Neufeld, A. H., & Levy, H. M. (1969) *J. Biol. Chem. 244*, 6493–6497.
Norby, J. G., & Jensen, J. (1971) *Biochim. Biophys. Acta 233*, 104–116.
Peters, K., & Richards, F. M. (1977) *Annu. Rev. Biochem. 46*, 523–551.
Post, R. L., Sen, A. K., & Rosenthal, A. S. (1965) *J. Biol. Chem. 240*, 1437–1445.
Post, R. L., Kume, S., Tobin, T., Orcutt, B., & Sen, A. K. (1969) *J. Gen. Physiol. 54*, 3065–3265.
Post, R. L., Hegyvary, C., & Kume, S. (1972) *J. Biol. Chem. 247*, 6530–6540.
Robinson, J. D. (1967) *Biochemistry 6*, 3250–3258.
Robinson, J. D. (1974) *Biochim. Biophys. Acta 341*, 232–247.
Rogers, T. B., & Lazdunski, M. (1979) *FEBS Lett. 98*, 373–376.
Singer, S. J. (1974) *Annu. Rev. Biochem. 43*, 805–833.
Skou, J. C. (1974) *Ann. N.Y. Acad. Sci. 242*, 168–184.
Skou, J. C. (1979) *Biochim. Biophys. Acta 567*, 421–435.
Skou, J. C., & Esmann, M. (1979) *Biochim. Biophys. Acta 567*, 436–444.
Sweadner, K. J. (1977) *Biochem. Biophys. Res. Commun. 78*, 962–969.

Proton Magnetic Resonance Study of Crambin, a Hyperstable Hydrophobic Protein, at 250 and 600 MHz†

Miguel Llinás,* Antonio De Marco,‡ and Juliette T. J. Lecomte

ABSTRACT: Crambin is a 44 amino acid, molecular weight 5000, water-insoluble protein of such crystalline thermal ordering that it yields a high-resolution X-ray diffraction pattern to the interplanar spacing of 0.88 Å [Teeter, M. M., & Hendrickson, W. A. (1979) *J. Mol. Biol. 127*, 219–223]. The protein represents a unique model system to investigate the physical properties of hydrophobic polypeptides, ubiquitous in biological membranes. Crambin has been studied by using 250- and 600-MHz ^{1}H NMR spectroscopy. Organic solvents of intermediate polarity had to be used to dissolve the protein in a homogeneous phase. We find that both acetic acid, a weak Brønsted proton donor, and dimethylformamide, a weak Lewis nucleophile, perform satisfactorily in preserving essentially identical globular structures. The whole aliphatic spectrum exhibits a high degree of nondegeneracy, indicating a rather rigid structure. Most of the 24 methyl peaks can be accounted for as they yield relatively sharp signals within an ~1.4-ppm range, well dispersed at a magnetic field of 14.09 T. The phenylalanyl and the two tyrosyl side chains are well resolved and show line widths and resonance frequencies indicative of immobilization. In acetic-d_3 acid-*d*, crambin exhibits such a remarkable low rate of ^{1}H–^{2}H exchange that, at room temperature, deuteration is still incomplete after 20 days in solution. By this criterion, a minimum of 28 to 29 peptidyl amides are shown to be solvent protected, which reflects a large extent of intramolecular hydrogen bonding and sheltered locations. A few stable amide signals appear at field positions higher than the aromatic resonances, suggesting that a number of NH groups are occluded in a hydrophobic matrix. All of these features indicate that crambin is a protein of unprecedented structural stability.

Although high-resolution nuclear magnetic resonance (NMR)[1] spectroscopy is the method of choice for the study of conformational and dynamic features of proteins in solution, its application to such problems is often limited by the size of the molecule or its state of aggregation. Even with the present day availability of high-field spectrometers, the technique is of limited usefulness beyond M_r ~10–15 K as the spectra become extremely complex in terms of the number of mostly overlapping signals and much of the fine structure information is lost due to dipolar line broadening. This poses a serious problem in the study of hydrophobic proteins which are normally water insoluble and tend to aggregate when dissolved in less polar solvent systems.

Because of their fundamental role as components of biological membranes, hydrophobic polypeptides are increasingly becoming the focus of attention when attempting to understand basic cellular processes. However, due to the water insolubility that these proteins exhibit, structural studies have thus far been mostly limited to diffraction techniques of various sorts, centered on crystalline or highly ordered lamellar systems. It is clear that the wealth of information provided by such approaches would be well supplemented if dynamic features, relevant to their physiological roles, could also be investigated by, e.g., NMR spectroscopy. Although this appears to be a reasonable goal, the first requirement to be met is the need to solubilize the protein in a suitable, nondenaturing solvent that preserves the protein as a monomeric unit.

Crambin is an M_r 5000 hydrophobic protein found in the seed of the plant *Crambe abyssinica* (Van Etten et al., 1965). It has the peculiarity of readily crystallizing out of aqueous ethanol solutions in space group $P2_1$ needles that diffract X-rays strongly to the interplanar spacing limit of 0.88 Å; the resolution thus afforded is characteristic of crystals from small organic molecules (Van Etten et al., 1965; Teeter & Hen-

† From the Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213. *Received September 17, 1979.* This research was partly supported by the Petroleum Research Fund, administered by the American Chemical Society, Grant No. 9781-G6, and the National Institutes of Health, Grant GM 25213. The NMR Facility is supported by National Institutes of Health Grant RR 00292.

‡ A.D.M. held a NATO Senior Fellowship. Permanent address: Istituto di Chimica delle Macromolecole, Consiglio Nazionale delle Ricerche, Via A. Corti 12, 20133 Milano, Italy.

[1] Abbreviations used: DMF, dimethylformamide-d_7; Me_2SO, dimethyl-d_6 sulfoxide; M_r, molecular weight; NMR, nuclear magnetic resonance; ppm, parts per million; Me_4Si, tetramethylsilane; F_3AcOD, trifluoroacetic acid-d_1.

0006-2960/80/0419-1140$01.00/0

drickson, 1979). Such extreme thermal ordering is unknown for a polypeptide of this size and suggests a high rigidity for the protein. The amino acid composition (Van Etten et al., 1965) indicates the lack of histidine, lysine, methionine, and tryptophan and the presence of, approximately, 1 mol each of leucyl, phenylalanyl, and glutamyl residues, 2 mol each of arginyl, seryl, tyrosyl, and valyl residues, 4 mol each of glycyl, isoleucyl, and prolyl residues, and 5 mol each of threonyl, alanyl, and aspartyl residues. Three of the six carboxylated side chains are amidated. The protein also contains six cysteinyl residues which form three –S–S– bridges. Out of the 44 residues only 17 are polar, which is consistent with its water insolubility. Hydrodynamic and X-ray diffraction data suggest that the protein is, overall, globular (Van Etten et al., 1965; Teeter & Hendrickson, 1979).

The fact that the crambin primary through tertiary structures are presently being actively investigated (M. M. Teeter, personal communication; Teeter & Hendrickson, 1979) combined with the availability of the protein in high yield and purity by extraction from the seed source and crystallization prompted us to initiate spectroscopic studies of this most simple protein as a suitable model to foster understanding of the conformational dynamics characterizing hydrophobic polypeptides. In this communication we report initial high-resolution 250- and 600-MHz ^{1}H NMR spectroscopic investigations on crambin indicative that, indeed, the protein is endowed with a rather unique structural stability.

Materials and Methods

Crambin was extracted and crystallized to high purity from *C. abyssinica* seeds as reported (Van Etten et al., 1965). For the 250-MHz NMR studies, 5–7 mg of the protein was dissolved in 0.4 mL of solvent and contained in 5-mm NMR tubes. For the 600-MHz experiments, the solutions contained 15 mg of protein. Several solvent systems were tried. Besides water, the protein was found to be insoluble in acetone, acetonitrile, absolute ethanol, ethyl ether, hexanol, and 1,1,2-trichlorotrifluoroethane. Although the protein is soluble in aqueous ethanol, aqueous acetone, trifluoroethanol, pyridine, and 1:1 ethanol–Me_2SO, in these solvents the spectra are extremely broad, suggesting extensive aggregation. Satisfactory spectra were obtained for the protein dissolved in Me_2SO, acetic acid, DMF, and F_3AcOH. Spectra for these four solvents are discussed in the text. Aqueous acetic acid yields a spectrum suggestive of partial denaturation. Deuterated NMR solvents were purchased from Merck Sharp & Dohme Canada, Ltd., while the protonated, reagent grade, acetic acid originated from J. T. Baker Chemical Co.

The ^{1}H NMR spectra were recorded in the correlation mode (Dadok & Sprecher, 1974; Gupta et al., 1974) using the 250- and 600-MHz NMR spectrometers of the NMR Facility for Biomedical Research at Carnegie-Mellon University. Chemical shifts are expressed in parts per million from internal Me_4Si, which also served to provide the homonuclear lock signal. All experiments at 250 MHz were run at ~28 °C while the 600-MHz spectra correspond to a probe temperature of ~22 °C. When resolution-enhanced, the spectra were deconvoluted by using the three-parameter "optimal filter" (Ernst, 1966).

Results and Discussion

The ^{1}H NMR spectrum of crambin, aliphatic region, for the protein dissolved in acetic-d_3 acid-d, is shown in Figure 1. Such a spectrum is characteristic of undenatured proteins (McDonald & Phillips, 1967; Dwek, 1973; James, 1975; Wüthrich, 1976) and is indicative that in this solvent crambin

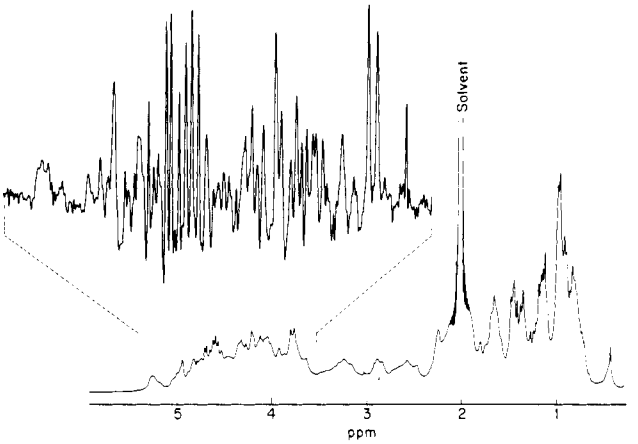

FIGURE 1: ^{1}H NMR spectrum of crambin at 250 MHz, aliphatic region. The insert expands, after resolution enhancement, the region between 3.5 and 5.3 ppm mostly populated by H^α resonances. Solvent acetic-d_3 acid-d; $T \approx 28$ °C; digital resolution 0.52 Hz; 1500 scans. The intense peak at ~2 ppm arises from residual protons in the solvent methyl group.

is not in a random coil conformation. The insert to Figure 1 shows an expansion of the 3.5–5.3-ppm region, after resolution enhancement. This region contains $^1H^\alpha$ resonances, and the lack of degeneracy exhibited by such a spectrum points toward the location of each proton in environments that are not averaged by conformational fluctuations. In particular, the variety of line widths depicted by the $^1H^\alpha$ region suggest that some groups are more buried and immobilized than others according to their positions in the molecular framework.

Although glacial acetic acid is a relatively weak Brønsted acid ($pK_a = 4.75$), it imposes a rather extreme environment for protein solvation. For the purpose of testing the extent of preservation of original conformational features, spectra were obtained for the protein dissolved in Me_2SO-d_6 and DMF-d_7, solvents that behave as weak Lewis bases and provide nucleophilic environments of properties opposed to that afforded by acetic acid. Figure 2 compares the high-field methyl region spectrum for crambin dissolved in Me_2SO-d_6 (B), DMF-d_7 (C), acetic-d_3 acid-d (D) with that of the protein dissolved in F_3AcOH-d_1 (A), a strongly denaturing solvent (McDonald & Phillips, 1969; Bak et al., 1968). Except for chemical shift differences well-known for the two solvents (Wüthrich, 1976), simple inspection shows common features for the spectra recorded from the Me_2SO and the F_3AcOD solutions while the overall appearance of the DMF and acetic acid spectra are indicative of conformational analogy for the protein in the two solvents. We interpret this as suggesting that in DMF and in acetic acid the protein lies in conformations that are closely related to the native structure, as otherwise the forces imposed by the two complementary solvents are more likely to lead to distinct final structures, globular or not, that should yield dissimilar spectra, which is not observed. The protein in DMF is currently being investigated and will be the subject of a separate communication. At this point we would like to indicate that the broad appearance of the DMF and Me_2SO spectra reflects aggregation that can be broken at higher temperatures.

On going from low to high fields, the F_3AcOD spectrum, Figure 2A, shows a doublet at ca. 1.67 ppm [(a), area = 1] and a group of signals clustered at ca. 1.58 [(b), area $\simeq$ 5], 1.46 [(c), area $\simeq$ 4], and 1.05 ppm [(d), area $\simeq$ 14]. These areas agree remarkably well with the side-chain methyl content of crambin and the expected methyl chemical shifts in trifluoroacetic acid as reported in the literature (Table I) so that

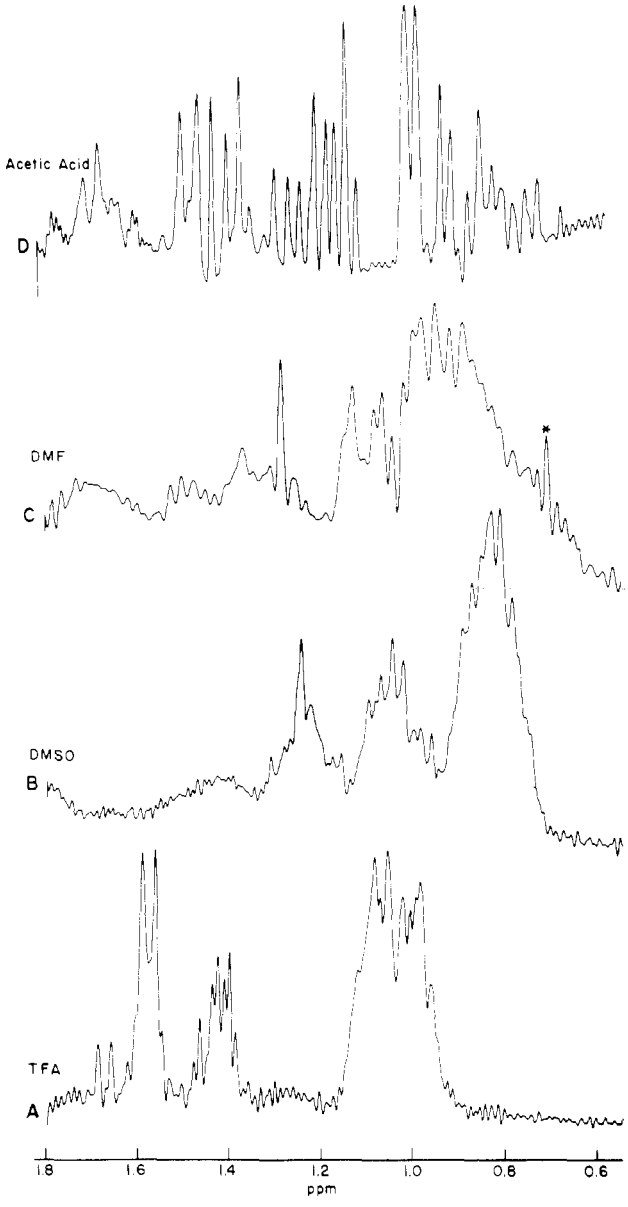

FIGURE 2: ^{1}H NMR spectra of crambin at 250 MHz, methyl region, after resolution enhancement. The solvents for each spectrum are (A) trifluoroacetic acid-*d*, (B) dimethyl-d_6 sulfoxide, (C) dimethylformamide-d_7, and (D) acetic-d_3 acid-*d*. $T \approx 28$ °C; digital resolution 0.52 Hz; 400–1500 scans. (*) Me_4Si spinning sideband.

if (a) is tentatively assigned to the shifted N-terminal threonyl methyl (J. Mazer and M. Teeter, personal communication), (b) is assigned to the five alanyl methyls, (c) to the four internal threonyl methyls, and (d) to the fourteen resonances stemming from the four valyl, eight isoleucyl, and two leucyl methyl groups, we conclude that the high-field spectrum of the denatured protein is consistent with its reported amino acid composition.

The denaturing effect of Me_2SO is not surprising. Thus, e.g., concentrations higher than 60% Me_2SO–^2H_2O cause denaturation of *Clostridium pasteurianum* ferredoxin, as detected by ^{1}H NMR (McDonald et al., 1973), while the redox and spectroscopic properties of the *Chromatium* high-potential iron–sulfur protein are drastically affected at 80% Me_2SO–H_2O (Cammack, 1973). Even the very stable basic pancreatic trypsin inhibitor fully denatures in Me_2SO (Wüthrich, 1976). Still, although the methyl and aromatic data for crambin in this solvent suggest a less than complete unfolding (see below), a reasonable agreement is found for the methyl resonance

Table I: ^{1}H Chemical Shifts for Denatured Crambin: Comparison with Other Random Coil Polypeptides[a]

	trifluoroacetic acid-*d*		dimethyl-d_6 sulfoxide	
	Bak et al. (1968)	crambin	Bundi et al. (1975)	crambin
methyls				
γ-Val, γ,δ-Ile, δ-Leu	1.02	0.95–1.15	0.84–0.98	0.74–0.90
γ-Thr	1.45	1.38–1.49	1.07	0.96–1.08
β-Ala	1.63	1.54–1.65	1.22	1.16–1.30
γ-Thr (N terminus)		1.68		
aromatics				
Tyr	6.98	6.94	6.64	6.64–6.61
	7.19	7.18	7.02	7.03–7.01
Phe	7.22	7.22	7.26	7.21
	7.33	7.33		

[a] The resonance positions are given in parts per million from internal Me_4Si. The crambin data correspond to the spectra shown in Figures 2 and 5 and are not meant to indicate resonance assignments but overall consistency with random coil chemical shifts.

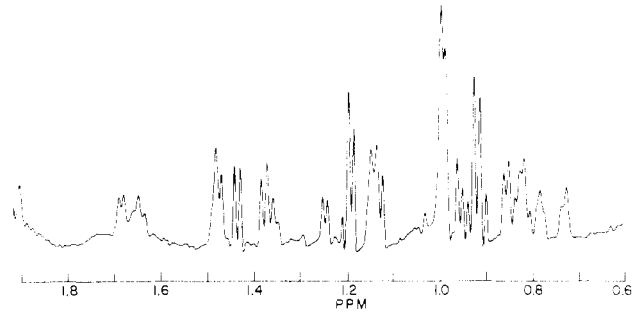

FIGURE 3: ^{1}H NMR spectrum of crambin at 600 MHz, methyl region, after resolution enhancement. Solvent acetic-d_3 acid-*d*; $T \approx 22$ °C; digital resolution 0.29 Hz; 200 scans.

positions with methyl chemical shifts reported for short, flexible peptides in Me_2SO (Table I). It should be remarked that even in F_3AcOH the denaturation is not spontaneous, requiring a few hours for the spectrum to achieve complete random coil characteristics. The acetic acid spectrum (Figure 2D) contains a number of relatively sharp peaks that span a frequency range analogous to that exhibited by other globular proteins such as basic pancreatic trypsin inhibitor (De Marco et al., 1977). The methyl resonances in acetic acid are shown more dispersed in Figure 3, recorded at 600 MHz, 22 °C. Due to solvent viscosity, this temperature does not take full advantage of motional narrowing. However, even under such conditions, at least 18 individual methyl signals can be recognized. The variety of line breadths exhibited by the methyl resonances is indicative of different regimes of rotational motion due to different degrees of packing and steric hindrance.

Figure 4 introduces the low-field region of the 600-MHz spectrum, containing aromatic and amide transitions not shown in Figure 1. The spectrum was recorded 2 h after dilution in acetic-d_3 acid-*d* so that although several of the exposed peptidyl NH protons have exchanged for deuterons, most of the amide signals are still present due to slow exchange. Typically, the amides spread from ~6.6 to ~9.4 ppm while the aromatics, two tyrosyl and one phenylalanyl side chain, appear between ~6.7 and ~ 7.4 ppm. The solvent effects discussed for the methyl resonances are also apparent in the low-field region. As illustrated in Figure 5A, the spectrum in F_3AcOD exhibits no NH signal, due to fast exchange of the amide protons with

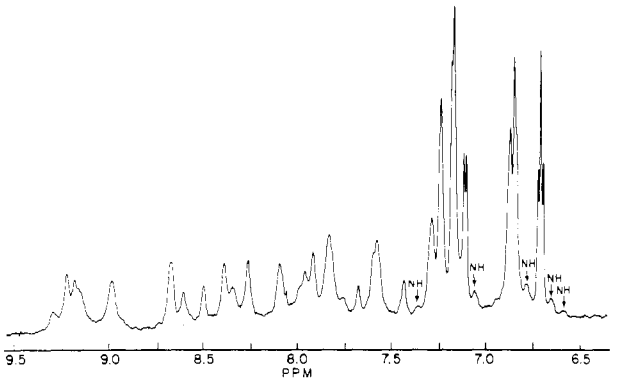

FIGURE 4: ^{1}H NMR spectrum of crambin at 600 MHz, amide and aromatic region. Solvent acetic-d_3 acid-d; $T \approx 22$ °C; digital resolution 0.29 Hz; 700 scans. The spectrum was recorded starting 2 h after dissolving the sample. Expanded displays of the aromatic and amide NH resonances after resolution enhancement are shown in Figures 6 and 8, respectively.

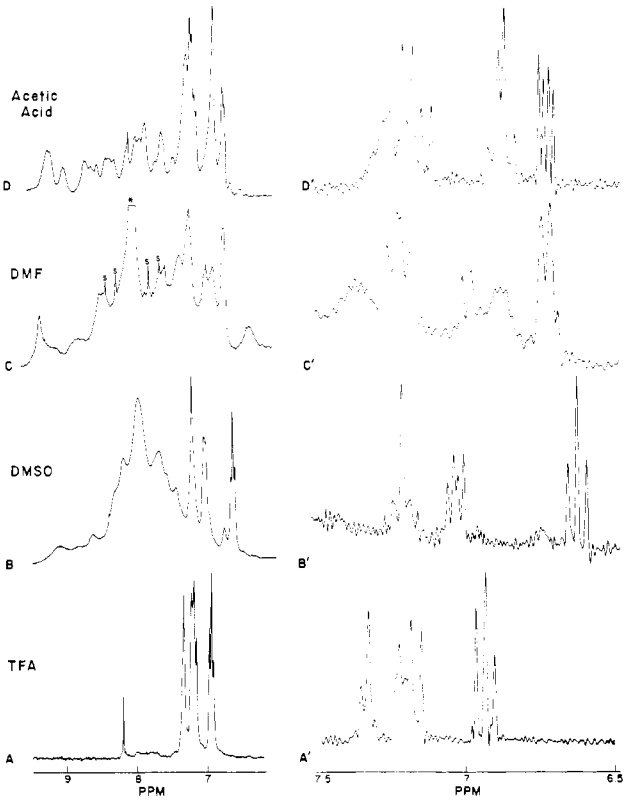

FIGURE 5: ^{1}H NMR spectra of crambin at 250 MHz, amide and aromatic region. In (A) trifluoroacetic acid-d, (B) dimethyl-d_6 sulfoxide, (C) dimethylformamide-d_7, and (D) acetic-d_3 acid-d, accumulated during 0.5 h immediately after dilution. (A′–D′) Expanded aromatic resonances, after ^{1}H–^{2}H exchange, with resolution enhancement; same solvents as for (A)–(D). $T \approx 28$ °C; digital resolution 0.4 Hz; 500–1500 scans. (*) Denotes residual ^{1}H (solvent peak) in DMF.

solvent deuterons. In contrast, as mentioned above, the acetic-d_3 acid-d spectrum (Figure 5D) shows that this solvent does not spontaneously unfold the protein to expose exchange-labile groups. Save for line breadth effects, spectrum C, corresponding to DMF solution, preserves features of the acetic acid spectrum D, while in Me_2SO, spectrum B, most NH resonances are clustered at ~8 ppm. Since in Me_2SO exposed NH's appear at ~9 ppm while typical amide–amide NH···O=C hydrogen bonds show up at ~8 ppm (Llinás & Klein, 1975), the cluster of signals at ~8 ppm in spectrum B is suggestive of extensive hydrogen bonding, as would be the case for an aggregated state. Also noteworthy is that while for a fixed peptide conformation the peptidyl NH chemical shifts in Me_2SO and DMF are almost identical (Llinás & Klein, 1975), such is not the case for crambin spectra C and D, which further supports the view sustaining different conformations for the protein in the two solvents.

Parts A′–D′ of Figure 5 show expanded versions of the aromatic spectra in the four solvents. While the overall similarity of the DMF (C′) and acetic acid (D′) spectra is apparent, the spectra of the F_3AcOD (A′) and Me_2SO (B′) solutions match excellently those expected for exposed, random polypeptides (Table I), again, as observed for the methyl resonances in these solvents. The aromatic spectra can be accounted for in terms of the aromatic residue content of crambin. Thus, the four tyrosyl ring protons each generate a doublet resonance due to $^3J(^1H_{ortho}-^1H_{meta})$ spin–spin couplings, each pair being an independent AX spin system. However, when the tyrosyl ring rotates freely at rates that are high in the NMR time scale, only two doublets are observed, each of area = 2 protons due to fast ortho–ortho and meta–meta rotational averaging. This is exemplified by the F_3AcOD spectrum A′, showing, because there are two tyrosyl rings, pairs of doublets at ~6.94 ppm ($^1H_{meta}$) and at ~7.18 ppm ($^1H_{ortho}$). Due to a chemical shift difference of the order of ^3J, the two upper field doublets partially overlap to appear as a "triplet", the difference in resonance frequency for the two tyrosyl rings being related to their different locations within the primary structure. As indicated in Table I, the agreement with the random coil spectrum in F_3AcOD is also excellent for the phenylalanyl resonances, which appear slightly to lower fields and partially overlapping with the low-field tyrosyl resonances.

In acetic-d_3 acid-d the tyrosyl spectrum is less degenerate (Figure 5D′); an AA′BB′ pattern, corresponding to one of the tyrosyl side chains, shows up at 6.85 ppm. Surprisingly, the other tyrosyl ring resolves into two pairs of doublets, one at ~6.71 ppm and the other at ~7.14 ppm. Due to chemical shift differences of the order of $^3J/2$, the high-field pair appears as a "quartet" while the two low-field doublets are well resolved at 7.11 and 7.17 ppm. Since the area of the 6.71-ppm "quartet" yields approximately half the area of the Tyr_I 6.85-ppm AA′BB′ spin system, assumed to have area = 4 protons, it might be concluded that the more spread tyrosyl signals arise from a single residue, Tyr_{II}, in stoichiometric 1:1 relation with the Tyr_I 6.85-ppm spin system. This pattern is rather surprising and leads to the conclusion that *either* the Tyr_{II} ring is structurally locked within the molecular framework, resulting in different magnetic environments for each of the four ring protons, *or* we are dealing with two slightly different, but equally populated, protein conformations in slow dynamic interconversion (Campbell et al., 1975; Snyder et al., 1975; Wagner et al., 1976). The 1:1:1:1 ratio of Tyr_{II} doublets would favor the first interpretation. The only phenylalanyl ring yields broad resonances at ~7.24 ppm, partly overlapping with the low-field Tyr_{II} signals centered at ~7.14 ppm.

A clearer view of the aromatic region in acetic-d_3 acid-d is illustrated in Figure 6, recorded at 600 MHz and shown with resolution enhancement. Now the high-field Tyr_{II} pair of doublets, a "quartet" at 250 MHz, becomes a triplet because of enhanced dispersion which makes the chemical shift difference, in hertz, of the order of ^3J. Similarly, the low-field pair of doublets appears better resolved from the phenylalanyl signals. The width of the phenylalanyl resonances, as well as those of the Tyr_I ring at 6.85 ppm, suggests that these two side chains are relatively immobile. The same is apparent for the

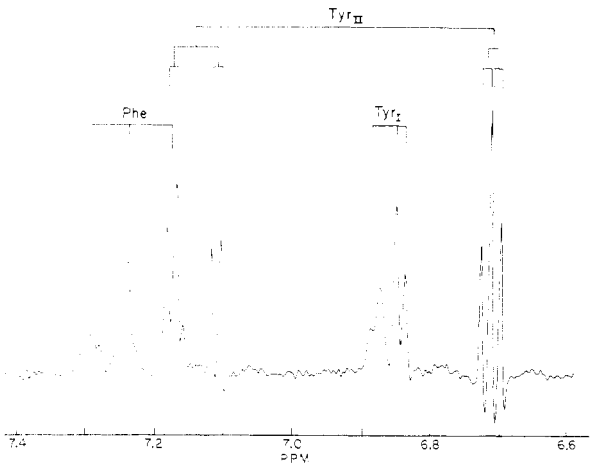

FIGURE 6: ^1H NMR spectrum of crambin at 600 MHz, aromatic region. The spectrum is an expanded version of the 6.58–7.40-ppm region in Figure 4, after resolution enhancement. The diagram identifies the individual spin systems.

Tyr_{II} low-field doublets at ~7.14 ppm, somewhat broader than their high-field counterpart at ~6.71 ppm, suggesting that two of the Tyr_{II} ring protons dipole interact with other protein protons in a stronger fashion than the other two do, in agreement with the hypothesis that despite its overall "free tyrosine" spectral appearance, the Tyr_{II} side chain is obstructed from rotation.

Prior to recording spectra B′ and C′ (Figure 5), a trace of ^2H_2O was added to the samples which were then heated to 85 (Me_2SO) or 50 °C (DMF) for 5 and 10 min, respectively, and then left for several hours at room temperature (Me_2SO) or at 40 °C (DMF). This was necessary in order to preexchange labile protons and eliminate the background-obscuring line shapes of the aromatic signals. However, even under such drastic conditions the exchange was incomplete, probably because of the aggregation responsible for the observed broad line spectra and also because of preservation of globular structure in DMF. *Surprisingly, in acetic-d_3 acid-d at 50 °C, the exchange was incomplete after several hours.*

As discussed above, the narrow line features of the acetic acid spectrum point clearly toward a nonaggregated state for crambin in this solvent. Hence, the slow exchange of labile protons indicates that a large number of them are protected from direct interaction with the solvent by the secondary and tertiary structures and that the conformational fluctuations responsible for exposure and ^1H–^2H exchange are minimal (Hvidt & Nielsen, 1966). This is striking as one would predict that the Brønsted acidity of acetic acid would favor structural breathing.

The time course of isotope exchange for crambin dissolved in acetic-d_3 acid-d is illustrated in Figure 7, which shows that after 20 days the exchange was still incomplete. The spectrum labeled zero time was recorded for the protein dissolved in *protonated* acetic acid and hence contains signals from all labile protons this chemical shift range. Thus, arginyl guanido groups yield a characteristic broad background centered at ~6.6 ppm, underneath the aromatic resonances (Wüthrich & Wagner, 1979). The sequence of spectra shown in Figure 7 indicates both the immediate disappearance of these broad signals after dilution in the deuterated solvent and the presence of a large number of NH resonances remaining in the protonated form. Interestingly, the amide resonances that disappear first are those that in the zero-time spectrum are clustered in the 7.5–8.0-ppm range. From previous studies in acidic Brønsted solvents, it is known that this chemical shift

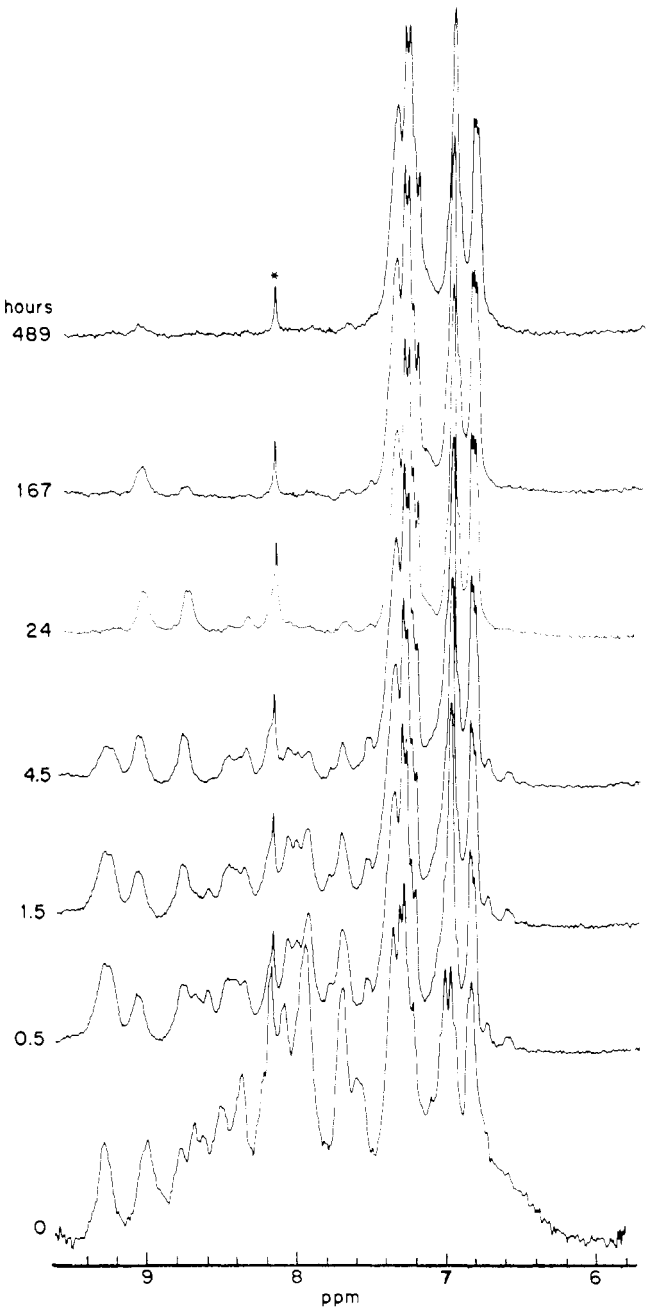

FIGURE 7: ^1H NMR spectra of crambin at 250 MHz. Time course for ^1H–^2H exchange of amide and guanido protons. Each spectrum was accumulated for 25 min starting at the indicated times after dissolving the sample in acetic-d_3 acid-d. The zero-time spectrum corresponds to a sample dissolved in protonated acetic acid. (*) Denotes an impurity. $T \approx 28$ °C; digital resolution 0.37 Hz; 1000 scans.

range is characteristic of exposed CO–NH peptidyl groups (Llinás & Klein, 1975). Not surprisingly, the most slowly exchanging amides are those near 9 ppm, the position typical of intramolecular NH···O=C hydrogen bonds (Llinás & Klein, 1975). Although no other signals were detected at fields lower than ~9.3 ppm, the strong solvent –COOH peak at ≳11.2 ppm precludes observation of signals in this region where unusually strong hydrogen-bonded amides are known to resonate (Llinás et al., 1970, 1972; Llinás & Klein, 1975). Most noteworthy are the two transitions at ca. 6.2 and 6.4 ppm, shown in the 0.5-, 1.5-, and 4.5-h spectra. A number of these high-field shifted NH's are also indicated in Figure 4. Such signals indicate intramolecular occlusion of peptide backbone groups within a hydrophobic matrix (Llinás & Klein, 1975).

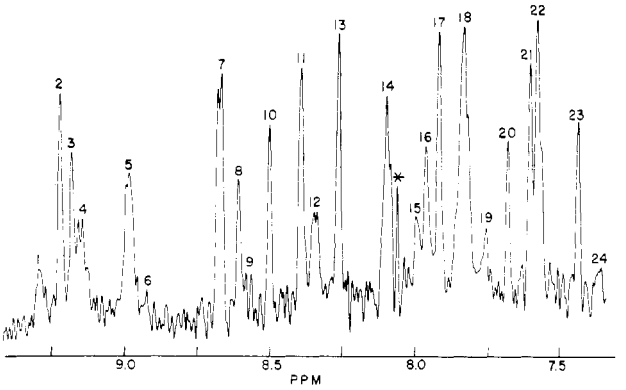

FIGURE 8: ^1H NMR spectrum of crambin at 600 MHz, amide resonances. The spectrum represents an expanded display of the slowly exchanging amides shown in Figure 4, after resolution enhancement. (*) Denotes an impurity.

A resolution-enhanced 600-MHz spectrum of the 7.3–9.3-ppm region, averaged during 1.25 h, is shown in Figure 8. Because the spectrum was accumulated starting 2 h *after* dissolving in acetic-d_3 acid-*d*, it only depicts signals arising from slowly exchanging amides that are truly significant. As indicated, ~24 NH peaks can be counted to lower fields from the aromatic resonances. These, combined with the four transitions visible in the aromatic region (Figure 4), and since each peak represents *at least* one amide proton, indicate that out of a total of 44 residues, a minimum of 28 peptidyl NH groups are protected from direct solvent exposure. Such a situation is unknown for proteins in acidic solvents and represents a conformational stability only comparable to that exhibited by the ferrichrome cyclohexapeptides that in ^2H_2O take days for some amides to half-exchange (Emery, 1967; Llinás et al., 1973) and that even in F_3AcOH preserve their tertiary structure (Llinás & Klein, 1975). The fact that crambin slowly unfolds in F_3AcOH can be explained from the knowledge that besides seryl and threonyl residues, cysteinyl residues also react with this solvent (Wüthrich, 1976) so that the denaturation is likely to represent not a dynamic conformational drift but rather an irreversible chemical modification. Indeed, besides overwhelming hydrophobic interactions, much of the crambin conformational rigidity should be ascribed to the presence of three disulfide bridges as well as to its content of four prolyl residues.

The intrinsic high resolution exhibited by its crystalline X-ray diffraction pattern (Teeter & Hendrickson, 1979) combined with its unprecedented structural stability in solution makes crambin a unique system to investigate in detail the interactions responsible for biological hydrophobic behavior.

Acknowledgments

We are thankful to Dr. Wayne A. Hendrickson for bringing crambin to our attention and to Dr. Martha M. Teeter for valuable preliminary information on the crystallographic and sequence investigation. Dr. C. H. Van Etten provided us with a generous supply of the crystalline protein and *C. abyssinica* seeds.

References

Bak, B., Dambmann, C., Nicolaisen, F., Pedersen, E. J., & Bhaca, N. S. (1968) *J. Mol. Spectrosc. 26*, 78–97.

Bundi, A., Grathwohl, Ch., Hochmann, J., Keller, R. M., Wagner, G., & Wüthrich, K. (1975) *J. Magn. Reson. 18*, 191–198.

Cammack, R. (1973) *Biochem. Biophys. Res. Commun. 54*, 548–554.

Campbell, I. D., Dobson, C. M., & Williams, R. J. P. (1975) *Proc. R. Soc. London, Ser. B 189*, 503–509.

Dadok, J., & Sprecher, R. F. (1974) *J. Magn. Reson. 13*, 243–248.

De Marco, A., Tschesche, H., Wagner, G., & Wüthrich, K. (1977) *Biophys. Struct. Mech. 3*, 303–315.

Dwek, R. A. (1973) *Nuclear Magnetic Resonance (N.M.R.) in Biochemistry*, Clarendon Press, Oxford.

Emery, T. F. (1967) *Biochemistry 6*, 3858–3866.

Ernst, R. R. (1966) *Adv. Magn. Reson. 2*, 1–135.

Gupta, R. K., Ferretti, J. A., & Becker, E. D. (1974) *J. Magn. Reson. 13*, 275–290.

Hvidt, A., & Nielsen, S. O. (1966) *Adv. Protein Chem. 21*, 287–386.

James, T. L. (1975) *Nuclear Magnetic Resonance in Biochemistry*, Academic Press, New York.

Llinás, M., & Klein, M. P. (1975) *J. Am. Chem. Soc. 97*, 4731–4737.

Llinás, M., Klein, M. P., & Neilands, J. B. (1970) *J. Mol. Biol. 52*, 399–414.

Llinás, M., Klein, M. P., & Neilands, J. B. (1972) *J. Mol. Biol. 68*, 265–284.

Llinás, M., Klein, M. P., & Neilands, J. B. (1973) *J. Biol. Chem. 248*, 924–931.

McDonald, C. C., & Phillips, W. D. (1967) in *Magnetic Resonance in Biological Systems* (Ehrenberg, A., Malmstroḿ, B. G., & VänngÅrd, T., Eds.) pp 3–24, Pergamon Press, London.

McDonald, C. C., & Phillips, W. D. (1969) *J. Am. Chem. Soc. 91*, 1513–1521.

McDonald, C. C., Phillips, W. D., Lovemberg, W., & Holm, R. H. (1973) *Ann. N.Y. Acad. Sci. 222*, 789–799.

Snyder, G. H., Rowan, R., Karplus, S., & Sykes, B. D. (1975) *Biochemistry 14*, 3765–3777.

Teeter, M. M., & Hendrickson, W. A. (1979) *J. Mol. Biol. 127*, 219–223.

Van Etten, C. H., Nielsen, H. C., & Peters, J. E. (1965) *Phytochemistry 4*, 467–473.

Wagner, G., De Marco, A., & Wüthrich, K. (1976) *Biophys. Struct. Mech. 2*, 139–158.

Wüthrich, K. (1976) *NMR in Biological Research: Peptides and Proteins*, North-Holland Publishing Co., Amsterdam.

Wüthrich, K., & Wagner, G. (1979) *J. Mol. Biol. 130*, 1–18.