Simulations of Peptide Conformational Dynamics and Thermodynamics

Charles L. Brooks III* and David A. Case*

Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037

Received August 5, 1993 (Revised Manuscript Received September 1, 1993)

Contents

1. Introduction

The past few years have seen impressive advances in the power and scope of computer simulation techniques for biomolecules, and many important aspects of the behavior of peptides in solution can now be studied with microscopic simulations. Calculations in this area are still quite difficult, however, both because of the difficulty in estimating energetics involved and because the floppy character of linear peptides makes averaging over conformational degrees of freedom quite difficult. In this review, we describe recent progress in carrying out simulations on short linear peptides in water. Our goal is both to describe the insights and results obtained in the calculations and to present some critical evaluations of the expected reliability of various approaches.

We begin in section 2 with a brief review of molecular dynamics methods as applied to peptide simulations, looking at both the forms of commonly used potential functions and at aspects of the methods used to perform the simulations. This is followed by a synopsis of the current state of theoretical methods for calculating basic features of peptide conformational energetics and thermodynamics. We then review a number of recent calculations on the alanine dipeptide model system, a prototypical system for studies of conformational and solvation energetics.

In section 4 we begin a survey of recent results from solvated simulations of peptides in solution, considering in turn models for helices, tight turns, and sheets. We conclude in section 5 with an overview of perspectives for further advances, and discuss the implications of simulation methods for our understanding of peptide conformational preferences and protein folding.

2. Molecular Dynamics

In this section we briefly review the equations underlying the classical molecular dynamics technique. We provide a description of the basic scheme used in molecular dynamics and then discuss the form of the empirical potential used in simulations of biopolymers.

In a molecular dynamics (MD) simulation, the classical equations of motion for the system of interest (e.g., a biopolymer in solution) are integrated numerically by solving Newton's equations of motion:[1]

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = -\nabla_i[U(\mathbf{r}_1,\mathbf{r}_2,...\mathbf{r}_N)] \qquad i = 1,N \tag{1}$$

From the solution of these equations, the atomic positions and velocities as a function of time are obtained (here m_i and $\mathbf{r}_i$ represent the mass and position of particle i and U is the potential-energy surface, which depends on the positions of the N particles in the system). Knowledge of the time history or trajectory of the atoms permits the computation of properties such as structure, folding pathways, diffusion, and thermodynamics to be studied. The key steps in the numerical solution of the classical equations of motion may be divided into two parts: the evaluation of energies and forces and the propagation of atomic positions and velocities.[1]

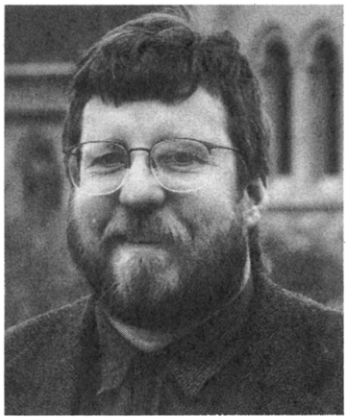
The functional form of the potential-energy function for biological molecules includes energy terms to represent chemical bonds, angles, and improper torsions as well as rotations about bonds (dihedrals) and pairwise additive nonbonded interactions (van der Waals and Coulombic). One form for the overall potential which is often used is indicated below:[2,3]

$$U = \sum_{\text{bonds}} {}^1/_2 k_b(b-b_0)^2 + \sum_{\text{angles}} {}^1/_2 k_\theta(\theta-\theta_0)^2 + \sum_{\text{impropers}} k_\omega(\omega-\omega_0)^2 + \sum_{\text{torsions}} k_\Phi[1+\cos(n\varphi-\delta)] + \sum_{\text{nonbonded pairs}} \left\{4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] + \frac{q_i q_j}{\epsilon r}\right\} \tag{2}$$

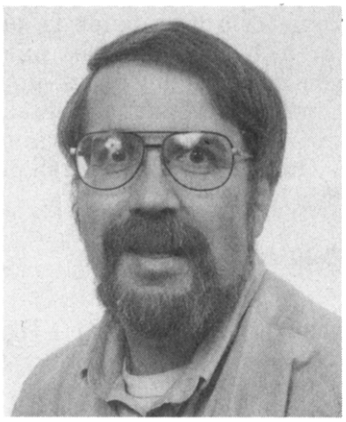
In this equation the constants k_b, k_θ, k_ω, and k_Φ are the force constants for deformation of bonds, angles, impropers, and dihedrals, respectively. The equilibrium values of bond distance, valence angle, and improper torsion correspond to b_0, θ_0, and ω_0. In the dihedral terms, n is the periodicity of the underlying torsional potential, e.g., $n = 3$ representing the 3-fold rotational minima around a C–C single bond, and δ is a phase factor. The constants ϵ_{ij}, σ_{ij}, and q_i represent the atomic (i,j) pair Lennard-Jones well depth and diameter and the partial electrostatic charges.

For simulations of peptides in water it turns out generally not to be sufficient to simply place a few shells of water molecules around the solute, since the surface tension of the resulting water–vacuum interface can appreciably affect the properties of the system. These boundary conditions can be modified to mimic the effects of the remaining solvent that is not explicitly included,[4,5] or "periodic" boundary conditions can be employed in which the system is replicated in three

0009-2665/93/0793-2487$12.00/0



Charles Brooks received his B.S. degree in Chemistry and Physics from Alma College in 1978 and his Ph.D. degree in Physical Chemistry, under the direction of Stephen Adelman, from Purdue University in 1982. He was a postdoctoral fellow at Harvard University from 1982 to 1985 with Martin Karplus and held a NIH postdoctoral fellowship during 1983–1985. He joined the faculty of Carnegie Mellon University as an Assistant Professor in 1985 and was promoted to Associate Professor in 1989 and to Professor in 1993. During the 1992–93 year, Dr. Brooks was a sabbatical visitor at the Center for Structural Biochemistry, Karolinska Institute in Stockholm, Sweden, and at The Scripps Research Institute, Department of Molecular Biology. Dr. Brooks was awarded a Visiting Scholar Grant from the Swedish National Research Council in 1992. He was the recipient of an A. P. Sloan Foundation Fellowship in (1990–1993). His research interests are in the area of molecular biophysics and include the simulation and statistical mechanics of protein and peptide conformational dynamics and thermodynamics.



David Case received a B.S. degree in chemistry from Michigan State University in 1970 and a Ph.D. degree in chemical physics, under the direction of Dudley Herschbach and Martin Karplus, from Harvard University in 1977. He then joined the chemistry department at the University of California, Davis as an Assistant Professor and was promoted to Associate Professor in 1982 and to Professor in 1986. He joined the Department of Molecular Biology at The Scripps Research Institute in 1986. Dr. Case has received Woodrow Wilson and Alfred P. Sloan Foundation fellowships. His research interests are in theoretical chemistry and deal with structural and dynamical properties of proteins and nucleic acids, with emphasis on macromolecular NMR spectroscopy.

dimensions so that no water–vacuum interfaces remain.[1] Because of the long-range nature of electrostatic interactions, it can be difficult to establish innocuous boundary conditions, and the question of how best to carry out solvated simulations is still an active area of research.[6,7]

From eq 2 it is clear that the internal energy U is a complicated function of the configuration of the many-body system which must be differentiated with respect to the Cartesian positions of each atom for the numerical solution of Newton's equation of motion. This tedious but straightforward calculation is well suited to automatic computation. In practice, the numerical solution of Newton's equations of motion are most often achieved using a very simple propagation algorithm initially adopted by Verlet:[8]

$$r_i(t + \delta t) = 2r_i(t) - r_i(t - \delta t) - \delta t^2 \nabla_i U(t)/m_i \quad (3)$$

which has excellent stability properties appropriate for long integrations. Step sizes δt are generally chosen to be around 10^{-15}s. The resulting trajectories describe the time course of the system, and it is currently feasible to simulate peptides in solution for times on the order of nanoseconds, i.e. for about 10^6–10^7 integration steps. In addition to this time-dependent behavior, the statistical properties of the configurations visited during even a short simulation can be related to time-*independent* thermodynamic quantities, as we discuss below.

When extensive holonomic constraints are to be applied to the dynamics, e.g., to keep all water molecules at their equilibrium geometry, an additional set of coupled constraint equations may be solved via an iterative self-consistent method known as SHAKE.[9] The most common form of the applied constraints used in simulations of pure solvents and biopolymer solutions represents the distance constraints between a pair of atoms as:

$$r_{ij}^2 - d_{ij}^2 = 0 \quad (4)$$

where r_{ij} is the instantaneous separation between atoms i and j and d_{ij} is the reference constraint value. These constraints are implemented in molecular dynamics by first taking an unconstrained step using eq 3, then adding a displacement vector, which represents the displacement due to the forces of constraint, to satisfy the constraints specified by eq 4. The solution of the constrained equations of motion is written as

$$r_i(t + \Delta t) = r_i'(t + \Delta t) + \Delta r_i(t) \quad (5)$$

where $r_i'(t + \Delta t)$ is the position vector of the ith atom after an unconstrained step and $\Delta r_i(t)$ is the displacement vector required to satisfy the set of coupled constraint equations

$$\Delta r_i(t) = \frac{2(\Delta t)^2}{m_i} \sum_j \lambda_{ij} r_{ij}(t) \quad (6)$$

The λ_{ij} in this equation represent the Lagrangian multipliers which, upon substitution of eqs 4 and 5 into eq 3 yielding a quadratic equation in λ_{ij}, are solved by linearization and iteration to give the $\Delta r_i(t)$.[9]

2.1. Free-Energy Methods

The basic elements just outlined constitute the necessary components of molecular dynamics programs. The numerical solution of classical equations of motion provides the primary tool for studies of peptide conformational dynamics and thermodynamics. However, the time scales of many interesting processes that occur in proteins and peptides are determined by the rates of overcoming free-energy barriers. Since the free-energy barriers to these processes typically range from a few to many tens of kilocalories per mole, the time scales range from hundreds of picoseconds to milli-

seconds and longer.[10] At present, it is only feasible to carry out MD simulations of proteins or small peptides in solution for several hundred picoseconds or, occasionally, a few nanoseconds. Thus, it is generally impossible to obtain a statistically valid sample of the relevant conformations using standard MD simulation techniques. To circumvent this problem in the case of conformational transitions, specialized sampling techniques such as umbrella sampling,[11] MD with holonomic constraints, and thermodynamic perturbation (TP) theory,[12,13] and techniques of "chemical" perturbations through thermodynamic cycles[14] are employed. Below we consider each of these methods in turn and provide a brief overview of their usage in calculations of conformational free energies. The methods described here are not the only ones possible, and other recent reviews outline some of the alternatives.[14,15]

The Helmholtz free-energy difference between one system in two differing conformational states, say one in which the reaction coordinate is $\zeta = \zeta_0$ and another in which $\zeta = \zeta_1$, is given by[16]

$$\Delta A \equiv A(\zeta_1) - A(\zeta_0) = -\beta^{-1} \ln \frac{\rho(\zeta_1)}{\rho(\zeta_0)} \tag{7}$$

where the probability density, $\rho(\zeta)$, of the conformational coordinate, ζ, is

$$\rho(\zeta) = \frac{\int \delta(\zeta' - \zeta) \exp[-\beta U(\zeta')]\, d\mathbf{x}'}{\int \exp[-\beta U(\zeta')]\, d\mathbf{x}'} \equiv \langle \delta(\zeta' - \zeta) \rangle \tag{8}$$

In eqs 7 and 8, U is the potential energy, $\beta = 1/k_B T$, T is the absolute temperature, $\delta(y)$ is the Dirac delta function, the $3N$ Cartesian coordinates of the system are represented by the vector $\mathbf{x}$, and the $\langle ... \rangle$ denotes a canonical ensemble average. (It is also generally possible to carry out constant pressure simulations to obtain Gibbs free energies.) The potential energy depends on all $3N$ degrees of freedom, but for the sake of simplicity we have only indicated its dependence on ζ. Furthermore, in what follows we assume that $\rho(\zeta)$ is normalized such that $\rho(\zeta_0) = 1$ and we drop the subscript on ζ_1. Hence eq 7 becomes

$$\Delta A = -\beta^{-1} \ln \rho(\zeta) \equiv W(\zeta) \tag{9}$$

In the MD application of the umbrella sampling approach,[11,16] a biasing potential, U^*, is used to enhance the sampling of the average in eq 8. When this is the case one can show that

$$\rho(\zeta) = \frac{\rho^*(\zeta) \exp[\beta U^*(\zeta)]}{\langle \exp(\beta U^*) \rangle^*} \tag{10}$$

where

$$\rho^*(\zeta) = \langle \delta(\zeta' - \zeta) \rangle^* \tag{11}$$

is the biased probability density of ζ, and the $\langle ... \rangle^*$ denotes an average over the biased ensemble. Combining eqs 10 and 11, the following equation for the free-energy surface in terms of the biased probability density is obtained

$$W(\zeta) = -\beta^{-1} \ln \rho^*(\zeta) - U^*(\zeta) + \beta^{-1} \ln \langle \exp(\beta U^*) \rangle^* \tag{12}$$

In principle, eq 12 could be used to compute the surface over any desired range of ζ if one could find a single suitable biasing potential, $U^*(\zeta)$, to add to the potential energy function of the system. This is generally not possible. Instead, a series of simulations is carried out in which the minimum in U^* is systematically varied. This procedure yields a set of biased distributions $\rho_i^*(\zeta)$, centered around different values of ζ. With suitable choices for the minima in U^*, a series of overlapping biased distributions which span the desired range of ζ is obtained. Finally, the W_i are connected together to construct the overall free-energy surface.[17,18] Note that the final term in eq 12 is independent of ζ, so that it may be obtained either by direct calculation or by the requirement that the resulting potential be smooth in overlap regions between various simulations; the latter approach appears to allow larger distances between the minima of successive biasing potentials.[18-21]

Within the formalism of thermodynamic perturbation (TP) theory,[22] the Helmholtz free-energy difference, ΔA, between one system, in which the conformational coordinate of interest (e.g. an internal coordinate) is equal to ζ, and another system, in which the coordinate has been "perturbed" by the amount $d\zeta$, is given by the formally exact equation:

$$\Delta A \equiv A(\zeta + d\zeta) - A(\zeta) = -\beta^{-1} \ln \langle \exp\{-\beta[U(\zeta + d\zeta) - U(\zeta)]\} \rangle_\zeta \tag{13}$$

where the $\langle ... \rangle_\zeta$ denotes a canonical ensemble average over over the "reference" ensemble in which the coordinate is equal to ζ. Since the average is over the reference ensemble, the system must be constrained to keep the value of the coordinate of interest, ζ, fixed at all times during the simulation. In other words, a holonomic constraint must be imposed during the integration of the equations of motion.[23] It is also possible to use more flexible restraints to confine simulations to particular regions of configuration space.[11]

If the free-energy difference between the systems with conformations ζ_0 and ζ_1, or the free-energy barrier separating them, is greater than about $2k_B T$, then a single simulation is not generally sufficient to determine accurately the free-energy difference.[24] This problem may be circumvented by breaking up the range of the coordinate, $\zeta_0 \rightarrow \zeta_1$, into several segments or "windows", and running a series of simulations where the free-energy differences, $\Delta A_i = A(\zeta_i \pm d\zeta)$ are computed in the ith simulation. The free-energy surface over the range $\zeta_0 < \zeta < \zeta_1$ is then constructed by connecting the ΔA_i from the various windows.

Using the two methods just described, one may compute free-energy surfaces for conformational changes in peptides and proteins, and integration over relevant ranges of these surfaces can yield relative populations of differing conformers. However, information about conformational equilibria may also be obtained using methods of "chemical" perturbations and thermodynamic cycles. We illustrate the use of this method for a simple problem, turn formation in alanine and glycine peptides.

The free energy difference for the formation of a type I turn for the sequences X-Ala-Y and X-Gly-Y are given by $\Delta A_{(i)}$(fold) and $\Delta A_{(ii)}$(fold), corresponding to:

$$(\text{X-Ala-Y})_{\text{extended}} \rightarrow (\text{X-Ala-Y})_{\text{turn}} \tag{i}$$

$$(\text{X-Gly-Y})_{\text{extended}} \rightarrow (\text{X-Gly-Y})_{\text{turn}} \quad \text{(ii)}$$

The relative ability of alanine to form a turn at this position (compared to glycine) is given by $\Delta\Delta A(\text{Ala} \rightarrow \text{Gly}) = \Delta A_{(i)}(\text{fold}) - \Delta A_{(ii)}(\text{fold})$. This quantity can be computed using the conformation free-energy methods described above. However, it is also possible to use chemical free-energy methods (thermodynamic perturbation theory) to estimate this difference.[15] By noting that the combined processes (i) and (ii) form a thermodynamic cycle, one can write

$$\Delta\Delta A(\text{Ala} \rightarrow \text{Gly}) = -\Delta A_{\text{turn}}(\text{Ala} \rightarrow \text{Gly}) + \Delta A_{\text{extended}}(\text{Ala} \rightarrow \text{Gly}) \quad (14)$$

where the two quantities on the right-hand-side represent the processes:

$$(\text{X–Ala–Y})_{\text{turn}} \rightarrow (\text{X-Gly-Y})_{\text{turn}} \quad \text{(iii)}$$

and

$$(\text{X-Ala-Y})_{\text{extended}} \rightarrow (\text{X-Gly-Y})_{\text{extended}} \quad \text{(iv)}$$

To obtain $\Delta A_{(iii)}$ and $\Delta A_{(iv)}$ a "hybrid" potential-energy function may be defined to map the system from the "reactant" state to the "product" state via the reaction progress coordinate λ:

$$U(\lambda) = (1 - \lambda)U_{\text{R}} + \lambda U_{\text{P}} + U_{\text{env}} \quad (15)$$

where U_{R} and U_{P} are the potential energies of the "reactant" (e.g. the alanine-based peptide) and "product" (e.g. the glycine-based peptide), respectively, U_{env} is the potential energy of the "environment", the part of the system whose identity does not change during the process (e.g. the solvent and remaining peptide), and λ is a coupling parameter ($\lambda = 0$ and $\lambda = 1$ correspond to the reactant and product systems, respectively). TP theory is used to compute incremental Helmholtz free-energy differences, ΔA_i, for small changes, $d\lambda$, in the coupling parameter:

$$\Delta A = A(\lambda_i + d\lambda) - A(\lambda_i) = -\beta^{-1} \ln \langle \exp\{-\beta[U(\lambda_i+d\lambda)-U(\lambda_i)]\}\rangle_{\lambda_i} \quad (16)$$

In eq 16, the $\langle \ldots \rangle_{\lambda_i}$ denotes an average over the canonical ensemble where $\lambda = \lambda_i$. The sum of ΔA_i from a series of simulations with the λ_i and $d\lambda$ chosen to span the full range (0 to 1) of λ yields the free-energy difference for the transformation of the reactant to the product.

In order to understand the relative free energies of the various conformations in solution in terms of microscopic interactions, the free-energy differences may be decomposed into energetic and entropic contributions arising from differences in peptide–peptide and peptide–water interactions. The Helmholtz free-energy difference between two conformations (e.g. two points on the reaction coordinate) may be written:[25]

$$\Delta A = \langle \Delta U_{\text{uu}} \rangle + \langle \Delta U_{\text{uv}} \rangle + \langle \Delta U_{\text{vv}} \rangle - T\Delta S_{\text{c,uv}} - T\Delta S_{\text{vv}} \quad (17)$$

where $\langle \Delta U_{\text{uu}} \rangle$, $\langle \Delta U_{\text{uv}} \rangle$, and $\langle \Delta U_{\text{vv}} \rangle$ are differences in the average peptide–peptide, peptide–water, and water–water interaction energies, respectively, $\Delta S_{\text{c,uv}}$ is the sum of the differences in the configurational and peptide–water entropies, and ΔS_{vv} is the difference in the water–water entropy. Yu and Karplus[26] have shown that the solvent–solvent energy and entropy contributions exactly cancel each other in the solvation free energy for each conformation. Thus, eq 17 may be rewritten as follows:

$$\Delta A = \langle \Delta U_{\text{uu}} \rangle + \langle \Delta U_{\text{uv}} \rangle - T\Delta S_{\text{c,uv}} \quad (18)$$

It is straightforward to compute the interaction energies but difficult to calculate the entropy directly from the simulation data. Since $T\Delta S_{\text{c,uv}}$ is the only missing term in eq 18, it can be obtained indirectly as the difference between ΔA and the average interaction energies. In many cases the computation of energies and entropies can lead to quantitative understanding of folding forces. However, the direct calculation of these quantities is known to be hampered by poor convergence, leading to estimates with large statistical uncertainties. The basic problem is that entropy and enthalpy estimates require averages over the complete potential energy, with correspondingly large fluctuations, whereas free-energy differences can be expressed as averages over a relatively small number of terms in the potential-energy function.

Before proceeding to the next section, where we discuss features of the alanine dipeptide potential-energy surface as a model for interactions in peptides, we note that the accuracy and precision of calculations of conformational thermodynamic quantities as outlined in this section are less well understood than their "chemical" counterparts. Questions regarding the nature and extent of sampling of conformational space are only now being carefully addressed. Furthermore, for quantities like the energies and entropies described in eqs 17 and 18, the precision is low: these quantities respond to fluctuations in the total energy (rather than energy differences), and represent thermodynamic derivatives that are intrinsically "noisier" than are free-energy differences. Given these caveats, we note that information consistent with findings from experiment and other simulation methods are pointing to the power of conformational free-energy methods to provide insight into the detailed nature of mechanism and kinetics in peptide folding.

3. The Alanine Dipeptide

The "alanine dipeptide" [α-(formylamino)propanamide] has for a long time served as one of the primary models for theoretical studies of backbone conformational equilibria in proteins and peptides. The dipeptide species has many of the structural features of a polypeptide backbone: flexible ϕ and ψ angles, two peptide groups whose NH and CO groups are capable of participating in hydrogen bonds with each other or with solvent molecules, and a methyl group attached to the C^{α} position. Although there are only two "floppy" degrees of freedom aside from methyl rotations, the potential energy surface is still reasonably complex, with about six local minima and a dozen saddle points. Calculations on this species thus can teach important lessons about the level of accuracy to be expected for various conformational analysis techniques.

3.1. Nomenclature

The conformations of this molecule can approximately be specified by the values of the conventional torsion angles ϕ and ψ about the central carbon atom. Several "allowed" regions have local minima in most

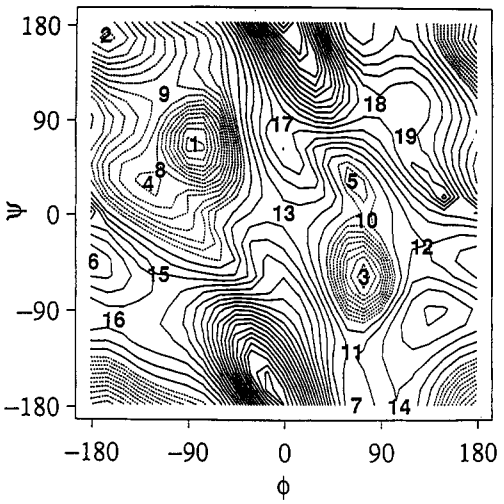


Figure 1. Hartree–Fock 3-21G energy surface for a truncated alanine dipeptide analogue. The fully relaxed HF energies were computed on a 15° grid. The dashed contours are drawn at 0.5 kcal/mol intervals, extending from the zero of energy at C_7^{eq} (labeled 1) to 7 kcal/mol. Solid contours are drawn at 1 kcal/mol intervals thereafter. Numbers label stationary points. Adapted from ref 30.

calculations of the potential energy surface: two conformers that have 7-atom rings closed by a hydrogen bond are C_7^{eq} near (ϕ,ψ) = (−80°,70°) and C_7^{ax} near (70°,−65°). A nearly extended structure forms a (strained) internal hydrogen bond in a 5-atom ring: C_5 near (−160°,170°). Conformers that could lead to α helices in larger peptides are labeled α_R near (−60°,−40°) and α_L near (+60°,+40°); the former often does not correspond to a local minimum in estimates of the gas-phase surface. It is also common to refer to a polyproline P_{II} conformation, at (−80°,+150°) or sometimes (−80°,+120°). In some calculations, especially in earlier work, conformations were generated by simply rotating ϕ and ψ, keeping other bonds and angles at "standard" values. It is now more common to relax all other degrees of freedom by energy minimization, but it can still be confusing to know if particular calculations report energies at accurately-determined stationary points, or simply energies at "interesting" regions of the surface.

3.2. Quantum Mechanical Results

Since our experimental knowledge of gas-phase conformations of the alanine dipeptide is quite limited, it is of interest to see what predictions can be made from quantum-mechanical calculations. Although such calculations go back many years,[27-29] it is only recently that ab initio calculations with flexible basis sets have been feasible, and we shall concentrate on these more recent results here.[30-33] Summaries of early quantum mechanical calculations on the alanine dipeptide are given elsewhere.[32,34]

Figure 1 shows a Hartree–Fock (HF) 3-21G basis set calculation for the entire (relaxed) potential energy surface, computed at 15° spacing in ϕ and ψ.[30] This shows fairly broad minima in the C_7^{eq} (labeled 1), C_7^{ax} (3) and C_5 (2) locations, with narrower and higher minima at α_L (5) and at a "β_2" position (4) that one might imagine could be distorted down into the α_R location through some modification of the surface. As with some other conformational properties, results at the HF/3-21G level turn out to be in agreement with correlated results using larger basis sets, so that this map provides a useful overview of the properties of the dipeptide Ramachandran map, although the details are not expected to be fully reliable.

It is of particular interest to analyze the differences in energy at various minima and important points for the dipeptide, especially for calibration of molecular mechanics potentials. Some features of interest are described in the following paragraphs.

3.2.1. $C_7^{eq} \rightarrow C_5$ Transition

At the 3-21G level, the C_7^{eq} conformer is favored by about 1.2 kcal/mol over C_5. This difference drops to about 0.2 kcal/mol in larger basis sets (see Table I), but then rises to ca. 1.5–2.0 kcal/mol when correlation effects at the MP2 level are included. It has been recognized for some time that correlation corrections are likely to be significant for these conformational energy differences, since the "compactness" of the molecule changes significantly from one minimum to another. Using an empirical dispersion correction, Weiner et al.[29] estimated that electron correlation would raise the energy of C_5 relative to the more compact C_7^{eq} by about 1.7 kcal/mol. This change is in rough agreement with MP2 and nonlocal density functional (NLDF) calculations listed in Table I, and the best current estimates suggest that C_5 should be above C_7^{eq} by 1.5–2.0 kcal/mol.

3.2.2. $C_7^{eq} \rightarrow C_7^{ax}$ Transition

Both of these conformations have an internal 7-membered hydrogen bond, but the latter is more crowded sterically and corresponds to a backbone conformation that is only rarely encountered in proteins. For this reason, there has been considerable discussion over the years about whether the absence of C_7^{ax} conformers in proteins arises from energetics intrinsic to the (gas-phase) dipeptide or from more cooperative interactions involving solvent or other peptide groups.[35] The results in Table I suggest that both explanations contribute: the C_7^{ax} conformations of the dipeptide are indeed disfavored relative to C_7^{eq}, but only by 2.0–2.5 kcal/mol. It is impossible to be sure about the intrinsic accuracy of these calculations, but general experience suggests that they are unlikely to be seriously in error, i.e. by more than 1 kcal/mol. The errors from basis set deficiencies are likely to be less than this, and it is encouraging that three estimates of correlation effects (MP2, nonlocal density functionals, and an empirical dispersion correction) all give similar answers.

3.2.3. $C_7^{eq} \rightarrow \alpha_R$ Transition

Quantum mechanical investigations of the alanine dipeptide surface have often ignored this region of dihedral angle space since, at most levels of theory, there is no local minimum: helical regions in proteins are stabilized by solvent interactions and by the formation of 10- or 13-membered hydrogen bond rings that are not available in the dipeptide. Nevertheless, it is of considerable importance for simulations of proteins and peptides to estimate the local penalty paid to move the dipeptide itself to the α_R region, and a few calculations have studied this by computing energies at (ϕ,ψ) = (−60°,−40°). The MP2 and NLDF results

Table I. Relative Energies in the Alanine Dipeptide in Vacuum

authors	ref	method	C_7^{eq}	C_7^{ax}	C_5	α_R
		Hartree–Fock Results				
Gould and Kollman	33	HF/3-21G	0.0	2.8	1.1	
	33	HF/TZVP//HF/6-31G**	0.0	3.0	0.2	4.2[b]
Bohm and Brode	31	HF/DZP	0.0	3.0	0.5	3.2
Head-Gordon et al.	30	HF/6-31+G*	0.0	2.6	0.2	
Frey et al.	32	HF/6-311G**	0.0		0.2	
		Estimates of Correlated Energies				
Weiner et al.	29	HF/4-31G+DC[a]	0.0	2.2	1.9	
Head-Gordon et al.	30	MP2/6-31+G**//HF/6-31+G*	0.0	2.2	1.1	
Frey et al.	32	MP2/6-311G**//HF/6-311G**	0.0		1.3	
Frey et al.	32	MP2/6-311G**	0.0		1.7	
Gould and Kollman	33	MP2/TZVP//HF/6-31G**	0.0	2.1	1.5	4.0[b]
St. Amant et al.		NLDF[c]	0.0	2.1	1.3	5.3[b]
		Empirical Potential Functions				
Zimmerman et al.	38	ECEPP	0.0	8.8	0.4	1.1
Roteman et al.	35	ECEPP/2	0.0	7.3	0.7	0.8
Weiner et al.	29	AMBER united atom	0.0	1.1	4.0	4.7
Schiffer et al.	39	AMBER all-atom	0.0	1.3	4.8	4.7
Jorgensen and Tirado-Rives	40	AMBER/OPLS	0.0	2.5	1.5	
Pettit and Karplus	41	"model 4"	0.0	0.3	4.8	
Tobias and Brooks	44	CHARMM-19	0.0	2.0	1.8	6.8
this work		CHARMM-22	0.0	2.2	1.0	5.4

[a] DC is an empirical correction for dispersion, based on the AMBER molecular mechanics potential. [b] Geometry in which the (ϕ,ψ) angles were constrained to (−60°,−40°), and the geometry was minimized at the HF/6-31G** level. [c] Nonlocal density functional calculations, computed by A. St. Amant, W. Cornell, T. Halgen, and P. A. Kollman (personal communication). The Gaussian basis set was (7111/411/1*) for heavy atoms and (41/1*) for hydrogens; exchange and correlation was modeled with with the "Becke 88" gradient expansion for exchange and the "Perdew 86" model for correlation. Details will be published elsewhere.

shown in Table I suggest that α_R in the bare dipeptide is disfavored by 4–5 kcal/mol relative to C_7^{eq}.

3.3. Molecular Mechanics Estimates

Successful "molecular mechanics" force fields incorporate many of the features that influence conformational transitions in complex molecules, and are widely used in the analysis of peptide and protein problems.In spite of their general overall success, there is little reason to expect such empirical potentials to be more accurate than the quantum results quoted above; indeed, in many cases, important parts of the input to the parameterization process come from quantum calculations, typically with smaller basis sets than those described above. It is nevertheless of interest to examine predictions from widely-used force fields, in order to gain some insight into the different predictions that are made, and to help gauge the reliability of such calculations when applied to larger peptide and protein systems.

Nguyen and Case[36] have made a careful analysis of the alanine dipeptide surface obtained in the AMBER all-atom force field,[37] characterizing 6 local minima and 11 saddle points, all within 10 kcal/mol of the global minimum. The general features are nearly the same as for the Hartree–Fock surface shown in Figure 1, except that the β_2 minimum found in the quantum calculation has been pushed over into the α_R region of space. Relative energies for some of the low-lying minima in this and other popular potentials are shown in the bottom part of Table I. In addition, Roterman et al.[35] have reported $\phi - \psi$ maps for the ECEPP/2, AMBER all-atom, and CHARMM19 potentials with a variety of assumptions about electrostatic interactions and the extent of flexibility allowed in the minimization procedure.

One measure of the complexity of this seemingly simple problem is the variation in nominally identical calculations reported by different groups. This is partially due to differences in minimization criteria and procedures, and partially due to ambiguities in the literature (and in computing codes) about what the "default" or "standard" parameters are, and about what names should be applied to particular combinations of parameters. In addition, some published force fields have not been explicit about how to handle the acetyl and *N*-methyl blocking groups found in the dipeptide. Nevertheless, there are some general trends that appear that are worth noting.

(1) One significant difference between the ECEPP potentials[35,38] and the other results has to do with the relative energy of the C_7^{ax} conformer relative to its less crowded analogue C_7^{eq}. The high value of C_7^{ax} in the ECEPP potential appears to be largely a result of its assumptions of rigid bond lengths and angles: for example, the AMBER all-atom potential gives a value of 4.5 kcal/mol for the $C_7^{eq} - C_7^{ax}$ difference if a pseudorigid geometry of internal coordinates is maintained.[35] Roterman et al. have argued that the low energies of the C_7^{ax} conformers in the non-ECEPP force fields represents a flaw in these potentials, in that they allow too much variability in bond angles, especially in opening up the C′–N–C$^\alpha$ bond angle in the C_7^{ax} region.[35] As mentioned above, however, quantum mechanical calculations support the notion that in the gas phase this energy gap is relatively small.

(2) The 1984 and 1986 AMBER potentials (with united CH groups and all atoms, respectively) predict too large a penalty for the extended C_5 configuration, but give a reasonable energy for α_R. Addition of explicit torsion terms in newer versions of these potentials leads to a more accurate reproduction of the quantum results.[39] The CHARMM and AMBER/OPLS[40] give lower (and presumably better) estimates for this difference.

Table II. Relative Energies in the Alanine Dipeptide in Water[a]

authors	ref	method	α_R	P_{II}	C_7^{eq}	C_7^{ax}	C_5
		Relative Solvation Free Energies in Water					
Pettitt and Karplus	43	ext. RISM	0.0	−1.5	9.9	9.2	5.8
			(−68,−56)	(−56,−171)	(−66,70)	(63,−69)	(−177,180)
Mezei et al.	17	Monte Carlo	0.0	0.4	3.6		
			(−70,−50)	(−80,150)	(−90,90)		
Tobias and Brooks	44	MD-free energy	0.0	8.8		10.0	
			(−80,−60)	(−80,120)		(60,−80)	
Jean-Charles et al.	45	MD-FEP	0.0		7.8		8.6
	45	PB/OPLS	0.0		7.2		8.2
Still et al.	46	GB	0.0		6.7		6.5
Ösapay et al.	*b*	PB/CHARMM19	0.0		6.2	6.2	4.9
			(−70,−40)		(−80,90)	(60,−70)	(−150,160)
Ösapay et al.	*b*	PB/CHARMM22	0.0		5.2	5.3	4.2
			(−90,−40)		(−90,70)	(70,−60)	(−150,160)
		Relative Free Energies in Water					
Pettitt and Karplus	43	ext. RISM	0.0	−1.8	−1.3	−1.8	−1.7
			(−68,−56)	(−56,−171)	(−66,70)	(63,−69)	(−177,180)
Tobias and Brooks	44	MD-free energy	0.0	−0.2		3.3	
			(−80,−60)	(−80,120)		(60,−80)	
Anderson and Hermans[c]	49	MD-free energy	0.0	−1.5		1.2	
			(−90,−40)	(−100,120)		(80,−80)	

[a] Free energies relative to α_R are given in kcal/mol; values in parenthesis are for the (ϕ,ψ) dihedral angles. Abbreviations are as follows: MD, molecular dynamics; FEP, free-energy perturbation; PB, Poisson–Boltzmann; GB, generalized Born equation; OPLS, optimized potential for liquid systems; CHARMM, chemistry at Harvard using molecular mechanics; RISM, reduced interaction-site method. [b] K. Ösapay, D. Bashford, D. A. Case, unpublished material. [c] Angles for the Anderson and Hermans results are approximate centers of the corresponding low-energy regions; see the text.

(3) As illustrated in Figure 1, gas-phase potential energy surfaces often fail to show a local minimum in the α_R region of conformational space even though this region is highly populated in protein structures. Even though there is no local minimum, it is clear from Figure 1 that structures in this region are roughly 5 kcal/mol above the C_7^{eq} region, and this difference agrees reasonably well with estimates from empirical force fields, as shown in Table I. An exception is the ECEPP results, which predict the α_R conformer to be only slightly disfavored. As we discuss below, solvation contributions stabilize the α_R region relative to other portions of the surface, and it seems likely that the ECEPP results should be viewed more as an effective potential for proteins than as an estimate of the bare gas-phase surface. The CHARMM19 force field gives a high penalty to the α_R conformation, but this is lowered in a more recent parameterization from the Harvard group.

3.4. Estimates of Gas-Phase Free Energies

In addition to intramolecular energies, more information is needed to estimate gas-phase equilibrium populations. This includes contributions to free energies arising from differences in zero-point energy and entropy between conformers. An estimate of these (using the harmonic oscillator–rigid rotor approximation and the AMBER united-atom force field)[29] shows that these corrections favor the C_5 and α_R conformations (relative to C_7) by about 1.1 kcal/mol. Pettitt and Karplus[41] arrived at a value of 0.6 kcal/mol for the vibrational contribution to this difference, using a different empirical force field. These differences arise primarily because the C_7 conformers have the strongest internal hydrogen bond, and hence the least internal flexibility. These estimated corrections are based on empirical potential energy functions, but it is likely that the (gas-phase) corrections would be similar for other energy surfaces as well.

3.5. Solvent Effects

Solvation is known to appreciably influence the equilibrium behavior of peptides in solution, and the calculations discussed above do not include such effects, except to the extent that empirical fitting of parameters in the molecular mechanics force fields represents in some average way experimental data derived from solution measurements. It is only recently that theoretical methods with secure foundations have been available to estimate solvent effects on conformational equilibria, and it is clear that the level of confidence that can be placed in these theoretical estimates is much lower than in the gas-phase results discussed above. Nevertheless, two general features are almost universally observed in such studies when water is the solvent. First, the α_R region of conformation space is significantly stabilized relative to the extended forms, a result that can qualitatively be explained in terms of peptide group dipole moments, as we discuss below. Second, the low-energy portions of the surface become much broader, so that (in most calculations) the C_7^{eq} and C_5 minima found in the gas phase merge into a single broad well in the β region of the map. Table II collects some results from the literature that are discussed in the following paragraphs.

3.5.1. Estimates of Relative Solvation Free Energies

Several techniques have been used to estimate the relative solvation energies of conformers of the alanine dipeptide, and results are summarized in the top part of Table II. Early calculations used the extended RISM integral equation approach and Monte Carlo free-energy simulations.[42,43] More recently, Tobias and Brooks have used the molecular dynamics-based free-energy methods discussed above to calculate relative hydration free energies for several states on the surface.[44] The Poisson–Boltzmann (PB) results shown in Table II[45] use a continuum electrostatic model to estimate the

electrostatic energy required to move the molecule from vacuum into a dielectric of 80. This molecular description employs partial charges taken from molecular mechanics potentials (OPLS or CHARMM) and uses a solvent-accessible surface based on empirical radii to create the division between the low-dielectric region used for the molecular interior and the high-dielectric region appropriate to the solvent. The generalized Born theory works with the same dielectric model, but uses an approximate equation, rather than a numerical solution, to estimate the electrostatic contribution to the solvation energy.[46] Finally, free-energy perturbation calculations have been reported in which the free-energy change computed is for changing the partial charges from zero to those in the OPLS force field. Both these calculations and those employing the Poisson–Boltzmann equation estimate only the electrostatic portion of the hydration energy. It is plausible to assume that the nonelectrostatic component, which involves nonpolar interactions between a molecule with no internal partial charges and solvent, will be reasonably independent of conformation, so that differences in electrostatic energy will mimic differences in total solvation energy; this assumption has not yet been directly tested for the alanine dipeptide.

Much of the conformational dependence of solvation energies seen in Table II can be rationalized by considering the relative orientations of the two peptide dipole groups as a function of the ϕ and ψ torsion angles. Along the "antidiagonal" of the Ramachandran plot, where ϕ and ψ are about equal in magnitude but opposite in sign, the two peptide dipoles approximately cancel, so that the net molecular dipole moment is small. On the other hand, in regions where $\phi + \psi$ is about $\pm\pi$ the dipoles are in phase and yield a large net molecular dipole moment. Hence, the α_R and α_L regions tend to show more favorable solvation energies than do C_7^{eq}, C_7^{ax}, and C_5, all of which have $\phi \approx -\psi$. Estimated differences between α_R and these latter three conformers range from 4 to 10 kcal/mol in the calculations reported in Table II. These differences arise both from variations in the force field parameters (especially for the partial charges) and from differences in the computational methods used. Unfortunately, there are no secure values for the "correct" answer, and the wide range of predicted values suggests that it is not yet clear which of the theoretical estimates is the most reliable.

Estimates for the solvation free energy of the "P_{II}" conformer, relative to α_R appear especially variable in Table II, but this may be a consequence of the different definitions of the geometry for this conformer. Note that the greatest predicted stabilization (for the extended RISM calculation) reflects a conformation for which ϕ and ψ are quite different in magnitude, whereas the least stable of the conformers labeled P_{II} (using MD free-energy techniques) is for a conformer much closer to the unfavorable region where dipoles nearly cancel.

The calculations shown in Table II all assume that the charge distribution is independent of conformation and is unaffected by solvation. Grant et al. have examined this assumption by using a quantum mechanical description of the charge density, with and without the effects of a reaction field that arises from placing the molecule in a high dielectric continuum.[47] The results suggest that polarization of the solute by the reaction field can be substantial; for example, the α_R conformation has a gas-phase dipole moment of 8.2 D in this model, which increases to 10.6 D in the presence of solvent, an effect that can have a substantial impact on predictions of solvation energetics. Further, these dipole moment changes are conformation dependent: the C_7^{eq} conformation, for example, has a smaller dipole moment change (from 3.1 to 3.5 D) upon solvation. Grant et al. report only estimates of solvation energies (rather than free energies), so their results cannot be directly compared to those in Table II. The general conclusions, however, about the stabilization of α_R and P_{II} conformations relative to C_7^{eq} and C_5 parallel those discussed above. The results do suggest that careful evaluations of solvation energies will require attention to the effects of solute polarization and its conformational dependence.

In related work, Sharp has studied the conformation dependence of solvation energies for the alanine dipeptide using a DIEMOND model in which the reaction field arising from solvent is represented as a surface charge density at the solvent-accessible surface.[48] Each piece of this density is assigned to the nearest nucleus, effectively leading to a set of molecular partial charges that varies with molecular conformation. The interaction of the corresponding reaction field with the nuclear charges also provides an estimate of solvation free energies. As with the calculations listed in Table II, this study showed the α_R region to be stabilized relative to C_7^{eq}, but the latter remained the global minimum, and the broadening of the conformational region in the β region was not as pronounced as in other calculations. The effective charges in a β conformation were on average about 12% different from those in the α_R conformation.

3.5.2. Estimates of Free-Energy Differences in Solution

By using the molecular dynamics with free-energy methods discussed above, as well as the extended RISM integral equation approach, comparable free-energy differences for conformations of the alanine dipeptide emerge from the work of Anderson and Hermans,[49] Tobias and Brooks,[44] and Pettitt and Karplus.[42,43] These results are summarized in the bottom portion of Table II. Each of these groups used different models for the peptide internal energetics, water interactions, and peptide–water interactions. Such model differences most likely contribute to the difference in absolute numbers for conformational free energies. However, similar trends exist. For example, the P_{II} conformation is lowest in free energy in the calculations from all studies. Furthermore, the $P_{II} - C_7^{ax}$ free-energy difference favors the P_{II} conformation by 2.8 and 3.5 kcal/mol (differing by only 0.7 kcal/mol) in the calculations of Andersen and Hermans and Tobias and Brooks, respectively. Pettitt and Karplus, however, find that the P_{II} and C_7^{ax} conformations have very similar free energies. This is most likely a reflection of the underlying gas-phase energies used: this early potential had only a 0.3 kcal/mol gap between C_7^{eq} and C_7^{ax} in the gas phase. We also note the α_L conformation is disfavored over the C_7^{ax} by 0.5, –0.3, and 1.1 kcal/mol in calculations from Brooks and Tobias, Anderson and Hermans, and Pettitt and Karplus, respectively. Thus,

calculations from these three groups illustrate that solvation preferentially stabilizes the α and extended "β" conformations (relative to the predominant gas-phase minima at the C_7 and C_5 conformations).

4. Thermodynamics and Kinetics of Conformational Transitions in Peptides

The demonstration that stable populations of secondary structure in small peptides can be identified and studied has motivated several molecular dynamics simulations performed to explore the time scale, mechanisms, and thermodynamics of folding/unfolding of peptides in aqueous solution. The simulation studies have focused on helical peptides, chain reversals or turn forming peptides, and β-sheet structural elements. Peptide unfolding to varying degrees was observed in all of these calculations, and a number of observations regarding the mechanism of unfolding have been made. More general aspects of thermodynamic stability, sequence influences in conformational preference and, in some instances, kinetic "processing" of peptides can be studied using techniques of conformational and mutational free-energy methods. In this section, we review progress in calculations aimed at using these methods to explore—correlate, rationalize, and predict—trends in peptide conformational stability and kinetics.

Both kinetic and structural models of protein folding involve the early formation of locally stable structural elements which then rearrange and associate to form compact tertiary structures.[50-52] The formation of the secondary structural elements itself is believed to be fast, and the further rearrangement of these involve the slow, rate-determining steps of protein folding. Hence it is of interest to study the initial formation of helical structures and their stability in aqueous solution. Interest in this subject has been heightened by recent experimental studies (primarily using circular dichroism and NMR spectroscopies) which indicate that relatively short linear peptides may have substantial helical character in aqueous solution[52-60] and that novel host–guest experiments are possible (based on such peptides) that yield new insights into sequence preferences for helix stabilization.[61-63] In spite of this wealth of empirical information, the origins of this sequence specificity are not well understood, nor is there much good information about the mechanism(s) by which local pieces of secondary structure are formed and destroyed.

Two general strategies have been adopted to explore the sequence dependence of secondary structure preferences in linear peptides. The first considers systematic variation of synthetic, generally repetitive, sequences in which relative helical stabilities can be measured,[54,62-66] or estimated from dynamics simulations.[67-70] A second strategy studies peptides having sequences of regions of secondary structure in particular proteins, in an attempt to understand features that lead to preferences for secondary structure in very nonrepetitive sequences. This second approach is illustrated by recent NMR studies on peptides derived from myoglobin, ribonuclease, plastocyanin, and myohemerythrin,[53,55,57,71,72] as well as by theoretical simulations on the ribonuclease and myoglobin peptides.[73-75] These simulations illustrate that, even within a largely

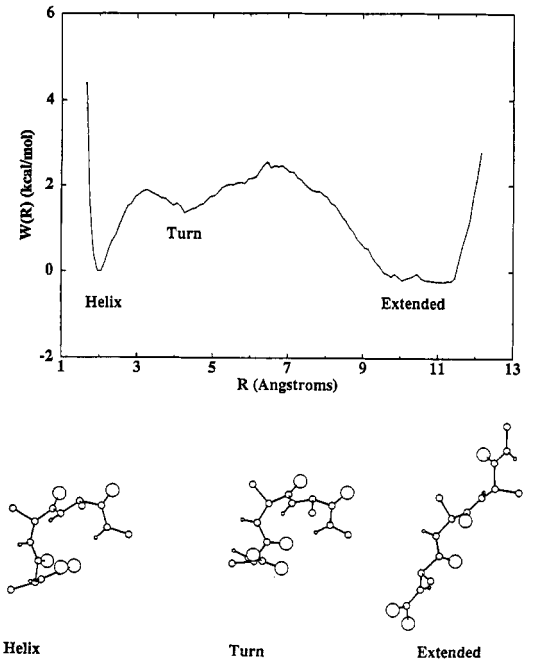


Figure 2. Free-energy surface for the formation of an α helix in Ac-$(Ala)_3$-NHMe, plotted as a function of R, the distance between O_1 and H_5 atoms involved in helix hydrogen bonding. The lower portion depicts conformations typical of helical, turn, and extended configurations. Adapted from ref 76.

helical peptide, there are often significant deviations from ideal geometries, and that helical hydrogen bonds can break and spontaneously re-form without requiring any "global" breakdown of the helix. The free-energy barriers between the helical hydrogen-bonded state and unfolded alternatives are apparently low enough to allow several such transitions to take place within a nanosecond of simulation.

4.1. Helix/Coil Thermodynamics and Kinetics

Simulation studies aimed at exploring the thermodynamics of helix formation have been performed by a number of groups using both conformational free-energy methods and "chemical" perturbations utilizing thermodynamic cycles. It has been found that the breakup of individual hydrogen bonds occurs through the "toggling" between $(i, i+4)$ and $(i, i+3)$ interactions before proceeding to a more extended conformation in which water strongly solvates the hydrogen-bond donor and acceptor groups. The metastable nature of chain reversal structures on the pathway of helix formation/dissolution is illustrated in Figure 2 for the folding of a simple blocked alanine tripeptide, which is of sufficient length to form one helical hydrogen bond. These results are from calculations of Tobias and Brooks.[76] In the top portion of the figure the underlying free-energy surface is shown. This surface suggests the presence of three distinct free-energy minima on the folding pathway for helical hydrogen bond formation: the helical manifold of states centered around $r = 2$ Å, the chain reversal states near $r = 5$ Å, the extended manifold beyond $r = 7$ Å. The lower portion of Figure 2 illustrates the nature of the structures which occur in the manifold of states described by the three free-energy minima. The observation that helix unfolding occurs through turn structures is in agreement with the findings of Sundaralingam and Sekharudu[85] based on a survey of

Table III. Kinetic Parameters, ΔA, for Helix Unfolding[a]

authors	temperature	helix → nonhelix	nonhelix → helix
Tobias and Brooks[b]	298	1.9	2.7
Daggett and Levitt[c]	278	−2	0.6
	298	−2	0.8
	323	−2	1.0
	373	−2	1.4
	423	−2	1.8
	473	−2	2.3

[a] Free-energy barriers for helical/nonhelical transitions, in kcal/mol; temperature in kelvins. [b] Tobias and Brooks (ref 76) also find a free-energy minimum for turnlike intermediates. Their barrier heights are: helix → turn, 1.9; turn → helix, 0.6; turn → coil, 1.2; and coil → turn, 2.7 kcal/mol. [c] Reference 67.

solvated helical structures in well refined protein structures. It is also evident from molecular dynamics studies, as described below, that this mechanism of helix unfolding is seen in longer peptides.

From simulation studies as those described above, information regarding kinetics (barrier crossing) can also be obtained. In the studies of Tobias and Brooks,[76] barrier heights have been used to estimate time scales for transitions between helical and turn states as well as for turn to extended states in simple alanine and valine helices. Studies of helix unfolding for alanine helices by Daggett and Levitt have also been used to investigate this question.[67] In Table III, we collect results from calculations of these sets of studies. In calculations by Tobias and Brooks, thermodynamically derived barrier heights are used to estimate rates for local helix unfolding and folding, i.e., a single hydrogen bond. In calculations of Daggett and Levitt, molecular dynamics simulations of a longer alanine helix are carried out at several temperatures in vacuum and solution. Here we focus on their simulations in aqueous solution. They estimate kinetic rates and barrier heights by simply counting the number of transitions occurring in their simulations at various temperatures. They use van't Hoff relationships to obtain thermodynamic information. The agreement between their findings and those of Tobias and Brooks, for both the unfolding barrier and the refolding barrier, are remarkably good if one assumes that the full barrier for refolding is seen only at higher temperatures, which takes the local helix from turnlike conformations, over the barrier, to more extended conformations. Their calculations indicate that helix unfolding proceeds through states with increasing extended character as temperature increases, yielding the more extended conformations studied by Tobias and Brooks as the "unfolded" state for simulations at the highest temperatures.

Both the explicit unfolding studies and the thermodynamic calculations indicate that the time scale associated with breakup of individual helical hydrogen bonds is on the order of a hundred picoseconds under ambient conditions for helices composed of a single amino acid, i.e., alanine or valine, and refolding is significantly longer.

Free-energy simulations on the initiation and propagation of folding in model valine, alanine, and proline helices,[76,81,82] as well as studies of changes in stability of longer helices due to amino acid substitutions (mutational free-energy calculations) by Hermans and coworkers,[68,69,77] approach the issue of sequence specific effects in helix stability. Tobias and Brooks find that valine-rich helices destabilize helix initiation over alanine-based analogues through a favorable solvation component in the intermediate turn state. Entropic factors from side-chain "immobilization" in the helical state most likely also contribute to the destabilization.[77,78] The studies by Tobias and Brooks are unable to discern this contribution.

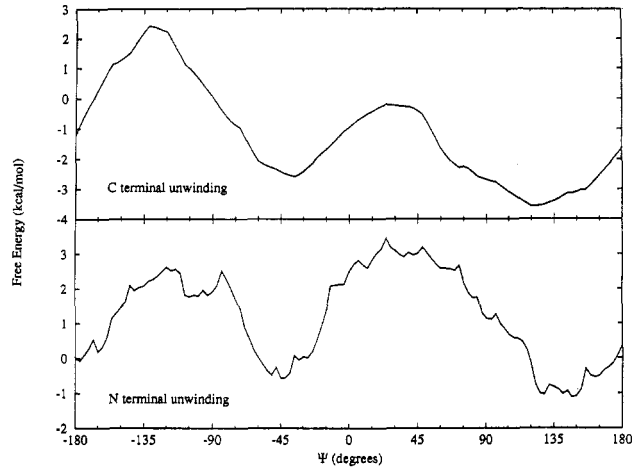


Figure 3. Free-energy surfaces for winding/unwinding the ends of an α helix in Ac-(Ala)$_4$-NHMe, plotted as functions of ψ_2 (N-terminal unwinding) and ψ_5 (C-terminal unwinding), in water. Adapted from ref 76.

To address questions beyond those of helix initiation, Tobias et al. have examined the free-energy surface (pmf) for unfolding one-turn of an α helix which has one other intact 1–5 helical hydrogen bond.[79,80] These calculations employed conformational "perturbation" methods of free-energy simulations.[13] The objective here was to contrast the free energies of helix initiation and propagation. This two hydrogen bond helix is the minimum unit which permits such an examination. The potentials of mean force for unwinding one turn of the helix from either the C-terminal on N-terminal end are shown in Figure 3. The free energy as a function of the ψ dihedral angle shows very similar barriers for unwinding one of the turns of the α-helix from an α-helical conformation ($\psi \approx -60$) to an extended conformation ($\psi \approx 180$) for both "unwinding directions". Furthermore, the free-energy difference between the helical structures and the partially unwound ($\psi \approx 180$) is quite similar to that found from helix initiation studies on the "one-turn" alanine helix. This suggests that the propagation parameter, which in some sense measures the cooperativity of adding another hydrogen bond, is close to unity or even possibly less than one for these short peptides. Similar conclusions have been obtained for the local unfolding of alanine helices by Daggett and Levitt; leading to the suggestion that helix initiation and propagation parameters are highly dependent on helix length.

The role of proline in helices is another sequence effect of general interest. Proline is statistically overrepresented in the N1 position of an α helix, but little is known about its mechanistic role in helix stabilization. There are three mechanisms by which an N1 proline could stabilize α helices: (a) stabilize helix initiation, (b) stabilize C-terminal helix propagation, or (c) inhibit

N-terminal propagation. Studies underway by Karpen and Brooks are focusing on these issues. In preliminary work from these researchers, helix initiation in the blocked tripeptide acetyl-Pro-Ala-Ala-NHMe was studied by determining the free-energy difference between the extended conformation and a single turn of a helix, with Pro in the N1 position. The reaction coordinate used was the distance between the carbonyl oxygen of the N-terminal blocking group (O1) and the amide proton of the C-terminal blocking group (H5), analogous to the helix initiation calculations on alanine and valine helices described above. The results from Karpen and Brooks show that proline does not stabilize helix initiation when compared to an all alanine tripeptide. The change in free energy of Pro-Ala-Ala, as its conformation goes from a turn of a helix to an extended conformation, is −3.5 kcal/mol; this change is −0.9 kcal/mol for the Ala-Ala-Ala tripeptide.[81]

C-terminal propagation is also under study by Karpen and Brooks. These calculations have been performed using the blocked tetrapeptide Pro-Ala-Ala-Ala. The peptide was constrained to form one turn of an α helix, forming a hydrogen bond between the carbonyl oxygen of the N-terminal blocking group (O1) and the amide proton of the third alanine (H5). The reaction coordinate used was the backbone ψ angle of the C-terminal alanine, where the residue adopts an extended conformation at values near $\psi = 130$ and an α-helical conformation (extending the helix to a second turn) at values near $\psi = -45$. A full 360° rotation of ψ was used to assess the reversibility of the free-energy simulation. It was found that proline in the N1 position has no statistically discernible effect on C-terminal propagation when compared with the previously studied Ala-Ala-Ala-Ala peptide. Both peptides gave a slight preference for the extended state over the helical state for the C-terminal residue. Thus proline must stabilize helices either by terminating N-terminal propagation or by acting in concert with another residue or side chain, not present in this peptide, to stabilize helix initiation and/or C-terminal helix propagation.[82]

The calculations just described have utilized free-energy methods to explore conformational thermodynamics as they influence conformational "processing". However, they have been focused on small model peptides. The relationships between such model peptides and the longer, and heterogeneous, peptides studied experimentally are yet to be fully established. Nevertheless, such studies are lending insights into time scales and mechanisms which are not currently available from similar studies on longer peptides.

Studies of helix stability by Hermans and co-workers have used thermodynamic cycle free-energy methods to explore the relative stability of helical structures in which blocks of specific amino acids are substituted in longer helices.[69] They have examined several substitutions in the interior and ends of the helix and have found generally good agreement with experimental results. In particular, this work compares the helix propensities, estimated from the free-energy perturbation methods of eq 14–16, with experimental results from the DeGrado and Kallenbach groups.[62,63] The computed trend in helix propensity (α-aminoisobutyric acid, alanine, α-amino-n-butyric acid, valine, D-alanine, glycine, leucine, proline) is the general agreement with the experimental rankings. Hermans and co-workers have also examined the effects on helix initiation and propagation when alanine is replaced by proline at the N1 position.[68] Although this study used a "chemical perturbation" approach, the findings show that proline is favored over alanine by only a small amount (ca. 0.5 kcal/mol) at the N1 position, in agreement with the C-terminal propagation studies of Karpen and Brooks.

4.2. Mechanisms of Folding–Unfolding Transitions in Helices

The tendency noted above for helical $i \rightarrow i + 4$ hydrogen bonds to switch to bifurcated forms or to $i \rightarrow i + 3$ bonds as an intermediate step in helix unwinding has been seen in a variety of simulations, such as on polyalanine helices[70] and in ribonuclease- and myoglobin-based helices.[73–75] For example, the existence of picosecond intermediates involving alternative internal hydrogen bonds can be seen clearly in the transient fraying and refolding of the C-terminal portion of the ribonuclease C-peptide helix of a recently reported simulation.[75] Such events are of interest from the points of view of the unfolding mechanism and of proton exchange with the solvent. For example, a significant deviation from helical structure in this simulation occurs in the stretch 725–795 ps. Figure 4 shows the changing pattern of internal hydrogen bonding and hydration of the backbone. Hydrogen bonds are represented by arrows going from the donor to the acceptor atom, and water molecule numbers are enclosed in circles. The first snapshot shows the hydrogen bonding scheme at 725 ps. All the seven possible $i \leftarrow i + 4$ hydrogen bonds for the stretch of the chain 3–13 are present, and in addition, an $i \leftarrow i + 3$ hydrogen bond, $7 \leftarrow 10$. The backbone N-H of Arg-10 is thus involved in "bifurcated" hydrogen bonding. Out of the six water molecules shown, the four on the left act as donors, forming hydrogen bonds with the backbone carbonyls of residues 5, 6, 8, and 10. The oxygen of W41 acts as an acceptor in a hydrogen bond with residue 14 (the C-terminal amide group). W410 is inserted into the $11 \leftarrow 14$ hydrogen bond, forming a bridged structure. Such bifurcated $\alpha/3_{10}$-helical and water-inserted hydrogen bonds are believed to be common intermediates in the formation/unwinding of α helices, as discussed above. At 735 ps the bifurcated hydrogen bonds have shifted two residues down the chain. A second water (W582) then bridges the NH and CO sides of the helix, and by 765 ps all of the internal hydrogen bonds are broken at the C-terminal end, and the NH and CO groups are well solvated. A spontaneous near-reversal of this unwinding then takes place between 765 and 795 ps, again involving several transient $i \leftarrow i + 3$ hydrogen bonds. Structural snapshots of this local fraying and refolding are given in Figure 5. These show that the breaking of amide hydrogen bonds (in the third and fourth snapshots) is associated with an unwinding of the helical backbone, drawn as a four-stranded ribbon.

This sequence of snapshots has captured a number of intermediate structures which might be observed in the unfolding/refolding of α-helical structures in water. The key role played by hydration in reversible unfolding is clear from these diagrams. It is instructive to note that the most hydrated structures have the most distortion from an α-helix: e.g., the snapshot at 765 ps

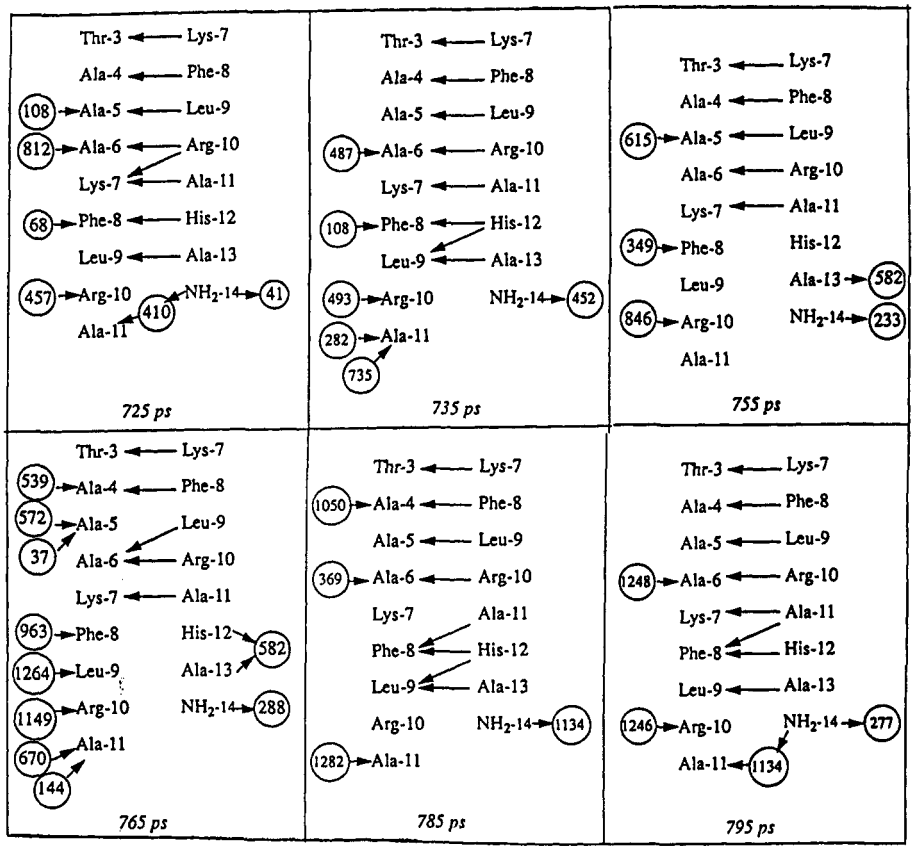


Figure 4. Evolution of backbone hydration and helical hydrogen bonding during a period of simulation of an analogue of the ribonuclease C-peptide. Residue numbers of water molecules are shown in circles, and hydrogen bonds are indicated by arrows. Potential carbonyl acceptors are shown in the left column, and potential NH donors are shown at the right. Adapted from ref 75.

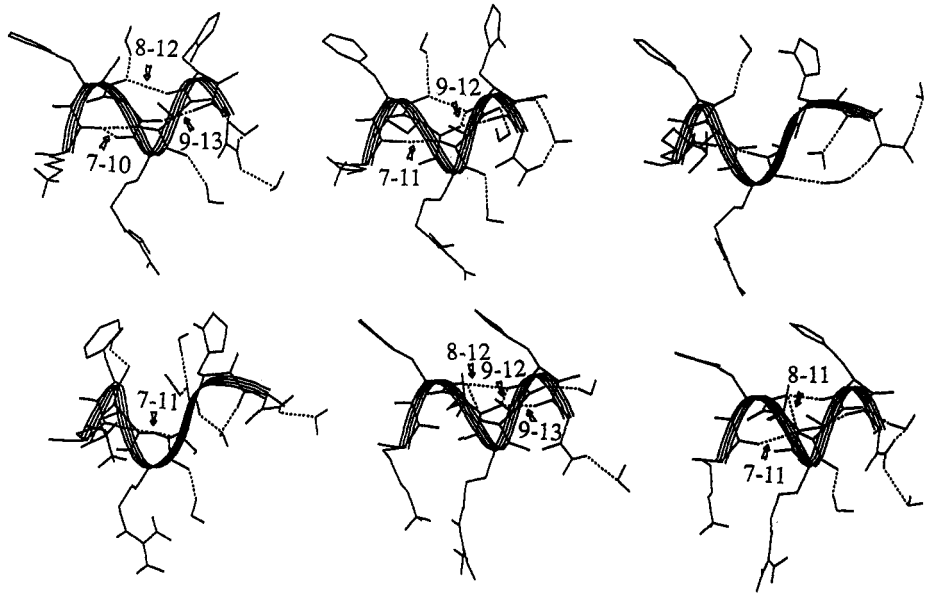


Figure 5. Structural snapshots corresponding to configurations of Figure 4. The numbers identify residues involved in α or 3_{10} hydrogen bonds.

which has 10 water molecules hydrogen bonded to the peptide backbone has the largest deviation from α-helical conformation (see Figure 5). The intermediates observed here are rather the same type as the ones in a recent myoglobin H-helix simulation (3_{10}, bifurcated, and water-inserted hydrogen bonds), where the helix unfolds almost completely.[74] In spite of the complexity of shifting hydrogen bond patterns and hydration, these local helix unwinding events are apparently quite facile reactions, with relatively low barriers between helical and coil states. Daggett and Levitt[67] and Tobias et al.[79,80] estimate the activation free energy for a local helix → coil transition to be about 2 kcal/mol for polyalanine, and this estimate is in rough accord with the frequency of such transitions seen in other simulations (cf. the discussion above). Water appears to act as a denaturant, stabilizing the coil form relative to helix, but the kinetics of the helix → coil

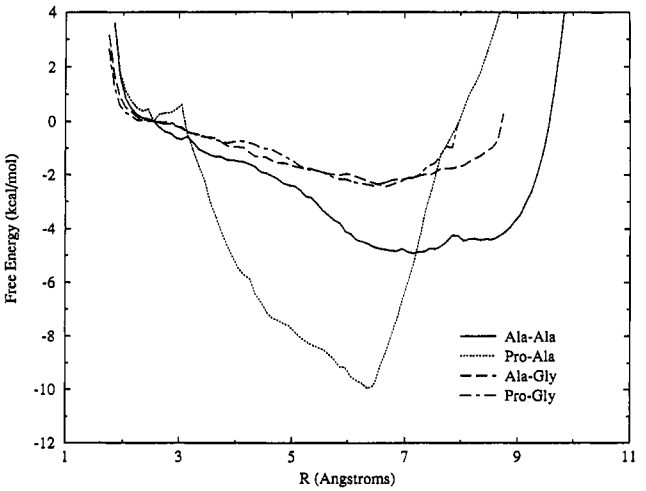


Figure 6. Free-energy surfaces as functions of the reaction coordinate R for type I reverse-turn folding/unfolding in Ac-Ala-Ala-NHMe (solid curve), Ac-Pro-Ala-NHMe (dotted curve), Ac-Ala-Gly-NHMe (dashed curve), and Ac-Pro-Gly-NHMe (dashed–dot curve) in water. The reaction coordinate is the distance between O_1 and H_4, the atoms involved in turn hydrogen bonding. Adapted from ref 88.

motion appears to be approximately 10^{10} s^{-1} in both vacuum and water.

The basic features of the unfolding events described above involve the insertion of water into helical hydrogen bonds and transient intermediates that have bifurcated or 3_{10}-like hydrogen bonds. These ideas have surfaced in other studies. Turnlike intermediates are seen as metastable states on helix-unfolding potentials of mean force, as discussed above. In addition, Czerminski and Elber have systematically searched for low-energy paths for forming helices using a vacuum model for the potential energy, and found intermediates that involve $i \leftarrow i + 3$ hydrogen bonds as in turns and 3_{10} helices.[83,84] Finally, peptide crystal structures have also been interpreted as providing evidence for the stability of intermediates like those seen in simulations.[85-87] It thus seems likely that this is a general feature of the formation and breaking of α helices.

4.3. Thermodynamics of Turn Formation

Reverse turns have been implicated in a causal fashion to influence folding initiation in proteins. These structural elements also form stable structures in peptide models in isolation. Studies which attempt to gain insight into factors determining the stability of chain reversals have been performed for a number of short peptide models. These calculations have utilized umbrella sampling and conformational "perturbation" methods to study the formation of simple type I turns, as well as the conversion of type I to type II turns, for blocked dipeptides of sequence X–Y, with X being Ala or Pro and Y being Ala or Gly, in aqueous solution.

The resulting potentials of mean force (pmfs) for the formation of turns in the sequences Ala-Ala, Pro-Ala, Ala-Gly, and Pro-Gly are shown in Figure 6. The calculations presented here are from work of Brooks and co-workers.[88,89] The first feature that emerges from these calculations is that all peptides are "intrinsically" unstable as turns in aqueous solution. These researchers conclude that, for the type I turns studied, significant contributions to stability must come from explicit side-chain contacts which are not explicitly represented in their studies. In addition, from a comparison between the peptides containing Ala in the second position to those with Gly, it appears that Gly dramatically destabilizes the manifold of extended states (at longer ranges, 5–9 Å) relative to the folded conformations at 2 Å (or conversely stabilizes the turn), thereby making the turn structures for the Gly-containing peptides more stable. The free-energy difference between turn and extended regions for both of the Gly-containing peptides is on the order of 3 kcal/mol favoring extended. This is to be compared with 5–10 kcal/mol for the Ala-containing peptides.

Table IV. Decomposition of Free-Energy Differences between Ideal Type II and Type I Turns[a]

peptide	$\Delta A \equiv A(II) - A(I)$	ΔU_{uu}	ΔU_{uv}	$-T\,\Delta S_{c,uv}$
Ala-Ala	3.3 ± 0.5	−3.6 ± 4.0	14.1 ± 9.8	−7.2 ± 11
Ala-Gly	−3.2 ± 0.5	−7.1 ± 3.8	6.9 ● 8.5	−3.0 ± 9.3
Pro-Gly	0.2 ± 0.5	−5.9 ± 4.0	5.9 ± 7.8	0.2 ± 8.8

[a] Energies in kcal/mol. Tabulated uncertainties were determined by error propagation from standard deviations in the averages. Data is from ref 79.

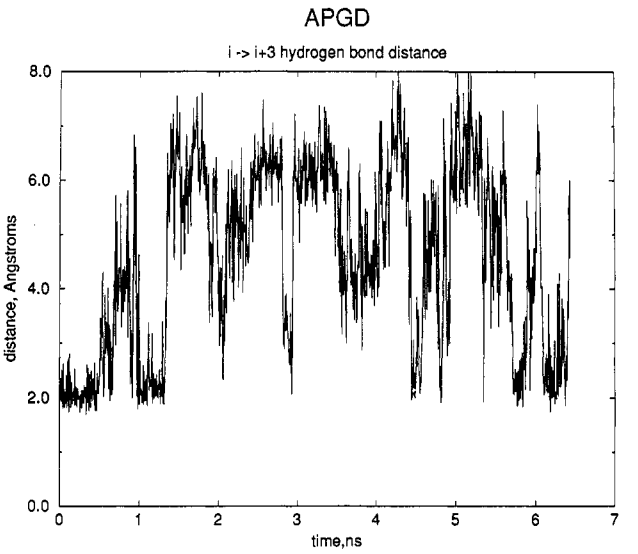


Figure 7. Turn hydrogen bond distance (from the O atom of alanine to the H atom of aspartate) in a solvated simulation of Ac-APGD-NHMe. Adapted and extended from ref 90.

The stability of type II turns for three of these peptides, Ala-Ala, Ala-Gly, and Pro-Gly, has also been examined. The findings of Tobias et al. for type I versus type II turn stability are summarized in Table IV, together with the thermodynamic decomposition of these free energies into solvation and intrapeptide components.[79] Results from these calculations indicate that the Ala-Gly blocked dipeptide is marginally stable in the type II turn conformation. However, all turns examined remain intrinsically unstable, or possess only marginal stability.

Some of the ideas outlined above can also be discerned in unbiased molecular dynamics simulations of short peptides in water. Figure 7 shows results of a 6.5-ns simulation of Ac-APGD-NHMe, showing the fluctuation of the $i \leftarrow i + 3$ hydrogen bond length.[90] In this particular run, the turn hydrogen bond breaks and reforms every few hundred picoseconds. There is about equal time spent in the turn and in extended forms, and the barrier between the two is evidently quite small. Potentials of mean force evaluated using umbrella

sampling suggest that the extended states are actually somewhat more populated at equilibrium than they are in the time period shown in Figure 7, so that the free-energy profile is much like that shown in Figure 6. As with helices, the "turn" structures seen often involve bifurcated hydrogen bonds in which the carbonyl of residue 1 is in close contact with the NH groups of both residues 4 and 5.

Recently, a simulation of similar length has been carried out for Ac-AYPYD-NHMe, with a *cis* peptide linkage at the proline position (D.A. Case, unpublished). Sequences of this motif exhibit high turn populations (>50%) as measured by NMR.[60] As with APGD, spontaneous breaking and reforming of the turn conformation is observed, with the "folded" conformers typically showing bifurcated hydrogen bonds. The time scale for making and breaking the turn is considerably slower here, however, with the turns existing for more than 2 ns on two separate occasions. Free-energy surface estimates are underway, but the preliminary results from the unbiased simulations suggest that this sequence can form type VI turns more readily than the sequences discussed above form type I or type II turns. The NMR experiments suggest that the aromatic groups adjacent to the proline promote turn stability, and it will be of interest to see if such trends can be reproduced in these microscopic simulations. Similar effects have been seen in explicit dynamics simulations of YPGDV, which point to the importance of small apolar clusters involving the Tyr side chain and peptide backbone as well as the Pro and Val portions of the molecule.[91,92]

4.4. Thermodynamics of Formation for a Model β Sheet

β sheets represent another common, repetitive secondary structural motif present in folded proteins.[93] To develop a quantitative understanding of the contribution of β sheets to the stability of folded proteins and folding intermediates and to investigate the possible role of small sections of β sheets as folding initiation structures, Tobias et al. investigated a simple model β sheet formed by two "alanine dipeptides" (blocked alanine residues, e.g. Ac-Ala-NHMe), with two hydrogen bonds between the peptides in the closely spaced antiparallel arrangement.[94] The reaction coordinate for the formation of the β sheet was defined as the H2···O5 and O2···H5 distances, i.e., the cross-strand hydrogen-bond distances, which were constrained to be equal and fixed within each window of their simulations. To sample the free energy along a path which increased these distances, they used the thermodynamic perturbation method with seven equally spaced windows to sample the range, $1.6 \leq r \leq 5.1$ Å.

In Figure 8, the free-energy surface, as a function of this reaction coordinate, for formation of the model β sheet in water is shown along with the free energy for forming a single hydrogen-bonded species between two formamide molecules ("half" of the β-sheet model).[95] The two surfaces have been placed so that the free energy at the minimum corresponding to the hydrogen bonded species (at an OH separation of ~2 Å) equals the binding free energy in water (5.5 kcal/mol for the sheet and 0.34 kcal/mol for formamide dimer), e.g. the free energy of the infinitely separated pair molecules in water is defined as zero. The overall shapes of the

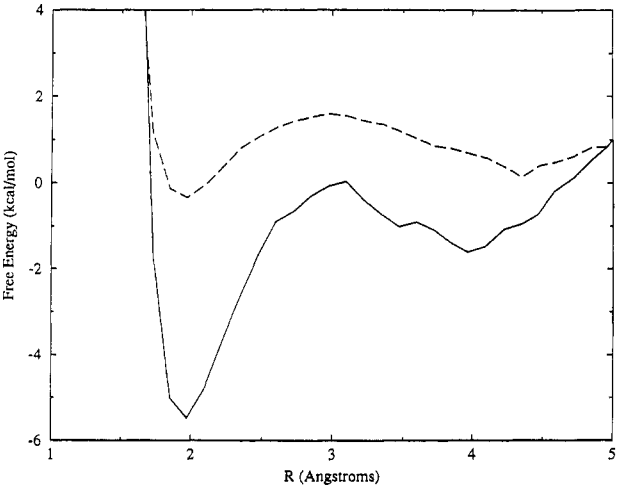


Figure 8. Free-energy surfaces for the formation of the model β sheet with closely space hydrogen bonds between two alanine dipeptide molecules in water (solid curve) and the formation of a linear amide hydrogen bond between two formamide molecules in water (dashed curve). Adapted from ref 94.

two free-energy surfaces in Figure 8 are similar. Each has two minima corresponding to stable hydrogen-bonded and solvent-separated species. The hydrogen-bonded species is the most stable on both surfaces. In addition, on each surface there is a small barrier to formation of the hydrogen bonded species from the solvent-separated species. One obvious difference is noted in comparing these two surfaces, the hydrogen-bonded species is thermodynamically much more stable than the solvent-separated species on the β-sheet free-energy surface. Moreover, since the barrier to breaking the hydrogen bond is much greater for the sheet, the sheet is also kinetically more stable.

It is worth noting that similar calculations by Jorgensen on the *N*-methylformamide dimer in water and chloroform suggest a similar picture of significantly reduced "hydrogen bond" strength in aqueous solution.[96] These studies suggest, however, that the preferred mode of association in water (in the absence of any geometric constraints from a protein environment) is a stacked configuration, with more linear hydrogen bonds being formed in less polar solvents.

By using transition state theory and the heights of the free-energy barriers in Figure 8, an estimate of the time scales for the formation of the hydrogen bonded species from the solvent-separated species can be made. They are 1 ps for formamide and 14 ps for the dipeptide. The time scales for the reverse processes, formation of the solvent-separated species from the hydrogen bonded species, are 24 ps for formamide and 10 ns for the dipeptide. Thus, these calculations predict that amide hydrogen bonds between the formamide molecules are rapidly formed and broken in water whereas the hydrogen bonds between the peptide molecules in the model β sheet are rapidly formed but relatively long-lived in water.

The fact that the model β sheet, which has two linear amide hydrogen bonds, is significantly more stable than the formamide dimer, which has one, was also examined by Tobias et al.[94] The thermodynamic decompositions of the free-energy differences between hydrogen-bonded and solvent-separated species shed some light on why this is so. Accompanying the transition from the

hydrogen-bonded to the solvent-separated arrangement is a large, positive change in the peptide–peptide energy (due to a loss of peptide–peptide hydrogen bonds), a large, negative change in the peptide–solvent energy (due to a gain of peptide–water hydrogen bonds), and a large, positive change in the peptide–solvent entropy (because the solvent-separated species binds more water molecules) for both the sheet and the formamide dimer. The changes in the peptide–peptide energy and the entropy are roughly twice as large for the sheet as they are for the formamide dimer. However, the magnitude of the peptide–water energy difference for the sheet is less than twice (by about 3.5 kcal/mol) that for the formamide dimer. Evidently, the presence of the side chains prevents the amide groups in the sheet from being solvated as favorably in the separated arrangement as in the formamide dimer, where the amide groups are completely exposed to the solvent. This interesting observation may be overestimated due to the simplified model of the single hydrogen bond, but may also reflect real differences between hydrogen-bonding interactions in sheets versus relatively exposed hydrogen bonds. It would be interesting to see what happens for other geometries of β sheets.

5. Summary

In this review, we have discussed a range of applications of molecular simulation methods to problems of peptide conformational dynamics and thermodynamics, along with a brief overview of the methods themselves. It is reasonable to ask what general conclusions can be drawn from the applications discussed in this review. Many interesting ideas are contained in the original papers that have not been discussed here, but the following observations consider some of the more general points we have found in preparing this article.

(1) Quantum chemistry calculations that use large basis sets and include some description of electron correlation give a reasonably consistent picture of the nature of the (gas-phase) potential-energy surface for small dipeptide models, and recent empirical force fields reproduce this behavior with good fidelity. Uncertainties in relative conformational energies are probably still larger than room temperature thermal energies, so that calculations that involve Boltzmann averages over various conformations must still be interpreted with caution. There is also a large literature, not discussed here, on empirical and quantum chemistry calibrations of force fields for amino acid side chains. The possible intrinsic errors in the resulting potentials are harder to evaluate, but may be substantial, particularly for polar and charged groups.

(2) Estimates of solvation contributions to conformational energy differences are more difficult to make, both because it is difficult to find direct and detailed experimental tests and because various theoretical estimates give somewhat disparate results. The general features at the dipeptide level (stabilization of the α_R conformation and broadening and merging of the C_7^{eq} and C_5 conformations) are reproduced with many models.

(3) Free-energy calculations have become a powerful tool for mapping out conformational thermodynamics in a series of interesting models for helices, turns, and sheets. Although there are still some unresolved questions about the level of statisticial convergence, results using different approaches appear to be in good qualitative accord. Theoretical estimates of helical propensities for different amino acids (and different locations in the helix) can be compared to experimental estimates in model peptide and protein systems, and a combination of theory and experiment provides good insight into the origins of sequence-dependent effects in these systems. We can anticipate continuing insights from simulations in this area.

(4) Free-energy calculations on forming NH···OC hydrogen bonds in short turns or for an isolated linear hydrogen bond show only a very small hydrogen-bond stabilization in aqueous solution. Greater stabilization is found for a β-sheet model with two adjacent hydrogen bonds, a result that is attributed to the poorer solvation of the separated strands relative to the individual amide groups. Results of this sort can provide important clues to the interpretation of the driving force for secondary structure formation in proteins.

(5) Molecular dynamics simulations have also provided interesting insights into the time scales and intermediates involved in formation of helices and turns from more random peptide conformations. Bifurcated hydrogen bonds and $i \leftarrow i + 3$ interactions appear to be common intermediates, and local transitions (involving just a few residues in a turn, or at the end of a helix) occur with residence times in the nanosecond range. Again, these broad conclusions do not appear to depend upon details of the potential functions or simulation protocols used.

It is also of interest to speculate what the future might hold. Clearly, computer hardware and software will continue to improve, making calculations like those discussed here easier to perform on a routine basis. This enhancement will certainly broaden the range of applications to consider more sequence-specific effects as well as more complicated geometrical and topological factors. This increased breadth should facilitate more direct comparisons to experiments on peptide and protein folding. We also expect to see developments in technical aspects of free-energy calculations, as in methods to reduce the spurious effects on long-range force truncation. In addition, there is an increasing focus on questions of convergence of sampling in free-energy simulations, which should lead to a better understanding of the errors to be expected. Finally, we note continuing progress on the experimental side toward greater temporal and spatial resolution in studies of peptides: NMR experiments can now probe peptide conformation on a residue-by-residue basis, and developments in time-resolved optical and IR spectroscopies allow more direct and detailed comparisons between theory and experiment. These considerations lead us to a optimistic outlook for the role of simulation methods in studies of peptide conformational dynamics and thermodynamics.

Acknowledgments. This work was supported by NIH grants GM38794 (D.A.C.) and GM37554 (C.L.B.). We thank Don Bashford, Mary Karpen, Klara Ösapay, and Alain St. Amant for helpful discussions and for communicating results prior to publication.

References

(1) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.
(2) Harvey, S.; McCammon, J. A. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.
(3) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Adv. Chem. Phys.* **1988**, *71*, 1–249.
(4) Brooks, C.; Brünger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843–865.
(5) Warshel, A.; Russell, S. T. *Quart. Rev. Biophys.* **1984**, *17*, 283–422.
(6) Schreiber, H.; Steinhauser, O. *Biochemistry* **1992**, *31*, 5856–5860.
(7) Schreiber, H.; Steinhauser, O. *J. Mol. Biol.* **1992**, *228*, 909–923.
(8) Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.
(9) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
(10) Sneddon, S. F.; Brooks, C. L., III. In *Molecular Structures in Biology*; Prout, K., Ed.; Oxford University Press: Oxford, 1991.
(11) Valleau, J. P.; Torrie, G. M. In *Modern Theoretical Chemistry, Vol. 5: Statistical Mechanics, Part A, Equilibrium Techniques*; Berne, B. J., Ed.; Plenum Press: New York, 1977.
(12) Tobias, D. J.; Brooks, C. L., III. *Chem. Phys. Lett.* **1987**, *142*, 472–476.
(13) Tobias, D. J.; Brooks, C. L., III. *J. Chem. Phys.* **1988**, *89*, 5115.
(14) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407–435.
(15) Hermans, J. *Curr. Opin. Struct. Biol.* **1993**, *3*, 270–276.
(16) Northrup, S. H.; Pear, M. R.; Lee, C.-Y.; McCammon, J. A.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4035–4039.
(17) Mezei, M.; Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1985**, *107*, 2239–2245.
(18) Boczko, E. M.; Brooks, C. L., III. *J. Phys. Chem.* **1993**, *97*, 4509–4513.
(19) Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 7690–7697.
(20) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195–1198.
(21) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(22) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
(23) Goldstein, H. *Classical Mechanics*; Addison-Wesley: Reading, MA, 1980.
(24) Beveridge, D. L.; DiCapua, F. M. In *Computer Simulations of Biomolecular Systems*; VanGunsteren, W. F., Weiner P. K., Eds.; ESCOM; Leiden, 1989.
(25) Tobias, D. J.; Sneddon, S. F.; Brooks, C. L., III. *J. Mol. Biol.* **1991**, *216*, 783.
(26) Yu, H.-A.; Karplus, M. *J. Chem. Phys.* **1988**, *89*, 2366–2379.
(27) Pullman, B.; Pullman, A. *Adv. Protein Chem.* **1974**, *28*, 348.
(28) Scarsdale, J. N.; Van Alsenoy, C.; Klimkowski, V. J.; Schäfer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438–3445.
(29) Weiner, S. J.; Singh, U. C.; O'Donnell, T. J.; Kollman, P. A. *J. Am. Chem. Soc.* **1984**, *106*, 6243–6245.
(30) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989–5997.
(31) Böhm, H.-J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129–7135.
(32) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369–5377.
(33) Gould, I. R.; Kollman, P. A. *J. Phys. Chem.* **1992**, *96*, 9255–9258.
(34) Case, D. A. In *Conformational Analysis of Medium-Sized Heterocycles*; Glass, R. S., Ed.; VCH Publishers: New York, 1988; pp 1–34.
(35) Roterman, I. K.; Lambert, M. H.; Gibson, K. D.; Scheraga, H. A. *J. Biomol. Struct. Dyn.* **1989**, *7*, 421–453.
(36) Nguyen, D. T.; Case, D. A. *J. Phys. Chem.* **1985**, *89*, 4020–4026.
(37) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.
(38) Zimmerman, S. S.; Pottle, M. S.; Némethy, G.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1–9.
(39) Schiffer, C. A.; Caldwell, J. W.; Stroud, R. M.; Kollman, P. A. *Protein Sci.* **1992**, *1*, 396–400.
(40) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1671.
(41) Pettitt, B. M.; Karplus, M. *J. Am. Chem. Soc.* **1985**, *107*, 1166–1173.
(42) Pettitt, B. M.; Karplus, M. *Chem. Phys. Lett* **1985**, *121*, 194.
(43) Pettitt, M.; Karplus, M. *J. Phys. Chem.* **1988**, *92*, 3994–3997.
(44) Tobias, D. J.; Brooks, C. L., III. *J. Phys. Chem.* **1992**, *96*, 3864–3870.
(45) Jean-Charles, A.; Nicholls, A.; Sharp, K.; Honig, B.; Tempczyk, A.; Henderson, T. F.; Still, W. C. *J. Am. Chem. Soc.* **1991**, *113*, 1454–1455.
(46) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
(47) Grant, J. A.; Williams, R. L.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 929–949.
(48) Sharp, K. *J. Comput. Chem.* **1991**, *12*, 454–468.
(49) Anderson, A. G.; Hermans, J. *Proteins: Struct. Funct. Genet.* **1988**, *3*, 262–265.
(50) Karplus, M.; Weaver, D. L. *Nature* **1976**, *260*, 404–406.
(51) Kim, P. S.; Baldwin, R. L. *Ann. Rev. Biochem.* **1982**, *51*, 459–489.
(52) Wright, P. E.; Dyson, H. J.; Lerner, R. A. *Biochemistry* **1988**, *27*, 7167–7175.
(53) Dyson, H. J.; Rance, M.; Houghten, R. A.; Wright, P. E.; Lerner, R. A. *J. Mol. Biol.* **1988**, *201*, 201–217.
(54) Padmanabhan, S.; Marqusee, S.; Ridgeway, T.; Laue, T. M.; Baldwin, R. L. *Nature* **1990**, *344*, 268–270.
(55) Osterhout, J. J., Jr.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1989**, *28*, 7059–7064.
(56) Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95–118.
(57) Waltho, J. P.; Feher, V. A.; Lerner, R. A.; Wright, P. E. *FEBS Lett.* **1989**, *250*, 400–404.
(58) Finkelstein, A. V.; Badretdinov, A. Y.; Ptitsyn, O. B. *Nature* **1990**, *345*, 300.
(59) Shoemaker, K. R.; Fairman, R.; Schultz, D. A.; Robertson, A. D.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Biopolymers* **1990**, *29*, 1–11.
(60) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519–538.
(61) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 5286–5290.
(62) O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646–651.
(63) Lyu, P. C.; Liff, M. I.; Marky, L. A.; Kallenbach, N. R. *Science* **1990**, *250*, 669–673.
(64) Scheraga, H. A. *Pure Appl Chem.* **1978**, *50*, 315.
(65) Merutka, G.; Stellwagen, E. *Biochemistry* **1991**, *30*, 1591–1594.
(66) Bradley, E. K.; Thomason, J. F.; Cohen, F. E.; Kosen, P. A.; Kuntz, I. D. *J. Mol. Biol.* **1990**, *215*, 607–622.
(67) Daggett, V.; Levitt, M. *J. Mol. Biol.* **1992**, *223*, 1121–1138.
(68) Yun, R. H.; Anderson, A.; Hermans, J. *Proteins: Struct. Funct. Genet.* **1991**, *10*, 219–228.
(69) Hermans, J.; Anderson, A. G.; Yun, R. H. *Biochemistry* **1992**, *31*, 5646–5653.
(70) DiCapua, F. M.; Swaminathan, S.; Beveridge, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 6145–6155.
(71) Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 795–817.
(72) Dyson, H. J.; Sayre, J. R.; Merutka, G.; Shin, H.-C.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 819–835.
(73) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1991**, *30*, 3864–3871.
(74) Soman, K. V.; Karimi, A.; Case, D. A. *Biopolymers* **1991**, *31*, 1351–1361.
(75) Soman, K. V.; Karimi, A.; Case, D. A. *Biopolymers* **1993** *33*, 1567–1580.
(76) Tobias, D. J.; Brooks, C. L.; III. *Biochemistry* **1991**, *30*, 6059–6070.
(77) Yun, R. H.; Hermans, J. *Prot. Eng.* **1991**, *4*, 761–766.
(78) Creamer, T. P.; Rose, G. D. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 5937–5943.
(79) Tobias, D. J.; Sneddon, S. F.; Brooks, C. L., III. In *Advances in Biomedical Simulations*; American Institute of Physics, Obernai: France, 1991; pp 174–199.
(80) Tobias, D. J.; Brooks, C. L., III. Helix propagation in a two-turn α helix. Manuscript in preparation.
(81) Karpen, M. E.; Brooks, C. L., III. Proline influences in helix initiation. Manuscript in preparation.
(82) Karpen, M. E.; Brooks, C. L., III. Proline influences in helix propagation and termination. Work in progress.
(83) Czerminski, R.; Elber, R. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 6963–6967.
(84) Czerminski, R.; Elber, R. *J. Chem. Phys.* **1990**, *92*, 5580–5601.
(85) Sundaralingham, M.; Sekharuda, Y. C. *Science* **1989**, *244*, 1333–1337.
(86) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Proteins: Struct. Funct. Genet.* **1990**, *7*, 62–73.
(87) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Biopolymers* **1993**, *33*, 827–837.
(88) Tobias, D. J.; Sneddon, S. F.; Brooks, C. L., III. *J. Mol. Biol.* **1990**, *216*, 783–796.
(89) Tobias, D. J.; Boczko, E. M.; Sneddon, S. F.; Brooks, C. L., III. Glycine effects on reverse turn formation. Manuscript in preparation.
(90) Wright, P. E.; Dyson, H. J.; Feher, V. A.; Tennant, L. L.; Waltho, J. P.; Lerner, R. A.; Case, D. A. In *Frontiers of NMR in Molecular Biology*; Live, D.; Armitage, I. M.; Patel, D., Eds.; Alan R. Liss: New York, 1990; pp 1–13.
(91) Tobias, D. J.; Mertz, J. E.; Brooks, C. L., III. *Biochemistry* **1991**, *30*, 6054–6058.
(92) Karpen, M. E.; Tobias, D. J.; Brooks, C. L., III. *Biochemistry* **1993**, *32*, 412–420.
(93) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167.
(94) Tobias, D. J.; Sneddon, S. F.; Brooks, C. L., III. *J. Mol. Biol.* **1992**, *227*, 1244–1252.
(95) Sneddon, S. F.; Tobias, D. J.; Brooks, C. L., III. *J. Mol. Biol.* **1989**, *209*, 817–820.
(96) Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 3770–2771.