

Jorge A. Vila¹

Daniel R. Ripoll²

Harold A. Scheraga³

¹ Universidad Nacional de
San Luis,
Facultad de Ciencias Físico
Matemáticas y Naturales,
Instituto de Matemática
Aplicada San Luis
Consejo Nacional de
Investigaciones Científicas y
Técnicas,
Ejército de los Andes 950,
5700 San Luis, Argentina

² Cornell Theory Center,
Ithaca, NY 14853-3801, USA

³ Baker Laboratory of
Chemistry and Chemical
Biology,
Cornell University,
Ithaca, NY 14853-1301, USA

Received 30 May 2000;
accepted 29 June 2000

Influence of Lysine Content and pH on the Stability of Alanine-Based Copolypeptides

Abstract: To account for the relative contributions of lysine and alanine residues to the stability of α -helices of copolymers of these two residues, conformational energy calculations were carried out for several hexadecapeptides at several pHs. All the calculations considered explicitly the coupling between the conformation of the molecule and the ionization equilibria as a function of pH. The total free energy function used in these calculations included terms that account for the solvation free energy and free energy of ionization. These terms were evaluated by means of a fast multigrid boundary element method. Reasonable agreement with experimental values was obtained for the helix contents and vicinal coupling constants ($^3J_{HN\alpha}$). The helix contents were found to depend strongly on the lysine content, in agreement with recent experimental results of Williams et

Correspondence to: Harold A. Scheraga; email: has5@cornell.edu

Contract grant sponsor: National Institutes of Health (NIH), NIH National Center for Research Resources (NIHCRR), National Science Foundation (NSF), National Foundation for Cancer Research, National Research Council of Argentina, and Universidad Nacional de San Luis—Argentina (UNSLA)

Contract grant numbers: GM-14312 and 5 R03 TW00857-03 (NIH), P41RR-04293 (NIHCRR), MCB95-13167 (NSF), and P-328402 (UNSLA)

Biopolymers, Vol. 58, 235–246 (2001)

© 2001 John Wiley & Sons, Inc.

al. (Journal of the American Chemical Society, 1998, Vol. 120, pp. 11033–11043) In the lowest energy conformation computed for a hexadecapeptide containing 3 lysine residues at pH 6, the lysine side chains are preferentially hydrated; this decreases the hydration of the backbone CO and NH groups, thereby forcing the latter to form hydrogen bonds with each other in the helical conformation. The lowest energy conformation computed for a hexadecapeptide containing 6 lysine residues at pH 6 shows a close proximity between the NH_3^+ groups of the lysine side chains, a feature that was previously observed in calculations of short alanine-based oligopeptides. The calculation on a blocked 16-mer of alanine shows a 7% helix content based on the Boltzmann averaged vicinal coupling constants computed from the dihedral angles ϕ , consistent with previous experimental evidence on triblock copolymers containing a central block of alanines, and with earlier theoretical calculations. © 2001 John Wiley & Sons, Inc. Biopoly 58: 235–246, 2001

Keywords: lysine; alanine; α -helices; copolymers; conformational energy calculations; electrostatic interactions; hydration; hexadecapeptides; pH dependence.

INTRODUCTION

It is very important to understand the factors that lead to the formation of α -helices in peptides and proteins. Conformational folding, in particular, is often described as being driven by short-, medium-, or long-range interactions, depending on how many residues separate the strongest interacting residues. The effect of short-range interactions within an amino acid, i.e., the interactions between a side chain and its own backbone, lead to the *intrinsic conformational propensities* of the amino acid. The additional interactions, i.e., those involving more than one residue, are denoted as *extrinsic propensities*. There are reasons to believe that intrinsic interactions play a dominant role in helix formation.^{1–3} Consequently, it is of interest to know the intrinsic helix-forming tendency of each of the 20 naturally occurring amino acids in order to be able to predict the locations of helices in globular proteins.

In principle, it should be possible to obtain a complete characterization of the *intrinsic propensities* for each of the 20 naturally occurring amino acids by examining the thermally induced helix–coil transition of homopolyamino acids in water. However, homopolymers of most of the amino acids are not soluble in water or, if soluble, do not form helical structures between 0 and 100°C. Therefore, copolymers with solubilizing residues have been used to investigate the helix–coil transition. The approaches that were followed include the use of regular-repeating sequence polymers, block copolymers, specific-sequence copolymers, and random copolymers.³ Random copolymers, in which the long-range interactions are averaged out,⁴ were used to determine the *intrinsic propensities* of each of the 20 naturally occurring amino acids using a so-called host–guest technique.⁵ For some amino acids, such as alanine, the *intrinsic propensities* to form a helix were confirmed by other

experimental approaches, such as the use of triblock copolymers.⁶ All this experimental evidence shows that alanine is not a strong helix former—i.e., in terms of the Zimm–Bragg propagation parameter, it has a value of s equal to 1.08 at 0°C,⁷ which means that short oligopeptides should have a very low helix content as discussed by Ingwall et al.⁶ Recently, Kemp and co-workers^{8–11} reported the results of a series of studies of alanine-rich sequences linked N-terminally to a synthetic helix-inducing template and found good agreement for the intrinsic propensities for alanine with the value previously reported by Platzer et al.⁷

On the other hand, Baldwin and co-workers have investigated an important set of specific-sequence copolymers in which charged groups were introduced to solubilize them. In such specific-sequence copolymers, the long-range interactions are not averaged out.⁴ The large observed helix content of 72% at 1°C in a 16-residue copolymer acetyl-AAAAKAAA-KAAA-amide, i.e., 3K(I),¹² is surprising in light of experimental evidence showing that short homopolymers of alanine exhibited no helix content.⁶ From the observation of a 72% helix content, Marqusee et al. concluded that alanine has an unusually high intrinsic propensity to form a helix. We have already examined this problem¹³ using conformational energy calculations, taking electrostatic hydration effects into account by using a finite difference approach^{14–16} to solve the Poisson–Boltzmann equation, assuming fixed protonation states at each pH. Indeed, we found a helix content for 3K(I) in reasonable agreement with the value observed by Marqusee et al.¹² On the other hand, a very similar calculation with a 16-mer of alanine¹³ led to a very low helix content, in agreement with our previous result for a 10-mer in a triblock copolymer.⁶

More recently, experimental work by Williams et al.¹⁷ has shown that three medium-sized helical pep-

tides, which differ only in the number of widely spaced lysine residues present in an alanine-rich region, display a helix content that increases when up to three alanine residues are replaced by lysines. In spite of this evidence, Rohl et al.¹⁸ claimed that helix formation in both standard peptides and template-peptide conjugates can be attributed to the large intrinsic propensities of alanine to form a helix. Scheraga³ has discussed the arguments of Rohl et al. in conjunction with the experimental evidence of William et al.

Despite the plethora of experimental and theoretical evidence about the intrinsic propensities of L-alanine, discrepancies in the interpretation of these results remain. This is due to the fact that the interplay of electrostatic and hydrophobic interactions in the stability of the helical structures remains poorly understood. Proper assessment of the factors that promote helix formation is very important because the intrinsic propensities of each of the naturally occurring amino acids are the basis of methods for attempting to predict secondary structure in proteins. To assess the relative contributions of various factors to helix formation, we have carried out simulations on the following: (a) 3K(I) at pH 6, 11, and 14; and (b) acetyl-AKAAKAKAAKAKAAKA-amide, i.e., 6K(I), at pH 6. In both cases, in contrast to our earlier procedure¹³ in which the charges were assumed to be independent of conformation, our present theoretical study considered the coupling between structure and ionization equilibria¹⁹ by examining the pH-dependent conformational preferences of these polypeptides, including the contribution from the conformational entropy. A comparison is made with existing experimental data from CD and NMR, for the average helix content of both 3K(I) and 6K(I) and for the vicinal coupling constants in solution for 3K(I), respectively. Moreover, simulations for a blocked alanine 16-mer, acetyl-(A)₁₆-amide, were carried out considering that the solvation free energy due to its transfer from the gas phase to solvent can be decomposed into two parts: (a) a term that accounts for the creation of a cavity and (b) a term that accounts for the solvent polarization effects due to the partial charges, calculated using the boundary element method of Vorobjev and Scheraga.²⁰

This approach enabled us: (a) to compare the results with previous simulations for peptides with non-ionizable groups in which the solvation effects due to both nonpolar and partially charged groups were included through a solvent-accessible surface area model²¹; (b) to obtain internal consistency with the calculations of both 3K(I) and 6K(I) in which the contributions from partial charges and nonpolar groups are

computed using the same procedure, and (c) to obtain more detailed information about how solvent polarization due to partial charges is modified when charged groups are introduced into the sequence. This breakdown of the contributions from solvation of the partial and full charges has enabled us to obtain a deeper understanding of the effect of lysine on the helix content of copolymers of lysine and alanine, and to obtain the intrinsic conformational propensities of L-alanine.

METHODS

In previous theoretical work,¹³ we studied the peptide 3K(I) assuming (a) full protonated lysines and (b) unprotonated lysines, respectively. In both cases, we did not consider the coupling between conformation and the process of proton binding/release or the conformational entropy contributions. The total conformational free energy was calculated as a sum of three contributions:

$$E_{\text{tot}} = E_1 + E_2 + E_3 \quad (1)$$

where E_1 is the ECEPP/2 energy; E_2 is a solvation energy estimated with an empirical solvent-accessible surface area model²¹ that includes solvation effects arising from both nonpolar and polar constituents of the polypeptide. However, because this approach does not account for the electrostatic free energy of solvation of the fully charged groups, an additional component, E_3 , was added. E_3 represents the electrostatic contribution to the solvent effects due to charged lysines. This contribution was computed by using the program DELPHI of Honig and co-workers.²² For a detailed description of these calculations, the reader is referred to the paper by Vila et al.,¹³ and references therein. It should be noted that the pH was not considered explicitly in those calculations. The pH dependence was assumed only through the protonated (or unprotonated) state of the lysine side chains. The component E_3 was assumed to be zero when no net charge was present in the sequence, i.e., the solvent polarization contribution to the free energy was given by the term E_2 . In addition, no contribution from conformational entropy was considered in those calculations.

In the present work, the evaluation of the conformational energy follows the procedure recently published^{19,20,23}; i.e., the total free energy, $E(\mathbf{r}_p, \text{pH})$, associated with the conformation, \mathbf{r}_p , of the molecule in aqueous solution at a given pH, can be defined by considering a three-step thermodynamic process (cavity creation, polarization of the solvent, and alteration of the state of proton binding) involved in transferring the neutral polypeptide from the gas phase to the aqueous solution, as

$$E(\mathbf{r}_p, \text{pH}) = E_{\text{int}}(\mathbf{r}_p) + F_{\text{vib}}(\mathbf{r}_p) + F_{\text{cav}}(\mathbf{r}_p) + F_{\text{solv}}(\mathbf{r}_p) + F_{\text{inz}}(\mathbf{r}_p, \text{pH}) \quad (2)$$

where $E_{\text{int}}(\mathbf{r}_p)$ is the internal conformational energy of the molecule in the absence of solvent, assumed to correspond to the ECEPP/3 energy^{24–27} of the neutral molecule; $F_{\text{vib}}(\mathbf{r}_p)$ is the conformational entropy contribution; $F_{\text{cav}}(\mathbf{r}_p)$ is the free energy associated with the process of cavity creation when transferring the molecule from the gas phase into the aqueous solution; $F_{\text{solv}}(\mathbf{r}_p)$ is the free energy associated with the polarization of the aqueous solution; and $F_{\text{inz}}(\mathbf{r}_p, \text{pH})$ is the free energy associated with the change in the state of charge of the ionizable groups due to the transfer of the molecule from the gas phase to the solvent, at a fixed pH value.

The contribution to the total free energy from the conformational entropy of the molecule, $F_{\text{vib}}(\mathbf{r}_p)$, has been approximated by the harmonic vibrational contribution^{28,29} of each conformation obtained by using the ECEPP/3 potential function. $F_{\text{cav}}(\mathbf{r}_p)$ describes the free energy of creation of a cavity to accommodate a *zero-charged* peptide molecule, i.e., all partial atomic charges are set to *zero*. As shown previously,^{30,31} $F_{\text{cav}}(\mathbf{r}_p)$ can be considered as the free energy of transfer of a nonpolar molecule from the gas phase to water. This free energy is proportional to the solvent-accessible surface of the molecule and is given by the following relation:

$$F_{\text{cav}}(\mathbf{r}_p) = \gamma S_{\text{acc}}(\mathbf{r}_p) + b \quad (3)$$

where γ and b are empirical parameters derived from an optimum fitting of Eq. (3) to the experimental free energy of transfer of small straight-chain alkanes into water.³² The numerical values assigned to the parameters γ and b are 0.005 kcal/Å² and 0.860 kcal, respectively.³⁰

The term $F_{\text{solv}}(\mathbf{r}_p)$, corresponding to the free energy of electrostatic solvation, is the difference between the electrostatic polarization free energies of the peptide in the gas phase and in the solvent environment. This term is expressed as

$$F_{\text{solv}}(\mathbf{r}_p) = W_{\text{polz}}(\mathbf{r}_p; \text{solv}) - W_{\text{polz}}(\mathbf{r}_p; \text{gas}) \quad (4)$$

where $F_{\text{solv}}(\mathbf{r}_p)$ is the free energy associated with the polarization of the aqueous solution by the peptide in a microstate of fixed zero charge; $W_{\text{polz}}(\mathbf{r}_p; \text{solv})$ and $W_{\text{polz}}(\mathbf{r}_p; \text{gas})$ are the electrostatic polarization free energy of the protein in the solvent and the gas-phase environment, respectively. The electrostatic polarization free energies can be calculated by using the dielectric continuum model^{22,33,34} to obtain a solution of the Poisson equation³⁵ for the potential of the reaction field. To calculate this reaction field potential, we used the fast Multigrid Boundary Element (MBE) method developed by Vorobjev and Scheraga.²⁰ The charges and atomic radii parameters for the solvation free energy from the PARSE algorithm³⁰ are used in the MBE method. The PARSE parameters were developed specifically for dielectric continuum methods and were obtained from fittings to experimental solvation free energies of small organic molecules. These atomic radii and

charges were selected to reproduce solvation free energies of simple functional groups.

Using a general multisite titration formalism,^{36–38} the term $F_{\text{inz}}(\mathbf{r}_p, \text{pH})$ is calculated by the MBE method as

$$F_{\text{inz}}(\mathbf{r}_p, \text{pH}) = -k_{\text{B}}T \ln Z \quad (5)$$

where Z is the partition function over all 2^N ionization states of the polypeptide, N being the number of ionizable groups in the molecule, k_{B} is the Boltzmann constant, and T is the absolute temperature.

The partition function Z is computed as

$$Z = \sum_{n=1}^{2^N} \exp[-\Delta G(PS, \mathbf{x}_n)/k_{\text{B}}T] \quad (6)$$

where $\mathbf{x}_n = \{[(x_1^n, \dots, x_i^n, \dots, x_N^n)]; (x_i^n = 0, 1)\}$ is the n th protonation microstate, and $\Delta G(PS, \mathbf{x}_n)$ is the free energy of ionization of the n th microstate given by

$$\Delta G(PS, \mathbf{x}_n) = \sum_{i=1}^N x_i(\text{pH} - \text{pK}_i^{\circ} - \Delta \text{pK}_i)(\ln 10)k_{\text{B}}T + \sum_{i,j=1;i < j}^N W_{ij}x_i x_j \quad (7)$$

where W_{ij} is the pairwise potential of mean force (Vorobjev et al.,³⁸) between the i th and j th ionizable groups of the protein; K_i° is the dissociation constant of the i th single, isolated (and charged) residue in the solvent (the value of 10.5 was adopted for pK_a° of the lysine residue); and ΔpK_i is the pK_a shift of this particular residue due to its transfer from the solvent to the protein environment.

In the present study, the pK shift of the i th ionizable group is computed as

$$\Delta \text{pK}_i = \text{pK}_{i,\text{intr}}^{\circ} - \text{pK}_i^{\circ} = \{[E(PS_i^+) - E(PS_i^{\circ})] - [E(S_i^+) - E(S_i^{\circ})]\}/[(\ln 10)k_{\text{B}}T] \quad (8)$$

where $E(PS_i^+)$ and $E(PS_i^{\circ})$ are the total energies of the i th residue in the charged and neutral state, respectively, in the protein environment; $E(S_i^+)$ and $E(S_i^{\circ})$ are the total electrostatic energies of the i th single isolated residue in the charged and neutral state, respectively, in the solvent; pK_i° is the pK_a of the i th single, isolated residue in the solvent; and $\text{pK}_{i,\text{intr}}^{\circ}$ is the pK_a that the i th residue would have if all other ionizable groups in the protein environment were neutral. The free energy $E(PS_i^+)$ is obtained under the assumption that all the ionizable residues, other than residue i , are in a state of zero charge.

The average protonation state of the i th site, $\langle x_i \rangle$, is found as a Boltzmann average of all protonation states of the protein as

$$\langle x_i \rangle = \frac{\sum_{n=1}^{2^N} x_i^n \exp[-\Delta G(PS, \mathbf{x}_n)/k_B T]}{Z} \quad (9)$$

The MBE method provides an accurate and stable calculation of both (a) the potential of mean force, W_{ij} , between ionized groups of the protein, and (b) the pK_a shifts of the ionizable groups as a function of the protein environment. It should be noted that the effect of ionic strength (assumed here to be less than 0.1M) was not included in the present study for the reasons described earlier.¹⁹

The distinctive features of the present calculations are the following: (a) we take into account the coupling between the process of proton binding/release and conformation adopted by the molecule at a given fixed pH [using Eq. (5)], (b) for sequences that do not contain ionizable residues, such as the 16-mer of alanine, the solvation effects arising from both nonpolar and polar constituents of polypeptides are taken into account using Eq. (4), and (c) contributions to the total free energy from the conformational entropy are included. (d) Our calculations of the solvation free energy due to the partial charges of a 16-mer of alanine were carried out using Eq. (2) for consistency with the present calculations for 3K(I) and 6K(I). It should be noted that, in the case of the all-alanine peptide, the contribution from Eq. (5) is zero, but the partial charges produce a non-zero contribution from F_{solv} . Previous calculations on the same peptide¹³ were carried out by assuming that F_{solv} plus F_{cav} were taken into account through a surface area model.

Evaluation of the Helix Content

To compute the helix content we have used two different approaches:

1. We used an empirical relation introduced by Pardi et al.³⁹ based on the identification of helical secondary structure by segments of several successive spin-spin coupling constants ${}^3J_{\text{HN}\alpha}$ less than 6 Hz. In this relation, residues with Boltzmann averaged vicinal coupling constants (${}^3J_{\text{HN}\alpha}$) less than 6 Hz contribute with weights of 0.90, 0.80, and 0.75 if the number of consecutive residues is four, three, or two, respectively, while a single residue contributes with a weight of 0.55. For example, for a hypothetical polypeptide of 20 residues, with groups of six, four, and three consecutive residues having coupling constants less than 6 Hz, and two additional residues having coupling constants less than 6 Hz (each group or single residue surrounded by neighboring residues with coupling constants higher than six), the helix content would be $[(6)(0.9) + (4)(0.8) + (3)(0.75) + (2)(0.55)]/20 = 0.5975 \sim 0.60$.

Adopting this criterion enabled us to distinguish the contributions between isolated residues and short sequences to the total helix content. Values of the

helix content obtained by using this approach are referred to as θ_{coupling} .

2. A second value for the helix content was computed as a ratio between the Boltzmann averaged number of residues in the helical conformation and the total number of residues (16 for the peptides under consideration). The assignment of a residue to a helical state was based on its dihedral angles, ϕ and ψ . A residue was considered to be in the helical state if both ϕ and ψ assumed the canonical values (-60 ± 20 , -40 ± 20). Adopting this criterion enabled us to compare the results with our previous calculations.¹³ Values of the helix content obtained by using this approach are referred to as θ_{dihedral} .

The Conformational Search

For each sequence described in Table I, 18,000 to 43,700 conformations were generated by using a modified version¹⁹ of the electrostatically driven Monte Carlo (EDMC) method.^{40–43} During each of these runs, the generated conformations were energy minimized using the Secant Unconstrained Minimization Solver (SUMSL) algorithm⁴⁴ in combination with ECEPP/3 plus a surface solvation model (SRFOPT),²¹ and their *free energies* were computed by using Eq. (2). Only a small set of low-energy conformations were actually stored. These corresponded to the accepted conformations from the Monte Carlo path followed by the EDMC method in each of the runs. The objective of these Monte Carlo runs was to sample the low-energy regions of the free energy $E(\mathbf{r}_p, \text{pH})$. The solvation free energy of the conformations was always included in the free energy $E(\mathbf{r}_p, \text{pH})$. Further details of the procedure can be found in an earlier publication.¹⁹ It should be pointed out that the acceptance rate (see Table I) is low (4%) for all the runs in our simulations. However, as was already noted,¹³ this acceptance ratio is characteristic of the EDMC procedure. During an EDMC run, generated conformations that lead to the same energy as the current conformation are rejected. This feature, plus the minimization procedure for every generated conformation, eliminates many of them. As an example, if the potential surface had an exact parabolic form, no additional conformations other than the minimum energy would be accepted because of the modified Metropolis criterion.

RESULTS AND DISCUSSION

In the present study, we carried out EDMC runs for the three different polypeptides sequences described in Table I. One sequence corresponds to a 16-mer of alanine while the other two contain lysine and alanine residues in the specific patterns described by Marqusee et al.¹² as 3K(I) and 6K(I). In all the sequences, the end groups were acetyl ($\text{CH}_3\text{CO}-$) and amino ($-\text{NH}_2$). For each of the runs described in Table I,

Table I Summary of the EDMC Runs for the Peptides 3K(I), 6K(I), and a 16-mer of Alanine

| Peptide Sequence | Number of Generated Conformations ^a | Number of Accepted Conformations | Lowest Free Energy (kcal/mol) | Helix Fraction θ_{coupling} (%) | Helix Fraction θ_{dihedral} (%) | Experimental Helix Fraction (%) |
|---------------------|--|----------------------------------|-------------------------------|---|---|----------------------------------|
| 3K(I) ^b | 23,724 | 953 | -264.16 | 56 | 60 (68 ^c) | 72 ^d /82 ^e |
| 3K(I) ^f | 24,570 | 869 | -294.07 | 52 | 56 | Not measured |
| 3K(I) ^g | 31,359 | 937 | -358.21 | 17 | 19 (31 ^c) | Not measured |
| 6K(I) ^h | 18,525 | 849 | -315.53 | 19 | 35 | 19 ^d |
| ALA-16 ⁱ | 43,752 | 1345 | -255.84 | 7 | 12 (6 ^j) | 0 ^k |

^a These values correspond to the number of generated conformations (including solvent) for the runs using the procedure described in Methods.

^b Values in this row were computed at pH 6.

^c In parentheses, the theoretical value determined previously¹² by assuming fixed net charges of the lysine residues: 1.0 e.c.u. at low pH values and 0.0 e.c.u. at high pH.

^d Value determined from CD experiments.¹²

^e Value of θ_{coupling} computed from the experimentally determined values of the NMR vicinal coupling constant.⁴⁶

^f Values in this row were computed at pH 11.

^g Values in this row were computed at pH 14.

^h Values in this row were computed at pH 6.

ⁱ Since this sequence contains no ionizable groups, the results are independent of pH.

^j In parentheses, the theoretical value determined previously,¹² by assuming that the solvation effects arising from both polar and nonpolar constituents of the polypeptide are described by a surface area model. The difference between 12 and 6 arises from differences between the method of Vila et al.¹² and that used in this paper.

^k This value is an estimate derived from the helix content measured for the 10-residue alanine central segment in a triblock copolymer.⁶

more than 18,000 conformations were generated following the procedure described in the Methods section, and the total free energy was computed by using Eq. (2).

Table II shows the Boltzmann averaged degrees of charge and the pK shifts for each lysine residue at pH 11 in the 3K(I) peptide. The corresponding values for the runs at pH 6 for both 3K(I) and 6K(I) were omitted because their Boltzmann averaged degrees of charge correspond to those of fully charged residues.

Table II Values of the Computed $pK_{i,\text{intr}}$, ΔpK_a of the Lys Residues^a and the Corresponding Degrees of Charge in the Lowest Energy Conformation of Ac-AAAAK₅AAAAK₁₀AAAAK₁₅A-NH₂ at pH = 11

| Residue | $pK_{i,\text{intr}}$ ^b | ΔpK_a ^b | Degree of Charge ^c |
|-----------------|-----------------------------------|----------------------------|-------------------------------|
| K ₅ | 10.41 | -0.09 | 0.21 |
| K ₁₀ | 10.44 | -0.06 | 0.22 |
| K ₁₅ | 10.69 | 0.19 | 0.33 |

^a Boltzmann averaged values computed by using the 869 accepted conformations of the simulation. The value of 10.5 adopted for pK_a of the lysine residues was computed as an average from the data of Perrin.⁵⁵

^b ΔpK_a and $pK_{i,\text{intr}}$ were computed by using Eq. (8).

^c The degree of charge was computed by using Eq. (9).

These degrees of charge are in agreement with the well known *Null* or *Zero Interaction* model.^{37,45} This model, in which all ionizable residues are assumed to titrate independently at their standard pK_a s, leads to results¹³ that are consistent with those found in our simulations at pH 6. In other words, our previous simulations for 3K(I) were carried out by keeping the charge distribution fixed during the whole simulation by assuming the validity of the Null model. However, particular attention must be addressed to the results shown in Table II, showing that this assumption is no longer valid for the calculations at pH 11—i.e., the degree of charge of each lysine residue depends on its location within the chain and on the conformation of the chain.

Finally, the Boltzmann averaged values over all accepted conformations for the vicinal coupling constants, $^3J_{HN\alpha}$ for each residue of the different sequences, and for the helix content obtained for each run are displayed in Table III.

pH Dependence of the Average Helix Content for the 3K(I) Peptide

The Boltzmann averaged values of the vicinal coupling constants $^3J_{HN\alpha}$, obtained from the simulations of the 3K(I) peptide at pH 6, are consistent with

Table III Computed^a Boltzmann Average of the Vicinal Coupling Constant (${}^3J_{N\alpha}$)

| Sequence | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|-------------------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 3K(I) ^b | 3.7 | 4.2 | 4.3 | 4.8 | 5.0 | 4.7 | 4.7 | 4.5 | 5.0 | 5.0 | 4.7 | 4.8 | 4.8 | 5.2 | 6.4 | 5.9 |
| 3K(I) ^c | 7.3 | 6.7 | 6.9 | 6.8 | 6.1 | 6.0 | 5.1 | 5.0 | 5.9 | 5.2 | 5.2 | 4.9 | 5.1 | 5.5 | 5.3 | 5.5 |
| 3K(I) ^d | 5.5 | 5.7 | 6.9 | 6.9 | 7.8 | 5.3 | 7.2 | 6.9 | 6.8 | 4.7 | 4.3 | 5.5 | 5.1 | 5.7 | 5.5 | 5.6 |
| 3K(I) ^e | 5.7 | 6.9 | 6.9 | 7.2 | 4.4 | 6.9 | 5.9 | 8.3 | 5.3 | 6.9 | 7.0 | 5.3 | 7.0 | 7.0 | 7.0 | 7.6 |
| 6K(I) ^f | 6.9 | 6.8 | 6.8 | 6.6 | 6.9 | 8.1 | 9.6 | 7.0 | 6.9 | 9.0 | 6.9 | 5.4 | 5.4 | 6.1 | 5.7 | 5.7 |
| Ac-A ₁₆ -NH ₂ | 5.2 | 6.9 | 6.2 | 7.4 | 6.9 | 6.7 | 6.9 | 6.3 | 9.2 | 7.1 | 7.1 | 7.9 | 6.8 | 6.9 | 6.9 | 5.4 |

^a The theoretical values of the coupling constants at pH 6 were computed from the calculated values of ϕ by using the Karplus relation:^{56,57} ${}^3J_{N\alpha} = 6.4 \cos^2\delta - 1.4 \cos\delta + 1.9$, where $\delta = |\phi - 60|$ (in degrees). These values of ${}^3J_{N\alpha}$ were then Boltzmann averaged. Values of the Boltzmann averaged coupling constants less than 6 Hz in the first and second rows are highlighted in boldface type to facilitate the comparison between theory and experiment.

^b Experimental value of $\langle {}^3J_{N\alpha} \rangle$ as determined by NMR at pH 5, assumed to be the same⁴⁶ at pH 6.

^c Values of $\langle {}^3J_{N\alpha} \rangle$ corresponding to the peptide 3K(I) computed at pH 6.

^d Values of $\langle {}^3J_{N\alpha} \rangle$ corresponding to the peptide 3K(I) computed at pH 11.

^e Values of $\langle {}^3J_{N\alpha} \rangle$ corresponding to the peptide 3K(I) computed at pH 14.

^f Values of $\langle {}^3J_{N\alpha} \rangle$ corresponding to the peptide 6K(I) computed at pH 6.

residues in α -helical conformations, leading to a Boltzmann averaged helix content, θ_{coupling} , of 56% for the whole peptide. In addition, use of a Boltzmann average of the residues in helical conformations to compute the helix content leads to a value of θ_{dihedral} of 60%, in agreement with a previous theoretical value¹³ of 68%. Both procedures used to compute the helix content, i.e., θ_{coupling} and θ_{dihedral} , yield values that are comparable to the experimental results from CD (72%).¹² While the two procedures are not independent (θ_{coupling} is computed from the values of the dihedral angles ϕ only, and θ_{dihedral} considers all residues with ϕ and ψ values in the α -helical region), we opted to present both values since θ_{coupling} should be compared directly with the experimentally measured value.

A comparison between the experimentally determined vicinal coupling constants, ${}^3J_{HN\alpha}$, for each residue obtained from NMR experiments at pH 5 by Millhauser et al.⁴⁶ and the corresponding theoretically determined Boltzmann averages at pH 6 is presented in Table III. The agreement for the range of residues from Lys-5 to Ala-16 is quite good, with only three residues, Lys-5, Ala-6, and Lys-15, having values that differ by more than 1.0 Hz. It must be pointed out that, although the simulations and experiments have been conducted at two different (but close) pHs, the experimental titration results for 3K(I)¹² show that the helix content is constant between pH 3 and 10, validating our comparison.

Using the experimental data for the NMR derived vicinal coupling constants and the criterion of Pardi et al.,³⁹ an estimate of the experimental helix content, θ_{coupling} , is 82% that, on the other hand, is greater than the 72% experimentally determined value by CD by

Marqusee et al.¹² Even though the experimentally determined NMR values of ${}^3J_{HN\alpha}$ for the N-terminal residues (1–5) shown in Table III are consistent with the suggested values for a canonical helix,⁴⁷ the NH(*i*)—NH(*i* + 1) nuclear Overhauser effect evidence shows a reduced intensity at both helix termini.⁴⁶ These results have been interpreted by Millhauser et al. as an indication that the N-terminus is unfolded. Although the three N-terminal residues are locked in conformations with negative values of the dihedral angles ϕ within the range $[-65, -55]$, the values for the dihedral angles ψ are different from those corresponding to a helical conformation. Taking into consideration that the three N-terminal residues are not helical, a better estimate of the helix content from the NMR experiment is 65%. As seen in Table III, the theoretical values of the Boltzmann averaged ${}^3J_{HN\alpha}$ of residues 1–5 are greater than 6.0 Hz. While consistent with an unfolded N-terminus, these results show that our simulations are not able to reproduce the conformational properties of the N-terminal portion of the polypeptide.

Calculations for 3K(I) at pH 11 were also carried out, and the results are shown in Tables II and III. The Boltzmann averaged helix content is comparable to the value obtained for 3K(I) at pH 6 (values of $\theta_{\text{coupling}} = 52\%$, and $\theta_{\text{dihedral}} = 56\%$ were obtained). It should be pointed out that the set of conformations used to obtain the averages reported here were obtained from two independent runs. The lowest free energy conformations at pH 6 and 11 are somewhat similar, both having the C-terminal fragment folded in an α -helical conformation. As already noted, the Boltzmann averaged charge on the lysine residues at pH 11 is not zero. On the other hand, at pH 14, the

helix content of 3K(I) of $\theta_{\text{coupling}} = 17\%$ and $\theta_{\text{dihedral}} = 19\%$ are substantially lower than the values at pH 6 and 11, indicating that charged or partially charged lysines are responsible for the higher helix content observed for 3K(I) at pHs 6 and 11.

The average helix content obtained for 3K(I) at pH 14 ($\theta_{\text{dihedral}} = 19\%$) is roughly comparable (within the error of the method) to a previously computed value of 31% obtained from simulations assuming unprotonated lysines.¹³ Thus, both, the old and new simulations consistently show that the Boltzmann averaged helix content is sensitive to the degree of charge of the lysine residues that diminishes when the lysine charges are turned off (as in runs at pH 14, Table III). In particular, the present calculation can be considered more accurate than the previous one because it shows a real pH dependence, i.e., no assumption about the state of charge of the ionizable residues is made at any stage of the simulation. It should be noted that the experiments of Marqusee et al.¹² show an increase of the helix content for 3K(I) for pHs beyond 10. However, as already noted by the authors, they did not investigate the possibility of aggregation at pH > 10. Due to the discrepancy between theory and experiment for pH > 10, we previously suggested¹³ that aggregation might be responsible for the observed increase in the average helix content for pH values > 10.

Helix Content of 6K(I) at pH 6

For the 6K(I) peptide, our calculation at pH 6 reveals a low helix content ($\theta_{\text{coupling}} = 19\%$ or $\theta_{\text{dihedral}} = 35\%$), i.e., comparable to the experimentally observed value (19%) from CD, obtained by Marqusee et al.¹² Inspection of the low-energy conformation found for this peptide shows a characteristic pattern of attractive interactions among lysine side chains, previously observed in simulations of short oligopeptides.⁴⁸ Attractive interactions between ionizable groups of the same type have been observed in x-ray crystal structures of proteins.⁴⁹ Ion pair interactions are the subject of intense studies.⁵⁰ As shown in Fig. 1 (top), the features of the interaction between lysine side chains involve (a) hydrophobic interactions between lysine side chains (Lys-10 and Lys-12), and (b) a close proximity between NH_3^+ groups. These theoretical results for 6K(I) at pH 6 are consistent with the experimental evidence that charged polylysine is not helical, even at high salt concentrations⁵¹ or methanol concentrations below 85%.^{38,52,53}

Helix Content of a 16-mer of Alanine

Computation of the Boltzmann averaged value over all accepted conformations shows a helix content of

$\theta_{\text{coupling}} = 7\%$ and $\theta_{\text{dihedral}} = 12\%$, consistent with both previous¹³ theoretical calculations ($\theta_{\text{dihedral}} = 6\%$), and experimental evidence from triblock copolymers,⁶ the host-guest technique,⁷ and recent experimental studies¹¹ on alanine-rich sequences linked N-terminally to a synthetic helix-inducing template. All of these results are also consistent with alanine being a weak helix-forming residue.

Charge Distribution and Conformational Preference

To help understand the underlying interactions that dominate the formation of the α -helical conformation, we have analyzed the distribution of charged and polar groups in the lowest energy conformations of these peptides. Figure 2 displays the accessible surface of the lowest energy conformation for the 16-mer of alanine. A tendency of the polar groups to interact with the solvent is observed in this conformation but, noticeably, groups with similar partial charges tend to be close to each other in clusters. The lowest energy conformation adopts a toroidal shape with the HN atoms of residues 2–7 and 11–14 forming the inner face of the torus, and most of the CO groups pointing outward. This interesting arrangement seems to maximize solvation of these polar groups, thus making helix formation unfavorable.

Similar analysis of the lowest-energy conformation for 3K(I) at pH 6, shown in Figure 3, reveals that the NH_3^+ groups of the lysine side chains tend to be well separated and exposed to the solvent. Some CO and NH groups form clusters (red and dark yellow patches in Figure 3) that interact with the solvent, but many (7) CO groups have paired with the NH groups in a helical segment. The HN atoms involved in the helix are completely buried while the O atoms are partially exposed to the solvent.

A comparison of the arrangements of the CO and NH groups in the lowest energy conformations for peptide 3K(I) at pH 6, and for the 16-mer of alanine, is quite revealing: the solvation preference of the charged lysine side chain groups in the 3K(I) sequence appears to dominate over that of the CO and NH groups. The conformational preference of the backbone is strongly determined by the solvation preference of the charged groups of lysine, while the CO and NH groups of the backbone are not able to compete with the charged groups for this hydration. As a result of this competition, COs and NHs tend to group in pairs forming hydrogen bonds, leading to an arrangement that favors α -helix conformation. In other words, the charges and preferential hydration of the lysine side chains diminish the hydration of the

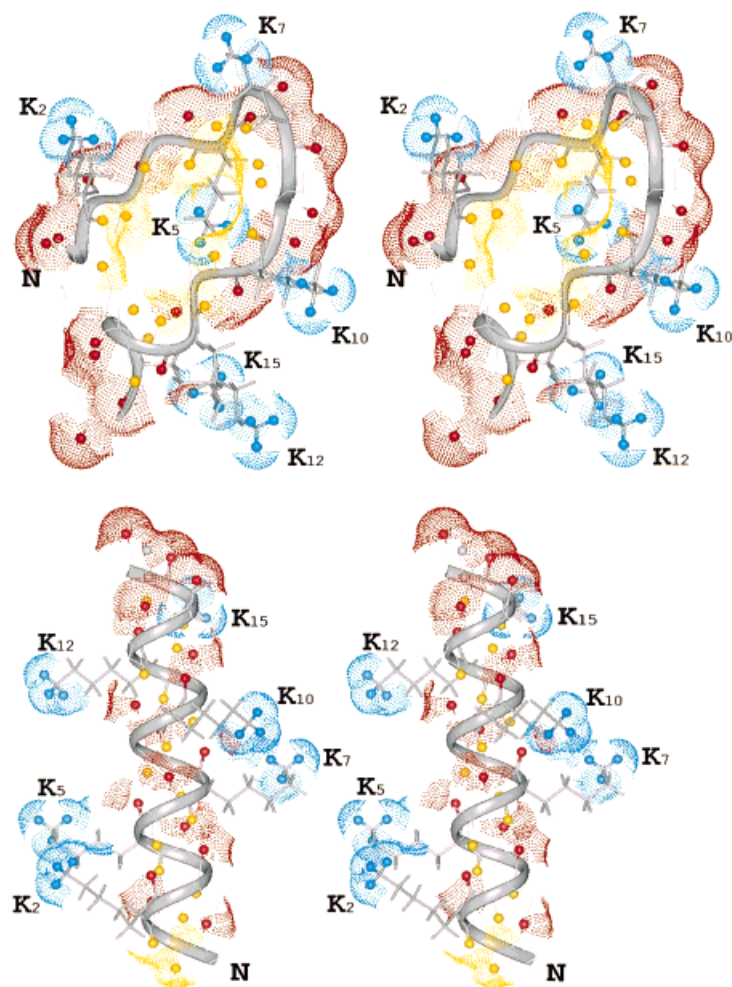


FIGURE 1 (Top) Stereo view of the lowest energy conformation of the 6K(I) peptide in water at pH 6. The backbone of the chain is traced with a gray ribbon. All the atoms have been colored in gray, except for the carbonyl O (red), amino HN (dark yellow), and NH_3^+ (blue) atoms. In addition, the solvent-accessible surfaces⁵⁸ of the oxygen and polar hydrogen atoms are displayed with dots. In close resemblance to previous calculations,⁴⁸ the lysine side chains appear in close proximity. Distances between the N_ϵ atoms of the residues Lys-2–Lys-7, Lys-5–Lys-15, and Lys-10–Lys-12 are 9.5, 9.0, and 8.1 Å respectively. (Bottom) Stereo view of the peptide 6K(I) in an α -helix generated by using standard helical values for the backbone ϕ and ψ dihedral angles (-66° , -40°) followed by local minimization of the ECEPP/3 energy with a surface-solvation model (SRFOPT).²¹ The backbone of the chain is traced with a gray ribbon. All the atoms have been colored in gray, except for the carbonyl O (red), amino HN (dark yellow), and NH_3^+ (blue) atoms. In addition, the solvent-accessible surface of the oxygen and polar hydrogen atoms are displayed using dots. Comparison with Figure 1 (top) shows a complete loss of exposure of the backbone HN atoms after forming hydrogen bonds with the carbonyl O atoms, with the oxygen atoms, on the other hand, remaining partially solvated even when they are involved in hydrogen bonds.

backbone CO and NH groups, thereby forcing the latter to form hydrogen bonds with each other in the helical conformation. The interactions between partial charges of the backbone and side-chain net charges for proteins in water, as those found here, were already discussed by Spassov et al.⁵⁴ These authors suggested that this kind of interaction may be an

important driving force that leads to native side-chain conformations in proteins.

At pH 6, the NH_3^+ groups of the lysine side chains of the 6K(I) peptide are exposed to the solvent but the separation between them, on the other hand, is less than in the case of 3K(I) at pH 6. The lowest-energy conformation adopts the shape of a U with a single

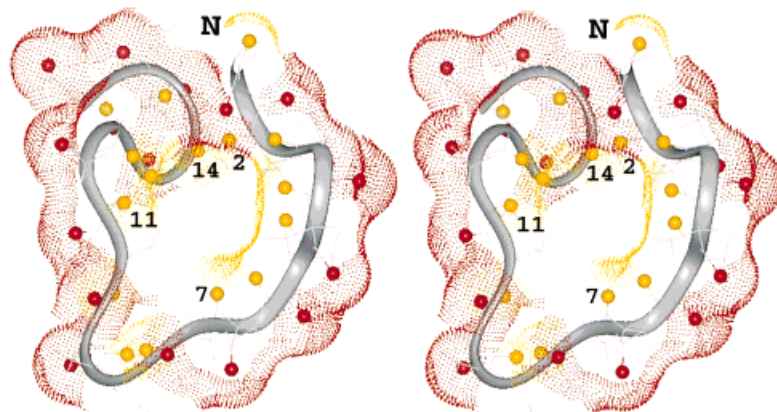


FIGURE 2 Stereo view of the lowest energy conformation of the 16-mer alanine polypeptide in water. All the atoms have been colored in gray, except for the backbone carbonyl O (red) and amino HN (dark yellow) atoms. The solvent-accessible surfaces of the oxygen and polar hydrogen atoms are displayed with dots. The backbone of the chain is traced with a gray ribbon.

helical turn at the C-terminus [Figure 1 (top)]. As in the case of the 16-mer of alanine, the backbone CO and NH groups are grouped in clusters, with the HN pointing toward the internal face of the U and the COs toward the exterior. At pH 14, on the other hand, the solvation effect of the lysine side chains, now unprotonated, is not dominant and helical conformations are less favorable than at pH 6.

Figure 1 (bottom) shows a canonical 6K(I) α -helix (not from simulations; the helical conformation was

generated with the side chains extending toward the solvent). It can be seen that the NH atoms involved in hydrogen bonds are completely buried while the carbonyl CO groups appear to be partially solvated after forming intramolecular hydrogen bonds.

CONCLUSIONS

By considering explicitly the coupling between the conformation of the molecule and the ionization equi-

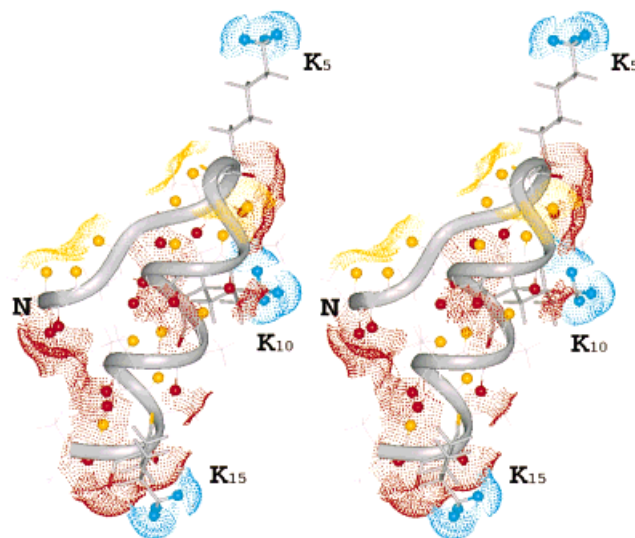


FIGURE 3 Stereo view of the lowest energy conformation of the 3K(I) peptide obtained from simulations in water at pH 6. All the atoms have been colored in gray, except for the carbonyl O (red), amino HN (dark yellow), and NH^ζ (blue) atoms. The backbone of the chain is traced with a gray ribbon and the solvent-accessible surfaces of the oxygen and polar hydrogens are displayed using dots. Only a fraction of the backbone HN groups, not involved in intramolecular hydrogen bonds, remain exposed to the solvent. All the lysine side chains extend into the solvent.

libria at a given pH, we have been able to simulate the conformational ensemble of a series of alanine-based peptides. The results of these theoretical calculations on 3K(I), 6K(I), and on a 16-mer alanine peptide are consistent with the available experimental evidence derived from CD and NMR for 3K(I) and 6K(I) as a function of the pH and with previous experimental evidence on a triblock copolymer of alanine, as well as with previous theoretical calculations.

Our theoretical results consistently show that a low number of charged lysines on an alanine template, such as in the 3K(I) peptide, confers stability to the helical conformation. This effect seems to be due mainly to the disturbance of the solvation preferences of the CO and NH groups by the charged lysine side chains. On the other hand, both deprotonation of the lysine residues at high pH, or an increase in the number of fully protonated lysine residues in the sequence [e.g., 6K(I)], lead to a significant decrease in the helix content. The solvation preference of the protonated side chains of the lysine residues in these alanine-based sequences overcomes that of the polar CO and NH groups, and appears to dictate the preferred conformations by forcing the NH and CO groups to pair among themselves. In the absence of charged groups, i.e., in a 16-residue poly-L-alanine chain, the backbone CO and NH groups tend to interact mostly with the solvent, preventing helix formation.

The most important observations from our simulations are as follows: (a) The backbone CO and NH groups in a polyaniline chain with uncharged side chains interact more favorably with the solvent than among themselves. To interact with the solvent, each kind of group tends to form separate clusters giving rise to extended interfaces; (b) Charged lysine side chains seem to affect considerably the surrounding solvent distribution in both helical and nonhelical conformations of the polypeptides. The charged NH_3^+ groups of the lysine side chains compete with the backbone CO and NH groups and among themselves to gain favorable interactions with the surrounding solvent. In the case of 3K(I), the hydration of lysine side chains is dominant and leads to conformations in which they are well separated from each other while forcing the NH and CO groups to pair among themselves, and hence shift the conformational equilibria to an α -helical structure. In the case of 6K(I), on the other hand, lysine side chains are constrained by the sequence to be close to each other. Their search for favorable arrangements leads to conformations in which they are forced to share their hydration shells ($\text{NH}_3^+ \cdots \text{NH}_3^+$ distance less than 9.5 Å) with other lysines while allowing the NH and CO groups to

group into separate clusters, thus shifting the equilibria to nonhelical conformations. (c) Alanine does not have an unusually high *intrinsic propensity* to form helices, contrary to the proposal by Marqusee et al.¹²; rather, it is the solvation effect due to the lysine residues, as described above, that is responsible for the high helix content of 3K(I).

Work is in progress to determine whether the presence of other ionizable or highly soluble residues lead to similar solvation effects.

Note added in Proof: Such evaluations have been carried out, and are now in press.⁵⁹

We are indebted to G. L. Millhauser for providing us with the values of the vicinal coupling constants for the 3K(I) peptide. This research was supported by grants from the National Institutes of Health (GM-14312 and 5 R03 TW00857-03), the NIH National Center for Research Resources (P41RR-04293), and the National Science Foundation (MCB95-13167). Support was also received from the National Foundation for Cancer Research, the National Research Council of Argentina (CONICET) and Project No P-328402 of the Universidad Nacional de San Luis-Argentina. This research was conducted using the resources of the Cornell Theory Center, which receives funding from Cornell University, New York State, the National Center for Research Resources at the National Institutes of Health (NIH), and members of the Theory Center's Corporate Partnership Program.

REFERENCES

1. Scheraga, H. A. *Pure Appl Chem* 1973, 36, 1–8.
2. Scheraga, H. A. *Pure Appl Chem* 1978, 50, 315–324.
3. Scheraga, H. A. In Vijayan, M.; Yathindra, N.; Kolasakar, A. S., Eds.; *Perspective in Structural Biology*; Indian Academy of Science: Universities Press, India, 1999; pp 275–292.
4. Scheraga, H. A. *Proc Natl Acad Sci USA* 1985, 82, 5585–5587.
5. Wojcik, J.; Altmann, K. H.; Scheraga, H. A. *Biopolymers* 1990, 30, 121–134.
6. Ingwall, R. T.; Scheraga, H. A.; Lotan, N.; Berger, A.; Katchalski, E. *Biopolymers* 1968, 6, 331–368.
7. Platzer, K. E. B.; Ananthanarayanan, V. S.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* 1972, 5, 177–187.
8. Kemp, D. S.; Boyd, J. G.; Muendel, C. C. *Nature* 1991, 352, 451–454.
9. Kemp, D. S.; Curran, T. P.; Boyd, J. G.; Allen, T. J. *J Org Chem* 1991, 56, 6683–6697.
10. Kemp, D. S.; Allen, T. J.; Oslick, S. L. *J Am Chem Soc* 1995, 117, 6641–6657.
11. Kemp, D. S.; Allen, T. J.; Oslick, S. L.; Boyd, J. G. *J Am Chem Soc* 1996, 118, 4240–4248.

12. Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc Natl Acad Sci USA* 1989, 86, 5286–5290.
13. Vila, J.; Williams, R. L.; Grant, J. A.; Wojcik, J.; Scheraga, H. A. *Proc Natl Acad Sci USA* 1992, 89, 7821–7825.
14. Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K. A.; Honig, B. *Proteins Struct Funct Genet* 1986, 1, 47–59.
15. Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J Comput Chem* 1988, 9, 327–335.
16. Gilson, M. K.; Honig, B. H. *Proteins Struct Funct Genet* 1988, 3, 32–52.
17. Williams, L.; Kather, K.; Kemp, D. S. *J Am Chem Soc* 1998, 120, 11033–11043.
18. Rohl, C. A.; Fiori, W.; Baldwin, R. L. *Proc Natl Acad Sci USA* 1999, 96, 3682–3687.
19. Ripoll, D. R.; Vorobjev, Y. N.; Liwo, A.; Vila, J. A.; Scheraga, H. A. *J Mol Biol* 1996, 264, 770–783.
20. Vorobjev, Y. N.; Scheraga, H. A. *J Comput Chem* 1997, 18, 569–583.
21. Vila, J.; Williams, R. L.; Vásquez, M.; Scheraga, H. A. *Proteins Struct Funct Genet* 1991, 10, 199–218.
22. Nicholls, A.; Honig, B. *J Comp Chem* 1991, 12, 435–445.
23. Ripoll, D. R.; Vila, J. A.; Villegas, M. E.; Scheraga, H. A. *J Mol Biol* 1999, 292, 431–440.
24. Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J Phys Chem* 1975, 79, 2361–2381.
25. Némethy, G.; Pottle, M. S.; Scheraga, H. A. *J Phys Chem* 1983, 87, 1883–1887.
26. Sippl, M. J.; Némethy, G.; Scheraga, H. A. *J Phys Chem* 1984, 88, 6231–6233.
27. Némethy, G.; Gibson, K. D.; Palmer, K. A.; Yoon, C. N.; Paterlini, G.; Zagari, A.; Rumsey, S.; Scheraga, H. A. *J Phys Chem* 1992, 96, 6472–6484.
28. Gō, N.; Scheraga, H. A. *J Chem Phys* 1969, 51, 4751–4767.
29. Zimmerman, S. S.; Scheraga, H. A. *Biopolymers* 1977, 16, 811–843. Erratum: *Biopolymers* 1977, 16, 1385.
30. Sitkoff, D.; Sharp, K. A.; Honig, B. *J Phys Chem* 1994, 98, 1978–1988.
31. Simonson, T.; Brünger, A. T. *J Phys Chem* 1994, 98, 4683–4694.
32. Chothia, C. *J Mol Biol* 1976, 105, 1–14.
33. Zauhar, R. J.; Morgan, R. S. *J Comp Chem* 1988, 9, 171–187.
34. Rashin, A. A. *J Phys Chem* 1990, 94, 1725–1733.
35. Vorobjev, Y. N.; Grant, J. A.; Scheraga, H. A. *J Am Chem Soc* 1992, 114, 3189–3196.
36. Bashford, D.; Karplus, M. *Biochemistry* 1990, 29, 10219–10225.
37. Yang, A.-S.; Honig, B. *J Mol Biol* 1993, 231, 459–474.
38. Vorobjev, Y. N.; Scheraga, H. A.; Hitz, B.; Honig, B.; *J Phys Chem* 1994, 98, 10940–10948.
39. Pardi, A.; Billeter, M.; Wüthrich, K. *J Mol Biol* 1984, 180, 741–751.
40. Ripoll, D. R.; Scheraga, H. A. *Biopolymers* 1988, 27, 1283–1303.
41. Ripoll, D. R.; Scheraga, H. A. *J Protein Chem* 1989, 8, 263–287.
42. Ripoll, D. R.; Liwo, A.; Scheraga, H. A. *Biopolymers* 1998, 46, 117–126.
43. O'Donnell, T. J.; Hotovy, S. G.; Pottle, M. S.; Ripoll, D. R.; Scheraga, H. A. In *Lecture Notes in Computer Science*, Liddell, H., Colbrook, A., Hertzberger, B., Sloat, P., Eds.; Springer-Verlag, Berlin, Heidelberg, New York 1996, 1067, 365–372.
44. Gay, D. M. *ACM Trans Math Software* 1983, 9, 503–524.
45. Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J Mol Biol* 1994, 238, 415–436.
46. Millhauser, G. L.; Stenland, C. J.; Bolin, K. A.; van de Ven, F. J. M. *J Biomol NMR* 1996, 7, 331–334.
47. Wüthrich, K. *NMR of Proteins and Nucleic Acids*, John Wiley & Sons: New York, 1986; Chap 9, pp 166–168.
48. Vila, J. A.; Ripoll, D. R.; Villegas, M. E.; Vorobjev, Y. N.; Scheraga, H. A. *Biophys J* 1998, 75, 2637–2646.
49. Magalhaes, A.; Maigret, B.; Hoflack, J.; Gomes, J. N. F.; Scheraga, H. A. *J Protein Chem* 1994, 13, 195–215.
50. Cho, K.-H.; No, K. T.; Scheraga, H. A. *J Phys Chem* 2000, A104, 6505–6509.
51. Barskaya, T. V.; Puitsyn, O. B. *Biopolymers* 1971, 10, 2181–2197.
52. Joubert, F. J.; Lotan, N.; Scheraga, H. A. *Physiol Chem Phys* 1969, 1, 348–354.
53. Liem, R. K. H.; Poland, D.; Scheraga, H. A. *J Am Chem Soc* 1970, 92, 5717–5724.
54. Spassov, V. Z.; Ladenstein, R.; Karshikoff, A. D. *Protein Sci* 1997, 6, 1190–1196.
55. Perrin, D. D. *Dissociation Constants of Organic Bases in Aqueous Solution*, 390; London: Butterworths, 1965.
56. Karplus, M. *J Chem Phys* 1959, 30, 11–15.
57. Karplus, M. *J Am Chem Soc* 1963, 85, 2870–2871.
58. Connolly, M. L. *J Appl Crystallogr* 1983, 16, 548–558.
59. Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. *Proc Natl Acad Sci USA* 2000, in press.