Biochemistry

 Volume 34, Number 11 March 21, 1995

Articles

Solution Structure of a Polypeptide Containing Four Heptad Repeat Units from a Merozoite Surface Antigen of *Plasmodium falciparum*[†,‡]

Terrence D. Mulhern,[⊥,§] Geoffrey J. Howlett,[∥] Gavin E. Reid,[#] Richard J. Simpson,[#] Damian J. McColl,[▽] Robin F. Anders,[▽] and Raymond S. Norton*[,⊥]

NMR Laboratory, Biomolecular Research Institute, 381 Royal Parade, Parkville, Victoria 3052, Australia, Department of Biochemistry, University of Melbourne, Parkville, Victoria 3052, Australia, Joint Protein Structure Laboratory, Ludwig Institute for Cancer Research and The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3050, Australia, and Department of Medicine, University of Queensland, Princess Alexandra Hospital, Queensland 4102, Australia

Received October 28, 1994[⊗]

ABSTRACT: The *Plasmodium falciparum* antigen SPAM (secreted polymorphic antigen associated with merozoites) contains an unusual set of heptad repeat units with alanine at the a and d positions. Twelve heptads with the consensus sequence AXXAXXX occur in three blocks of four, linked by short nonrepetitive sequences. A 38-residue polypeptide comprising the first block of four heptad units and five flanking residues at either end, SPAM-H1, has been synthesized and its structure in aqueous solution determined from ^{1}H NMR data. Sedimentation equilibrium showed the peptide to be monomeric in aqueous solution. Its structure was determined from ^{1}H NMR-derived distance and dihedral angle constraints by using distance geometry calculations, restrained simulated annealing, and conjugate gradient energy minimization in the CHARMm force field. The polypeptide contains an α-helix extending from Ser10 (position e of the first heptad) to at least Lys32 (position f of the fourth heptad) and possibly as far as Val35. The helix is bent, partly as result of a kink around residues 19−20. The conformations of the nine N-terminal residues and the six C-terminal residues are not well defined by the NMR data. The rms deviation from the average of the 20 best structures over the well-defined region (residues 11−31, which have backbone angular order parameters > 0.8) was 1.56 Å for backbone heavy atoms (N, C^α, and C) and 2.12 Å for all heavy atoms. ^2H_2O exchange experiments identified slowly exchanging amide protons near the C-terminus and the last two turns of the helix. The unusual stability of the C-terminus reflects the presence of a new C-capping motif, which may involve the side chain of an asparagine in a position external to the C-cap residue. Possible interactions of the H1 sequence with the other two heptad repeat units in the intact merozoite antigen are discussed.

Malaria is a major cause of mortality in many tropical regions of the world. The disease is caused by protozoan parasites of the genus *Plasmodium*, which are transmitted to humans by the bite of the *Anopheles* mosquito. The increasing resistance of the parasites to conventional drug therapy has emphasized the importance of developing

[†] T.D.M. is supported by an Australian Postgraduate Research Award. This work was supported in part by the National Health & Medical Research Council of Australia.

[‡] Coordinates for the 20 structures, together with the final set of NMR restraints, have been deposited at the Protein Data Bank, Brookhaven National Laboratory, Upton, NY 11973 (Bernstein et al., 1977) (filename 1PSM).

* To whom correspondence should be addressed. Fax: +61-3-903 9655. e-mail: ray@mel.dbe.csiro.au.

[⊥] Biomolecular Research Institute.
[§] University of Queensland.
[∥] University of Melbourne.
[#] Ludwig Institute for Cancer Research.
[▽] The Walter and Eliza Hall Institute of Medical Research.

[⊗] Abstract published in *Advance ACS Abstracts*, February 15, 1995.

0006-2960/95/0434-3479$09.00/0

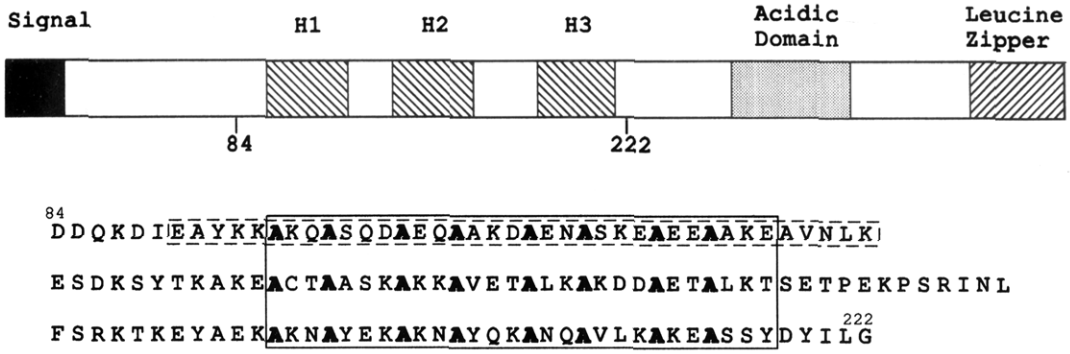

FIGURE 1: Schematic representation of SPAM showing the location and sequence of the alanine heptad repeats. H1, H2, and H3 refer to the three blocks of four heptad repeats. The partial sequence shown spans all three blocks of heptads, which are aligned and enclosed in a box (solid line). The peptide SPAM-H1 is also boxed (broken line). SPAM-H1 includes all four heptads in the H1 block and an additional five residues on both the N-terminal and C-terminal ends.

prophylactic vaccines. It is recognized that the asexual merozoite stage of the parasite's life cycle is vulnerable to attack by the immune system during its brief extracellular exposure prior to invasion of a new erythrocyte. Thus, the potential of vaccines based on merozoite surface antigens has stimulated considerable interest in their identification and characterization (Anders & Saul, 1994).

Merozoite surface antigens include integral membrane proteins, such as MSA-1 (Holder, 1988) and MSA-2 (Smythe et al., 1988), which are both GPI-anchored proteins, and erythrocyte binding antigen (Camus & Hadley, 1985) and apical membrane antigen 1 (Peterson et al., 1989), both typical type I membrane proteins. A number of secreted parasite proteins are also known to associate with the merozoite surface, but the biological significance of these interactions is unknown (Ardeshir et al., 1987; Borre et al., 1991).

One such protein is the secreted polymorphic antigen associated with merozoites (SPAM[1]). The complete genomic sequences of SPAM from two isolates of *Plasmodium falciparum* have been described recently (McColl et al., 1994), and the conserved features of the deduced polypeptide sequences are shown in Figure 1. Repetitive sequence motifs, common to many plasmodial antigens, are also found in SPAM. While SPAM does not contain extensive tandem repeats, such as those in the S-antigen (Anders et al., 1993) and MSA-2 (Smythe et al., 1988), it does contain 12 heptad repeats in three separate blocks, each containing four heptads. Heptad repeat patterns $(abcdefg)_n$, where hydrophobic residues are conserved at the a and d positions and the other positions are occupied by hydrophilic residues, are predictive of the amphipathic α-helices which form the basis of the coiled-coil and helical bundle motifs commonly found in protein structures (Cohen & Parry, 1990). In coiled coils, up to four amphipathic helices wrap around each other with a left-handed supercoil such that residues at the a and d positions are maintained at the hydrophobic interface. The a and d positions are normally occupied by the more hydrophobic residues such as leucine, isoleucine, or valine (Cohen & Parry, 1990; Lupas et al., 1991). Depending on the identities of the hydrophobic residues at the a and d positions and the potential for interhelical salt bridges between residues at the e and g positions, peptides made up of heptad repeats can associate to form tight, coiled-coil dimers or three- or four-stranded helical bundles (O'Shea et al., 1991; Lovejoy et al., 1993; Harbury et al., 1993, 1994). The heptad repeats in SPAM are unusual in that the a and d positions are almost exclusively occupied by alanine. The only other naturally occurring proteins we are aware of that have heptad repeats based on alanine residues are the M proteins of *Streptococcus pyogenes* (Mouw et al., 1988) and the gp17 tail fiber of bacteriophage T7 (Steven et al., 1988).

Because of the novel features of the heptad repeats in SPAM, we have prepared a synthetic peptide encompassing the first heptad repeat block from the sequence of SPAM FC27 clone D10 and investigated its state of self-association by analytical ultracentrifugation and its structure in aqueous solution by NMR. These studies have shown that the peptide is largely α-helical in aqueous solution and does not self-associate.

In the course of the NMR studies, we observed considerable differences in the relative exchange rates of amide protons at the N- and C-termini of the helix, suggesting that some stabilizing influence was present at the C-terminus. A mechanism for the stabilization of helix termini through capping interactions was proposed by Presta and Rose (1988), who showed that α-helices in proteins are usually flanked by clusters of polar residues (N-terminal and C-terminal bounds) having side chains capable of forming hydrogen bonds with the exposed backbone NH protons at the N-terminal turn of the helix or the exposed peptide carbonyl oxygens at the C-terminal turn, respectively. They proposed that the presence of N- and C-terminal bounds, together with suitable intervening sequences, would be sufficient to generate helical conformations without the need for stabilizing tertiary interactions. The helix termini and their flanking residues were labeled as follows:

$$N''\text{-}N'\text{-}N\text{-}N1\text{-}N2\text{-}N3\text{-}\cdots\text{-}C3\text{-}C2\text{-}C1\text{-}C\text{-}C'\text{-}C''$$

where N1-N2-N3 to C3-C2-C1 participate in the helical backbone hydrogen-bonding network and have helical backbone dihedral angles. N and C are the N-cap and C-cap residues, respectively, which participate in the hydrogen-

[1] Abbreviations: SPAM, secreted polymorphic antigen associated with merozoites; NMR, nuclear magnetic resonance; 1D, one-dimensional; 2D, two-dimensional; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; DQF-COSY, double-quantum-filtered scalar-correlated spectroscopy; TOCSY, total correlation spectroscopy; E-COSY, exclusive correlation spectroscopy; HMQC, heteronuclear multiple quantum coherence; TFE, trifluoroethanol; rms, root mean square.

bonding network but do not necessarily have helical backbone angles. Richardson and Richardson (1988) and Rose and Wolfenden (1993) analyzed the frequencies of occurrence of particular residue types at specific positions at and around the N- and C-caps, leading to identification of the hydrogen-bonding patterns and sequence determinants of one N-capping motif, the N-capping box (Harper & Rose, 1993), and two C-capping motifs, the Schellman motif and the α_L motif (Aurora et al., 1994). SPAM-H1 appears to contain a C-capping motif which is different from both the Schellman and α_L motifs and may involve side chain to backbone as well as backbone to backbone hydrogen bonding.

MATERIALS AND METHODS

Materials. The 38-residue peptide EAYKKAKQASQ-DAEQAAKDAENASKEAEEAAKEAVNLK (SPAM-H1) was synthesized using *tert*-butyloxycarbonyl amino acids, 2-(1*H*-benzotriazolyl)-1,1,3,3-tetramethyluronium tetrafluoroborate, and an optimized solid-phase peptide synthesis procedure, as described previously (Reid & Simpson, 1992). The crude peptide was purified to homogeneity by preparative reversed-phase high-performance liquid chromatography and assessed by on-line capillary LC/electrospray mass spectrometry (Ji et al., 1994). The observed molecular mass of SPAM-H1 (4064.6 Da) was in excellent agreement with the theoretical mass of 4064.4 Da. Deuterated solvents were obtained from Cambridge Isotope Laboratories, Woburn, MA.

Analytical Ultracentrifugation. Sedimentation equilibrium data were collected using a Beckman XL-A analytical ultracentrifuge. Distributions were measured at 230 nm and 293 K with a rotor speed of 35 000 rpm, for solutions containing 0.1 mM SPAM-H1, 10 mM K_3PO_4, pH 7.6, with and without 150 mM NaCl. A series of SPAM-H1 concentrations (0.05, 0.5, and 5.0 mM) was analyzed at 277 K in 50 mM sodium acetate, pH 4.9, with a rotor speed of 20 000 rpm. The wavelengths used for 0.05, 0.5, and 5.0 mM SPAM-H1 were 230, 240, and 290 nm, respectively. The 0.05 and 0.5 mM samples were also analyzed at 293 K and 240 nm with a rotor speed of 35 000 rpm. A series of experiments was performed under similar conditions to the NMR experiments, with solutions containing 1.1, 2.2, and 4.4 mM SPAM-H1 in 90% ^1H_2O/10% ^2H_2O at pH 5.0, being analyzed at 277 K with a wavelength of 245 nm and a rotor speed of 40 000 rpm. All data were analyzed using nonlinear regression to yield a molecular mass and standard error. The analysis assumed a single species and a value for the partial specific volume of the peptide (0.709 mL/g) calculated from the peptide composition.

NMR Spectroscopy. NMR samples were prepared by dissolving lyophilized peptide in 90% ^1H_2O/10% ^2H_2O or 99.96% ^2H_2O. The pH was adjusted with small additions of 0.1 M NaO^2H or ^2HCl. Reported pH values were measured at room temperature and are uncorrected for deuterium isotope effects. Sample concentrations were in the range 4−5 mM except where noted. ^1H chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at 0 ppm, either directly against an internal reference or via the chemical shift of the residual H_2O resonance, and ^{13}C chemical shifts to *p*-dioxane at 69.4 ppm (Wishart & Sykes, 1994).

NMR experiments were performed on Bruker AMX-500 or AMX-600 spectrometers. 2D spectra were recorded in phase-sensitive mode using time-proportional phase incrementation (Marion & Wüthrich, 1983). 2D nuclear Overhauser effect spectroscopy (NOESY) experiments (Anil-Kumar et al., 1980; Macura et al., 1981) were recorded with mixing times in the range 50−300 ms. Total correlation spectroscopy (TOCSY) experiments (Braunschweiler & Ernst, 1983) were recorded using the DIPSI-2 spin-lock sequence (Rucker & Shaka, 1989) with spin-lock times of 24 and 54 ms. Double-quantum-filtered scalar-correlated spectroscopy (DQF-COSY) experiments (Rance et al., 1983) were recorded in 90% ^1H_2O/10% ^2H_2O and 99.96% ^2H_2O. Exclusive correlation spectroscopy (E-COSY) (Griesinger et al., 1987) and heteronuclear (^{13}C) multiple quantum coherence (HMQC) (Bax et al., 1983) experiments were carried out on samples dissolved in 99.96% ^2H_2O. Typically, 500 t_1 increments were acquired, with 96 scans at each increment. ^1H sweep widths were 4087.7 Hz at 500 MHz and 5263.2 Hz at 600 MHz, and spectra were acquired over 4K data points. Solvent suppression was achieved by selective, low-power irradiation of the water signal during the relaxation delay (typically 1.8 s) and during the mixing time in NOESY experiments. The probe temperature was controlled using a B-VT1000E unit and a Haake cooling bath and calibrated using methanol (van Geet, 1970). Spectra were processed using UXNMR (Bruker) and analyzed using FELIX (version 2.05, Biosym). Sine-squared window functions, with phase shifts of 60°−90°, were applied in both dimensions prior to Fourier transformation.

Scalar $^3J_{NHC\alpha H}$ coupling constants were measured from 1D, NOESY, and DQF-COSY spectra recorded at 600 MHz. Rows were extracted from NOESY spectra, inverse Fourier transformed, zero filled to 32K, and multiplied by a Gaussian window function prior to Fourier transformation. Intraresidue $C^\alpha H-NH$ NOESY cross-peaks were excluded from this analysis because of the possibility of modulation due to *J*-effects (Szyperski et al., 1992). Peak shapes were then simulated and coupling constants obtained using FELIX. Cross sections along ω_2 of antiphase DQF-COSY doublets were analyzed using inverse Fourier transformation, as described above. The dispersive peak shapes were simulated to take account of the effect of broad line widths on small coupling constants using FELIX and an in-house program, COUPLING.

The exchange properties of amide protons were investigated at pH 4.9, 5.0, and 5.5 at 275 K (pH was measured at room temperature at the end of the experiment). The pH of each sample was adjusted in 90% ^1H_2O/10% ^2H_2O, and then it was freeze-dried and redissolved in precooled ^2H_2O. A series of 1D experiments were recorded, followed by a short duration TOCSY experiment to identify slowly exchanging backbone and side-chain amide protons.

Structural Restraints. Interproton upper bound distance restraints were derived mainly from the 200 ms mixing time NOESY spectrum of a 5 mM sample in 90% ^1H_2O/10% ^2H_2O at pH 4.9 and 275 K. Some additional distance restraints were obtained from NOESY spectra recorded at pH 5.5 and 281 K. Only one pair of nondegenerate $C^\beta H$ protons (Ser10) was free of spectral overlap. As a result, volumes were calibrated using six cross-peaks between protons from the backbone of the helical section of the peptide, *viz.* $d_{\alpha N}(15,18)$, $d_{\alpha N}(15,19)$, $d_{\alpha N}(19,22)$, $d_{\alpha N}(22,25)$, $d_{NN}(18,19)$, and $d_{NN}(19,-$

Table 1: Proton Chemical Shifts of SPAM-H1 at pH 5.5 and 281 K in 90% ^1H_2O/10% ^2H_2O[a]

residue	NH	$C^\alpha H$	$C^\beta H$	others
Glu1	–	4.05	2.10	$C^\gamma H_2$ 2.36
Ala2	8.86	4.37	1.39	
Tyr3	8.47	4.53	3.03	C(2,6)H 7.16; C(3,5)H 6.87
Lys4	8.28	4.26	1.69	$C^\gamma H_2$ 1.37; $C^\delta H_2$ 1.76; $C^\epsilon H_2$ 3.02; $N^\zeta H^+_3$ 7.66
Lys5	8.33	4.23	1.85	$C^\gamma H_2$ 1.50; $C^\delta H_2$ 1.76; $C^\epsilon H_2$ 3.06; $N^\zeta H^+_3$ 7.71
Ala6	8.50	4.30	1.43	
Lys7	8.50	4.29	1.82	$C^\gamma H_2$ 1.50; $C^\delta H_2$ 1.73; $C^\epsilon H_2$ 3.05; $N^\zeta H^+_3$ 7.66
Gln8	8.55	4.34	2.14, 2.05	$C^\gamma H_2$ 2.43; $N^\epsilon H_2$ 7.69, 7.02
Ala9	8.63	4.36	1.47	
Ser10	8.49	4.45	3.99, 3.92	
Gln11	8.59	4.36	2.19, 2.08	$C^\gamma H_2$ 2.44; $N^\epsilon H_2$ 7.69, 6.99
Asp12	8.43	4.61	2.81, 2.74	
Ala13	8.47	4.25	1.51	
Glu14	8.44	4.22	2.14	$C^\gamma H_2$ 2.41, 2.33
Gln15	8.35	4.09	2.17	$C^\gamma H_2$ 2.47; $N^\epsilon H_2$ 7.91, 7.03
Ala16	8.39	4.27	1.50	
Ala17	8.07	4.29	1.53	
Lys18	8.07	4.22	1.94	$C^\gamma H_2$ 1.50; $C^\delta H_2$ 1.60; $C^\epsilon H_2$ 3.03; $N^\zeta H^+_3$ 7.72
Asp19	8.63	4.55	2.79	
Ala20	8.15	4.30	1.54	
Glu21	8.23	4.23	2.18	$C^\gamma H_2$ 2.45, 2.35
Asn22	8.45	4.65	2.92	$N^\delta H_2$ 7.91, 7.04
Ala23	8.30	4.33	1.54	
Ser24	8.30	4.42	4.05	
Lys25	8.25	4.28	1.97	$C^\gamma H_2$ 1.50; $C^\delta H_2$ 1.61; $C^\epsilon H_2$ 3.03; $N^\zeta H^+_3$ 7.70
Glu26	8.42	4.18	2.12	$C^\gamma H_2$ 2.48, 2.40
Ala27	8.23	4.30	1.54	
Glu28	8.29	4.21	2.18	$C^\gamma H_2$ 2.48, 2.35
Glu29	8.32	4.19	2.13	$C^\gamma H_2$ 2.45
Ala30	8.26	4.25	1.51	
Ala31	8.04	4.29	1.51	
Lys32	7.98	4.24	1.95	$C^\gamma H_2$ 1.51; $C^\delta H_2$ 1.62; $C^\epsilon H_2$ 3.03; $N^\zeta H^+_3$ 7.69
Glu33	8.17	4.27	2.14, 2.05	$C^\gamma H_2$ 2.46, 2.39
Ala34	8.15	4.34	1.48	
Val35	7.98	4.12	2.17	$C^\gamma H_3$ 1.02, 1.00
Asn36	8.43	4.78	2.91, 2.80	$N^\delta H_2$ 7.73, 7.02
Leu37	8.28	4.41	1.69	$C^\gamma H$ 1.51; $C^\delta H_3$ 0.98, 0.92
Lys38	7.98	4.21	1.88, 1.78	$C^\gamma H_2$ 1.45; $C^\delta H_2$ 1.72; $C^\epsilon H_2$ 3.05; $N^\zeta H^+_3$ 7.62

[a] Chemical shifts are given in ppm and are referenced to DSS (0 ppm). For methylene protons, two chemical shifts are given only when two resonances could be resolved. Stereospecifically assigned shifts are underlined. The first shift is for the proton with the lower branch number, e.g., $H^{\delta 21}$, $H^{\delta 22}$ for Asn.

20). Distances of 3.4, 4.2, and 2.8 Å were used for $d_{\alpha N}(i,i+3)$, $d_{\alpha N}(i,i+4)$ and $d_{NN}(i,i+1)$ connectivities, respectively (Wüthrich, 1986). The calibration constant obtained in this way yielded a distance of 2.1 Å for the $C^\beta H$ protons of Ser10, compared with the theoretical distance of 1.8 Å. Where cross-peak volumes could not be measured accurately because of peak overlap, the NOEs were classified as strong, medium, or weak and assigned upper bound distances of 3.7, 4.2, or 4.7 Å, respectively, by comparison with peaks of known volumes. Where the cross-peak volume could not even be estimated reliably, an upper bound of 5.0 Å was assigned. The volumes of cross-peaks representing degenerate protons were divided by the appropriate factor before conversion to distances. Pseudoatom corrections (Wüthrich et al., 1983) were added where necessary. Corrections of 0.5 and 1.0 Å were added to distance constraints involving only backbone protons and at least one side-chain proton, respectively, to allow for conformational averaging and errors in volume integration. All lower bound distance constraints were set to 1.79 Å.

Backbone dihedral angle constraints were inferred from $^3J_{NHC\alpha H}$ values measured from spectra recorded at pH 5.5 and 281 K as follows: $^3J_{NHC\alpha H} < 5$ Hz, $\phi = -60° \pm 30°$; 5 Hz $< {}^3J_{NHC\alpha H} < 6$ Hz, $\phi = -60° \pm 40°$; $^3J_{NHC\alpha H} > 8$ Hz, $\phi = -120° \pm 40°$. Where 6 Hz $< {}^3J_{NHC\alpha H} < 8$ Hz, ϕ angles were not constrained. Nondegenerate $C^\beta H$ resonances were observed for nine residues at pH 4.9 and 275 K. Where possible, $^3J_{C\alpha HC\beta H}$ coupling constants were measured from passive couplings as displacements in E-COSY spectra or peak splittings in DQF-COSY spectra. The relative intensities of intraresidue $d_{\alpha\beta}$ and $d_{N\beta}$ NOEs were measured in a 50 ms mixing time NOESY spectrum measured under the same conditions. None of the patterns fitted those expected for any of the three staggered conformations ($\chi 1 = -60°$, 60°, or 180°) (Wagner et al., 1987), as a result of which all $\chi 1$ angles were left unconstrained. The side-chain amide protons of asparagine and glutamine residues were stereospecifically assigned on the basis of the intensities of NOESY cross-peaks from the side-chain amide protons to their respective $C^\beta H$ or $C^\gamma H$ protons (Montelione et al., 1984). $H^{\delta 1}$ of Asn and $H^{\epsilon 1}$ of Gln (trans to the amide carbonyl oxygen) will always be closer to the protons of the preceding methylene carbon regardless of the intervening χ angle.

Structure Calculations. Distance geometry structures were calculated using the program DIANA (Güntert et al., 1991). Preliminary structures were calculated using an initial set of 301 distance constraints. These structures were then used in the resolution of ambiguous NOESY cross-peaks. The final structures were calculated using a set of 312 upper bound distance constraints (one long range ($|i - j| \geq 5$), 125

medium range ($2 \leq |i - j| \leq 4$), 118 sequential, and 68 intraresidue constraints) and 27 backbone dihedral angle constraints. Three REDAC cycles (Güntert & Wüthrich, 1991) were performed with the maximal target function value per residue for locally acceptable segments set to 0.4 Å^2.

The 50 structures with the lowest target function from DIANA were then refined in X-PLOR, version 3.1 (Brünger, 1992). Restrained simulated annealing refinement was carried out using a method adapted from the standard simulated annealing protocol in X-PLOR, refine.inp. The starting temperature was set to 2000 K, and the system was cooled in steps of 50−100 K over 10 ps, with a time step of 1 fs. Each DIANA structure was subjected to 10 rounds of refinement; if the value of the overall energy function decreased during one round of refinement, the resulting coordinates were used as the starting point for the next round, but if the overall energy was higher, the coordinates from the previous round were used. Generally, no further improvement was observed after 6 to 8 rounds. At the end of this refinement procedure, the structures were energy minimized using the CHARMm force field, version 2.2 (Brooks et al., 1983). During the final minimization in CHARMm the electrostatic function was computed with a distance-dependent dielectric. The 20 structures with the lowest overall energies were retained.

RESULTS

Solution Properties of SPAM-H1. Initial 1D and 2D spectra showed that the helical content of the peptide was temperature dependent, with both the amide proton shift dispersion and the number and relative strength of d_{NN} NOESY cross-peaks increasing as the temperature was lowered below ambient. Once at low temperature, however, the conformation was stable in the acid pH range, with spectra recorded at pH 2.5, 4.9, and 5.5 at 275 K showing no major differences.

Addition of trifluoroethanol (TFE), which is often required to stabilize helical structures (Nelson & Kallenbach, 1986; Morton et al., 1994), caused the peptide to aggregate. At 298 K the observable peptide resonances remained sharp but decreased in intensity by a factor of about 7 when TFE was added to 10% (v/v). The signal continued to decrease upon addition of further TFE up to 50% (v/v), dropping about 2-fold with each subsequent addition of 10% TFE. Particulate material was observable at 30% TFE and above. The backbone amide proton chemical shift dispersion (excluding the shift of the Ala2 resonance) increased from 0.52 ppm in the absence of TFE to 0.98 ppm in 50% TFE at 298 K, which is comparable to the 0.83 ppm dispersion observed in the absence of TFE at 275 K. Thus, it appears that TFE promotes formation of a helical structure for SPAM-H1 but that the resulting helix aggregates in this solvent. As the helical conformation was stable at lower temperatures without the addition of TFE, these conditions were used for subsequent NMR analysis.

Molecular Mass Analysis. The apparent molecular mass of SPAM-H1 in aqueous solution was measured by sedimentation equilibrium under a variety of conditions (described in Materials and Methods), including those of the NMR experiments used for structure determination. The average molecular mass observed for SPAM-H1 was 4280 ± 200 Da for all experiments and 4050 ± 150 Da in 90% ^1H_2O/10% ^2H_2O at pH 5.0 and 277 K. This is in very good agreement with the theoretical value of 4064.4 Da and shows that the molecule is monomeric under the conditions used for structure determination.

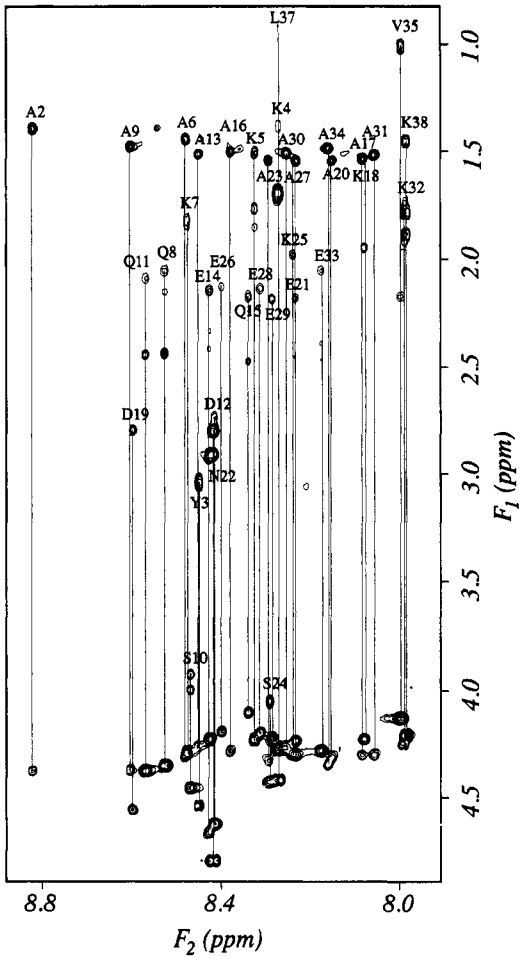

FIGURE 2: Fingerprint region of a 54 ms spin−lock time TOCSY spectrum of 4.4 mM SPAM-H1 at pH 5.5 in 90% ^1H_2O/10% ^2H_2O at 281 K. The vertical lines connect the resonances in each spin system. The spin systems are labeled with the standard single-letter abbreviation for the residue type and a number representing the position of the residue in the sequence.

NMR Resonance Assignments. Sequence-specific assignments were made using standard procedures (Wüthrich, 1986). Ambiguities were resolved by analyzing data from TOCSY and NOESY spectra under a variety of pH and temperature conditions, as follows: pH 2.5, 4.9, 5.5, and 6.1 at temperatures in the range 275−308 K. The assignments listed in Table 1 are for pH 5.5 and 281 K. Figure 2 shows the fingerprint region of the TOCSY acquired under these conditions, while the NH−C$^\alpha$H and NH−NH regions of the NOESY spectrum under these conditions are shown in Figure 3.

No stereospecific assignments were possible for nondegenerate geminal C$^\beta$H protons, but the $^3J_{C\alpha HC\beta H}$ coupling constants and intensities of intraresidue $d_{\alpha\beta}$ and $d_{N\beta}$ NOE values gave some information on the degree of side-chain conformational averaging. The expected $^3J_{C\alpha HC\beta H}$ values for staggered conformations are either both 3.4 Hz ($\chi 1 = 60°$) or 12.9 and 3.4 Hz ($\chi 1 = 180°$ or $-60°$) (Smith et al., 1991). Slight fluctuations around the $\chi 1$ angle will reduce the large coupling constants and increase the low ones (Hoch et al., 1985). Nondegenerate C$^\beta$H protons were observed for nine

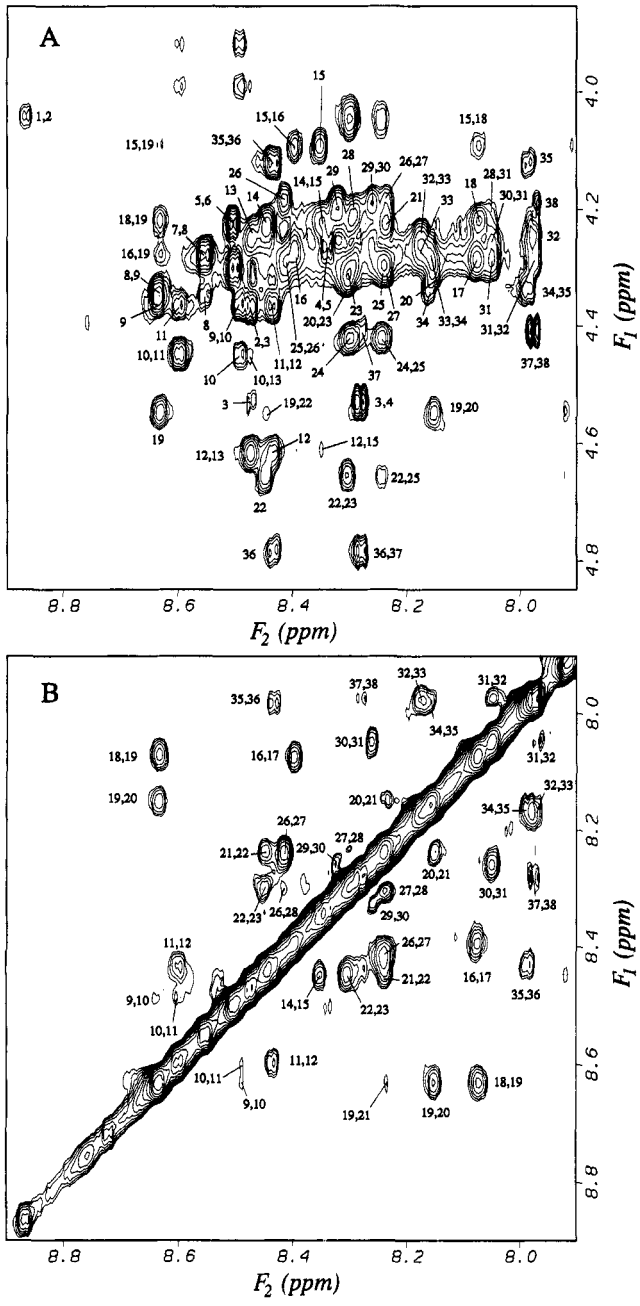

FIGURE 3: Regions of the 200 ms mixing time NOESY spectrum of SPAM-H1 at pH 5.5 in 90% ^1H_2O/10% ^2H_2O and 281 K. A: $C^\alpha H-NH$ region. Intraresidue cross-peaks are labeled with a single number. Sequential and medium range cross-peaks are labeled with two numbers indicating the residue contributing the $C^\alpha H$ and NH protons, respectively. B: NH−NH region. NH−NH cross-peaks are labeled with two numbers identifying the sequence position of the interacting residues.

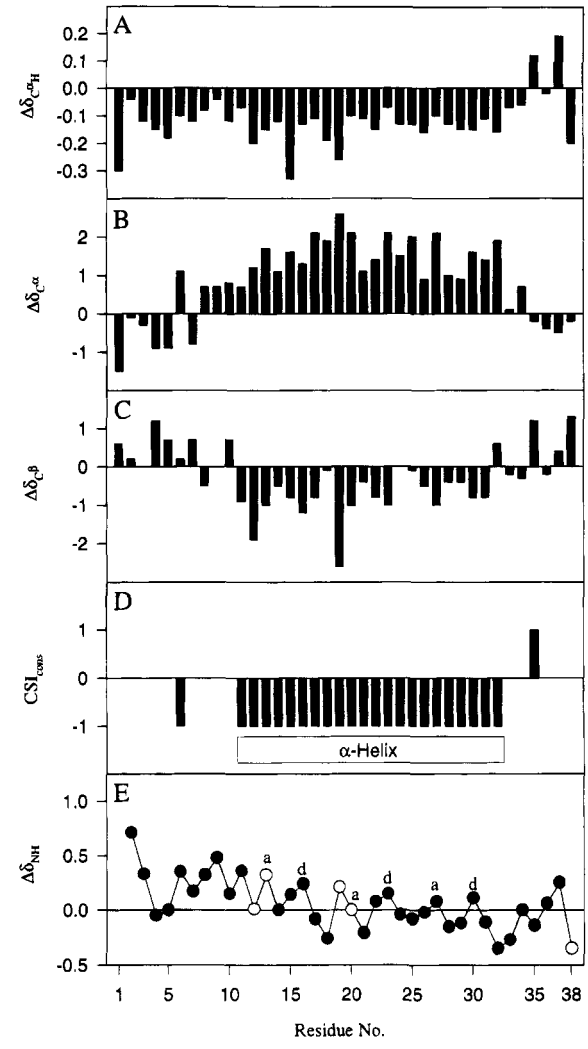

FIGURE 4: Chemical shift analyses for SPAM-H1 at pH 5.5 and 275 K. ^1H chemical shifts were measured from 2D homonuclear spectra and ^{13}C chemical shifts from a HMQC spectrum measured at natural abundance. In parts A−C and E the deviations from random coil shifts of the ^1H or ^{13}C chemical shifts of SPAM-H1 are plotted as a function of residue number. Random coil ^1H and ^{13}C values were taken, respectively, from Wishart et al. (1992) and Wishart and Sykes (1994). ^1H chemical shifts were referenced to DSS (0 ppm) and ^{13}C chemical shifts to *p*-dioxane (69.4 ppm). A: $\Delta\delta_{C\alpha H}$. B: $\Delta\delta_{C\alpha}$. C: $\Delta\delta_{C\beta}$. D: Chemical shift index consensus (CSI_{cons}). Individual $C^\alpha H$, C^α, and C^β chemical shift indices (Wishart et al., 1992; Wishart & Sykes, 1994) were calculated from the chemical shift deviations displayed in A−C. CSI_{cons} was assigned −1 when at least two of the three indices per residue indicated α-helix and, where applicable, the third index was not contradictory. CSI_{cons} was assigned 1 when at least two indices per residue indicated β-sheet and, where applicable, the third index was not contradictory. In all other cases CSI_{cons} was set to 0. Secondary structure termination points (at either end) are indicated by either the occurrence of an index of the opposite sign or two consecutive indices of 0. The helical residues based on this analysis (11−32) are indicated. E: $\Delta\delta_{NH}$. NH protons influenced by local functional group ionization (Asp12, Ala13, Asp19, Ala20, and Lys38) are indicated by an open circle. Residues at heptad a and d positions within the helix are labeled.

residues at pH 4.9 and 275 K, and $^3J_{C\alpha HC\beta H}$ values were obtained for four of these, *viz.* Tyr3 (8.0, 8.0), Ser10 (6.7, 4.6), Gln11 (6.1, 8.9), and Asn36 (5.9, 9.0), where the numbers in parentheses are $^3J_{C\alpha HC\beta H}$ in hertz and the first value is for the lower field $C^\beta H$ resonance. Although values could not be obtained for the other five residues (Gln8, Gln15, Asp19, Glu33, and Lys38) due to overlap, splitting patterns in the DQF-COSY spectrum indicated that for Gln8 and Gln15 the coupling constants to the lower field $C^\beta H$ resonances were smaller than those to the upfield resonances. From the $^3J_{C\alpha HC\beta H}$ coupling constants and the relative intensities of intraresidue $d_{\alpha\beta}$ and $d_{N\beta}$ NOEs for Tyr3, Gln8, Ser10, Gln11, Gln15, and Asn36 measured in a 50 ms mixing time NOESY spectrum, it appears that these six residues are undergoing internal rotation about the $C^\alpha-C^\beta$ bond.

Chemical Shift Analyses. The $C^\alpha H$, C^α, and C^β chemical shift values were compared with the values for residues in random coil peptides (Wishart et al., 1992), as shown in Figure 4. A chemical shift index analysis, carried out to

FIGURE 5: Summary of NMR data for SPAM-H1 at pH 4.9 and 5.5 and 275−281 K. The intensities of $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ connectivities are represented by the thickness of the respective line as strong, medium, or weak. An asterisk indicates that the presence of an NOE could not be confirmed unambiguously due to overlap. Values of $^3J_{NHC\alpha H}$ are those measured at 281 K, pH 5.5, and are indicated by ↓ if <6 Hz, ↑ if > 8 Hz, and a double-headed vertical arrow if between 6 and 8 Hz; blanks indicate that a value could not be measured due to overlap. The relative exchange rates of backbone NH protons are indicated in the row labeled NH, based on the strength of cross-peaks in ^2H_2O exchange TOCSY experiments at pH 4.8−5.0. Slowly exchanging amides are indicated by filled circles; open circles indicate intermediate exchange. Unambiguous medium range NOEs are indicated by solid lines connecting the two residue positions. Ambiguous medium range NOEs are indicated by dashed lines. Approximately one-half of the unambiguous medium range NOEs are summarized in this figure, the rest being to side chains. The only long range NOE observed was from $C^{\alpha}H$ of Ala31 to the side chain of Asn36.

determine a consensus estimate of secondary structure (Wishart & Sykes, 1994), showed that residues 11−32 adopted a helical structure (Figure 4D).

The deviations of backbone NH chemical shifts from random coil values are shown in Figure 4E. The NH protons showed a gradual upfield shift when going from the N- to C-terminus, consistent with the shielding effect of a helix macrodipole (Wishart et al., 1991). Superimposed on this gradual upfield shift was a periodic variation. Taking into account the NH shifts influenced by local functional group ionization, a three−four repeat pattern of local maxima that coincides with the a and d positions of the helix heptad pattern was observed. NH shift variation of this type has been observed for other isolated amphipathic α-helices (Zhou et al., 1992) and attributed to differences in i to $i + 4$ C=O to backbone NH hydrogen bond lengths, which are expected to be shorter on the apolar face of an amphipathic α-helix than on its polar face (Wishart et al., 1991). The fact that the three−four repeat pattern breaks down at residues 19−20 may be due to distortion of the helix or effects of the Asp19 side-chain carboxylate on the chemical shifts of both NH resonances (Bundi & Wüthrich, 1979).

Amide Exchange. The majority of the slowly exchanging backbone amide protons were found in the last two turns of the helix and the C-terminal tail of the peptide (residues 27−32, 34, 35, 37, and 38), as summarized in Figure 5. The NH of the C-terminal residue (Lys38) had the slowest exchange rate of all at pH 4.9−5.5 and 275 K and remained visible in 1D spectra recorded at 298 K and pH values up to 9.4, while none of the other amide protons was detectable at pH > 7.3. The amide exchange rates showed considerable sensitivity to pH; at pH 4.9 all of the NH resonances indicated in Figure 5 were observable up to 3 h after ^2H_2O exchange commenced, while at pH 5.0 the Lys38 NH was the only amide still present 1.5 h into the experiment, and at pH 5.5 it was the only NH proton visible after 10 min. The ^2H_2O exchange TOCSY experiments indicated that $H^{\delta 1}$ of Asn36 exchanged more slowly than $H^{\epsilon 1}$ of Gln8 and Gln11 ($H^{\delta 1}$ of Asn22 and $H^{\epsilon 1}$ of Gln15 were obscured by overlap), although still more quickly than the slowest of the backbone amides. A strong peak that could contain signals from both $H^{\delta 2}$ of Asn36 and $H^{\epsilon 2}$ of Gln8 was also detected. The peaks for $H^{\delta 2}$ of Asn22 and $H^{\epsilon 2}$ of Gln11 were both very weak. These observations suggest that $H^{\delta 1}$ (and possibly $H^{\delta 2}$) of Asn36 are hydrogen bonded. Due to extensive overlap there were too few distance constraints to define the hydrogen bond partner of Asn36 in the calculated structures (see below), but a weak cross-peak was detected between $H^{\delta 1}$ of Asn36 and $C^{\alpha}H$ of Ala31.

Structure Determination. A summary of the sequential and medium range NOE connectivities, $^3J_{NHC\alpha H}$ coupling constants, and slowly exchanging backbone amide protons is presented in Figure 5. The presence of an α-helix extending from residue 10 to at least residue 32 is indicated by the low $^3J_{NHC\alpha H}$ coupling constants (<5 Hz), the numerous $d_{\alpha N}(i,i + 3)$, $d_{\alpha N}(i,i + 4)$, and $d_{\alpha\beta}(i,i + 3)$ NOEs, and the locations of slowly exchanging backbone amide protons. The low $^3J_{NHC\alpha H}$ values indicate that the helical structure does not undergo extensive conformational averaging at temperatures below 280 K (Dyson & Wright, 1991). The ultracentrifugation studies demonstrate that SPAM-H1 does not aggregate under the conditions of the NMR experiments and thus that the observed NOEs are intramolecular and do not reflect intermolecular contacts or the conformational preference of some higher molecular weight oligomer. Distance constraints were calculated from NOESY spectra and dihedral angle constraints were generated from $^3J_{NHC\alpha H}$ coupling constants as described under Materials and Methods. Structures were determined using distance geometry calculations and refined by restrained simulated annealing and energy minimization.

Structural Analysis and Description. Analysis of the angular order parameters (S) (Hyberts et al., 1992; Pallaghy et al., 1993) of the 20 final structures indicates that residues 11−31 and 33−34 are well defined, with $S > 0.9$ for both ϕ and ψ angles (except for the ψ of residue 31, which has a value of 0.8). A Ramachandran plot (Figure 6) places the average backbone dihedral angles of these residues in sterically allowed regions. Several residues (Ala6, Lys7, Ala9, Asn36, and Lys38) have well-defined ϕ angles ($S >$

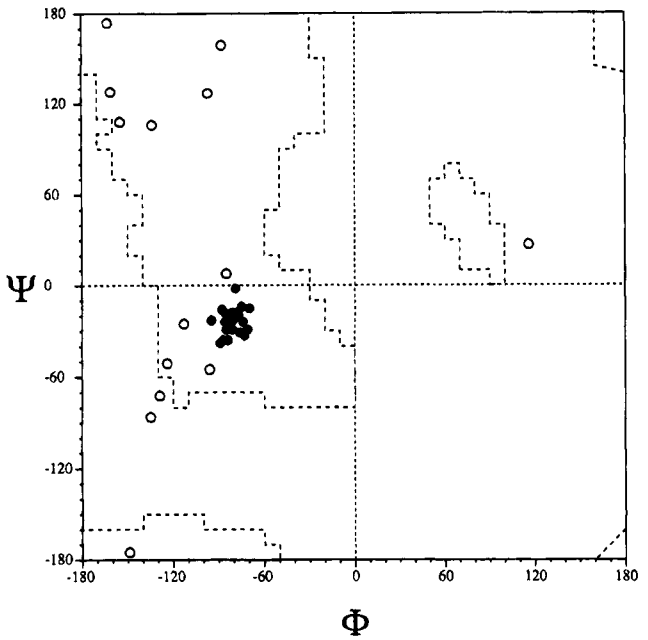

FIGURE 6: Ramachandran plot for the angular average of the 20 structures of SPAM-H1. Bold symbols denote residues with well-defined backbone dihedral angles (S > 0.9).

0.9) but poorly defined ψ angles. Most of the residues with well-defined ϕ angles had their ϕ angles constrained in the structure calculations. Residues Val35 and Leu37 have well-defined ψ angles. The average ϕ angle of residue 37 is positive, but it is not well defined (ϕ 115°, ψ 29°; S_ϕ 0.60, S_ψ 0.94). Over the well-defined region indicated by the angular order parameters, the rms differences (mean ± SD) from the average structure for the backbone heavy atoms N, C^α, and C and all heavy atoms, respectively, were 1.56 ± 0.70 and 2.12 ± 0.71 Å. These structures had no NOE restraint violations >0.2 Å, no dihedral angle violations >3°, an rms deviation for bond lengths of 0.014 ± 0.0005 Å, and an rms deviation for bond angles of 2.82° ± 0.14°. The energy associated with the NOE restraints was 2.22 ± 0.86 kcal mol^{-1} and the total energy −498 ± 39 kcal mol^{-1}.

The overall conformation of SPAM-H1 is evident in Figure 7, where the backbone atoms of the best 20 structures are superimposed. The molecule is clearly α-helical from residue 10 to at least residue 32. While the ϕ and ψ angles of Lys32 are not well defined by the NMR data, the following evidence suggests that the helix extends up to Val35: (i) Glu33 and Ala34 have well-defined helical backbone geometry; (ii) a $d_{\alpha N}(i,i+4)$ NOE was observed between Glu29 and Glu33 and a $d_{\alpha N}(i,i+3)$ between Lys32 and Val35; (iii) there is a stretch of strong, uninterrupted d_{NN} connectivities up to and including Val35 (Figure 5). Many of the structures also contain turn-like conformations between Val35 and Lys38, but the NOEs, coupling constants, and solvent-protected NH protons do not fit the descriptions for any of the standard tight turns (Richardson, 1981).

In the structures shown in Figure 7 the helix is bent such that the face formed by the polar residues is concave. Upon closer examination, it appears that this apparent curvature is due at least in part to a kink in the region of residues 19−20, with the ψ angle of Asp19 being −1° (S 0.97). Indeed, if the regions of helix on either side of this region are superimposed, significantly lower rms differences for the backbone heavy atoms N, C^α, and C are obtained, as follows: residues 11−19, 0.39 ± 0.14 Å, and residues 20−31, 0.97 ± 0.29 Å.

DISCUSSION

The polypeptide SPAM-H1 forms a stable helical structure in aqueous solution without the need for addition of TFE and is monomeric under a range of conditions. It appears that the presence of alanine residues at the a and d positions of four heptad repeats does not provide a sufficiently large hydrophobic surface to drive a coiled-coil-like association. Zhou et al. (1992, 1993) found that an 18-residue synthetic peptide with exclusively alanine at what could be considered the a, d, and e positions of a series of heptad repeats was monomeric, while an analogous peptide with exclusively leucine at these positions aggregated in aqueous solution.

Another factor contributing to the monomeric nature of SPAM-H1 is the lack of suitable residues to form interhelical salt bridges. Coiled-coil structures are generally stabilized by such salt bridges (O'Shea et al., 1991; Hu et al., 1993; Lovejoy et al., 1993; Harbury et al., 1993; Monera et al., 1994; Zhou et al., 1994), usually between the side chains of residues in the e and g positions. There is no possibility of this type of interaction occurring in SPAM-H1 dimers, trimers, or tetramers as the e positions are occupied by either Ser or Ala and the g positions by Glu or Asp (Figure 8). The only potential interhelical salt bridge could come from a g−b pairing of Asp12 with Lys7. A g−b salt bridge was found in the up−up−down, three-stranded coiled coil formed by the coil-Ser peptide (Lovejoy et al., 1993), and four g−b salt bridges occurred in the parallel, four-stranded coiled coil formed by p-LI, a mutant of the GCN4 sequence (Harbury et al., 1993), but in both cases the g−b interactions were accompanied by e−g interactions.

Structure of the Complete Antigen. We have demonstrated here that the H1 heptad repeat of SPAM adopts a relatively stable helical structure in solution. Taking into account the sequence similarities between the H1, H2, and H3 blocks (Figure 1), it is likely that H2 and H3 can also form helices. These helical units could then associate in several different ways to give coiled-coil structures. In addition to the highly conserved Ala residues at the a and d positions, the residues in the e positions of all three heptads tend to be apolar, providing a broad apolar surface on each helix made up of the side chains from the a, d, and e residues. In H1 the residues at the e position are small (Ala and Ser), while in H2 and H3 they tend to be residues with larger side chains (Val, Leu, and Tyr).

Our experience with SPAM-H1 in solution suggests that interhelical salt bridges will play a significant role in directing self-association. A further consideration is that there is a Cys residue near the start of the H2 block (Figure 1) which is conserved in all variants of SPAM sequenced to date (McColl et al., 1994), but there are no other Cys residues in the protein. As SPAM has been detected in the parasitophorous vacuole space in infected erythrocytes, which is an oxidizing environment, it is unlikely that this Cys exists in the free thiol state. Thus, this residue may mediate either the dimerization of two SPAM molecules or the association with another protein through a disulfide linkage.

We consider first the structure of a disulfide-linked coiled coil between two SPAM molecules. A parallel arrangement seems unlikely because of the relatively small number of

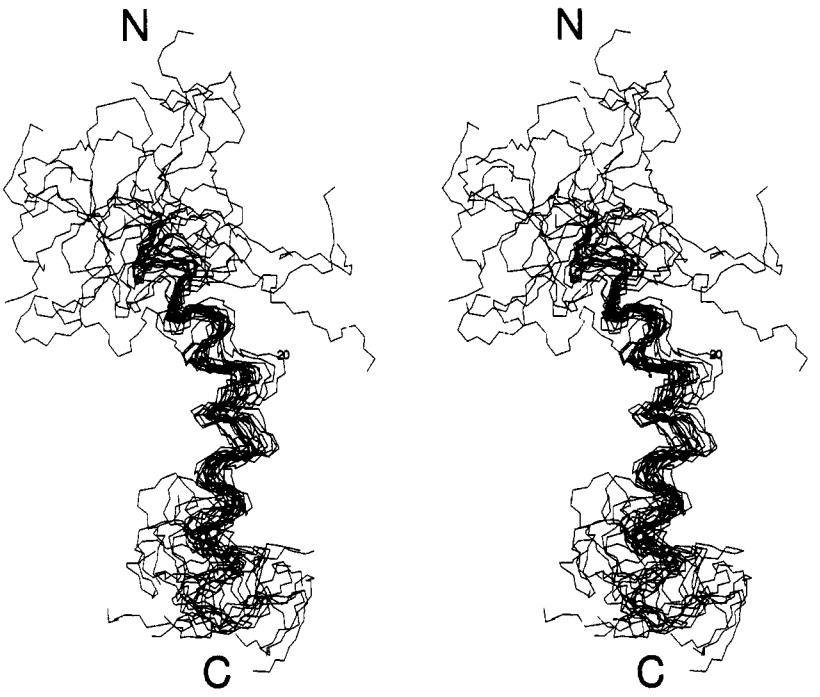

FIGURE 7: Stereoview of the 20 structures with the lowest energies after refinement and energy minimization, superimposed on the backbone heavy atoms of residues 11−31 of the average coordinates (not shown). Only the backbone atoms are shown.

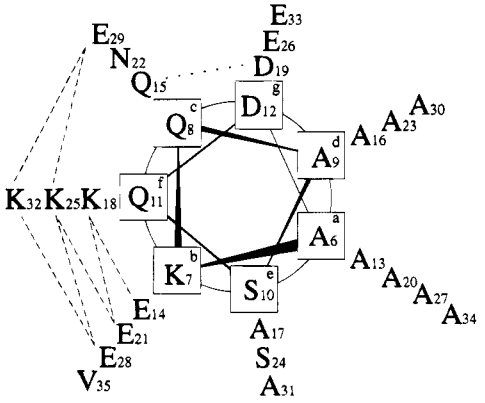

FIGURE 8: Helical wheel projection for the four heptad repeats in SPAM-H1 (residues 6−33) and residues 34 and 35. Heptad positions are labeled with a lowercase letter. Potential intrahelical (i to $i + 3$ or 4) salt bridges are shown with dashed lines. A potential Gln15→Asp19 hydrogen bond is shown as a dotted line, while a potential Gln8→Asp12 hydrogen bond is omitted because Gln8 is not in the helix.

interhelical salt bridges that would be formed, and we have already found that the isolated H1 repeat units have little affinity for one another. An antiparallel arrangement would necessarily be staggered because the Cys residue on the H2 repeat is located near its N-terminal end, but the intervening loops between H1, H2, and H3 may be flexible enough to allow the H1 and H3 units to interact, in which case a larger number of salt bridges could form. Thus, if a homodimeric, disulfide-linked coiled coil forms, it is likely to be antiparallel.

The other likely structure for the intact SPAM molecule is an intramolecular coiled coil, which may then form a disulfide link with SPAM or another protein. The overall topology of the three-helical bundle would be expected to be dictated by the length of the H1−H2 and H2−H3 loops and the nature of salt bridges involving residues at the b and g positions. The H1−H2 loop could be no longer than 16 residues, which would allow both parallel and antiparallel orientations. However, the presence of additional conserved Ala residues at the end of the H1 block and the start of the H2 block could extend the helices and thus shorten the intervening loop to 12 residues or less (an extended H1 helix in SPAM is supported by our findings for the C-terminal region of the H1 peptide), in which case an antiparallel topology is more likely. The H2−H3 loop is longer, having at least 23 residues in different SPAM isolates (unpublished results), and a parallel orientation of the H2 and H3 helices is therefore possible. Figure 9 shows a schematic representation of H1, H2, and H3 from the D10 sequence in a left-handed, up−down−down coiled coil. In a simple knobs-in-holes packing arrangement, the Ala residues at the a, d, and d positions in H1, H2, and H3, respectively, would form one layer at the apolar interface and the d, a, and a positions in H1, H2, and H3, respectively, the other layer. Because the Ala side chains are smaller than the Leu and Ile side chains commonly found in coiled-coil structures (Lovejoy et al., 1993; Harbury et al., 1993, 1994), closer packing of the strands may be possible, allowing salt bridge formation between the b positions in H1 and H2 (Figure 9). Equivalent salt bridges are not observed in other coiled coils, presumably because these side chains are too far apart.

All possible topologies for SPAM were aligned to maximize the number of potential interhelical salt bridges, and the up−down−down topology was found to give the greatest number. If the H3 helix were inverted to give an up−down−up topology, the H3 helix would make only one less potential salt bridge, and this arrangement could be favored by other parameters such as packing considerations at the apolar interface.

Another potential packing arrangement could be envisaged in which residues at the a, d, and e positions *all* contributed to the apolar interface, with residues at the a position aligned in a staggered fashion down the central axis of the bundle. In terms of the coiled-coil parametrization of Crick (1953), the angle (ϕ) between R_1 (the α-helix radius to the a position) and R_0 (the supercoil radius) would be 0°. This type of

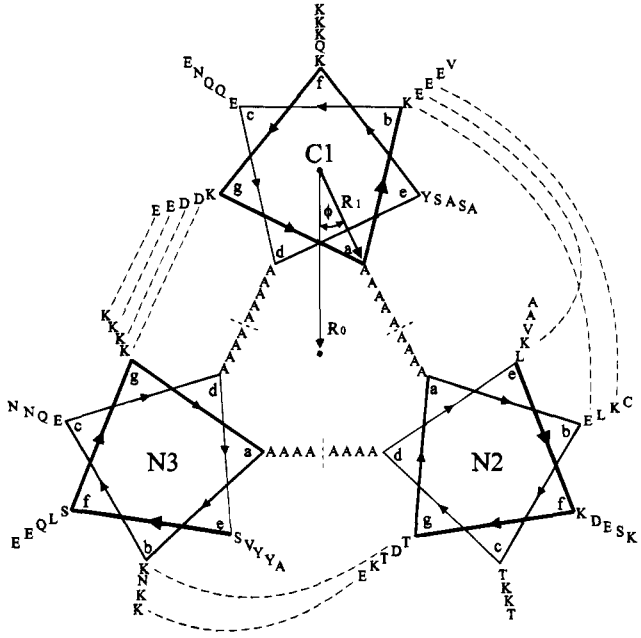

FIGURE 9: Helical wheel projection for a putative triple-stranded coiled coil of SPAM in a left-handed, up−down−down configuration. The heptad positions are labeled a−g, with the outermost residues closest to the viewer. ϕ is the angle between vectors R_1, to C^α of the a position, and R_0, to the axis of the coiled coil (Crick, 1953); in the illustrated structure it is 26°. Potential interhelical salt bridges are indicated by dashed lines. The first helix is oriented with its C-terminus closest to the viewer, and the second and third are oriented with their N-termini closest. The first helix is shown as residues 90−124, the second as residues 136−166, and the third as residues 187−216 of the D10 sequence (McColl et al., 1994).

arrangement brings oppositely charged residues in the g and b positions of all three heptads closer together than in typical coiled-coil trimers (Figure 9) but might be disfavored by close contact between residues at the e positions of H1 and H2, where some larger side chains (Val, Leu, Lys, and Tyr) are found.

Helix Bending. Curvature is a common feature of helices in globular and fibrous proteins (Barlow & Thornton, 1988; Cohen & Parry, 1990) and is associated with shorter i to $i + 4$ hydrogen bonds on the hydrophobic face than on the hydrophilic face (Blundell et al., 1983; Baker & Hubbard, 1984). Kinks also occur in helices at proline residues or because of side-chain packing considerations (Barlow & Thornton, 1988). Figure 10 shows the structure closest to the average of the ensemble in Figure 7. The structure is bent, causing the face formed by the polar residues to be concave. One reason for this is the presence of a kink around residues 19−20, but there is also some curvature of the helical segments 11−19 and 20−32. We believe that the presence of the kink and the degree of curvature in the SPAM-H1 structures reflect to some extent the nature of the NMR restraints used to determine the structures, as the alanine residues show considerable overlap in their $C^\beta H$ chemical shifts (Table 1), making it impossible to determine accurate volumes for NOESY cross-peaks involving them in most cases. As a result, default volumes equivalent to distances of 5.0 Å were assigned, and when pseudoatom corrections and the 1.0 Å added to all side-chain NOEs were included, the resulting upper bounds were 7.0 Å. By contrast, the residues away from the apolar face showed considerably more chemical shift dispersion, thus reducing overlap and allowing their NOESY volumes to be measured more accurately. Upper bound distance constraints for NOEs between these residues, therefore, were generally lower than those on the apolar face. Using model structures, we have investigated the possibility of curvature due to data bias (unpublished experiments). By introducing a systematic uncertainty to the data (e.g., simulating residue-type-based overlap), curved helices can be obtained from distance constraints that would otherwise describe linear structures. Thus, our data should be construed as defining an α-helical structure for SPAM-H1 but not necessarily one that is curved away from the apolar (Ala-containing) face. Similarly, the kink around residues 19−20 is influenced by side-chain to side-chain NOEs between Lys18 and Asn22; while these NOEs have been measured accurately and are satisfied in the final structures, the lack of "counterbalancing" NOEs involving residues on the opposite face of the helix may cause the kink to be more pronounced in the calculated structures than in solution.

Because alanine is only weakly hydrophobic, an amphipathic α-helix such as SPAM-H1 would be expected to show very little curvature (Zhou et al., 1992). This is supported by the pattern of NH shift variation (Figure 4E). Similar NH shift variations have been seen in other isolated amphipathic α-helices with hydrophobic residues at the a, d, and e positions of the heptad repeats (Zhou et al., 1992). When the a, d, and e positions were occupied by Leu, there was a clear four−three repeat pattern, with significant downfield displacements of the backbone amide resonances from the a and e positions. These positions are at the center of the apolar face, and the amide is hydrogen bonded to the carbonyl of another Leu. The difference between the $\Delta\delta_{NH}$ values on the apolar and polar faces of about 1 ppm was attributed to differences in the i to $i + 4$ carbonyl−NH hydrogen bond lengths between the faces. When the a, d, and e positions were occupied by Ala, however, the variation between the faces was considerably less (~0.2 ppm) and the maxima no longer coincided with just the a and e positions, the shifts for the a, d, and e positions being similar. This was interpreted as showing that the Ala-substituted peptide did not have hydrogen bond length distortions associated with helical curvature (Zhou et al., 1992). The pattern and magnitude of variation of $\Delta\delta_{NH}$ values in SPAM-H1 are very similar to those in the Ala-subsituted peptide, suggesting that SPAM-H1 would not be curved to any great extent.

The breakdown of the three−four repeat pattern in the $\Delta\delta_{NH}$ pattern at residues 19 and 20 (Figure 4E) lends support to the notion of a kink at this position. However, the NMR description of the conformation around residue 19 is strongly dependent on the accuracy of distance restraints derived from NOEs involving Ala16, Ala17, Ala20, and Ala23, so the angle between the two helical segments observed in the structures presented here should not be regarded as accurately defined.

C-Terminal Capping. The pattern of slowly exchanging amide protons in SPAM-H1 indicates that the C-terminus of the helix is more stable than its N-terminus. Helices can be stabilized by intrahelical salt bridges involving oppositely charged side chains positioned three or four residues apart (Gans et al., 1991) and by hydrogen bonds between Gln and carboxyl side chains in an i to $i + 4$ orientation. These hydrogen bonds can be as strong as the hydrogen-bonding component of i to $i + 3$ or $i + 4$ salt bridges (Scholtz et al.,

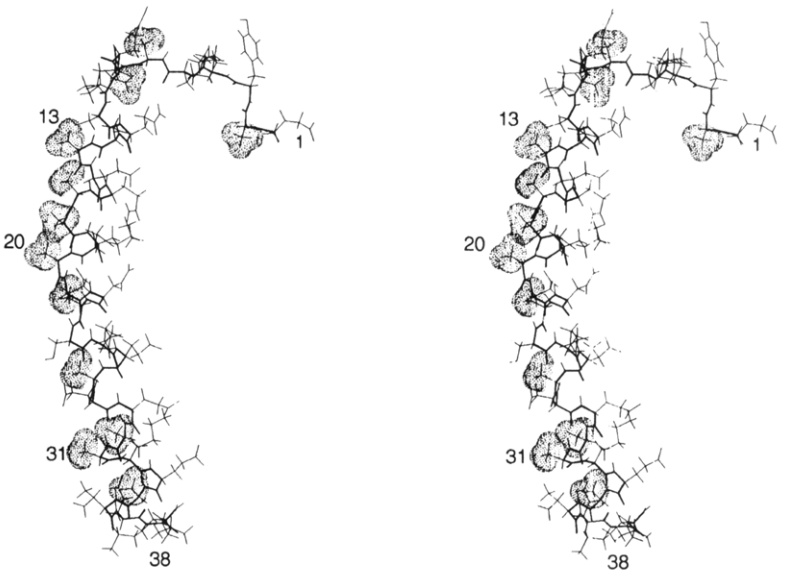

FIGURE 10: Stereoview of the structure closest to the average coordinates of the 20 structures with the lowest energies after refinement and energy minimization. The backbone is shown in bold, and the van der Waals surfaces of the Ala methyl groups are shaded.

1993). Figure 8 shows that these potential stabilizing interactions are relatively evenly spaced along the SPAM-H1 helix and thus would be expected to stabilize it along its entire length, not just in the last two turns. Interactions with the helix dipole may play a role. The nine residues at the N-terminus of the peptide are unstructured, so there is no negatively charged residue in a suitable position to interact with the helix dipole, but at the C-terminus, the helix finishes closer to the end of the peptide and some interaction with Lys38 may occur. His (Armstrong & Balwin, 1993) and Arg (Fiori et al., 1994) near the C-termini of helices have been found to exert stabilizing influences. Lys38 may be contributing to the stability of the helix by electrostatic interactions with the dipole, but it is difficult to imagine how charge−helix dipole interactions alone could be responsible for the slow exchange of the NH protons of residues 37 and 38, which are not part of the helix. It seems more likely that the slow exchange of amide protons in the last turns of the helix and C-terminal to the helix is due to the presence of some sort of C-capping motif, which may well include interactions with the side chain of Lys38.

The sequence determinants of two C-capping motifs have been described recently (Aurora et al., 1994). These C-capping motifs, the α_L motif and the Schellman motif, are both characterized by a positive ϕ angle at the C′ position. The α_L motif cannot be present in SPAM-H1 as Gly is the only residue sterically allowed at the C′ position of this motif (Creamer et al., 1995). In the Schellman motif the positive ϕ angle is typically provided by a Gly residue, although in some cases Asn or Lys has been found. In this motif hydrogen bonds occur between the NH of C′ and the carbonyl of C2 and between the NH of C″ and the carbonyl of C3, with the latter hydrogen bond generally being accompanied by a hydrophobic interaction between their side chains. Kallenbach et al. (1994) have reported a C-capping interaction involving a hydrogen bond between the side chain of an Asn at the C-cap and the carbonyl of a residue at position C4.

The position of the C-cap in SPAM-H1 is not clearly defined in the calculated structures. As discussed above, it is likely that the helix extends to Val35, but the $^3J_{NHC\alpha H}$ coupling constants and the $C^\alpha H$ and C^β chemical shifts of Val35 and subsequent residues indicate that the backbone is definitely nonhelical from this point on. Thus, while Asn36 cannot be the C-cap residue, it may occupy the C′ position if Val35 represents the C-cap residue. In proteins Asn is found quite often in this position (Rose & Wolfenden, 1993). While there were too few experimental constraints to define the C-terminal region well in our structures, the relatively slow exchange of the $N^\delta H$ protons of Asn36, compared with the other side-chain amide protons, the presence of a NOESY cross-peak from $N^{\delta 1}H$ of Asn36 to $C^\alpha H$ of Ala31, and the somewhat constrained nature of the $\chi 1$ angle of Asn36 all suggest that it may be involved in this type of capping interaction, most probably with the carbonyl of Lys32 (the residue at the C3 position). The slow exchange of the NH of Lys38, however, suggests that it is also involved in either a capping hydrogen bond or a hydrogen bond that stabilizes the conformation of the capping motif. It is interesting to note that the slow exchange of the Lys38 amide was clearly evident at 298 K, where little NMR evidence for helix formation was detected. If the slow exchange of Lys38 NH were due to a hydrogen bond that was an integral part of the C-capping motif, it may be that this hydrogen bond forms independently and provides an initiation point for helix folding.

Regardless of where the C-cap is in SPAM-H1, the slowly exchanging NH protons in the C-terminal section of the peptide cannot be fitted to the hydrogen-bonding pattern of the Schellman or α_L motifs and Asn36 in SPAM-H1 is not the C-cap residue. Therefore, it is also distinct from the C-cap of Kallenbach et al. (1994). Further experiments are underway to define the structural and sequence determinants of what appears to be a new class of C-capping motif.

Analysis of NOEs and chemical shifts places the N-cap at Ser10. However, there is no evidence from the NOEs to suggest that it is involved in a well-defined capping interaction similar to that in N-capping box motifs (Lyu et al., 1993; Zhou, H. X. et al., 1994). That the helix is less stable at the N-terminus may be because Ser10 is able to provide only transient hydrogen-bonding stabilization to the N1, N2, or N3 residues. Transient hydrogen bonding is supported by

the observation that there is not completely free rotation about the $C^{\alpha}-C^{\beta}$ bond of Ser10.

In conclusion, the heptad repeats in SPAM are the first example of a repetitive sequence motif from a merozoite surface antigen which can be directly correlated to a structural element. It remains to be seen how, if at all, the H1 heptad repeat interacts with the H2 and H3 repeats in the intact antigen and to define the role of the conserved cysteine residue in the structure. The SPAM-H1 peptide itself forms an isolated α-helix which is considerably more stable at its C-terminus than its N-terminus, apparently due to the presence of a specific C-capping motif. The hydrogen-bonding pattern in the capping interaction in SPAM-H1 appears different from those identified in previously documented C-capping motifs, and further investigations are required to determine the exact sequence and structural determinants of this motif.

ACKNOWLEDGMENT

We are grateful for the provision of software by David Smith (SSTRUC) and David Craik (COUPLING), for assistance with computing from Richard Ford and Lisa Cowen, and for many helpful discussions with Paul Pallaghy.

REFERENCES

Anders, R. F., & Saul, A. J. (1994) in *Molecular Immunological Considerations in Malaria Vaccine Development* (Good, M. F., & Saul, A. J., Eds.) pp 169−193, CRC Press, Boca Raton, FL.

Anders, R. F., McColl, D. J., & Coppel, R. L. (1993) *Acta Trop. 53*, 239−253.

Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.

Ardeshir, F., Flint, J. E., Richman, S. J., & Reese, R. T. (1987) *EMBO J. 6*, 493−499.

Armstrong, K. M., & Baldwin, R. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11337−11340.

Aurora, R., Srinivasan, R., & Rose, G. D. (1994) *Science 264*, 1126−1130.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.

Barlow, D. J., & Thornton, J. M. (1988) *J. Mol. Biol. 201*, 601−619.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Reson. 55*, 301−315.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Blundell, T., Barlow, D., Borkakoti, N., & Thornton, J. (1983) *Nature 306*, 281−283.

Borre, M. B., Dziegiel, M., Hogh, B., Petersen, E., Rieneck, K., Riley, E., Meis, J. F., Aikawa, M., Nakamura, K., Harada, M., Wind, A., Jakobsen, P. H., Cowland, J., Jepsen, S., Axelsen, N. H., & Vuust, J. (1991) *Mol. Biochem. Parasitol. 49*, 119−131.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Brünger, A. T. (1992) *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*, Yale University, New Haven, CT.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 299−311.

Camus, D., & Hadley, T. J. (1985) *Science 230*, 553−556.

Cohen, C., & Parry, D. A. D. (1990) *Proteins 7*, 1−15.

Creamer, T. P., Srinivasan, R., & Rose, G. D. (1995) in *Techniques in Protein Chemistry VI* (Crabb, J. W., Ed.) Academic, San Diego, CA (in press).

Crick, F. H. C. (1953) *Acta Crystallogr. 6*, 685−689.

Dyson, H. J., & Wright, P. E. (1991) *Annu. Rev. Biophys. Chem. 20*, 519−538.

Fiori, W. R., Lundberg, K. M., & Millhauser, G. L. (1994) *Nature Struct. Biol. 1*, 374−377.

Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W., & Kallenbach, N. R. (1991) *Biopolymers 31*, 1605−1614.

Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.

Güntert, P., & Wüthrich, K. (1991) *J. Magn. Reson. 96*, 403−407.

Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.

Harbury, P. B., Zhang, T., Kim, P. S., & Alber, T. (1993) *Science 262*, 1401−1407.

Harbury, P. B., Kim, P. S., & Alber, T. (1994) *Nature 371*, 80−83.

Harper, E., & Rose, G. D. (1993) *Biochemistry 32*, 7605−7609.

Hoch, J. C., Dobson, C. M., & Karplus, M. (1985) *Biochemistry 24*, 3831−3841.

Holder, A. A. (1988) *Prog. Allergy 41*, 72−97.

Hu, J. C., Newell, N. E., Tidor, B., & Sauer, R. T. (1993) *Protein Sci. 2*, 1072−1084.

Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.

Ji, H., Whitehead, R. H., Reid, G. E., Moritz, R. L., Ward, L. D., & Simpson, R. J. (1994) *Electrophoresis 15*, 391−405.

Kallenbach, N. R., Zhou, H., Lyu, P., Wemmer, D. E., Mayne, L., & Englander, S. W. (1994) *Protein Sci. 3* (Suppl. 1), 55.

Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., & Eisenberg, D. (1993) *Science 259*, 1288−1292.

Lupas, A., van Dyke, M., & Stock, J. (1991) *Science 252*, 1162−1164.

Lyu, P. C., Wemmer, D. E., Zhou, H. X., Pinker, R. J., & Kallenbach, N. R. (1993) *Biochemistry 32*, 421−425.

Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259−281.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.

McColl, D. J., Silva, A., Foley, M., Kun, J. F. J., Favaloro, J. M.,Thompson, J. K., Marshall, V. M., Coppel, R. L., Kemp, D. J.,& Anders, R. F. (1994) *Mol. Biochem. Parasitol. 68*, 53−67.

Monera, O. D., Kay, C. M., & Hodges, R. S. (1994) *Biochemistry 33*, 3862−3871.

Montelione, G. T., Arnold, E., Meinwald, Y. C., Stimson, E. R., Denton, J. B., Huang, S.-G., Clardy, J., & Scheraga, H. A. (1984) *J. Am. Chem. Soc. 106*, 7946−7958.

Morton, C. J., Simpson, R. J., & Norton, R. S. (1994) *Eur. J. Biochem. 219*, 97−107.

Mouw, A. R., Beachey, E. H., & Burdett, V. (1988) *J. Bacteriol. 170*, 676−684.

Nelson, J. W., & Kallenbach, N. R. (1986) *Proteins 1*, 211−217.

O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539−544.

Pallaghy, P. K., Duggan, B. M., Pennington, M. W., & Norton, R. S. (1993) *J. Mol. Biol. 234*, 405−420.

Peterson, M. G., Marshall, V. M., Smythe, J. A., Crewther, P. E., Lew, A., Silva, A., Anders, R. F., & Kemp, D. J. (1989) *Mol. Cell. Biol. 9*, 3151−3154.

Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632−1640.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.

Reid, G. E., & Simpson, R. J. (1992) *Anal. Biochem. 200*, 301−309.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−335.

Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648−1652.

Rose, G. D., & Wolfenden, R. (1993) *Annu. Rev. Biophys. Biomol. Struct. 22*, 381−415.

Rucker, S. P., & Shaka, A. J. (1989) *Mol. Phys. 68*, 509−517.

Scholtz, J. M., Qian, H., Robbins, V. H., & Baldwin, R. L. (1993) *Biochemistry 32*, 9668−9676.

Smith, L. J., Sutcliffe, M. J., Redfield, C., & Dobson, C. M. (1991) *Biochemistry 30*, 986−996.

Smythe, J. A., Coppel, R. L., Brown, G. V., Ramasamy, R., Kemp, D. J., & Anders, R. F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5195−5199.

Steven, A. C., Trus, B. L., Maizel, J. V., Unser, M., Parry, D. A. D., Wall, J. S., Hainfeld, J. F., & Studier, F. W. (1988) *J. Mol. Biol. 200,* 351−365.
Szyperski, T., Güntert, P., Otting, G., & Wüthrich, K. (1992) *J. Magn. Reson. 99*, 552−560.
van Geet, A. L. (1970) *Anal. Chem. 42,* 679−680.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Gõ, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611−639.
Wishart, D. S., & Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169,* 949−961.
Zhou, H. X., Lyu, P. C., Wemmer, D. E., & Kallenbach, N. R. (1994) *Proteins 18*, 1−7.
Zhou, N. E., Zhu, B.-Y., Sykes, B. D., & Hodges, R. S. (1992) *J. Am. Chem. Soc. 114*, 4320−4326.
Zhou, N. E., Kay, C. M., Sykes, B. D., & Hodges, R. S. (1993) *Biochemistry 32*, 6190−6197.
Zhou, N. E., Kay, C. M., & Hodges, R. S. (1994) *J. Mol. Biol. 237*, 500−512.

BI9425260