View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Role of Hydrophobicity and Solvent-Mediated Charge-Charge Interactions in Stabilizing α-Helices

Jorge A. Vila,*# Daniel R. Ripoll,§ Myriam E. Villegas,* Yury N. Vorobjev,¶ and Harold A. Scheraga#

*Universidad Nacional de San Luis, Facultad de Ciencias Físico Matemáticas y Naturales, and Instituto de Matemática Aplicada San Luis, Consejo Nacional de Investigaciones Científicas y Técnicas, 5700 San Luis, Argentina; #Baker Laboratory of Chemistry and Chemical Biology and §Cornell Theory Center, Cornell University, Ithaca, New York 14853 USA; and ¶Department of Biochemistry, FLOB, The University of North Carolina, Chapel Hill, North Carolina 27599 USA

ABSTRACT A theoretical study to identify the conformational preferences of lysine-based oligopeptides has been carried out. The solvation free energy and free energy of ionization of the oligopeptides have been calculated by using a fast multigrid boundary element method that considers the coupling between the conformation of the molecule and the ionization equilibria explicitly, at a given pH value. It has been found experimentally that isolated alanine and lysine residues have somewhat small intrinsic helix-forming tendencies; however, results from these simulations indicate that conformations containing right-handed α-helical turns are energetically favorable at low values of pH for lysine-based oligopeptides. Also, unusual patterns of interactions among lysine side chains with large hydrophobic contacts and close proximity (5–6 Å) between charged NH_3^+ groups are observed. Similar arrangements of charged groups have been seen for lysine and arginine residues in experimentally determined structures of proteins available from the Protein Data Bank. The lowest-free-energy conformation of the sequence Ac-(LYS)$_6$-NMe from these simulations showed large pK_α shifts for some of the NH_3^+ groups of the lysine residues. Such large effects are not observed in the lowest-energy conformations of oligopeptide sequences with two, three, or four lysine residues. Calculations on the sequence Ac-LYS-(ALA)$_4$-LYS-NMe also reveal low-energy α-helical conformations with interactions of one of the LYS side chains with the helix backbone in an arrangement quite similar to the one described recently by Groebke et al., 1996 (*Proc. Natl. Acad. Sci. U.S.A.* 93:4025–4029). The results of this study provide a sound basis with which to discuss the nature of the interactions, such as hydrophobicity, charge-charge interaction, and solvent polarization effects, that stabilize right-handed α-helical conformations.

INTRODUCTION

One of the fundamental problems in the biophysical chemistry of proteins is the mechanism by which the amino acid sequence determines the three-dimensional structure of a protein. This is known as the protein folding problem (Gibson and Scheraga, 1988; Vásquez et al., 1994). Proteins are commonly built from secondary structural elements, mainly α-helices, β-structures, and turn conformations. Research has been focused on the nature of the factors that stabilize these elements of secondary structure to gain insights into the mechanisms involved in protein folding (Richardson, 1976; Zimmerman and Scheraga, 1977a; Chou et al., 1990; Wilmot and Thornton, 1990; Kemp et al., 1991a).

The α-helix conformation is not only the most abundant element of secondary structure observed in proteins (Creighton, 1983) but is also the most studied, both theoretically (Poland and Scheraga, 1965, 1970; Finkelstein et al., 1991) and experimentally (Padmanabhan et al., 1990; Bradley et al., 1990; Wojcik et al., 1990; O'Neil and DeGrado, 1990; Lyu et al., 1990). The data obtained in these investigations are very useful for understanding the primary events that take place during the protein folding process. One goal of these studies is to be able to identify short fragments in protein sequences that can form α-helices in isolation. These fragments may act as possible chain-folding initiation sites in the whole protein (Anfinsen, 1972; Finkelstein and Ptitsyn, 1976; Matheson and Scheraga, 1978; Presta and Rose, 1988; Montelione and Scheraga, 1989).

In an attempt to understand the roles that the amino acid side chains and the solvent play in determining the structure and stability of α-helices, we developed the host-guest technique (Von Dreele et al., 1971a,b; Ananthanarayanan et al., 1971), which was used to evaluate the helix-coil stability constants for each of the 20 naturally occurring amino acids in water (Wojcik et al., 1990). In particular, the Zimm-Bragg parameters (Zimm and Bragg, 1959) σ and s for L-alanine, derived from host-guest studies (Platzer et al., 1972) agree substantially well with those derived from studies of tri-block co-polymers (Ingwall et al., 1968). The value of $s = 1.08$ (at 20°C) for the intrinsic helix-forming tendency of L-alanine implies that short oligopeptides of L-alanine should have a very low helix content. In fact, the observed helix content for a tri-block co-polymer with a central block of 10 residues of L-alanine was indistinguishable from zero (Ingwall et al., 1968). Recently, Kemp and co-workers (1991a,b, 1995, 1996) reported the results of a series of studies of conjugates Ac-Hel$_1$-(ALA$_n$-LYS-ALA$_m$)-NH$_2$ and analogs in aqueous solution. In agreement

Received for publication 6 July 1998 and in final form 27 August 1998.

¶Dr. Scheraga is on leave from the Novosibirsk Institute of Bioorganic Chemistry, Novosibirsk, 630090, Russia.

Address reprint requests to Dr. Harold A. Scheraga, Dept. of Chemistry and Chemical Biology, Cornell University, Baker Laboratory, Ithaca, NY 14853-1301. Tel.: 607-255-4034; Fax: 607-254-4700; E-mail: has5@cornell.edu.

© 1998 by the Biophysical Society

0006-3495/98/12/2637/10 $2.00

with the results from the host-guest study, they found s to be close to 1 for L-alanine at 25°C.

However, because many naturally occurring amino acids such as L-alanine are not soluble in water, charged residues have been inserted inside synthetic oligopeptides to solubilize them in studies designed to ascertain the relative helix stability of these amino acids in water (Marqusee et al., 1989; Padmanabhan et al., 1990; O'Neil and DeGrado, 1990; Lyu et al., 1990; Merutka et al., 1990). Among others, Marqusee et al. (1989) found that a synthetic 16-residue alanine-based oligopeptide with the sequence acetyl-AAAAKAAAAKAAAAKA-amide (3K(I)) adopts conformations with an unusually high α-helix content. The tendency of this oligopeptide to adopt helical conformations was attributed to a high helical potential of the L-alanine residue. This assumption of a high helical potential for L-alanine is in contrast with the results obtained from host-guest and block co-polymers experiments (Platzer et al., 1972; Ingwall et al., 1968). In an attempt to account for this apparent discrepancy, we carried out simulations (Vila et al., 1992) on the oligopeptides 3K(I) and acetyl-AAQAAAAQAAAAQAAY-amide (AQY) and obtained helix contents in agreement with experimental values. From these computations, we concluded that the lysine and glutamine residues that solubilize 3K(I) and AQY, respectively, confer stability on the α-helix conformation. By contrast, our theoretical analysis involving both statistical and molecular mechanics calculations on a 16-residue poly (L)-alanine with no interior lysines or glutamines showed very low helix content, in complete agreement with predictions of the helix fraction based on the s values reported by several authors (Platzer et al., 1972; Ingwall et al., 1968; Kemp et al., 1991a,b, 1995, 1996).

Despite these important results, it appears that a number of interesting questions had not been answered, either theoretically or experimentally. Among them, if both alanine ($s = 1.08$ at 25°C (Platzer et al., 1972)) and lysine ($s = 0.94$ at 25°C (Dygert et al., 1976)) are not strong helical formers, as determined by the host-guest technique, why is it that sequences containing charged lysine residues do not prefer extended conformations, as is intuitively assumed? And if the α-helical conformation is stabilized by charges as we proposed (Vila et al., 1992), is it possible to identify the nature of the stabilizing interactions?

The roles of electrostatic and hydrophobic interactions for the stability of α-helical structures are not yet fully understood. To accomplish this goal, we have carried out simulations of lysine- and alanine-based oligopeptides. Among the sequence patterns studied are Ac-KAK-NMe and Ac-KAAAAK-NMe present in 6K(I) and 3K(I) oligopeptides, respectively, and Ac-K_n-NMe, with $n = 2$, 3, 4 and 6. In all cases, a pH value of 2 was assumed. The current theoretical study considers the coupling between structure and ionization equilibria by examining the pH-dependent conformational preferences of these short oligopeptides (Ripoll et al., 1996).

METHODS

Evaluation of the conformational energy

The evaluation of the conformational energy follows the procedure recently published (Ripoll et al., 1996; Vorobjev and Scheraga, 1997); 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:

$$E(\mathbf{r}_p, \text{pH}) = E_{int}(\mathbf{r}_p) + F_{vib}(\mathbf{r}_p) + F_{cav}(\mathbf{r}_p) + F_{solv}(\mathbf{r}_p) + F_{inz}(\mathbf{r}_p, \text{pH}), \quad (1)$$

where $E_{int}(\mathbf{r}_p)$ is the internal conformational energy of the molecule in the absence of solvent, assumed to correspond to the ECEPP/3 energy (Momany et al., 1975; Némethy et al., 1983, 1992; Sippl et al., 1984) of the neutral molecule; $F_{vib}(\mathbf{r}_p)$ is the conformational entropy contribution; $F_{cav}(\mathbf{r}_p)$ is the free energy associated with the process of cavity creation when transferring the molecule from the gas phase into the aqueous solution; $F_{solv}(\mathbf{r}_p)$ is the free energy associated with the polarization of the aqueous solution; and $F_{inz}(\mathbf{r}_p, \text{pH})$ is the free energy associated with the change in the state of ionization 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, $F_{vib}(\mathbf{r}_p)$, has been approximated by the harmonic vibrational contribution (Gō and Scheraga, 1969; Zimmerman and Scheraga, 1977b) of each conformation obtained by using the ECEPP/3 potential function. $F_{cav}(\mathbf{r}_p)$ 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 (Sitkoff et al., 1994; Simonson and Brünger, 1994), $F_{cav}(\mathbf{r}_p)$ 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.

The pK shift of the ith ionizable group is computed as:

$$\Delta pK_i = \{[E(PS_i^+) - E(PS_i^o)] - [E(S_i^+) - E(S_i^o)]\}/[(\ln 10)\, k_B T], \quad (2)$$

where $E(PS_i^+)$ and $E(PS_i^o)$ are the total energies of the ith residue in the ionized and neutral state, respectively, in the protein environment, and $E(S_i^+)$ and $E(S_i^o)$ are the total electrostatic energies of the ith single isolated residue in the ionized and neutral state, respectively, in the solvent. 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. It should be noted that the effect of ionic strength (assumed here to be less than 0.1 M) was not included in the present study for the reasons described previously (Ripoll et al., 1996). The MBE method provides an accurate and stable calculation of both the potential of mean force between ionized groups of the protein and the pK shifts of the ionizable groups as a function of the protein environment. Additional details of the calculation of the electrostatic properties are provided by Ripoll et al. (1996).

Generation of the oligopeptide chains

A large part of the computational work was devoted to identifying low-energy structures that constitute a representative ensemble of the conformations in solution. The conformations corresponding to the different sequences under investigation were constructed using the ECEPP/3 algorithm (Momany et al., 1975; Némethy et al., 1983, 1992; Sippl et al., 1984). The program considers the complete set of backbone and side-chain dihedral angles as the independent variables while the bond lengths and bond angles of the oligopeptide chain are maintained fixed at their ECEPP/3 values.

The free energy terms, $F_{solv}(\mathbf{r}_p)$ and $F_{inz}(\mathbf{r}_p, \text{pH})$, associated with electrostatic solvation in Eq. 1 are very costly to compute. A full search for the global minimum of the function represented by Eq. 1, requiring the energy-minimization of thousands of conformations, is beyond current computational capabilities. For this reason, a protocol that produces a reasonable sampling of the conformational space defined by $E(\mathbf{r}_p, \text{pH})$ without minimizing this particular function, was used.

The conformational search

For each sequence, 200 or more local-minimum conformations were obtained from a series of conformational search runs carried out by using a modified version (Ripoll et al., 1996) of the electrostatically driven Monte Carlo (EDMC) method (Ripoll and Scheraga, 1988, 1989; Ripoll et al., 1998; O'Donnell et al., 1996). During each of these runs, no less than 5000 conformations were generated. After energy minimization using the secant unconstrained minimization solver (SUMSL) algorithm (Gay, 1983) in combination with 1) ECEPP/3 or 2) ECEPP/3 plus a surface solvation model (SRFOPT) (Vila et al., 1991), their free energies were computed by using Eq. 1. The objective of these Monte Carlo runs was to sample the low-energy regions of the free energy $E(\mathbf{r}_p, \text{pH})$. It is worth noting that steps 1 and 2 are associated only with the generation procedure and are intended to enhance the diversity of the conformations generated. The solvation free energy of the conformations was always included in the free energy $E(\mathbf{r}_p, \text{pH})$. Additional details of the procedure can be found in an earlier publication (Ripoll et al., 1996).

Evaluation of the helix content

As only very short sequences are considered here, a common definition of helix content (Ripoll et al., 1996) is not quite appropriate. Consequently, we have computed two quantities that represent the preference of the residues for the right-handed α-helical conformation: 1) F_A, the fraction of residues in the A region of the ϕ-ψ map as defined by Zimmerman et al. (1977) ($-110° \leq \phi < -40°$ and $-90° \leq \psi < -10°$) and 2) F_{Jcoup}, the fraction of residues with a Boltzmann-averaged vicinal coupling constant ($^3\mathbf{J}_{HN\alpha}$) less than or equal to 6 Hz (Wüthrich, 1986).

Both quantities were computed with the equation:

$$F_X = n_X / N_{res}, \tag{3}$$

where the subscript X stands for A or J_{coup}; n_X is the number of residues in the A region of the ϕ-ψ map in definition 1 or the number of residues with the Boltzmann-averaged coupling constant ($^3\mathbf{J}_{HN\alpha}$) less than or equal to 6 Hz in definition 2. N_{res} is the total number of residues in the oligopeptide chain.

RESULTS

In the present study, EDMC runs were carried out for the several oligopeptide sequences described in Table 1. Four sequences correspond to homo-oligopeptides of the lysine residue. Two other sequences contain lysine and alanine residues in some of the patterns present in sequences studied by Marqusee et al. (1989), whereas the remaining sequences correspond to homo-oligopeptides of alanine. All the EDMC runs were started from randomly generated conformations; i.e., all the initial dihedral angles were produced by a random number generator. For each of the runs for the sequences containing lysine residues, more than 5000 conformations were generated, and the total free energy given by Eq. 1 was computed (see Table 1). For the runs with sequences of homo-oligopeptides of alanine, the numbers of generated conformations exceeded 11,000.

TABLE 1 Summary of the EDMC runs on various oligopeptide sequences

Sequence*	Number of generated conformations	Number of accepted conformations[#]	Lowest energy[§] (kcal/mol)	Duration of the run[¶]
KK	5996	200	−95.10	80.9
KKK	5990	200	−130.73	109.2
KKK[‖]	4010	200	−127.78	31.6
KAK	5283	200	−110.60	81.4
KAK[‖]	3902	200	−106.39	25.0
AAA[‖]	12305	1000	−20.19	1.0
KKKK	5896	201	−166.58	131.4
KKKK[‖]	5380	200	−161.39	58.8
KKKKKK[‖]	9052	270	−225.29	164.6
KKKKKK	5765	200	−231.34	101.0
KAAAAK	5436	201	−143.16	55.2
KAAAAK[‖]	15539	443	−138.23	162.9
AAAAAA[‖]	11305	1000	−34.95	8.2

*In all sequences, the backbone was terminated at the α-amino end by an acetyl (CH_3CO-) group and at the α-carboxyl end by a methylamide ($-NHCH_3$) group. The degree of ionization of the LYS residues is a pH- and conformation-dependent quantity and is computed by the MBE algorithm. However, as the calculations were carried out by assuming pH = 2, the LYS residues were fully ionized in practically all of the EDMC accepted conformations.

[#]These are the conformations accepted by using the Metropolis criterion (Metropolis et al., 1953) during the EDMC runs.

[§]See Table 2 for conformations of lowest energy.

[¶]The CPU time for the run for sequence AAA (5201 s ~1.4 hs) was taken as the reference. The CPU times for the other runs are expressed as multiples of the reference CPU time. Calculations were carried out on an IBM SP/2 computer using a parallel code with up to 16 processors and on a single-processor SGI Impact.

[‖]Energy minimization of the conformations in this run was carried out by using the ECEPP/3 plus a solvation surface model (SRFOPT). In all of the other runs, energy minimization of the conformations was carried out by using only ECEPP/3. In all cases, the minimization procedure was carried out by assuming a neutral sequence; i.e., lysine residues were assumed to be uncharged during minimization, but partial charges were placed on every atom.

The calculations for lysine homo-oligopeptides composed of two, three, four, and six residues at pH 2 show that these sequences tend to prefer conformations with a high α-helix content as can be seen in Table 2. For two and three lysine residues, all the residues of the lowest-energy conformations (shown in Figs. 1 and 2, respectively) lie in region A of the ϕ-ψ map (Zimmerman et al., 1977). However, conformations with all the residues in the A* region ($110° \geq \phi > 40°$ and $90° \geq \psi > 10°$), corresponding to the left-handed α-helix, are also energetically very favorable. For the two-lysine oligopeptide, the best-optimized conformation with both residues in the A* region is 1.3 kcal/mol higher than the global minimum (with both residues in the A region) whereas, for the three-lysine oligopeptide, the all-A* conformation is only 0.07 kcal/mol higher than the all-A conformation. The Boltzmann-averaged values for the vicinal coupling constants, $^3\mathbf{J}_{HN\alpha}$, computed over all accepted conformations, and the computed helix content of

TABLE 2 Lowest-energy conformations and overall helix content

Sequence*	Zimmerman et al. code#/⟨$^3J_{HN\alpha}$⟩§ Residue 1	Residue 2	Residue 3	Residue 4	Residue 5	Residue 6	F_A¶/$F_{J\ coup}$‖ (%)
KK	A/5.8	A/6.2					100/50
KKK	A/6.3	A/6.0	A/6.1				100/33
KAK	A/5.5	A/5.4	A/5.5				100/100
AAA	A/7.1	A/7.0	A/7.0				100/0
KKKK	A/5.8	A/5.7	A/5.2	A/5.0			100/100
KKKKKK	A/6.4	A/5.0	A/4.9	A/5.2	A/5.7	C/6.9	83/67
KAAAAK	A*/7.0	A/5.3	A*/6.3	A/5.6	A/4.9	A/5.8	67/67
AAAAAA	C/6.2	G/6.6	D/6.5	G/7.5	A/7.0	A/6.7	33/0

*In all sequences, the backbone was terminated at the α-amino end by an acetyl (CH_3CO-) group and at the α-carboxyl end by a methylamide ($-NHCH_3$) group.

#The characters on the left side of columns 2 to 7 correspond to the Zimmerman et al. (1977) codes of each residue in the lowest-energy conformation encountered for the sequence indicated in column 1.

§The values on the right side of columns 2 to 7 correspond to the computed Boltzmann-averaged vicinal coupling constant ($^3J_{HN\alpha}$) of each residue in the sequence indicated in column 1. The average was computed by using the total number of accepted conformations from all of the EDMC runs. For sequences for which more than one EDMC run was carried out, the average was computed using the accepted conformations from all of the runs. The theoretical values for the coupling constant (in Hz) were computed by using the Karplus relation (Karplus, 1959, 1963): $^3J_{HN\alpha} = 6.4\cos^2\delta + 1.9$, where $\delta = |\phi - 60|$ (in degrees). A residue with $\langle ^3J_{HN\alpha}\rangle$ less than or equal to 6 Hz is considered to be in an α-helical conformation (Wüthrich, 1986).

¶Helix content of the lowest-energy conformation for the sequences given in column 1. The values were computed by using Eq. 3 and the data on the left side of columns 2 to 7.

‖Helix content for the sequences given in column 1. The values were obtained by using Eq. 3 and the Boltzmann-averaged vicinal coupling constants for each residue (data on the right side of columns 2 to 7).

each sequence, are listed in Table 2. The low values of F_{Jcoup} for the sequences KK and KKK (50% and 33%, respectively) arise from the contribution of coupling constants greater than 6 Hz (from the left-handed α-helical conformations) to the Boltzmann average $\langle ^3\mathbf{J}_{HN\alpha}\rangle$.

The homo-oligopeptide sequence with four lysine residues also exhibits a large preference for α-helical conformations (see Table 2). In this case, the lowest-energy conformation (−166.58 kcal/mol) is fully α-helical; however, a second conformation (A*AAA) quite close in energy (−166.46 kcal/mol) also exists. In the latter conformation, both hydrophobic interactions between side chains of residues in positions 1 and 4 and charge-charge solvent-mediated interactions of all NH_3^+ groups are optimized (see Fig. 3, which also shows the solvent-accessible surface (Connolly, 1983a,b). In particular, it is interesting to observe the arrangements of the side chains of residues 1 and 4. These side chains are packed together, enhancing the hydrophobic interactions described above, whereas the NH_3^+ groups at the tips of the side chains tend to separate from each other by pointing in opposite directions (Nζ −Nζ distance = 8.1

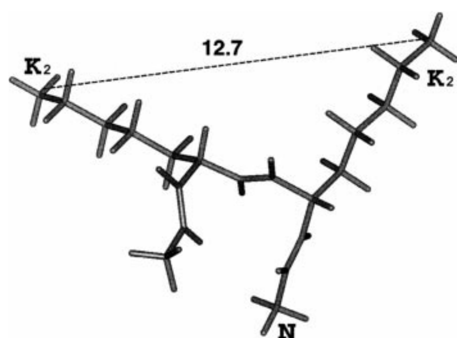

FIGURE 1 Lowest energy conformation (E = −95.10 kcal/mol) found for the sequence KK. The amino terminus of the chain is indicated with the letter N, and a dashed line is used to show the distance (in angstroms) between the NH_3^+ groups of the two lysines. Both LYS residues are in the A (Zimmerman et al., 1977) region of the ϕ-ψ map.

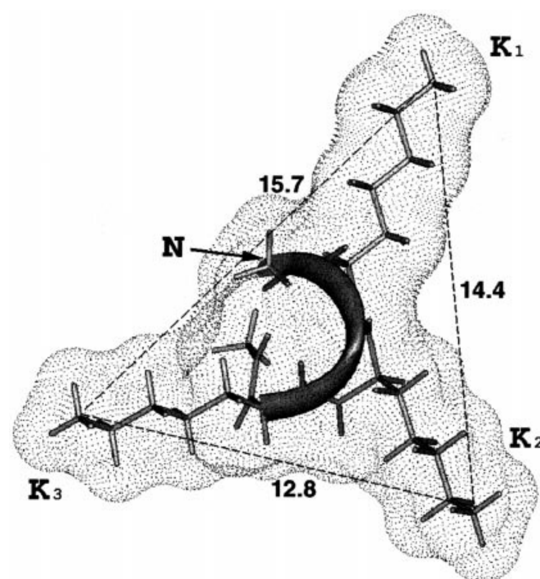

FIGURE 2 Lowest energy conformation (E = −130.73 kcal/mol) for the sequence KKK obtained from a conformational search. A ribbon is used to trace the backbone of the oligopeptide, and the amino terminus of the chain is indicated with the letter N. Dashed lines are used to show the distances (in Å) between the NH_3^+ groups of the LYS residues. The solvent-accessible surface (Connolly, 1983a,b) of the molecule is displayed by dots. The conformation corresponds to a complete α-helical turn with a single hydrogen bond between the NH of the carboxyl-terminal group (−NMe) and the CO of the amino-terminal group (Ac−).

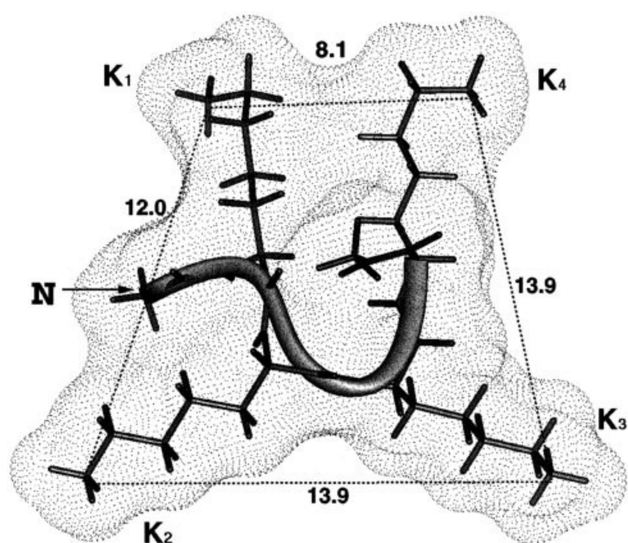

FIGURE 3 Conformation with the second-lowest energy ($E = -166.46$ kcal/mol) found for the sequence KKKK. A ribbon is used to trace the backbone of the oligopeptide, and the amino terminus of the chain is indicated with the letter N. This conformation contains three residues in the A region (Zimmerman et al., 1977) of the ϕ-ψ map forming an α-helical turn and a single hydrogen bond between the NH of the carboxyl-terminal group (−NMe) and the CO of residue LYS-1. Strong hydrophobic interactions and good solvent polarization effects bring the LYS side chains of residues LYS-1 and LYS-4 together. To achieve this close packing, the backbone of LYS-1 is forced out of the A region into the A* region. The solvent-accessible surface (Connolly, 1983a,b) is displayed by dots to indicate the compactness of the conformation. The NH_3^+ groups at the tips of the side chains show good solvent exposure. The interactions among LYS side chains placed the NH_3^+ groups at distances between closest neighbors in space within the range of 12–14 Å in three of four cases; the NH_3^+ groups of LYS-1 and LYS-4 are at a distance of ~8 Å. The conformation is 0.12 kcal/mol higher in energy than the full α-helix, which is the lowest-energy conformation found for the sequence.

Å). Such favorable close approach of charged amino groups in water, due to polarization of the intervening water, has been demonstrated by ab initio quantum mechanical calculations (Cho et al., 1998) and by observations (Magalhaes et al., 1994) of proteins in the Protein Data Bank (Bernstein et al., 1977).

For six LYS residues, the Zimmerman et al. (1997) codes for residues LYS-1 to LYS-5 are in the A region of the ϕ-ψ map; i.e., the lowest-energy conformation is mostly α-helical, as shown in Table 2. The Boltzmann-averaged values for the vicinal coupling constants, $\langle {}^3\mathbf{J}_{HN\alpha} \rangle$, indicate the presence of other conformers close in energy with LYS-1 in conformations other than the A region.

Analysis of the lowest-energy conformation for the KKKKKK sequence shows that arrangements of the lysine side chains, similar to those found for the sequence with four lysine residues, are the result of optimization of both hydrophobic and charge-charge solvent-mediated interactions. The packing between the lysine side chains of residues 1–5 and 2–6 can be seen in Fig. 4. It is worth noting that residues 3 and 4 remain in the α-helical region with the side chains well exposed to the solvent, as shown in Fig. 4. This conformational preference of the six-residue sequence is associated with a large ΔpK_a of all the ionizable groups (in the range of 1.5–2.5 pK units). Even though the ΔpK_a values are not very accurate because the calculations at pH 2 were carried out by assuming zero salt concentration, nevertheless, these values indicate repulsive charge-charge interactions among the ionizable groups that are counterbalanced by increasing the hydrophobic packing in conformations other than the α-helix. It is important to point out that we have not found large shifts in the pK_a values of the lysine residues in the lowest-energy conformations from runs with shorter sequences of lysine residues. The fact that conformations other than the α-helix are the most probable for the sequence KKKKKK is consistent with experimental observations showing that poly-lysine is not helical at this pH (Applequist and Doty, 1962).

As already mentioned, the two-lysine oligopeptide has a tendency to adopt conformations compatible with the α-helical structure (both left- and right-handed). The sequence KKK, shown in Table 2, behaves similarly. This behavior is altered by inclusion of one alanine residue between two

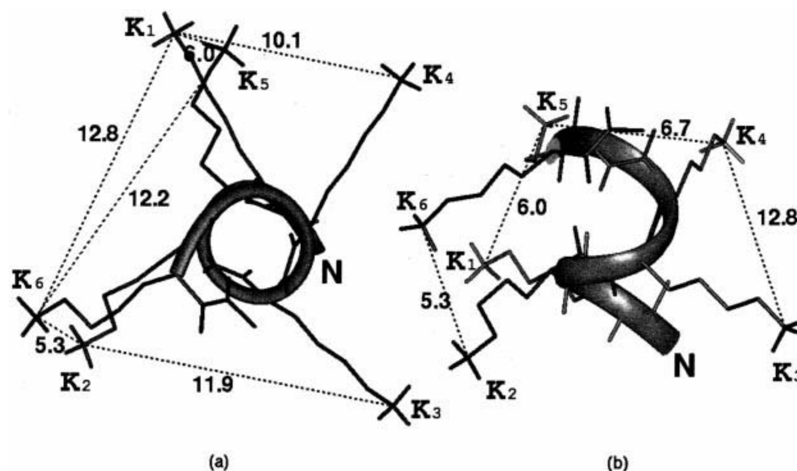

FIGURE 4 Top (*a*) and side (*b*) views of the lowest-energy conformation for the sequence KKKKKK. A ribbon is used to trace the backbone of the oligopeptide, and the amino terminus of the chain is indicated with the letter N. This conformation has an 83% helix content (5/6 of all residues are in the A region (Zimmerman et al., 1977) of the ϕ-ψ map). The interactions among LYS side chains position the NH_3^+ groups at distances within the ranges 9.0–13.0 Å and 5.3–7.0 Å.

lysine residues (sequence KAK, shown in Table 2). In the latter case, only the right-handed α-helical conformations is preferred. Inclusion of four alanine residues between two lysine residues (see results for the sequence KAAAAK in Table 2) shows reduced α-helix content compared with the observations for the KAK, KKKK, and KKKKKK sequences. The analysis of the pattern KAAAAK is important for three reasons. First, this pattern appears in the sequences identified as 3K(I), 3K(II), and 4K in the studies by Marqusee et al. (1989) on short 16-residue alanine-based peptides. Second, a similar pattern has been used in more than 50 lysine-containing peptides in studies by Chakrabartty et al. (1994) aimed at determining helix propensities of the naturally occurring amino acids. Furthermore, this pattern is present in synthetic peptides used to determine the amino- and carboxyl-capping preferences of all 20 amino acids (Doig and Baldwin, 1995). And third, based on experimental studies of a series of alanine-rich peptides containing a single lysine residue (conjugates Ac-Hel_1-(ALA_n-LYS-ALA_m)-NH_2), Groebke et al. (1996) proposed a plausible structural model for the stabilization of the α-helical conformation by the lysine residue.

Among all the accepted conformations obtained in our simulations for the sequence KAAAAK, there are some that correspond to full α-helical structures. A graphical inspection of these α-helical conformations shows at least one, displayed in Fig. 5, in which one of the lysine residues (K_6) interacts with the α-helical backbone in a pattern that quite closely resembles the one described by Groebke et al. (1996). However, this conformation, having a total energy $E = -139.39$ kcal/mol, is not the lowest energy found. In fact, the lowest-energy conformation ($E = -143.16$ kcal/mol) for the sequence KAAAAK, shown in Fig. 6, is not fully α-helical (see Zimmerman et al. (1977) code in Table 2). Perhaps the most important observation is the tendency of lysine side chains to adopt packed conformations with optimal hydrophobic and charge-charge solvent-mediated interactions, as observed for the other oligopeptides previously described. The interactions between the two lysine residues provide the extra stabilization energy that makes this conformation preferred over the full α-helix.

The Boltzmann-averaged values of the vicinal coupling constants, $\langle {}^3\mathbf{J}_{HN\alpha} \rangle$, for the residues of the sequence KAAAAK (Table 2) indicate that four of six residues are in the α-helical region of the ϕ-ψ map. Right-handed α-helix contents of 67% are obtained for both F_A and F_{Jcoup}, using the definitions given by Eq. 3. These results are consistent with a weak helix-forming tendency of L-alanine (Vila et al., 1992).

Runs for the homo-oligopeptides of alanine show a lesser tendency to adopt α-helical conformations than those sequences in which lysine was included. As shown in Table 2, the lowest-energy conformation for the sequence AAAAAA has only 33% and 0% helix content, using the definitions for

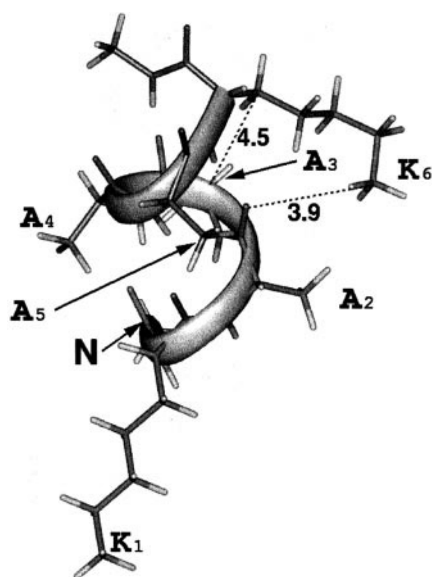

FIGURE 5 An α-helical conformation with relatively low energy ($E = -139.39$ kcal/mol) encountered in the conformational search for the sequence KA_4K, with the amino terminus of the chain indicated with the letter N. A ribbon is used to trace the backbone of the oligopeptide. Residue LYS-6 interacts with the α-helical backbone in the manner suggested by Groebke et al. (1996) based on NOEs observed in ^{1}H NMR spectra of selectively deuterated analogs of Ac-Hel_1-(ALA_n-LYS-ALA_m)-NH_2. The distances 4.5 and 3.9 Å between C^β of LYS-6 and C^α of ALA-3 (*a*) and N^ζ of LYS-6 and the carbonyl O of ALA-2 (*b*), respectively, are indicated.

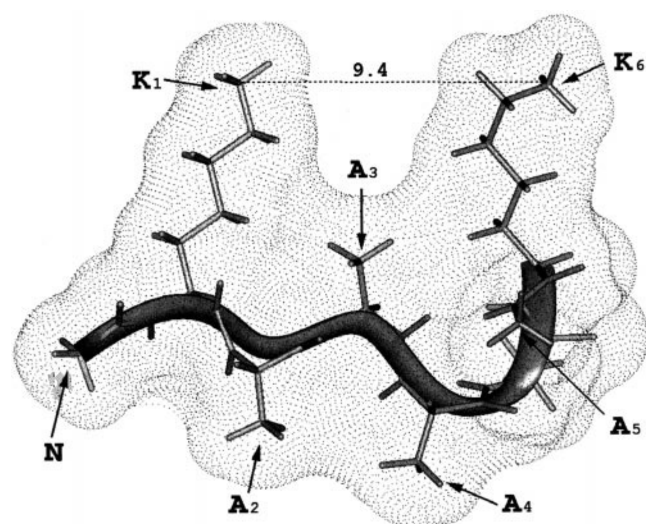

FIGURE 6 Lowest-energy conformation ($E = -143.16$ kcal/mol) encountered in the EDMC run for the sequence KA_4K. This sequence is the repeating unit in the peptide 3K(I) of Marqusse et al. (1989). The backbone of the oligopeptide is traced with a ribbon, and the amino terminus of the chain is indicated with the letter N. The solvent-accessible surface (Connolly, 1983a,b) is displayed by dots to indicate the compactness of the conformation. The distance between Nζ of LYS-1 and Nζ of LYS-6 is 9.4 Å. The conformation has a helix content of 67% using the definition for F_A (or F_{Jcoup}) given by Eq. 3. The characteristic pattern of side-chain/side-chain interactions between LYS residues is again observable. The LYS side chains show a preference for packed conformations instead of being extended and completely exposed to the solvent.

F_A and F_{Jcoup}, respectively. For the sequence AAA, on the other hand, the lowest-energy conformation is fully α-helical, but a significant number of conformers similar in energy but with residues in non-α-helical regions are present. These conformers contribute to produce a helix content, F_{Jcoup}, of zero for the AAA sequence.

Table 1 shows that runs for the alanine homo-oligopeptides were terminated after 1000 accepted conformations whereas the runs for oligopeptides containing lysine residues show a smaller number of accepted conformations. The reason for these differences is that the former runs required considerably less CPU time. For example, the CPU requirements for one of the runs with the sequence KKK is more than 100 times that of the run for the sequence AAA, whereas the former run led to five times fewer accepted conformations.

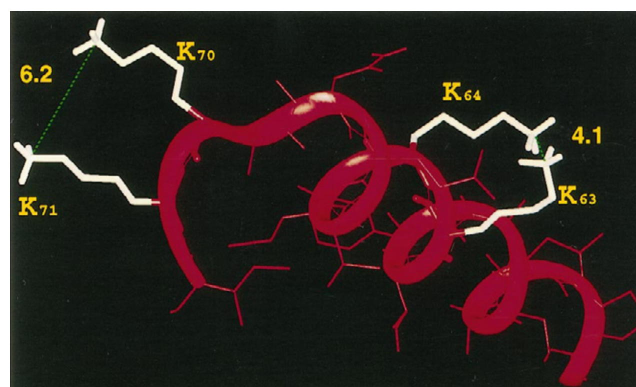

FIGURE 7 Fragment from the x-ray crystallographic structure of staphylococcal nuclease (Wynn et al., 1997), containing two interacting pairs of lysine residues. The backbone of the fragment, involving residues G55 to I72, is highlighted with a red ribbon. The side chains of the lysine residues are shown in white. In close resemblance with the conformations encountered in our simulations, the aliphatic portions of the lysine side chains form strong hydrophobic contacts bringing the Nζ atoms of neighboring side chains at distances of 4.1 and 6.2 Å for the K63–K64 and K70–K71 pairs, respectively. In particular, the pair K70–K71 is well exposed to the solvent, and all the side-chain heavy atoms have low temperature factors. The x-ray crystallographic structure of staphylococcal nuclease shows 11 pairs of lysine residues with distances less than 10 Å between the NH_3^+ groups.

DISCUSSION

The novel and distinctive effects observed in our simulations are the following. First, the LYS side chains in the lowest-energy conformations for all these sequences show a clear tendency to pack among themselves. As the conformations are generated through energy minimization with ECEPP/3 (or ECEPP/3 plus a surface solvation term), the nonbonded interactions (that contribute largely to E_{int} ($\mathbf{r}_p$)) play a major role in defining the side-chain packing. The lysine side-chain arrangements show both good hydrophobic contact and good solvated conformations for the NH_3^+ groups. Typical distances between NH_3^+ groups fall into the two ranges: 5.5–8.0 Å and 9.5–13 Å.

Second, the low-energy conformations of these oligopeptides consistently show a preference for the lysine side chains to interact among themselves rather than adopt conformations far from each other, as is commonly assumed. Although this could be seen as an artifact of the simulations, we note that there are unusual arginine-arginine contacts (Magalhaes et al., 1994) and also lysine-lysine contacts, similar to the ones observed for lysine residues in our simulations, in many protein molecules from the Protein Data Bank (Bernstein et al., 1977). Calculations carried out by Magalhaes et al. (1994) on the guanidinium/guanidinium ion pair and by Cho, K.-W., K. T. No, and H. A. Scheraga (unpublished) on pairs of ionized methylamines reveal that a possible explanation for these arrangements is the bridging role of the water molecules that keeps the positively charged groups close to each other. Fig. 7 illustrates one example of lysine-lysine pairs in a protein, similar to the pair arrangements found in the simulations presented here.

Third, short alanine sequences containing three or six residues do not show a strong preference for α-helical conformations (as seen in Table 2), when compared with the behavior of lysine oligopeptides, in complete agreement with our previous experimental (Ingwall et al., 1968) and theoretical (Vila et al., 1992) results. Although a β_1-α_4 interaction free energy (a hydrophobic bond formed between the C^β of the ith residue and the C^α of the $(i + 3)$th residue, with i increasing toward the amino terminus) of −0.3 kcal/mol exists in a poly(L-alanine) α-helix (Némethy and Scheraga, 1962), the strength of this interaction overcomes the non-hydrogen-bonded defects at the ends of an α-helix only when the helix becomes sufficiently long (Bixon et al., 1963); this is consistent with the absence of helix content in a 10-residue poly(L-alanine) chain (Ingwall et al., 1968) but the presence of helix to the extent of ~75% (at 0°C) when the chain length attains a value of 160 (Ingwall et al., 1968).

Fourth, for the pattern KAAAAK included in the 3K(I) peptide of Marqusee et al. (1989), only a fraction of the chain remains α-helical. Thus, according to our calculations, inclusion of alanine residues does not appear to stabilize the helical conformation in short oligopeptides. An outstanding feature of the lowest-energy conformation for the sequence KAAAAK is the arrangement of a pairwise interaction between the LYS-LYS side chains similar to the one described in the first point above (see Fig. 6).

It is worth mentioning with respect to the fourth point above that studies of conjugates Ac-Hel$_1$-(ALA$_n$-LYS-ALA$_m$)-NH$_2$ (with n = 2–5 and m = 1–7) and analogs in aqueous solution by Groebke et al. (1996) show that the intrinsic helix-forming tendency of L-alanine is close to 1.0 (in the range of 1.01–1.02) at 25°C. Besides being in close agreement with the intrinsic helix-forming tendency of L-alanine proposed by Scheraga and co-workers (Platzer et al., 1972; Ingwall et al., 1968) and with model predictions by Vila et al. (1992), the results of Groebke et al. (1996) also indicate a significant population of packed conformations

compatible with interactions between lysine and the helix backbone; i.e., the CH_2 groups of the side chain of lysine (at position *i*) interact with the α-CH of alanine at position (*i*–3), and the protons of the NH_3^+ group of lysine interact with the carbonyl oxygen of residue (*i*–4), with *i* increasing toward the α-carboxyl terminus. Based on these results, Groebke et al. (1996) concluded that "helix stability is controlled primarily by interactions of the lysine side chain with the helix barrel, and only passively by the alanine matrix." The type of interaction described by Groebke et al. (1996) was observed in our simulations for the sequence KAAAAK (Fig. 5). However, our lowest-energy conformation for KAAAAK is not fully helical. The presence of a second LYS residue in KAAAAK led to other conformations with close packing between the lysine side chains, resulting in lower energy. This side-chain packing seems to be due to both strong hydrophobic interactions and optimal solvent polarization effects (Fig. 6). A theoretical investigation of the behavior of alanine-based sequences with a single ionizable residue, i.e., lysine, glutamic acid, arginine, etc., such as the one studied by Groebke et al. (1996), is currently in progress.

With regard to the interaction of lysine with the helix backbone, although quite important, this may not be the only explanation for the unusual helix formation in the short alanine-based peptides studied by Marqusee et al. (1989), as substitution of three LYS residues in the sequence of 3K(I) by GLU residues (with the sequence identified as 3E) leads to an oligopeptide with almost the same helix content as 3K(I). However, as the side chain of a GLU residue is shorter than that of a LYS residue, the stabilization effects due to side-chain/helix/backbone interactions are expected to be smaller.

Our discrimination between the conformational preferences for alanine and lysine residues leads us to infer that the unusual high helix content observed by Marqusee et al. (1989) in short alanine-based peptides is due, mainly, to solvent polarization effects and possibly hydrophobic side-chain/side-chain interactions between the lysine residues.

SUMMARY

In sequences containing less than seven lysine residues, two or more lysine side chains seem to prefer conformations in which they interact with each other, optimizing hydrophobic interactions as well as the solvent polarization effects due to the ionizable groups. These interactions force these oligopeptides to adopt specific conformational patterns; in particular, the α-helix is among the most commonly (low-energy) found. As a consequence, a significant reduction in the number of allowed conformations appears to take place when charged groups are included in the sequences. On the other hand, sequences that include alanine residues show a weaker tendency to adopt α-helical conformations. These results are in agreement with experimental evidence (Ingwall et al., 1968; Wojcik et al., 1990; Kemp et al., 1991a,b, 1995, 1996; Groebke et al., 1996) showing that alanine is not a strong helix former, contrary to the conclusions of Marqusee et al. (1989). The source of the unusual stability of the helical conformation in short 16-residue alanine-based peptides seems to be associated with the conformational preference of ion pairs of charged lysine residues introduced into the sequences to make them water soluble. The stability of the ion pairs of charged lysines is due, mainly, to hydrophobic interactions and solvent polarization effects that contribute with low free energy of solvation to bring the lysine side chains close together.

We thank F. Aversa for help with searches of the PDB and A. Liwo and K. Gibson for helpful discussion.

This research was supported by grants from the National Institutes of Health (GM-14312 and 1 R03 TW00857-01), the National Institutes of Health National Center for Research Resources (P41RR-04293 and RR-08102), the National Science Foundation (MCB95-13167), the Agencia Nacional de Promoción Científica y Tecnológica Argentina (PICT0030), and Project P-328402 of the Universidad Nacional de San Luis. M.E.V. acknowledges support from CONICET. Support was also received from the National Foundation for Cancer Research. 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, and members of the Center's Corporate Partnership Program.

REFERENCES

Ananthanarayanan, V. S., R. H. Andreatta, D. Poland, and H. A. Scheraga. 1971. Helix-coil stability constants for the naturally occurring amino acids in water. III. Glycine parameters from random poly(hydroxybutylglutamine-co-glycine). *Macromolecules.* 4:417–424.

Anfinsen, C. B. 1972. The formation and stabilization of protein structure. *Biochem. J.* 128:737–749.

Applequist, J., and P. Doty. 1962. α-helix formation in poly-ϵ-carbobenzoxy-L-lysine and poly-L-lysine. *In* Polyamino Acids, Polypeptides and Proteins. M. A. Stahmann, editor. University of Wisconsin Press, Madison, WI. 161–177.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The protein data bank: a computer based archival file for molecular structures. *J. Mol. Biol.* 112:535–542.

Bixon, M., H. A. Scheraga, and S. Lifson. 1963. Effect of hydrophobic bonding on the stability of poly-L-alanine helices in water. *Biopolymers.* 1:419–429.

Bradley, E. K., J. F. Thomason, F. E. Cohen, P. A. Kosen, and I. D. Kuntz. 1990. Studies of Synthetic helical peptides using circular dichroism and nuclear magnetic resonance. *J. Mol. Biol.* 215:607–622.

Chakrabartty, A., T. Kortemme, and R. L. Baldwin. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* 3:843–852.

Chou, K. C., G. Némethy, and H. A. Scheraga. 1990. Energetics of interactions of regular structural elements in proteins. *Acc. Chem. Res.* 23:134–141.

Connolly, M. L. 1983a. Solvent-accessible surfaces of proteins and nucleic acids. *Science.* 221:709–713.

Connolly, M. L. 1983b. Analytical molecular surface calculation. *J. Appl. Crystallogr.* 16:548–558.

Creighton, T. E. 1983. Proteins. Freeman, New York. 182–186.

Doig, A. J., and R. L. Baldwin. 1995. N- and C-capping preferences for all 20 amino acids in α-helical peptides. *Protein Sci.* 4:1325–1336.

Dygert, M. K., G. T. Taylor, F. Cardinaux, and H. A. Scheraga. 1976. Helix-coil stability constants for the naturally occurring amino acids in

water. II. Lysine parameters from random poly(hydroxybutylglutamine-co-L-lysine). *Macromolecules.* 9:794–801.

Finkelstein, A. V., A. Y. Badretdinov, and O. B. Ptitsyn. 1991. Physical reasons for secondary structure stability: α-helices in short peptides. *Proteins Struct. Funct. Genet.* 10:287–299.

Finkelstein, A. V. and O. B. Ptitsyn. 1976. A theory of protein molecule self-organization. IV. Helical and irregular local structures of unfolded protein chains. *J. Mol. Biol.* 103:15–24.

Gay, D. M. 1983. Algorithm 611: subroutines for unconstrained minimization using a model/trust-region approach. *ACM Trans. Math. Software.* 9:503–524.

Gibson, K. D. and H. A. Scheraga. 1988. The multiple-minima problem in protein folding. *In* Structure and Expression, Vol. 1: From Proteins to Ribosomes. R. H. Sarma and M. H. Sarma, editors. Adenine Press, Guilderland, NY. 67–94.

Gō, N. and H. A. Scheraga. 1969. Analysis of the contribution of internal vibrations to the statistical weights of equilibrium conformations of macromolecules. *J. Chem. Phys.* 51:4751–4767.

Groebke, K., P. Renold, K. Y. Tsang, T. J. Allen, K. F. McClure, and D. S. Kemp. 1996. Template-nucleated alanine-lysine helices are stabilized by position-dependent interactions between the lysine side chain and the helix barrel. *Proc. Natl. Acad. Sci. U.S.A.* 93:4025–4029.

Ingwall, R. T., H. A. Scheraga, N. Lotan, A. Berger, and E. Katchalski. 1968. Conformational studies of poly-L-alanine in water. *Biopolymers.* 6:331–368.

Karplus, M. 1959. Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.* 30:11–15.

Karplus, M. 1963. Vicinal proton coupling in nuclear magnetic resonance. *J. Am. Chem. Soc.* 85:2870–2871.

Kemp, D. S., T. J. Allen, and S. L. Oslick. 1995. The energetics of helix formation by short templated peptides in aqueous solution. I. Characterization of the reporting helical template Ac-Hel_1. *J. Am. Chem. Soc.* 117:6641–6657.

Kemp, D. S., T. J. Allen, S. L. Oslick, and J. G. Boyd. 1996. The structure and energetics of helix formation by short templated peptides in aqueous solution. II. Characterization of the helical structure of Ac-Hel_1-Ala_6-OH. *J. Am. Chem. Soc.* 118:4240–4248.

Kemp, D. S., J. G. Boyd, and C. C. Muendel. 1991a. The helical *s* constant for alanine in water derived from template-nucleated helices. *Nature.* 352:451–454.

Kemp, D. S., T. P. Curran, J. G. Boyd, and T. J. Allen. 1991b. Studies of N-terminal templates for α-helix formation: synthesis and conformational analysis of peptides conjugates of (2S, 5S, 8S, 11S)-1-acetyl-1,4-diaza-3-keto-5-carboxy-10-thiatricyclo[2.8.1.$0^{4,8}$]-tridecane (Ac-Hel_1-OH). *J. Org. Chem.* 56:6683–6697.

Lyu, P. C., M. I. Liff, L. A. Marky, and N. R. Kallenbach. 1990. Side-chain contributions to the stability of α-helical structure in peptides. *Science.* 250:669–673.

Magalhaes, A., B. Maigret, J. Hoflack, J. N. F. Gomes, and H. A. Scheraga. 1994. Contribution of unusual arginine-arginine short-range interactions to stabilization and recognition in proteins. *J. Protein Chem.* 13: 195–215.

Marqusee, S., V. H. Robbins, and R. L. Baldwin. 1989. Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. U.S.A.* 86:5286–5290.

Matheson, R. R. Jr., and H. A. Scheraga. 1978. A method for predicting nucleation sites for protein folding based on hydrophobic contacts. *Macromolecules.* 11:819–829.

Merutka, G., W. Lipton, W. Shalongo, S. H. Park, and E. Stellwagen. 1990. Effect of central-residue replacements on the helical stability of a monomeric peptide. *Biochemistry.* 29:7511–7515.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1092.

Momany, F. A., R. F. McGuire, A. W. Burgess, and H. A. Scheraga. 1975. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, non-bonded interactions, hydrogen bond interactions and intrinsic torsional potentials for the naturally occurring amino-acids. *J. Phys. Chem.* 79:2361–2381.

Montelione, G. T. and H. A. Scheraga. 1989. Formation of local structures in protein folding. *Ac. Chem. Res.* 22:70–76.

Némethy, G., K. D. Gibson, K. A. Palmer, C. N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, and H. A. Scheraga. 1992. Energy parameters in polypeptides. X. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm with application to proline-containing peptides. *J. Phys. Chem.* 96:6472–6484.

Némethy, G., M. S. Pottle, and H. A. Scheraga. 1983. Energy parameters in polypeptides. IX. Updating of geometrical parameters, nonbonded interactions and hydrogen bond interactions for the naturally occurring amino acids. *J. Phys. Chem.* 87:1883–1887.

Némethy, G., and H. A. Scheraga. 1963. The structure of water and hydrophobic bonding in proteins. III. The thermodynamic properties of hydrophobic bonds in proteins. *J. Phys. Chem.* 66:1773–1789. [Erratum in *J. Phys. Chem.* 67:2888]

O'Donnell, T. J., S. G. Hotovy, M. S. Pottle, D. R. Ripoll, and H. A. Scheraga. 1996. Implementation of ECEPP-based programs on the IBM SP2 and the SGI Power Challenge computers. *In* Lecture Notes in Computer Science, Vol. 1067: High-Performance Computing and Networking. H. Liddell, A. Colbrook, B. Hertzberger, and P. Sloot, editors. Springer-Verlag, Berlin. 365–372.

O'Neil, K. T., and W. F. DeGrado. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino-acids. *Science.* 250:646–651.

Padmanabhan, S., S. Marqusee, T. Ridgeway, T. M. Laue, and R. L. Baldwin. 1990. Relative helix-forming tendencies of nonpolar amino acids. *Nature.* 344:268–270.

Platzer, K. E. B., V. S. Ananthanarayanan, R. H. Andreatta, and H. A. Scheraga. 1972. Helix-coil stability constants for the naturally occurring amino acids in water. IV. Alanine parameters from random poly(hydroxypropyl-glutamine-co-L-alanine). *Macromolecules.* 5:177–187.

Poland, D. C. and H. A. Scheraga. 1965. Statistical mechanics of noncovalent bonds in polyamino acids. *Biopolymers.* 3:275–419, 593.

Poland, D. C. and H. A. Scheraga. 1970. Theory of Helix-Coil Transitions in Biopolymers. Academic Press, New York.

Presta, L. G. and G. D. Rose. 1988. Helix signals in proteins. *Science.* 240:1632–1641.

Richardson, J. S. 1976. Handedness of crossover connections in β-sheets. *Proc. Natl. Acad. Sci. U.S.A.* 73:2619–2623.

Ripoll, D. R., A. Liwo, and H. A. Scheraga. 1998. New developments of the electrostatically driven Monte Carlo method: test on the membrane-bound portion of melittin. *Biopolymers.* 46:117–126.

Ripoll, D. R. and H. A. Scheraga. 1988. On the multiple-minima problem in the conformational analysis of polypeptides. II. An electrostatically driven Monte Carlo method: tests on poly (L-alanine). *Biopolymers.* 27:1283–1303.

Ripoll, D. R. and H. A. Scheraga. 1989. On the multiple-minima problem in the conformational analysis of polypeptides. III. An electrostatically driven Monte Carlo method: tests on enkephalin. *J. Protein Chem.* 8:263–287.

Ripoll, D. R., Y. N. Vorobjev, A. Liwo, J. A. Vila, and H. A. Scheraga. 1996. Coupling between folding and ionization equilibria: effects of pH on the conformational preferences of polypeptides. *J. Mol. Biol.* 264: 770–783.

Simonson, T., and A. T. Brünger. 1994. Solvation free energies estimated from macroscopic continuum theory: an accuracy assessment. *J. Phys. Chem.* 98:4683–4694.

Sippl, M. J., G. Némethy, and H. A. Scheraga. 1984. Intermolecular potentials from crystal data. VI. Determination of empirical potentials for O-H. . .O-C hydrogen bonds from packing configurations. *J. Phys. Chem.* 88:6231–6233.

Sitkoff, D., K. A. Sharp, and B. Honig. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* 98:1978–1988.

Vásquez, M., G. Némethy, and H. A. Scheraga. 1994. Conformational energy calculations on polypeptides and proteins. *Chem. Rev.* 94: 2183–2239.

Vila, J., R. L. Williams, J. A. Grant, J. Wojcik, and H. A. Scheraga. 1992. The intrinsic helix-forming tendency of L-alanine. *Proc. Natl. Acad. Sci. U.S.A.* 89:7821–7825.

Vila, J., R. L. Williams, M. Vásquez, and H. A. Scheraga. 1991. Empirical solvation models can be used to differentiate native from near-native

conformations of bovine pancreatic trypsin inhibitor. *Proteins Struct. Funct. Genet.* 10:199–218.

Von Dreele, P. H., N. Lotan, V. S. Ananthanarayanan, R. H. Andreatta, D. Poland, and H. A. Scheraga. 1971b. Helix-coil stability constants for the naturally occurring amino acids in water. II. Characterization of the host polymers and application of the host-guest technique to random poly-(hydroxypropylglutamine-co-hydroxybutyl-glutamine). *Macromolecules.* 4:408–417.

Von Dreele, P. H., D. Poland, and H. A. Scheraga. 1971a. Helix-coil stability constants for the naturally occurring amino acids in water. I. Properties of copolymers and approximate theories. *Macromolecules.* 4:396–407.

Vorobjev, Y. N., and H. A. Scheraga. 1997. A fast adaptive multigrid boundary element method for macromolecular electrostatic computations in a solvent. *J. Comput. Chem.* 18:569–583.

Wilmot, C. M., and J. M. Thornton. 1990. β-turns and their distortions: a proposed new nomenclature. *Protein Eng.* 3:479–493.

Wojcik, J., K. H. Altmann, and H. A. Scheraga. 1990. Helix-coil stability constants for the naturally occurring amino acids in water. XXIV. Half-cysteine parameters from random poly (hydroxybutylglutamine-co-S-methylthio-L-cysteine). *Biopolymers.* 30:121–134.

Wüthrich, K. 1986. NMR of Proteins and Nucleic Acids. John Wiley and Sons, New York. 166–168.

Wynn, R., P. C. Harkins, F. M. Richards, and R. O. Fox. 1997. Comparison of straight chain and cyclic unnatural amino acids embedded in the core of staphylococcal nuclease. *Protein Sci.* 6:1621–1626.

Zimm, B. H. and J. K. Bragg. 1959. Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* 31: 526–535.

Zimmerman, S. S., M. S. Pottle, G. Némethy, and H. A. Scheraga. 1977. Conformational analysis of the 20 naturally occurring amino acid residues using ECEPP. *Macromolecules.* 10:1–9.

Zimmerman, S. S. and H. A. Scheraga. 1977a. Local interactions in bends of proteins. *Proc. Natl. Acad. Sci. U.S.A.* 74:4126–4129.

Zimmerman, S. S. and H. A. Scheraga. 1977b. Influence of local interactions on protein structure. I. Conformational energy studies of *N*-acetyl-*N'*-methylamides of Pro-X and X-Pro dipeptides. *Biopolymers.* 16: 811–843. [Erratum in *Biopolymers* 1997, 16:1385]