Determination of the Three-Dimensional Solution Structure of Ragweed Allergen *Amb t* V by Nuclear Magnetic Resonance Spectroscopy[†,‡]

William J. Metzler,[*,§] Kathleen Valentine,[∥] Marianne Roebber,[⊥] Mark S. Friedrichs,[§] David G. Marsh,[⊥] and Luciano Mueller[§]

Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, New Jersey 08543-4000, Department of Chemistry, Princeton University, Princeton, New Jersey 08544, and Johns Hopkins Asthma and Allergy Center, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21224

Received December 3, 1991; Revised Manuscript Received March 24, 1992

ABSTRACT: Analysis of two-dimensional NMR experiments has afforded essentially complete assignment of all proton resonances in the allergenic protein *Amb t* V. Conformational constraints were obtained from the NMR data in three forms: interproton distances derived from NOE cross-peak intensities of NOESY spectra, torsion angle constraints derived from *J*-coupling constants of COSY and PE-COSY spectra, and hydrogen bond constraints derived from hydrogen-exchange experiments. Conformations of *Amb t* V with low constraint violations were generated using dynamic simulated annealing in the program XPLOR. The refined structures are comprised of a C-terminal α-helix, a short stretch of triple-stranded antiparallel β-sheet, and several loops. In addition, the cystine partners of the four disulfide linkages (for which there are no biochemical data) have been assigned. The refined structures of *Amb t* V will allow us to suggest surface substructures for the *Amb* V allergens that are likely to participate in B cell epitopes and will assist us in defining the Ia/T cell epitopes that interact with the MHC class II (or Ia) molecule and the T cell receptor leading to the induction of the immune response to *Amb t* V.

Atopic allergy toward inhaled allergens provides an ideal model for studying the genetic and molecular basis of human immune responsiveness (Marsh, 1975; Marsh et al., 1990). This model is favored by the ultralimiting low-dose conditions of immunization that normally prevail during natural exposure to most inhaled allergens. Under such conditions, it has been suggested that only certain immunodominant epitopes are recognized at the B cell (antibody) and T cell levels (Marsh, 1975). Immunodominant B cell epitopes tend to be primarily surface structures on native protein molecules (Benjamin et al., 1984), i.e., portions of the molecule that have been subject to rapid evolutionary change (Richardson, 1981). As such, these structures are often species-specific. "Ia/T cell" epitopes are usually peptide fragments of the antigen that are displayed on the antigen-presenting cell (APC) within a groove of the MHC class II (or Ia) molecule (Bjorkman et al., 1987; Brown, J. H., et al., 1988) in sufficient quantity to permit recognition by a specific receptor of a helper T cell, leading to a specific immune response (Unanue & Allen, 1987; Sette et al., 1987). Recent crystallographic data of a related MHC class I molecule have shown that peptides are bound in an extended conformation (Madden et al., 1991). The nature and location of the immunodominant Ia/T cell epitopes within the *native* antigen molecule are less clear. On the basis of data primarily from mouse experiments, Margalit et al. (1987) proposed that amphipathic structures, usually amphipathic α-helices, may be the preferred conformation of such epitopes. On the basis of the allergy model, Marsh (1986) postulated that immunodominant Ia/T cell epitopes may be relatively conserved regions of the protein molecule. These regions would be similar to the rigid cores of protein molecules that are often found in β-pleated sheets and/or disulfide-bonded structures (Marsh, 1986).

The *Ambrosia* (ragweed) pollen allergens (*Amb* V, formerly Ra5) constitute particularly good models in which to test the structure–function relationships for both B cell and Ia/T cell epitopes in the human immune response. *Amb a* V and *Amb t* V, isolated from the two most commonly occurring species of ragweed, *Ambrosia artemisiifolia* (short) and *Ambrosia trifida* (giant), are single-chain polypeptides of 45 and 40 amino acids, respectively (Lapkoff & Goodfriend, 1974; Mole et al., 1975; Goodfriend et al., 1985; Roebber et al., 1985). These extremely heat-stable molecules each contain four disulfide bonds. The two allergens are known to be virtually non-cross-reactive at the B cell level, suggesting a dissimilarity in their surface structures (Roebber et al., 1985). However, there is concordance in the immune recognition of each allergen by the same HLA-DR2/Dw2-associated class II molecule, HLA-DR(α,β_1*1501), formerly DR2B or HLA-DRα,$\beta_1$2.2 (Marsh et al., 1982, 1990; Zwollo et al., 1991; Huang et al., 1991), suggesting that the two *Amb* V molecules possess similar Ia-binding epitopes.

To define precisely the Ia/T cell and B cell epitopes of the *Amb* V molecules, it is essential to know the proteins' detailed three-dimensional structures. Despite the keen interest in these proteins, little is known about their conformations in solution. In fact, although these proteins are highly cross-linked, no biochemical data are currently available discerning the cystine partners participating in the disulfide linkages. Thus we decided to perform a full structural analysis of several of these allergens by nuclear magnetic resonance (NMR) spectroscopy. NMR, in combination with computational techniques, has proven to be an invaluable tool for the elucidation of the three-dimensional structure of proteins and nucleic acids in

[†] This work was supported in part by a grant from the NIH to D.G.M. (AI19727).

[‡] Atomic coordinates for the derived *Amb t* V structure have been deposited in the Brookhaven Protein Data Bank.

* To whom correspondence should be addressed.

[§] Bristol-Myers Squibb Pharmaceutical Research Institute.

[∥] Princeton University.

[⊥] The Johns Hopkins University School of Medicine.

0006-2960/92/0431-5117$03.00/0

solution [for reviews, see Clore and Gronenborn (1989) and Wuthrich (1986)]. In this report, the solution structure of the ragweed protein *Amb t* V is presented.

Materials and Methods

Sample Preparation. *Amb t* V was isolated from *A. trifida* pollen (Greer Laboratories, Lenoir, NC) essentially according to the procedure of Roebber et al. (1985), except that HPLC was used as a final purification step. In brief, 250 g of pollen was extracted (at 1:10 w/v) with 0.05 M Tris-HCl buffer at pH 7.6 for 16 min at 23 °C. The resultant extract was subjected to ultrafiltration through an Amicon H1P10 hollow-fiber cartridge. The filtrate (containing *Amb t* V) was concentrated by pressure filtration with a YM2 membrane (through which *Amb t* V does not pass), and the concentrate was depigmented and partially purified on a DEAE-cellulose column. The fraction eluted with 0.05 M Tris-HCl buffer at pH 7.6 was further fractionated on a CM-Sephadex column, followed by gel filtration (twice) on a Bio-Gel P10 column. The resultant *Amb t* V-containing fraction was further purified by reverse-phase HPLC with a Waters RCM 8 × 10 radially compressed Nova-Pak C_{18} column in H_2O/0.1% TFA, with an increasing gradient of acetonitrile/0.1% TFA. The final *Amb t* V sample eluted as a single peak on the same HPLC column and gave single protein-staining bands on SDS–PAGE and by agarose gel electrophoresis at pH 8.6 [cf. Roebber et al. (1985)].

For the NMR experiments in H_2O, 9.0 mg of highly purified *Amb t* V was dissolved in 600 μL of NMR buffer (50 mM acetate, 0.1 mM EDTA, 0.1 mM NaN_3, pH 5.0 at 23 °C, with 10% ^2H_2O included as an internal lock) to yield a final concentration of 3.5 mM. For measurement of amide-exchange rates, the sample was lyophilized, dissolved in 99.8% ^2H_2O, and placed immediately into the spectrometer. Samples were later exchanged into ^2H_2O by two cycles of lyophilization followed by dissolution of the sample in 99.8% ^2H_2O, with a final lyophilization and dissolution step in 99.996% ^2H_2O before transferring the sample to a 5-mm NMR tube.

NMR Spectroscopy. The NMR experiments were performed on a Varian Unity 600 spectrometer operating at 599.5 MHz. 2D NMR spectra were recorded in the phase-sensitive mode with quadrature detection in both dimensions by the hypercomplex method (Mueller & Ernst, 1979; States et al., 1982). Water presaturation was achieved by applying continuous low-power irradiation from the transmitter during the 1.2-s recycle delay, followed by a scuba pulse train to allow saturated α protons to recover (Brown, S. C., et al., 1988). The 2D data were collected as follows: 2048 complex data points in t_2 with 350–512 complex FIDs in t_1. A total of 16–32 transients were signal-averaged for each point in t_1, and a spectral width of 7700 Hz was used in both dimensions. Data sets were acquired at a temperature of 30 °C.

Double-quantum-filtered correlated spectroscopy (DQF-COSY; Piantini et al., 1982; Rance et al., 1983), total correlated spectroscopy (clean TOCSY; Braunschweiler & Ernst, 1983; Bax & Davis, 1985), primitive exclusive COSY (PE-COSY; Mueller, 1987; Marion & Bax, 1988), "no-diagonal" COSY (ND-COSY; Friedrichs et al., 1991), and 2D nuclear Overhauser enhancement spectroscopy (NOESY; Macura & Ernst, 1980; Kumar et al., 1980) were acquired by standard methods. The clean TOCSY experiments employed an average spin lock field of 8.3 kHz for 70 ms, consisting of a 3-ms trim pulse followed by an MLEV-17 composite pulse sequence. NOESY data were acquired with mix times of 80, 120, 160, and 300 ms in ^2H_2O and with mix times of 100, 150, and 300 ms in H_2O. A composite 180° pulse was included in the middle of the mix time to minimize recovery of the solvent peak.

All NMR data were transferred to a Silicon Graphics Iris 4D/240 VGX computer and processed with the program FELIX (Hare Research, Inc.). The first complex points in both the t_1 and t_2 dimensions were scaled to eliminate ridge artifacts (Otting et al., 1986). A low-frequency deconvolution filter (Marion et al., 1989) was applied in t_2 to eliminate the residual water peak. All spectra, except for the PE-COSY, were Fourier transformed into a 2048 × 2048 matrix by zero filling, with both time domains in each data set multiplied by a 30–90° phase-shifted skewed sine bell (COSY-type spectra) or a Lorentzian/Gaussian (NOESY and TOCSY spectra) apodization function before Fourier transformation. The PE-COSY spectrum was transformed into a 4096 × 4096 matrix as above.

Cross-Peak Assignments. Sequence-specific resonance assignments were made by the standard method (Wuthrich, 1986). Cross-peaks were classified according to spin system with the through-bond experiments DQF-COSY and TOCSY. Sequential connectivities between spin systems were then made using the NOESY data.

Nonsequential NOE cross-peaks were assigned by using a procedure similar to that described by Kline et al. (1990). A large number of cross-peaks could be unambiguously assigned on the basis of their unique chemical shifts in uncrowded regions of the spectra. Distance constraints were then calculated from the cross-peak volumes, and preliminary structures were generated (see below for complete details). These initial structures allowed further assignments to be made with the aid of a modification of the program IDNOE (Weber and Mueller, personal communication; Kline et al., 1990) and the use of spin diffusion cross-peaks derived from spectra acquired with longer mix times. With this protocol, 114 nonsequential NOE cross-peaks were identified.

Derivation of Structural Constraints. Three types of constraints were derived from the NMR data: interproton distance constraints, torsion angle constraints, and hydrogen bond constraints.

A total of 437 distance constraints were found by using the NOESY spectra. These included 180 intraresidue constraints, 123 sequential constraints, 55 medium-range constraints ($|i - j| < 5$), 59 long-range constraints, and 20 constraints defining 10 hydrogen bonds. Each interresidue NOE was converted into an interproton distance by normalizing its integrated cross-peak volume in either the 150-ms (H_2O) or 160-ms (^2H_2O) mix time NOESY spectrum against a calibrated volume. For the calibration, the average volume of nine resolved Hβ1–Hβ2 cross-peaks was used to calculate a reference distance based on the two-spin approximation. [After the structures had been generated, cross-peak volumes were recalibrated using the volume of the average sequential αN connectivity of residues in the β-sheet and the volume of the βH–γ(CH_3) connectivity of Val. Target distances were not substantially affected.] Each cross-peak was then classified as strong, medium, or weak on the basis of an interproton distance of less than 2.7 Å, between 2.7 and 3.5 Å, or greater than 3.5 Å. Constraint ranges were generated by adding ±0.5 Å, −0.5 +1.0 Å, and ±1.0 Å to the calculated distance of the strong, medium, and weak cross-peaks, respectively. Intraresidue NOEs were derived from the 80-ms NOESY spectrum and classified as above. For cross-peaks in which one of the protons was an amide, the lower distance bound was set to 1.8 Å (for stronger cross-peaks) or 2.0 Å (for weaker cross-peaks) to allow for the possibility of amide exchange reducing the

intensity of the NOE, unless there was strong evidence that rapid exchange was not occurring. For distance constraints involving methyl protons, 1.0 Å was added to the upper distance bound.

Constraints on the backbone Φ torsion angles were derived from the value of $^3J_{HN\alpha}$ measured in a high-resolution DQF-COSY spectrum. Torsion angle constraint ranges for Φ were set to −120° ± 30° (for $^3J_{HN\alpha} > 9.0$ Hz), −120° ± 40° (for 9.0 Hz > $^3J_{HN\alpha} > 8.3$ Hz), and −60° ± 30° (for $^3J_{HN\alpha} < 4.5$ Hz and when medium-range NOEs indicated the presence of an α-helix). Stereospecific assignments for the β-methylene protons were made by analysis of the $^3J_{\alpha\beta}$ (measured in the PE-COSY spectrum) and the intraresidue NH–β and α–β NOEs (Hyberts et al., 1987). In this way, 14 of the possible 22 nondegenerate β-methylene protons were assigned, and torsion constraints for χ_1 (either 60°, −60° or 180° ± 30°) were obtained for 14 residues. Constraints for the χ_1 angles of residues with a single β-methylene proton (Thr, Val, and Ile) were obtained by performing a grid search of the available conformational space using the program HABAS (Guntert et al., 1991). In addition to providing several more χ_1 constraints, the program yielded numerous constraints for the Ψ torsion angles and the remaining Φ torsion angles. It should be noted that the majority of these latter constraints were very loose and did not necessarily restrict the available conformational space significantly. A total of 32 Φ, 24 Ψ, and 18 χ_1 torsion angle constraints were obtained.

Hydrogen bond constraints were derived from analysis of amide proton exchange rates. Constraints were included only after initial structure calculations provided unambiguous identification of the hydrogen bond acceptor group. In total, 20 constraints defining 10 hydrogen bonds were included; for each hydrogen bond, the HN–O and the N–O distances were constrained to be 2.0 ± 0.4 and 3.0 ± 0.4 Å, respectively.

Structure Calculation. All structure calculations were performed in the program XPLOR on a SGI 4D/240 VGX. An initial set of 10 structures was generated by randomly selecting values for the Φ and Ψ torsion angles. These random structures were then refined by using only unambiguous constraints to produce a set of preliminary structures. Examination of these structures allowed many of the ambiguities in the assignment of NOE cross-peaks to be resolved and enabled the definition of Ψ torsion angles that fell only into one of two possible constraint ranges. These preliminary structures also permitted the cystine partners composing the disulfide bridges to be determined. In addition, hydrogen bond acceptors for some of the slowly exchanging amide protons (in the helix and sheet) could be identified. These additional constraints were then used to refine further the preliminary structures, and the process was repeated. After all possible constraints had been extracted, a set of 50 new random structures was generated and refined against the complete collection of constraints.

The refinement procedure used in the program XPLOR has been described in detail (Nilges et al., 1988). Briefly, the procedure involves an initial Powell minimization of the random structures to remove any bad steric interactions, followed by extensive dynamic simulated annealing and a final round of Powell minimization. During the initial stage of dynamic simulated annealing, constraints are weighted so that the penalty assessed to the largest violations is reduced. The reduction of the penalty in effect allows the local structure to form before the global structure and helps prevent "knotting" of the structure. In addition, a dramatically reduced van der Waals repulsive potential is used, allowing constraints to be rapidly satisfied. Next, the van der Waals repulsive potential is increased, as is the penalty assessed to the largest constraint violations. A final round of annealing ensures that the structure is at equilibrium with respect to all potentials applied (vdW repulsion, bonds, angles, dihedrals, impropers, NOE and torsion angle constraints). Refined structures were analyzed for the best fit to the NMR constraints and the overall energy.

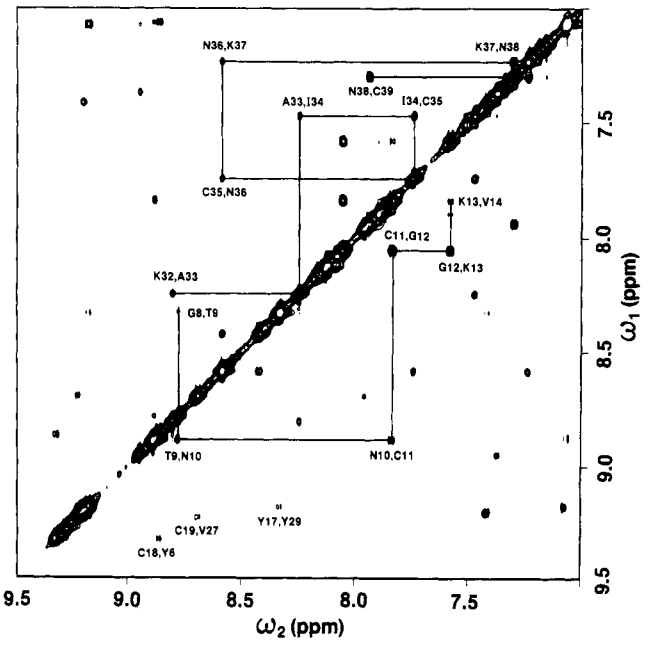
Back-Calculation. The 2D NOE spectrum for a given structure was calculated with modified versions of the programs BKCALC and GNOE (Hare Research, Inc.). The procedure for back-calculation has been described (Banks et al., 1989). It involves numerical integration of the Bloch equations to explicitly account for the effects of spin diffusion. These calculations require values for the cross-relaxation scaling factor and the "Z leakage" rate. The cross-relaxation scaling factor corresponds to a correlation time parameter, and it was assumed to be constant for the whole molecule. The value for the cross-relaxation scaling factor was estimated by determining the best fit of the cross-peak/diagonal peak intensity ratios of several well-resolved cross-peaks which result from protons that are separated by a fixed distance. The Z leakage rate accounts for the loss of all Z magnetization during the mix time for an individual spin, except for that arising from NOE cross-relaxation, and thus corresponds to a T1-type parameter. The Z-leakage rate (0.8 s) was also assumed to be constant for the whole molecule and was estimated by fitting the decay of the total intensity of the 2D spectrum at each mix time to an exponential form.

Results

Resonance Assignments. Assignment of the proton resonances of *Amb t* V was performed by standard procedures as described in Wuthrich (1986). The first step is to associate scalar-coupled resonances into specific spin systems using DQF-COSY and TOCSY experiments, thus classifying the resonances according to their amino acid type. NOESY spectra are then used to assign the spin systems to specific amino acids in the protein sequence. Numerous spin systems were identified that provided various starting points for the sequential connectivities. These included five glycines, two threonines, two valines, and a unique alanine at position 33. In addition, analysis of the NOESY spectrum enabled the aromatic protons of the four tyrosines and the δ-amide protons of the three asparagines to be linked to their respective β protons, reducing the degeneracy of the 20 AMX spin systems and providing additional markers in the sequence. A NOESY spectrum at 300 ms was used to locate the αN, βN, and NN connectivities necessary to make interresidue connectivities. Figure 1 shows contour plots of the sequential connectivities for the αN and NN regions of the NOESY. The NOEs used in making the sequential connectivities are summarized in Figure 2. Connectivities for all residues except residue 1 could be found. Table I lists the proton assignments for *Amb t* V.

Secondary Structure Analysis. The presence of slowly exchanging amide protons provides evidence for a stable secondary structure. Secondary (and tertiary) structural features protect amide protons from exchange with the solvent (Englander & Kallenbach, 1984). Several slowly exchanging amide protons were observed after dissolution of the lyophilized protein in D_2O (30 °C) and acquisition of TOCSY data. The use of the TOCSY experiments allowed for unambiguous assignment of these amide protons (Figure 2). In fact, the amide protons of residues 29 and 19 remained unexchanged even after 2 months in ^2H_2O at 4 °C ($p^2H = 5.0$), a clear sign of a stable, regular secondary structure.

Qualitative analysis of the medium-range NOEs (Figure 2) shows the existence of a stretch of $d_{\alpha\beta}(i,i+3)$ and $d_{\alpha N}(i,i+3)$



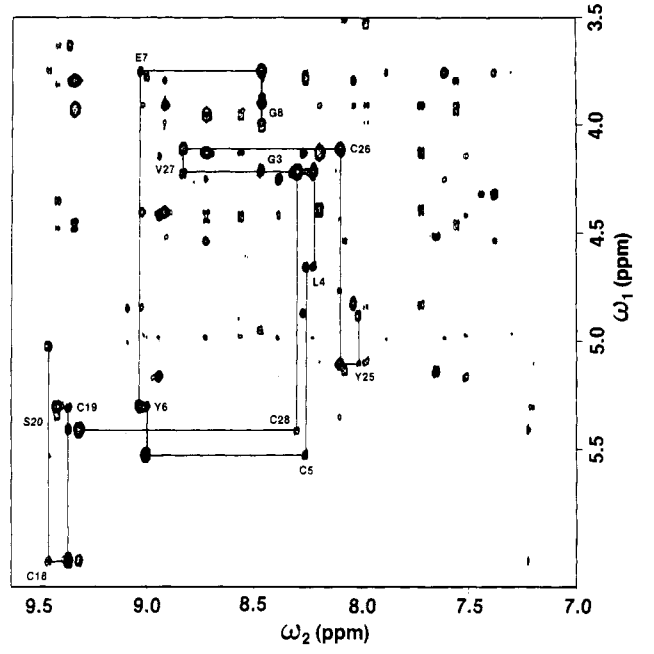


FIGURE 1: Expansion of the amide proton (A, top) and fingerprint (B, bottom) regions of the NOESY spectrum acquired with a 300-ms mix time. Representative sequential connectivities are highlighted.

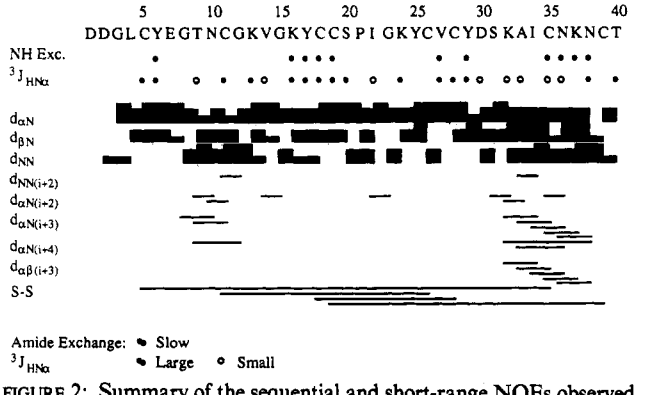


FIGURE 2: Summary of the sequential and short-range NOEs observed for *Amb t* V: (●) residues with slowly exchanging amide protons or with a large (≥8.3 Hz) value for $^3J_{HN\alpha}$; (○) residues with a small (≤4.5 Hz) value for $^3J_{HN\alpha}$. For the sequential NOEs, the relative intensity of the NOE is represented by the thickness of the bar. S–S indicates the cystine partners of the disulfide bridges.

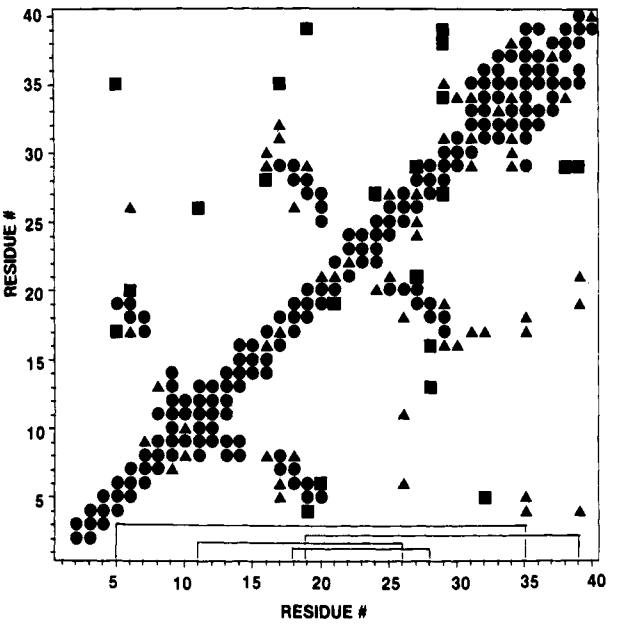


FIGURE 3: Diagonal plot of the NOEs observed for *Amb t* V. The presence of at least one NOE between protons of the two residues in the sequence position is indicated by (●) two backbone (HN or Hα) protons, (▲) one backbone proton and one side-chain proton, and (■) two side-chain protons. Preference is given to the symbol of the NOE with the highest number of backbone atoms. Data from 150- and 300-ms NOESY spectra are plotted to the upper left and lower right of diagonal, respectively.

connectivities from residues 31–38, suggesting the presence of helical conformation in this region. This conclusion is confirmed by the small values for $^3J_{HN\alpha}$ (which are expected for a helix) and the stretch of slowly exchanging amide protons.

Because of their extended conformation, β-sheets are characterized by large $^3J_{HN\alpha}$ coupling constants, short sequential $d_{\alpha N}$ distances, and long intraresidue $d_{N\alpha}$ distances. An example of such a stretch is depicted in the sequential connectivities in Figure 1b. A series of close contacts between α protons remote in the primary structure indicate that the β-sheet is antiparallel. Figure 3 presents the observed NOEs in a diagonal plot, with residue numbers plotted on both axes. The two stretches of NOEs running orthogonal to the diagonal suggest the β-sheet is triple stranded. Inspection of the NOESY spectrum shows that all of the NOEs expected for a β-sheet are present. Additional evidence for a β-sheet is provided by the presence of the very slowly exchanging amide protons mentioned above.

Structure Determination. A set of 50 randomly generated starting structures was refined by dynamic simulated annealing with the program XPLOR as described under Materials and Methods. After refinement, 31 structures had no distance constraint violations greater than 0.5 Å, and most of these 31 structures only had four or fewer distance constraint violations greater than 0.3 Å. In addition, there were essentially no violations of the torsion angle constraints for these structures. For the complete structural analysis, the 15 structures that gave the lowest rms distance deviation from the input distance constraints were used. (The number of structures analyzed, 15, was arbitrarily chosen as a cutoff.) An average refined structure was calculated from these 15 structures, and its energy was minimized in XPLOR.

Superposition of the backbone atoms for the 15 final refined structures is shown in Figure 4. Structural statistics are given in Table II. With the exceptions of the N-terminal residues 1–4 and the C-terminal residue 40, most regions of *Amb t* V were well-defined by the NMR data. Analysis of the rms

Table I: Proton Chemical Shifts (ppm)[a]

residue	NH	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	other
D1		4.29	2.89, 2.81			
D2	7.99	4.18	0.90, 1.58			
G3	8.36	3.97				
L4	8.11	4.41	1.68	1.66	0.92, 0.90	
C5	8.15	5.28	3.54^R, 2.71^S			
Y6	8.89	5.06	3.11, 3.02		7.09	$C_\epsilon H$ 6.79
E7	8.92	3.51	2.19^S, 1.76^R	2.54, 2.07		
G8	8.35	3.76, 3.64				
T9	8.81	3.67	4.16	1.19		
N10	8.91	4.59	2.84^S, 2.77^R			$N_\delta H$ 7.38, 6.87
C11	7.87	4.85	3.29^S, 2.86^R			
G12	8.08	4.15, 3.88				
K13	7.61	4.58	1.55, 1.37	1.07	1.29	$C_\epsilon H$ 2.85
V14	7.93	3.56	1.89	1.06, 0.90		
G15	9.22	4.22, 3.69				
K16	7.45	4.70	2.00, 1.70	1.25, 1.15	1.36	$C_\epsilon H$ 2.83
Y17	8.36	4.78	2.54, 2.36		6.69	$C_\epsilon H$ 6.86
C18	9.34	5.76	2.74^R, 2.35^S			
C19	9.25	5.06	3.39, 3.12			
S20	9.30	5.10	4.24^S, 4.11^R			
P21		4.63	2.36, 1.99	2.10, 1.87	3.57, 3.47	
I22	8.16	3.89	1.82	1.62	1.26	$C_\gamma H_3$ 0.93
G23	8.61	4.18, 3.69				
K24	8.45	4.64	1.47^R, 1.34^S	1.77	1.65, 1.54	$C_\epsilon H$ 2.98; $N_\epsilon H$ 7.55
Y25	7.90	4.86	2.92		7.10	$C_\epsilon H$ 6.79
C26	7.99	3.87	3.19^S, 2.69^R			
V27	8.72	3.97	2.22	0.85, 0.83		
C28	8.19	5.16	2.81^R, 2.52^S			
Y29	9.20	4.77	3.20^S, 2.55^R		7.11	$C_\epsilon H$ 6.70
D30	8.98	4.60	2.88, 2.59			
S31	7.41	4.93	4.17^R, 3.90^S			
K32	8.84	3.19	1.27^S, 0.30^R	1.07, 0.68	1.36	$C_\epsilon H$ 2.80
A33	8.26	4.01	1.31			
I34	7.50	3.52	1.63	1.53	0.92	$C_\gamma H_3$ 1.06
C35	7.77	3.12	3.12, 2.99			
N36	8.61	4.29	2.66			$N_\delta H$ 7.22, 6.98
K37	7.25	4.07	1.77	1.59	1.42	$C_\epsilon H$ 2.92; $N_\epsilon H$ 7.48
N38	7.33		2.81^S, 1.96^R			$N_\delta H$ 7.19, 6.80
C39	7.97	4.89	3.26^S, 2.91^R			
T40	7.54	4.28	4.28	1.14		

[a] Assignments for *Amb t* V in 50 mM acetate buffer (pH = 5.0) in ^2H_2O at 30 °C. Amide assignments in the same buffer in H_2O at 30 °C. *R* and *S* denote stereospecifically assigned protons.

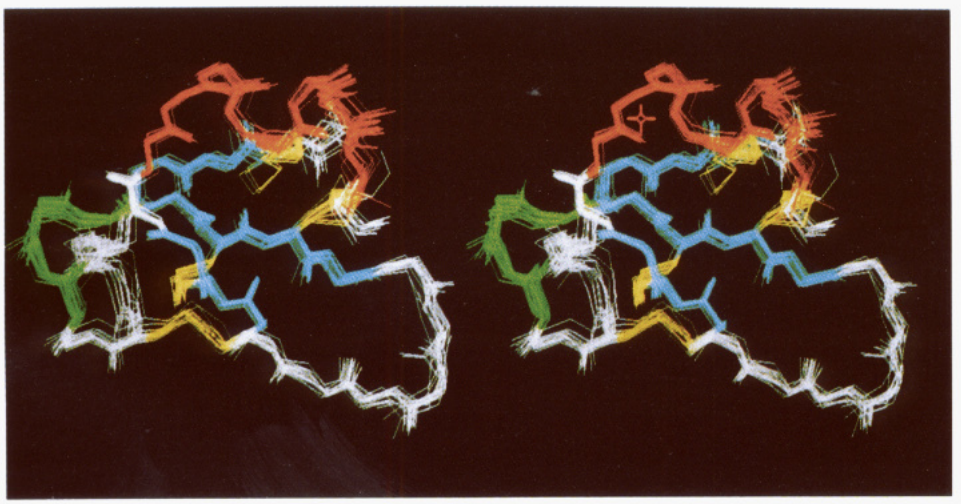


FIGURE 4: Stereoview of the 15 final refined structures of *Amb t* V. The structures were superimposed to minimize the rmsd of the backbone atoms (N, Cα, C, O) of residues 5–38. N-Terminal residues 1–4 and C-terminal residues 39–40 are not shown.

distance deviations and torsion angle deviations confirms this conclusion. The average pairwise rms difference for the backbone atoms (N, Cα, C, O) of residues 5–39 is 0.64, with an average angular rms difference for the torsion angles Φ and Ψ of 10°. Figure 5 plots the atomic rms distance and torsion angle deviations of the 15 final refined structures from the minimized average structure. Several regions of the protein are better defined than others. For example, the α-helix and β-sheet are well-defined, whereas the turn and loop areas exhibit higher conformational variability (see below).

Structural Details. On a coarse structural level, *Amb t* V can be viewed as two segments held together by two disulfide bonds: Cys_{18}–Cys_{28} and Cys_{11}–Cys_{26}. The N-terminal segment (residues 1–30) contains a small region of triple-stranded

Table II: Structural Statistics

	⟨SA⟩	$(SA)_r$
rmsd from experimental NOE constraints[a]		
all (409)	0.051	0.047
intraresidue (157	0.052	0.055
sequential (110)	0.036	0.040
medium range (45)	0.064	0.055
long range (97)	0.058	0.036
energy[b]		
F_{NOE} (kcal/mol)	51.3	48.0
F_{TOR} (kcal/mol)	0.74	0.74
F_{REP} (kcal/mol)	21.4	18.5
E_{L-J} (kcal/mol)	−140	−139
deviations from idealized covalent geometry		
bonds	0.005	0.004
angles	1.95	2.78
impropers[c]	1.04	1.06

[a] ⟨SA⟩ refers to the 15 individual final refined structures. $(SA)_r$ is the average, energy-minimized structure obtained as described in the text. The number of constraints are given in parentheses. [b] Square-well NOE and torsion angle potentials were calculated in the program XPLOR with force constants of 50 kcal mol^{-1} $Å^{-2}$ and 80 kcal mol^{-1} rad^{-2}, respectively [see eq 2 and 3 in Clore et al. (1986)]. The quartic van der Waals repulsion term was calculated with a force constant of 4 kcal mol^{-1} $Å^{-4}$ with the hard-sphere van der Waals radii set to 0.8 times the standard values used in the CHARMM (Brooks et al., 1983) empirical energy function [see eq 5 in Nilges et al. (1988)]. The Lennard-Jones van der Waals energy was calculated with the CHARMM empirical energy function. [c] Impropers refer to planarity and chirality constraints.

β-sheet, with two irregular loops connecting the strands. Two disulfide bonds stabilize this substructure. The C-terminal segment (residues 31–40) is helical. The two segments are held together by two additional disulfide bonds: Cys_5–Cys_{35} and Cys_{19}–Cys_{39}.

Secondary structural analysis of NOE cross-peak patterns suggested that a helix was present in the C-terminus of the protein. However, at this stage, it was not possible to determine whether it was an α-helix or a 3–10 helix. An examination of the initial refined structures indicated that the hydrogen exchange data could only be explained by the presence of hydrogen bonds between CO_i and NH_{i+4}, suggesting an α-helix. This conclusion was supported by the detection of $d_{\alpha N}(i,i+4)$ NOEs. Nevertheless, two model structures were refined in which a hydrogen bond constraint for each slowly exchanging amide was included. Whereas the structure with hydrogen bond constraints between CO_i and NH_{i+4} satisfied the NOE constraints, the structure with hydrogen bond constraints between CO_i and NH_{i+3} resulted in substantial constraint violations. This refinement confirmed that *Amb t* V contained an α-helix. All subsequent refinements included hydrogen bond constraints between the CO_i and NH_{i+4} of residues$_i$ 31–34. A superposition of the backbone atoms of the α-helices in the final refined structures is shown in Figure 6A. The backbone is well-defined, with an average pairwise rms distance deviation of 0.26 Å. The helix extends from residues 31 through 38. C-Terminal residues 39 and 40 lie in a more extended conformation, with residue 40 being disordered.

The secondary structure analysis predicted a small section of triple-stranded antiparallel β-sheet. The primary indicators for this were the large NOEs between the α protons of residues 18 and 28, 7 and 17, and 5 and 19. Figure 7 summarizes the sequential and long-range backbone NOEs observed in this region. Amide protons with decreased exchange rates are also shown. A large number of NOEs involving side-chain protons are observed which help define the side-chain conformations. As was the case with defining the helix, no hydrogen bond constraints were added until after the initial round of structural refinement, and even then only when the hydrogen bond acceptor was completely unambiguous. No hydrogen bond constraint was included for the amide proton of residue 6 despite the presence of the CO of residue 18 as a potential acceptor because hindered amide exchange was not observed for this proton. A superposition of the sheet is shown in Figure 6B. The three strands of the β-sheet run from residues 5 to 7, 17 to 20, and 27 to 29. Again, this region of *Amb t* V is well-defined, with a pairwise rms distance deviation of the backbone atoms for the 15 final refined structures of 0.30 Å.

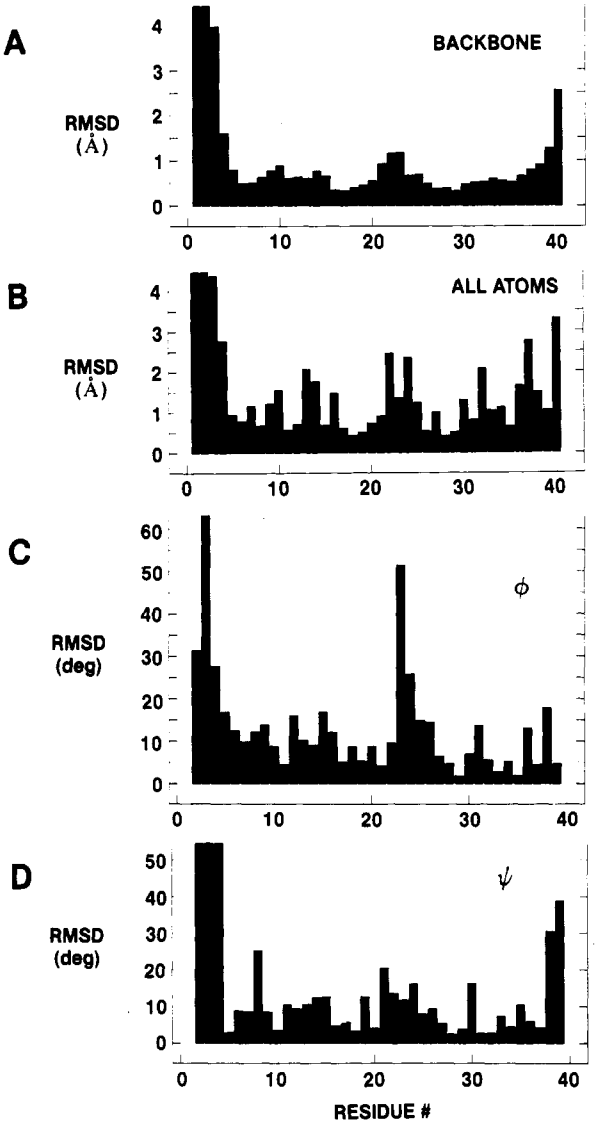


FIGURE 5: Plots of the average pairwise rmsd of the 15 final refined structures of *Amb t* V vs sequence position: (A) rmsd in the atomic position for the backbone atoms (N, Cα, C, O); (B) rmsd in the atomic position for all atoms; (C) rmsd in the angular distribution of torsion angle Φ; (D) rmsd in the angular distribution of torsion angle Ψ. The atomic rmsds were determined by superimposing the 15 structures to minimize the pairwise rmsd of the backbone atoms of residues 5–38 and then calculating the appropriate RMSD for the each residue.

The three strands of β-sheet are connected by two irregular loops. The loop from residues 8 through 16, however, actually contains a type II β-turn (Richardson, 1981) between residues 13–16. A superposition of this turn from the 15 structures indicates how well-defined it is (Figure 6C). The backbone rms distance deviation of this region is 0.23 Å. The loop is characterized by a strong αN NOE between residues 14 and 15, a medium αN NOE between residues 14 and 16, and a strong NN connectivity between residues 15 and 16. The carbonyl oxygen of residue 13 was found to be in position for

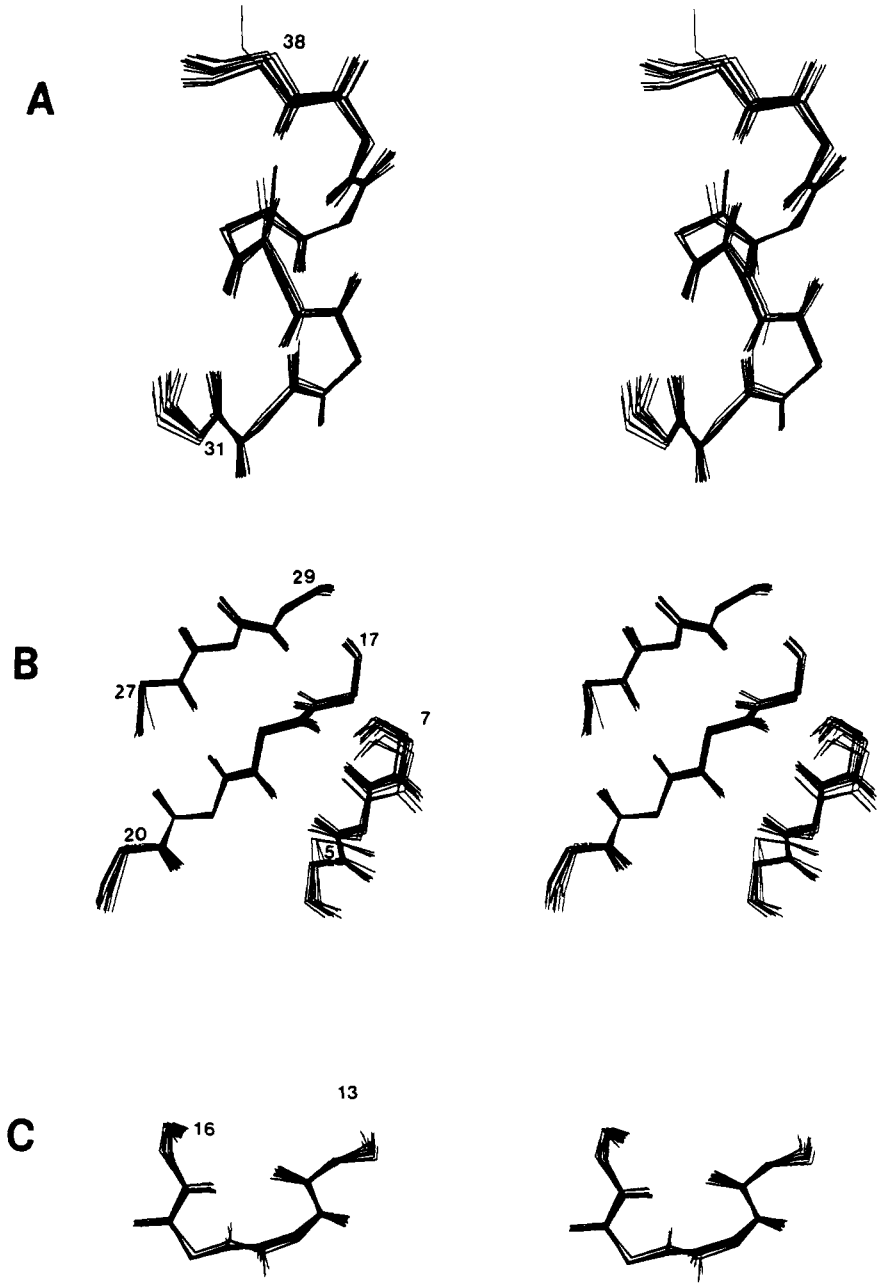


FIGURE 6: Stereoviews of the (A) helix, (B) β-sheet, and (C) turn of the 15 refined structures of *Amb t* V. The structures were superimposed by minimizing the rmsd of the backbone atoms (N, Cα, C, O) of the regions shown: residues 31–38 of the helix; residues 5–7, 17–20, and 27–29 of the β-sheet; and residues 13–16 of the turn.

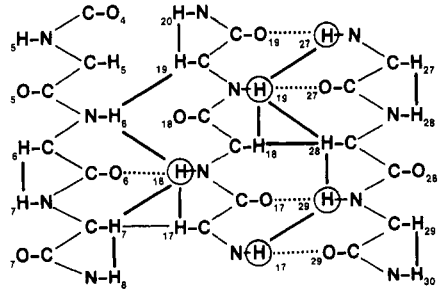


FIGURE 7: Short sequential and long-range backbone NOEs observed for the β-sheet region of *Amb t* V. NOEs are indicated by thick bars, and slowly exchanging amides are circled. Hatched bars indicate hydrogen bonds included in the structure determination.

formation of a hydrogen bond with the slowly exchanging amide proton of residue 16. This hydrogen bond was included as a constraint in the last stages of refinement. The remainder of the loop (residues 8–12) and the loop between strands 2 and 3 of the β-sheet (residues 21–26) are also well-defined but do not fall within the standard classification of turns (Richardson, 1981). The nonstandard conformations of these two loops may be a result of the disulfide bond between cystines 11 and 26. Residues 8–12 are characterized by a series of large NH–NH NOEs (Figure 1A). The loop from residues 21 to 26 shows higher conformational variability (rmsd = 0.47 Å) than the loop from residues 8 to 12 (rmsd = 0.30 Å) and appears to hinge on the psi torsions of glycine 23 (see Figure 5C).

N-Terminal residues 1–4 are quite disordered, which is not unexpected because there are no medium- or long-range NOEs constraining these residues. In addition, the $^3J_{HN\alpha}$ values for residues 2 and 4 are 6.5 and 7.1 Hz, respectively, and the chemical shifts for the α protons of glycine 3 are degenerate, suggesting that the observed torsion angles are conformationally averaged in solution.

There are four disulfide bonds holding the secondary structure units of *Amb t* V together. At the beginning of the

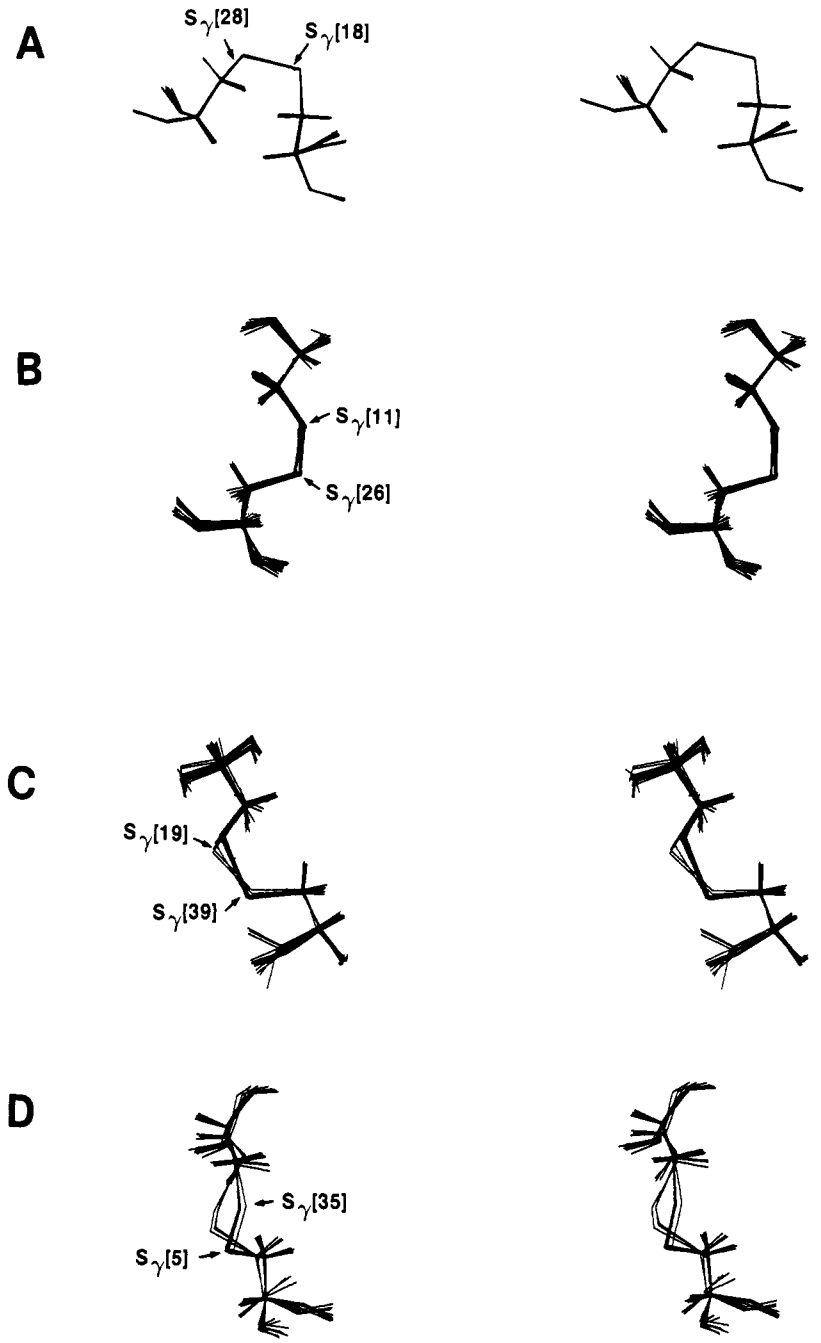


FIGURE 8: Stereoviews of the disulfide bridges of the 15 refined structures of *Amb t* V: (A) Cys_{18}–Cys_{28}; (B) Cys_{11}–Cys_{26}; (C) Cys_{19}–Cys_{39}; and (D) Cys_5–Cys_{35}. The structures were superimposed by minimizing the rmsd of all atoms of the disulfide bridge shown.

structure calculations, the disulfide pairs were unknown. Assignment of the partners in two disulfides in the N-terminal region, Cys_{18}–Cys_{28} and Cys_{11}–Cys_{26}, became apparent after qualitative analysis of the long-range NOE data. NOEs were found correlating the cross-disulfide β protons, and spin-diffusion NOEs were found between the α and β protons in experiments collected at long mix times: constraints for these disulfides were included in subsequent refinements. For the remaining two disulfides, assignment of the partners was not initially unambiguous. The β protons for Cys_5 had NOEs to the β protons of both Cys_{19} and Cys_{35}. In addition, the β protons of Cys_{39} had NOEs to both Cys_{35} and Cys_{19}. Initially, structure calculations were performed without any disulfide bond constraints. Analysis of the Cys Cβ distances was insufficient to descriminate between the various possible disulfides. To determine the correct disulfide pairings, three structure calculations were performed (one each for structures containing the disulfides), (1) Cys_5–Cys_{35} and Cys_{19}–Cys_{39}, (2) Cys_5–Cys_{19} and Cys_{35}–Cys_{39}, and (3) Cys_5–Cys_{39} and Cys_{19}–Cys_{35}. Only the calculation for the disulfide pairs Cys_5–Cys_{35} and Cys_{19}–Cys_{39} produced structures without substantial violations of the input NMR constraints. These disulfide bonds were then included as constraints in later stages of refinement. All disulfide bridges are well-defined in the final refined structures (Figure 8). Three of the disulfides fall roughly (±20° of the preferred torsion angles) into one of the families found for disulfide bridges in proteins (Srinivasan et al., 1990): Cys_{11}–Cys_{26} into that of the "left-handed spiral" (Richardson, 1981); Cys_{18}–Cys_{28} bridging the antiparallel β-sheets; and Cys_5–Cys_{35} bridging an extended strand to helix. Although the disulfide bridge Cys_{19}–Cys_{39} is well-defined, it could not be categorized into one of the predefined classes.

Several of the *Amb t* V side-chain conformations are

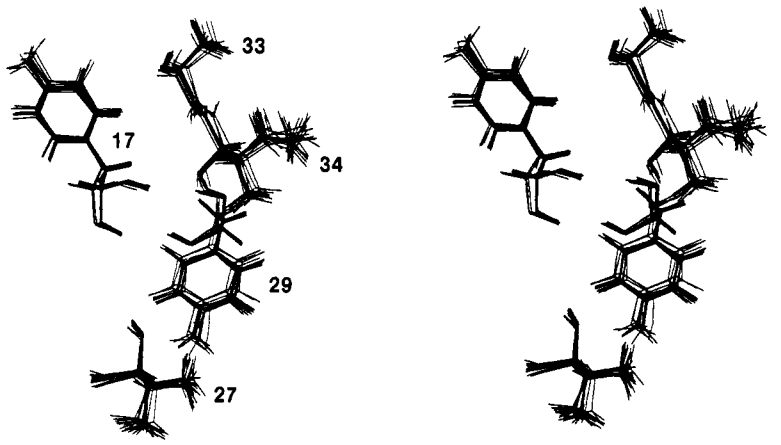


FIGURE 9: Stereoview showing the side-chain atoms of residues Tyr_{17}, Val_{27}, Tyr_{29}, Ala_{33}, and Ile_{34} for the 15 refined structures of *Amb t* V. The structures were superimposed by minimizing the RMSD of all atoms in the residues shown.

well-defined in the final refined structures. These make up part of a small hydrophobic core that lies between the sheet and the helix. Figure 9 shows an example of a superposition for the side chains of Tyr_{17}, Val_{27}, Tyr_{29}, Ala_{33}, and Ile_{34}.

Quality of Structural Calculations. Although the degree of convergence of a structural refinement can be determined by analyzing the number of distance and torsion angle constraint violations and the rmsd for distances and torsion angles among the set of structures, these criteria do not necessarily demonstrate that a "good" structure has been determined. We analyzed the four structures with four additional criteria. First, the structures do not deviate significantly from their idealized covalent geometry (Table II), indicating that satisfaction of the pseudo energy potentials for the NMR constraints has not led to unrealistic geometries. Second, the Leonard-Jones energies are both large and negative (Table II), despite the fact that no attractive forces were explicitly included in the structure calculations. Thus, satisfaction of the NMR constraints allows for favorable van der Waals contacts to form. Third, Ramachandran plots (Figure 10) illustrate that the Φ and Ψ torsion angles for the structures lie in sterically allowed regions. Although the Φ torsion angle for Cys_{26} is unusual in that it resides in positive Φ space, this conformation has been seen previously in the crystal structures of numerous proteins (for example, α-chymotrypsin). This region of Φ,Ψ space is of low energy and corresponds to a region where one often finds the Φ torsion angle of Asp and Asn residues involved in turns. Fourth, good agreement is found between experimental NOESY data and NOESY data calculated for the energy-minimized, average structure. An extensive comparison was made between the experimental NOEs used as constraints in the structure calculation and calculated NOEs derived from the average structure. Although no attempt has been made to match peak intensities, the comparison does indicate that nearly all NOEs expected are present in the calculated data. Only experimental NOEs from protons that are farther apart than the cutoff distance of 4.5 Å are not observed in the calculated spectrum. Numerous predicted NOEs were found in the calculated spectrum which did not have corresponding experimental constraints. Many of these NOEs were a result of spectral overlap that precluded their assignment in the experimental data. Other calculated NOEs were observed in the 150-ms experimental data but were not included as constraints because their small intensity made them difficult to quantitate. These NOEs were well above the noise level at the longer mixing time (300 ms). Several additional cross-peaks appear in the calculated spectrum. These were generally due to the less well-defined regions of the protein, which become fixed when the spectrum is calculated. A complete list of the cross-peaks has been included as supplementary material.

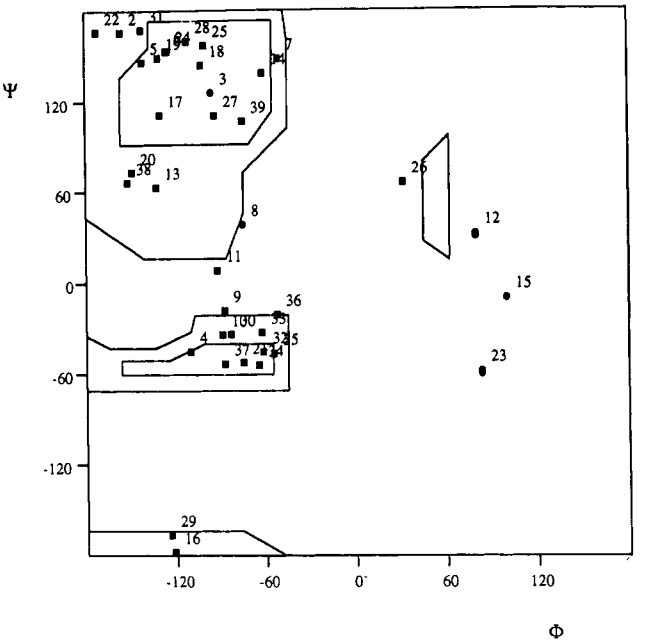


FIGURE 10: Ramachandran plot for the geometrically averaged, energy minimized structure of *Amb t* V. Each amino acid is labeled with its residue number: (●) glycines; (■) all other amino acids.

The derived structure of a protein should also be able to account for the observation of anomalous chemical shifts. For example, the β protons of Lys_{32} are shifted substantially upfield. Examination of the structure of *Amb t* V indicates that these protons lie directly beneath the aromatic ring of Tyr_{17}, a residue in the hydrophobic core of the protein. Thus the anomalous upfield chemical shift of the β protons of Lys_{32} is a reflection of this ring current effect. The chemical shifts of Asp_2 are also anomalous. However, this part of the protein is poorly defined by the NMR data, and we can find no explanation for this anomaly.

Discussion

The *Amb* V allergens from pollens of the genus *Ambrosia* (ragweeds) constitute a good model system for exploring certain structural features relevant in understanding the mechanism of the human immune response. Interestingly, because these small *Amb* V molecules (ca. 40–45 amino acids) usually contain four disulfide bonds, they are particularly

useful in exploring the nature of immunological epitopes containing disulfide bonds, which are generally not amenable to analysis with synthetic peptides. To date, no biochemical data have been published on the cystine partners of the disulfide linkages for *Amb t* V or any of its homologues. Two distinct disulfide pairings for *Amb a* V have been proposed previously. In 1980, on the basis of modeling studies of wheat germ agglutinin and snake venom erabutoxin, Drenth et al. proposed disulfide pairings for *Amb a* V that would correspond to residues 5–19 and 35–39 of *Amb t* V. In 1985, on the basis of 1D NMR experiments with *Amb a* V, Vidusek et al. suggested a disulfide pairing corresponding to residues 5–39 and 19–35 in *Amb t* V. The refined structures presented here agree with neither of these but indicate that disulfide bridges exist between residues 5–35 and 19–39. The remaining two disulfides, corresponding to residues 11–26 and 18–28 in *Amb t* V, agree with both of the previously proposed pairings.

The structure of *Amb t* V calculated here contains a rigid core composed of a short segment of β-pleated sheet that is stabilized by two disulfide bridges. An α-helix, located at the C-terminal region of the molecule, is tethered to the sheet by two more disulfide bonds, with charged residues exposed on the surface of the molecule. It is noteworthy that this C-terminal helical region is amphipathic and is implicated as being the Ia/T cell epitope (Huang & Marsh, 1991). Such a conformation is in agreement with the model of Margalit et al. (1987). However, we must emphasize that although we find that the putative Ia/T cell epitope to be helical in the native molecule, this does not mean to imply that it is also helical when it is bound as a peptide fragment by the MHC class II molecule. Recent crystallographic data of a distinct, yet related, MHC class I molecule indicate that peptides are bound in an extended conformation (Madden et al., 1991). These crystallographic data also suggest that the conformation of the antigenic peptide fragment is induced by the MHC molecule. However, some significant differences have been observed in the physical properties of peptides that bind to MHC class I and class II molecules. Thus, we believe it remains to be determined whether the extended conformation will be the rule for how all antigenic peptides are bound by MHC molecules or just one of the many vehicles by which this immunological recognition takes place. Therefore, we are attempting to clone a truncated MHC class II molecule that will be of size and solubility suitable for NMR analysis of the peptide–MHC class II molecule complex.

Although proteins in the *Amb* V family are homologous, the immune responses they elicit are often quite distinct. Experimental results suggest that *Amb t* V and *Amb a* V possess similar MHC class II binding regions but that the residues that interact with the T cell receptors (TcRs) are dissimilar (Huang & Marsh, 1991). In addition, *Amb t* V and *Amb a* V have been shown to be virtually non-cross-reactive at the B cell (antibody) level (Roebber et al., 1985). These data suggest that class II-binding regions are derived from evolutionarily conserved parts of the molecules, whereas the T cell receptor-binding and antibody-binding regions come from nonconserved parts. The NMR structural data we report here should be valuable in identifying the nature and location of the *Amb t* V epitopes. We plan to determine the three-dimensional structures of several members of the *Amb* V family, including some mutated variants, which will allow us to define the portions of these homologous molecules that comprise Ia/T cell and B cell epitopes.

Supplementary Material Available

NMR-derived constraints including all NOE constraints with cross-peak volumes used in the structure calculations, all torsion angle constraints with measured *J*-values; hydrogen bond constraints derived from the exchange data, disulfide bond constraints with a listing of all NOEs used in the assignment or confirmation of the disulfide bonds, spin-diffusion cross-peaks observed in the 300-ms NOESY spectra, and discrepant cross-peaks predicted by back-calculation but not included in the experimental constraints (19 pages). Ordering information is given on any current masthead page.

References

Banks, K. M., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 6996–7010.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Benjamin, D. C., Berzofsky, J. A., East, I. J., Gurd, F. R., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., Prager, E. M., Reichlin, M., Sercarz, E. E., Smith-Gill, S. J., Todd, P. E., & Wilson, A. C. (1984) *Annu. Rev. Immunol. 2*, 67–101.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987) *Nature 329*, 506–512.

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.

Brown, J. H., Jardetzky, T., Saper, M. A., et al. (1988) *Nature 332*, 845–850.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

Clore, G. M., & Gronenborn, A. M. (1989) *CRC Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.

Clore, G. M., Brunger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *J. Mol. Biol. 105*, 1–14.

Drenth, J., Low, B. W., Richardson, J. S., & Wright, C. S. (1980) *J. Biol. Chem. 255*, 2652–2655.

Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521–655.

Friedrichs, M., Metzler, W. J., & Mueller, L. (1991) *J. Magn. Reson. 95*, 178–183.

Goodfriend, L., Choudhury, A. M., Klapper, D. G., Coulter, K. M., Dorval, G., Del Carpio, J., & Osterland, C. K. (1985) *Mol. Immunol. 22*, 899–906.

Guntert, P., Braun, W., & Wuthrich, K. (1991) *J. Mol. Biol. 217*, 517–530.

Huang, S. K., & Marsh, D. G. (1991) *Immunology 73*, 363–365.

Huang, S. K., Zwollo, P., & Marsh, D. G. (1991) *Eur. J. Immunol. 21*, 1469–1473.

Hyberts, S. G., Marki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625–635.

Jeener, J., Meier, B. H., Backmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Kline, T. P., Brown, F. K., Brown, S. C., Jeffs, P. W., Kopple, K. D., & Mueller, L. (1990) *Biochemistry 29*, 7805–7813.

Kumar, A., Ernst, R. R., & Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Lapkoff, C. B., & Goodfriend, L. (1974) *Int. Arch. Allergy Appl. Immunol. 46*, 215–229.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Madden, D. R., Gorga, J. C., Strominger, J. L., & Wiley, D. C. (1991) *Nature 353*, 321–325.

Margalit, H., Spouge, J. L., Cornette, J. L., De Lisi, C., & Berzofsky, J. A. (1987) *J. Immunol. 138*, 2213–2229.

Marion, D., & Bax, A. (1988) *J. Magn. Reson. 80*, 528–533.

Marion, D., Ikura, M., & Bax, A. (1989) *J. Magn. Reson. 84*, 425–428.
Marsh, D. G. (1975) in *The Antigens* (Sela, M., Ed.) Vol. III, pp 271–359, Academic Press, New York.
Marsh, D. G. (1986) *J. Allergy Clin. Immunol. 76* (Suppl.), 242–248.
Marsh, D. G., Hsu, S. H., Roebber, M., Kautzky, E. E., Freidhoff, L. R., Meyers, D. A., Pollard, M. K., & Bias, W. B. (1982) *J. Exp. Med. 155*, 1439–1451.
Marsh, D. G., Zwollo, P., Huang, S. K., Ghosh, B., & Ansari, A. A. (1990) *Cold Spring Harbor Symp. Quant. Biol. 54*, 459–470.
Mole, L. E., Goodfriend, L., Lapkoff, C. B., Kehoe, J. M., & Capra, J. D. (1975) *Biochemistry 14*, 1216–1220.
Mueller, L. (1987) *J. Magn. Reson. 72*, 191.
Mueller, L., & Ernst, R. R. (1979) *Mol. Phys. 38*, 963–992.
Nilges, M., Gronenborn, A. M., & Clore, G. M. (1988) *FEBS Lett. 239*, 129–136.
Otting, G., Widmer, H., Wanger, G., & Wuthrich, K. (1986) *J. Magn. Reson. 66*, 187–193.
Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wanger, G., Ernst, R. R., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 69*, 979–987.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.
Roebber, M., Klapper, D. G., Goodfriend, L., Bias, W. B., Hsu, S. H., & Marsh, D. G. (1985) *J. Immunol. 134*, 3062–3069.
Sette, A., Buus, S., Colon, S., Smith, J. A., Miles, C., & Grey, H. M. (1987) *Nature 1328*, 395–399.
Srinavasan, N., Sowdhamini, R., Ramakrisnan, C., & Balaram, P. (1990) *Int. J. Pept. Protein Res. 36*, 147–155.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Unanue, E. R., & Allen, P. M. (1987) *Science 236*, 551–557.
Vidusek, D. A., Roberts, M. F., & Goodfriend, L. (1985) *Biochemistry 24*, 2747.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
Zwollo, P., Ehrlich-Kautzky, E., Ansari, A. A., Scharf, S. J., Erlich, H. A., & Marsh, D. G. (1991) *Immunogenetics 33*, 141–151.

Kinetics and Mechanism of the Folding of Cytochrome *c*†

Kenneth M. Pryse,[‡,§] Thomas G. Bruckman,[§,∥] Bruce W. Maxfield,[⊥,#] and Elliot L. Elson*[,‡,§]
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, and Departments of Chemistry and Physics, Cornell University, Ithaca, New York 14853
Received January 7, 1992; Revised Manuscript Received March 25, 1992

ABSTRACT: The reversible folding of cytochrome *c* in urea at pH 4.0 was investigated by repetitive pressure perturbation kinetics and by equilibrium spectroscopic methods. Two folding reactions were observed in the 1 ms to 10 s time range. The rates and amplitudes of these reactions depend on urea concentration in a complex manner, which is different for each process. The absorbance spectra of the kinetic amplitudes of the two reactions also differ from each other. A model with a three-state mechanism can quantitatively account for all of the kinetic and equilibrium data, and it enables us to determine the rate constants and volume changes of the two steps. If a rapid protonation step is added to the mechanism, the analysis can be extended to calculate the pH dependence of the rate and amplitude of the faster folding step. This pH dependence is in excellent agreement with previously published data [Tsong, T. Y. (1977) *J. Biol. Chem. 252*, 8778–8780]. Kinetic experiments in the 695-nm band show clearly that the axial ligand methionine-80 is involved in the slow folding process and the other axial ligand, histidine-18, is involved in the fast process. Additional experiments with a cyanogen bromide fragment of the protein, and fluorescence detection of the folding kinetics of the intact protein, support an interpretation of the model in terms of known structural elements of cytochrome *c*. This work provides new information about the mechanism of the folding of cytochrome *c*, resolves conflicts in earlier interpretations, and demonstrates the applicability of the repetitive pressure perturbation kinetics method to protein folding.

The role of hydrophobic interactions in stabilizing the folded conformations of proteins is a long-standing problem of protein physical chemistry (Kauzmann, 1959; Privalov & Gill, 1988). The thermodynamic properties of these interactions are not yet well understood. Although solvent transfer experiments provide a useful model for the heat capacity changes associated with protein folding [e.g., Baldwin (1986)], they are less successful in accounting for the corresponding volume and compressibility changes [e.g., Kauzmann (1987)]. Nevertheless, the response of the protein conformation to changes in pressure, which is governed by the differences among the volumes of the conformational states, can usefully complement measurements of the response of the protein to temperature changes, which is governed by differences in enthalpy. Both enthalpy and volume changes are thought to be dominated by changes in interactions of water molecules with each other due to the exposure of amino acids to the solvent during unfolding (Edsall & McKenzie, 1983).

† This work was supported by grants from NIH (GM-38838) and the Lucille P. Markey Charitable Trust to E.L.E.
* Author to whom correspondence should be addressed at Department of Biochemistry and Molecular Biophysics, Box 8231, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110.
‡ Department of Biochemistry and Molecular Biophysics, Washington University.
§ Department of Chemistry, Cornell University.
∥ Current address: Betz Laboratories, Inc., Somerton Rd., Trevose, PA 19047.
⊥ Department of Physics, Cornell University.
Current address: Industrial Sensors, 400 Hester St., San Leandro, CA 94577.

0006-2960/92/0431-5127$03.00/0