Huxley, H. E. (1972) *Cold Spring Harbor Symp. Quant. Biol. 37*, 361–376.
Ishii, Y., & Lehrer, S. S. (1984) *Biophys. J. 45*, 392a.
Ishiwata, S. (1978) *Biochim. Biophys. Acta 534*, 350–357.
Jost, P. C., & Griffith, O. H. (1978) *Methods Enzymol. 49*, 369–418.
Lamkin, M., Tao, T., & Lehrer, S. S. (1983) *Biochemistry 22*, 3053–3058.
Lehrer, S. S. (1978) *J. Mol. Biol. 118*, 209–226.
Lehrer, S. S., & Morris, E. P. (1982) *J. Biol. Chem. 257*, 8073–8080.
Lehrer, S. S., Graceffa, P., & Betteridge, D. (1981) *Ann. N.Y. Acad. Sci. 366*, 285–299.
McLachlan, A. D., & Stewart, M. (1975) *J. Mol. Biol. 98*, 293–304.
Nauss, K. M., Kitagawa, S., & Gergely, J. (1969) *J. Biol. Chem. 244*, 755–765.
Parry, D. A. D., & Squire, J. M. (1973) *J. Mol. Biol. 75*, 33–55.
Pato, M. D., & Smillie, L. B. (1978) *FEBS Lett. 87*, 95–98.
Phillips, G. N., Jr. (1984) *Biophys. J. 45*, 392a.
Phillips, G. N., Jr., Lattman, E. E., Cummins, P., Lee, K. Y., & Cohen, C. (1979) *Nature (London) 278*, 413–417.
Phillips, G. N., Jr., Fillers, J. P., & Cohen, C. (1980) *Biophys. J. 32*, 485–500.
Potekhin, S. A., & Privalov, P. L. (1978) *Biofizika 23*, 219–223.
Satoh, A., & Mihashi, K. (1972) *J. Biochem. (Tokyo) 71*, 597–605.
Spudich, J. A., & Watt, S. (1971) *J. Biol. Chem. 246*, 4866–4871.
Tanaka, H. (1972) *Biochim. Biophys. Acta 285*, 459–466.
Tao, T., & Lamkin, M. (1984) *FEBS Lett. 168*, 169–173.
Tawada, Y., Ohara, H., Ooi, T., & Tawada, K. (1975) *J. Biochem. (Tokyo) 78*, 65–72.
Taylor, K. A., & Amos, L. A. (1981) *J. Mol. Biol. 147*, 297–324.
van Eerd, J.-P., & Kawasaki, Y. (1973) *Biochemistry 12*, 4972–4980.
Vibert, P., & Craig, R. (1982) *J. Mol. Biol. 157*, 299–319.
Wakabayashi, T., & Toyoshima, C. (1981) *J. Biochem. (Tokyo) 90*, 683–701.
Wakabayashi, T., Huxley, H. E., Amos, L. A., & Klug, A. (1975) *J. Mol. Biol. 93*, 477–497.
Weeds, A. G., & Pope, B. (1977) *J. Mol. Biol. 111*, 129–157.
Wegner, A., & Walsh, T. P. (1981) *Biochemistry 20*, 5633–5642.
Woods, E. F. (1969) *Int. J. Protein Res. 1*, 29–43.
Woods, E. F. (1976) *Aust. J. Biol. Sci. 29*, 405–418.
Woods, E. F. (1977) *Aust. J. Biol. Sci. 30*, 527–542.
Wu, C.-S. C., & Yang, J. T. (1976) *Biochemistry 15*, 3007–3014.
Zeidan, H., Watanabe, K., Piette, L. H., & Yasunobu, K. T. (1980) *J. Biol. Chem. 255*, 7621–7626.

500-MHz ^{1}H NMR Studies of Ragweed Allergen Ra5

David A. Vidusek and Mary F. Roberts*
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Lawrence Goodfriend
Department of Medicine, McGill University, Montreal, Quebec, Canada
Received September 4, 1984

ABSTRACT: The solution conformation of short ragweed allergen Ra5, a protein of 45 amino acid residues cross-linked with four disulfide bridges, has been investigated by ^{1}H NMR spectroscopy at 500 MHz. The aromatic region, which contains resonances from three tyrosines and two tryptophans, has been partially assigned. Two tyrosines titrate with a pK of 10.2; a third tyrosine is buried under the tryptophan resonances, and its pK could not be determined. The two tryptophans reside in different microenvironments; the resonances of one are very similar to those found in random coil structures while the other has dramatically shifted peaks. Nuclear Overhauser effect (NOE) difference spectroscopy is used to define two distinct spin-diffusion systems for the aromatic residues and to further identify several methyl-containing amino acids involved in these systems. Assignments in the methyl region are based on selective decoupling, chemical shifts, NOE difference spectra, and 2-D J-resolved and 2-D J-correlated spectroscopy (COSY) methodology. A unique ring-current-shifted methyl doublet in the Ra5 spectrum titrates into the bulk methyl region with a pK of 10.2. Examination of the COSY map suggests that this resonance belongs to either leucine-1 or isoleucine-38. Chemical removal of the N-terminal leucine did not affect the ring-current-shifted methyl. Therefore, this unique resonance has been assigned to the methyl of isoleucine-38. With this assignment and the spin-diffusion behavior of the aromatic residues, it is possible to suggest disulfide assignments as well as specific structural features of Ra5 consistent with the toxin–agglutinin fold proposed by Drenth and co-workers [Drenth, J., Low, B. W., Richardson, J. S., & Wright, C. S. (1980) *J. Biol. Chem. 255*, 2652–2655].

In recent years there has been an increasing interest in the structure and activity of environmentally derived proteins that on inhalation or injestion provoke immediate allergic reactions in predisposed individuals [for reviews, see Marsh (1975), King (1976), and de Weck (1977)]. For the North American population, the most important single source of these provocative proteins, and the most studied, is the pollen of *Ambrosia elatior*, commonly known as short ragweed. This pollen

0006-2960/85/0424-2747$01.50/0

Table I: Amino Acid Sequence[a] and Sequence-Based Predictions[b] of α-Helical, β-Strand, and β-Turn Conformations of Ra5

no.	residue	conformation	no.	residue	conformation
1	Leu	–[c]	24	Arg	–
2	Val (Leu)	β-turn	25	Tyr	–
3	Pro	(P_t =	26	Cys	–
4	Cys	1.3 ×	27	Pro	–
5	Ala	10^{-4})	28	Trp	
6	Trp	–	29	Gln	internal
7	Ala	–	30	Val	β-strand
8	Gly	–	31	Val	($[A]_{p=2}$
9	Asn	β-turn	32	Cys	$[H_L]_{n=4}$
10	Val	(P_t =	33	Tyr	$[A]_{q=1}$)
11	Cys	1.4 ×	34	Glu	
12	Gly	10^{-4})	35	Ser	β-turn
13	Glu	–	36	Ser	(P_t =
14	Lys	–	37	Glu	1.7 ×
15	Arg	internal	38	Ile	10^{-4})
16	Ala	β-strand	39	Cys	
17	Tyr	($[A]_{p=2}$	40	Ser	
18	Cys	$[H_L]_{n=3}$	41	Lys	α-helix
19	Cys	$[A]_{q=0}$)	42	Lys	(Z frag-
20	Ser	–	43	Cys	ment)[d]
21	Asp	–	44	Gly	
22	Pro	–	45	Lys	
23	Gly	–			

[a] Mole et al., 1975. [b] α-Helical and β-strand conformations predicted by the method of Lim (1974) and β-turns by the method of Chou & Fasman (1974). Terminology follows the respective authors. [c] Dash denotes residues not assigned within the ordered conformations considered here. These include β-turns initiated at Cys-19, Asp-21, and Gly-23. Although predicted with high probability (P_t = 2.1 × 10^{-4}, 5.5 × 10^{-4}, and 1.6 × 10^{-4}, respectively), these turns were not utilized in molecular modeling of Ra5 (see Discussion) in favor of a crossover loop, Ser-20 to Pro-27, joining the two predicted β-strands and conforming to the corresponding regions of the toxin–agglutinin fold (Drenth et al., 1980). [d] The x fragment Cys-39 to Cys-43 was extended to the C-terminus to form a Z fragment in view of the presence of (1–5) electrostatic pair Lys-41 to Lys-45.

Table II: Amino Acid Compositions of Native and Transaminated Ra5[a]

	no. of residues/mol of Ra5[b]	
amino acid	native Ra5	transaminated Ra5
cysteic acid	0.04	0.09
Asp	1.74 (2)[c]	1.97
Ser	3.16 (4)	3.75
Glu	3.24 (4)	4.41
Pro	2.24 (3)	2.06
Cys–Cys	2.82 (4)	3.00
Gly	3.48 (4)	4.31
Ala	3.00 (3)	3.00
Val	2.47 (3.7)	2.25
Ile	0.69 (1)	0.75
Leu	0.89 (1.3)	0.19
Tyr	2.41 (3)	2.25
Lys	3.38 (4)	3.75
Arg	1.68	1.55
Trp	*d* (2)	*d*

[a] Based on analysis of single 24-h acid hydrolysates of unoxidized samples in a Beckman 120C with Model 126 data system. [b] Assuming 3 Ala/mol of Ra5. [c] Values in parentheses give the number of residues in Ra5 by sequence determination (Mole et al., 1975). [d] Not determined.

contains numerous allergens (Løwenstein & Marsh, 1981), some of which have been isolated to homogeneity for study of their chemical and immunological properties (Marsh & Goodfriend, 1976; King, 1976; Løwenstein et al., 1981), and of these, protein Ra5 (Lapkoff & Goodfriend, 1974) is of particular interest. Owing to its low molecular weight (M_r 5000) and restricted incidence of activity (10–15%) within the ragweed allergen sensitive population, this protein has emerged as a fecund reagent for study of the genetic and immunochemical basis of the allergic and immune response in man (Marsh et al., 1975, 1982; Goodfriend, 1976; Bias et al., 1978).

Apart from its amino acid sequence (Mole et al., 1975), little is known about the Ra5 structure. The protein consists of a single polypeptide chain of 45 amino acid residues with no detectable carbohydrate or lipid. As shown in Table I, the chain sequence is homogeneous except at the second N-terminal position due to a 30% valine to leucine substitution. The major and minor variants appear to be allergenically equivalent since the major form prepared by peptide synthesis is equiactive on a weight basis with native Ra5 (Choudhury & Goodfriend, 1983). Table I also shows results (L. Goodfriend, unpublished data) of Ra5 sequence analysis by sequence predictive methods (Lim, 1974; Chou & Fasman, 1974), which suggest that 70% of the amino acid residues participate in ordered conformation. The most striking feature of Ra5 is the presence of 8 half-cystines in a sequence of 45 residues: these form 4 intrachain disulfide links, but the Cys–Cys pairing has not been specified. From the close correspondence in sequence position of the half-cystines in Ra5, in one of the domains of wheat germ agglutinin, and in erabutoxin, Drenth et al. (1980) have postulated a common toxin–agglutinin fold for these proteins. Evaluation of this conformational feature for Ra5 is difficult, however, in the absence of information concerning conformational features of the molecule. Fluorescence and low-temperature phosphorescence studies (Galley et al., 1978) of Ra5 suggest microenvironmental differences between the two tryptophans in the protein.

We have initiated a 500-MHz ^{1}H NMR study of Ra5 by examining distinctive resonances of the protein as a function of pH and using NOE[1] difference spectroscopy and various two-dimensional techniques for residue identification. The ^{1}H NMR parameters provide useful information for specific residues, which is used to derive secondary (disulfide cross-links) and tertiary structural features of the molecule. This structural information, coupled with the simple predictive methods of Table I, permits evaluation of the agglutinin–toxin fold for Ra5 postulated by Drenth et al. (1980).

Experimental Procedures

Ra5 was isolated from short ragweed pollen (Greer, NJ) and purified to disc electrophoretic homogeneity as described by Lapkoff & Goodfriend (1974) with minor modifications. Des-Leu^1-Ra5 was prepared by transamination (Dixon & Field, 1972) according to Webster & Offord (1972) with minor modifications (L. Goodfriend, unpublished experiments). Transamination led to an 80% reduction in Leu content with no major changes in other amino acid residues (Table II). Assuming an actual Leu content of 1.3 residues/mol of Ra5 due to the proportion of the Leu-Val and Leu-Leu Ra5 variants (Lapkoff & Goodfriend, 1974; Mole et al., 1975), the reduction indicated virtually complete removal of the N-terminal Leu. This was confirmed by Edman degradation (Edman, 1970) and reverse-phase HPLC analysis of the PTH's, which gave as N-terminal residues only Leu before and Val (with trace Leu) after transamination of Ra5. It should be noted that with HPLC these PTH derivatives are well separated (several minutes difference in retention times). For NMR studies, the salt-free proteins were dissolved in D_2O (to

[1] Abbreviations: NOE, nuclear Overhauser effect; DSS, 3-(trimethylsilyl)propanesulfonic acid sodium salt; TSP, sodium 3-(trimethylsilyl)tetradeuteriopropionate; FID, free induction decay; 2-D, two dimensional; COSY, 2-D NMR homonuclear *J*-correlated spectroscopy; HPLC, high-pressure liquid chromatography; PTH, phenylthiohydantoin.

0.5 mM; at higher concentrations the Ra5 tends to aggregate), lyophilized, and reconstituted with 99.996% D_2O (Aldrich). The apparent pH was adjusted by the addition of NaOD or DCl; no buffers were used. The pH values are reported without correction for isotope effects.

^1H NMR spectra were recorded at 500 MHz by using the home-built spectrometer and data-reduction system of the Francis Bitter National Magnet Laboratory, M.I.T., Cambridge, MA. Between 40 and 1000 scans were accumulated for each spectrum with a repetition rate of 2.5 s and a flip angle of 90°. Resonance positions were measured from both the internal DSS peak and the HDO peak, assumed to be 0.004 and 4.67 ppm from TSP, respectively. Resolution enhancement, achieved by convolution difference of the FID (Campbell et al., 1973), used a broad multiplication factor of 3 Hz and a narrow factor of 1 Hz. NOE difference spectra were acquired with similar parameters but with a repetition rate of 5 s. The resonance of interest was presaturated for 300 ms with the appropriate radio frequency prior to acquisition. The power level was adjusted to just null the upfield Tyr doublet under these conditions. Frequency selectivity was such that irradiation 75 Hz away from upfield tyrosine doublets yielded a null difference spectrum. Irradiation 25 Hz (the chemical shift difference of the two upfield tyrosines) to the high-field side of the upfield tyrosine doublet caused at most a 10% change in the intensity of that tyrosine. Any NOEs observed for resonances close to the irradiation frequency were adjusted for this effect. The Fourier transformed difference spectra were generated by subtracting the FID of a control proton acquisition that was irradiated in a region with no resonances from the FID resulting from the presaturated acquisition. An exponentiation factor of 1 Hz was applied to the FID, which was then Fourier transformed and phased. Spin–lattice relaxation times, T_1 values, for several resonances were determined by inversion–recovery (Vold et al., 1968).

The two-dimensional *J*-correlated spectroscopy (COSY) and *J*-resolved experiments are relatively standard and have been described in detail elsewhere (Aue et al., 1976; Bax, 1982). The two-dimensional homonuclear *J*-correlated experiment that was utilized is phase sensitive with the following pulse sequence: T–$(\pi/2)_x$–t_1–$(\pi/2)_x$–ACQ. The pulse sequence used for the two-dimensional *J*-resolved experiment is T–$(\pi/2)_x$–$t_{1/2}$–$(\pi)_y$–$t_{1/2}$–ACQ. Data reduction of the two-dimensional data set used a Gaussian window function to enhance resolution.

Molecular modeling of Ra5 was conducted on a Vector General Inc. Model 3404 graphics display system attached to a Digital Equipment Corp. PDP-11/60 computer system. The software used on the graphics system is the FRODO model building and refinement program described by T. A. Jones (1978).

Results

500-MHz ^1H NMR Spectrum of Ra5—General Features. The ^1H NMR spectrum of Ra5 is well resolved (Figure 1). The aromatic region is readily identifiable (6.4–7.6 ppm); one can easily pick out two tyrosine doublets in the upfield region and the complex resonances from the tryptophan residues on the downfield side. Many of the resonances in the methyl region (0.4–1.2 ppm) of Ra5 can be assigned to particular types of amino acids. A single ring-current-shifted methyl group is observed at 0.4 ppm. Several of these spectral regions have been analyzed to suggest specific assignments.

Aromatic Residues. Figure 2 shows a convolution difference spectrum of the aromatic region (7.5–6.5 ppm) of the 500-MHz ^1H NMR spectrum of Ra5 and assignments of the

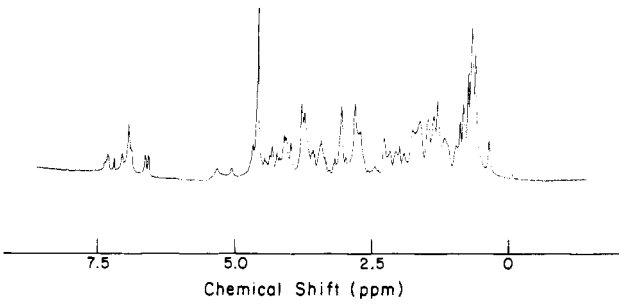

FIGURE 1: 500-MHz ^1H NMR spectrum of 0.5 mM ragweed allergen Ra5 in D_2O at 27 °C and apparent pH 7.5. A sweep width of 5000 Hz with 4096 data points was used for acquisition of 64 transients.

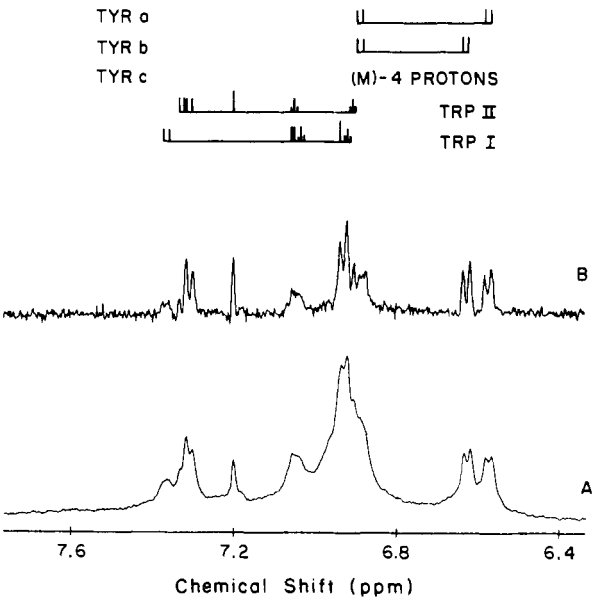

FIGURE 2: (A) Expansion of the aromatic region of Ra5; (B) convolution difference spectrum showing resolved splitting patterns. Partial assignments for the aromatic protons are indicated. For the convolution difference spectrum a broad exponential factor of 3 Hz and a narrow factor of 1 Hz were used.

resonances. The allergen contains three tyrosines (Tyr-17, Tyr-25, and Tyr-33) and two tryptophans (Trp-6 and Trp-28) as the only aromatic residues. The upfield pair of sharp doublets (6.57 and 6.62 ppm), each integrating to two protons, is assigned to the C-3 and C-5 protons of two of the tyrosines on the basis of their characteristic chemical shift and their pH shift behavior. Assignment of the other aromatic resonances required the additional information obtained from single-frequency decoupling and/or 2-D techniques such as COSY. The resonance pattern of one tryptophan (Trp-I) consists of a broad doublet at 7.38 ppm integrating to one proton, a complex multiplet centered at 7.03 ppm integrating to two protons, and resonances within the peak envelope centered at 6.92 ppm also integrating to two protons. The other tryptophan (Trp-II) has a similar pattern, consisting of a pseudo-triplet resulting from the C-4 and C-7 ring protons centered at 7.33 ppm, a singlet at 7.20 ppm from the C-2 proton, a broad peak at 7.03 ppm, and another proton signal under the resonance envelope at 6.92 ppm. These last two resonances correspond to the C-6 and C-5 protons.

The remaining tyrosyl resonances are buried under the peak envelope at 6.92 ppm (on the basis of integration of normal spectra). The resolved tyrosine patterns are characteristic of two tyrosines in similar environments, while the third tyrosine is in a quite different and distinct environment. The two tryptophan residues also appear to be in different microen-

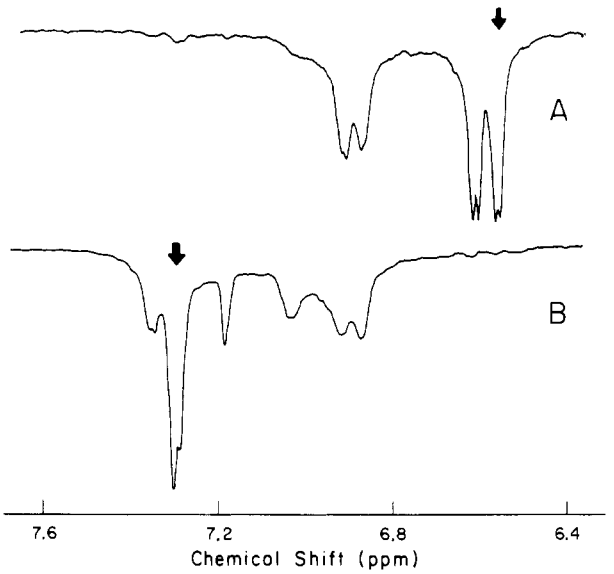

FIGURE 3: (A) NOE difference spectrum with irradiation at 6.55 ppm showing the Tyr-a/Tyr-b spin system. The boldface arrow indicates the irradiation frequency corresponding to the upfield tyrosyl doublet of Tyr-a. (B) NOE difference spectrum with irradiation at 7.30 ppm (arrow indicating irradiation of a pair of tryptophan doublets of Trp-II) showing the Trp-I/Trp-II/Tyr-c spin-diffusion system.

vironments, as deduced from the resonance patterns. Trp-II appears to have greater mobility, with shifts similar to tryptophan in random coil structures. Trp-I, on the other hand, appears to be restricted in mobility, as the chemical shifts for the C-4 and C-7 protons are well separated from one another. In addition, the singlet of Trp-I is shifted upfield to the region where the C-5 and C-6 proton resonances occur (ca. 7.2 ppm). This would also seem to indicate a degree of restriction for Trp-I or proximity to another aromatic ring system. The coupling patterns as described are shown above the convolution difference spectrum of this region (Figure 2).

The above interpretation is consistent with earlier results based on fluorescence and low-temperature phosphorescence (Galley et al., 1982). Those studies indicated that the two tryptophans in Ra5 are located in two distinct environments and that only one is exposed to polar solvent.

As a result of phenol deprotonation, both clearly resolved tyrosine doublets (6.57 and 6.62 ppm) shift upfield with increasing pH. The third tyrosine, although it is located within a large envelope of resonances, does not appear to shift as a function of pH to the same extent as the other tyrosines. Tyr-a and Tyr-b both have pK values of approximately 10.2.

T_1 values for aromatic protons are in the range 1.5–2.5 s. (A more detailed examination of Ra5 relaxation behavior will be presented elsewhere.) NOE difference experiments on the aromatic region of Ra5 suggested a number of structural features. NOE difference spectra and integration results are consistent only if Tyr-a and Tyr-b are part of one highly efficient spin-diffusion system (Figure 3A), while Trp-I, Trp-II, and Tyr-c are part of another efficient spin-diffusion system (Figure 3B). Moreover, the Tyr-a/Tyr-b spin system is completely distinct and therefore must be spatially resolved from the Trp-I/Trp-II/Tyr-c spin system.

Methyl Region. Methyl groups occur on amino acids threonine, alanine, valine, leucine, isoleucine, and methionine. Of these, Ra5 contains neither threonine nor methionine. There are potentially 15 discrete methyl groups in the Ra5 methyl region (Figure 4). Of these, 3 belong to alanine, 7.3 to valine, 2.7 to leucine, and 2 to isoleucine. The alanine methyl groups can be identified by irradiation at the appropriate α-CH region, ~4.4 ppm. Only alanine methyl groups will be coupled to resonances around 4 ppm. Upon irradiation at that frequency, three discrete resonances in the methyl region are affected. The chemical shifts of these alanine methyl groups are determined to be 0.95, 0.90, and 0.80 ppm.

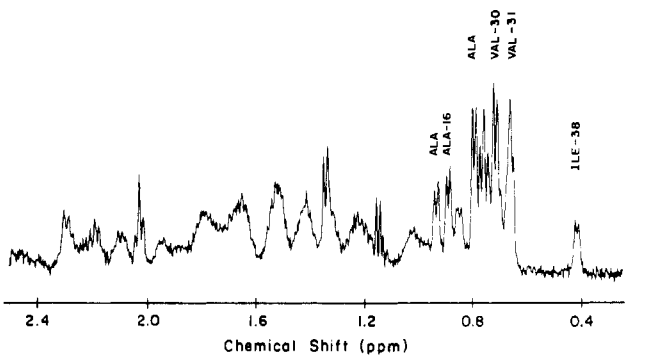

FIGURE 4: Convolution difference spectrum of the methyl/methylene region of Ra5. Tentative assignments are indicated.

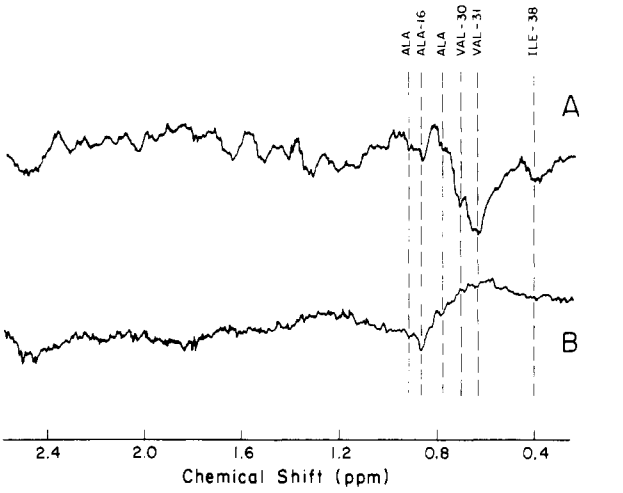

FIGURE 5: (A) NOE difference spectrum of the aliphatic region of Ra5 upon irradiation of the Tyr-a/Tyr-b spin system at 6.55 ppm; (B) NOE difference spectrum of the aliphatic region of Ra5 upon irradiation of the Trp-I/Trp-II/Tyr-c spin system at 7.30 ppm. Methyl resonances that have been assigned are indicated. The irradiation frequencies are the same as those shown in Figure 3.

Observation of the primary sequence immediately suggests an NOE experiment to identify valines-30 and -31. These residues are sandwiched between Trp-28 and Tyr-33 and adjacent to a disulfide cross-link (Trp-Gln-Val-Val-Cys-Tyr). The valine adjacent to cystine would be expected to have some motional constraints. The other valines in Ra5 are not close to the aromatics in the primary sequence. Irradiation of the upfield Tyr-a/Tyr-b resonances (and possibly the Trp/Tyr-c spin-diffusion system) should perturb the neighboring valines (presumably Val-31 more than Val-30). This does, in fact, occur (Figure 5). Moreover, a specific assignment can be made, as the NOE difference spectra show a greater enhancement for one pair of methyl resonances over another. On this basis, we have assigned the methyl resonance for Val-31 to be 0.68 ppm and that for Val-30 to be 0.73 ppm. (It is the former resonance that shows the larger magnitude NOE, hence the assignment.)

A unique upfield-shifted methyl doublet is observed at 0.42 ppm. This resonance titrates into the normal methyl region with a basic pK of approximately 10.2 and an acidic pK of approximately 4 and is linked to the Tyr-a/Tyr-b spin-diffusion system. The methyl region, with the exception of this ring-current-shifted methyl and a resonance at 0.9 ppm (identified

as an alanine methyl group), is not perturbed by changes in pH. Since the ring-shifted methyl and the 0.9 ppm resonance exhibit similar pH behavior, it is reasonable that both methyl groups are near Tyr-a or Tyr-b (where a tyrosine is hydrogen bonded to an acidic residue). T_1 values for methyl resonances are in the region 0.4–0.7 s.

There are several residues that could give rise to the ring-shifted methyl group: leucine, alanine, valine, or isoleucine. The three alanine methyl resonances have already been identified by homonuclear irradiation of the α-CH at approximately 4.4 ppm. The doublet pattern of the ring-shifted methyl group indicates that the methyl is adjacent to a CH group; this eliminates the δ-CH_3 of isoleucine.

A COSY map (Figure 6) shows a correlation of the ring-current-shifted methyl resonance (indicated by a along the diagonal in the figure) with a resonance at 1.45 ppm (c along the diagonal). This correlated peak shows two other connectivities: one to a resonance at 2.7 ppm (d), which is appropriate for an α-CH, and one to a resonance at 1.1 ppm (b) in the right range for a γ-CH_2. This also confirms that the ring-shifted methyl does not belong to alanine (where the 1.45 ppm peak would have to be an α-CH) or to valine (where the 1.1 ppm peak would have to belong to a methyl group, and we can already account for all methyl groups in the region 1.0–0.4 ppm). The β-CH of valine in model peptides is at 2.1 ppm. If the peak correlated to the ring-current-shifted methyl were a valine β-CH, then the β-CH would also be ring-current-shifted about 0.6 ppm, while only one of the two methyl groups showed a large upfield shift—this would represent an unusual geometry. Furthermore, we have identified the methyl groups from Val-30 and Val-31 with NOE experiments. A methyl group from Val-2 would not yield an area corresponding to three protons, but rather two (because of the positional heterogeneity). All these arguments tend to rule out valine as the amino acid whose methyl group is at 0.4 ppm. This leaves two possible amino acids: leucine, of which there are $1^1/_3$ (Leu-1 and the heterogeneous residue at position 2), or the sole isoleucine at position 38. Integration of the ring-shifted methyl and comparison to the tyrosine doublets indicate three protons, so that leucine at the heterogeneous 2-position is not an option.

The final two possibilities are the γ-CH_3 of Ile-38 and the two δ-CH_3's of Leu-1. Random coil parameters of isoleucine predict a chemical shift for the β-CH of 1.89 ppm. This is somewhat far from the 1.45 ppm observed but not out of line with an expected ring-current shift of 0.4 ppm. We could not easily discern a cross-peak for the γ-CH_2/CH_3 interaction, i.e., a cross-peak clearly linking the resonance at 1.1 ppm with a methyl resonance. Given the spectral noise and the proximity of the γ-CH_2/CH_3 shifts, one expects such a correlation to be difficult to detect. Another argument in favor of the assignment of the ring-shifted methyl to Ile-38 is the spin–spin coupling value of the doublet. The random coil value for J_{CH_3CH} is 6.9 Hz—very close to the observed value of 7.1 Hz that was obtained from a 2-D J-resolved spectrum.

The alternate possibility, viz., that the ring-shifted methyl belongs to the Leu-1 residue, is not very consistent with the correlations detected in the COSY experiment. While the experimental and random coil chemical shifts for the γ-CH of leucine agree quite well (1.45 and 1.65 ppm, respectively), the additional correlations to resonances at 1.1 and 2.7 ppm are not appropriate for the β-methylene or CH_3 groups. Another argument against the leucine-1 assignment is the spin–spin coupling value observed (J_{CH_3CH} = 7.1 Hz, measured by 2-D J-resolved spectroscopy) vs. random coil predictions

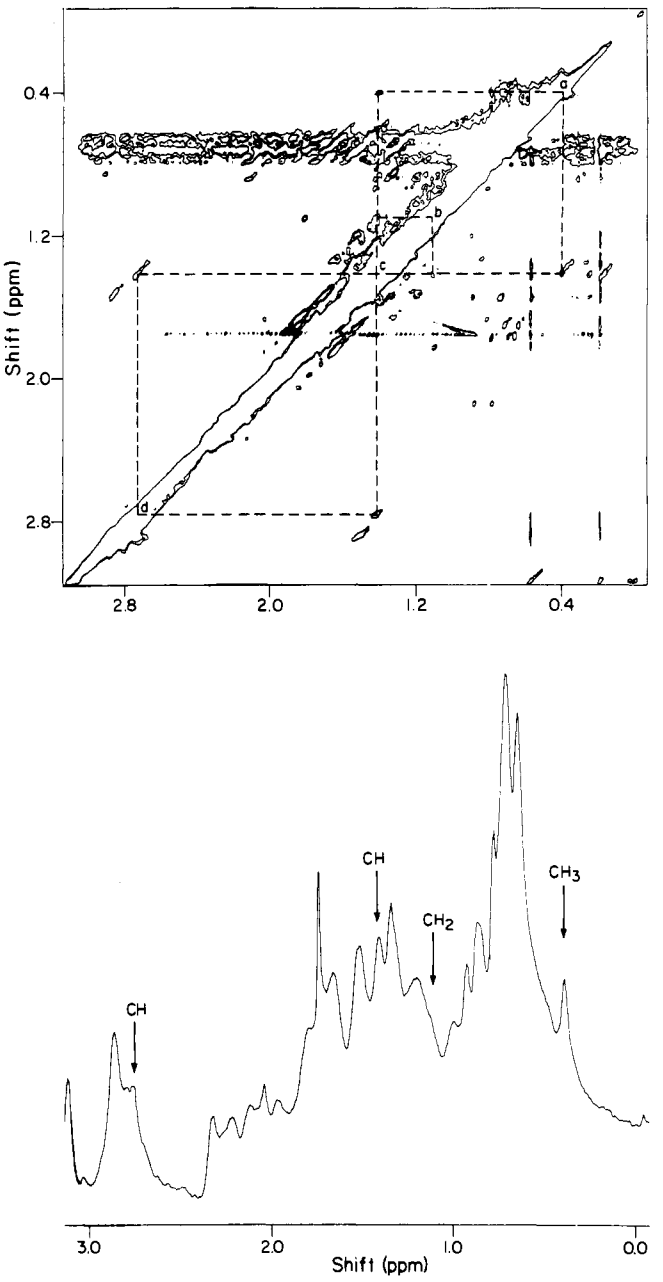

FIGURE 6: (Top) Expansion of the contour plot of the COSY spectrum of Ra5 showing the methyl/methylene region. The ring-current-shifted methyl (peak a) at 0.4 ppm is connected by two off-diagonal cross-peaks with a resonance at 1.4 ppm (labeled c). That resonance is in turn connected to one at 1.1 ppm (labeled b) and another at 2.7 ppm (labeled d). The chemical shift pattern of these resonances is used in identifying the ring-shifted methyl as part of isoleucine-38. (Bottom) The location and identification of these isoleucine protons in the 1-D spectrum.

(6.3 Hz)—a disagreement of nearly 1 Hz where variations of 0.5 Hz are rarely seen, except in highly strained systems.

To confirm this assignment for Ile-38, we prepared the des-Leu[1] derivative of Ra5. A spectrum of the methyl region of this derivative is shown in Figure 7. The chemical shift and intensity of the ring-current-shifted methyl have not been altered. Therefore, this uniquely shifted resonance belongs to Ile-38.

Discussion

A number of interesting resonances are observed in Ra5 and point to specific structural features of the protein. In particular, the detection of two distinct NOE spin-diffusion

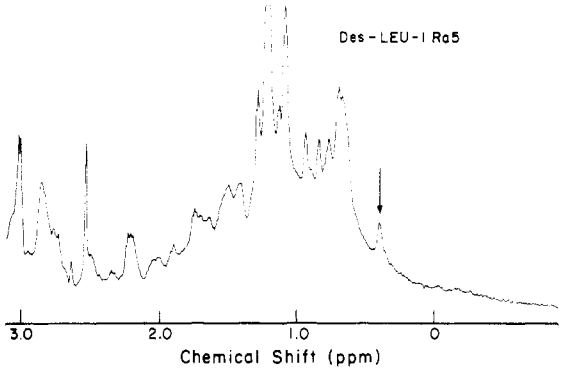

FIGURE 7: 500-MHz ^1H NMR spectrum of 0.15 mM des-Leu^1-Ra5 in D_2O at 25 °C and apparent pH 7.0. A sweep width of 5000 Hz, 4096 data points, 90° pulse, 4-s recycle time, and a line broadening of 2 Hz were used for acquisition and processing of 1000 transients.

systems and the identification of the ring-current-shifted methyl can be used to suggest reasonable disulfide cross-links in this protein and to evaluate the structure on the basis of the toxin–agglutinin fold (Drenth et al., 1980).

The ring-current-shifted methyl must be positioned above a tyrosine ring whose hydroxyl is hydrogen bonded to an acidic residue. Given our experimental results that the ring-shifted methyl is Ile-38, a number of extrapolations can be made. The ring-shifted methyl is an integral part of the Tyr-a/Tyr-b spin-diffusion system. Irradiation of either of the tryptophans (and hence Tyr-c by spin diffusion) shows no saturation of the ring-shifted methyl. These same two tyrosines are also spatially close to Val-30 and Val-31, as seen in the NOE difference spectrum. The Tyr-a/Tyr-b spin-diffusion system shows enhancement of only one alanine (presumably Ala-16 at 0.90 ppm) whereas the tryptophan spin-diffusion system shows enhancement of one of the other two alanines (both alanines-5 and -7 flank Trp-6). Ala-5 is expected to be held in a relatively rigid section of Ra5 due to the proline and cystine residues and would be expected to show such enhancement. Ala-7, in a more flexible segment of the sequence, might not be part of the spin-diffusion system and hence would show little enhancement. Of the three alanine methyl groups, the one at 0.95 ppm is tentatively assigned to Ala-5, the 0.90 ppm resonance to Ala-16, and the 0.80 ppm resonance to Ala-7.

On the basis of this information, a tentative assignment of a disulfide bridge between Cys-(18,19) and Cys-32 can be made. Such a cross-link would produce the Tyr-a/Tyr-b/ring-shifted methyl/Val-30,31/Ala-16 spin-diffusion system that is observed. It is interesting to note that Ile-38 is adjacent to a cysteine residue. Since it is involved in the same rigid system as the two tyrosines, a covalent disulfide cross-link would provide the most efficient spin diffusion. This forms the basis for another tentative disulfide assignment between Cys-(19,18) and Cys-39. That particular disulfide arrangement brings into proximity glutamic acids (Glu-34 and Glu-37) required for the hydrogen bonding to the tyrosine. The first of these disulfide pairs is consistent with the data presented by Drenth et al. (1980). The second is not within their proposed pairing scheme but is more suitable from the NMR data.

Drenth et al. (1980) have proposed that proteins which show pronounced sequence homologies (particularly in cysteine content) to wheat germ agglutinin contain a similar toxin–agglutinin fold. Typically, these proteins tend to be small (40–70 amino acids) and disulfide rich. The four small proteins compared all contain a core of four disulfide linkages. The crystallographic structures for two of these, wheat germ

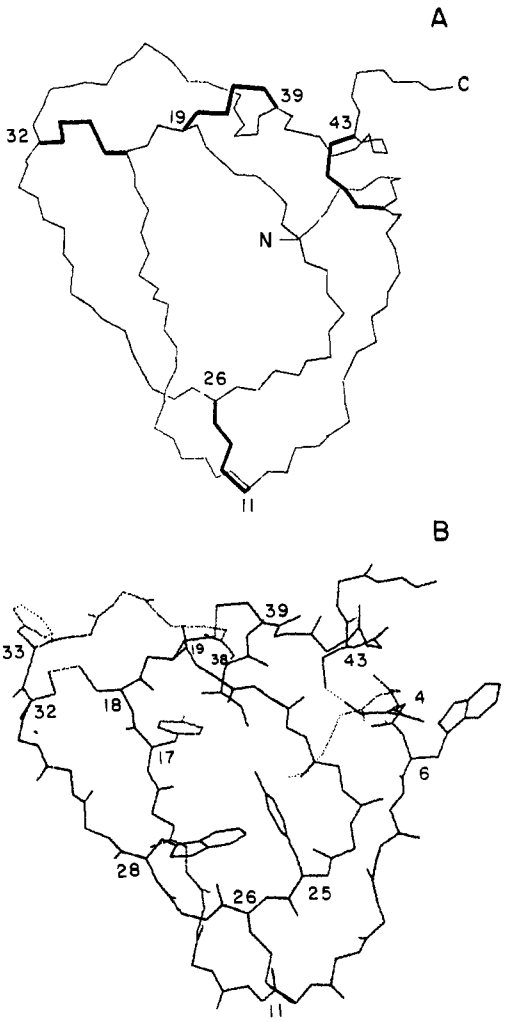

FIGURE 8: Proposed structure of Ra5 based loosely on the toxin–agglutinin fold proposed by Drenth et al. (1980) and modified as described in the text to be consistent with ^1H NMR data. This hypothetical model generated with a computer graphics system illustrates (A) the polypeptide backbone and disulfide linkages of Ra5 and (B) the spatial relationships of the various aromatic residues and the ring-current-shifted isoleucine methyl group.

agglutinin and erabutoxin, show similar arrangements of the polypeptide backbone. Ragweed allergen Ra5 and hevein, whose crystal structures are not known, show sequence homologies in the cysteine residues that deviate at less than two amino acid locations from those in wheat germ agglutinin and erabutoxin. A comparison of authentic and suggested disulfide pairings shows that the link between Cys-11 and Cys-26 is the most highly conserved. Given our two Ra5 disulfide cross-links suggested by the NMR data [Cys-(18,19)–Cys-32 and Cys-(19,18)–Cys-39] and the highly conserved disulfide seen by Drenth et al. (1980) (Cys-11–Cys-26), the final disulfide bridge in Ra5, Cys-4–Cys-43, would tie the N- and C-termini together.

On the strength of these disulfide assignments and the identification of two distinct spin domains, and using the secondary structure predictions (Table I), we generated a solution structure for Ra5 (Figure 8) that is consistent with the toxin–agglutinin fold model proposed by Drenth et al. (1980) and incorporates all the structural features we detect by NMR. The disulfide pairing is emphasized in Figure 8A while side chains critical to interactions detected by NMR are shown in Figure 8B. In the model, the ambiguity in the Cys-18 and Cys-19 pairing has been resolved. Unless Cys-18 is paired with Cys-32 and a Cys-19–Cys-39 disulfide formed, the po-

lypeptide chain cannot be constrained to form a disulfide between Cys-4 and Cys-43. Looking at the arrangement of side chains, we see that a methyl group of Ile-38 is ring-current-shifted by a tyrosine (Tyr-17) that is spatially close (and tied by an adjacent disulfide) to another tyrosine (Tyr-33) and the two adjacent valines (Val-30,31). The methyl resonance for Ala-16 is in the vicinity and would also be expected to show pH titration behavior indicative of the adjacent tyrosine interactions. Several glutamates could be spatially close to Tyr-17 (Glu-34, Glu-37) and possibly Asp-21, and one of these could be involved in a hydrogen bond with the aromatic residue. The environments of the two tryptophans are distinctly different: one (Trp-28) is juxtaposed to Tyr-c (Tyr-25). Such a spatial relationship would yield environments for both of these amino acids dramatically different from random coil parameters. Trp-6, on the other hand, would be expected to behave as if in a random coil. Another interesting feature of the proposed Ra5 structure is that the two β-strands form a parallel β-sheet; for the known structures with the toxin–agglutinin structure the β-strands are antiparallel.

This modified Drenth model of Ra5, it should be emphasized, is only a possible conformation Ra5 may acquire in solution, based on disulfide assignments, secondary structure predictions, and NMR constraints. While the details may differ from an actual crystal structure, we suggest these overall features will be retained. They may be conserved in related allergens as well. Preliminary ^{1}H NMR studies of giant Ra5 from *Ambrosia trifida* show a ring-current-shifted methyl group interacting with a tyrosine with pK_a values of 4 and 10.5.[2] If the sequence of this allergen is aligned with that of small Ra5, there is an isoleucine in giant Ra5 homologous to isoleucine-38. Similar ^{1}H NMR and more detailed NOE studies should allow us to compare and refine solution structures for Ra5 and related allergens. This information will be crucial to understanding the allergic and immune response in man.

Registry No. Isoleucine, 73-32-5.

REFERENCES

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Bax, A. (1982) *Two-Dimensional Nuclear Magnetic Resonance in Liquids*, Reidel, Boston, MA.

Bias, W. B., Hsu, S. H., Meyers, D. A., Goodfriend, L., & Marsh, D. G. (1979) *Transplant. Proc. 11*, 1853–1857.

Campbell, I. D., Dobson, C. M., Williams, R. J. P., & Xavier, A. V. (1973) *J. Magn. Reson. 11*, 172–181.

Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 211–244.

Choudhury, A. M., & Goodfriend, L. (1983) *J. Protein Chem. 2*, 321–333.

de Weck, A. L. (1977) *Monogr. Allergy II*, 3–18.

Dixon, H. B. F., & Fields, R. (1972) *Methods Enzymol. 25*, 409–419.

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

Edman, P. (1970) in *Protein Sequence Determination* (Needleman, S. B., Ed.) pp 211–255, Springer-Verlag, New York.

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

Goodfriend, L. (1976) *Int. Congr. Ser.—Excerpta Med., 1976 No. 414*, 151–155.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–273.

King, T. P. (1976) *Adv. Immunol. 23*, 77–105.

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

Lim, V. I. (1974) *J. Mol. Biol. 88*, 873–894.

Løwenstein, H., & Marsh, D. G. (1981) *J. Immunol. 126*, 943–948.

Løwenstein, H., King, T. P., Goodfriend, L., Hussain, R., Roebber, M., & Marsh, D. G. (1981) *J. Immunol. 127*, 637–642.

Marsh, D. G. (1975) *Antigens 3*, 271–359.

Marsh, D. G., Bias, W. B., Santilli, J., Jr., Schachter, B., & Goodfriend, L. (1975) *Immunochemistry 12*, 539–543.

Marsh, D. G., Hsu, S. H., Roebber, M., Ehrlich-Kautzky, E., Freidhoff, L. R., Meyers, D. A., Pollard, M. K., & Bias, W. B. (1982) *J. Exp. Med. 155*, 1439–1451.

Mole, L. E., Goodfriend, L., Lapkoff, C. B., Kehoe, J. M., & Capra, J. D. (1975) *Biochemistry 14*, 1216–1220.

Vold, R. L., Waugh, J. S., Klein, M. P., & Phelps, D. E. (1968) *J. Chem. Phys. 48*, 3831–3832.

Webster, D., & Offord, R. E. (1972) *Biochem. J. 130*, 315–317.

[2] M. F. Roberts and L. Goodfriend, unpublished results.