Articles

Proton Resonance Assignments and Three-Dimensional Solution Structure of the Ragweed Allergen *Amb a* V by Nuclear Magnetic Resonance Spectroscopy[†,‡]

William J. Metzler,[*,§] Kathleen Valentine,[∥] Marianne Roebber,[⊥] 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 May 5, 1992

ABSTRACT: Essentially complete assignment of the proton resonances in the allergenic protein *Amb a* V has been made by analysis of two-dimensional NMR experiments. Conformational constraints were obtained 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 a* 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 small segment of antiparallel β-sheet, and several loops. A hydrophobic core exists at the interface of the α-helix and β-sheet. The derived structure accounts for the several anomalous proton chemical shifts that are observed. The structure determined here for *Amb a* V is topologically similar to the structure determined previously for the homologous allergenic protein *Amb t* V [Metzler, W. J., Valentine, K., Roebber, M., Friedrichs, M. S., Marsh, D., & Mueller, L. (1992) *Biochemistry 31*, 5117–5127]; however, significant differences exist in the packing of side chains in the hydrophobic core of the molecules. Comparison of the detailed structural features of these two proteins will allow us to suggest surface substructures for the *Amb* V allergens that are likely to participate in B cell epitopes.

The *Amb* V *Ambrosia* (ragweed) pollen allergens provide particularly good models in which to test the structure–function relationships for both B cell and MHC class II (Ia)/T cell epitopes in the human immune response. *Amb a* V and *Amb t* V are proteins isolated from the two most commonly occurring species of ragweed, *Ambrosia artemisiifolia* (short) and *Ambrosia trifida* (giant). Although these two proteins are over 45% homologous, the immune responses they elicit are quite distinct. *Amb a* V and *Amb t* V are virtually non-cross-reactive at the B cell (antibody) level (Roebber et al., 1985), suggesting that significant differences exist in their surface structures. In contrast, 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) (Marsh et al., 1982, 1990; Zwollo et al., 1991; Huang et al., 1991), suggesting that the two *Amb* V molecules possess similar MHC class II-binding epitopes. These epitopes are likely to be derived from evolutionarily conserved parts of the molecules. Structural information complementing the well-characterized immunological properties of these allergens has been lacking. Therefore, we have initiated a project aimed at determining the three-dimensional structures of several members of the *Amb* V family. As a first step, we determined the structure of *Amb t* V (Metzler et al., 1992); in this report, we describe the three-dimensional structure of *Amb a* V.

Amb a V is a single-chain polypeptide of 45 amino acids (Lapkoff & Goodfriend, 1974; Mole et al., 1975). This protein contains four disulfide bonds and exhibits the resistance to denaturation commonly associated with small disulfide-rich proteins (Lord et al., 1985). Although some structural features of *Amb a* V have been probed with fluorescence, phosphorescence (Galley et al., 1982), and laser Raman spectroscopy (Lord et al., 1985), little is known about the conformation of *Amb a* V in solution. The small size of *Amb a* V makes it particularly amenable for structural analysis by nuclear magnetic resonance (NMR) spectroscopy. One-dimensional ^{1}H NMR studies (Vidusek et al., 1985) have suggested that *Amb a* V folds into an ordered three-dimensional structure; however, their data were insufficient to experimentally determine this structure. In this paper, we have used two-dimensional ^{1}H NMR, in combination with computational techniques, to determine experimentally the three-dimensional structure of *Amb a* V in solution. The structure is compared with the structure previously determined for *Amb t* V (Metzler et al., 1992).

MATERIALS AND METHODS

Sample Preparation. Amb a V was isolated from *A. artemisiifolia* ragweed pollen (Greer Laboratories, Lenoir, NC) by methods previously described (Roebber et al., 1985; Metzler et al., 1992). For the NMR experiments in H_2O, 8.5 mg of highly purified *Amb a* V was dissolved in 600 μL of buffer (50 mM acetate, 0.1 mM EDTA, and 0.1 mM NaN_3, pH 5.0 at 23 °C, with 10% ^2H_2O included as an internal lock)

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

‡ Atomic coordinates for the derived *Amb a* V structures 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.

0006-2960/92/0431-8697$03.00/0

Table I: Structural Constraints for *Amb a* V

type of constraint	no. of constraints
interproton distances[a]	
total	385
intraresidue	145
sequential ($\|i-j\| = 1$)	102
medium range ($\|i-j\| < 5$)	35
long range ($\|i-j\| \geq 5$)	103
torsion angles[b]	
total	71
Φ	28
Ψ	21
χ_1	22
hydrogen bonds[c]	12
disulfide bonds	4

[a] Interproton distances are derived from nuclear Overhauser enhancements. *i* and *j* refer to the *i*th and *j*th residues in the molecule. [b] Torsion angles are derived from measured coupling constants (Φ) and from a systematic search of available conformational space (Ψ, χ_1). [c] Two constraints were use to constrain each of the six hydrogen bonds.

to yield a final concentration of 2.8 mM. 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 samples to a 5-mm NMR tube.

NMR Spectroscopy. The NMR experiments were performed on either a Varian Unity 600 spectrometer or a JEOL 500 spectrometer operating at 599.5 and 498.2 MHz, respectively. Data sets were acquired at temperatures of 20 and 30 °C. Readers are referred to Metzler et al. (1992) for details of data acquisition, data processing, resonance assignment, derivation of structural constraints, and structure calculations. In brief, 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. Sequence-specific resonance assignments were made by the standard method (Wuthrich, 1986). Nonsequential NOE cross-peaks were assigned as described by Metzler et al. (1992). Three types of constraints were derived from the NMR data: interproton distance constraints, torsion angle constraints, and hydrogen bond constraints. These are summarized in Table I. A complete listing of all constraints is provided as supplementary material. All structure calculations were performed in the program XPLOR on a SGI 4D/240VGX essentially as described (Nilges et al., 1988). Refined structures were analyzed for the best fit to the NMR constraints and the overall energy.

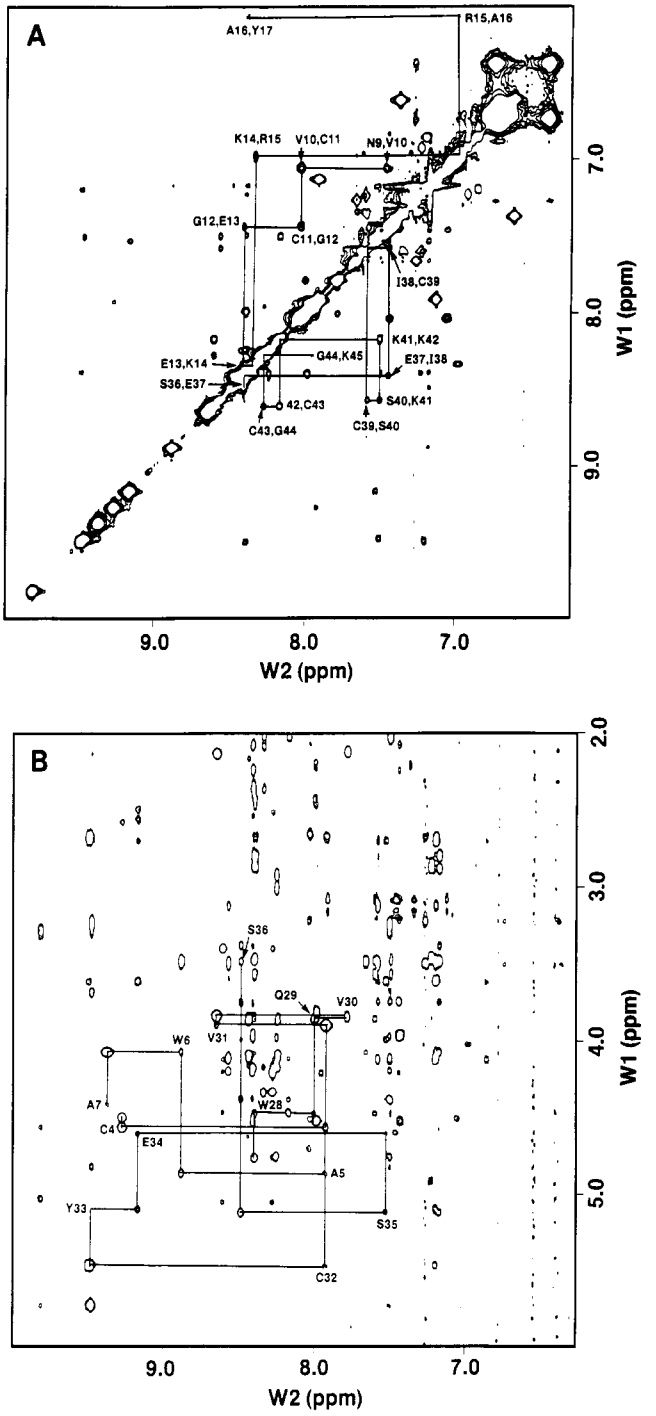

FIGURE 1: Expansion of the amide proton (A) and fingerprint (B) regions of the NOESY spectrum of *Amb a* V acquired with τ_{mix} = 0.32 s. Representative sequential connectivities are highlighted.

RESULTS

Resonance Assignments. The first step in making sequential resonance assignment was 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 were 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 three glycines, four valines, and two alanines. In addition, analysis of the NOESY spectrum enabled the aromatic protons of the three tyrosines and two tryptophans and the δ- and ϵ-amide protons of the asparagine and glutamine to be linked to their respective β-protons, reducing the degeneracy of the 18 AMX spin systems and providing additional markers in the sequence. A NOESY spectrum at 320 ms was used to locate the $d_{\alpha N}$, $d_{\beta N}$, and d_{NN} connectivities necessary to make interresidue connectivities. Figure 1 shows contour plots with examples 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 could be found except glycine 8, where no resonances could be found. Because mass spectral analysis confirmed the full-length *Amb a* V sequence (unpublished results), our inability to detect resonances for glycine 8 is likely to be a result of unfortunate

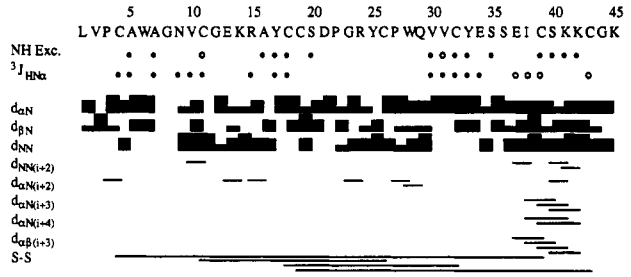

FIGURE 2: Summary of the sequential and short-range NOEs observed for *Amb a* V. Symbols: (●) residues with slowly exchanging amide protons or with a large (≥8.3 Hz) value for $^3J_{HN\alpha}$; (○) residues with a small (≤4.5 Hz) value for $^3J_{HN\alpha}$. For the sequential NOEs, the relative intensity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

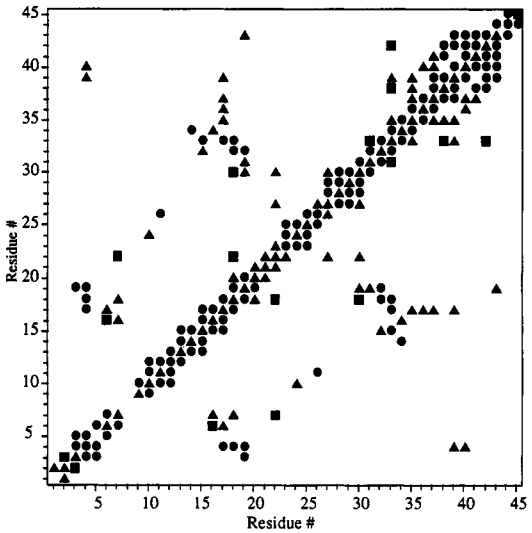

FIGURE 3: Diagonal plot of the NOEs observed for *Amb a* V. The presence of at least one NOE between protons of the two residues in the sequence position is indicated by (●) two backbone (HN or H_α) protons, (▲) one backbone proton and one side-chain proton, and (■) two side-chain protons. Preference is given to the symbol of the NOE with the highest number of backbone atoms. Data from 160-ms NOESY spectra are plotted.

resonance degeneracy and/or rapid exchange with solvent of its amide proton. Table II lists the proton assignments of *Amb a* V.

Structure Determination. Table I lists the type and number of constraints used in the structure calculations. A complete listing of all constraints is provided in the supplementary material. In addition, stereospecific assignments for the β-methylene protons of 11 residues and for the C_γ-methyl groups for three valines were made by analysis of the $^3J_{\alpha\beta}$ and NH–β/γ and α–β/γ NOEs (Hyberts et al., 1987; Zuiderweg et al., 1985). The stereospecifically assigned protons are indicated in Table II by a superscripted *R* or *S*. Figure 3 presents the observed NOEs in a diagonal plot, with residue numbers plotted on both axes. The stretch of NOEs running parallel to the diagonal in upper right corner of the plot suggests that presence of a C-terminal helix. The existence of this helix is confirmed by qualitative analysis of the medium-range NOEs (Figure 2), which shows a stretch of d_{NN}, $d_{\alpha\beta(i,i+3)}$, and $d_{\alpha N(i,i+3)}$ connectivities from residues 36 through 42. Additional evidence for the helix was provided by the small values for $^3J_{HN\alpha}$ and a stretch of slowly exchanging amide protons. Examination of initial refined structures indicated that the hydrogen-exchange data could only be explained by the presence of hydrogen bond between CO_i and NH_{i+4}, and therefore the helix is an α-helix. This conclusion is supported by the detection of $d_{\alpha N(i,i+4)}$ NOEs. All subsequent refinements included hydrogen bond constraints between the CO_i and NH_{i+4} of residues$_i$ 35 through 38.

The diagonal plot of Figure 3 also shows two stretches of NOEs running orthogonal to the diagonal, suggesting the presence of a triple-stranded β-sheet. Additional evidence for a β-sheet is provided by the measurement of large $^3J_{HN\alpha}$ coupling constants, short sequential $d_{\alpha N}$ distances, and the presence of the very slowly exchanging amide protons. Close contacts between α-protons remote in the primary structure indicated that the β-sheet is antiparallel. Inspection of the NOESY spectrum, however, indicated that some of the NOEs expected for a β-sheet were not present, and therefore hydrogen bonds were included as constraints only when the evidence for the acceptor carbonyl was completely unambiguous. This was the case for the slowly exchanging amides of residues 17 and 33; therefore, hydrogen bonds were included as constraints for these amides. No hydrogen bond constraints were included for the slowly exchanging amides of residues 5, 7, 18, and 20 even though reasonable guesses could be made as to their acceptor carbonyls because a sufficient number of corroborating NOEs could not be identified.

A set of 50 randomly generated starting structures was refined by dynamic simulated annealing with the program XPLOR. After refinement, 34 structures had no distance constraint violations greater than 0.5 Å. Nineteen of these structures only had two or fewer distance constraint violations greater than 0.3 Å with essentially no violations of the torsion angle constraints. An average refined structure was calculated from these 19 structures, and its energy was minimized in XPLOR. Superposition of the backbone atoms for the 19 final refined structures is shown in Figure 4. Structural statistics are given in Table III. Analysis of the root mean square distance (rmsd) deviations and torsion angle deviations suggests that the structures determined for *Amb a* V are well-defined by the NMR data. The average rms difference for the backbone atoms (N, C_α, C, O) from the energy-minimized average structure is 1.10 Å, and the average angular rms difference for the torsion angles Φ and Ψ is 17°. Figure 5 plots the atomic rms distance and torsion angle deviations of the 19 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 loop areas exhibit higher conformational variability.

Topologically, the structure of *Amb a* V (Figure 4; see also Figure 8) is quite similar to the structure of *Amb t* V previously determined (Metzler et al., 1992). *Amb a* V contains a C-terminal α-helix which is tethered to the N-terminal segment of the protein by two disulfide bonds. The N-terminal segment contains a small region of "β-sheet", with an irregular loop connecting the strands. Only two β-strands comprise the β-sheet of our current structure for *Amb a* V. These two β-strands are held together by a disulfide bond between Cys_{18} and Cys_{32}. An N-terminal β-strand may also be involved in the sheet, although it is less well aligned. The helix lies across the top half of the sheet, forming a hydrophobic core at their interface.

A superposition of the backbone atoms of the α-helix in the final refined structures is shown in Figure 6A. The backbone is well-defined, with an average pairwise rms distance deviation of 0.33 Å. The helix extends from residues 36 through 42. C-Terminal residues 44 and 45 lie in a more extended conformation and are not well-defined (Figure 5). A superposition of the sheet is shown in Figure 6B. Although the third region of β-strand (residues 4–7) does not form a neat

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

residue	NH	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	other
L1		4.13	1.73, 1.66		0.94	
V2	8.17	4.54	2.36	1.06^R, 0.96^S		
P3		4.28	3.15, 2.14	2.14, 1.85	3.90	
C4	9.16	4.36	3.39^S, 3.25^R			
A5	7.82	4.66	1.52			
W6	8.79	3.91	3.29, 3.28		7.08	$C_{\epsilon 1}H$ 10.80, $C_{\zeta 2}H$ 7.49, $C_{\eta 2}H$ 7.22, $C_{\epsilon 3}H$ 7.55, $C_{\zeta 3}H$ 7.16
A7	9.26	4.18	1.11			
G8						
N9	7.39	4.38	2.95^S, 2.88^R			$N_\delta H$ 7.81, 7.05
V10	6.95	4.29	1.87	0.89^S, 0.83^R		
C11	7.93	4.50	2.99^S, 2.45^R			
G12	7.32	4.29, 3.73				
E13	8.30	4.30	1.97, 1.51			
K14	8.23	4.17	1.89, 1.82	1.58, 1.49	1.70	$C_\epsilon H$ 3.82
R15	6.87	3.82	1.22^R, 1.01^S	2.10, 1.24	3.27, 3.00	$C_\epsilon H$ 7.14 $C_\eta H$ 6.63, 6.28
A16	5.97	3.07	1.02			
Y17	8.27	4.80	2.65, 2.59		7.07	$C_\epsilon H$ 7.07
C18	9.72	5.50	3.09, 3.05			
C19	9.35	4.61	3.46^S, 3.00^R			
S20	7.39	4.53	3.92^S, 3.76^R			
D21	8.15	4.55	2.77, 2.73			
P22		4.41	1.96, 1.88	2.01	4.02, 3.95	
G23	8.33	3.89, 3.65				
R24	8.14	3.96	1.16		2.98	$C_\epsilon H$ 6.99
Y25	8.30	4.70	3.24^S, 2.65^R		7.12, 7.10	$C_\epsilon H$ 6.82, 6.80
C26	7.88	4.30	4.11, 4.11			
P27		4.55	2.26, 2.23	2.02, 1.85	3.64, 3.58	
W28	8.29	4.26	3.35, 3.26		7.36	$C_{\epsilon 1}H$ 10.09, $C_{\zeta 2}H$ 7.47, $C_{\eta 2}H$ 7.23, $C_{\epsilon 3}H$ 7.49, $C_{\zeta 3}H$ 7.11
Q29	7.89	3.64	1.59^R, 1.34^S	2.02, 1.99		
V30	7.68	3.62	1.92	0.82, 0.84		
V31	8.54	3.68	1.93	0.96^S, 0.87^R		
C32	7.81	5.24	2.49, 2.46			
Y33	9.35	4.89	3.39^S, 2.49^R		7.10, 7.09	$C_\epsilon H$ 6.75, 6.74
E34	9.06	4.41	2.37, 2.29	2.54, 2.54		
S35	7.42	4.88	4.17, 3.79			
S36	8.38	3.28	3.56, 3.23			
E37	8.30	3.98	2.04, 1.88			
I38	7.33	3.53	1.59	1.50, 1.04	0.89	$C_\gamma H_3$ 0.57
C39	7.47	3.92	3.29^S, 2.98^R			
S40	8.46	3.99	3.91, 3.89			
K41	7.39	4.18	1.81, 1.62	1.51	1.45	$C_\epsilon H$ 3.11, 3.03
K42	8.06	4.26	1.71^S, 1.43^R	1.21, 1.08	1.49	$C_\epsilon H$ 2.93, 2.85 $C_\zeta H$ 7.57
C43	8.50	4.83	3.18, 3.17			
G44	8.17	4.05, 4.01				
K45	7.84	4.32	1.75, 1.40	1.89		$C_\epsilon H$ 3.00

[a] Assignments for *Amb a* V are in 15 mM acetate buffer (pH = 5.0) in D_2O at 30 °C. Amide assignments are in the same buffer in H_2O at 30 °C. *R* and *S* denote stereospecifically assigned protons.

FIGURE 4: Stereoview of the 19 refined structures of *Amb a* V. The structures were superimposed to minimize the rmsd of the backbone atoms (N, C_α, C, O) of residues 2–43. The α-helix (red), β-sheet (blue), and disulfide bonds (yellow) are highlighted.

secondary structure, we have included this strand in this figure of the sheet for those interested in comparing it with the sheet found in *Amb t* V (Metzler et al., 1992). Again, this region of *Amb a* V is well-defined, with a pairwise rms distance deviation of the backbone atoms for the 19 final refined structures of 0.64 Å.

There are four disulfide bonds holding the secondary structure units of *Amb a* V together. At the beginning of the structure calculations, the cystine partners in the disulfide pairs were unknown, although comparison with the homologous protein *Amb t* V suggested probable candidates. Spectroscopic evidence verifying the assignment of the partners in the di-

Table III: Structural Statistics

	⟨SA⟩	SA	$(SA)_m$	best
rmsd from experimental NOE constraints[a]				
all (385)	0.043	0.26	0.038	0.039
intraresidue (145)	0.050	0.31	0.049	0.048
sequential (102)	0.043	0.10	0.042	0.033
medium range (35)	0.035	0.026	0.015	0.026
long range (103)	0.031	0.34	0.015	0.024
no. of violations > 0.3 Å	2	2[b]	2	1
energy[c]				
E_{NOE} (kcal/mol)	45.8	1415	38.4	37.3
E_{TOR} (kcal/mol)	1.60	12.4	0.18	0.26
E_{REP} (kcal/mol)	17.2	$>10^6$	9.3	11.0
E_{L-J} (kcal/mol)	−150	$>10^6$	−158	−153
deviations from idealized covalent geometry				
bonds (Å)	0.006	0.006	0.006	0.006
angles (deg)	2.25	19.4	2.22	2.23
impropers (deg)[d]	0.96	1.52	0.83	0.94
rmsd for backbone (N, C_α, C, O) atoms	1.43	0.98	1.10	

[a] ⟨SA⟩ refers to the 19 individual final refined structures. SA is the structure obtained by averaging the coordinates for these 19 structures. $(SA)_m$ is the energy-minimized SA structure. Best refers to the individual refined structure which best fits the NMR and covalent geometry constraints.[b] The SA structure had 24 violations greater than 0.3 Å; however, all of these except two were violations of intraresidue constraints of long side chains which were poorly defined. [c] 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. [d] Impropers refer to planarity and chirality constraints.

sulfide linkages (Cys_{18}–Cys_{32}, Cys_{11}–Cys_{26}, Cys_4–Cys_{39}, and Cys_{19}–Cys_{43}) was provided by 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. A list of the NOEs used in confirming disulfide bonds is given in the supplementary material. Measurement of the Cys–Cys C_β distances in structures generated without any disulfide bond constraints confirmed the assignment of the Cys partners in disulfide linkages. Constraints for these disulfides were included in subsequent refinements.

Several of the *Amb a* V side-chain conformations are 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 6C shows an example of a superposition for the side chains of Tyr_{17}, Cys_{18}, Cys_{19}, Val_{30}, Val_{31}, Cys_{32}, Tyr_{33}, Ile_{38}, Cys_{39}, and Cys_{43}.

The structures determined for *Amb a* V fulfill several criteria which we currently use to judge the "goodness" of the structure. There are very few violations of the experimentally derived constraints. The structures do not deviate significantly from their idealized covalent geometry (Table III), indicating that satisfaction of the pseudo-energy potential for the NMR constraints has not led to unrealistic geometries. Calculation of the Leonard-Jones energies shows that they are both large and negative (Table III), 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. Ramachandran plots (Figure 7) illustrate that the Φ and Ψ torsion angles for the structures lie in sterically allowed regions. The Φ torsion angle for Cys_{26} is unusual in that it resides in positive Φ space. A similar conformation was found for the Φ torsion angle for Cys_{26} in *Amb t* V and 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. Finally, good agreement is found between experimental NOESY data and NOESY data calculated for the energy-minimized average structure.

DISCUSSION

Anomalous Chemical Shifts. The chemical shift of a proton resonance is highly dependent upon the proton's local electronic environment. Unusual chemical shifts provide sensitive probes of a protein's structure; the derived structure should be able to account for any anomalies observed in the chemical shift. The proton chemical shift assignments of *Amb a* V given in Table II indicate that the resonance frequencies of a number of the protons deviate significantly from expected values. Several of these chemical shift anomalies are uninteresting, merely reflecting ring current effects from sequential aromatic residues. However, the chemical shifts of Trp_6 $C_\epsilon H$, Ala_{16} NH and $C_\alpha H$, and Ile_{38} $C_\gamma H_3$ are affected by residues remote in the primary structure of *Amb a* V, and the origin of their unusual chemical shifts points to some interesting features of the molecule.

The chemical shifts for the $C_\epsilon H$ protons of the two tryptophans of *Amb a* V are quite different, suggesting that their aromatic rings lie in distinct environments. The structure determined for *Amb a* V confirms this. Trp_{28} is fully exposed to solvent and, as such, possesses proton chemical shifts close to the value expected for a tryptophan in a random coil. Conversely, while on the surface of the molecule, Trp_6 is packed next to several hydrophobic residues. The $C_\epsilon H$ of Trp_6 lies in the plane of the aromatic ring of Tyr_{17}, resulting in a large downfield shift of this proton's resonance frequency. Previously, other groups probed the conformation of these two tryptophans. On the basis of solvent relaxation effects in fluorescence spectra and heterogeneity in low-temperature phosphorescence spectra, Galley et al. (1982) proposed that the two tryptophans of *Amb a* V lie in different environments on the surface of the molecule. On the basis of laser Raman spectra, Lord et al. (1985) suggested that at least one of the tryptophans is interacting with hydrophobic residues. Both of these results are consistent with the structure we have determined for *Amb a* V.

The amide nitrogen proton of Ala_{16} resonates at 5.97 ppm, over 2.2 ppm upfield of its expected random coil value. Likewise, the $C_\alpha H$ proton of Ala_{16} is shifted 1.3 ppm upfield. Such upfield shifts can often be attributed to shielding of the

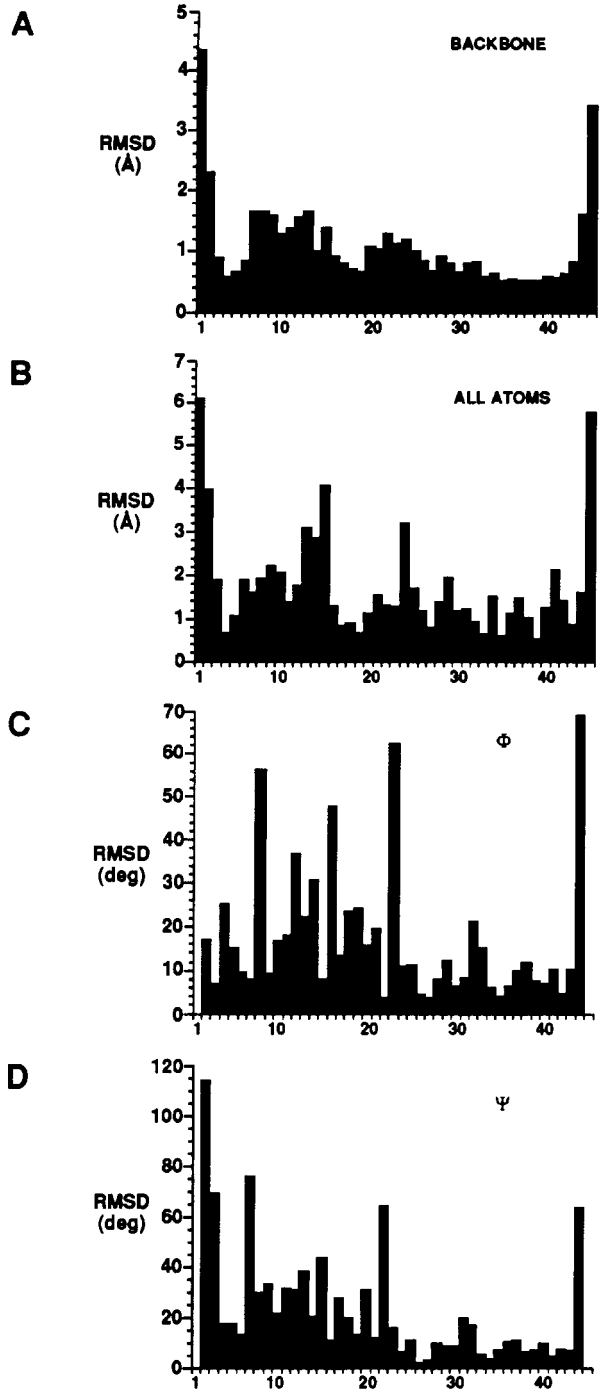

FIGURE 5: Plots of the average pairwise rmsd of the 19 final refined structures of *Amb a* V vs sequence position: (A) rmsd in atomic position for the backbone atoms (N, C_α, C, O); (B) rmsd in atomic position for all atoms; (C) rmsd in angular distribution of torsion angle Φ; (D) rmsd in angular distribution of torsion angle Ψ. The atomic rmsds were determined by superimposing the 19 structures to minimize the pairwise rmsd of the backbone atoms of residues 2–44 and then calculating the appropriate rmsd for each residue.

proton due to ring current effects. Examination of the *Amb a* V structure indicates that these protons of Ala_{16} are nestled between two aromatic rings, Trp_6 and Tyr_{17}, accounting for the observed chemical shift anomalies. Examination of the *Amb a* V structure would also predict that the methyl protons of Ala_{16} should absorb the largest effects of the aromatic ring currents. This is not observed in the NMR spectra; thus the actual orientation of Ala_{16} with respect to the above aromatic rings will be slightly different than that we have found. The likely cause for this minor discrepancy is dynamics, which is differentially effecting the relative intensities of NOEs to the Ala_{16} amide and methyl protons. Because of the proximity of the Ala_{16} NH to the two rings and the sensitivity of chemical shift to environment, small fluctuations in the position of the NH relative to the rings (or the converse) can have large effects on its chemical shift. The resultant averaging of the chemical shift can lead to a decrease in the observed resonance intensity for the amide proton and explains the reduced intensity observed for the Ala_{16} NH. More importantly, this averaging would also lead to a large reduction in the intensities of any NOE to the Ala_{16} NH and, thus, an underestimation of the closeness of the NH to the rings. The intensities of the observed NOEs to the Ala_{16} methyl protons, which on the basis of its ring current shift is expected to be further from the rings, would not be reduced by the effects of chemical shift averaging. The result is that, solely on the basis of the NOE data, the methyl group appears to be closer to the aromatic rings than does the NH, leading to the minor distortion of the structure in this region. We are currently performing chemical shift calculations and examining the dynamics of *Amb a* V to obtain a better understanding of the structural details of these residues.

The resonance of Ile_{38} $C_\gamma H_3$ is also shifted to higher field, the result of being juxtaposed to the center of the aromatic ring of Tyr_{33}. Interestingly, although these two residues are uniquely conserved in the homologous protein *Amb t* V, the corresponding $C_\gamma H_3$ in *Amb t* V (Ile_{34}) is *not* shifted upfield. Instead, it is the C_β protons of Lys_{32} of *Amb t* V which feel the affects of the tyrosine ring currents (Metzler et al., 1992). This result indicates that the packing of the hydrophobic core in *Amb a* V is different than it is in *Amb t* V.

Examination of Hydrogen-Exchange Data. The presence of slowly exchanging amide protons provides evidence for a stable protein structure. Secondary structural features such as α-helices and β-sheets involve networks of hydrogen bonding which protect amide protons from exchange with the solvent (Englander & Kallenbach, 1984). Amide protons also can be made inaccessible to solvent by the tertiary structure of a molecule. Thus the observation of amide protons with hindered exchange with solvent should be consistent with the derived molecular structure. Several slowly exchanging amide protons were detected 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). Secondary structural features account for the hindered exchange rates of the amide protons of residues 39 through 42 (α-helix) and residues 17 and 33 (β-sheet). Extension of the β-sheet two more residues with a bulge at residue 19 would account for the exchange rates of residues 20 and 31. Slowly exchanging amides are also observed for residues 5, 7, 16, and 18. These residues may form the third strand of β-sheet in *Amb a* V. However, because a sufficient number of unambiguous NOEs could not be found to define uniquely the hydrogen bond acceptors, no hydrogen bond constraints for these amide protons were included in structure calculations. Addition of hydrogen bond constraints for these amide protons results in structures which remain consistent with the NOE data and are much better defined (data not shown), but we believe the inclusion of such constraints would be an overinterpretation of our NMR data. The result of this is that the third strand is not well aligned with the remaining two strands of sheet. The amides of residues 11, 30, 32, and 35 are also slowly exchanging. The carbonyl oxygen of residue 27 lies in close proximity (1.9 Å in the minimized average structure) to the amide proton of residue 30 and, thus, is the likely candidate for a hydrogen

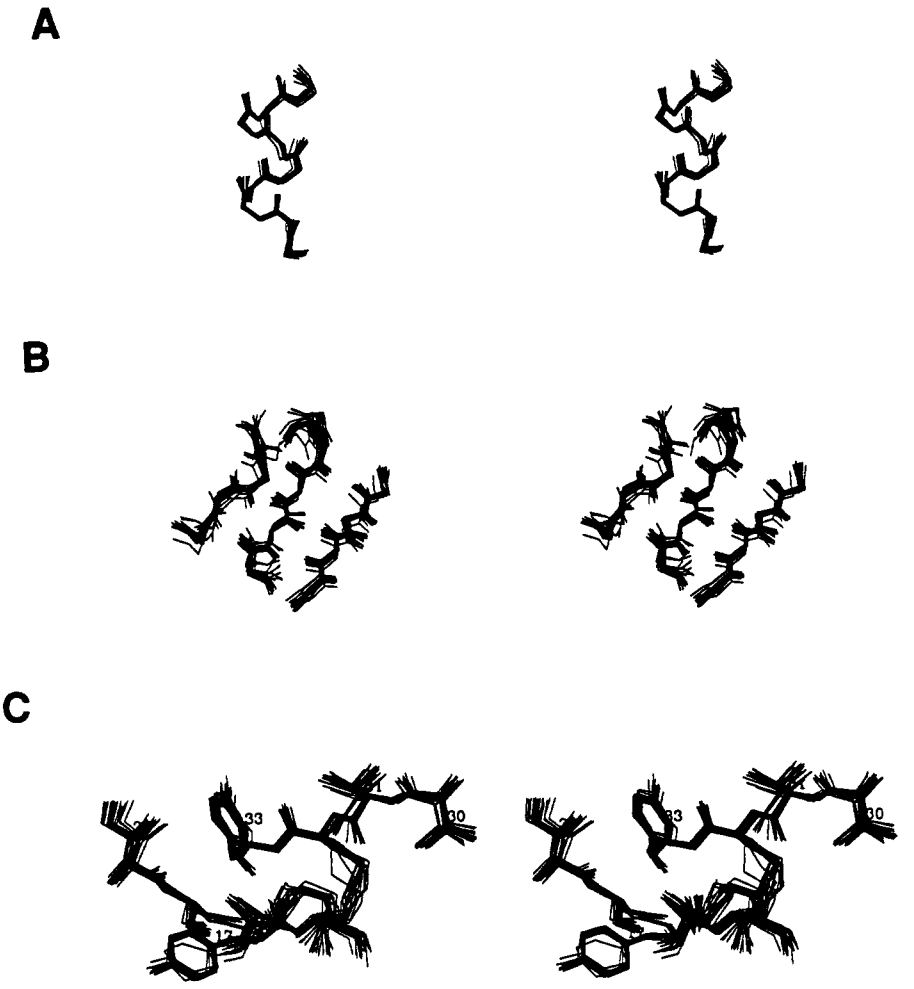

FIGURE 6: Stereoviews of the (A) helix, (B) β-sheet, and (C) some side chains of the 19 refined structures of *Amb a* V. The structures were superimposed by minimizing the rmsd of the backbone atoms (N, C_α, C, O) of the regions of residues 35–43 for the helix and residues 4–7, 17–20, and 31–33 for the β-sheet. For the side chains, the structures were superimposed by minimizing the rmsd of all atoms in residues 17–19, 30–33, 38–39, and 43.

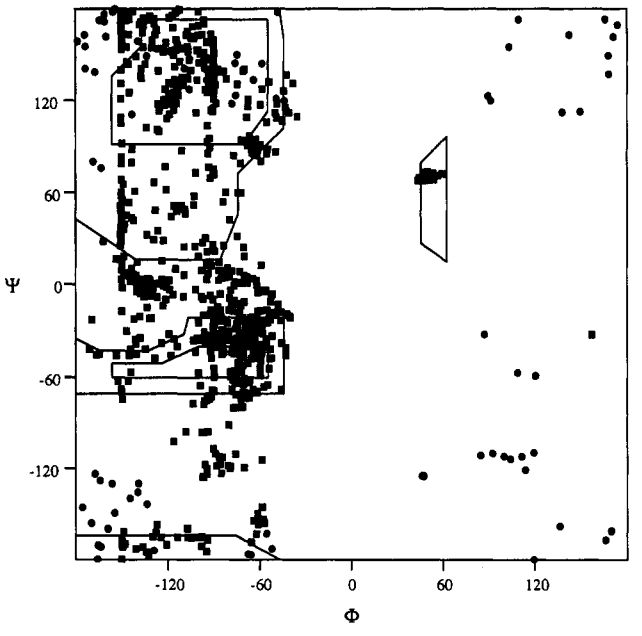

FIGURE 7: Ramachandran plot for the 19 refined structures of *Amb t* V. Each amino acid is labeled with its residue number. Symbols: (●) glycines; (■) all other amino acids.

bond acceptor. The amide protons of residues 11 and 35 are both inaccessible to solvent, which likely accounts for their diminished exchange rates. The amide proton of Cys_{32}, however, is completely accessible to solvent, and the derived structure of *Amb a* V provides little insight into the cause of its slow exchange rate. A search for potential hydrogen bond acceptors reveals at least one possible candidate, the side-chain amide oxygen of Gln_{29}. Although not currently in position to form a hydrogen bond, the side chain can be rotated to place it within reasonable hydrogen-bonding distance. The suggestion of Gln_{29} being the hydrogen bond acceptor for the amide proton of Cys_{32} is only speculation. It is interesting to note, however, that *Amb t* V contains a four-residue deletion with respect to the *Amb a* V sequence which removes this Gln. No slowly exchanging amide is observed for the corresponding Cys amide proton of *Amb t* V.

Comparison of the Amb a V and Amb t V Structures. Topologically, the structure determined for *Amb a* V is quite similar to that determined previously for *Amb t* V. Both proteins contain a C-terminal α-helix which lies atop a small section of antiparallel β-sheet. However, in contrast to *Amb t* V in which the β-sheet is comprised of three well-characterized strands, only two strands of the β-sheet form a neat secondary structure in *Amb a* V. The third strand is less well aligned. It is noteworthy that the analysis of laser Raman spectra has led Lord et al. (1985) to suggest that *Amb a* V would have less β-structure than *Amb t* V.

The interface of the helix and sheet forms a hydrophobic core which is held together with disulfide bonds. As mentioned above, observable differences exist in the packing of the hydrophobic residues of this core, particularly in the orientation

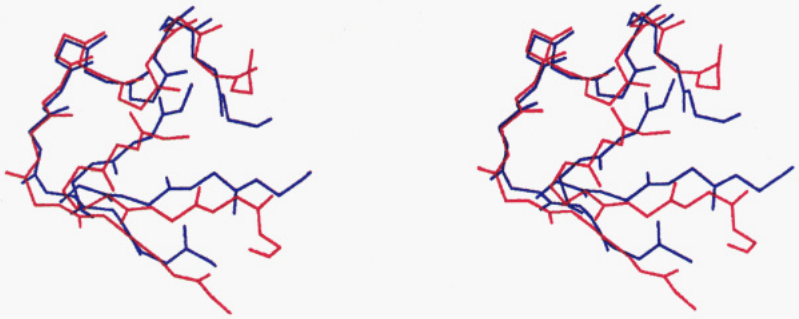

FIGURE 8: Stereoview of the superposition of *Amb a* V (red) and *Amb t* V (blue). The superposition was performed to minimize the rmsd between the backbone N, C_α, C, and O atoms of residues 4–7, 17–20, and 31–43 (*Amb a* V numbering scheme).

of Tyr_{33} (*Amb a* V numbering). Nevertheless, superposition of *Amb a* V on *Amb t* V illustrates the structural similarity of these two proteins (Figure 8). The rmsd deviation for the backbone atoms (N, C_α, C, O) of the helix and sheet regions (residues 17–20 and 31–43 in *Amb a* V) is 1.45 Å and rises to only 1.86 Å if the third strand (residues 4–7 in *Amb a* V) is included. The most significant deviation between the two structures is in the region of residues 21–30. However, this is not surprising because this region contains a four-residue insertion (residues 27–30) in *Amb a* V which is not present in *Amb t* V. A detailed comparison of the structures of *Amb a* V and *Amb t* V will be made after direct relaxation matrix (Yip & Case, 1991) and chemical shift (Osapay & Case, 1991) refinement of the structures is performed.

Implications for Epitopes. The small size of the *Amb* V molecules (ca. 40–45 amino acids) limits the number of possible epitopes available for recognition by the immune system; thus these proteins provide good model systems for exploring the structural features important in the human immune response. Immunological studies have shown that there is concordance in the immune recognition of *Amb a* V and *Amb t* V by the same MHC class II molecule, suggesting that the two *Amb* V molecules possess similar binding agretopes (Huang & Marsh, 1991). By synthesizing peptide analogues (Ala was substituted for Cys to facilitate synthesis) spanning two regions of *Amb a* V and measuring their ability to inhibit T cell responses to the native *Amb a* V, Huang and Marsh (1991) showed that a peptide containing residues 31–44 binds to the *Amb a* V-specific MHC class II molecule, HLA–DR-(α,β_1*1501). This is the region of the protein that is structurally the most conserved between *Amb a* V and *Amb t* V. Examination of the structure for *Amb a* V reveals that residues 35–42 of this peptide form an amphipathic helix in the *native* structure. The synthetic Cys → Ala substitutions would increase the helicity of this region. These results are consistent with the models proposed by the Berzofsky (DeLisi & Berzofsky, 1985; Margalit et al., 1987) and Rothbard groups (Hill et al., 1991; Rothbard & Gefter, 1991), who suggest that peptide epitopes with the propensity to form helices tend to favor T cell recognition. However, recent crystallographic data of a related MHC class I molecule have shown that peptides are bound in an extended conformation. Thus it appears that the helix propensity that has been correlated with T cell recognition is likely to be important at a stage of immunorecognition prior to the actual binding of the peptide by the MHC molecule. For example, secondary structural features of the native antigen, such as helical conformations, may cause the peptide to have a decreased capacity for being degraded into peptides and thereby increase the probability that it will be presented to the T cell receptor. Further experiments are in progress to identify more precisely the sites interacting with the MHC and T cell receptor molecules.

In contrast to the results for MHC class II-binding regions, *Amb a* V and *Amb t* V show virtually no antibody cross-reactivity, suggesting that the binding regions for B cell receptors and antibodies are dissimilar (Roebber et al., 1985). This is not surprising when one considers that antibody binding tends to be a surface property of the antigen, and although *Amb a* V and *Amb t* V are 45% homologous, most of the conserved residues are located in the interior of the molecule, being involved in disulfide bonds or hydrophobic interactions.

The NMR structural data we report here should be valuable in identifying the nature and location of the *Amb* 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

Four tables of 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, and disulfide bond constraints (12 pages). Ordering information is given on any current masthead page.

REFERENCES

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

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.

De Lisi, C., & Berzofsky, J. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7048.

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.

Galley, W. C., Williams, R. E., & Goodfriend, L. (1982) *Biochemistry 21*, 378–383.

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.

Hill, C. M., Hayball, J. D., Allison, A. A., & Rothbard, J. B. (1991) *J. Immunol. 147*, 189.

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.

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.

Lord, R. C., Petsko, G. A., Seaton, B. A., & Goodfriend, L. (1985) *Spectrochim. Acta 41*, 199–203.

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

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.

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.

Metzler, W. J., Valentine, K., Roebber, M., Friedrichs, M. S., Marsh, D. G., & Mueller, L. (1992) *Biochemistry 31*, 5117–5127.

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.

Nilges, M., Gronenborn, A. M., & Clore, G. M. (1988) *FEBS Lett. 239*, 129–136.

Osapay, K., & Case, D. A. (1991) *J. Am. Chem. Soc. 113*, 9436–9444.

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.

Roebber, M., Klapper, D. G., Goodfriend, L., Bias, W. B., Hsu, S. H., & Marsh, D. G. (1985) *J. Immunol. 134*, 3062–3069.

Rothbard, J. B., & Gefter, M. (1991) *Annu. Rev. Immunol. 9*, 527.

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.

Yip, P., & Case, D. A. (1989) *J. Magn. Reson. 83*, 643–648.

Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601–611.

Zwollo, P., Ehrlich-Kautzky, E., Ansari, A. A., Scharf, S. J., Erlich, H. A., & Marsh, D. G. (1991) *Immunogenetics 33*, 141–151.