PROTEINS: Structure, Function, and Genetics 19:85–97 (1994)

RESEARCH ARTICLES

α-Helix-Forming Propensities in Peptides and Proteins

Trevor P. Creamer and George D. Rose

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

ABSTRACT **Much effort has been invested in seeking to understand the thermodynamic basis of helix stability in both peptides and proteins. Recently, several groups have measured the helix-forming propensities of individual residues (Lyu, P.C., Liff, M.I., Marky, L.A., Kallenbach, N.R. Science 250:669–673, 1990; O'Neil, K.T., DeGrado, W.F. Science 250:646–651, 1990; Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T.M., Baldwin, R.L. Nature (London) 344:268–270, 1990). Using Monte Carlo computer simulations, we tested the hypothesis that these differences in measured helix-forming propensity are due primarily to loss of side chain conformational entropy upon helix formation (Creamer, T.P., Rose, G.D. Proc. Natl. Acad. Sci. U.S.A. 89:5937–5941, 1992). Our previous study employed a rigid helix backbone, which is here generalized to a completely flexible helix model in order to ensure that earlier results were not a methodological artifact. Using this flexible model, side chain rotamer distributions and entropy losses are calculated and shown to agree with those obtained earlier. We note that the side chain conformational entropy calculated for Trp in our previous study was in error; a corrected value is presented. Extending earlier work, calculated entropy losses are found to correlate strongly with recent helix propensity scales derived from substitutions made within protein helices (Horovitz, A., Matthews, J.M., Fersht, A.R. J. Mol. Biol. 227:560–568, 1992; Blaber, M., Zhang, X.-J., Matthews, B.M. Science 260:1637–1640, 1993). In contrast, little correlation is found between these helix propensity scales and the accessible surface area buried upon formation of a model polyalanyl α-helix. Taken in sum, our results indicate that loss of side chain entropy is a major determinant of the helix-forming tendency of residues in both peptide and protein helices.**

Key words: protein conformation, secondary structure, protein folding, helix stability, helix formation, conformational entropy

INTRODUCTION

The α-helix was first introduced by Pauling et al.[1] and has been studied extensively ever since. Studies have included empirical predictions,[2] experimental "host–guest" measurements,[3] and theoretical approaches.[4] Slightly more than a decade ago, Bierzynski et al.,[5] expanding upon earlier work of Brown and Klee,[6] demonstrated the existence of short stable helices in water. Following this observation, several groups employed isolated helical peptides as models of helix formation. In such studies, a heterogeneous, synthetic peptide, known to form a stable helix in solution, is used as a host system. Guest residues are then introduced systematically into this host and their effect on helix stability measured. In this way, Padmanabhan et al.[7] determined the helix-forming tendencies of five naturally occurring residues, using circular dichroism (CD) spectroscopy of an alanine-based host in water. Lyu et al[8] used CD measurements together with a Zimm–Bragg[4]-based theoretical treatment to derive the free energy contribution to helix formation for 10 naturally occurring residues in a highly polar host peptide. Both groups also measured the helix-forming tendencies of several nonnatural residues.[9,10] O'Neil and DeGrado[11] obtained the free energy contributions to helix formation for all 20 naturally occurring residues by measuring changes in the helix–coil equilibrium constant upon substi-

Received September 10, 1993; revision accepted February 2, 1994.

Address reprint requests to George D. Rose, Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, 660 S. Euclid Avenue, St. Louis, MO 63110.

tution of guest residues within a coiled-coil peptide host.

These three scales of helix-forming tendencies yield similar though not identical results. Differences are probably due to particulars of the individual host systems (e.g., interactions between side chains). All three scales differ markedly from a scale obtained in earlier host–guest studies employing a host peptide comprised of nonnatural residues.[3]

It has been suggested by the authors of all three recent scales[7,8,11] that differences in the experimentally observed helix-forming tendencies are due, at least in part, to the loss of side chain conformational entropy that accompanies helix formation. This entropic contribution arises from unfavorable steric interactions between the side chain and the helix backbone. We tested this hypothesis in our own host–guest studies using Monte Carlo computer simulations[12] and found that side chain conformational entropy loss is indeed a major determinant in the observed helix-forming tendencies of the hydrophobic residues (Ala, Val, Ile, Leu, Met, Phe, Tyr, and Trp). Nevertheless, corresponding enthalpic contributions to the observed helix propensities remain a possibility.

The conformational entropy of a side chain measures the rotational freedom of the side chain. In turn, the rotational freedom can be estimated from side chain rotamer distributions,[12,13] which are known to be dependent upon backbone conformation.[14,15]

In our earlier study,[12] we used differences in the rotamer distributions between the helical state and a tripeptide standard state to calculate the side chain conformational entropy loss that accompanies helix formation. In that work, a polyalanyl host helix was modeled as a rigid body, with backbone atoms fixed in space and only the side chain atoms of the guest residue allowed to move. This rigid model is overly constrained; some side chain rotamers that are observed occasionally in X-ray elucidated protein structures are eliminated altogether in the model. To relax the rigid model, the van der Waals radii of all –CH, $-CH_2$, and $-CH_3$ groups were scaled to 90% of their original values. After scaling, excellent agreement was obtained between observed and simulated rotamer populations.

This present study is a generalization of our previous work using a rigid helix model.[12] Here, a flexible helix model is developed that can reproduce the full range of backbone conformational behavior observed in protein helices and is suitable for both Monte Carlo and molecular dynamics simulations. The flexible model requires no van der Waals scaling approximations to be made. Using this model, all previous side chain conformational entropy simulations were repeated, and the resulting side chain rotamer distributions and entropy losses shown to be in good agreement with those obtained in our earlier study. Both models support the conclusion that loss of side chain entropy is a major determinant of the helix-forming tendency of hydrophobic residues.

Results obtained using the flexible helix model not only underscore our previous conclusion,[12] but also show that a simple rigid body model is a sufficient representation of the helix backbone for studies of side chain conformational behavior. Importantly, the use of a rigid helix backbone reduces computation time by an order of magnitude.

It has been suggested by Blaber et al.[16,17] that burial of hydrophobic surface area makes a significant favorable contribution to the measured helix-forming propensities. To test this assertion, we computed the average accessible surface area buried by side chains upon α-helix formation (from the standard state). For the hydrophobic side chains, no discernible correlation between side chain surface area buried and helix-forming propensity is found. This observation suggests that the burial of hydrophobic surface area is *not* a determining factor in the helix-forming propensities of peptide models. However, hydrophobic interactions between side chains can and do make significant contributions to the stability of helices within proteins.

Another aspect of the present study addresses the issue of unwanted correlation between adjacent pairs of conformations generated in computer simulations, a matter of crucial concern.[18,19] Typically, conformations generated in a simulation are not independent of one another, but rather are strongly correlated. To derive valid statistics from a simulation, such correlation should be minimized. Steps taken to accomplish this objective are described below.

METHODS

Side Chain Conformational Entropy Calculations

The change in conformational entropy, ΔS, for a side chain upon helix formation is the difference between the entropy in some standard state, S^0, and the entropy in an α-helix, S^α, i.e.

$$\Delta S = S^\alpha - S^0. \quad (1)$$

The conformational entropy in any given state is reckoned from the distribution of side chain rotamers using

$$S = -R\sum_i p_i \ln p_i \quad (2)$$

where p_i is the probability of the side chain being in rotamer class i, the sum being taken over all rotamer classes[12]; R is the gas constant. Rotamer distributions are generated from Monte Carlo simulations.

To determine the probability of a side chain having a given conformation, the conformational space

of the side chain must be exhaustively partitioned into rotamer classes. As defined in our previous study,[12] the rotamer dihedral angle limits used are (1) χ angles involving rotation about a bond between two tetrahedral carbons, or a tetrahedral carbon and the sulfur of Met, are partitioned into three preferred rotamer positions: *trans* ($\chi > 120°$ or $\chi \leq -120°$), *gauche+* ($0° < \chi \leq 120°$), and *gauche−* ($-120° < \chi \leq 0°$). (2) χ angles involving rotation about a bond between a tetrahedral carbon and trigonal carbon are partitioned into four preferred rotamer positions: *trans* ($\chi > 135°$ or $\chi \leq -135°$), $+90°$ ($45° < \chi \leq 135°$), $0°$ ($-45° < \chi \leq 45°$), and $-90°$ ($-135° < \chi \leq -45°$). Following the IUPAC-IUB convention,[20] *trans* corresponds to $\chi = 180°$ and clockwise rotations are considered positive.

The rotation of the Tyr hydroxyl hydrogen about the Cζ to Oη bond is treated in the same manner as rotation about a tetrahedral carbon to trigonal carbon bond. This convention differs from that used in our previous study where this rotation was treated as a tetrahedral carbon to tetrahedral carbon bond. Changing the partitioning of a rotamer in this way alters the value of side chain conformational entropy [S in Eq. (2)], but the change in entropy [ΔS in Eq. (1)] is unaffected.

To illustrate the rotamer classes, consider the side chain of a Leu residue: side chain rotation can occur at either or both of two bonds, each of the tetrahedral carbon to tetrahedral carbon type. With three rotamer positions for each bond, there are nine possible rotamer classes for Leu. Phe also has two bonds about which rotation may occur. In this case, one bond is of the tetrahedral carbon to tetrahedral carbon type, the other of the tetrahedral carbon to trigonal carbon type. Hence, there are three rotational positions for the first bond and four for the second, a total of 12 possible classes.

Monte Carlo Simulations

The conformational properties of peptides, including side chain rotamer distributions, are obtained readily from Monte Carlo simulations using Metropolis sampling.[21] The Metropolis algorithm is a widely used method of generating configurations of a molecular system from the Boltzmann distribution. In brief, the algorithm involves randomly perturbing a system from some initial configuration in order to generate a new trial configuration. The difference in energy, ΔE, between the trial and initial configurations is then calculated. If (1) $\Delta E < 0$ (the trial configuration has lower energy than the initial) or (2) $\Delta E > 0$ and a normal random deviate between 0 and 1 is less than the Boltzmann-weighted energy, $e^{-\Delta E/RT}$, the trial configuration is accepted and becomes the new initial configuration. If neither condition is satisfied, the trial configuration is rejected, and the initial configuration is retained. This scheme is repeated iteratively with data being collected at the end of each cycle, regardless of whether a new configuration was accepted or rejected. The resulting distribution of configurations will satisfy the Boltzmann distribution.

The atomic interaction energies required by the Metropolis algorithm were calculated using the AMBER/OPLS forcefield.[22,23] This forcefield employs the united atom approximation; polar hydrogens are included explicitly; nonpolar hydrogens are treated by increasing the radii of the heavy atoms to which they are bonded.

In this work, α-helices were modeled as acetylated (Ace) and methyl-amidated (NMe) peptides of sequence Ace-$(Ala)_n$-NMe, where n is the total number of residues. For simulations testing helix constraints (see below), polyalanine helices with $n = 12$ were employed. For side chain entropy calculations, helices of length $n = 11$ were used, with the central residue, Xaa, as the guest (Xaa = Val, Ile, Leu, Met, Phe, Tyr, or Trp).

In simulations of the helical state, all atoms were allowed to undergo small random displacements ($\leq$ 0.005 Å). Side chain conformations were generated by a combination of small atomic displacements and rotation about randomly selected side chain χ angles by a randomly selected amount ($\leq 120°$). Rotations were not attempted about the helical backbone dihedral angles. This algorithm is similar to our earlier formulation,[12] except that the random angle bends used previously were found to be unnecessary and were therefore eliminated.

The standard state used in the entropy calculations was modeled as an acetylated and methyl-amidated peptide of sequence Ace-Ala-Xaa-Ala-NMe, with Xaa as the guest residue whose entropy is to be calculated. Both the backbone and side chain in the tripeptide are flexible and can undergo small atom displacements ($\leq$ 0.005 Å) and random torsion rotations ($\leq 120°$).

Each simulation modeled a single isolated peptide. Water was treated as a dielectric continuum. In simulations of flexible polyalanine helices used to test the helix conformational constraints, two dielectrics were used: $\epsilon = 2$ and 78. In the simulations used to calculate side chain conformational entropy loss, the dielectric of bulk water ($\epsilon = 78$) was employed. Since the residues being studied in this work are all hydrophobic, the choice of dielectric will have no effect on the side chains. The system temperature was kept at 298K. In all simulations, an equilibration period of 50,000 Monte Carlo iterations was used, with data collection commencing after this interval.

α-Helix Conformational Constraints

Hydrogen bond constraints were imposed to prevent α-helices from unfolding. These constraints maintain reasonable hydrogen bond length by applying a forcing potential when the bond length ex-

ceeds some threshold value. The constraint potential used, adapted from the work of Yun and Hermans,[24] is of the form

$$U = \begin{cases} 0 & d<d_U \\ U_{\text{Hbond}}(3-2\zeta)\zeta^2,\ \zeta = |(d-d_U)|/d_R & d_U<d<d_U+d_R \\ U_{\text{Hbond}} & d_U+d_R<d \end{cases} \quad (3)$$

where d is the distance between the amide hydrogen and carbonyl oxygen atoms that is being constrained. U_{Hbond} is the applied potential, d_U is the upper distance limit for the hydrogen bond, and d_R is the distance over which the potential is increased. This potential takes the form of a flat minimum, with no applied energy, when the hydrogen bond donor (amide hydrogen) and acceptor (carbonyl oxygen) are within a specified distance, $d < d_U$, of each other. If the donor–acceptor distance exceeds this limit, a potential "ramp" is applied that increases smoothly over a specified distance, d_R, after which a fixed potential U_{Hbond} is applied. The hydrogen bond constraint parameters used are: $U_{\text{Hbond}} = 15.0$ kcal mol^{-1}, $d_U = 1.4$ Å, and $d_R = 4.0$ Å. These values of the parameters were found to lead to good agreement with experimentally observed helix hydrogen bond lengths (see Results).

To maintain helical values of the dihedral angles, ϕ, ψ, in the helix models, a dihedral constraint potential is applied to each ϕ and ψ torsion in the helix.[24] This potential is similar in form to the hydrogen bond potential above, viz.

$$U = \begin{cases} U_{\text{Dihedral}} & \theta < \theta_L - \theta_R \\ U_{\text{Dihedral}}(3-2\zeta)\zeta^2,\ \zeta = |(\theta-\theta_L)|/\theta_R & \theta_L - \theta_R < \theta < \theta_L \\ 0 & \theta_L < \theta < \theta_U \\ U_{\text{Dihedral}}(3-2\zeta)\zeta^2,\ \zeta = |(\theta-\theta_U)|/\theta_R & \theta_U < \theta < \theta_U + \theta_R \\ U_{\text{Dihedral}} & \theta_U + \theta_R < \theta \end{cases} \quad (4)$$

where θ is the current dihedral angle, U_{Dihedral} is the applied potential, θ_L is the lower limit of the dihedral angle, and θ_U is the upper limit. θ_R is the dihedral angle range over which the potential is increased. Like the hydrogen bond constraint potential in Eq. (3), the dihedral constraints consist of a flat minimum when the dihedral angle, θ, falls between two set limits ($\theta_L < \theta < \theta_U$). A smoothly increasing potential is applied over a set range, θ_R, when the dihedral angle exceeds those limits, and a constant potential is applied when this range is exceeded. The parameters used are (1) for ϕ dihedrals $U_{\text{Dihedral}} = 15.0$ kcal mol^{-1}, $\theta_L = -60.0°$, $\theta_U = -60.0°$, and $\theta_R = 60.0°$ and (2) for ψ dihedrals $U_{\text{Dihedral}} = 30.0$ kcal mol^{-1}, $\theta_L = -10.0°$, $\theta_U = -60.0°$, and $\theta_R = 10.0°$. For ϕ dihedrals, this choice of parameters forces values near $-60°$ in the range where the potential is applied. The use of these parameters for ϕ and ψ leads to backbone dihedral distributions that closely mimic the distributions observed for protein helices (see Results).

In the helix simulations employing conformational constraints, all backbone hydrogen bonds were constrained using the potential from Eq. (3). In addition, the oxygen of the N-terminal acetyl group and the amide hydrogen of the C-terminal methylamide group are constrained to hydrogen bond to the backbone amide hydrogen of residue 4 and backbone carbonyl oxygen of residue $(n-3)$, respectively. In simulations where backbone dihedral constraints are applied, all backbone ϕ and ψ torsions in the helices were constrained using Eq. (4).

Solvent Accessible Surface Area Calculations

Surface area calculations were performed using the analytical method of Richmond.[25] A probe radius of 1.4 Å was employed.

Data Correlation and Standard Deviations

In a typical simulation (Monte Carlo or molecular dynamics) not every configuration generated is actually retained for analysis. Instead, one of two approaches is taken: either the conformation is sampled once for every n_c configuration generated, or the average over a block of n_c configurations is taken (the so-called "block averages" approach). The former approach was used in this work. Thus, with N_{Total} configurations generated, $N_{\text{Conf.}} = N_{\text{Total}}/n_c$ are retained for analysis.

The ensemble of configurations, N_{Total}, can be strongly correlated because each is generated by a small perturbation of the previous configuration. In a molecular system where the conformation is changing gradually, such as the simulation of a flexible α-helix, consecutive configurations will be inescapably correlated. For this reason, the choice of the number of configurations generated between each sampled conformation, n_c, is of crucial importance. In particular, to derive valid statistics n_c must be large enough to suppress correlation between successive sampled conformations in order. The degree of correlation can be determined by the autocorrelation at lag 1, r_1.[18] For some property, X, of the system, the autocorrelation at lag 1 is

$$r_1 = \frac{\sum_{i=2}^{N_{\text{Conf.}}} (X_i - \langle X\rangle)(X_{i-1} - \langle X\rangle)}{\sum_{i=1}^{N_{\text{Conf.}}} (X_i - \langle X\rangle)^2} \quad (5)$$

where X_i is the value of the property for the ith conformation and $\langle X\rangle$ is the average value over the ensemble $N_{\text{Conf.}}$. The autocorrelation will lie in the range $-1 \leq r_1 \leq 1$, with $r_1 = 0$ indicating no correlation and $r_1 < 0$ indicating negative correlation. Ideally the choice of n_c would push $r_1 \to 0$. In practice, this ideal results in excessively large values of n_c requiring unrealistically long simulations. Kolafa[19] proposed that an autocorrelation of $r_1 < 0.7$ should be adequate. However, in our work a more conservative value of $r_1 < 0.5$ was chosen.

TABLE I. Simulation Length Data and the Autocorrelations at Lag 1 of the Conformation Energies for the Flexible Polyalanine α-Helices Used to Test the Conformational Constraints

	No applied constraints		Hydrogen bond constraints*		Full constraints†	
	$\epsilon = 2$	$\epsilon = 78$	$\epsilon = 2$	$\epsilon = 78$	$\epsilon = 2$	$\epsilon = 78$
N_{Total}‡ ($\times 10^{-6}$)	10	10	10	10	10	20
n_c§	2000	2000	2000	2000	2000	2000
r_1**	0.404	0.342	0.288	0.321	0.262	0.288

*Hydrogen bond constraints only are applied in these simulations.
†Both hydrogen bond and backbone dihedral constraints are applied in these simulations.
‡Total number of configurations generated in the simulation.
§Number of configurations between each stored conformation.
**Autocorrelation at lag 1 of the configuration energy.

When sampled conformations are uncorrelated, it is possible to estimate the standard deviation, σ, for some property of the system, X, from the zeroth-order variance, V_0 (ref. 18) as

$$V_0 = \frac{1}{N_{Conf.} - 1} \sum_{i=1}^{N_{Conf.}} (X_i - \langle X \rangle)^2. \quad (6)$$

The standard deviation in $\langle X \rangle$ is then given by

$$\sigma = \sqrt{V_0}. \quad (7)$$

Our standard deviations are calculated using Eqs. (6) and (7).

RESULTS

α-Helix Conformational Constraints

Helix conformational constraints were tested by simulation of polyalanine helices at two dielectrics of $\epsilon = 2$ and 78. The number of configurations generated and sampled and the autocorrelation at lag 1 of the configuration energy are given in Table I.

Typically, the autocorrelation at lag 1 of the configuration energy is significantly less than 0.5 (Table I). This value ensures that adjacent pairs of stored conformations are, on average, not correlated, enabling standard deviations to be calculated from the zeroth-order variance [Eq. (6)].

The average φ and ψ backbone dihedral angles of the central four helix residues are given in Table II. Also shown are the average backbone amide nitrogen to carbonyl oxygen distances, d_{N-O}, for the central three hydrogen bonding groups. For comparison, the table includes data from simulations of unconstrained helices as well as the corresponding values for α-helices in X-ray elucidated proteins.

Central residues in unconstrained helices simulated at $\epsilon = 2$ maintain reasonably helical geometry over the length of the simulation (Table II). However, the average d_{N-O} is somewhat longer and the average φ and ψ angles differ from those observed in protein helices. In addition, helix ends tend to fray, particularly at the C-terminus; the average φ and ψ dihedrals of the last two residues in the peptide are (−85° ± 14°, −13° ± 18°) and (−107° ± 21°, −32° ± 22°), respectively.

At $\epsilon = 78$, the unconstrained helix is far less helical. Helical geometry deteriorates over the course of the simulation, as assayed by the average d_{N-O} and dihedral angles of the central residues (Table II). At $\epsilon = 2$ the stronger electrostatic forces between backbone amide hydrogens and carbonyl oxygens maintain better helical hydrogen bonds, which, in turn, improve backbone geometry. However, the backbone geometry of the helix simulated at $\epsilon = 2$ still differs markedly from that observed in actual protein helices.

Hydrogen bond distance constraints were proofed in simulations at the two different dielectrics. With hydrogen bond constraints, good agreement is observed between the average simulated d_{N-O} and that observed in protein helices (Table II). These constraints have only a marginal effect upon the geometry of central residues at $\epsilon = 2$; hydrogen bonds are slightly shorter and dihedral angles have slightly smaller standard deviations in comparison with the unconstrained helix. Average backbone dihedral angles shift toward those in protein helices. Further, the applied constraints inhibit fraying because residues near the termini are constrained to make helical hydrogen bonds. At $\epsilon = 78$, the imposition of constraints prevents unfolding, similar to results obtained at $\epsilon = 2$. While the imposition of hydrogen bond constraints succeeds in preventing helices from unfolding, the φ and ψ dihedral angles at both dielectrics still fail to resemble those observed in protein helices.

Imposition of dihedral angle and hydrogen bonding constraints results in closer agreement between average φ and ψ angles obtained in simulations and those observed in protein helices (Table II). The φ dihedrals are forced toward −60° by the applied potential ramps, overcoming the tendency of the AMBER/OPLS forcefield to give average φ angles near −70°. In contrast, the ψ dihedrals simply need boundary "walls" to obtain good agreement with protein data.

TABLE II. The Conformational Results From the Simulations Testing the Hydrogen Bonding and Backbone Dihedral Constraints Using 12 Residue Polyalanine α-Helices

	Protein helices*	No applied constraints		Hydrogen bond constraints†		Full constraints‡	
		$\epsilon = 2$	$\epsilon = 78$	$\epsilon = 2$	$\epsilon = 78$	$\epsilon = 2$	$\epsilon = 78$
ϕ§	−63.8° ± 6.6°	−75.4° ± 10.1°	−94.6° ± 16.6°	−70.3° ± 8.6°	−73.0° ± 9.5°	−62.8° ± 4.4°	−63.2° ± 4.5°
ψ§	−41.0° ± 7.2°	−28.4° ± 11.9°	−14.6° ± 17.0°	−35.7° ± 9.1°	−33.6° ± 9.9°	−43.0° ± 6.2°	−42.8° ± 6.5°
d_{N-O} (Å)**	3.02 ± 0.20	3.39 ± 0.40	3.85 ± 0.50	2.96 ± 0.15	3.09 ± 0.17	2.92 ± 0.13	3.03 ± 0.14

*From Presta and Rose.[26]
†Hydrogen bond constraints only are applied in these simulations.
‡Both hydrogen bond and backbone dihedral constraints are applied in these simulations.
§Averages of the dihedral angles from residues in protein helices and of the middle four residues of simulated helices.
**Protein helix d_{N-O} distances are obtained by averaging over all α-helices in the 26 proteins examined by Presta and Rose.[26] Simulation d_{N-O} distances are obtained by averaging over the central four hydrogen bonds in helices.

The distributions of ϕ and ψ angles obtained in the AMBER/OPLS simulation at $\epsilon = 78$ with both hydrogen bond and dihedral constraints are plotted in Figure 1. Distributions for X-ray elucidated proteins from Presta and Rose[26] are also shown. The close agreement between these distributions is not attainable using hydrogen bond constraints alone (data not shown).

Side Chain Conformational Entropy Loss

The number of configurations generated in each simulation, the number of configurations between each sampled conformation, and the autocorrelation at lag 1 of the configuration energy are given in Table III. The autocorrelations of the conformation energies are all less than 0.5, the target adopted for this work. Not surprisingly, autocorrelations are smaller for standard state peptides than for helices. In simulations of standard state peptides, the backbone is unconstrained, allowing conformations to evolve rapidly and requiring shorter simulations. The number of configurations generated in each simulation was judged to be sufficient when the side chain conformational entropy calculated from the first half of the simulation matched the value obtained over all conformations.

The side chain rotamer distributions obtained in these simulations using a flexible helix are found to mimic those obtained in our earlier study using a rigid helix, with the exception of Trp (see below). Examples are shown in Figure 2 where the χ_1 side chain rotamer distributions obtained with both flexible and rigid helices are plotted for Val (Fig. 2a) and Leu (Fig. 2b) guest residues. The hydrogen bond and dihedral torsion conformational constraints, described in the previous section, were used in simulations of the flexible helix.

In all cases except side chains with rings, the rotamer distributions from simulated helices compare well with the distributions from protein helices.[14] As seen in our earlier simulations, residues with rings (Phe, Tyr, and Trp) are found to prefer χ_1 values of *gauche*−, a conformation that leads to favorable van der Waals contacts between the ring moiety and the helix backbone. By contrast, the ring-containing side chains favor χ_1 values of *trans* in protein helices.

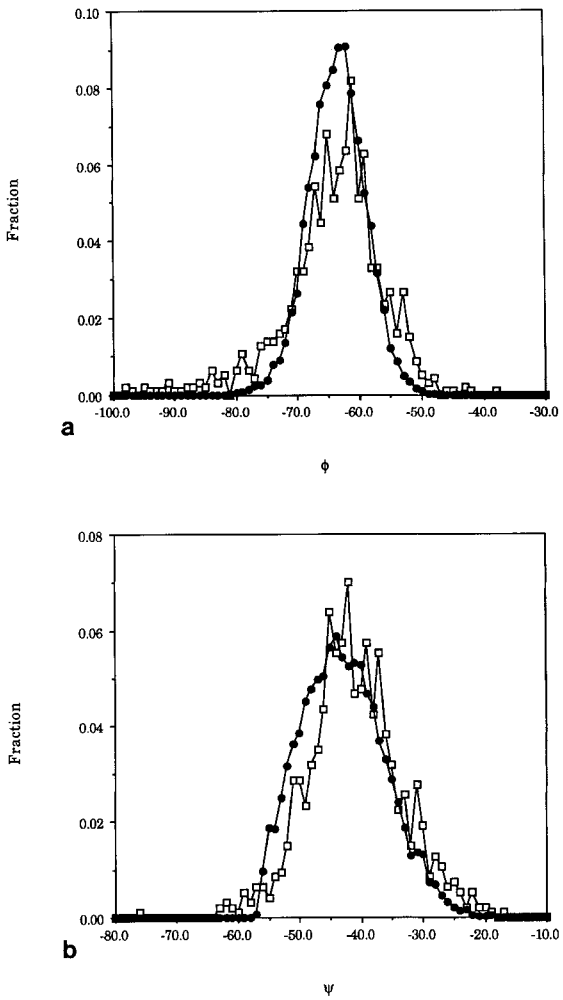

Fig. 1. (**a**) The backbone ϕ dihedral distributions and (**b**) the backbone ψ dihedral distributions, from simulation and protein helices. The simulation results are shown using solid circles and the protein helix data are shown using open squares.

TABLE III. Simulation Length Data and the Autocorrelations at Lag 1 of the Conformation Energies for the α-Helices and Standard States Used for the Side Chain Conformational Entropy Loss Calculations

	α-Helix			Standard state		
Residue	N_{Total}* ($\times 10^{-6}$)	n_c†	r_1‡	N_{Total}* ($\times 10^{-6}$)	n_c†	r_1‡
Val	20	4000	0.476	3	1000	0.335
Ile	30	4000	0.468	6	1000	0.355
Leu	30	4000	0.443	6	1000	0.331
Met	30	4000	0.310	10	1000	0.303
Phe	40	4000	0.412	20	1000	0.330
Tyr	40	4000	0.365	20	1000	0.326
Trp	40	4000	0.443	20	1000	0.385

*Total number of configurations generated in the simulation.
†Number of configurations between each stored conformation.
‡Autocorrelation at lag 1 of the configuration energy.

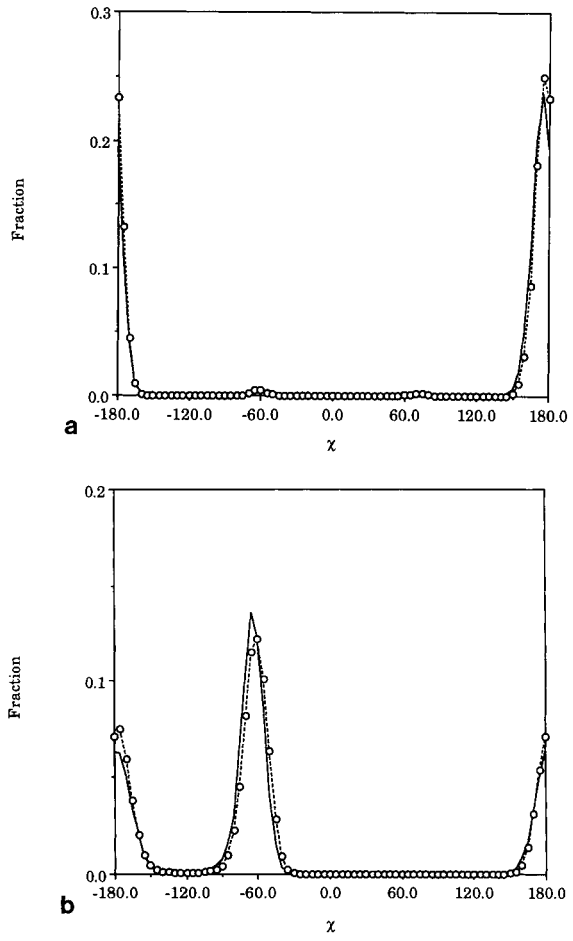

Fig. 2. The side chain χ_1 distributions of (**a**) Val and (**b**) Leu in α-helices. The results of the flexible helix simulations are shown as open circles connected by dashed lines and the results of the rigid helix simulations[12] are shown as solid lines.

The side chain rotamer distributions obtained with the acetylated- and methyl-amidated tripeptides are virtually identical to those obtained in the earlier simulations using unblocked tripeptides.[12] These distributions also resemble the overall rotamer distributions observed in proteins by Ponder and Richards[27] and MacGregor et al.[14]

Calculated values of side chain conformational entropy are summarized in Table IV. All simulations were run at a dielectric of $\epsilon = 78$. Results from our earlier study[12] (which used rigid backbone geometry and united atom radii scaled to 90%) are also listed in Table IV. Agreement between the two sets of results is generally good, with the exception of Leu (see below).

The side chain conformational entropy for Trp in a rigid helix has been found to be in error in our previous study.[12] The corrected value of S^{α}/R for Trp is 1.564. Using this value in Eq. (1) yields a corrected $-T\Delta S_{Rigid}$ of 0.29 kcal mol^{-1} for Trp (Table IV). The linear model for predicting the $\Delta\Delta G$ values measured by O'Neil and DeGrado[11] from calculated values of $-T\Delta S_{Rigid}$ then becomes

$$\Delta\Delta G_D = -1.49T\Delta S_{Rigid} + 0.03 \quad \rho = 0.94 \quad (8)$$

where ρ is the correlation coefficient.

The flexible helix values ($-T\Delta S_{Flexible}$) are plotted against the rigid helix values ($-T\Delta S_{Rigid}$) in Figure 3. Excluding Leu (open square in Fig. 3), a linear model yields an excellent correlation:

$$T\Delta S_{Flexible} = 0.94T\Delta S_{Rigid} - 0.01 \quad \rho = 0.98. \quad (9)$$

The entropy loss calculated for Leu ($-T\Delta S_{Flexible}$ = 0.15 kcal mol^{-1}) is considerably larger than that obtained using the rigid helix model in our previous study ($-T\Delta S_{Rigid}$ = 0.04 kcal mol^{-1}; Creamer and Rose[12]). This difference is due mainly to a decreased side chain entropy in the helical state (S^{α}/R = 0.713 in the flexible helix vs. S^{α}/R = 0.831 in the rigid helix). As seen in Figure 2b, the χ_1 rotamer distributions for Leu do not differ significantly between rigid and flexible helices. In α-helices, Leu populates two rotamer classes preferentially: *trans*,

TABLE IV. Side Chain Conformational Entropies Calculated Using the Flexible and Rigid α-Helices and the Flexible Standard States

Residue	S^{α}/R (α-helix)	S^{0}/R (standard state)	$-T\Delta S_{\text{Flexible}}$* (kcal mol^{-1})	$-T\Delta S_{\text{Rigid}}$† (kcal mol^{-1})
Ala‡	—	—	0.0	0.0
Val	0.048	0.706	0.39	0.39
Ile	0.750	1.216	0.28	0.29
Leu	0.679	0.936	0.15	0.04
Met	2.360	2.574	0.13	0.15
Phe	1.773	2.143	0.22	0.26
Tyr	2.497	2.952	0.27	0.33
Trp	1.536	1.953	0.25	0.29

*Side chain configurational entropy loss at T = 298 K from simulations in this work.
†Side chain configurational entropy loss at T = 298 K calculated using a rigid helix with united atom radii scaled to 90% of their van der Waals values, from Creamer and Rose.[12] The previously published value[12] for $-T\Delta S_{\text{Rigid}}$ of Trp has been corrected.
‡Since Ala has no side chain dihedrals, it loses no entropy upon transfer from the standard state to an α-helix.

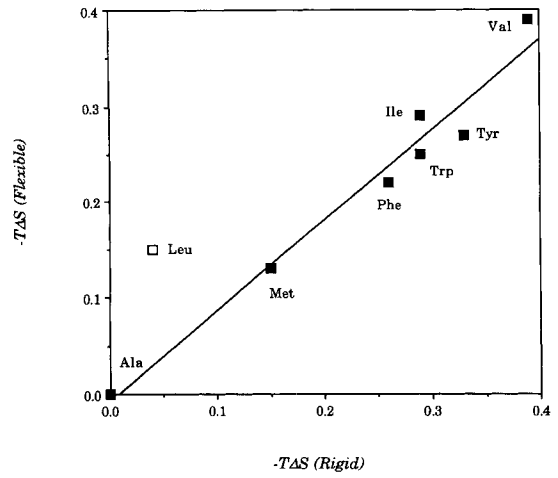

Fig. 3. Side chain conformational entropy losses calculated from the flexible helix simulations ($-T\Delta S_{\text{Flexible}}$) plotted against the entropy losses from the rigid helix simulations ($-T\Delta S_{\text{Rigid}}$) from Creamer and Rose.[12] The corrected value of $-T\Delta S_{\text{Rigid}}$ for Trp is used. The outlier given by the open square is Leu (see text).

gauche+ (*tg*+), and *gauche*−, *trans* (*g* − *t*) (ref. 14). In the rigid helix, the populations in these classes are 33.8% in *tg*+ and 62.7% in *g* − *t*. Corresponding populations in the flexible helix are 33.6 and 65.8%, respectively. The small increase in the population of the *g* − *t* rotamer class occurs at the expense of other sparsely populated rotamer classes, viz. *tt*, *tg*−, *g*−*g*−, and *g*−*g*+, leading to a reduction in S^{α}/R and a consequent increase in $-T\Delta S_{\text{Flexible}}$. The increased population of *g* − *t* rotamers results from closer van der Waals contacts between the Leu Cγ atom and the carbonyl oxygen of the residue at position *i*−4 (with Leu at position *i*). The $C_{\gamma l}O_{i-4}$ distance is approximately 2.9–3.0 Å, within van der Waals contact. While Leu in the rigid helix model can also make this contact, it is not as energetically favored because the Cγ radius is scaled to 90% of its van der Waals value.

As found in our earlier study,[12] strong correlations exist between calculated side chain entropies and experimentally measured helix propensities. Linear models for predicting the experimental values of $\Delta\Delta G$ (normalized with respect to Ala) yield the following equations and correlation coefficients: for O'Neil and DeGrado,[11]

$$\Delta\Delta G_{\text{D}} = -1.75T\Delta S_{\text{Flexible}} - 0.01 \quad \rho = 0.93 \quad (10)$$

and for Lyu et al.,[8]

$$\Delta\Delta G_{\text{K}} = -1.19T\Delta S_{\text{Flexible}} + 0.02 \quad \rho = 0.97. \quad (11)$$

The correlation with the O'Neil and DeGrado[11] free energies, $\Delta\Delta G_{\text{D}}$, is similar to that obtained using a rigid helix and scaled united atom radii, with the intercept closer to zero and the slope increased. The correlation with the Lyu et al.[8] free energies, $\Delta\Delta G_{\text{K}}$, remains high, with an intercept slightly closer to zero.

Comparison with the ellipticities measured by Padmanabhan et al.[7] for Ala, Val, Ile, Leu, Met, and Phe yields

$$-[\theta]_{222} = 57{,}077T\Delta S_{\text{Flexible}} + 27{,}713 \quad \rho = 0.85. \quad (12)$$

This correlation improves when Phe is excluded from the analysis (ρ = 0.92). In this case, the correlation is not quite as high as that obtained using the rigid helix (ρ = 0.95; ref. 12), due to the larger value of $-T\Delta S_{\text{Flexible}}$ obtained for Leu in this work. We note that Padmanabhan et al.[7] pointed out that the CD spectrum of the Phe-containing peptide does not pass through the isodichroic point in their study.

The correlation of $-T\Delta S_{\text{Flexible}}$ with the Chou and Fasman[28] indices of helical preference is weak (ρ = 0.62). However, exclusion of Tyr yields a much improved correlation (ρ = 0.95), better than that observed using a rigid helix model (ρ = 0.90, using the corrected Trp entropy loss). Correlation with the s constants derived from host–guest studies by Sueki

et al.[3] is very weak; a linear regression model for $-T\Delta S_{\text{Flexible}}$ vs. ln(s) yields a correlation of only $\rho = 0.50$. This value does not differ significantly from the rigid helix result[12] (with the corrected Trp value).

The excellent correlation between calculated side chain conformational entropy losses and experimentally measured helix propensity scales[7,8,11] validates the hypothesis that $-T\Delta S$ is a major determinant of helix propensity in isolated helical peptides. Recently, Horovitz et al.[29] published a free energy scale for substitutions made in a helix in barnase (residue 32). A linear model for predicting these $\Delta\Delta G_{\text{B}}$ values (normalized with respect to Ala) yields

$$\Delta\Delta G_{\text{B}} = -2.62T\Delta S_{\text{Flexible}} + 0.03 \quad \rho = 0.95. \quad (13)$$

Figure 4 plots barnase free energies against our calculated side chain entropy losses. Blaber et al.[16] have published similar scales for substitutions at two helical sites in bacteriophage T4 lysozyme (residues 44 and 131). Linear models for predicting these $\Delta\Delta G_{\text{L}}$ values (normalized with respect to Ala) are

$$\Delta\Delta G_{\text{L44}} = -0.92T\Delta S_{\text{Flexible}} + 0.04 \quad \rho = 0.71 \quad (14)$$

and

$$\Delta\Delta G_{\text{L131}} = -0.50T\Delta S_{\text{Flexible}} + 0.04 \quad \rho = 0.81. \quad (15)$$

The good correlations obtained with these scales indicate that the loss of side chain entropy is also a major determinant of helix propensity in proteins. This assertion is underscored by the good correlation obtained with the Chou–Fasman indices of helical preference,[28] discussed above. The wide variation in the slopes obtained in Eqs. (13), (14), and (15) are probably due to context-dependent effects in the protein helices (e.g., interactions of the guest residue side chain with other side chains within the helix). The two bacteriophage T4 lysozyme scales are notable in that they both have narrower energy ranges than those measured in other systems. For example, the differences in free energy between Ala and Val are 0.33 and 0.25 kcal mol^{-1} for substitution sites 44 and 131, respectively, in contrast to 0.63 and 0.45 kcal mol^{-1} in the peptide scales of O'Neil and DeGrado[11] and Lyu et al.,[8] respectively. These narrower energy ranges result in slopes that are less than unity in Eqs. (14) and (15).

Pickett and Sternberg[13] and Blaber et al.[17] have recently presented side chain entropy scales for residues in the unfolded state. These scales were derived from global surveys of the rotamer distributions of nonhelical residues in databases of protein structures. Both scales partition side chain rotamers into rotamer classes differently from each other and from the scheme used in this and our previous work.[12] Entropy values calculated from Eq. (2) are sensitive to the partitioning scheme, so it is necessary to repartition our rotamer distributions in order to effect comparisons with these scales. After recalculation of the standard state entropies for the hydrophobic side chains using the respective partitioning schemes of Pickett and Sternberg[13] and of Blaber et al.,[17] comparison with the two scales yields for Pickett and Sternberg,[13]

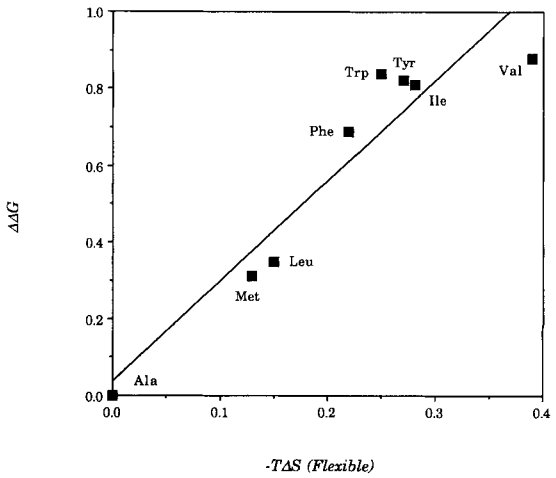

Fig. 4. Side chain conformational entropy losses plotted against the free energies measured for substitutions in an α-helix in barnase.[29]

$$S_{\text{PS}}/R = 0.99S^0/R + 0.12 \quad \rho = 0.98 \quad (16)$$

and for Blaber et al.,[17]

$$S_{\text{B}}/R = 0.86S^0/R + 0.16 \quad \rho = 0.91. \quad (17)$$

Figure 5a and b plots the recalculated standard state entropies against the scales of Pickett and Sternberg[13] and of Blaber et al.,[17] respectively. The remarkable correlation obtained with these scales indicates that the ensemble of side chain rotamers observed for nonhelical residues in proteins strongly resembles that obtained from a simple tripeptide model of the unfolded state.

Blaber et al.[17] also present a scale of side chain entropies for residues within protein helices and a scale of estimated loss of side chain entropy. Comparing this latter scale with our calculated entropy losses (after suitable repartitioning) yields only a modest correlation:

$$T\Delta S_{\text{B}} = 0.68T\Delta S_{\text{Flexible}} + 0.01 \quad \rho = 0.63. \quad (18)$$

The entropy losses of Blaber et al.[17] are plotted against our recalculated entropy losses in Figure 6. The main outliers in the linear model are side chains containing rings (i.e., Phe, Tyr, and Trp). This finding is not surprising since it was noted both in this and our previous work[12] that the ring-containing side chains in simulated helices have rotamer distributions that differ significantly from those observed in actual protein helices. If Phe, Tyr, and

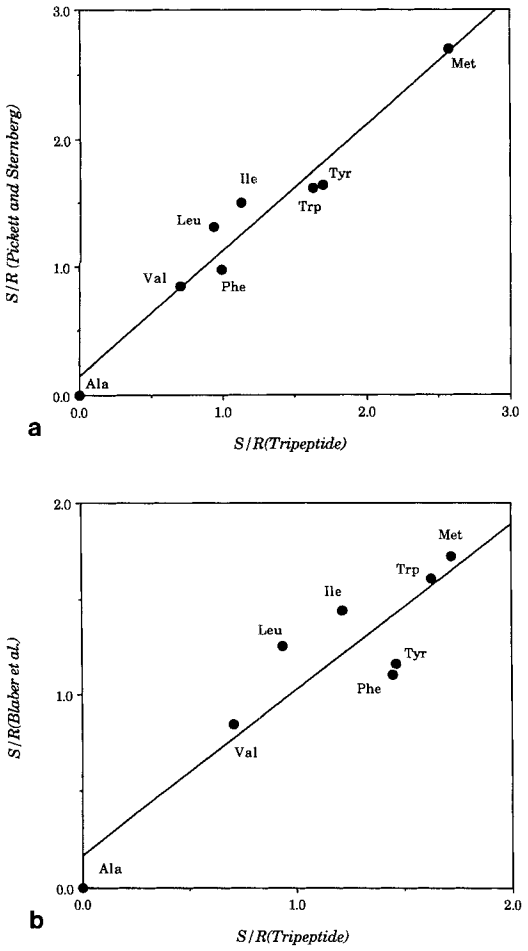

Fig. 5. Side chain conformational entropy in the standard state plotted against side chain entropies for the unfolded state derived from protein structures: (**a**) from Pickett and Sternberg[13] and (**b**) from Blaber et al.[17]

Trp are excluded from the linear model, an improved correlation is obtained:

$$T\Delta S_{B} = 0.83T\Delta S_{\text{Flexible}} + 0.03 \quad \rho = 0.83. \quad (19)$$

Solvent Accessible Surface Area Calculations

The calculated average solvent accessible area of each side chain in both standard and helical states is given in Table V, along with the average amount of side chain surface area buried. These accessibilities are Boltzmann-weighted averages from the simulations. Notably, all side chains bury similar amounts, with little or no correlation between area buried and any helix propensity scale either in peptides,[7,8,11] or in proteins.[13,16] Lack of correlation does not imply that hydrophobic interactions fail to play an important role in stabilizing peptide or protein helices. However, interactions between side chains in an α-helix would depend critically upon their local environment, which cannot be anticipated in models such as the polyalanyl-based helices employed in

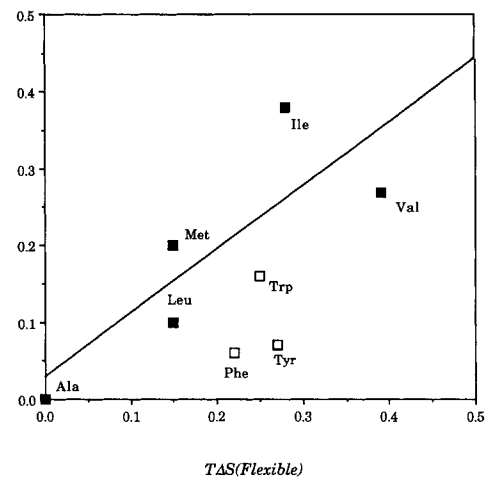

Fig. 6. Calculated side chain entropy losses plotted against side chain entropy losses estimated from rotamer distributions within proteins.[17] The line of best fit excluded the Phe, Tyr, and Trp points (open squares).

this study. Further, the tripeptide standard state is an inadequate representation of the unfolded state for surface accessibility calculations; tripeptide side chains are more exposed to solvent than corresponding side chains in a longer peptide where issues of excluded volume play a role. Hence, the buried surface area estimates in Table V should be viewed as *overestimates*.

DISCUSSION

The flexible helix model was developed to validate our previous work[12] by ensuring that the observed correlations are not an adventitious consequence of the rigid model. Care has been taken in all of our simulations to ensure that consecutive data points are not highly correlated. As noted by Smith and Wells[18] and Kolafa,[19] this precaution is mandatory in computer simulations if standard deviations are to be accurately derived from the zeroth-order variance [Eq. (6)]. The strategy employed here, using the autocorrelation at lag 1, is applicable to both Monte Carlo and molecular dynamics simulations when calculating configurational properties of molecular systems.

α-Helix Hydrogen Bond and Dihedral Angle Constraints

The helix conformational constraints used in this work are an adaptation of constraints developed by Yun and Hermans.[24] Of course, there are simpler ways of achieving the same flexible helix behavior. For example, using the Monte Carlo technique it would be possible to sample backbone conformations directly from the experimentally observed range of helix backbone distributions. However, the constraints derived in this work were developed in order to provide generally applicable criteria for use in

TABLE V. Average Side Chain Accessible Surface Areas in the Standard State and in the Flexible Helix Model

Residue	$\langle A^0 \rangle$* (Å^2)	$\langle A^\alpha \rangle$† (Å^2)	$\langle A^\alpha \rangle - \langle A^0 \rangle$ (Å^2)
Ala	69.6	51.5	−18.1
Val	125.5	106.6	−19.5
Ile	150.5	123.9	−26.6
Leu	155.3	130.8	−24.5
Met	160.8	135.1	−25.7
Phe	171.9	150.8	−21.1
Tyr	186.2	164.9	−21.3
Trp	211.6	183.3	−28.3

*Average side chain solvent accessible surface area in the standard state model.
†Average side chain solvent accessible surface area in the flexible α-helix model.

both Monte Carlo and molecular dynamics simulations.

The optimum choice of an experimentally determined reference system for comparison with simulated α-helices would be the (ϕ,ψ) distributions and hydrogen bond lengths of isolated peptide α-helices in water. Unfortunately, such data are currently unavailable; by default, protein helices are the best available option. Peptide helices may in fact be more flexible than their protein counterparts. In addition, isolated peptide helices are known to fray significantly at their termini,[10] while the constrained helices modeled here are not permitted to do so. However, this work is concerned with the behavior of central residues in the helix, where fraying has little if any effect.

Evidence that fraying does not affect the central residues significantly is provided by comparing the experimental systems themselves. O'Neil and DeGrado[11] determined helix propensities as free energies using a helical coiled-coil, an all-or-none system in which the peptide termini are not expected to be frayed. When folded, peptides in the coiled-coil are virtually 100% helical, as assessed by circular dichroism; guest residues occupy central, solvent-exposed positions. Yet, residue free energies measured in this system compare well with those determined by Lyu et al.,[8,10] with guest residues in the center of a monomeric helical peptide that is known to be frayed.

It has been shown in this work that both hydrogen bond and backbone dihedral constraints are needed to reproduce the conformational distributions observed for protein helices; unconstrained simulations resulted in helices with geometry far from ideal (see Table II). Simulations employing hydrogen bond constraints alone resulted in helices that were prevented from unfolding, but with only slightly improved geometry (Table II).

Our constraints force ϕ to be almost normally distributed around −60°, regardless of the natural minimum in the underlying forcefield (see Fig. 1a). In contrast, the existing forcefield need not be modified for ψ dihedrals; constraints that function as potential "walls" and confine the ψ dihedrals within a specified range are adequate in this case. This awkward strategy for ϕ angles contrasts with the preferred approach of Yun and Hermans,[24] using the Cedar forcefield, where barrier potentials can be imposed on (ϕ,ψ) positions that already have acceptable energy minima.

Side Chain Entropy Loss and Helix-Forming Propensity Scales

When modeling side chain rotamers, good agreement is obtained between simulations of flexible and rigid helices. Of course, a rigid helix with scaled atomic radii is preferred in cases where it proves equally effective; flexible helix simulations are at least an order of magnitude more expensive in CPU time than corresponding rigid helix simulations. As it turns out, the rigid helix model[12] would have been adequate for conformational entropy calculations. The largest difference in the calculated entropy loss occurs with Leu, which can make slightly better van der Waals contacts in the flexible helix model. Even so, rotamer distributions for Leu in a flexible helix are not very different from those obtained in a rigid helix.

In fact, it is possible to derive the rigid helix results from exhaustive conformational search of side chain conformations, a much less demanding method than Monte Carlo simulation. However, exhaustive search becomes unwieldy when exploring conformational behavior in the standard state.

There is a remarkable degree of similarity between the calculated side chain rotamer distributions in tripeptides and rotamer distributions derived from surveys of non-helical residues in X-ray elucidated proteins.[13–15,17] Consequently, excellent agreement exists between calculated standard state entropies and those derived from distributions in proteins[13,17] [see Eqs. (16) and (17)]. Two main conclusions follow. First, the tripeptide standard state effectively models conformational behavior of side chains in the unfolded state. However, this standard state is not suitable for studies involving solvent accessibility because areas calculated from the model are overestimates. Second, the α-helix is unique in being the secondary structure type that most perturbs side chain rotamer distributions. This observation, also noted by MacGregor et al.,[14] implies that side chain entropy loss may not be as strong a determinant for other secondary structure types. This issue begs further examination.

Bai and Englander[30] find a remarkable correlation between side chain blocking factors and β-sheet propensities measured using a zinc-finger host peptide.[31] These side chain blocking factors are ob-

tained using hydrogen exchange techniques and are a measure of the degree to which a side chain obstructs hydrogen bonding to solvent in the backbone. The blocking factors correlate strongly with *β-sheet* propensities ($\rho = 0.96$) but show little correlation with helix-forming propensities.[30] This finding implies that β-sheet propensities are determined in part by the tendency of side chains to impede solvent access to backbone polar atoms and thereby promote the formation of intramolecular hydrogen bonds. The lack of correlation with helix propensity scales implies that blocking factors do not play an equally important role in this case. Clearly, the formation of both α-helix and β-sheet is determined by side chain size, shape, and flexibility, and these factors affect the peptide backbone accordingly.

We note that a strong correlation exists between the experimentally determined helix propensities in isolated peptides and the entropy a side chain loses upon transfer from a flexible tripeptide to a flexible helix. Consequently, the overall conclusion reached in earlier work[12] persists, viz., loss of side chain conformational entropy is a major determinant of a residue's helix-forming tendency, although in at least one carefully modeled case, enthalpic contributions to the helix propensity of valine were found to be of comparable magnitude.[24]

This conclusion would also appear to apply to α-helices in proteins, as indicated by the strong correlation obtained with Chou–Fasman indices of helical preference,[28] as well as with the measured free energies from barnase[29] and bacteriophage T4 lysozyme.[16] At face value, the agreement with barnase data [Eq. (13)] is surprising. The barnase substitution site, residue 32, is in the last turn of a small helix, not a central turn in a longer helix, as modeled in simulations. However, helix side chains lose entropy principally through steric conflict with the backbone of the preceding turn of helix, and, in the absence of fraying, these interactions will not differ from the middle to the last turn. Such would not have been the case had the substitution site been in the first turn of helix. The T4 lysozyme sites, residues 44 and 131, are located in more central portions of longer helices. In all three cases the substitution sites are on the surface, are well-exposed to solvent, and appear to have little interaction with surrounding side chains.

Only modest correlation is found between calculated entropy losses and those derived from rotamer distributions in X-ray elucidated proteins[17] [see Eq. (18)]. However, this correlation is improved markedly by excluding ring-containing residues from the linear fit [Eq. (19)]. Simulated Phe, Tyr, and Trp rotamer distributions in isolated peptides differ from those observed in protein helices, and consequently the calculated entropies are not expected to resemble those derived from protein structures. The differing rotamer distributions are probably a result of two factors. First, our polyalanyl helix model precludes interactions between the bulky rings and other surrounding side chains. Second, ring-containing residues are typically buried in protein structures, facilitating van der Waals contacts with residues in the protein core and resulting in preferred χ_1 values of *trans*. In an isolated peptide, van der Waals contacts with the protein backbone are the only ones possible, resulting in preferred χ_1 values of *gauche*−.

Little correlation between buried surface area and helix-forming propensities was found in this work. However, we note that these calculations are model-limited; more surface area would be buried in a longer unfolded peptide. Nonetheless, it would appear that loss of side chain conformational entropy is the major determining factor in the helix-forming propensities of helical peptides.[7,8,11] Burial of hydrophobic surface is expected to influence the stability of substitutions in protein α-helices; no doubt this factor contributes to the variability in the slopes of Eqs. (13), (14), and (15). However, Lin et al.[32] have shown that hydrophobic interactions augment the contribution from the helix-forming propensities, in contrast to the argument of Blaber et al.[16,17] that the hydrophobic contributions are an integral component of the propensities. In sum, although contrary arguments have been proposed,[16,17] the loss of side chain conformational entropy also appears to be a major determinant in the helix-forming tendencies of residues within protein helices.

Entropy loss always reduces stability, so the conformational entropy loss of residue side chains must destabilize helices. Therefore, the primary source of stabilization energy has to originate in backbone interactions (e.g., amide-carbonyl hydrogen bonds) (see Shirley et al.[33] and references therein). Since all residues (except Pro and Gly) have identical backbones but not all sequences are helical, side chains must confer specificity upon helix formation, determining why some sequences are helical and others are not.[34]

ACKNOWLEDGMENT

We thank Rajgopal Srinivasan, Rajeev Aurora, and Jeffrey Seale for a critical reading of the manuscript. We also thank Walter Englander and Brian Matthews for copies of manuscripts prior to publication and Jay Ponder for accessible surface area subroutines. We are particularly indebted to Jan Hermans for his deep insights and high standards. This work was supported by GM 29458 from the NIH.

REFERENCES

1. Pauling, L., Corey, R.B., Branson, H.R. The structures of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. Proc. Natl. Acad. Sci. U.S.A. 37: 205–210, 1951.
2. Fasman, G.D., ed. "Prediction of Protein Structure and the

Principles of Protein Conformation." New York: Plenum, 1989.
3. Sueki, M., Lee, S., Powers, S.P., Denton, J.B., Konishi, Y., Scheraga, H.A. Helix-coil stability constants for the naturally occurring amino acids in water. XXII. Histidine parameters from poly[(hydroxybutyl) glutamine-*co*-L-histidine]. Macromolecules 17:148–155, 1984.
4. Zimm, B.H., Bragg, J.K. Theory of the phase transition between helix and random coil in polypeptide chains. J. Chem. Phys. 31:536–535, 1959.
5. Bierzynski, A., Kim, P.S., Baldwin, R.L. Local secondary structure in ribonuclease A denatured by guanidine·HCl near 1 degree C. Proc. Natl. Acad. Sci. U.S.A. 79:2470–2474, 1982.
6. Brown, J.E., Klee, W.A. Helix-coil transition of the isolated amino terminus of ribonuclease. Biochemistry 10: 470–476, 1971.
7. Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T.M., Baldwin, R.L. Relative helix-forming tendencies of nonpolar amino acids. Nature (London) 344:268–270, 1990.
8. Lyu, P.C., Liff, M.I., Marky, L.A., Kallenbach, N.R. Side chain contributions to the stability of alpha-helical structure in peptides. Science 250:669–673, 1990.
9. Padmanabhan, S., Baldwin, R.L. Straight-chain non-polar amino acids are good helix-formers in water. J. Mol. Biol. 219:135–137, 1991.
10. Lyu, P.C., Sherman, J.C., Chen, A., Kallenbach, N.R. Alpha helix stabilization by natural and unnatural amino acids with alkyl side chains. Proc. Natl. Acad. Sci. U.S.A. 88:5317–5320, 1991.
11. O'Neil, K.T., DeGrado, W.F. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. Science 250:646–651, 1990.
12. Creamer, T.P., Rose, G.D. Side-chain entropy opposes alpha-helix formation but rationalizes experimentally determined helix-forming propensities. Proc. Natl. Acad. Sci. U.S.A. 89:5937–5941, 1992.
13. Pickett, S.D., Sternberg, M.J.E. Empirical scale of side-chain conformational entropy in protein folding. J. Mol. Biol. 231:825–839, 1993.
14. MacGregor, M.J., Islam, S.A., Sternberg, M.J.E. Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. J. Mol. Biol. 198: 295–310, 1987.
15. Dunbrack, R.L., Jr., Karplus, M. Backbone-dependent rotamer library for proteins. J. Mol. Biol. 230:543–574, 1993.
16. Blaber, M., Zhang, X-J., Matthews, B.W. Structural basis of amino acid alpha helix propensity. Science 260:1637–1640, 1993.
17. Blaber, M., Zhang, X-J., Lindstrom, J.D., Pepoit, S.D., Baase, W.A., Matthews, B.M. Determination of α-helix propensity within the context of a folded protein: Sites 44 and 131 in bacteriophage T4 lysozyme. J. Mol. Biol. 235: 600–624, 1994.
18. Smith, E.B., Wells, B.H. Estimating errors in molecular simulation calculations. Mol. Phys. 52:701–704, 1984.
19. Kolafa, J. Autocorrelation and subseries averages in Monte Carlo simulations. Mol. Phys. 59:1035–1042, 1986.
20. IUPAC-IUB Commission on Biochemical Nomenclature. J. Mol. Biol. 52:1–17, 1970.
21. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E. Equation of state calculations by fast computing machines. J. Chem. Phys. 21:1087–1092, 1953.
22. Jorgensen, W.L., Tirado-Rives, J. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. J. Am. Chem. Soc. 110: 1657–1666, 1988.
23. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr., Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. J. Am. Chem. Soc. 106:765–784, 1984.
24. Yun, R.H., Hermans, J. Conformational equilibria of valine studied by dynamics simulation. Protein Eng. 4:761–766, 1991.
25. Richmond, T.J. Solvent accessible surface area and excluded volume in proteins. J. Mol. Biol. 178:63–89, 1984.
26. Presta, L.G., Rose, G.D. Helix signals in proteins. Science 240:1632–1641, 1988.
27. Ponder, J.W., Richards, F.M. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. J. Mol. Biol. 193:775–791, 1987.
28. Chou, P.Y., Fasman, G.D. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47:45–148, 1978.
29. Horovitz, A., Matthews, J.M., Fersht, A.R. α-Helix stability in proteins. II. Factors that influence stability at an internal position. J. Mol. Biol. 227:560–568, 1992.
30. Bai, Y., Englander, S.W. Hydrogen bond strength and β-sheet propensities: The role of a side chain blocking effect. Proteins 18:262–266, 1994.
31. Kim, C.A., Berg, J.M. Thermodynamic b-sheet propensities measured using a zinc-finger host peptide. Nature (London) 362:267–270, 1993.
32. Lin, L., Pinker, R.J., Kallenbach, N.R. α-Helix stability and the native state of myoglobin. Biochemistry 32: 12638–12643, 1993.
33. Shirley, B.A., Stanssens, P., Hahn, U., Pace, C.N. Contribution of hydrogen bonding to the conformational stability of ribonuclease T1. Biochemistry 31:725–732, 1992.
34. Lattman, E.E., Rose, G.D. Protein folding—what's the question? Proc. Natl. Acad. Sci. U.S.A. 90:439–441, 1993.