases the negative intensity of the 191-nm peak in the CD spec-

Fig. 9. Proposed interactions in F4-6 in the Type A region of the MD simulation of the left-handed helical peptide. The structure shows possible helix-stabilizing interactions including Lys-Asp and Tyr-Tyr pairings, as well as a potential Asp-Tyr interaction.

trum of the Asn derivative to about $-12,000\text{ deg cm}^2\text{ dmol}^{-1}$ (not illustrated). Further evidence for the relative decrease in stability of the F4-Asn analogue comes from the observation that the peptide undergoes thermal disruption in a less cooperative manner at temperatures below 20°C (Fig. 10D).

Because Gly appears as every fifth amino acid and from inspection of the DNA sequence it is clear that this codon has been strongly maintained by selection we deduced that this residue probably plays a critical role in the F4-6 structure. Gly was therefore replaced with the two optical isomers of Ala. The spectrum of the L-isomer ana-

Table 2. Amino acid sequences of the F4-6 analog peptides^a

F4-6	(Gly-Tyr-Asp-Lys-Tyr) ₆
F4-Ser	Lys-Tyr- Gly-Ser-Asp-Lys-Tyr- Gly-Tyr-Glu-Lys- Gly-Tyr-Asp-Lys-Tyr- Gly-Ser-Asp-Lys-Tyr- Gly-Tyr-Glu-Lys- Gly-Tyr-Asp-Lys-Tyr- Gly-Ser-Asp
F4-Asn	(Gly-Tyr-Asn-Lys-Tyr) ₆
F4-L-Ala	(Ala-Tyr-Asp-Lys-Tyr) ₆
F4-D-Ala	([D-Ala]-Tyr-Asp-Lys-Tyr) ₆
F4-Glu	(Gly-Tyr-Glu-Lys-Tyr) ₆

^a All residues are L-amino acids, except where noted.

Fig. 10. CD of F4-6 and analogue peptides. **A:** Far-UV CD spectra of F4-6 (solid line), F4-Ser (dotted line) and F4-Asn (dashed line). **B:** Far-UV CD spectra of F4-D-Ala (solid line) and F4-L-Ala (dotted line). The near-UV CD spectra are displayed in the insets for both A and B. The peptides were dissolved in 20 mM phosphate buffer, pH 6.5, at a peptide concentration of 15 μ M. The data were recorded at 10 $^{\circ}$ C with path lengths of 2 and 10 mm for the far- and near-UV spectra, respectively. **C:** The spectra of F4-Glu at pH 4.5 (solid line) and 2.1 (dashed line) at 10 $^{\circ}$ C; refer to the text for further details. **D:** Thermal dependence of the CD at 190 nm for F4-Asn (O), F4-D-Ala (●) and at 196 nm for F4-Glu (▲). Data points are 30-s averages; other conditions are as in A–C.

logue (F4-L-Ala) manifests only weak ellipticity at 190 and 228 nm and in the near UV (Fig. 10B) but a strong Tyr signal at 200 nm. Furthermore, F4-L-Ala displays no FTIR peak near 1,668 cm^{-1} in the amide I region (Fig. 11D) and no evidence of a thermal transition. Thus, this peptide seems to exhibit little ordered structure. In contrast, replacement of Gly with the D-optical isomer of Ala (F4-D-Ala), which would be predicted to enhance any tendency toward left-handed helix formation, does indeed produce this result when compared to F4-6, F4-L-Ala, and the F4-Ser peptides. Compared to F4-6, stronger CD signals are seen at 188, 228, and 274 nm (Fig. 10B) and in the IR at 1,668 cm^{-1} (Fig. 11C). This peptide displays a well-defined thermal transition comparable to F4-6 (Fig. 10D).

Finally, an analogue was prepared in which Asp was substituted by Glu (F4-Glu). Rather than the usual negative peak near 190 nm, this peptide produces a strong positive ellipticity maximum at 196 nm with an accompanying structured negative signal between 210 and 240 nm

(Fig. 10C). Deprotonation of the Glu residues by lowering the pH to 2.1 results in extensive loss of both the positive and negative CD signals with only the 200-nm Tyr band now apparent. The low solubility of F4-Glu did not permit amide I FTIR and near-UV CD spectra to be obtained. The F4-Glu analogue also displays a large thermal transition between 10 and 80 $^{\circ}$ C (Fig. 10D). Unlike F4-6 and several of the other structured peptides, however, this transition is very broad, suggestive of a noncooperative conformational alteration.

The proton NMR spectra of the F4-D-Ala and F4-Asn peptides show well-dispersed resonance peaks. In contrast, the F4-L-Ala and F4-Ser spectra show little dispersion and sharper linewidths. In agreement with the CD data, this narrowing suggests that little stable secondary structure is present in these latter two peptides.

The CD spectrum of each of the analogues was also obtained in TFE. As seen with F4-6 itself, the spectra were inverted relative to their water counterparts, although not exact mirror images. As an example, the spectra of F4-

Fig. 11. Deconvoluted amide I FTIR spectra of F4-6 analogue peptides. Spectra are shown for F4-Ser (A), F4-Asn (B), F4-D-Ala (C), and F4-L-Ala (D). Samples were prepared as films cast on AgCl discs.

D-Ala and F4-L-Ala in TFE are shown in Figure 12. Although the far-UV spectrum of the D-Ala form is much lower in intensity than that of the L-Ala peptide, both exhibit a positive peak near 196 nm and a negative signal in the 210–240-nm region.

To further investigate the preference of the various analogues for a particular helix sense, energy minimizations

were performed in vacuo for all analogue peptides. The energy differences between α_L - and α_R -helical conformations of each peptide after energy minimization are illustrated in Figure 13 and the resultant structures in Kinemages 1–6. Both F4-6 and F4-D-Ala minimized structures have lower energies in the α_L -helical conformation than when in an α_R -helix. F4-Asn and F4-L-Ala show only a weak preference for the α_R -conformation. In contrast to the other peptides, F4-Glu shows a very strong preference for the α_R -conformation over the α_L -helical form (Fig. 13).

Fig. 12. Far-UV CD spectra of F4-6 alanine derivatives in TFE at 10 °C; F4-D-Ala, solid line; F4-L-Ala, dotted line.

Discussion

The CD spectral features of F4-6 are most consistent with left-handed β -turns (e.g., type I') or left-handed helices. Nevertheless, the value of the ellipticity at 190 nm suggests that only 30–40% of the peptide is highly ordered if the magnitude of this peak is used as a measure of the helical state (based on a value of +75,000 deg cm² dmol^{−1} for a pure α_R -helix). As the dynamics simulations suggest, the low helical content may be the result of “frayed” ends as well as transient salt-bridge disruptions. Although

Fig. 13. Preference for sense of α -helix in F4-6 and analogue peptides. Each peptide was placed in either an α_R - or α_L -helix and minimized in vacuo until convergence was obtained. The abscissa is the difference in the final total energies of the right-handed and left-handed α -helices. The positive difference values indicate a preference for the α_L -helix, and negative values suggest that the α_R -helix is favored.

Tyr side chains contribute to the negative peak at 188 nm, the sensitivity of this signal to denaturing conditions suggests that secondary structure is responsible for the majority of this negative ellipticity. This is supported by the observations that extremes of pH, temperature, and ionic strength decrease the negative ellipticity at 188 nm to a value expected for the Tyr residues alone. This conclusion is further supported by the NMR data in which the increase in the chemical shift dispersion and the resonance linewidths coincide with changes seen in the CD spectra as a function of pH and temperature. We cannot exclude the possibility that induced optical activity in phenolic rings as an intermediate consequence of structure formation makes a contribution to the observed ellipticity signals. Examination of Tyr-containing amino acid polymers in regular conformations, however, suggests that the presence of enhanced phenolic transitions themselves do not necessarily obscure the peptide bond transitions in the CD spectra. For example, in poly-(Tyr₃₅,Ala₃₁,Glu₃₄), the spectral features of its α_R -helical nature are clearly evident. Because tyrosine rings have no transition in the 210–220-nm range, the signal in this region also probably arises from the conformation of the peptide backbone. Right-handed α -helices, β -sheets, and type I β -turns all produce negative ellipticities in the 210–225-nm region (Bandeekar et al., 1982). In contrast, a positive signal in this region could arise from either α_L -helices or certain types of β -turns (Chang et al., 1978; Applequist, 1981).

The formation of aggregates does not appear to cause the unusual CD features or the NMR linewidth changes seen in the spectra of F4-6. Indeed, the spacing of hydrophobic residues in F4-6 does not favor association (Hodges et al., 1990), nor can the F4-6 amino acid sequence form an amphipathic α -helix, a potentially asso-

ciative structure. A change in solvent polarity appears to invert the CD spectral features, suggesting a change in conformation rather than, for example, dissociation of coiled-coils (Tachibana & Kambara, 1967; Lau et al., 1984). This spectral inversion does not represent an exact mirror image. This is expected, however, if the local chirality of the amino acids is not perfectly inverted or there is helix distortion (Chen & Woody, 1971; Manning & Woody, 1991). A similar solvent-dependent effect has also been reported for the synthetic polymer poly- β -benzyl-L-aspartate (Giancotti et al., 1972). This inversion is also observed in polypeptides devoid of aromatic amino acids (Drake et al., 1988), so it is unlikely that this effect is solely caused by the interaction of aromatic rings.

The FTIR spectra also are consistent with some type of left-handed helical conformation for F4-6. The 1,667 cm^{-1} amide I and 1,530 cm^{-1} amide II peaks are closer to the position seen with α_L -helices (Bradbury et al., 1962; Hashimoto & Arakawa, 1967) than any other commonly found secondary structure and are near those seen for some types of β -turns. Although one might expect, a priori, left- and right-handed helices to have similar amide I peak positions, the distinct geometry of the left-handed form may be reflected in the observed frequency shift. The presence of additional peaks at 1,642 and 1,655 cm^{-1} are presumably due to random forms of the F4-6 peptide (Venyaminov & Kalnin, 1990b), although helical bands are theoretically possible in the 1,630–1,640 cm^{-1} region (Torii & Tasumi, 1992).

Both energy minimization and MD simulations suggest that the F4-6 peptide can exist in an α_L -helix. The formation of Lys_{*i*}–Asp_{*i*+4} salt bridges may stabilize the α_L -helix. This interaction has been reported to stabilize an α_R -helix in either order, but the ability to do so is less than pairings of other charged side chains (Marqusee & Baldwin, 1990; Merutka & Stellwagen, 1991). In agreement with the MD simulation, the presence of charge on the Asp and Lys side chains appears to enhance secondary structure suggesting that salt bridges may stabilize the conformation (but see below concerning F4-Asn). Interaction with the basic Lys side chain can account for the decrease of the Asp pK_a observed in the potentiometric titration. Initial minimization studies suggest that hydrogen bonding between the Asp_{*i*} carboxyl group and the Tyr_{*i*+4} phenolic hydroxyl hydrogen may also contribute to the stability of F4-6. The multiple transitions and shifts of the pK_a 's suggest that some Tyr residues may be interacting with the Asp side chains, while others are not perturbed by either of the charged groups. This stabilization could be analogous to the Glu_{*i*}–Tyr_{*i*+4} interaction responsible for α_R -helix formation in poly-(Tyr₃₅,Ala₃₁,Glu₃₄) (Ramachandran et al., 1971). The high Tyr content may also allow Tyr–Tyr interactions to stabilize the α_L -conformation as suggested by the modeling (Fig. 9).

Examination of variants of the F4-6 peptide also suggests that it exists in some type of left-handed structure.

The structure of F4-6 is apparently disrupted as detected by CD either in a natural sequence found in the F4 gene in which Asp $\rightarrow$ Glu and Tyr $\rightarrow$ Ser substitutions occur as well as alteration in the pentameric phasing induced by deletions (F4-Ser) or in an analogue in which the Gly residues have been replaced by L-Ala. The latter observation can be rationalized by the necessity of the flexibility imparted by Gly residues permitting left-handed conformations and the limitation in this mobility imposed by the L-Ala substitution. In contrast, substitution of D-Ala in this position might be expected to enhance any tendency toward left-handed structure as suggested by the CD of this peptide (Fig. 10B). Substitution of Glu for Asp every fifth residue appears to invert the conformational sense of the peptide and produce a CD spectrum very much like that of an α_R -helix (see above). This result is consistent with the formation of α_R -helices by a variety of peptides containing appropriately spaced Lys and Glu residues (Marqusee & Baldwin, 1987) and with the length of side chains dictating helix sense in copolyaspartates (Kusakawa et al., 1985). The CD spectrum of F4-Glu does not possess the properly positioned familiar double minima of an α_R -helix, but it does resemble the spectrum of poly- β -benzyl-L-aspartate in an α_R -conformation (Giancotti et al., 1972) possibly suggesting deviations in helix length or small deviations in the helix dihedral angles (Manning et al., 1988; Manning & Woody, 1991).

Because Lys_{*i*}-Asp_{*i+4*} salt bridges can be postulated to stabilize the F4-6 structure in a manner analogous to synthetic Ala-based peptides (Marqusee & Baldwin, 1987), the substitution of Asp by Asn was also examined. Surprisingly, elimination of this negatively charged residue does not appear to abolish secondary structure as monitored by CD, FTIR, and NMR spectra. Like F4-6, this structure can be disrupted by high temperatures and increased salt concentrations. The strong dependence on ionic strength suggests that an electrostatic component is still involved in stabilizing the secondary structure. A possible origin of this effect might be an ion-dipole interaction between the Asn and Lys side chains. Although, an Asn or Asp side chain can form a turn by hydrogen bonding to backbone peptide carbonyl oxygens (Abbadi et al., 1991), the interaction with a charged Lys side chain is probably more favorable. Furthermore, peptides that contain Asx turns require Pro to further support the unique turn conformation. The surprising stability of F4-Asn suggests other stabilizing factors such as Tyr aromatic-aromatic interactions (Burley & Petsko, 1985) or Tyr phenolic hydrogen bonding to a charged side chain may be involved in stabilizing this peptide as suggested for F4-6.

Energy minimizations of all the peptides as well as MD simulation of the F4-6 peptide suggest that an α_L -helix is a plausible structure for some of these peptides. The peptides with the most intense negative CD signals, F4-6 and F4-D-Ala, exhibit an energetic preference for the α_L -conformation. F4-Glu is the only peptide that shows an

α_R -helix-like spectrum, and energy calculations are consistent with this view.

From consideration of the structures of other naturally occurring peptides and proteins for which high resolution structural information is available, the component amino acids of F4-6 appear to have an inherent propensity to exist with torsion angles that are consistent with an α_L -helix. Analyses of protein crystallographic data show that Gly, Asn, and Lys are the most common residues having dihedral angles characteristic of an α_L -helix (Adzhubei et al., 1987; Nicholson et al., 1989). In fact, these same amino acids have recently been found to frequently occur in five-length clusters when the sequences of a large number of proteins were examined (Karpen et al., 1992). Asp and Tyr also have a greater tendency to adopt ϕ and ψ angles that define an α_L -helical conformation relative to the remaining amino acids (Adzhubei et al., 1987). Proteins and peptides with high Pro content often form left-handed helices (Applequist, 1981) as an immediate consequence of the unique stereochemistry of the rigid pyrrolidine ring although helices of the polyproline type have spectral characteristics somewhat different than those manifested by F4-6 (i.e., CD minimum near 200 rather than 188 nm). Channel-forming peptides composed of L-residues may also form left-handed β -helices (10–15 residues per turn), but are only stable in the apolar environment of membranes (Kennedy, 1978) and, as so far described, cannot accommodate the F4-6 sequence.

Naturally occurring left-handed helical structures can occur in peptides consisting entirely of D-amino acids (Blanc & Kaiser, 1984; Mortishire-Smith et al., 1991) and, as indicated above, have been proposed for some peptides possessing a high Pro content (Urry, 1974). Although both senses of α -helices were originally proposed as possible protein structures (Pauling et al., 1951), α_L -helices composed of L-amino acids are so uncommon that only two examples of a single turn have been found in proteins to date (Matthews et al., 1974; Kimura et al., 1992). Steric hindrance between β -carbon atoms and backbone atoms is generally considered to favor right-handed α -helices and forbid those of opposite sense (Ramachandran et al., 1966). Molecular modeling, however, suggests that this does not occur in F4-6.

Bush and collaborators (1981, 1984) provide CD evidence for a left-handed helix in an antifreeze glycoprotein composed of an L-Ala-L-Ala-L-Thr repeating unit. This protein is heavily glycosylated, however, and the carbohydrate groups probably contribute significantly to its unique structure.

The CD spectra of Pro-containing peptides (Wright et al., 1990) and cyclic Ala dipeptides (Bandekar et al., 1982), which exist in left-handed β -turns, have some features in common with that of F4-6. These include a minimum near 190 nm and a maximum near 208 nm. As previously mentioned, these peaks coincide with Tyr CD peaks, so small changes in the positions of the spectral

features are difficult to interpret. Secondary structure predictions (Chou & Fasman, 1974) from the amino acid sequences indicate that contiguous β -turn conformations are preferred for F4-6, F4-Ser, F4-Asn and, to a lesser extent, F4-L-Ala (not illustrated). The torsional angles defining β -turns (Chou & Fasman, 1977) and α -helices are similar such that small distortions of the dihedral angles in turns can approach angles that define an α_L -helix, a more probable conformation for an extended structure. Also, β -turn-like conformations have been proposed to be helix precursors during the folding process (Dyson et al., 1988; Perczel et al., 1992) further emphasizing the close relationship between these two structures. Both Lys_{*i*}-Asp_{*i*+1} and Glu_{*i*}-Lys_{*i*+1} have been reported to stabilize β -turns through side-chain interactions (Ishii et al., 1988; Otter et al., 1989). In contrast, Asp_{*i*}-Lys_{*i*+1} has not been shown to support a β -turn (Liu et al., 1992). Furthermore, helix formation seems to be favored over β -turns in peptides with adjacent Glu_{*i*}-Lys_{*i*+1} and Lys_{*i*}-Glu_{*i*+1} dipeptide pairs that allow the (*i*, *i*+4) salt bridge to form (Marqusee & Baldwin, 1987; Krstenansky et al., 1989). Thus, adjoining oppositely charged residues alone do not appear to dictate β -turns but do favor α -helices if optimally spaced. By analogy, helices of both senses are assumed to be affected in a similar manner. The contribution of noninteracting β -turns to the CD spectra should also be independent of the chain length. In fact, shorter lengths of F4-6 (1, 2, and 3 pentameric repeats) have been synthesized and manifest no structure as detected by CD (not illustrated). A predominantly β -turn conformation should be little affected by increases in chain length unless stabilized by some additional longer range interactions. The dependence of the secondary structure CD signal on the length of the peptide thus suggests a structure with longer range stabilizing interactions, such as a helix. This is strongly supported by the MD simulations, which suggest that all five Lys-Asp charge pairs are necessary to stabilize the α_L -helix in F4-6.

Although 3_{10} -helices and α -helices have similar CD spectra (Applequist, 1981), the 3_{10} -helix in a left-handed form is less likely to form in F4-6 than an α_L -helix because the latter structure is more stable than the 3_{10} -helix in terms of both local conformational energy and hydrogen-bond configuration for a long periodic structure (Richardson, 1981). However, short Ala-based peptides have been suggested to exist in a 3_{10} -helix rather than an α -helix (Mick et al., 1992), so both conformations for F4-6 may be possible. The α -helix is also favored energetically over its other variant counterparts (2.2_7 -, π -, and ω -helices) in both right- and left-handed conformations (Ramachandran & Sasisekharan, 1968). Another possible conformation may be an extended left-handed helix as proposed for charged polypeptides and poly-L-Pro (Tiffany & Krimm, 1972; for review see Woody, 1992), but the CD spectra of these peptides display minima in the 197–208-nm range, which is clearly distinct from the

188–190-nm minima seen for some of the F4 peptides. Also, the former set of peptides do not show a cooperative thermal transition as observed for F4-6.

Why might a left-handed helix be preferred over a right-handed one? The shorter side chain of the Asp that is involved in the salt bridge may inhibit its ability to interact with Lys in a right-handed helical conformation, unlike Glu, which easily supports α_R -helices. As observed in the MD simulations, the α_L -helix is energetically favored over the other structures examined when these salt bridges are formed. The residues constituting F4-6 are found with dihedral angles defining left-handed helices more than other L-amino acids suggesting an inherent propensity for these side chains to permit the required dihedral angles. The α_L -helix might also gain additional stability from Tyr-Asp interactions as hinted by the modeling studies and perturbation of the Tyr absorption peaks (Figs. 5B, 8). Because precedent shows that a Tyr interaction is possible with Glu in a helix (Ramachandran et al., 1971), Asp may impose the same restrictions here as it does when hydrogen bonding with Lys, i.e., interacting more favorably in an α_L -helical conformation. Lastly, the high content of Tyr residues may further enhance the stability of the left-handed helix by aromatic-aromatic interactions (Burley & Petsko, 1985).

Although the insolubility of the eggshells prevents recording CD spectra, the FTIR results suggest that the left-handed helical structure may be present in the highly cross-linked eggshell as well. The amide I maximum is near that of the F4-6 peptide and distinct from that predicted for other known secondary structure types.

In conclusion, at the very least, we postulate that the F4-6 peptide forms some type of left-handed structure. The data presented here indicate that it may be an α_L -helix, but the possibility of another closely related left-handed repetitive structure cannot yet be excluded.

Materials and methods

Unless specified otherwise, all residues mentioned in the text are L-amino acids.

Peptide synthesis and purification

Peptides were synthesized with an Applied Biosystems Model 430A peptide synthesizer employing Fmoc protection (Meienhofer et al., 1979). The sequences of the peptides employed in this study are shown in Table 1. Peptide purity was checked by reverse-phase HPLC, amino acid analysis, and mass spectroscopy. When necessary, the peptides were purified by gel filtration on tandem columns of Sephadex G-15 and G25SF (both 1.4 × 115 cm) run under nitrogen and HPLC using a C-18 column and a 40-mL gradient of 21–26% acetonitrile in 0.1% TFA run under helium. Random poly-(Tyr₃₅,Ala₃₁,Glu₃₄) (~20 kDa), random poly-(Tyr,Glu) (~30 kDa), L-Tyr and poly-

β -benzyl-L-aspartate (~13 kDa) were purchased from Sigma (St. Louis, Missouri).

UV absorption spectroscopy

Spectra were collected on an Aviv 14DS spectrophotometer at 10 °C. For all peptides except F4-Ser, a calculated extinction coefficient of $17,040 \text{ M}^{-1} \text{ cm}^{-1}$ at 276 nm was used for the determination of concentration ($1,420 \text{ M}^{-1} \text{ cm}^{-1}$ for each Tyr residue; see Creighton [1984]). A coefficient of $9,940 \text{ M}^{-1} \text{ cm}^{-1}$ was employed for F4-Ser. Similarly, an extinction coefficient of $1.80 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$ at 276 nm is used for the determination of concentration of poly-(Tyr₃₅,Ala₃₁,Glu₃₄), and $2.03 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$ is used for poly-(Tyr,Glu).

Circular dichroism spectroscopy

CD spectra were recorded on either an Aviv 62DS, JASCO J-500, or a JASCO J-720 spectropolarimeter. Far-UV scans were recorded in a cuvette with a 0.2- or 0.01-cm pathlength; a 1-cm pathlength cuvette was used for near-UV spectra. Unless otherwise specified, spectra were collected at 10 °C and at a peptide concentration of approximately $15 \mu\text{M}$ in 20 mM sodium phosphate buffer, pH 6.5. As a consequence of the limited solubility of the Glu-substituted peptide near neutral pH, this analogue was dissolved at pH 4.5.

Thermal unfolding experiments were conducted by equilibrating peptide solutions in 1-mm pathlength cuvettes for 3 min followed by signal averaging for 60 s. The cuvette was thoroughly washed between each temperature point to minimize adsorption effects because the F4-6 peptide has a strong tendency to adsorb to the walls of the quartz cuvette during prolonged experiments. The pH titrations were performed by adding $1\text{-}\mu\text{L}$ aliquots of 0.1 M HCl or 0.1 M NaOH.

FTIR spectroscopy

An aqueous solution of the F4-6 peptide was deposited onto a polished AgCl disk 2 mm thick by 25 mm in diameter. When a concentric area of dried film approximately 15 mm in diameter was obtained, a transmission FTIR spectrum was immediately recorded using a predetermined spectrum of a blank AgCl disk as the background. The spectrum in the absorbance mode was partially deconvoluted by the method of Kauppinen et al. (1981), using parameters of width at half-width (σ) = 17 and a resolution enhancement factor (K) = 1.9 (Kauppinen et al., 1981). Final deconvolution was performed as described previously (Copeland et al., 1991).

A BioRad Digilab FTS-60 equipped with a high intensity global source and a narrow range MCT detector was used to obtain transmission spectra. Spectra of the solid peptide and schistosome eggshells were recorded through

an Analect XAD microscope with an Analect 6260 FTIR spectrometer as previously described (Aird et al., 1989; Dong et al., 1991).

Potentiometric titration

Titrations were performed under nitrogen with a Radiometer TTA80/TTT80/ABU80/PHM82 titration apparatus at 20 °C. A peptide concentration of 1 mM in the presence of 0.15 M KCl was titrated with 0.15 M HCl and 0.15 M KOH.

Spectrophotometric titration

Absorbance spectra were recorded between 250 and 300 nm with a Hewlett Packard 8450A diode array spectrophotometer. The increase in absorbance at 293 nm was used as a measure of the degree of ionization of the Tyr phenolic hydroxyl group (Freifelder, 1982). Data were analyzed in the form of first derivatives.

NMR spectroscopy

Proton NMR spectra were recorded on a Varian VXR-500S spectrometer operating at 499.843 MHz. A solution of the F4-6 peptide was prepared in D₂O (99.96% atom D, Merck Isotopes) to give a peptide concentration of $150 \mu\text{M}$. The pH of the solution was determined by direct readings with a pH meter and an Ag/AgCl glass electrode at 23 °C. The pH was adjusted from 2 to 12 by adding small amounts of either 0.1 M NaOD or 0.1 M DCl. The sample temperature during data acquisition was maintained at 13 °C. The residual water signal was suppressed by selective presaturation of the H₂O resonance during the recycle delay. Typical data sets consisted of 1,200–2,000 transients collected with a sweep width of 8,000 Hz, 16 K points zero-filled to 64 K points, and a recycle delay plus acquisition time of 2 s.

Molecular modeling

Energy minimizations were performed by the conjugate gradient method with DISCOVER (Biosym) software or with Quanta/CHARMm software (Polygen) employing an adopted-basis set Newton-Raphson algorithm until convergence (approach of the energy to within 0.01 energy units between successive cycles). Starting structures included extended chain, β -strand, π -, 3_{10} -, α_R -, and α_L -helices.

Initial energetic studies done via QUANTA 3.0 used the CHARMm force field and conjugate gradient energy minimization done in vacuo on each conformer (Brooks et al., 1983). QUANTA default nonbonded parameters were used. Detailed atomistic and energetic studies were subsequently undertaken via MD simulations. The calculations were done using CHARMm 2.0 running on an

IBM 3090 3x computer. A shift nonbonded potential was employed to make the long-range interactions run smoothly to zero between 10.5 and 11.5 Å. All calculations used an integration time step of 0.1 fs, saving coordinates every 0.5 ps. Each series of MD calculations involved a 10-ps segment of slow heating to 300 K followed by 50 ps of equilibration at 300 K with a temperature window of 15 K. Final 100-ps production trajectories (at 300 ± 15 K) were produced.

The simulation protocol used to study each conformation of the peptide involved a four-step process. First, MD simulations for 160 ps (as described above) were done in vacuo holding the peptide backbone fixed (in the predefined secondary conformation) thus allowing the side-chain atoms to alter and optimize their positions. Second, each final in vacuo structure from phase 1 was then solvated with a sheath of water (TIPS3P; see Jorgensen [1981]) by centering a preequilibrated 15-Å sphere at each of the 30 amino acids in the peptide. Next, the solvated system was subjected to 160-ps MD simulations with the solvent and side-chain atoms freely movable while the peptide backbone was maintained in its initial secondary structure. For the fourth stage, the backbone constraints were removed and the solvated peptide underwent a final 160-ps simulation of the full system. Structural and energetic analyses were then performed by post-processing the final 100-ps production trajectory of phase 4.

Acknowledgments

This work was supported in part by grant N00014-90-J-1997 to J.S.C. from the Office of Naval Research. We thank Jeff Campbell and Dustin Landis for assisting in preparation of the figures, and Dr. Bruce L. Bush for helpful comments.

References

- Abbadì, A., Mcharfi, M., Aubry, A., Prémilat, S., Boussard, G., & Marraud, M. (1991). Involvement of side functions in peptide structures: The Asx turn. Occurrence and conformational aspects. *J. Am. Chem. Soc.* 113, 2729–2735.
- Adzhubei, A.A., Eisenmeyer, F., Tumanyan, V.G., Zinke, M., Brodzinski, S., & Esipova, N.G. (1987). Approaching a complete classification of protein secondary structure. *J. Biomol. Struct. Dyn.* 5, 689–704.
- Aird, S.D., Steadman, B.L., Middaugh, C.R., & Kaiser, I.I. (1989). Comparative spectroscopic studies of four crotoxin homologs and their subunits. *Biochim. Biophys. Acta* 997, 211–218.
- Applequist, J. (1981). Theoretical π - π^* circular dichroic spectra of helical poly(glycine) and poly(L-alanine) as functions of backbone torsion angles. *Biopolymers* 20, 387–397.
- Bandekar, J., Evans, D.J., Krimm, S., Leach, S.J., Lee, S., McQuie, J.R., Minisian, E., Nemethy, G., Pottle, M.S., Scheraga, H.A., Stimson, E.R., & Woody, R.W. (1982). Conformations of cyclo(L-alanyl-L-alanyl- ϵ -aminocaproyl) and of cyclo(L-alanyl-D-alanyl- ϵ -aminocaproyl); cyclized dipeptide models for specific types of β -bends. *Int. J. Pept. Protein Res.* 19, 187–205.
- Blanc, J.P. & Kaiser, E.T. (1984). Biological and physical properties of a β -endorphin analog containing only D-amino acids in the amphiphilic helical segment. *J. Biol. Chem.* 259, 9549–9556.
- Bobek, L., Rekosh, D.M., Van Keulen, H., & LoVerde, P.T. (1986). Characterization of a female-specific cDNA derived from a developmentally regulated mRNA in the human blood fluke *Schistosoma mansoni*. *Proc. Natl. Acad. Sci. USA* 83, 5544–5548.
- Bradbury, E.M., Brown, L., Downie, A.R., Elliot, A., Fraser, R.D.B., & Hanby, W.E. (1962). The structure of the ω -form of poly- β -benzyl-L-aspartate. *J. Mol. Biol.* 5, 230–247.
- Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., & Karplus, M. (1983). CHARMM—A program for macromolecular energy, minimization, and dynamics calculations. *Comput. Chem.* 4, 187–217.
- Bundi, A. & Wüthrich, K. (1979). Use of amide ^1H -NMR titration shifts for studies of polypeptide conformation. *Biopolymers* 18, 285–298.
- Burley, S.K. & Petsko, G.A. (1985). Aromatic-aromatic interaction: A mechanism of protein structure stabilization. *Science* 229, 23–28.
- Bush, C.A., Feeney, R.E., Osuga, D.T., Ralapati, S., & Yeh, Y. (1981). Antifreeze glycoprotein. Conformational model based on vacuum ultraviolet circular dichroism data. *Int. J. Pept. Protein Res.* 17, 125–129.
- Bush, C.A., Ralapati, S., Matson, G.M., Yamasaki, R.B., Osuga, D.T., Yeh, Y., & Feeney, R.E. (1984). Conformation of the antifreeze glycoprotein of polar fish. *Arch. Biophys. Biochem.* 232, 624–631.
- Chang, C.T., Wu, C.-S.C., & Yang, J.T. (1978). Circular dichroic analysis of protein conformation: Inclusion of β -turns. *Anal. Biochem.* 91, 13–31.
- Chen, A.K. & Woody, R.W. (1971). A theoretical study of the optical rotary properties of poly-L-tyrosine. *J. Am. Chem. Soc.* 93, 29–37.
- Chen, L.L., Rekosh, D.M., & Loverde, P.T. (1992). *Schistosoma mansoni* p48 eggshell protein gene: Characterization, developmentally regulated expression and comparison to the p14 eggshell protein gene. *Mol. Biochem. Parasitol.* 52, 39–52.
- Chou, P.Y. & Fasman, G.D. (1974). Conformational parameters for amino acids in helical, β -sheet, and random coil regions calculated from proteins. *Biochemistry* 13, 211–222.
- Chou, P.Y. & Fasman, G.D. (1977). β -Turns in proteins. *J. Mol. Biol.* 115, 135–175.
- Copeland, R.A., Ji, H., Halfpenny, A.J., Williams, R.W., Thompson, K.C., Herber, W.K., Thomas, K.A., Bruner, M.W., Ryan, J.A., Marquis-Omer, D., Sanyal, G., Sitrin, R.D., Yamazaki, S., & Middaugh, C.R. (1991). The structure of human acidic fibroblast growth factor and its interaction with heparin. *Arch. Biochem. Biophys.* 289, 53–61.
- Creighton, T.E. (1984). *Proteins*. W.H. Freeman, New York.
- Dong, Z., Lewis, R.V., & Middaugh, C.R. (1991). Molecular mechanism of spider silk elasticity. *Arch. Biochem. Biophys.* 284, 53–57.
- Drake, A.F., Siligardi, G., & Gibbons, W.A. (1988). Reassessment of the electronic circular dichroism criteria for random coil conformations of poly(L-lysine) and the implications for protein folding and denaturation studies. *Biophys. Chem.* 31, 143–146.
- Dyson, H.J., Rance, M., Houghten, R.A., Wright, P.E., & Lerner, R.A. (1988). Folding of immunogenic peptide fragments of proteins in water solution. II. The nascent helix. *J. Mol. Biol.* 201, 201–217.
- Freifelder, D. (1982). *Physical Biochemistry*, 2nd Ed., p. 512. W.H. Freeman, New York.
- Giancotti, V., Quadrioglio, F., & Crescenzi, V. (1972). Helix sense of poly- β -benzyl-L-aspartate. *J. Am. Chem. Soc.* 94, 297–298.
- Hagler, A.T., Huler, E., & Lifson, S. (1974). Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals. *J. Am. Chem. Soc.* 96, 5319–5327.
- Hashimoto, M. & Arakawa, S. (1967). Studies of poly- β -benzyl-L-aspartate helix. III. Infrared spectra of copolymers of β -benzyl-L-aspartate with β -p-methyl, chloro, cyano, or nitrobenzyl-L-aspartate in chloroform solution. *Bull. Chem. Soc. Jpn.* 40, 1698–1701.
- Hodges, R.S., Zhou, N.E., Kay, C.M., & Semchuk, P.D. (1990). Synthetic model proteins: Contribution of hydrophobic residues and disulfide bonds to protein stability. *Pept. Res.* 3, 123–137.
- Ishii, H., Okano, T., Maeji, N.J., Inoue, Y., & Chujo, R. (1988). The role of Lys side-chains in the stabilization of β -turn structure. *Polymer J.* 20, 65–72.
- Jackson, M. & Mantsch, H.H. (1991). Beware of proteins in DMSO. *Biochim. Biophys. Acta* 1078, 231–235.
- Johnson, K.S., Taylor, D.W., & Cordingley, J.S. (1987). Possible eggshell protein gene from *Schistosoma mansoni*. *Mol. Biochem. Parasitol.* 22, 89–100.

- Jorgensen, W.L. (1981). Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *J. Am. Chem. Soc.* 103, 335–340.
- Karpen, M.E., de Hasath, P.L., & Neet, K.N. (1992). Differences in the amino acid distributions of 3_{10} -helices and α -helices. *Protein Sci.* 1, 1333–1342.
- Kauppinen, J.K., Moffat, D.J., Mantsch, H.H., & Cameron, D.G. (1981). Fourier transforms in the computation of self-deconvoluted and first order derivative spectra of overlapped band contours. *Anal. Chem.* 53, 1454–1457.
- Kennedy, S.J. (1978). Structures of membrane proteins. *J. Membr. Biol.* 42, 265–279.
- Kimura, S., Kanaya, S., & Makamura, H. (1992). Thermostabilization of *Escherichia coli* ribonuclease HI by replacing left-handed helical Lys⁹⁵ with Gly or Asn. *J. Biol. Chem.* 31, 22014–22017.
- Krstenansky, J.L., Owen, T.J., Hagaman, K.A., & McLean, L.R. (1989). Short model peptides having a high α -helical tendency: Design and solution properties. *FEBS Lett.* 242, 409–413.
- Kusakawa, M., Tsujita, Y., & Takizawa, A. (1985). Optical rotary dispersion and circular dichroism studies of copolyaspartates with long alkyl side chain. *Polymer* 26, 848–852.
- Lau, S.Y.M., Taneja, A.K., & Hodges, R.S. (1984). Effects of high-performance liquid chromatography solvents and hydrophobic matrices on the secondary structure and quaternary structure of a model protein. *J. Chromatogr.* 317, 129–140.
- Lehrer, S.S. & Fasman, G.D. (1967). Ultraviolet irradiation effects in poly-L-tyrosine and model compounds. Identification of bityrosine as a photoproduct. *Biochemistry* 6, 757–767.
- Liu, X., Scott, P.G., Otter, A., & Kotovych, G. (1992). A sequence dependent ¹H-NMR study in the formation of β -turns in tetrapeptides containing charged residues. *Biopolymers* 32, 119–130.
- Manning, M.C., Illangsekare, M., & Woody, R.W. (1988). Circular dichroism of distorted α -helices, twisted β -sheets, and β -turns. *Biophys. Chem.* 31, 77–86.
- Manning, M.C. & Woody, R.W. (1991). Theoretical CD studies of polypeptide helices: Examination of important electronic and geometric factors. *Biopolymers* 31, 569–586.
- Marqusee, S. & Baldwin, R.L. (1987). Helix stabilization by Glu[−]...Lys⁺ salt bridges in short peptides of de novo design. *Proc. Natl. Acad. Sci. USA* 84, 8898–8902.
- Marqusee, S. & Baldwin, R.L. (1990). α -Helix formation by short peptides in water. In *Protein Folding – Deciphering the Second Half of the Genetic Code* (Gierasch, L.M. & King, J., Eds.), pp. 85–94. American Association for the Advancement of Science, Washington, D.C.
- Matthews, B.W., Weaver, L.H., & Kester, W.R. (1974). The conformation of thermolysin. *J. Biol. Chem.* 249, 8030–8044.
- Meienhofer, J., Waki, M., Heimer, E.P., Lambros, T.J., Makofske, R.C., & Chang, C.-D. (1979). Solid phase synthesis without repetitive acidolysis. *Int. J. Pept. Res.* 13, 35–42.
- Merutka, G. & Stellwagen, E. (1991). Effect of amino acid ion pairs on peptide helicity. *Biochemistry* 30, 1591–1594.
- Miick, S.M., Martinez, G.V., Fiori, W.R., Todd, A.P., & Millhauser, G.L. (1992). Short alanine-based peptides may form 3_{10} -helices and not α -helices in aqueous solution. *Nature* 359, 653–655.
- Mortishire-Smith, R.J., Drake, A.F., Nutkins, J.C., & Williams, D.H. (1991). Left-handed α -helix formation by a bacterial peptide. *FEBS Lett.* 278, 244–246.
- Nicholson, H., Söderlind, E., Tronrud, D.E., & Matthews, B.W. (1989). Contributions of left-handed helical residues to the structure and stability of bacteriophage T4 lysozyme. *J. Mol. Biol.* 210, 181–193.
- Otter, P.G., Scott, P.G., Liu, X., & Kotovych, G. (1989). A ¹H and ¹³C NMR study on the role of salt-bridges in the formation of a type I β -turn in N-acetyl-L-Asp-L-Glu-L-Lys-L-Ser-NH₂. *J. Biomol. Struct. Dyn.* 7, 455–476.
- Pauling, L., Corey, R.B., & Branson, H.R. (1951). The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. USA* 37, 205–211.
- Perczel, A., Foxman, B.M., & Fasman, G.D. (1992). How reverse turns may mediate the formation of helical segments in proteins: An X-ray model. *Proc. Natl. Acad. Sci. USA* 89, 8210–8214.
- Ramachandran, G.N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem.* 23, 283–438.
- Ramachandran, G.N., Venkatachalam, C.M., & Krimm, S. (1966). Stereochemical criteria for polypeptide and protein chain conformations. *Biophys. J.* 6, 849–872.
- Ramachandran, J., Berger, A., & Katchalski, E. (1971). Synthesis of physicochemical properties in aqueous solution of the sequential polypeptide poly(Tyr-Ala-Glu). *Biopolymers* 10, 1829–1851.
- Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 34, 167–339.
- Susi, H. & Byler, D.M. (1986). Resolution enhanced Fourier transform infrared spectroscopy of enzymes. *Methods Enzymol.* 130, 290–311.
- Tachibana, T. & Kambara, H. (1967). Enantiomorphism in the super helices of poly(γ -benzyl glutamate). *Kolloid-Z. Z. Polymere* 219, 40–42.
- Tiffany, M.L. & Krimm, S. (1972). Effect of temperature on the circular dichroism of polypeptides in the extended state. *Biopolymers* 11, 2309–2316.
- Torii, H. & Tasumi, M. (1992). Application of the three-dimensional doorway-state theory to analyses of the amide-I infrared bands of globular proteins. *J. Chem. Phys.* 97, 92–98.
- Urry, D.W. (1974). Arterial mesenchyme and arteriosclerosis. Studies on the conformation and interaction of elastin. *Adv. Exp. Med. Biol.* 43, 211–243.
- Venjaminov, S.Yu. & Kalnin, N.N. (1990a). Quantitative IR spectrophotometry of peptide compounds in water (H₂O) solutions. I. Spectral parameters of amino acid residue absorption bands. *Biopolymers* 30, 1243–1257.
- Venjaminov, S.Yu. & Kalnin, N.N. (1990b). Quantitative IR spectrophotometry of peptide compounds in water (H₂O) solutions. II. Amide absorption bands of polypeptides and fibrous proteins in α -, β -, and random coil conformations. *Biopolymers* 30, 1259–1271.
- von Lichtenberg, F. (1987). In *The Biology of Schistosomes: From Genes to Latrines* (Rollinson, D. & Simpson, A.J.G., Eds.), pp. 185–232. Academic Press, London.
- Wells, K.E. & Cordingley, J.S. (1991). *Schistosoma mansoni*: Eggshell formation is regulated by pH and calcium. *Exp. Parasitol.* 73, 295–310.
- Woody, R.W. (1992). Circular dichroism and conformation of unordered polypeptides. *Adv. Biophys. Chem.* 2, 37–79.
- Wright, P.E., Dyson, H.J., Feher, V.A., Tennant, L.L., Waltho, J.P., Lerner, R.A., & Case, D.A. (1990). Folding of peptide fragments of proteins in aqueous solution. In *Frontiers of NMR in Molecular Biology*, pp. 1–13. Alan R. Liss, New York.