Table III. Protein Hydration[a]

Protein	State[b]	Model[c]	Calcd[d]	Obsd[e]	Ref
Gelatin	N	FE	0.50	0.45	*l*
Myoglobin	N	FE	0.45	0.415	*f*
Bovine albumin	N	FE	0.445	0.40	*f, l, m*
Bovine albumin	D, urea	FE	0.445	0.44	*m*
Bovine albumin	D, pH 3	FE, T	0.32	0.30	*l*
Hemoglobin	N	FE	0.415	0.42	*l*
Chymotrypsinogen	N	FE	0.39	0.34	*f*
Lysozyme	N	FE	0.36	0.34	*f, l*
Lysozyme	N	FE-B	0.335	0.34	*f, l*
Ovalbumin	N	FE	0.37	0.33	*f*
TMV coat	Intact virus	FE[g]	0.365	0.25	*h*
TMV coat	Stripped virus, pH 10	FE[g]	0.365	0.36[i]	*h*
Chymotrypsin	N	FE	0.36	0.33	*f*
Protamine sulfate	N	FE[j]	0.28	0.28	*f*
1-Lysine, 1-glutamate copolymer	pH 4	FE	3.0[k]	2.4[k]	
	pH 11	FE	6.3[k]	7.5[k]	

[a] As grams of water per gram of protein. [b] N, native protein, 0.01 *M* KCl; D, denatured protein, as shown. [c] FE, all residues fully exposed; FE, T, residues fully exposed but all carboxylate ions assumed titrated; FE-B, corrected for buried residues. [d] Residue hydration taken from Table II. Amino acid analyses from ref 10 and 11. [e] Results from ref 1 have been revised slightly to yield bovine albumin hydration of 0.40, in agreement with later work. [f] This work, proteins run as described in ref 1, and 2. [g] RNA is assumed to be unhydrated. [h] W. Bostian and I. D. Kuntz, unpublished results. [i] Corrected for RNA hydration assumed to be 0.07 g of water per gram of intact virus. [j] Corrected for weight of sulfate as counterion. [k] As moles of water per mole of amino acid. [l] See ref 1. [m] See ref 2.

compositions are available,[10,11] few explicit assignments of "protected" and "exposed" residues have been made.[12] Thus, we have assumed that each amino acid is exposed to the solvent. This should lead to an overestimate of the hydration. The magnitude of this overestimate can be roughly evaluated using lysozyme where the protected residues have been identified.[12] Alternatively, one could measure the hydration of completely denatured proteins which should approximate the "fully" exposed state.[13] We have ignored all contributions from other "structural" features, such as grooves, channels, helical regions, etc.

Calculated and experimental (nmr) results are given in Table III. Remember that the absolute hydration of both the polypeptides and proteins has been based on the assumption that bovine albumin has a hydration of 0.40 ± 0.04 g of water/g of protein. The error in the protein measurements is on the order of 10%. The calculated values appear consistently high although within the experimental error. The correction for buried residues is about 10% in the case of lysozyme. This correction should increase as the size of the (globular) protein increases. It is encouraging that both the urea and acid-denatured samples of serum albumin agree with the simple predictions: *i.e.*, the urea exposes the hydrophobic groups which pick up only a small amount of water; acid, by titrating the carboxyl groups, dehydrates the protein. The only striking "structural" effect in Table III is a negative one. Tobacco mosaic virus is much less hydrated than the calculations indicate. When the coat protein is dissociated, the hydration increases to the expected value. The comparison of hemoglobin and myoglobin suggests that aggregation of subunits, *per se*, does not generate a major change in hydration (see also ref 14). We rationalize this finding using the "hydrophobic binding" model for subunit association. If only poorly hydrated hydrophobic regions are involved, one expects small (<10%) decreases in water binding. Conversely, the protein–RNA interface in TMV must contain a large percentage of ionic groups to cause such a precipitous decrease in hydration.

In conclusion, the nmr determination of hydration yields an internally consistent method of estimating the hydration of globular proteins in terms of the hydration of the appropriate polypeptides. The major part of the water is associated with the ionic side chains, although 25–30% is "bound" by the nonionic groups. We predict that the hydration of proteins should be quite sensitive to acid titration but should be insensitive to pH above pH 6 (assuming no major structural changes). Fully exposing the hydrophobic core of globular proteins should produce a small increase in hydration. Aggregation which involves hydrophobic sites should slightly decrease hydration, while aggregation which protects ionic sites should cause large decreases.

Further work is in progress on several of these topics.

I. D. Kuntz
Department of Chemistry, Princeton University
Princeton, New Jersey 08540
Received October 19, 1970

(10) M. O. Dayhoff, "Atlas of Protein Sequence and Structure," National Biomedical Research Foundation, Silver Spring, Md., 1969.
(11) G. R. Tristam and R. H. Smith, *Advan. Protein Chem.*, **18**, 227 (1963).
(12) I. Klotz, *Arch. Biochem. Biophys.*, **138**, 704 (1970), and references therein.
(13) C. Tanford, *Advan. Protein Chem.*, **23**, 122 (1968).
(14) R. Jaenicke and M. A. Lauffer, *Biochemistry*, **8**, 3083 (1969).

Hydration of Macromolecules. IV. Polypeptide Conformation in Frozen Solutions

Sir:

In a companion paper we have discussed the use of an nmr method of determining the hydration of water-soluble polypeptides.[1] In this communication we show that the nmr technique is sensitive to the polypeptide conformation, even though the hydration, *per se*, appears independent of the structure of the polypeptide.[1]

(1) I. D. Kuntz, *J. Amer. Chem. Soc.*, **93**, 514 (1971).

Our basic experiment observes the high-resolution proton nmr spectra of frozen, aqueous solutions of macromolecules. These spectra ordinarily consist of a single symmetrical signal which can be shown to come from water protons still capable of motion at −20 to −60°.[1,2] The area under the curve is proportional to the amount of water affected by the macromolecules. In addition we measure the line width (full width at half-height) of the water resonance. The line width is related to the spin–spin relaxation time of the water protons[3] and might be expected to be sensitive to the type and structure of the macromolecular components.[4,5] Line widths were read directly from 100-MHz spectra. External 5- and 10-kHz oscillators were used to reduce side-band interference. Standard variable temperature accessories were used, along with a thermistor, mounted in the sample tube, to monitor the temperature.[1,2,5] The sources, concentrations, and molecular weights of the polypeptides have been given.[1]

We find that the line width of the water signal is sensitive to the particular polypeptide chosen, the temperature of observation, and, for the ionic polymers, to the pH of the room temperature solution. The concentrations and molecular weights of the polypeptides are not important variables. Table I summarizes these results. Turning first to the temperature effects, we found that the line widths generally followed simple Arrhenius behavior over the accessible temperature range (−15 to −50°). Only polylysine (L and DL) and poly-L-ornithine near neutral pH (7–8) exhibited more complicated effects. Apparent activation enthalpies were calculated in the usual way. We caution that such enthalpies are plagued with large experimental uncertainties. Further, detailed interpretation requires an exact model of the various microscopic reactions contributing to the relaxation of the water protons. At present we report these numbers primarily as a convenient summary of the experimentally observed temperature dependences. Activation entropies cannot be determined without additional assumptions[5] and will not be considered here.

If one examines the line width and activation enthalpy data with the view of characterizing side chain or group line widths, the first impression is of the remarkable diversity of the results. The line widths span, at constant temperature, nearly two orders of magnitude, while the activation enthalpies differ by more than a factor of five. In fact, no two polypeptides studied show identical results, even the closely related poly-L-glutamate–poly-L-aspartate or poly-L-lysine–poly-L-ornithine pairs. This can be contrasted with the results for native, globular proteins which show less than a twofold range in both line widths and activation enthalpies.[2,5]

Two other negative statements should be made. First, in spite of a general grouping of narrow water signals with charged polypeptides, there is no general correlation of line width with ionic side chains. The exceptions which force us to discard this generalization are poly-L-arginine and the copolymer poly-L-lysine–poly-L-glutamic acid (4:6) at low pH. Furthermore, poly-L-valine also gives a narrow water line although it is nonionic. Second, the line width appears independent of net hydration.[1] This is particularly apparent for the basic polypeptides whose line width changes on going from low to high pH although the hydration remains constant.[1]

The above remarks make it clear that a general understanding of the results must await a detailed picture of the types of proton motion in the vicinity of macromolecular surfaces. Nonetheless an interesting pattern is present. We find, without exception, that all systems which, at *room temperature*, are thought to be random coils, exhibit sharper water lines than the corresponding system in the helical conformation. The changes in line width occur, in each case, over a pH range which is close to that expected from conventional solution studies (Table II). Thus it appears that the rapid freezing procedure we use is capable of trapping the solution conformations of the polypeptides.[6] Proceeding on the assumption that the solution conformations apply to the corresponding frozen solutions, we conclude that the width of the water resonance is dependent, at least in part, on the shape of the macromolecule. This conclusion is supported by the data for

Table I. Water Line Width and Activation Enthalpies (−35°)

Polypeptide[a]	pH[b]	Line width,[c] Hz	Activation[d] enthalpy, kcal/mol
	Basic Groups		
Arg	3	(2,500)	15
	10	(7,600)[e]	14[e]
Lys	3–6	750	9.4
	10.5–12	1,900	14
DL-Lys	3	370	7.1
	11	(2,900)	16
Orn	1.5–6	800	9.4
	11	(3,200)	15
	Acidic Groups		
Asp	4.5	1,500[e]	7.7[e]
	8.4	(150)	18
Glu	4.5	1,500[e]	8.2[e]
	7–12	(100)	13
Tyr	11.3	360[e]	7.0[e]
	11.5–12.2	200	3.5
	Nonpolar Groups		
DL-Ala		1,650	4.5
Gly		800	2.9
Pro		1,250	10.5
Val		(400)	
	Copolymers		
$Lys^{40}Glu^{60}$	2–4	2,200	8.1
	12	220	13
$Lys^{50}Phe^{50}$	2–6	800	7.6
	9	1,450[e]	3.5[e]

[a] Standard abbreviations, arranged alphabetically within groups. Concentrations, molecular weights, and suppliers are given in ref 1. "L" isomer unless otherwise noted. [b] Measured before and after freezing, no buffer added, uncertainty ±0.3 pH unit. [c] Reproducibility ±10% for line widths less than 2 kHz, ±20% from 2 to 3 kHz. Brackets indicate extrapolated values. [d] Uncertainties of order ±2 kcal/mol for values less than 8; of order ±4 kcal/mol for higher values. [e] Titration probably not complete; some mixture of conformations is likely to exist.

(2) I. D. Kuntz, T. S. Brassfield, G. Law, and G. Purcell, *Science*, **163**, 1329 (1969).

(3) W. S. Brey, T. E. Evans, and L. H. Hitzrot, *J. Colloid Sci.*, **26**, 306 (1968); M. E. Fuller and W. S. Brey, *J. Biol. Chem.*, **243**, 274 (1968).

(4) J. Glasel, *J. Amer. Chem. Soc.*, **92**, 372, 375 (1970).

(5) I. D. Kuntz and T. S. Brassfield, *Arch. Biochem. Biophys.*, in press.

(6) Alternatively, if we are monitoring an equilibrium process, the combined changes in pH and helix–coil enthalpies must compensate in such a way as to be independent of temperature over a 70° interval.

Table II. Helix–Coil Transitions

Polypeptide	Line width: Coil	Line width: Helix	Approximate pH of transition nmr (frozen)	pH midpoint solution studies
Asp	(150)	≥1,500[a]	6.0	
Glu	(100)	≥1,500[a]	5.3	5.4,[b] 5.6[c]
Tyr	200	≥360[a]	11.3	11.5[d]
Arg	(2,500)	(7,600)	≥8.5[a]	
Lys	750	1,900[a]	9.5	10[b]
DL-Lys	370	(2,900)	9.5	
Orn	800	(3,200)	≥8.5[a]	
LysGlu (4:6)	220	2,200	7 (broad)	5.5[e]
LysPhe (1:1)	800	≥1,450[a]	≥8.5[a]	9[f]

[a] Transition may not be complete. [b] Reference 8, p 550. [c] G. Barone, V. Crescenzi, and F. Quadrifoglio, *Biopolymers*, **4**, 529 (1966). [d] Reference 8, p 553. [e] Reference 7. [f] E. Peggion, A. S. Verdini, A. Cosani, and E. Scoffone, *Macromolecules*, **3**, 194 (1970).

copolymers. Both the lysine–phenylalanine and lysine–glutamic acid systems show line widths in agreement with the appropriate structure rather than composition. Thus, the latter polymer is thought to be helical at low pH (as polyglutamic acid) and random coil at high pH (as polyglutamic acid, but unlike polylysine).[7] The line widths at the limiting pH values parallel the glutamic acid results. In a similar way, the lysine–phenylalanine line widths and temperature dependences follow those of polylysine. This is really quite a different result from the hydration of the copolymers. The hydrations were essentially weighted averages of the appropriate homopolypeptide values.[1]

If we apply this rule (that a change in line width indicates a change in conformation) our results require us to assign a conformation transition to poly-DL-lysine at high pH. This is presumably a normal coil–helix transformation which can occur in such polymers because they are not truly optically random. That is, short runs of pure "L" or pure "D" isomer occur and are of sufficient length to assume helical structures under appropriate conditions.[8]

We can make a tentative effort to extend these ideas to the nonionic polymers. It is not easy to demonstrate conformational changes within these systems because their cooperative structures are independent of salt, pH, and temperature perturbations generally accessible in aqueous solutions. One useful point of reference, however, is the polyvaline polymer, which is considered to be sterically limited to the random coil form in solution.[9] As noted above, polyvaline shows a very narrow water line, with distinct side-chain resonances detectable as well. By contrast, poly-DL-alanine, polyproline, and polyglycine show considerable broadening of the water signal and no side-chain resonances below about −10°. We are thus tempted to suggest that these last three polymers assume some nonrandom structures in solution [presumably α helix, polyproline II, and β structure (?), respectively[8]] which serve to restrict the freedom of motion of the water protons in their immediate vicinity.

In sum, we feel there is considerable direct and indirect evidence to suggest that macromolecular conformation can influence the mobility and hence the nmr line width of "bound" water. There are important qualifications. (1) We cannot use the sign and/or magnitude of the line width change either to identify the conformation change involved or to determine the extent of the change from nmr data alone. In fact, we observe that the Arrhenius plots all show points of intersection at finite temperatures suggesting that line-broadening effects in one temperature region will become line-narrowing effects in some other temperature range. (2) There are no simple rules to predict line widths from compositions because we cannot predict conformation knowing only composition. Nor is it known how various structural components will interact with each other. Thus it is not possible to calculate, for example, the line widths of the water associated with globular proteins simply from their amino acid composition.

We hope that further experimentation will remove some of these restrictions, permitting fuller use of the large sensitivity of the nmr line widths for conformational studies.

I. D. Kuntz

Department of Chemistry, Princeton University
Princeton, New Jersey 08540

Received October 19, 1970

(7) E. R. Blout and M. Idelson, *J. Amer. Chem. Soc.*, 80, 4909 (1958).
(8) G. D. Fasman in "Poly-α-Amino Acids," G. D. Fasman, Ed., Marcel Dekker, New York, N.Y., 1967, Chapter 11.
(9) Reference 8, but see R. F. Epand and H. A. Scheraga, *Biopolymers*, 6, 1551 (1968).

Biosynthesis of Pyridoxine[1]

Sir:

Recently we presented evidence concerning the primary precursors of pyridoxine (= pyridoxol), one of the forms of vitamin B_6.[2] We demonstrated by Kuhn–Roth degradation of labeled pyridoxol obtained from cultures of the *Escherichia coli* B mutant WG2 (B_6-2),[3] which had been incubated with various radiomers of pyruvic acid, that the two-carbon fragment, C-2′,2, of pyridoxol is derived specifically from a C_2 unit, corresponding to the methyl and the carbonyl carbon atoms of pyruvate. The carboxyl carbon of pyruvate was not incorporated into the product. We suggested that pyridoxol was derived from three glycerol units, one of which was incorporated *via* pyruvate as a two-carbon fragment at the oxidation level of acetaldehyde, while the other two entered pyridoxol intact, possibly *via* triose phosphate (Scheme I). We now offer further evidence in support of this view.

(1) This investigation was supported by a grant from the National Research Council of Canada.
(2) R. E. Hill and I. D. Spenser, *Science*, **169**, 773 (1970).
(3) W. B. Dempsey and P. F. Pachler, *J. Bacteriol.*, **91**, 642 (1966); W. B. Dempsey, *ibid.*, **97**, 1403 (1969).