Helix Lifetimes within the Conformational Transition Region. A Random Walk Model[1]

R. L. Jernigan, J. A. Ferretti,* and G. H. Weiss

Physical Sciences Laboratory, Division of Computer Research and Technology, National Institutes of Health, Bethesda, Maryland 20014. *Received March 26, 1973*

ABSTRACT: The kinetics of the helix-random-coil intramolecular transition is treated in terms of two times: (1) Δt_2, the time for formation of a first helix unit from an all-coil molecule and (2) Δt_1, the time to add another helix unit to an existing helix region. Interpretation of experimental results implies that Δt_1 is significantly shorter than Δt_2. Thus, it is expected that the totally random-coil conformer is longer lived than any other form. Expressions are presented for the lifetime of any chain unit as helix with given initial positions of the helix-coil boundaries. Numerical results are given for chains with 150 and fewer units by averaging over all initial positions with a one helix sequence partition function using Zimm-Bragg polypeptide statistical weights. At the transition midpoint no units exist as helix for longer than 5×10^3 Δt_1 in a chain of 150 peptides. If Δt_2 is significantly larger than about 10^4 Δt_1, then the configurations may be grouped into two classes: (1) the totally random-coil form and (2) all-helix-containing conformations. In an experiment at an appropriate frequency, only these two groups would be distinguishable.

Nuclear magnetic resonance[2a] and relaxation[2b] measurements of synthetic polypeptides can be interpreted to indicate two distinct regions of relaxation times. Rapid reaction techniques have yielded a short time, which has been widely associated with the time, Δt_1, for adding to or melting one peptide unit from an existing helix sequence. A much longer time, Δt_2, is inferred from equilibrium nmr measurements. From among the variety of interpretations[2a] of this nmr time, we choose here to identify it as the time for initial formation of a helical segment (see Figure 1). The intramolecular helix-random-coil transition can be completely described in terms of the two times Δt_1 and Δt_2. If $1/\Delta t_1$ is much larger than the frequency of the experimental probe and if $1/\Delta t_2$ is similar in magnitude to the frequency of the experimental probe, then configurations which exchange at a rate of $1/\Delta t_1$ will be indistinguishable. Thus all helix-containing conformers are expected to exchange rapidly among themselves. The possibility remained that some internal residues, because of their long distance from chain termini, might evidence long lifetimes. The calculations here indicate this to be unlikely. We calculate average lifetimes for each chain unit to remain in the α helix conformation. The results are used in the following paper[3] to aid in the interpretation of the nmr spectra.

Random Walk Model of the Helix-Coil Transition

Conformational transitions in linear polymers, in which each unit can exist in several states, can be modeled in terms of one-dimensional random walks. This idea was suggested by Flory[4] and has been applied recently by Miller[5] to treat the present problem. The helix-random-coil transition in polypeptides is particularly suitable for such treatment. Under equilibrium melting conditions, the polypeptide will consist of only a few long sections of helix or sections of random-coil units. In terms of a random walk model, the locations of the walkers represent the positions of junctions between helical and coil sequences. Thus, treatment of a long-chain polypeptide with r distinct internal helical regions would require $2r$ random walkers. Solution of this problem for $r > 1$ is difficult, even if r were constant throughout the process. Under conditions where multiple helical segments are permitted, the number of helical regions in a very long melting chain would change during the transition. Throughout much of the normal experimental range, it is sufficient to consider only one helical region restricted to the interior of the chain.[6] The present model takes account of only one helical segment.

The mean lifetimes of each unit as helix are related to the equilibrium nmr spectra. Individual conformers or groups of conformers will appear in separate nmr peaks only if rates for transitions to other conformations are slow (*ca.* $<10^2$ sec^{-1}). Here, we have calculated mean lifetimes, as helix, for each unit in the chain.

Our treatment for two random walkers proceeds from probabilities derived for one random walker. This walker is restricted to a finite range between a reflecting barrier at 0 and an absorbing barrier at i. Absorption corresponds to the transition for the ith peptide unit from helix to coil. The reflection point corresponds to the end of the polymer.

In order to treat a chain of n monomers with one internal helix region, it is necessary to combine the results for two such random walks with respective lengths, i and $n - i + 1$. If $p_j(t)$ denotes the probability of the walker being at position j at time t, and forward and backward rates are given by k_f and k_b, then the following set of differential equations describes the random walk

$$\begin{aligned} dp_0/dt &= -k_f p_0 + k_b p_1 \\ dp_j/dt &= -(k_f + k_b)p_j + k_b p_{j+1} + k_f p_{j-1} \\ &\text{for } 1 < j < i - 1 \\ dp_{i-1}/dt &= k_f p_{i-2} - (k_b + k_f)p_{i-1} \\ dp_i/dt &= k_f p_{i-1} \end{aligned} \tag{1}$$

The first of eq 1 corresponds to complete reflection, and the last corresponds to absorption. Substitution of

$$p_j(t) = \alpha^{-j/2} \exp[-(k_f + k_b)t]\eta_j \tag{2}$$

$$(\text{for } 0 \leq j < i)$$

$$t = \tau/(2\sqrt{k_b k_f})$$

$$\alpha = k_b/k_f$$

(1) Presented at the 164th National Meeting of the American Chemical Society, New York, N. Y., Sept 1972; *Polym. Prepr., Amer. Chem. Soc., Div. Polym. Chem.*, 13, 941 (1972).

(2) (a) J. A. Ferretti, B. W. Ninham, and V. A. Parsegian, *Macromolecules*, 3, 34 (1970); J. A. Ferretti and B. W. Ninham, *ibid.*, 3, 30 (1970). (b) R. Lumry, R. Legare, and W. G. Miller, *Biopolymers*, 2, 434 (1964); G. G. Hammes and P. B. Roberts, *J. Amer. Chem. Soc.*, 91, 1812 (1969); A. Wada, T. Tanaka, and H. Kihara, *Biopolymers*, 11, 587 (1972); K. Saksena, B. Michels, and R. Zana, *J. Chim. Phys.*, 65, 597 (1968).

(3) J. A. Ferretti and R. L. Jernigan, *Macromolecules*, 6, 687 (1973).

(4) P. J. Flory, *J. Polym. Sci.*, 49, 105 (1961).

(5) W. G. Miller, *Polym. Prepr. Amer. Chem. Soc., Div. Polym. Chem.*, 12, 558 (1971); W. G. Miller, *Macromolecules*, 6, 100 (1973).

(6) D. Poland and H. A. Scheraga, "Theory of Helix-Coil Transitions in Biopolymers," Academic Press, New York, N. Y., 1970.

$$\cdots cc\underset{k_b}{\overset{k_f}{\rightleftharpoons}}\cdots ccc\cdots$$

processes identified with the time Δt_1

$$\cdots cccccc\cdots \longrightarrow \cdots cchhcc\cdots$$

process identified with the time Δt_2

Figure 1. Conformational transitions associated with experimental times. Model rates k_f and k_b are also indicated.

into the first i (eq 1) results in the set of equations

$$2d\eta/d\tau = \mathbf{A}\cdot\eta \tag{3}$$

where η is an i element column vector and $\mathbf{A}$ is the square array

$$\mathbf{A} = \begin{bmatrix} \sqrt{\alpha} & 1 & & & & \\ 1 & 0 & 1 & & & \\ & 1 & 0 & 1 & & \\ & & \cdot & \cdot & \cdot & \\ & & & & 0 & 1 \\ & & & & 1 & 0 \end{bmatrix}$$

The eigenvalues of $\mathbf{A}$ are given by $\cos\theta_s$, where the θ_s are solutions to the following equation

$$\tan i\theta_s = \sin\theta_s/(\sqrt{\alpha} - \cos\theta_s) \tag{4a}$$

If $(\alpha)^{1/2} > 1 + 1/i$, then there will be one imaginary solution, which is found by solving

$$\tanh i\theta_0 = \sinh\theta_0/(\sqrt{\alpha} - \cosh\theta_0) \tag{4b}$$

Thus, if $(\alpha)^{1/2} > 1 + 1/i$, then one solution will come from eq 4b and $i - 1$ from eq 4a; otherwise eq 4a will furnish all i solutions. Note that i in these equations is the chain index and not $(-1)^{1/2}$. Numerical solutions to these equations can be obtained straightforwardly. The desired probabilities are found to be

$$p_j = \alpha^{(m-j)/2}\ \exp[-\tau(\sqrt{\alpha} + \sqrt{1/\alpha})/2] \times \sum_s (u_s)_m(u_s)_j \exp(\lambda_s\tau/2) \tag{5}$$

in which λ_s and u_s are the eigenvalues and eigenvectors of the matrix $\mathbf{A}$. Expressed in terms of solutions to eq 4, this result becomes

$$p_j = \alpha^{(m-j)/2}\ \exp[-\tau(\sqrt{\alpha} + \sqrt{1/\alpha})/2] \times \{A_0^{-1}\sinh i\theta_0 \sinh m\theta_0 \exp(\cosh\theta_0\tau) + \sum_{s=1}^{i} A_s^{-1}\sin i\theta_s \sin m\theta_s \exp(\cos\theta_s\tau)\} \tag{5'}$$

If $(\alpha)^{1/2} \leq 1 + 1/i$, then the first term disappears and the sum assumes a lower index of 0. The coefficients are given by

$$A_0 = i/2 - \cosh(i+1)\theta_0 \sinh i\theta_0/2\sinh\theta_0$$

and

$$A_s = i/2 + \tfrac{1}{4} - \sin(2i+1)\theta_s/(4\sin\theta_s)$$

The probability density for absorption, from eq 1 and 5, is given by

$$f_{i,m}(t) = dp_i/dt = k_f p_{i-1} = k_f \alpha^{(m+1-i)/2} \sum_s (u_s)_m(u_s)_{i-1} \exp(\nu_s t) \tag{6}$$

where

$$\nu_s = \sqrt{k_b k_f}(\lambda_s - \sqrt{\alpha} - \sqrt{1/\alpha})$$

The probability of absorption at time t or later is therefore

$$F_{i,m}(t) = k_f \alpha^{(m+1-i)/2} \sum_s (u_s)_m(u_s)_{i-1}(\nu_s^{-1}) \exp(\nu_s t) \tag{7}$$

This result is valid for one walker with initial position m on a walk of length i. If the probability of absorption for the walker on the other side of the helical region is given by $G_{n-i+1,n-m'}(t)$ with an initial position m' on a walk of length $n - i + 1$, then the mean time until absorption for fixed initial conditions, or the lifetime of unit i as helix, is

$$\bar{T}_{nimm'} = \int_0^\infty F_{i,m}(t)G_{n-i+1,n-m'}(t)dt$$

By using this method of combining two random walks, we assume them to act independently of one another. The spatial domains of the two walks overlap only at the ith unit. Substitution from eq 7 yields

$$\bar{T}_{nimm'} = k_f^2 \alpha^{(m+m'+1-n)/2} \times \sum_{s,s'} (u_s)_m(u_s)_{i-1}(u'_{s'})_{m'}(u'_{s'})_{n-i}[(\nu_s + \nu'_{s'})\nu_s\nu'_{s'}]^{-1} \tag{8}$$

This is a general result for the lifetime of unit i as helix in a chain of length n with the two initial positions of the helix-coil boundaries at m and m'.

The result in eq 8 must be averaged over all initial conformations. If the all random coil conformation is eliminated and the conventional Zimm-Bragg[6,7] statistical weights are used, then the partition function is given by

$$Z = \sigma \sum_{m=0}^{n-1} \sum_{m'=m+1}^{n} s^{m'-m} \tag{9}$$

where σ and s represent the statistical weights for initiation and propagation of helix. Since the all-coil conformation has been excluded, at least unit i is required to be initially in a helical state. Combination of eq 8 and 9 yields the required average

$$\langle T_{ni}\rangle = Z^{-1}\sigma \sum_{m,m'} s^{m'-m}\bar{T}_{nimm'} \tag{10}$$

We have expressed the transition rates simply in terms of the helix propagation weight s as $k_f = 1/(1+s)$ and $k_b = s/(1+s)$. In the absence of detailed kinetic information this is a simple definition and clearly produces the correct ratio k_b/k_f. Thus it is found that $\alpha = s$. Equation 10 has been utilized to calculate the average lifetimes of units in polypeptide chains of various lengths up to 150 units. This method is limited computationally because the sums in eq 10 must be performed over all possible initial positions of the walkers. The eigenvalues and eigenvectors of all arrays $\mathbf{A}$ of order $\leq n$ are required in order to evaluate the sums in eq 8.

Results

The one-helix sequence partition function utilized for the present purpose is given by $Z' = Z + 1$, where Z is given by eq 9. The values of s at the transition midpoint (fraction of helix units equal 0.5) are found to lie in the range $(s)^{1/2} > 1 + 1/i$.

In Figure 2, the curve for $i = n/2$ represents the longest lifetimes of units in chains of various lengths calculated at values of s corresponding to the transition midpoints. The

(7) B. H. Zimm and J. K. Bragg, *J. Chem. Phys.*, 31, 526 (1959).

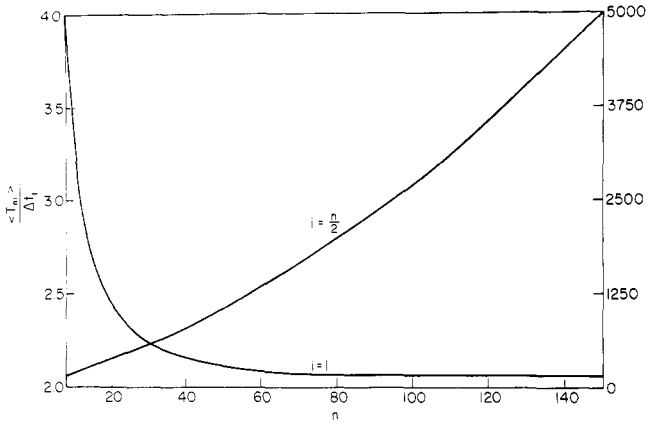

Figure 2. Terminal, $i = 1$, and central, $i = n/2$, helix unit lifetimes at transition midpoints for chains of n units. The partition function Z' was used to determine s at the transition midpoint. A value of $\sigma = 10^{-4}$ was chosen. Ordinate axis on the left is for the terminal unit; axis on the right corresponds to the central unit.

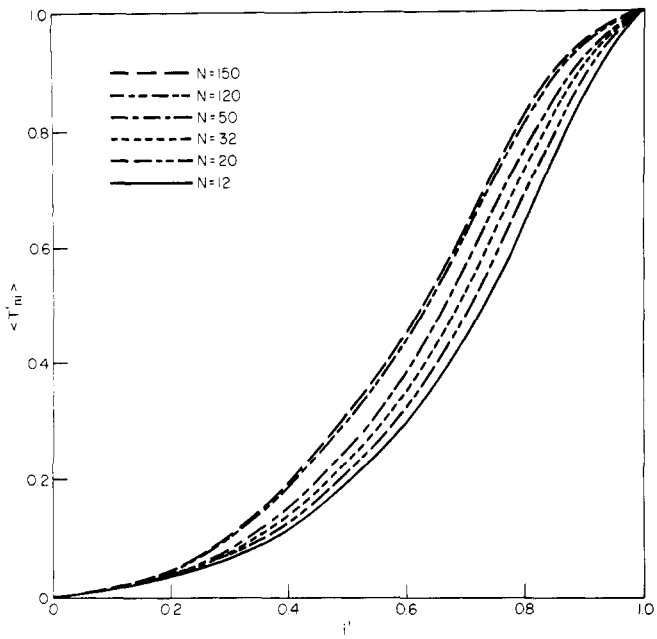

Figure 3. Scaled distributions of lifetimes of helix for all units in chains of various lengths.

ordinate axis scale for this curve is on the right side of the figure. A value[5,7] of $\sigma = 10^{-4}$ has been used throughout. The other curve for $i = 1$ shows the average lifetime of the terminal chain unit. Its ordinate axis scale is on the left. The value of the step time Δt_1 has not been specified here. The longest time for the middle unit of a 150-unit chain was found to be $5 \times 10^3\ \Delta t_1$. If Δt_1 is of order of 10^{-8} sec or less, then lifetimes of all units as helix in chains of even several hundred monomers will be short compared to the nmr observed long time of *ca.* 10^{-2} sec.[2a,3]

In Figure 3 the distributions of helix lifetimes over chain positions are presented. The axes have been normalized in the following manner

$$\langle T'_{ni} \rangle = (\langle T_{ni} \rangle - \langle T_{n1} \rangle)/(\langle T_{n,n/2} \rangle - \langle T_{n1} \rangle)$$

and $i' = (i - 1)/(n/2 - 1)$.

The distributions become significantly broader for longer chains but appear to be converging rapidly to a limiting form within the range of n presented.

The assumption of only one helix region is valid for chains of the size treated here. It is expected[6] to be valid for $n < 2/(\sigma)^{1/2}$. Miller[5] treated also the case with two helix regions and concluded that the one-sequence model gives good estimates of lifetimes within the transition region.

Miller's method for calculating helix lifetimes differs from ours in one important regard. He considered two independent random walks of semiinfinite range. In our method the ranges of walks are finite, with the chain end demarcated by a reflecting barrier. The present method permits the determination of the error introduced by Miller's approximation. For chains of lengths considered here, $n \leq 150$, this approximation has two effects. (1) Helix lifetimes at the chain ends are inaccurate whereas those in the middle of the chain are nearly identical, for values of s at the transition midpoint or above. (2) All lifetimes are less accurate as s approaches unity; for Miller's model lifetimes at $s = 1$ become infinite. The first effect occurs in Miller's approximation because the absence of reflecting barriers most strongly influences the closest units. Deviations of the second type arise because the bias of the walker toward the chain end decreases as s approaches one.[8] As an example, consider a 12-unit chain for which the transition midpoint is at $s = 2.0$. Both methods give a helix lifetime $\langle T_{12,6} \rangle/\Delta t_1$ for the two middle units as 2.3×10^2; however for the first unit our method gives $\langle T_{12,1} \rangle/\Delta t_1$ as 3.0, but Miller's method yields 6.0. At $s = 1.4$ the differences are larger, our values are $\langle T_{12,6} \rangle/\Delta t_1 = 5.4 \times 10^1$ and $\langle T_{12,1} \rangle/\Delta t_1 = 2.4$ and Miller's results are $\langle T_{12,6} \rangle/\Delta t_1 = 7.5 \times 10^1$ and $\langle T_{12,1} \rangle/\Delta t_1 = 7.8$. At the transition midpoint even midchain lifetimes for Miller's model deviate appreciably from ours for long chains. For $n \leq 150$ our calculations agree to sufficient accuracy with Miller's results at the transition midpoint for the lifetime of the middle chain unit. However, if lifetimes of other units or for other points in the transition were required, his method would not suffice.

Much longer lifetimes are calculated[9] at values of s far away from the transition midpoint. These agree with Miller's result; however, here we have presented only the results at the transition midpoints. Dimensionless lifetimes presented here are independent of the times Δt_1. The only parameters entering the calculations are σ and s. However, establishment of a connection with real time requires specification of Δt_1. From particular[3] interpretations of experiments we would expect values of Δt_1 to be 10^{-5}–10^{-8} and Δt_2 to be about 10^{-2}. These interpretations have not been proven.[3] Physically, one might expect the rate of helix initiation to be faster than 10^2 sec^{-1}.

The fraction of molecules in the all-random-coil conformation will decrease with increasing chain length; consequently nmr observation of a separate random-coil peak should become more difficult for longer chains. The degree of polymerization below which this peak is distinguishable will depend upon the conformational statistics and the experimental nmr parameters. If the two nmr peaks arise for a reason other than that favored here, the effect of a slow helix initiation step should still be evident in other experiments at frequencies perhaps somewhat different from the nmr experiments. Direct measurement of Δt_2 should be possible for shorter chains in which a significant fraction of molecules exist in the all-random-coil form. If Δt_2 is as long as we expect, it should be feasible to design an experiment to determine this time by conventional fast reaction methods.

More information is available concerning polypeptide conformations[6,10] than has been utilized in either the Zimm–Bragg model or the simple kinetic scheme present-

(8) W. Feller, "An Introduction To Probability Theory and Its Applications," Vol. 1, 3rd ed, Wiley, New York, N. Y., 1950.
(9) R. L. Jernigan and J. A. Ferretti, unpublished data.
(10) P. J. Flory, "Statistical Mechanics of Chain Molecules," Interscience Publishers, New York, N. Y., 1969.

ed here. For example, knowledge of the variety of random-coil configurations indicates that a better theory should include multiple random-coil states and rates for transitions among them. A general kinetic description of helix–random-coil transitions for an isolated polypeptide is not available. Any general formulation must also include the slower rate for initiation of helix in a random-coil region and perhaps the rate for the process of initiation of coil within a helix sequence.

We conclude that no long-lived helix-containing configurations are likely to exist, at least not for polypeptides comprising less than a few hundred residues. Existence of multiple helix regions will shorten average lifetimes; thus, lifetimes calculated for long chains with one helix sequence would serve as an upper bound for helix lifetimes in all longer chains in which multiple helix regions are important. The distinction between transitions occurring in two regions of time does, however, permit simple conclusions to be drawn. By nmr, a long transition time, as in the transformation of the all-random-coil conformation to any helix-containing conformation, is independent of the short times characteristic of the transitions of all conformations containing helix. Lifetimes of the order of Δt_1 represent rapidly exchanging species; whereas Δt_2 corresponds to slow exchange on the nmr time scale. Qualitatively we would expect two nmr peaks, one for the long-lived totally random-coil conformation and another which arises collectively from all other short-lived, helix-containing, conformations. These conclusions are in agreement with the experimental results.[2a,3]

Conformational Lifetimes in the Helix–Random-Coil Transition Region by Nuclear Magnetic Resonance with Application to Poly(γ-benzyl L-glutamate)[1]

J. A. Ferretti* and R. L. Jernigan

Physical Sciences Laboratory, Division of Computer Research and Technology, National Institutes of Health, Department of Health, Education, and Welfare, Bethesda, Maryland 20014.
Received March 26, 1973

ABSTRACT: Nuclear magnetic resonance spectra of three poly(γ-benzyl L-glutamate) samples of $\overline{DP}_w$ of 55,280 and 910 have been studied in the helix–random-coil transition region. Double peaks whose behavior is chain-length dependent are observed for the α-CH proton resonance. A two-site model for systems undergoing chemical exchange is used to obtain lifetimes and to fit the behavior of the two peaks throughout the transition region. These results are in good agreement with theoretical spectra calculated on the basis of a description of the time-dependent behavior of the transition.

High-resolution nuclear magnetic resonance (nmr) spectroscopy is often used to study the helix–random-coil transition in synthetic polypeptides.[2a,b,3] In a number of examples such as poly(γ-benzyl L-glutamate,[4] poly(L-alanine,[5] and poly(L-arginine),[6] the α-CH and peptide NH protons each show two peaks in the transition region. The detailed behavior of these separate peaks depends on the polypeptide, solvent conditions, and the chain length of the polymer. A variety of explanations have been suggested for this observation of separate peaks. J. H. Bradbury *et al.*[7] and Tam and Klotz[8] considered slow protonation of amide residues; Scheraga *et al.*[9] proposed slow solvation; Nagayama and Wada[10], E. M. Bradbury *et al.*,[11] and Ullman[12] suggest that polydispersity together with a variation in helicity as a function of monomer placement is the cause of the separate peaks. Goodman *et al.*[13] indicate that the double peaks arise from the presence of low molecular weight polypeptide oligomers. Previous investigations in our laboratory have led us instead to conclude that the two peaks are a result of slow nucleation of helix from randomly coiled peptide units.[14,15]

The observation of two separate nmr peaks for a given proton implies that only two conformations or two groups of conformations exist for times as long or longer than 10^{-3}–10^{-1} sec;[16] all other exchanges occur more rapidly. This slow rate determined by nmr differs from relaxation times measured by other techniques. Kinetic measurements such as temperature jump and ultrasonic relaxation techniques utilized in the transition region yield times in the range 10^{-5}–10^{-8} sec.[17-19] Theoretical investigations have been presented,[14,15,20] one in the previous paper,[15] to explain the kinetic and nmr experiments. These models of the transition demonstrate that the observation of two different time scales is not an inconsistency. In terms of these theories the experimentally measured fast times are related to the time for adding to or melting one helix unit

(1) J. A. Ferretti and R. L. Jernigan, paper presented to the Division of Polymer Chemistry, 164th National Meeting of the American Chemical Society, New York, N. Y., Sept 1972; *Polym. Prepr., Amer. Chem. Soc., Div. Polym. Chem.*, 13, 946 (1972).
(2) (a) J. A. Ferretti, *Chem. Commun.*, 1030 (1967). (b) E. M. Bradbury, C. Crane-Robinson, H. Goldman, and H. W. E. Rattle, *Nature (London)*, 217, 812 (1968).
(3) L. Paolillo, P. A. Temussi, E. M. Bradbury, and C. Crane-Robinson, *Biopolymers*, 11, 2043 (1972).
(4) J. A. Ferretti and B. W. Ninham, *Macromolecules* 3, 30 (1970).
(5) J. A. Ferretti and L. Paolillo, *Biopolymers*, 7, 155 (1969).
(6) M. Boublik, E. M. Bradbury, C. Crane-Robinson, and H. W. E. Rattle, *Eur. J. Biochem.*, 12, 258 (1970).
(7) J. H. Bradbury and M. D. Fenn, *Aust. J. Chem.*, 22, 357 (1969).
(8) J. W. O. Tam and I. M. Klotz, *J. Amer. Chem. Soc.*, 93, 1313 (1971).
(9) F. J. Joubert, N. Lotan, and H. A. Scheraga, *Biochemistry*, 9, 2197 (1970).
(10) K. Nagayama and A. Wada, *Chem. Phys. Lett.*, 16, 50 (1972).
(11) E. M. Bradbury, C. Crane-Robinson, and H. W. E. Rattle, *Polymer*, 11, 277 (1970).
(12) R. Ullman, *Biopolymers*, 9, 471 (1970).
(13) M. Goodman, F. Toda, and N. Ueyama, *Proc. Nat. Acad. Sci. U. S.*, 70, 331 (1973).
(14) J. A. Ferretti, B. W. Ninham, and V. A. Parsegian, *Macromolecules*, 3, 34 (1970).
(15) R. L. Jernigan, J. A. Ferretti, and G. H. Weiss, *Macromolecules*, 6, 684 (1973).
(16) J. W. Emsley, J. Feeney, and H. Sutcliffe, "High Resolution Nuclear Magnetic Resonance Spectroscopy," Pergamon Press, New York, N. Y., 1967, p 481.
(17) R. Lumry, R. Legare, and W. G. Miller, *Biopolymers*, 2, 434 (1964).
(18) G. G. Hammes and P. B. Roberts, *J. Amer. Chem. Soc.*, 91, 1812 (1969).
(19) A. Wada, T. Tanaka, and H. Kihara, *Biopolymers*, 11, 587 (1972).
(20) W. G. Miller, *Macromolecules*, 6, 100 (1973).