Kraal, B., & Hartley, B. S. (1978) *J. Mol. Biol. 124*, 551–564.
Kushiro, A., Shimizu, M., & Tomita, K.-I (1987) *Eur. J. Biochem. 170*, 93–98.
La Cour, T. F. M., Nyborg, J., Thirup, S., & Clark, B. F. C. (1985) *EMBO J. 4*, 191–219.
Laursen, R. A. (1977) *Methods Enzymol. 47*, 277–289.
Laursen, R. A., L'Italien, J. J., Nagarkatti, S., & Miller, D. L. (1981) *J. Biol. Chem. 256*, 8102–8109.
Leberman, R., & Egner, U. (1984) *EMBO J. 3*, 339–341.
Leberman, R., Antonsson, B., Giovanelli, R., Guariguata, R., Schumann, R., & Wittinghofer, A. (1980) *Anal. Biochem. 104*, 29–36.
Lechner, K., & Böck, A. (1987) *Mol. Gen. Genet. 208*, 523–528.
Lowe, P. N., & Beechey, R. B. (1982) *Bioorg. Chem. 11*, 55–71.
McCormick, F., Clark, B. F. C., la Cour, T. F. M., Kjeldgaard, M., Norskov-Lauritsen, L., & Nyborg, J. (1985) *Science (Washington, D.C.) 230*, 78–82.
Miller, D. L., & Weissbach, H. (1977) in *Molecular Mechanisms of Protein Biosynthesis* (Weissbach, H., & Pestka, S., Eds.) pp 323–373, Academic, New York.
Miyazaki, M., Uritani, M., Fujimura, K., Yamakatsu, H., Kageyama, T., & Takahashi, K. (1988) *J. Biochem. (Tokyo) 103*, 508–521.
Möller, W., Schipper, A., & Amons, R. (1987) *Biochimie 69*, 983–989.
Montandon, P.-E., & Stutz, E. (1984) *Nucleic Acids Res. 12*, 2851–2859.
Nakamura, S., Ohta, S., Arai, K.-I., Arai, N., Oshima, T., & Kaziro, Y. (1978) *Eur. J. Biochem. 92*, 533–543.
Ohama, T., Yamao, F., Muto, A., & Osawa, S. (1987) *J. Bacteriol. 169*, 4770–4777.
Ohta, S., Nakanishi, M., Tsuboi, M., Arai, K.-I., & Kaziro, Y. (1977) *Eur. J. Biochem. 78*, 599–608.
Rogers, S., Wells, R., & Rechsteiner, M. (1986) *Science (Washington, D.C.) 234*, 364–368.
Ruusala, T., Ehrenberg, M., & Kurland, C. G. (1982) *EMBO J. 1*, 75–78.
Salnikow, J., Lehmann, A., & Wittmann-Liebold, B. (1981) *Anal. Biochem. 117*, 433–442.
Seidler, L., Peter, M., Meissner, F., & Sprinzl, M. (1987) *Nucleic Acids Res. 15*, 9263–9277.
Sigal, I. S., Gibbs, J. B., D'Alonzo, J. S., & Scolnick, E. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4725–4729.
Stryer, L., & Bourne, H. R. (1986) *Annu. Rev. Cell Biol. 2*, 391–419.
Van Noort, J. M., Kraal, B., Sinjorgo, K. M. C., Persoon, N. L. M., Johanns, E. S. D., & Bosch, L. (1986) *Eur. J. Biochem. 160*, 551–561.
Wittinghofer, A., Warren, W. F., & Leberman, R. (1977) *FEBS Lett. 75*, 241–243.
Wittinghofer, A., Frank, R., & Leberman, R. (1980) *Eur. J. Biochem. 108*, 423–431.

^{1}H NMR Studies of Human C3a Anaphylatoxin in Solution: Sequential Resonance Assignments, Secondary Structure, and Global Fold†

Walter J. Chazin,*,‡ Tony E. Hugli,§ and Peter E. Wright*,‡
Departments of Molecular Biology and Immunology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, California 92037

Received May 5, 1988; Revised Manuscript Received August 4, 1988

ABSTRACT: The spin systems that comprise the ^{1}H nuclear magnetic resonance (NMR) spectrum of the complement fragment C3a (M_r 8900) have been completely identified by an approach which integrates data from a wide range of two-dimensional NMR experiments. Both relayed and multiple quantum experiments play an essential role in the analysis. After the first stage of analysis the spin systems of 60 of the 77 residues were assigned to the appropriate residue type, providing an ample basis for subsequent sequence-specific assignments. Elements of secondary structure were identified on the basis of networks of characteristic sequential and medium-range nuclear Overhauser effects (NOEs), values of $^3J_{HN\alpha}$, and locations of slowly exchanging backbone amide protons. Three well-defined helical segments are found. Gradients of increasing mobility in distinct segments of the C3a polypeptide are observed, with very high mobilities for several residues near the C- and N-termini, including the complete C-terminal receptor binding site pentapeptide LGLAR. The NMR data, combined with known disulfide linkages and a small number of critical long-range NOEs, provide the global folding pattern of C3a in solution. Identical solution structures were found for both the intact active protein and the largely inactive physiologic product des-Arg^{77}-C3a. Differences between the solution and crystal structures of C3a are observed, particularly in the N-terminal region. The relevance of these new observations is discussed with respect to physiologic responses that are elicited by the "local hormone-like" anaphylatoxin molecule.

Anaphylatoxin C3a is a small protein fragment (77 amino acids, M_r 8900) of complement component C3 released into the blood upon activation of the complement cascade. The C3a molecule is characterized biologically as a spasmogen capable of inducing smooth muscle contraction, increasing vascular permeability, and causing a skin wheal and flare reaction when injected intradermally. Many of the biologic responses are mediated via tissue mast cells that are activated either directly or indirectly by the anaphylatoxin. The protein has been studied by a variety of biochemical techniques [re-

† This work was supported by Grants GM 36643 (P.E.W.), HL 16411 (T.E.H.), and HL 25658 (T.E.H.) from the National Institutes of Health.
* To whom correspondence should be addressed.
‡ Department of Molecular Biology.
§ Department of Immunology.

0006-2960/88/0427-9139$01.50/0

viewed in Hugli (1981, 1984)] and by several physical methods (Hugli et al., 1975a,b; Muto et al., 1985, 1987) that characterize stability and conformation. An X-ray structure of human C3a was reported by Huber et al. (1980) which provided important insights into the folded structure and structure–function relationships of this molecule. However, some doubts remained concerning certain structural aspects of C3a in solution under physiologic conditions. Several important questions could not be answered by crystallography because coordinates for the first 14 N-terminal residues and of the C-terminal arginine, which is required for activity (Bokisch & Muller-Eberhard, 1970), could not be determined by X-ray analysis. To determine the structure in solution, we have undertaken two-dimensional ^{1}H NMR[1] studies of human C3a and its inactive des-Arg^{77} form. In the present paper we describe unambiguous assignments of the entire ^{1}H NMR spectrum, which allow us to describe in qualitative terms the structure of C3a in solution, and compare it with the crystal structure. In addition, we could compare in detail the solution structure of the active intact and inactive des-Arg^{77} forms of C3a.

Identification of the ^{1}H spin systems and sequence-specific assignment of C3a have been carried out with a systematic approach for identification of spin systems (Chazin et al., 1988) followed by the sequential resonance assignment procedure (Billeter et al., 1982). The identification of spin systems depends upon a wide range of ^{1}H NMR experiments to obtain an essentially *complete* set of scalar (through-bond) connectivities for each residue in the protein. A unique aspect of this approach is the identification of most spin systems in spectra acquired from ^1H_2O solution. The backbone amide proton rather than the C$^\alpha$ proton is used as the foundation for spin system identification (Chazin & Wright, 1987). The availability of complete spin system assignments greatly facilitates the sequential resonance assignment procedure. We have demonstrated the success of this approach on plastocyanin, composed largely of β-sheets and regular turns, and demonstrate here the feasibility for helical proteins. This is particularly important because the small value of $^3J_{HN\alpha}$ characteristic of helix significantly reduces the efficiency of long-range coherence transfer.

Materials and Methods

Intact human C3a and des-Arg^{77}-C3a were isolated and purified as described previously (Hugli et al., 1975a, 1982). Samples for NMR were dialyzed repeatedly against H_2O adjusted to pH 4 to remove residual formate, then against neutral H_2O, on a microdialysis apparatus. The solutions were then lyophilized. Most of the experiments were carried out on a single preparation of intact C3a. Experiments were first run at 5 mM protein and 10 mM NaCl at pH 5.5. This solution contained 5% ^2H_2O but for brevity will be referred to as a ^1H_2O solution. The pH was then titrated to 4.5 with microliter quantities of HCl, and additional experiments were run. The sample was then back-titrated to pH 5.5 with microliter quantities of NaOH and split in equal aliquots. The first aliquot was lyophilized and then dissolved in ^2H_2O to identify slowly exchanging amides. This sample was then heated to 60 °C for 10 min, cooled, lyophilized, and made up in "100%" ^2H_2O at pH* 5.6 (not corrected for isotope effects). The second aliquot was titrated to pH 6.8 for further experiments. A fresh sample was adjusted to 2.5 mM protein in 50 mM sodium phosphate at pH 6.0. One des-Arg^{77}-C3a sample was adjusted to 7 mM protein in 10 mM NaCl at pH 5.5 and a second to 4 mM protein in 150 mM sodium phosphate at pH 4.5. Special precautions were necessary during sample preparation due to the extreme basicity of the protein, including use of plastic apparatus rather than glass whenever possible and siliconization [Prosil-28, PCR Research Chemicals, Inc.] of the NMR sample tubes. Control experiments demonstrated that there were no peaks in the NMR spectrum which resulted from the siliconization process. C3a is extremely stable but has a high potential for proteolytic degradation from contamination by airborne organisms. Therefore, all buffer solutions contained sodium azide (0.5–1.0 mM), and air was excluded from the sample tubes by sealing under argon.

^{1}H NMR spectra were recorded at 283, 296, 300, and 310 K with Bruker AM-500 spectrometers equipped with Aspect 3000 computers and digital phase shifting hardware. The raw data were processed on an Aspect 3000 data station using Bruker software or a SUN 3/160C computer, with a version of the FTNMR program provided by Dr. Dennis Hare and modified for use on the SUN by Dr. James Sayre.

Two-dimensional NMR spectra were recorded, and data were processed as described in detail in Chazin et al. (1988). Experiments at pH 6.8 were run at 296 K; all others, at 300 K unless stated otherwise. COSY spectra from ^1H_2O solutions of C3a were acquired at pH 4.5, 5.0, 5.5, 6.0, and 6.8 and of des-Arg^{77}-C3a at pH 4.5 and 5.5. One R-COSY spectrum with τ = 38 ms was acquired from the ^1H_2O solution of des-Arg^{77}-C3a at pH 5.5. DR-COSY spectra with τ_1 = 22 ms and τ_2 = 38 or 40 ms (Chazin & Wüthrich, 1987) were acquired from ^1H_2O solutions of C3a at pH 4.5, 5.5, and 6.8 and of des-Arg^{77}-C3a at pH 5.5. An additional DR-COSY spectrum with τ_1 = 45 ms and τ_2 = 38 ms, a 2Q spectrum with a 33-ms excitation period, and three TOCSY spectra with 61-, 80- (310 K), and 101-ms spin-lock periods were acquired from the ^1H_2O solution of des-Arg^{77}-C3a at pH 5.5. TOCSY experiments were obtained with a modification described by Rance (1987). Pure absorption NOESY spectra with a mixing time of 200 ms were acquired from ^1H_2O solutions of C3a at pH 6.0, and at pH 5.5 and 283 K, and with a 100-ms mixing time from the ^1H_2O solution of des-Arg^{77}-C3a at pH 5.5. Two additional NOESY experiments with mixing times of 100 and 150 ms and with the observe pulse replaced by a $45°_{(x)}$–τ–$45°_{(-x)}$ sequence (Clore & Gronenborn, 1983) were acquired from the ^1H_2O solution at pH 6.8. Pure absorption Hahn-echo NOESY spectra (M. Rance, unpublished) with mixing times of 40, 80, 200, and 400 ms were acquired from the ^1H_2O solution of C3a at pH 5.5, with a 200-ms mixing time from the ^1H_2O solution at pH 4.5 and with a 400-ms mixing time from the ^1H_2O solution at pH 6.8. The following spectra were acquired from the ^2H_2O solution of C3a: 3QF-COSY; 2Q with a 30-ms excitation period; 3Q with a 22-ms excitation period; TOCSY with 64- and 102-ms (310 K) spin-lock periods; pure absorption Hahn-echo NOESY with mixing times of 40, 80, and 150 ms.

[1] Abbreviations: NMR, nuclear magnetic resonance; 1D, one dimensional; 2D, two dimensional; COSY, correlated spectroscopy; R-COSY, relayed COSY; DR-COSY, double relayed COSY; MQ (2Q, 3Q), multiple quantum (two quantum, three quantum) spectroscopy; MQF-COSY (2QF-COSY, 3QF-COSY), multiple quantum filtered (two quantum filtered, three quantum filtered) COSY; TOCSY, total correlation spectroscopy; FID, free induction decay; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; 3-spin, C$^\alpha$H–C$^\beta$H$_2$ spin system of serine, cysteine, aspartic acid, asparagine, and the aromatic amino acid residues; 5-spin, C$^\alpha$H–C$^\beta$H$_2$–C$^\gamma$H$_2$ spin subsystem of glutamic acid, glutamine, and methionine residues.

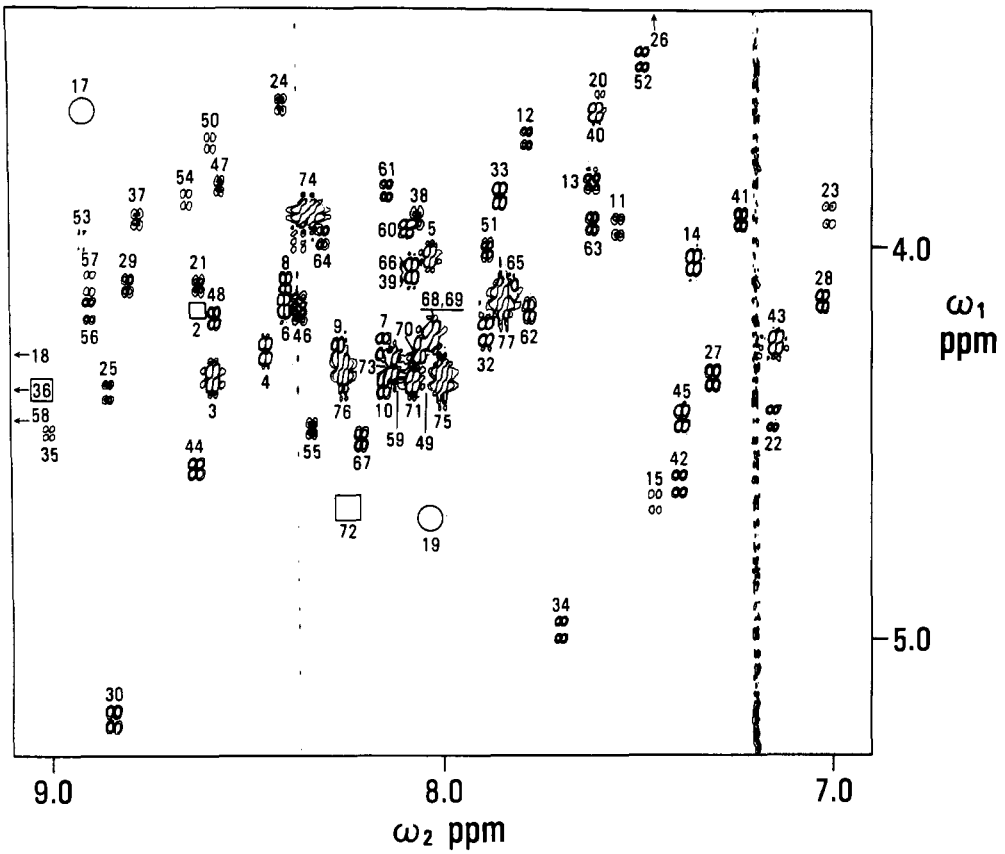

FIGURE 1: Backbone fingerprint region of a 500-MHz COSY spectrum of C3a at pH 5.5, 300 K. The ω_1 = C$^\alpha$H, ω_2 = NH cross-peaks are labeled with the sequence-specific assignment. Arrows indicate cross-peaks which appear outside the region shown. Boxes show the location of cross-peaks not observed in this spectrum, but which could be identified in other COSY spectra. Circles are drawn around cross-peaks that can only be observed at very low contour levels.

The slowly exchanging amide protons indicated in Figure 3 were identified on the basis of observation of at least one cross-peak in a 4-h Hahn-echo NOESY spectrum (τ_m = 200 ms) started 1 h after dissolution of the protein in ^2H_2O at pH* 5.6. Backbone NH/C$^\alpha$H coupling constants were measured as described by Marion and Wüthrich (1983). Except for the N- and C-terminal regions where line widths are narrow, the estimated accuracy of these measurements is adversely affected by the relatively large line widths (>10 Hz) for the amide proton resonances. On the basis of the simulations of Neuhaus et al. (1985), the lower limiting value that can be measured for the line separation is in the range of 6.0–6.5 Hz. The large line widths therefore constrain the definition of "small" coupling constant in Figure 3 to observations of line separation of ≤6.5 Hz. Confidence in these data is increased by the concomitant severe attenuation of cross-peak intensity for almost all of the residues with the minimum separation, an observation that is consistent with the actual values being substantially less than the observed separation (Neuhaus et al., 1985).

Results

Comparison of Intact C3a and Des-Arg^{77}-C3a. Initially, both 1D and 2D ^{1}H NMR spectra were obtained under identical conditions for the intact and des-Arg^{77} forms of the protein. The spectra were essentially identical, except for the loss of the resonances of Arg-77 and shifting of the resonances of Ala-76 in the spectra of the des-Arg^{77} form. Comparison of the COSY spectra provided assignments for the backbone NH/C$^\alpha$H cross-peaks of Arg-77 and Ala-76, and the C$^\alpha$H/C$^\beta$H$_3$ cross-peaks of Ala-76, and showed that the resonances in the Ser-1–Leu-75 polypeptide segment were identical in the intact and des-Arg^{77} forms. A residue by residue comparison of NOESY cross-peaks confirmed that the solution conformations were identical. Thus, the two forms could be used interchangeably for NMR experiments.

Spin System Identification. ^{1}H spin systems of C3a were identified according to the integrated assignment strategy described in detail in Chazin et al. (1988). In this strategy, relayed coherence transfer and multiple quantum experiments are utilized to make complete and unambiguous spin system assignments. In the following, we distinguish fragments of a spin system which show connectivities to the backbone amide protons (backbone amide based spin subsystems) from those which show connectivities to the side-chain terminal protons (side-chain terminal based spin subsystems). The resonances comprising each spin system are classified as unique (glycine, alanine, threonine, valine, isoleucine, leucine, arginine, lysine, and proline), 3-spin, or 5-spin on the basis of the pattern of characteristic proton scalar connectivities. This primary classification scheme can be refined on the basis of characteristic NOEs from methionine C$^\epsilon$H$_3$, asparagine and glutamine side-chain amides, and aromatic ring protons to the corresponding backbone and C$^\beta$ protons, providing residue-specific assignments for spin systems initially classified in the 3-spin and 5-spin categories. The objective of extensive spin system assignment is to provide as large a data base as possible to facilitate the subsequent sequential assignment procedure. It is important to note that assignment of a proton to a specific location along the side chain does not imply a stereospecific assignment.

(*A*) *Preliminary Analysis.* The characteristic backbone fingerprint region (ω_1 = 6.0–3.0 ppm, ω_2 = 10.5–6.5 ppm) of the COSY spectrum of C3a should contain cross-peaks for 74

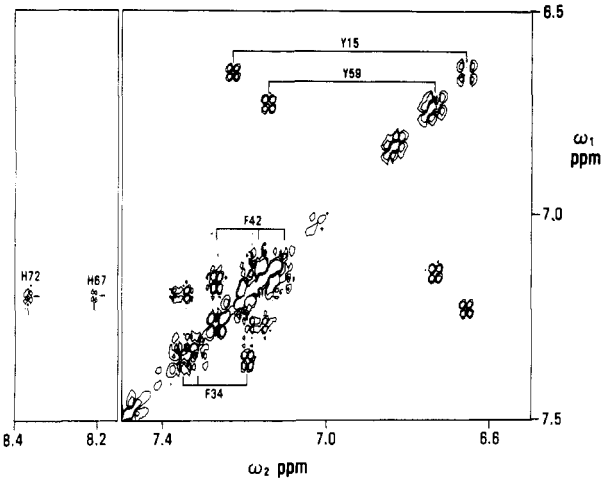

FIGURE 2: Aromatic ring fingerprint region of the same COSY spectrum as Figure 1. Cross-peaks are labeled with the sequence-specific assignments.

residues (77 residues minus the N-terminal serine and two prolines). At pH 5.5 and 300 K, cross-peaks for 71 residues are identified (Figure 1). Three cross-peaks are lost due to saturation transfer from irradiation of the H_2O signal but are observed at lower pH.

Seven cross-peaks were identified for the methyl protons of the four alanine and three threonine residues in the characteristic methyl region of the COSY spectrum. The valine/leucine/isoleucine fingerprint region exhibits two pairs of antiphase triplet cross-peaks for the two isoleucine C^{δ} methyl groups and 21 of the 22 methyl doublet resonances expected for the three valine, seven leucine, and two isoleucine residues. The absence of one cross-peak is due to the accidental degeneracy of the Leu-54 methyl group proton resonances.

In the aromatic region of the COSY spectrum (Figure 2), the ring proton spin systems of the two histidines, the two tyrosines, and two of the three phenylalanines could be identified in a straightforward manner. The assignments were complicated by near degeneracy of resonances but could be unambiguously identified in the 3Q spectrum (Chazin et al., 1988). The ring proton resonances of the third phenylalanine could not be specifically assigned, because they are highly degenerate (between 7.01 and 7.09 ppm) and do not give rise to peaks in any of the multiple quantum experiments.

In the backbone fingerprint region, a group of eight residues (subsequently identified as Val-2, Gln-3, and the segment His-72–Arg-77) exhibit very narrow line widths and intense cross-peaks. In addition, a second group (segments Leu-4–Glu-6 and Ala-68–Ser-71) exhibit line widths and cross-peak intensities in an intermediate range between the first group and the large majority of residues. The significance of these and corresponding observations in other regions of the spectra will be discussed below.

(*B*) *Identification of Backbone Amide Based Spin Systems.* For each spin system a substantial number (in many cases all) of the side-chain resonances were identified from a series of COSY, R-COSY, DR-COSY, and TOCSY spectra acquired from ^1H_2O solutions. For C3a, these assignments represent 88% of the proton resonances that could under ideal circumstances be identified through relayed connectivities to corresponding backbone amide proton resonances, including 130 (96%) of the 136 C^{β} and 89 (86%) of the 104 C^{γ} protons. On the basis of data obtained solely from spectra recorded in ^1H_2O solutions, 28 of the 41 unique spins systems could be directly identified, including the four glycines, four alanines, three threonines, three valines, six of seven leucines, six of eleven arginines,[2] and two of seven lysines. Two additional backbone amide-based spin subsystems could be unambiguously assigned to the two isoleucine residues on the basis of the relayed connectivities to the characteristic $C^{\gamma}H_3$ resonances, despite the fact that the complete spin systems could not be identified at this stage. Utilizing the scheme described in Chazin et al. (1988) for classifying spin systems, 20 of the remaining spin systems were initially assigned as one of the 22 3-spin side-chain residues, and 23 others were initially assigned to the 12 5-spin or the remaining arginine (five), leucine (two), and lysine (five) spin systems. There were two spin systems (subsequently identified as Lys-17 and Ser-35) in an indeterminate group because no relayed connectivities were observed in any spectrum. Due to the lack of backbone amide resonances, assignments for Ser-1, Pro-16, and Pro-31 were not yet available at this stage of the procedure.

(*C*) *Identification of Residues with Unique Spin Systems.* Two of four glycine spin systems were identified from their characteristic $\omega_1 = \omega_{\alpha}$, $\omega_2 = \omega_{NH}$ multiplet structures in the COSY spectra. These assignments were confirmed by observation of remote peaks at $\omega_1 = \omega_{\alpha} + \omega_{\alpha'}$, $\omega_2 = \omega_{NH}$ in the 2Q spectrum acquired from ^1H_2O solution. The 2Q spectrum also provided assignments for the remaining two glycine spin systems. Complete spin systems for the four alanine and three threonine residues were identified from R-COSY and DR-COSY spectra acquired from ^1H_2O solution, respectively. The three valine spin systems could also be completely identified in DR-COSY spectra acquired from ^1H_2O solution, from connectivities from the C^{α}, C^{β}, and C^{γ} protons to the backbone NH. The cross-peaks for Val-2 were observed only in the spectrum at pH 4.5. All of these assignments were confirmed in the 2QF-COSY and 2Q spectra acquired in ^2H_2O solution.

The complete spin systems of Leu-73 and Leu-75 were identified through observation of relayed connectivities from $C^{\delta}H_3$, $C^{\gamma}H$, $C^{\beta}H_2$, and $C^{\alpha}H$ to the backbone amide proton in DR-COSY spectra acquired from ^1H_2O solutions. The observation of relayed connectivities over six bonds from methyl group to backbone amide is possible because the $C^{\beta}H_2$ and $C^{\gamma}H$ resonances are degenerate for Leu-75 and strongly coupled for Leu-73 (Table I).[3] We note that the long-range relayed connectivities are stronger in the spectrum with τ_1 = 45 ms than in the spectra with τ_1 = 22 ms. The spin systems of Leu-4, Leu-45, Leu-54, and Leu-63 were identified from the complete set of side-chain connectivities relayed to the backbone amide in TOCSY spectra acquired with 80- and 100-ms mixing periods. The complete spin system of the remaining leucine (Leu-19) and both isoleucine residues (Ile-43 and Ile-60) were assigned by the coincidence of resonances in side-chain terminal based and backbone amide based spin subsystems identified via multistep relayed connectivities. The Leu-54 spin system could not initially be distinguished from the isoleucine spin systems because the methyl resonances are accidentally degenerate. However, both isoleucines can be identified unambiguously from $C^{\gamma}H_3/C^{\delta}H_3$ relayed connectivities. For all of the leucine and isoleucine spin systems unambiguous assignment of the $C^{\beta}H_2$ and $C^{\gamma}H_2$ resonances is obtained from analysis of the MQ and MQF-COSY spectra acquired in ^2H_2O.

We wish to stress that the identification of complete leucine and isoleucine spin systems depends on observation of relayed

[2] The side-chain NH of arginine is not included because its resonance is exchange broadened at pH >5.

[3] The corresponding extra step of relayed connectivity between methyl and C^{α} protons is observed in the R-COSY spectra. This phenomenon has been discussed in Kay et al. (1987).

Table I: ^{1}H NMR Chemical Shifts of C3a (pH 5.5, 300 K)[a]

residue	chemical shifts (ppm)					
	NH	C^α	C^β	C^γ	C^δ	other
S1	–	4.20	3.93, 3.97			
V2	8.62	4.15	2.05	0.89, 0.92		
Q3	8.58	4.33	1.98, 2.10	2.38, (2.38)[b]		6.82, 7.48 ($N^\epsilon H_2$)
L4	8.44	4.27	1.68, 1.68	1.55	0.76, 0.81	
T5	8.03	4.01	4.18	1.20		
E6	8.39	4.14	2.04, 2.04	2.26, (2.26)[b]		
K7	8.14	4.25	1.84, 1.89	1.48, (1.48)[b]	1.66, 1.72	2.97, 2.97 ($C^\epsilon H_2$)
R8	8.39	4.09	1.69, 1.93	1.80, 1.80	3.20, 3.20	7.33 ($N^\epsilon H$)
M9	8.25	4.26	2.09, 2.09	2.57, 2.69		2.04 ($C^\epsilon H_3$)
N10	8.14	4.35	2.60, 2.70			
K11	7.54	3.94	1.50, 1.89	[1.50, 1.50	0.87, 1.53	2.27, 3.02 ($C^\epsilon H_2$)][c]
V12	7.78	3.72	2.04	0.95, 0.98		
G13	7.60	3.84, 3.84				
K14	7.35	4.04	1.52, 1.56	1.35, 1.35	1.47, 1.47	2.86, 2.86 ($C^\epsilon H_2$)
Y15	7.45	4.64	2.60, 2.90		7.22	6.66 ($C^\epsilon H$)
P16	–	4.54	1.98, 2.48	2.07, 2.18	3.77, 4.06	
K17	8.92	3.65	1.83, 1.91	1.48, 1.48	1.71, 1.71	3.01, 3.01 ($C^\epsilon H_2$)
E18	9.80	4.25	2.10, 2.14	2.32, 2.39		
L19	8.00	4.71	1.65, 1.70	1.60	0.88, 0.99	
R20	7.60	3.61	1.79, 1.82	1.69, 1.89	3.09, 3.10	7.30 ($N^\epsilon H$)
K21	8.62	4.10	1.82, 1.90	1.49, 1.49	1.62, 1.71	3.02, 3.02 ($C^\epsilon H_2$)
C22	7.14	4.43	3.38, 3.80			
C23	7.00	3.91	2.39, 3.05			
E24	8.40	3.64	1.98, 2.08	2.25, 2.30		
D25	8.84	4.37	2.78, 2.97			
G26	7.44	2.25, 2.95				
M27	7.30	4.33	2.00, 2.14	2.28 (2.28)[b]		2.18 ($C^\epsilon H_3$)
R28	7.02	4.14	1.69, 1.90	1.82, 1.98	3.28, 3.30	7.39 ($N^\epsilon H$)
Q29	8.80	4.10	1.89, 1.96	2.27, (2.27)[b]		
N30	8.84	5.20	2.50, 2.80			7.43, 7.73 ($N^\delta H_2$)
P31	–	4.34	1.94, 2.30	1.91, 2.00	3.64, 3.82	
M32	7.88	4.22	1.18, 1.33	2.23, 2.37		1.99 ($C^\epsilon H_3$)
R33	7.84	3.86	1.78, 1.96	1.46, 1.46	3.12, 3.16	7.09 ($N^\epsilon H$)
F34	7.69	4.97	2.82, 3.20		7.19	7.35 ($C^\epsilon H$); 7.31 ($C^\zeta H$)
S35	9.01	4.47	4.12, 4.40			
C36	9.10	4.36	2.90, 3.00			
Q37	8.78	3.93	2.01, 2.06	2.42, (2.42)[b]		6.92, 7.56 ($N^\epsilon H_2$)
R38	8.06	3.93	1.89, 1.89	1.47, 1.47	2.81, 2.81	
R39	8.06	4.04	2.22, 2.37	1.33, 1.33	2.86, 3.02	6.76 ($N^\epsilon H$)
T40	7.60	3.65	4.12	1.21		
R41	7.23	3.93	1.42, 1.52	1.22, 1.33	2.93, 2.93	7.04 ($N^\epsilon H$)
F42	7.38	4.60	2.82, 3.43		7.28	7.16 ($C^\epsilon H$); 7.10 ($C^\zeta H$)
I43	7.14	4.24	1.78	0.98, 1.09	0.44	0.76 ($C^\gamma H_3$)
S44	8.62	4.56	3.79, 3.82			
L45	7.37	4.44	1.48, 1.51	1.50	0.69, 0.85	
G46	8.36	4.00, 4.16				
E47	8.56	3.84	1.98, 2.05	2.30, (2.30)[b]		
A48	8.57	4.18	1.46			
C49	8.06	4.33	2.88, 3.13			
K50	8.58	3.73	1.80, 1.86	1.32, 1.32	1.62, 1.62	2.88, 2.92 ($C^\epsilon H_2$)
K51	7.88	4.01	1.99, 2.03	1.48, 1.50	1.72, 1.79	3.01, 3.01 ($C^\epsilon H_2$)
V52	7.48	3.52	1.95	0.18, 1.09		
F53	8.91	3.97	2.92, 3.39			7.0–7.1 ($C^\delta H$, $C^\epsilon H$, $C^\zeta H$)
L54	8.64	3.88	1.76, 1.96	1.80	0.92, 0.92	
D55	8.32	4.46	2.83, 2.92			
C56	8.90	4.16	3.10, 3.65			
C57	8.89	4.10	2.71, 3.03			
N58	9.38	4.45	2.88, 3.03			6.73, 7.67 ($N^\delta H_2$)
Y59	8.08	4.37	3.25, 3.32		7.14	6.73 ($C^\epsilon H$)
I60	8.08	3.96	1.92	1.45, 1.52	0.99	1.18 ($C^\gamma H_3$)
T61	8.13	3.86	4.42	1.26		5.46 ($O^\gamma H$)
E62	7.77	4.16	2.08, 2.08	2.29, 2.32		
L63	7.61	3.94	1.52, 1.52	1.38	0.59, 0.72	
R64	8.30	3.97	1.85, 1.93	1.66, 1.66	3.09, 3.12	7.06 ($N^\epsilon H$)
R65	7.82	4.12	1.94, 1.94	1.61, 1.77	3.16, 3.20	7.50 ($N^\epsilon H$)
Q66	8.08	4.05	2.08, 2.11	2.32, 2.48		6.77, 7.58 ($N^\epsilon H_2$)
H67	8.20	4.48	3.13, 3.25		8.22	7.19 ($C^\epsilon H$)
A68	8.02	4.22	1.46			
R69	8.02	4.22	1.81, 1.88	1.64, 1.71	3.18, 3.18	7.21 ($N^\epsilon H$)
A70	8.05	4.26	1.39			
S71	8.08	4.34	3.78, 3.83			
H72	8.25	4.63	3.18, 3.28		8.37	7.19 ($C^\epsilon H$)
L73	8.11	4.30	1.63, 1.56	1.50	0.84, 0.88	
G74	8.33	3.93, 3.93				
L75	8.00	4.35	1.58, 1.58	1.57	0.83, 0.88	
A76	8.24	4.31	1.35			
R77	7.85	4.12	1.68, 1.83	1.56, 1.56	3.16, 3.16	7.16 ($N^\epsilon H$)

[a] Chemical shifts are referenced to the H_2O signal at 4.75 ppm and are generally accurate to ±0.01 ppm (±0.03 ppm for geminal protons separated by <0.1 ppm). [b] The degeneracy of the C^γ protons is inferred. [c] The side-chain spin subsystem of Lys-11 could not be directly correlated with the backbone spin subsystem.

cross-peaks in a region of the spectrum already crowded with the direct $C^{\beta}H/C^{\gamma}H_3$ and $C^{\gamma}H/C^{\delta}H_3$ cross-peaks (typically, ω_1 = 0.5–2.5, ω_2 = 0.0–1.2 ppm). As in a previous study (Chazin et al., 1988), phase-sensitive R-COSY and DR-COSY experiments were found to be superior to TOCSY for this analysis because partly overlapped cross-peaks are easier to discriminate when they have antiphase multiplet structure (R-COSY and DR-COSY) as opposed to pure in-phase character (TOCSY). However, for obtaining multistep relayed connectivities over five or more bonds, TOCSY is clearly more efficient.

While it is possible in some cases to identify complete arginine and lysine spin systems via observation of multistep relayed connectivities from all side-chain protons to the backbone amide proton [e.g., Chazin et al. (1988)], the strategy of matching side-chain terminal based and backbone amide based spin subsystems (Chazin et al., 1987) is of more general use. Six of the eleven arginine and two of the seven lysine residues in C3a were completely identified from relayed connectivities to the amide proton, while the others were identified with the dual strategy. Discrimination between the C^{β}, C^{γ}, and C^{δ} proton resonances relied upon MQF-COSY and MQ spectra (Chazin et al., 1987). The identification of arginine spin systems with the dual strategy was greatly aided by observation of relayed connectivities to the side-chain NH for nine of the eleven arginines. Resolution of the highly overlapped lysine and arginine side-chain terminal based spin subsystems was greatly facilitated by the spectral simplification obtained in the 3Q spectrum.

(*D*) *Identification of Spin Systems for Residues with 3-Spin Side Chains.* Seventeen of the 22 spin systems with 3-spin side chains were completely identified on the basis of relayed connectivities to the backbone amide proton. Each assignment established via relayed connectivity was verified through observation of at least one of the following: (i) cross-peaks in the 3QF-COSY spectrum; (ii) direct and/or remote peaks in 2Q spectra; (iii) direct peaks in the 3Q spectrum.

Four of the other 3-spin spin systems (Ser-35, Ser-44, Phe-53, Tyr-59) exhibit only one $C^{\beta}H$ relayed connectivity, and there are none for Ser-1 because it has no observable resonance for the free terminal NH_2. Assignments for the remaining $C^{\beta}H$ resonances were unambiguously obtained in the MQF-COSY and MQ spectra. Cross-peaks observed in the 3QF-COSY spectrum provide the assignments of the $C^{\beta}H$ resonances of Ser-35, Phe-53, and Tyr-59. Remote peaks at $\omega_1 = \omega_{\beta+}\omega_{\beta'}$, $\omega_2 = \omega_{\alpha}$ in the 2Q spectrum confirm these assignments and provide unambiguous assignment of the nearly degenerate $C^{\beta}H$ resonances of Ser-1 and Ser-44.

Four spin systems were initially assigned to the four serine residues on the basis of the average of the $C^{\beta}H$ chemical shifts (>3.8 ppm). This was confirmed in one case (Ser-35) by the small value of $^3J_{\beta\beta}$ measured in a COSY spectrum (Neuhaus et al., 1985). Ser-1 was specifically identified from the absence of a backbone amide resonance, particularly narrow line widths, and very intense cross-peaks in 2D spectra. Subsequent sequential assignments confirmed the serine spin system identification.

The assignments of two of the three phenylalanines, the two tyrosines, the two histidines, and four of the five asparagines were completed by observation of NOEs between side-chain amide or ring protons and the C^{β}, C^{α}, or backbone amide protons. A scalar connectivity was observed between the C^{δ} and C^{β} protons of His-72. For the first time, this was observed not only in 2Q spectra (Dalvit et al., 1987; Chazin et al., 1988) but also in a DR-COSY spectrum.

(*E*) *Identification of the Spin Systems of Residues with 5-Spin Side Chains.* For 11 of the 12 5-spin spin systems, side-chain resonances were identified on the basis of relayed connectivities to the backbone amide proton. These spin systems were distinguished from other long side chain residues by a pattern of chemical shifts with C^{γ} at lower field than C^{β}. The C^{β} protons were identified from $C^{\alpha}H/C^{\beta}H$ connectivities in MQ and MQF-COSY spectra, as described above for the 3-spin spin systems. The remaining relayed connectivities were then assigned to the C^{γ} protons. These assignments were verified, where possible, by observation of direct $C^{\beta}H/C^{\gamma}H$ connectivities in MQF-COSY and MQ spectra.

For the 12th spin system (Met-27), only two relayed connectivities to the backbone amide proton were observed, and these were unambiguously assigned to the C^{β} protons in MQ and MQF-COSY experiments. The C^{γ} protons were identified on the basis of a relayed connectivity to the C^{α} proton in TOCSY spectra acquired from ^2H_2O solutions.

The C^{γ} resonances appear to be degenerate for each of Gln-3, Glu-6, Met-27, Gln-29, Gln-37, and Glu-47. While no direct experimental evidence (e.g., MQ remote peaks) for degeneracy could be obtained, the assumption of degeneracy for C^{γ} resonances is supported by the small chemical shift dispersion of the C^{β} protons.

Three of the four glutamine and the three methionine spin systems were distinguished on the basis of characteristic intraresidue NOEs to side-chain amide protons and $C^{\epsilon}H_3$ resonances, respectively. The remaining spin systems in the 5-spin side-chain category must arise from the remaining glutamine and the five glutamic acid residues.

(*F*) *Identification of Proline Spin Systems.* The two proline spin systems were identified after the C^{α} and C^{β} protons of all other spin systems were assigned. The identification of the two proline $C^{\alpha}H$–$C^{\beta}H_2$ spin subsystems was straightforward because all other spin subsystems observed in the appropriate region of the spectra can be assigned to other amino acids. Complete assignments were obtained by observation of relayed connectivities to both the C^{α} and $C^{\delta}H$ proton resonances. Detailed analysis of 3QF-COSY, 2Q, and 3Q verified the spin system assignments based on relayed connectivities and unambiguously distinguished between C^{β} and C^{γ} proton resonances.

At the end of the spin system identification stage of analysis, all 74 backbone amide protons and 463 of the 469 nonlabile side-chain protons had been identified and assigned to a specific (but not stereospecific) proton. The remaining six nonlabile protons ($C^{\gamma}H_2$, $C^{\delta}H_2$, and $C^{\epsilon}H_2$ of Lys-11) could only be tentatively assigned. In addition, 10 of the 14 side-chain amide protons, 10 of the 11 arginine guanidinium N^{ϵ} protons, and the O^{γ} proton of Thr-61 were identified and assigned to the appropriate spin systems. A total of 60 spin systems could be assigned unambiguously to the appropriate amino acid, while the remaining 17 were assigned to one of three groups: ten 3-spin side-chain residues (one phenylalanine, one asparagine, two aspartic acid, six cysteine); six 5-spin side-chain residues (one glutamine, five glutamic acid); one lysine residue.

Sequence-Specific Assignment. Sequence-specific assignments for C3a were made according to the standard sequential assignment procedure (Billeter et al., 1982) and are listed in Table I. We adopt the shorthand notation of Wüthrich et al. (1984) for specifying short proton–proton distances and corresponding NOE connectivities. The sequential assignments were made by using data from NOESY spectra recorded with mixing times between 40 and 200 ms at four different pHs

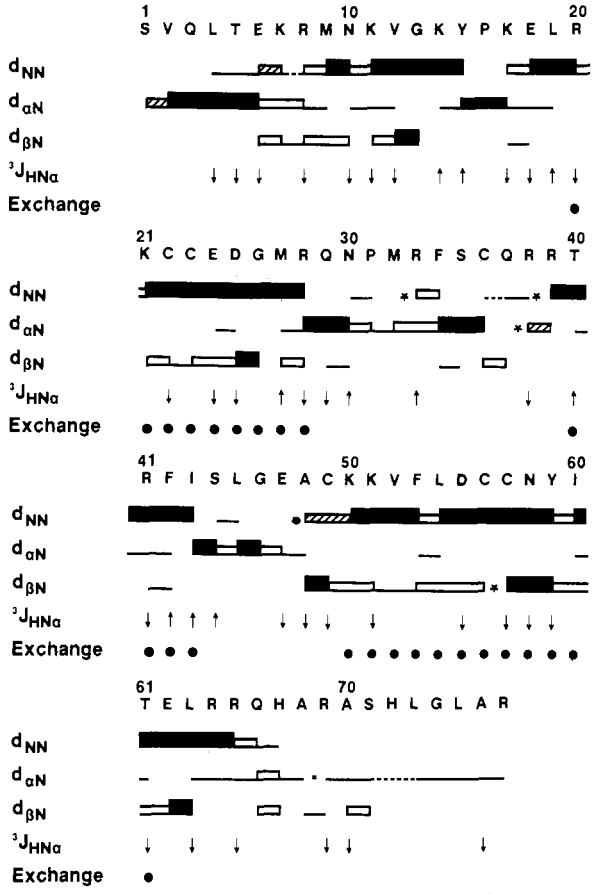

FIGURE 3: Summary of sequential NOE connectivities, backbone coupling constants ($^3J_{HN\alpha}$), and slow amide proton exchange rates observed in C3a and used to characterize secondary structure. The one-letter code for the sequence is given at the top. Characteristic NOE connectivities [$d_{NN}(i,i+1)$; $d_{\alpha N}(i,i+1)$; $d_{\beta N}(i,i+1)$] are indicated by lines or bars between the two residues. The height of the bars give a qualitative measure of the relative strength of the NOE in the τ_m = 40 ms NOESY spectrum. Connectivities that could not be identified in this spectrum, but were identified at another pH or temperature, are crosshatched or dashed, while those not identified due to degeneracy of resonances are labeled with an asterisk (*). Values of $^3J_{HN\alpha}$ are classified as small (<6.5 Hz, see text) or large (>8 Hz) as indicated by (↓) and (↑), respectively. Backbone amide protons that exchange slowly are indicated with a dot.

and three temperatures. These are estimated to provide cross-peaks for interproton distances up to 5 Å. A summary of the sequential NOE connectivities is included in Figure 3. Relative intensities were estimated in the τ_m = 40 ms spectrum. We stress that complete spin system identification and classification in terms of amino acid type greatly facilitate the sequential assignment procedure.

For each pair of consecutive residues in C3a, with the exception of A68/R69, at least one $d_{\alpha N}(i,i+1)$ or $d_{NN}(i,i+1)$ connectivity is observed (Figure 3). The identification of certain sequential connectivities is difficult at pH 5.5, due to the overlap of the corresponding pairs of amide protons. However, these could be obtained by varying pH and temperature. The majority of residues are in nonextended conformation, as distinguished by the observation of sequential $d_{NN}(i,i+1)$ and strong intraresidue $d_{N\alpha}(i,i)$ connectivities. Regions of the 80-ms NOESY spectrum containing d_{NN} connectivities for polypeptide segments Arg-8 to Tyr-15, Lys-17 to Arg-28, Thr-40 to Ile-43, and Lys-50 to Gln-66 are shown in Figure 4. Sequence-specific assignments for peptide segments with extended conformation were made on the basis of strong $d_{\alpha N}(i,i+1)$ NOE connectivities. An example is shown in Figure 5, for the C-terminal active pentapeptide Leu-73–Arg-77.

Secondary Structure. Helical regions can be identified on the basis of networks of sequential d_{NN} and medium-range [$(i,i+3)$, $(i,i+2)$, $(i,i+4)$] NOE connectivities, small values of $^3J_{HN\alpha}$, and amide proton exchange. The data for C3a are summarized in Figures 3 and 6. Helices are assigned to the segments Thr-5 to Tyr-15, Leu-19 to Arg-28, and Gly-46 to Ala-70, designated as helices I, II, and III, respectively. On the basis of the NOE connectivities, a single turn of distorted helix ($^3J_{HN\alpha}$ is large for three of the four residues) appears to be present in the segment Thr-40–Ile-43.

The ends of all three helices appear to fray open, as evidenced by increased values[4] of $^3J_{HN\alpha}$, higher rates of backbone amide proton exchange, a transition from strong sequential d_{NN} to $d_{\alpha N}$ connectivities, and the observation of few medium-range NOEs. This is particularly evident for the N- and C-terminal helices (helices I and III, respectively) which exhibit pronounced fraying from residue Arg-8 to Thr-5 and from Gln-66 to Ala-70. In these regions both d_{NN} and $d_{\alpha N}$ NOE connectivities are observed together with some weak medium-range NOEs, suggesting that the polypeptide chain adopts a significant population of helical structures which dynamically fray into more extended conformations.

Global Fold. On the basis of the known disulfide linkages, the elements of secondary structure determined from the NMR data and a few key tertiary NOEs, it is possible to give a schematic description of the global fold of C3a in solution (Figure 7). Building from the long C-terminal helix (helix III), helix II is positioned by the disulfide linkages from Cys-22 to Cys-49 and from Cys-23 to Cys-56. Since these helices are packed in a parallel fashion, the intervening irregular polypeptide (Gln-29 to Leu-45) is required to loop around from the end of helix II, run in the antiparallel direction for the length of helix III, and join helix III. The NMR data as summarized in Figures 3 and 6 indicate the presence of turns or loops in local segments of this antiparallel strand. However, the details of the conformation must await three-dimensional solution structure determination. We do note the presence of a single turn of distorted helix between Thr-40 and Ile-43. A disulfide linkage between Cys-36 and Cys-57 connects the Gln-29 to Leu-45 strand to helix III and aligns it relative to helices II and III. Helix I does not contain disulfide linkages but can be positioned by tertiary NOE connectivities observed between aromatic ring protons of Tyr-15 and side-chain protons of Cys-23 and Val-52 and between the methyl protons of Val-12 and side-chain protons of Cys-23, Glu-24, and Val-52. These NOEs indicate that helix I is in contact with and parallel to helices II and III, forming the mini helical bundle shown in Figure 7. Note that this model requires a reversal of the direction of the polypeptide chain over the polypeptide segment Tyr-15–Leu-19.

A clear gradient in the strength of the NOEs between helix I and the globular core of the protein is observed. NOEs to Tyr-15 are very strong and are clearly observed in NOESY spectra with short mixing times, those to Val-12 are significantly weaker and are first observed in spectra with 150-ms mixing times, and long-range contacts from the methyl protons of Met-9 to the side chain of Ile-60 are very weak and can only be observed in a NOESY spectrum acquired with a 400-ms mixing period. We interpret the NOE data as indicating fluctuations of helix I relative to the globular core, with a significant population of conformations in which the helix is

[4] Dynamic averaging with extended chain forms, for which $^3J_{HN\alpha}$ is greater than 8 Hz, will have the effect of increasing the observed coupling constant.

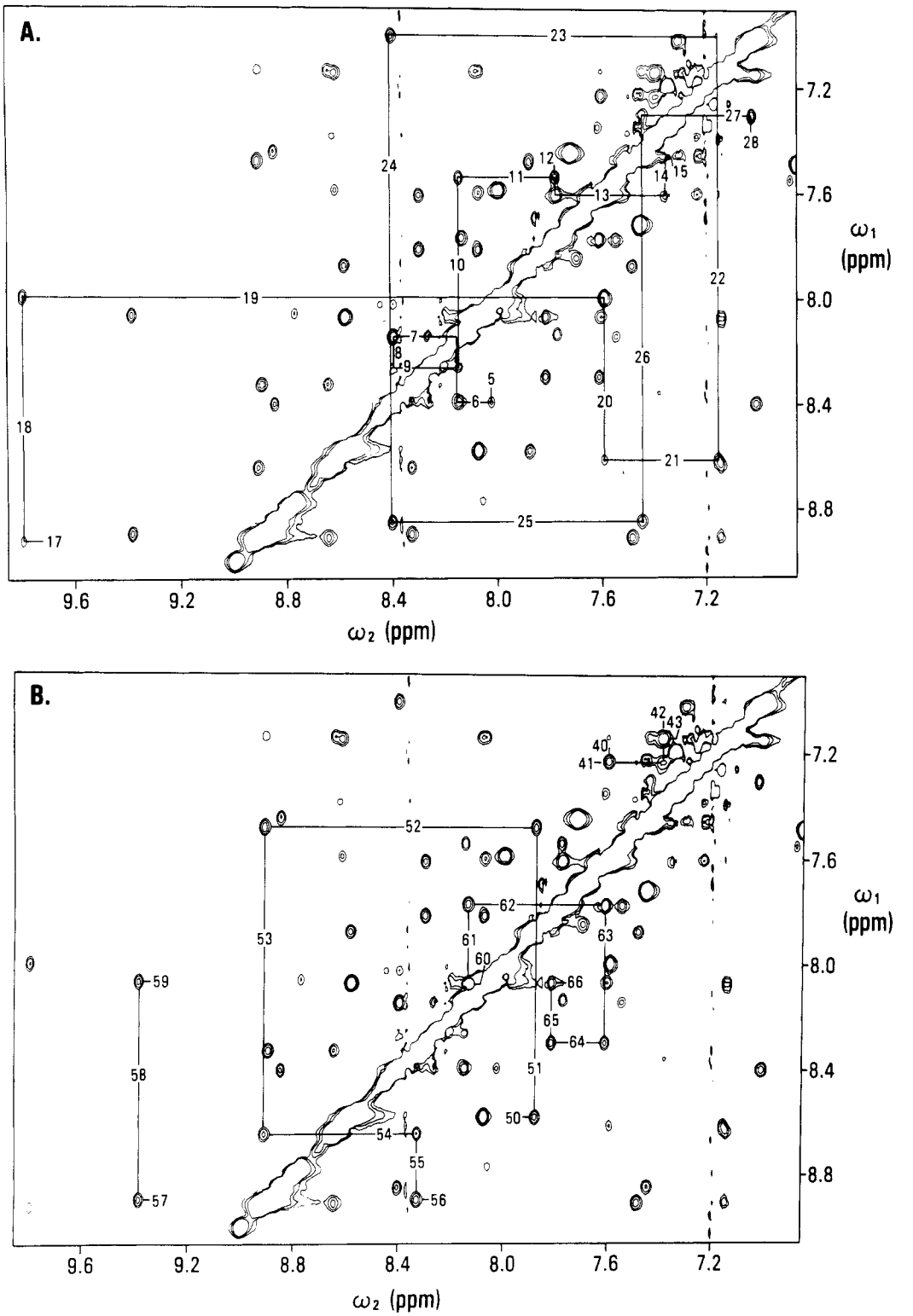

FIGURE 4: Sequential resonance assignments for polypeptide segments in helical conformation. Region of a 500-MHz Hahn-echo NOESY spectrum (τ_m = 80 ms) of a 5 mM solution of C3a at pH 5.5 and 303 K showing d_{NN} connectivities. Lines are drawn between pairs of $d_{NN}(i,i-1)$ and $d_{NN}(i,i+1)$ cross-peaks and labeled with the sequence-specific assignment. The experimental data are shown for Thr-5 to Tyr-15 and Lys-17 to Arg-28 in (A) and for Thr-40 to Ile-43 and Lys-50 to Gln-66 in (B).

packed into the cleft between helices II and III as shown in Figure 7. However, an alternative explanation is possible, wherein the packing is relatively static with helix I extending away from the globular core at an acute angle which results in progressively longer distances between the side chains of helix I and those of helices II and III. In view of the dynamic nature (high side-chain and backbone mobility and helix fraying) of the N-terminal end of helix I in solution, we favor the dynamic over the static model.

Discussion

These studies confirm that residues Gly-13–Arg-77, all of which were observed in the X-ray structure of C3a (Huber et al., 1980), adopt essentially the same global conformation in solution as in the crystal state. However, local differences in conformation are observed. Residues Ser-1–Val-12, which

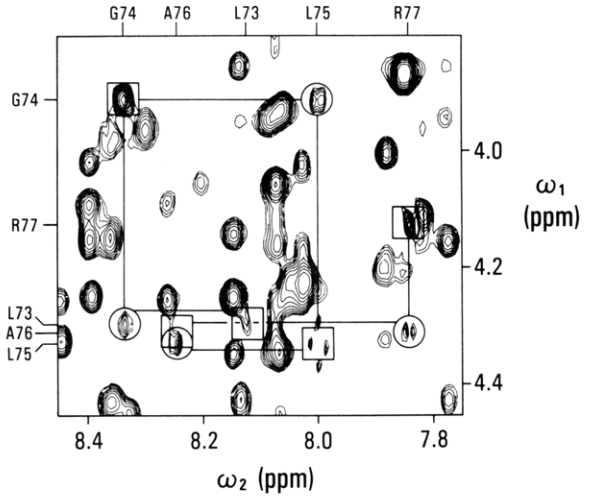

FIGURE 5: Sequential resonance assignments for the required "active" pentapeptide segment (Leu-73–Arg-77) of C3a. A portion of the $d_{\alpha N}$ region of the same spectrum as in Figure 4 is shown. Circles are drawn around cross-peaks corresponding to $d_{\alpha N}(i,i+1)$ NOE connectivities, which are joined by a vertical line to squares drawn to indicate the location of the corresponding COSY cross-peak for each residue. A horizontal line is drawn between the corresponding intraresidue COSY square and the sequential NOESY cross-peak. Sequence-specific assignments of the backbone NH resonances are indicated above and of the $C^{\alpha}H$ resonances to the left.

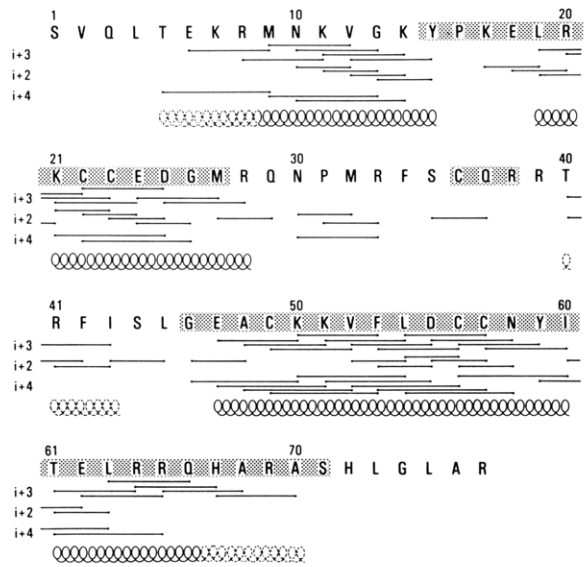

FIGURE 6: Summary of medium-range NOE connectivities used to identify the helical segments of C3a in solution. The one-letter code for the sequence is given at the top. The locations of helices in the X-ray structure are indicated by shading. Characteristic NOE connectivities [$d_{\alpha N}(i,i+3)$ or $d_{\alpha\beta}(i,i+3)$; $d_{\alpha N}(i,i+2)$; $d_{\alpha N}(i,i+4)$] are indicated by lines between the two residues. On the bottom line, the locations of helices identified in the solution structure are indicated by coils.

were not seen by X-ray analysis, are observed in solution, with a well-defined helical segment from Met-9 to Tyr-15. The NMR experiments provide evidence that an additional turn of helix, from Thr-5 to Arg-8, forms transiently in solution. The N-terminus appears to be quite flexible in solution, with dynamic fraying of the N-terminal helix and movement of the whole helix relative to the core of the protein. The long mixing time NOESY experiments indicate that a transient structure is formed in which the N-terminal helix (helix I) is positioned between the two other helices in the folded protein, as shown schematically in Figure 7. The absence of electron density for residues 1–12 in the X-ray structure (Huber et al., 1980)

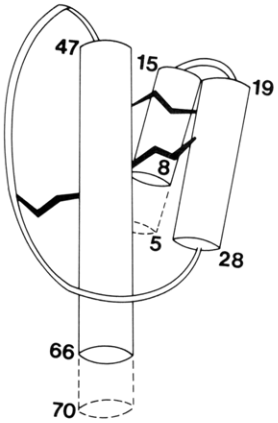

FIGURE 7: Schematic diagram of the chain folding of C3a in solution. The three helical elements of secondary structure are indicated by cylinders, nonregular structure by ribbon, and the disulfide linkages by lightning bolts. Dashed ends of cylinders identify dynamic portions of the helices that largely (but not completely) populate helical conformation. The Thr-5 to Tyr-15 helix is packed between the other two helices and behind them in this projection. Our model also includes a certain additional dynamic motion of this N-terminal helix; i.e., it spends significant periods of time not packed against the other two helices.

may be due to similar conformational disorder at the N-terminus in the crystalline state.

Subtle differences exist in the location of the helical segments in solution as compared to the crystal structure (Huber et al., 1980). In the crystal structure of C3a the C-terminal helix extends from Glu-47 to Ser-71, whereas in solution the helix extends only to Ala-70. In addition, the helix is progressively less well-defined in solution from Arg-66 to Ala-70, where we observe only a transient (but significant) population of helix. In the crystalline state, two C3a molecules associate closely to form a dimer with their C-terminal helices packed in an antiparallel arrangement (Huber et al., 1980). It is presumably these packing interactioins which stabilize the helix in the crystal. Other regions which differ between crystal and solution are the single turn of helix between Cys-36 and Arg-38 observed in the crystal but not in solution and the C-terminal end of helix II which extends one additional residue to Arg-28 in solution.

A more substantial difference is observed for the residues at the N-terminal end of helix II. In the crystal structure, helix II extends to Tyr-15, whereas in solution helix II terminates at Leu-19 and helix I extends from the N-terminal region to Tyr-15. In the crystal structure a chain reversal is observed just beyond the N-terminal end of helix II, between Gly-13 and Tyr-15, whereas in solution these residues are part of helix I. In solution, the reversal must occur between the ends of helix I (Tyr-15) and helix II (Leu-19). Differences between the crystal and solution structures of C3a are also characterized by the unique placement of the aromatic ring of Tyr-15. In the crystal this aromatic ring is oriented away from the surface of the protein, whereas in solution a large number of long-range NOE connectivities clearly indicate that the ring is packed into a hydrophobic pocket between helices I, II, and III. This placement of the Tyr-15 side chain is consistent with the radioiodination pattern of native C3a, which shows approximately a 5:1 preference for labeling Tyr-59 compared to Tyr-15 on the basis of distribution of ^{125}I in isolated cyanogen bromide fragments of human C3a (Hugli et al., 1975b). This labeling pattern suggests that in solution Tyr-15 is relatively inaccessible, in agreement with the NMR results.

The NMR data indicate a general tendency for the region of C3a that is adjacent to the putative receptor binding site

(i.e., Leu-73 to Arg-77) to assume helical conformation. Although the C-terminal region of the C3a molecule assumes less helical structure in solution than is evident in the crystalline form, helix is clearly identified as the predominant form of ordered structure. A number of studies have indicated the importance of helix formation for activity. Analogues of the fully active 21-residue C3a peptide Cys-57 to Arg-77 (Lu et al., 1984) were purposefully designed to explore the contribution of helix to activity. When peptides of 21 residues were synthesized with helix-promoting amino acids such as α-aminobutyric acid strategically placed to enhance helix formation, activity of the product was increased over that of analogues based on the natural sequence (Hoeprich & Hugli, 1986). Conversely, when several prolyl residues were incorporated in the 21-residue C3a fragment to disrupt helix formation, activity fell to the same low level as the pentapeptide Leu-73 to Arg-77. In aqueous media the C-terminal region of C3a appears somewhat less regular and more mobile than in the crystal, and in particular, the essential residues including Leu-73 to Arg-77 are freely mobile. These results suggest that the exact conformation of the C3a binding region that interacts with cellular C3a receptors may be induced by or selected for at the receptor surface. In addition, from the detailed residue by residue comparison of C3a and des-Arg^{77}-C3a which shows that in solution the des-Arg form adopts an identical conformation with that of the intact protein, it is now certain that the critical role of Arg-77 in activity is not based on conformational effects (in agreement with the peptide studies).

We have also examined the pH dependence of the solution conformation of C3a over the pH range 4.5–6.8. The NMR spectra show that the structure of C3a is relatively insensitive to pH changes over this range; changes in crystalline order over the pH range 4.5–7.5 (Huber et al., 1980) are thus unlikely to arise from conformational changes and probably reflect changes in crystal packing interactions.

Due to homology in primary structure, genetic origin, and biological activity, C3a should serve as a good structural model for C4a and C5a. Greer (1985) has modeled the structure of C5a on the basis of the crystal structure of C3a and provides convincing evidence for the structural homology of C3a and C5a. The most interesting aspect of his study with respect to our findings is the assignment of helical conformation for the N-terminal residues of C5a and the docking of this helix into a hydrophobic cleft between the long central helix and the helix from residue 17 to residue 27. This is precisely what we observe in our studies of C3a. Greer suggests that the corresponding N-terminal peptide of C3a is likely to be less helical and is unlikely to be docked into the cleft between helices I and II. The NMR studies partially support his conclusions in that the N-terminal helix of C3a in solution is clearly flexible, with significant fraying toward the N-terminus. However, a significant population of conformers is found in which the N-terminal helix packs against the globular core of C3a. Comparison of our results on C3a with ongoing NMR studies of C5a (E. Zuiderweg, private communication) should provide important insights into the relative stabilities of the corresponding N-terminal helices and the nature of the packing interactions with the globular core and may prove valuable for stimulating new therapeutic approaches to controlling the inflammatory action of the anaphylatoxins.

Added in Proof

Following submission of this manuscript, 2D NMR studies of C3a and C5a have been published by Zuiderweg and coworkers (Nettesheim et al., 1988; Zuiderweg et al., 1988). Their independent study of C3a reaches similar conclusions regarding secondary structure. However, the long-range NOEs that establish docking of the N-terminal helix onto the globular core were not identified.

Acknowledgments

We thank Dr. Mark Rance for continued support with experimental methods, Marleen Kawahara for carrying out the protein purifications, Dr. Keith Cross for acquiring preliminary NMR data, Dr. Claudio Dalvit for helpful discussions, Lonnie Harvey for assistance in preparation of the manuscript, and Prof. R. Huber for providing the refined coordinates of crystalline C3a.

Registry No. C3a, 80295-42-7.

References

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Bokisch, V. A., & Muller-Eberhard, H. J. (1970) *J. Clin. Invest. 49*, 2427–2436.

Chazin, W. J., & Wright, P. E. (1987) *Biopolymers 26(6)*, 973–977.

Chazin, W. J., & Wüthrich, K. (1987) *J. Magn. Reson. 72(2)*, 358–363.

Chazin, W. J., Rance, M., & Wright, P. E. (1987) *FEBS Lett. 222*, 109–114.

Chazin, W. J., Rance, M., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 603–626.

Clore, G. M., & Gronenborn, A. M. (1983) *J. Magn. Reson. 53*, 423–426.

Dalvit, C., Wright, P. E., & Rance, M. (1987) *J. Magn. Reson. 71*, 539–543.

Greer, J. (1985) *Science (Washington, D.C.) 228*, 1055–1060.

Hoeprich, P. D., & Hugli, T. E. (1986) *Biochemistry 25*, 1945–1950.

Huber, R., Scholze, H., & Deisenhoffer, J. (1980) *Hoppe-Seyler's Z. Physiol. Chem. 361*, 1389–1399.

Hugli, T. E. (1981) *CRC Crit. Rev. Immunol. 1(4)*, 321–366.

Hugli, T. E. (1984) *Springer Semin. Immunopathol. 7*, 192–219.

Hugli, T. E., Vallota, E. H., & Muller-Eberhard, H. J. (1975a) *J. Biol. Chem. 250*, 1472–1478.

Hugli, T. E., Morgan, W. T., & Muller-Eberhard, H. J. (1975b) *J. Biol. Chem. 250*, 1479–1483.

Hugli, T. E., Gerard, C., Kawahara, M., Scheetz, M. E., Barton, R., Briggs, S., Koppel, G., & Russell, S. (1982) *Mol. Cell. Biochem. 41*, 59–66.

Kay, L. E., Jones, P.-J., & Prestegard, J. H. (1987) *J. Magn. Reson. 72*, 392–396.

Lu, Z-xian, Fok, K. F., Erickson, B. W., & Hugli, T. E. (1984) *J. Biol. Chem. 259*, 7367–7370.

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

Muto, Y., Fukumoto, Y., & Arata, Y. (1985) *Biochemistry 24*, 6659–6665.

Muto, Y., Fukumoto, Y., & Arata, Y. (1987) *J. Biochem. (Tokyo) 102*, 635–641.

Nettesheim, D. G., Edalji, R. P., Mollison, K. W., Greer, J., & Zuiderweg, E. R. P. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5036–5040.

Neuhaus, D., Wagner, G., Vašák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257–273.

Rance, M. (1987) *J. Magn. Reson. 74*, 557–564.

Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.

Zuiderweg, E. R. P., Mollison, K. W., Henkin, J., & Carter, G. W. (1988) *Biochemistry 27*, 3568–3580.