PROTEINS: Structure, Function, and Genetics 10:219–228 (1991)

Proline in α-Helix: Stability and Conformation Studied by Dynamics Simulation

R.H. Yun, Amil Anderson, and Jan Hermans
Department of Biochemistry and Biophysics, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599-7260

ABSTRACT **Free-energy simulations have been used to estimate the change in the conformational stability of short polyalanine α-helices when one of the alanines is replaced by a proline residue. For substituting proline in the middle of the helix the change in free energy of folding ($\Delta\Delta G°$) was calculated as 14 kJ/mol (3.4 kcal/mol), in excellent agreement with the one available experimental value. The helix containing proline was found to be strongly kinked; the free energy for reducing the angle of the kink from 40° to 15° was calculated, and found to be small. A tendency to alternate hydrogen bonding schemes was observed in the proline-containing helix. These observations for the oligopeptide agree well with the observation of a range of kink angles (18–35°) and variety of hydrogen bonding schemes, in the rare instances where proline occurs in helices in globular proteins. For substituting proline at the N-terminus of the helix the change in free energy of folding ($\Delta\Delta G°$) was calculated as −4 kJ/mol in the first helical position (N1) and +6 kJ/mol in the second helical position (N2). The observed frequent occurrence of proline in position N1 in α-helices in proteins therefore has its origin in stability differences of secondary structure. The conclusion reached here that proline may be a better helix former in position N1 than (even) alanine, and thus be a helix initiator may be testable experimentally by measurements of fraction helical conformation of individual residues in oligopeptides of appropriate sequence. The relevance of these results in regards to the frequent occurrence of proline-containing helices in certain membrane proteins is discussed.**

Key words: proline, α-helix, kinked α-helix, molecular dynamics, computer simulation, peptide conformation stability, protein conformational stability, amino acid substitution, protein architecture, helix start/stop signal

INTRODUCTION

This paper describes the first of several studies of substitution of a central residue of a short oligoalanine α-helix, in which we estimate how strongly each substitution modifies the stability of the helix relative to the random coil, examine conformational effects of the substitutions in both helix and random conformations, and attempt to rationalize the differences in terms of specific interactions. As we shall here describe, we use molecular dynamics simulations, with the usual empirical force field (but complete with explicit representation of solvent water); we employ both free dynamics and dynamics with forcing methods, the former to calculate conformational distributions, the latter to calculate free energy differences.

The attractive side of such a study is the great detail, in terms of energetics, conformation, and conformational variability, with which the molecular dynamics model describes the system. A drawback is uncertainty about the accuracy of the results due to the empirical and approximate nature of the force field. We (and others) resolve this uncertainty by comparing (precisely) computed and observed properties of the system. For these we have chosen the *differences* in stability of the folded conformation (the helix) relative to the random coil state, in terms of differences ($\Delta\Delta G°_h$) of the free energy of folding ($\Delta G°_h$). (On the basis of the theory of the helix–coil transition, this is the free energy corresponding to the ratio of the helix growth parameters for the two residue types, at least if the substitution is made away from the helix ends.)

It has long been inferred that substitution of a single alanine with proline has significant consequences for the stability and the conformation of an α-helix. In globular proteins of known conformation, proline occurs rarely inside α-helices, and generally, the helices are kinked.[1] Furthermore, proline apparently destabilizes the α-helix so strongly that the model oligo- and polypeptides whose helix–coil equilibria are used to assess differences in helix stabil-

Received August 15, 1990; revision accepted December 3, 1990.

Address reprint requests to Dr. Jan Hermans, Department of Biochemistry, University of North Carolina, Chapel Hill, NC 27599-7260.

Amil Anderson's present address is Department of Chemistry, Wittenberg University, Springfield, OH 45501.

Abbreviations: Ace, *N*-acetyl; Mam, *N*-methylamine.

ity, demonstrate zero helicity. This is not surprising if one considers that proline lacks the N–H group needed to form an intrahelix hydrogen bond with the C–O group three residues back along the chain (the orphan carbonyl group).

Recently, O'Neill and DeGrado[2] have developed a model system, consisting of oligopeptide helices additionally stabilized by dimerization, in which they were able to measure the effect of replacing a central alanine residue by proline, as well as other naturally occurring amino acids. The value of $\Delta\Delta G^\circ_h$ (Ala⇒Pro) was estimated in this way at 14 kJ/mol (3.4 kcal/mol), considerably larger than for any other residue that they studied.

The theoretical model reproduces this experimental stability difference, and, furthermore, we have found that the model produces a series of conformations of the proline-containing helix, which are kinked to a varying extent and which appear to be easily interconvertible. The model has been found to possess a much greater stability if proline is the first helical residue than when it is the second helical residue; the latter is again a more stable molecule than one in which proline is interior to the helix.

METHODS

Program and Forcefield

Simulations were carried out with the program Cedar (developed in this laboratory). Cedar employs a forcefield in which the nonbonded parameters[3] are the same as those of the Gromos program, and a description of geometry and geometric deformations that was developed in this laboratory.[4] Cedar uses the Shake algorithm[5] in order to maintain bond-lengths rigorously constant, with a time step of 2 fsec. Peptide molecules were surrounded with SPC water,[6] and periodic boundary conditions were used. Mean temperature and pressure were maintained constant by small adjustments at each dynamics step of the kinetic energy and the dimensions of the periodic box.[7] All systems were initially equilibrated for at least 10 psec to a temperature of 300 K and a pressure of 1 atm. We used a central-atom representation for CH, CH_2, and CH_3 groups, and a 6 Å cutoff on nonbond interactions between formally neutral groups: CH_x, CO, NH, and H_2O.

Free-Energy Calculations

Two kinds of free-energy calculations were used: one to produce conformational forcing and the other to produce molecular replacement. Conformational forcing methods were used in order to assess free energy differences between different conformations (cf. earlier work[8,9]). Briefly, conformation changes were obtained by internal rotation about a single bond, in this case, a C_α–C bond in the polypeptide backbone, by including in the potential energy, U a forcing term of the form

$$U_f = (U_{f,0}/2)\,[1-\cos(\rho-\rho^*)] \qquad (1)$$

where ρ is the dihedral angle for the bond, which is thereby restrained to the neighborhood of ρ^*. The latter's value is slowly varied, and the conformation changes accordingly.

Molecular-replacement calculations were used in order to estimate free energy differences between molecules containing alanine and proline. When these are calculated for both the helix and the coil state, then their difference is equal to $\Delta\Delta G^\circ$, the same term that was defined above in terms of a difference of free energies for conformation change. This can be seen according to the following cyclic scheme[10–13] involving helical and random (coil) states of the alanine and proline variants:

$$\begin{array}{llll} [1] & \text{Helix (Ala-}i\text{)} & \Leftrightarrow \text{ Coil (Ala-}i\text{)} & [2] \\ & \Uparrow\Downarrow & \Uparrow\Downarrow & \\ [3] & \text{Helix (Pro-}i\text{)} & \Leftrightarrow \text{ Coil (Pro-}i\text{)} & [4] \end{array} \qquad (2)$$

The condition that the sum of the free energy changes for the entire cycle must be zero, can be rewritten:

$$\Delta\Delta G^\circ = \Delta G^\circ_{12} - \Delta G^\circ_{34} = \Delta G^\circ_{13} - \Delta G^\circ_{24} \qquad (3)$$

where the first difference is in terms of experimentally accessible free energies of folding, and the second difference is computed for the model by the molecular replacement method.[12]

In a molecular replacement calculation, the two alternate residues are both represented and the energy of the system is calculated with varying contributions from each. In the Cedar program this is implemented as follows. Where reasonable, equivalent atoms of each residue are maintained at identical positions; one or more atoms in at least one of the residues is "unique." (In this case, proline has the unique atoms, and hence simulation at $\lambda=1$ produces the proline-containing form, and at $\lambda=0$ the alanine-containing form.) Forces for geometric deformations are multiplied by a coupling parameter, λ for one residue and $1-\lambda$ for the other, for all terms which involve only shared atoms, and are fully counted when any unique atoms are affected. Nonbond force terms for atom pairs in which at least one atom is part of the residue containing the unique atoms, are multiplied by powers of the coupling parameter, λ^3 for the attractive Lennard–Jones term and the electrostatic term and λ^5 for the repulsive Lennard–Jones term,[14] while those for the other residue are multiplied by $(1-\lambda^3)$ and $(1-\lambda^5)$.[15] At the end points ($\lambda=0$ and $\lambda=1$), the inactive residue's geometry is maintained; the additional mass will perturb the dynamic, but not the equilibrium properties of the system.

The free energy change for either forced process, molecular replacement or conformation change, is

estimated as the work done on the system by the changing potential in a quasistatic process

$$\Delta G^\circ = \int\langle \partial U/\partial\lambda\rangle d\lambda, \quad \text{or}$$
$$\Delta G^\circ = \int\langle \partial U/\partial\rho^*\rangle d\rho^*. \qquad (4)$$

The $\langle\,\rangle$ signs indicate averages over a Boltzmann distribution (at each value of λ or ρ^*). Each integral is estimated in a single molecular dynamics simulation, in which the value of λ or ρ^* is changed by the same small increment (or decrement), $\delta\lambda$ or $\delta\rho^*$ after *every* step, as

$$\Delta G^\circ = \Sigma(\partial U/\partial\lambda)\delta\lambda, \quad \text{or} \quad \Delta G^\circ = \Sigma(\partial U/\partial\rho^*)\delta\rho^*. \qquad (5)$$

The system is equilibrated at constant λ or ρ^* following the simulation, and then the process is reversed. ΔG° for the forward and $-\Delta G^\circ$ for the reverse calculations are averaged; "reversibility" is determined by assessing the superposition of the forward and reverse progress curves, and from the sum of ΔG°s for the forward and reverse calculations. The latter should be equal to zero and is reported below as "closure error." It should be noted that the uncertainty in the mean ΔG° is much less than this closure error. From numerous examples for which we have calculated ΔG° successively with longer runs, i.e., each time more closely approximating a quasistatic process, we estimate the uncertainty as 20% of the closure error, provided the superposition of forward and reverse progress curves supports the hypothesis that the simulations are almost quasistatic. The reported free energy differences are precise to within 1 kJ/mol, or less.[8,9]

Conformational Restraints Prevent Conformation Change[16]

Conformational restraints were used where necessary to prevent conformation change, for example, to prevent unfolding of the ends of the α-helix. After some experimentation, we have selected a particularly effective scheme, by which the conformation may be confined, but the conformational probability distribution is unperturbed locally. The scheme consists of adding an artificial restraint potential which is zero in the part of conformation space that is of interest, but rapidly grows to a large value (say 40 kJ/mol) over a short interval wherever a (free) energy barrier exists via which the conformation might make an undesired change to an alternate, and perhaps more stable, conformation. Thus, the distribution is perturbed where the probability density is nearly zero (the barrier), and, of course, conformation change is prevented. We have applied this restraint successfully to dihedral angles and to interatomic distances. For example, the α-helical conformation can be maintained by adding a potential which rapidly increases from zero to 40 kJ/mol when the distance between a hydrogen-bonded pair (O and H) increases from 2.5 to 3.5 Å, a distance which is atypical of the α-helical conformation, but will occur if the helix changes conformation, either to a 3_{10}-helix, or to a (partially) extended conformation.

(This constraint potential is constant below and above a narrow range of the constrained geometric parameter, in this case dihedral angle or interatomic distance, having zero value on one side, and a large and positive value on the other, the "forbidden" side. Within the narrow range the function is described by a third-degree polynomial. The function is continuous and has a continuous first derivative.)

Conformational freedom of a molecule for which a free energy calculation is carried out causes special problems, as one can understand from the formal need to approximate a Boltzmann distribution during the entire simulation, in order to approximate the quasistatic process required if the adopted integration process is to produce a free energy difference. If the Boltzmann distribution includes different conformation states separated by energy barriers, which are spontaneously crossed very infrequently, then proper sampling will require a very long calculation. We avoid this difficulty by restraining the conformation to a part of conformation space within which conformational equilibration is rapid. It is then necessary subsequently to take into account also the free energy for imposition of this restraint on the beginning and ending states for the transformation for which the free energy difference was calculated, by adding the difference of the restraint free energies for the two states:

$$\Delta G^\circ_{AB} = \Delta G^\circ_{rest,A} + (\Delta G^\circ_{AB})_{rest} - \Delta G^\circ_{rest,B} \qquad (6)$$

Here $(\Delta G^\circ_{AB})_{rest}$ is the result of the calculation in the presence of the restraint and $\Delta G^\circ_{rest,A}$ and $\Delta G^\circ_{rest,B}$ the free energies for imposing the conformational restraint. These are calculated from the conformational probability distribution in the absence of the restraint, $P(\chi)$, χ standing for the *set* of degrees of freedom to which the restraints are applied, according to

$$\Delta G^\circ_{rest,A} = -kT \ln [\int \exp(-U_{rest}(\chi)/kT)\, P(\chi)\, d\chi] \qquad (7)$$

with $U_{rest}(\chi)$ the restraint energy.[8,9]

This approach effectively solves the problem of how to deal with conformational variability if there are few possible conformations. (For example, the alanine α-helix with a single conformation, the proline dipeptide with two, an alanine helix containing a single valine residue, which has three, by virtue of internal rotation of the side chain.) However, in this case, the problem includes the random-coil state of the oligopeptide, which almost by definition is distributed over a very large number of conformations. How to deal with this problem is discussed in the Results section.

The Problem of Ring-Imposed Restraint and Its Solution

A peculiar problem occurs in calculations in which alanine is replaced with proline due to the way in which the replacement calculation is executed. The cyclic structure of proline restricts the conformation of the alanine which it replaces, even when it is inactive ($\lambda = 0$) and the *nonbond* interactions of proline atoms are zero. Thus, the free energy of the alanine form of the molecule is raised. (The presence of alanine does not change the free energy of the proline form, when the nonbond interactions of alanine are zero.)

It is possible to negate the effects of this restriction by imposing an even stronger restraint on the dihedral angle ϕ of the replaced residue, of the form of Eq. (1), $(U_{f,0}/2)[1-\cos(\phi-\phi^*)]$ where ϕ^* corresponds to the free energy minimum for the proline-containing molecule. This restraint should be much tighter than that imposed by the ring structure, as it eliminates the latter's effect on the free energy in proportion to the ratio of the restraints' (effective) force constants. The restraint free energies for the two forms are calculated using Eq. (6), with the probability distributions estimated by simulations of the alanine- and proline-containing forms, respectively, using representations that do *not* include the other residue in any form.

Another way of estimating the effect of the presence of the ring on the free energy of the alanine is to estimate the distribution of the dihedral angle whose distribution is restricted by the presence of the ring, in two simulations: in one the proline residue is not represented, which gives $P(\phi)$, and in the other it is represented, but at $\lambda = 0$, which gives $P^*(\phi)$. If P and P^* are both normalized and ϕ_0 is a value of ϕ, for which the restraint is zero, then the restraint free energy is given by [16]

$$\Delta G^* = -kT \ln[P(\phi_0)/P^*(\phi_0)]. \quad (8)$$

RESULTS

Conformational Probability Distribution of the Proline Dipeptide

A conformational probability distribution of the proline dipeptide was computed for use in the calculation of the effects of the conformational restraints according to Eq. (6). As for the alanine dipeptide, this was done by first estimating conformational distributions near local free energy minima by free simulation, then calculating free energy differences between different minima by forced conformation change, and scaling the individual distributions according to the corresponding Boltzmann factors.

The internal rotation about the N–C_α bond of a proline residue is restricted by the presence of the five-membered ring formed by the backbone nitrogen and the C_β, C_γ, and C_δ atoms. The backbone dihedral angle, ϕ, can deviate only slightly from $-60°$. By analogy with the conformational probability distribution of the alanine dipeptide,[8,9] the distribution for proline is therefore expected to have only two minima, one for positive and the other for negative angles of the dihedral angle Ψ. The distribution was calculated for a molecule Ace-Pro-Mam in a cubic box containing also 161 water molecules. After constant-pressure equilibration, the volume of the box averaged at circa 5,040 Å^3. The two local probability distributions were estimated from dynamics trajectories (80 psec each) computed for $0<\Psi<180$ and for $-180<\Psi<0$, respectively. A forced dynamics calculation, via $\Psi = 0°$, was performed to estimate the free energy difference between the two conformations. (Closure error for forward and reverse calculations of 200 psec each was 1.2 kJ/mol.) A correction for the imposition/removal of the restraint used in the forcing calculation was computed according to Eq. (6), using the local probability distributions of proline in the β- and α_R-conformations. The resulting free energy difference was found to be 6.7 kJ/mol, with the β-conformation[17] the most stable. These results were combined to give the free energy map shown in Figure 1.

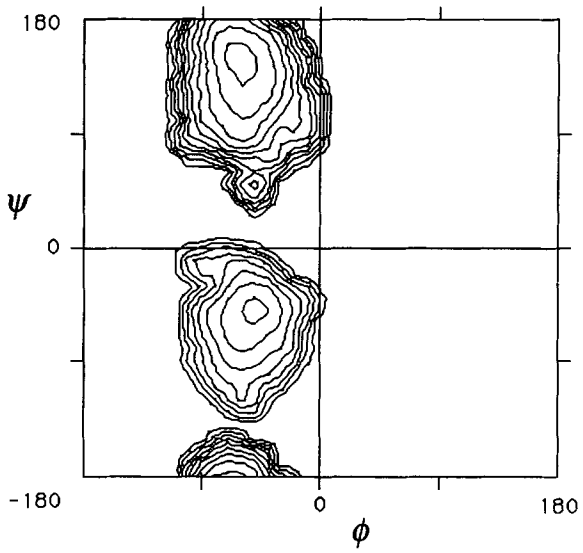

Fig. 1. Conformational free energy map of the proline dipeptide. Contours are at intervals of 2 kJ/mol. The free energy minimum in the upper left of the figure corresponds to the β-conformation. The first contour is 2 kJ/mol above this minimum. The minimum on the lower left is the α_R conformation. The lowest contour drawn for this minimum is at 8 kJ/mol.

Molecular Replacement for the Random-Coil State

The free energy for replacing alanine with proline in the random coil state, ΔG°_{24} [cf. Eqs. (2) and (3)] was computed as follows:

The tetrapeptides Ace-Ala-Ala-Ala-Mam and Ace-Ala-Pro-Ala-Mam were used to represent the non-helical state. These were constrained to the β-conformation[17] and converted one into the other. The answer is incorrect, because the conformation has been restrained. We correct the answer by assuming that the conformational probability distributions of residues in the random coil state are independent of one another (an assumption made some time ago by Brant and Flory[18] in a theory of the dimensions of random-coil peptides and protein chains). With this assumption, the effects of restraining the alanine residues which the two molecules have in common (1 and 3) cancel, and the free energy terms for restraining the central alanine/proline residue can be calculated from the known probability distributions of alanine and proline *in the dipeptides* (ref. 9 and Fig. 1).

The molecular replacement simulation was done with a molecule of the tetrapeptide in which the central alanine could be replaced with proline, in a cubic box with 145 water molecules. After equilibration the volume was 4750 Å^3 (as Ala). The free energy for conversion to proline was calculated in forward and reverse simulations of 200 psec each (closure error 3.1 kJ/mol). This calculation was done in the presence of a conformational restraint on the alanine residue with $U_{f,0}=250$ kJ/mol and $\phi^*=-55°$, according to the first of the two described methods to account for the ring-imposed restraint. The free energy for the replacement of alanine with proline was found to be $\Delta G^\circ_{24} = -17.5$ kJ/mol, as the sum of three terms: imposition of the restraints on the alanine dipeptide (9.7), replacement of alanine with proline in the restrained tetrapeptide (−22.3), and removal of the restraints from the proline dipeptide (−4.9).

(Use of the tetrapeptide as representative of the random state was suggested by the results of calculations for replacement of alanine with glycine, where only a very small difference was found when the dipeptide, rather than the tetrapeptide was used.)

Molecular Replacement in the α-Helix: Interior Residue

The free energy for replacing alanine with proline in the helix, ΔG°_{13} was computed as follows.

The α-helix was studied with a peptide containing 14 alanine residues, with acetyl and methylamide blocking groups. The ninth alanine residue was replaced with proline. This molecule was given an α-helical conformation, which was maintained by restraints on backbone dihedral angles and hydrogen bond distances. The backbone dihedral angles, ϕ and Ψ, were restrained to the interval −90° to −10° by potentials that rose from 0 to 125 kJ/mol over the 20° intervals −10° to 10° and −90° to −110°. Second, the helical hydrogen bond distances were retained to be below 2.5 Å, by a restraint potential ("latch") that rose from 0 to 125 kJ/mol when the H-to-O distance changed from 2.5 to 3.5 Å. The distance restraint was not applied to the hydrogen bonds Ala-4-O and Ala-8-H_N, Ala-5-O and Ala-9-H_N, and Ala-6-O and Ala-10-H_N (Ala-9 being the residue that was converted to Pro).

This molecule was placed in a periodic box $18\times 18\times 35$ Å^3, surrounded by 266 water molecules. The dimensions of the box were scaled isotropically to maintain constant mean pressure (final volume averaged near 9,200 Å^3). The long dimension of the helix was maintained approximately along a diagonal of the box, by imposition of positional restraints on the first and last carbon atoms of the molecule, with a relaxation time of 10 psec.

The free energy for the replacement of alanine with proline was calculated in forward and reverse simulations of 400 psec each (closure error 3.9 kJ/mol), and found to be −3.4 kJ/mol. An analysis showed that the statistical distribution of the backbone dihedrals ϕ and Ψ and the hydrogen bond distances *of the α-helix* in a free dynamics simulation are not perturbed by the restraint potentials, and therefore no correction is necessary. A comparison of the distributions of the backbone dihedral angle ϕ of Ala-9 in two dynamics simulations, one with and the other without representation of Pro-9 as an alternate residue, showed that the restraint imposed by the proline ring is insignificant relative to that naturally imposed by the helical conformation. Thus, the figure for replacement of alanine with proline in the helix needs no correction for conformational restraints ($\Delta G^\circ_{13} = -3.4$ kJ/mol).

The change in the free energy of forming an oligopeptide helix when an (interior) alanine residue is replaced with proline, calculated for the model, is found by subtracting ΔG°_{24} from ΔG°_{13}, which gives $\Delta\Delta G^\circ = 14.1$ kJ/mol.

We observed that the introduction of the proline residue produced a tendency for the formation of 3_{10} helical, as well as α-helical hydrogen bonding, if no constraints had been placed on H-to-O distances. We then also observed the expected concomitant tendency to a larger closure error in the replacement calculations.

Molecular Replacement in the α-Helix: N-Terminal Residues

In two calculations the *second* alanine residue in the model was replaced with proline. In one case, the first alanine residue was in the extended conformation, so that proline was the first helical residue, and in the other case all residues were in the helical conformation, so that proline was the second helical residue. In the nomenclature of Presta and Rose[19] proline was placed in position N1 in the first, and N2 in the second case. In the first case the α-helical hydrogen bond to the carbonyl oxygen of the acetyl

group which terminates our model molecule is absent because the first alanine residue does not have the helical conformation, in the second case it is present; the first alanine's oxygen atom is hydrogen bonded in both cases.

With proline as first helical residue, the free energy for substituting alanine with proline was found to be −22.2 kJ/mol (closure error 0.5 kJ/mol). Accordingly, $\Delta\Delta G°$ was found to be −3.6 kJ/mol. With proline as second helical residue, we obtained −12.2 kJ/mol (closure error 0.2 kJ/mol) for a $\Delta\Delta G°$ of +5.8 kJ/mol. (These $\Delta\Delta G°$ include corrections of 1.1 and 0.5 kJ/mol, respectively, which were computed by the second of the two described methods to account for the reduced conformational freedom of alanine in the presence of the ring restraint. In these calculations all α-helical hydrogen bonds were constrained by distance latches; in the first, the nonhelical alanine residue was kept in the β-conformation by confining the backbone dihedral angle Ψ to lie between 10° and 170°.)

Thus, an α-helix with proline as first helical residue is found to be considerably more stable than otherwise similar helices with proline as second, or as interior helical residue.

Conformation Change: Helix Kinking

From analysis with molecular graphics it was found that the helix changed conformation during the course of the substitution of alanine with proline, from an essentially straight to a kinked conformation. Examples of such conformations are shown in Figure 2. In order to systematically follow the conformation change, cylinders were fit to the atoms, except C_β, of residues 1 through 7, and of residues 10 through 14, by minimizing the sum of the squares of the distances to the cylinder surfaces, while varying the direction and location of the cylinder axes and the magnitude of the cylinder radii. This sort of a fit can be easily, but more slowly, accomplished by eye, using interactive graphics (cf. Fig. 2). Graphical and computed fits compared quite well. The cylinder axes may reasonably be called the helix axes.

The change of the angle between the two helix axes for a series of conformations observed during the changeover from alanine to proline, and back, is shown in Figure 3.

An additional calculation was performed in order to determine the ease with which the angle between the helix axes can be varied in a proline-containing helix. This calculation was done with additional quadratic conformational restraints, of the form $U_d = (K_d/2)(d-d_0)^2$, which allow control, and gradual forced change, of the Ala-4-O to Ala-8-H_N and Ala-5-O to Ala-9-H_N distances from the large values in the strongly kinked helix (constraint distance, $d_0 = 4.5$ Å), to the values typical of hydrogen-bonded pairs in an all-alanine helix ($d_0 = 1.8$ Å). The force constants for these constraints, K_d, had values of 125 kJ/mol. Figure 4 shows the results of this calculation, in terms of both the free energy change [computed as $\Sigma(\partial U_d/\partial d_0)\ \delta d_0$, cf. Eq. (5)], and the interhelix angle. It is seen in that for $d_0 > 3$ Å the free energy change is small, while the helix–helix angle assumes values that vary from 40 to 15°.

DISCUSSION

The simulations exactly reproduce the one available experimental estimate[2] of the change of helical stability when proline replaces an interior alanine. The agreement must be deemed somewhat fortuitous, because it is impossible to guess how easily the helical dimer studied by O'Neill and DeGrado accommodates the presence of a kink in both helices. With this reservation, the model may be considered to have performed exceedingly well in this case, in which helix structure and stability are strongly disturbed. A part of the change in stability should be attributed to the presence of an intramolecular hydrogen-bond to the carbonyl group four residues before the proline residue (Ala-5-O). The observed $\Delta\Delta G°$, of 14 kJ/mol much exceeds the 2–6 kJ/mol that might be expected for the absence of an intramolecular hydrogen bond. The difference is attributable to steric repulsion between the proline ring and the orphan carbonyl group, which distorts the helical structure; this must be considered as a second crucial factor in the large change in stability.

According to these results, proline is a poor helix former in nearly all positions. This is so if proline occupies an internal position and even if it is present at the N-terminus of a helix in positions in which no carbonyl group's hydrogen bonding is affected. The relative ease with which a proline can be introduced into position N1 is an important exception, first noted on the basis of conformational energy calculations by Schimmel and Flory.[20] In globular proteins, proline is found with significantly high frequency as the first helical residue, but quite rarely in the second or third position.[21] The calculations give approx. 9 kJ/mol as the free energy change opposing extension of a helix with proline in position N1 with another residue (whereby proline would move into position N2). Proline's role as a helix terminator can therefore be understood in terms of secondary structure, i.e., without any need to invoke tertiary interactions in the folded protein. The atomic interactions that work against the introduction of proline into position N2 are those of the proline δ-methylene group clashing with the β-methyl group of the N1 residue.

The calculations indicate that proline in position N1 gives a more stable helix than alanine. If this is indeed the case, then proline should be considered not merely for the passive role of helix *terminator*, but also for the active role of helix *initiator*. Experimental confirmation may be possible by solution

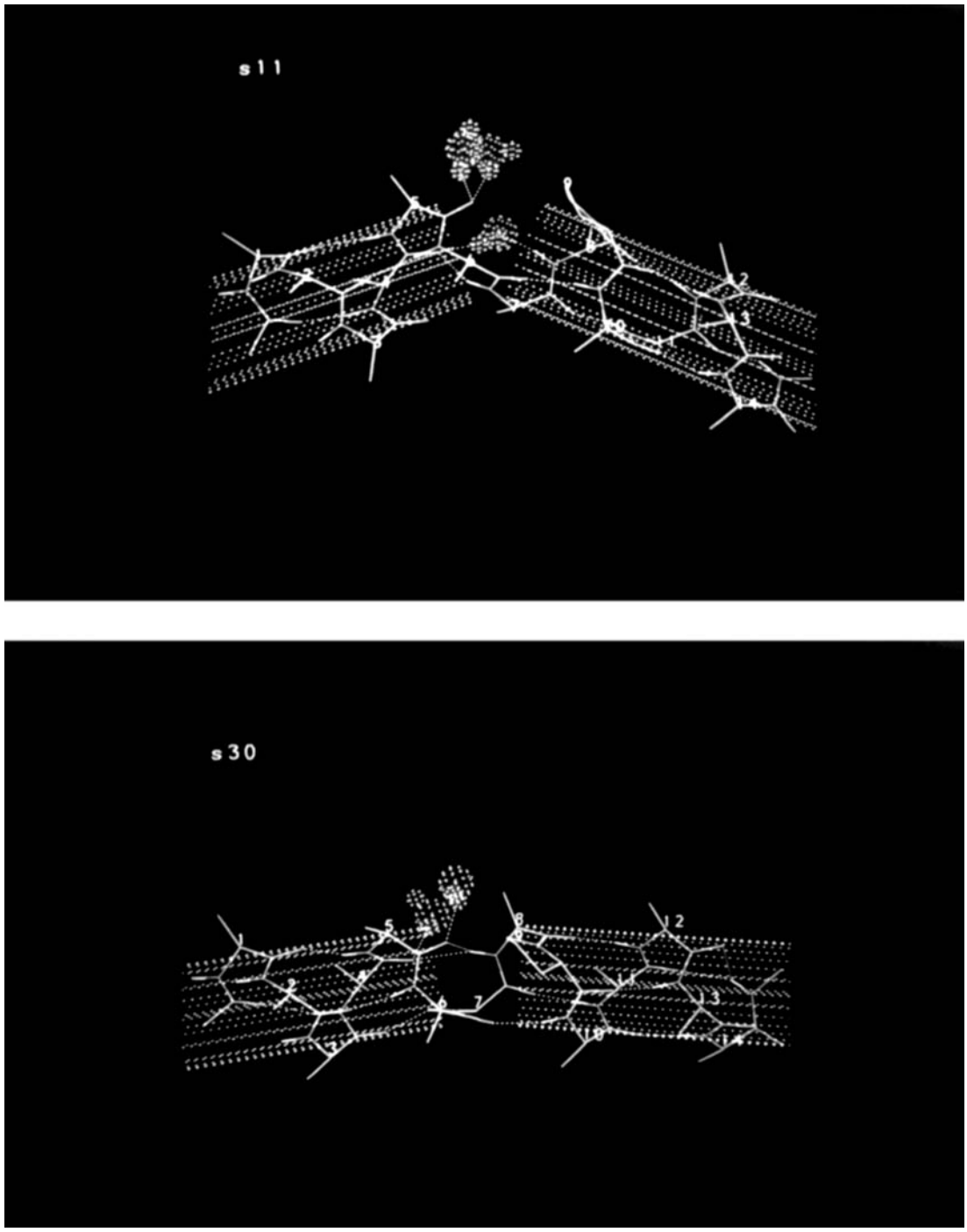

Fig. 2. Two typical conformations of a helix containing 13 alanine residues and a single proline residue (number 9), observed during the course of a molecular dynamics simulation in which the helix was gradually straightened by application of a restraint on the distance between atoms Ala-4-O and Ala-8-H_N. Top: in a strongly kinked conformation a water molecule forms a bridge between these two atoms and two water molecules are hydrogen bonded to Ala-5-O. Included are images of cylinders having 2 Å diameter, which were fit to the helices by eye using interactive graphics. Bottom: in a less strongly kinked conformation Ala-4-O and Ala-8-H_N are hydrogen bonded. In addition, Ala-8-H_N can be considered hydrogen bonded to Ala-5-O and water molecules from hydrogen bonds to Ala-4-O and Ala-5-O.

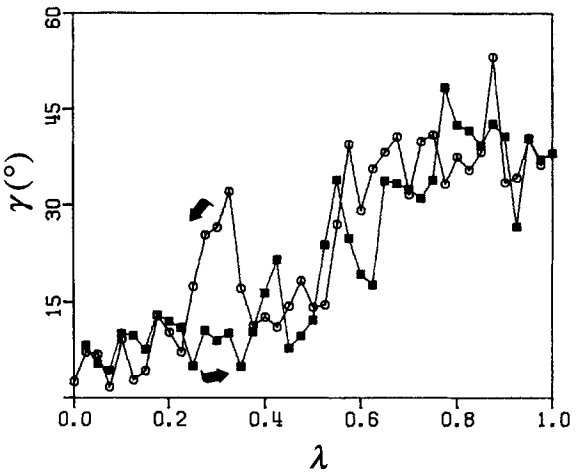

Fig. 3. Change of the angle, γ between the helix axes on either side of the proline residue in a 14-residue helix, during the course of replacement of alanine with proline, and the reverse, as a function of the coupling parameter, λ. (For λ = 0, the system represents the all-alanine helix.) An arrow by each curve indicates the direction in which λ was changed.

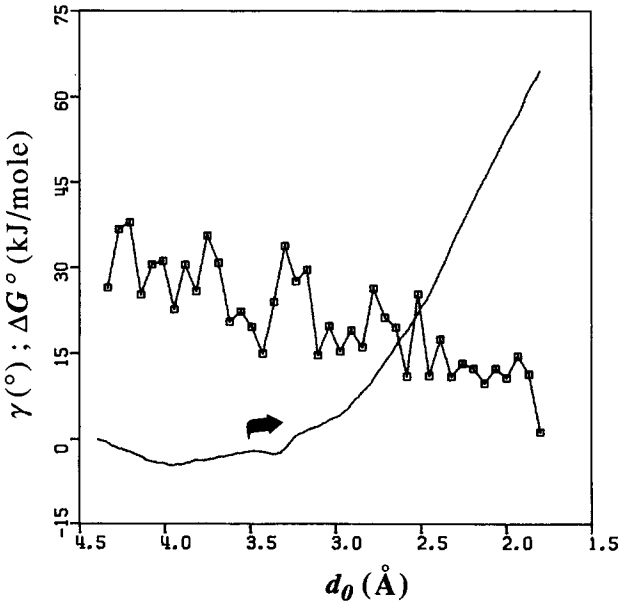

Fig. 4. Change of the conformational free energy (solid curve) and of the helix–helix angle, γ (squares) in a calculation in which the Ala-4-O to Ala-8-H_N and Ala-5-O and Ala-9-H_N distances were restrained to successively smaller values, d_0 by added quadratic potentials. (This calculation was done for the proline-containing helix, but in the presence of an inactive alanine residue as an alternate to Pro-9, i.e., at λ = 0.)

measurement of helical/nonhelical ratios of individual residues in a proline-containing oligopeptide.[22]

In has often been suggested that proline can stabilize folded peptide conformations by virtue of the limited conformation freedom of proline in the random coil. An analysis of the conformational free energy maps of the proline and alanine dipeptides (Fig. 1 and Ref. 8) as descriptive of the random coil state, indicates only a modest difference in conformational freedom, for a difference in free energy of 2 kJ/mol.[23] This accounts for much of the 3.6 kJ/mol by which the (model) helix is stabilized when proline replaces alanine in position N1. (We have not identified causes to which the remainder may be attributed.)

In globular proteins helices containing proline are (1) *rare* (surveys published in 1988 found 10 mid-helix prolines among 650 mid-helical residues[1,21]) and (2) generally *kinked,* by an angle that varies in the range of 19 to 35°.[1] The outside of the kink is exposed to solvent and solvent water molecules can and, presumably, do form hydrogen bonds with the protruding carbonyl oxygens. Hydrogen bonding in some of these helices is relatively uninterrupted, as in lamprey hemoglobin[24] (except of course for the orphan carbonyl group). In that case, the conformations assumed by these helices are quite similar to those of the less strongly kinked helices observed in the model in the presence of straightening forces (Fig. 2, bottom panel). In sperm whale myoglobin the hydrogen bond from the NH of the residue after the proline is stretched, to an H-to-O distance of 3.9 Å.[25] In other proteins, such as adenylate kinase,[26] the kink is accompanied by an irregular hydrogen bonding pattern.[27,28]

Qualitatively, these static features are consistent with our description of the dynamics of proline-containing helices. (The problems of accommodating proline in the α-helix have also been analyzed by energy minimization calculations with qualitatively similar results.[29]) Not found in globular proteins are the strongly kinked conformations with an inserted water molecule. However, the free energy required to convert the more strongly kinked conformations having an inserted water molecule, to straighter ones, was found to be small. Irregular and varying helical hydrogen bonding schemes were observed in some runs, and eliminated by the use of hydrogen bond latches in order to make the forcing calculations more easily reversible. One concludes that interactions with the structure of the entire protein decide which of several possible kinked structures is used.

We mention briefly the occurrence of proline in helices in apolar environments. Karle and coworkers have determined the structure of crystals of a very hydrophobic synthetic oligopeptide containing proline, grown from largely organic solution, and find one-and-a-half turns of α-helix on the N-terminal side of proline (residue 10), a kink at the proline, and then a beta-bend ribbon on the C-terminal side.[30] One of the four molecules of water in the unit cell forms a hydrogen bond with O7. Otherwise, both O7 and O6 are "orphans." Proline also occurs in the transmembrane helices of energy- and signal-transducing proteins.[31] Site-directed mutagenesis studies indicate that the functional properties of the transmembrane protein bacterio-

rhodopsin are sensitive to substitution of prolines: substitution of proline alters proton pumping capacity and spectral properties.[32] The proline-containing helices of bacterial rhodopsin have also been found to be kinked in a very recently obtained structure based on electron diffraction measurements at a resolution where polypeptide chains and large amino acid side chains are distinct.[33] It has been suggested that proline's function is to undergo a cis-trans isomerization as part of a protein conformation change that is presumably coupled to the light-induced conformation change of the chromophore.[31] An alternate possibility is that one or more of the kinks functions as a hinge in the coupled conformation change. The hydrogen-bonding capabilities of H_N of any residue replacing proline might alter the *energetics* of changing the kink angle, without necessarily abolishing the hinging function.

It is reasonable to consider a helix containing proline (except at position N1) as a *de*stabilizing feature of a protein's structure. The kinked proline-containing helix is a rare feature, somewhat of an oddity. A reasonable hypothesis is, however, that features that destabilize at the secondary structure level are not all that uncommon in globular proteins; we think, for example, of the unavoidable turns connecting segments of secondary structure which are unstable in terms of secondary structure, but which can be tolerated in the presence of an abundance of stabilizing (often hydrophobic) interactions. In membrane proteins, conformational stability is further enhanced by interactions of the protein surface with the hydrophobic part of the bilayer, and in these proteins destabilizing features can presumably be tolerated (even) more easily.

We are currently extending this work with a study of the effects of other amino acids (glycine, valine, D-alanine, and α-aminoisobutyric acid) on the stability of the α-helix and with a similar study of a β-turn model.

ACKNOWLEDGMENTS

This work was supported by a research grant from the National Science Foundation's Biophysics Program (DMB-8903584) and an instrumentation grant from the National Science Foundation's Biological Instrumentation Program (DIR-8908054). Computations were done on the University's Convex C240 and the North Carolina Supercomputer Center's Cray YMP. We thank Dr. William L. DeGrado and Dr. Neville R. Kallenbach for communicating unpublished results. J.H. thanks Dr. C.W. Carter, Jr. for frequently questioning the relevance of simulation studies.

REFERENCES

1. Barlow, D.J., Thornton, J. M. Helix geometry in proteins. J. Mol. Biol. 201:601–619, 1988.
2. O'Neill, K., DeGrado, W.L. A thermodynamic scale for the helix-forming tendencies of the commonly-occurring amino-acids. Science 250:646–651, 1990.
3. Hermans, J., Berendsen, H.J.C., van Gunsteren, W.F., Postma, J.P.M. A consistent empirical potential for water-protein interactions. Biopolymers 23:1513–1518, 1984.
4. Ferro, D.R., McQueen, J.E., McCown, J.T., Hermans, J. Energy minimizations of rubredoxin, J. Mol. Biol. 136:1–18, 1980.
5. Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. J. Comput. Phys. 23:327–341, 1977.
6. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Hermans, J. Interaction models for water in relation to protein hydration. In "Intermolecular Forces." Pullman, B., ed. Dordrecht, Holland: Reidel, 1981: 331–342.
7. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak J.R. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81:3684–3690, 1984.
8. Anderson, A., Carson, M., Hermans, J. Molecular dynamics simulation study of polypeptide conformational equilibria: A progress report. Ann. N.Y. Acad. Sci. 482:51–59, 1986.
9. Anderson A., Hermans J., Microfolding: Conformational probability map for the alanine dipeptide in water from molecular dynamics simulations. Proteins 3:262–265, 1988.
10. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. Statistical mechanics and molecular dynamics: The calculation of free energy. In "Molecular Dynamics and Protein Structure." Hermans, J., ed. Western Springs, IL: Polycrystal Book Service, 1985:43–46.
11. Tembe, T.L., McCammon, J.A. Ligand-receptor interactions. Comput. Chem. 8:281, 1984.
12. Mezei, M., Beveridge, D.L. Free energy simulations. Ann. N.Y. Acad. Sci 482:1, 1986. Beveridge, D.L., DiCaputa, F.M. Free energy via molecular simulation. In "Computer Simulation of Biomolecular Systems." Beveridge, D.L., van Gunteren, W.F., eds. Leiden: Escom, 1989:1–26. van Gunsteren, W.F. Methods for calculation of free energies and binding constants: Successes and problems. *ibid.*: 27–59.
13. Dang, L.X., Mertz, K.M., and Kollman, P.A. Free energy calculations on protein stability: Thr-157 → Val-157 mutation of T4 lysozyme. J. Am. Chem. Soc. 111:8505–8508, 1989.
14. Hermans, J., Pathiaseril, A., Anderson, A. Excess free energy of liquids from molecular dynamics simulations. Application to water models. J. Am. Chem. Soc. 110:5982–5986, 1988.
15. As mentioned, bondlengths are kept fixed with Shake. The use of λ^3 and λ^5 as multipliers is equivalent to varying both ϵ and σ^3 as λ when using the common ϵ–σ form of the Lennard–Jones equation.
16. Hermans, J., Yun, R.H., Anderson, A.G. Precision of free energies calculated by molecular dynamics simulations of peptides in solution. J. Phys. Chem., submitted for publication.
17. As in earlier work, we refer to the neighborhoods of the four free-energy minima of the alanine dipeptide as the β-, α_R, α_L, and C_{ax}^7 conformations, using already familiar terms less stringently. (Roughly speaking, each conformation corresponds to a different quadrant of the Ramachandran diagram, e.g., for the β-conformation $-180°<\phi<0$, $0<\Psi<180°$, for the α_R conformation $-180°<\phi<0$, $-180°<\Psi<0°$.)
18. Brant, D.A., Flory, P.J. The configuration of random polypeptide chains. II. Theory. J. Am. Chem. Soc. 87:2791–2800, 1965.
19. Presta, L.G., Rose, G.D. Helix signals in proteins. Science 240:1632–1641, 1988.
20. Schimmel, P.R., Flory, P.J. Conformational energies and configurational statistics of copolypeptides containing L-proline. J. Mol. Biol. 34:105–120, 1968.
21. Richardson, J.S., Richardson, D.C. Amino acid preferences for specific locations at the ends of α-helices. Science 240: 1648–1652, 1988.
22. Lyu, P.C., Liff, M.I., Marky, L.A., Kallenbach, N.R. Side chain contributions to the stability of α-helical structure in peptides. Science 250:669–673, 1990.

23. Computed as the difference for the alanine and proline dipeptides of free energies given by -kTln $\int\int \exp[-\Delta G^{\circ}/kT]d\phi d\Psi$, where Δq° represents conformational free energy.
24. Honzatko, R.B., Hendrickson, W.A., Love, W.E. Refinement of a molecular model of lamprey hemoglobin. J. Mol. Biol. 184:147–164, 1985.
25. Takano, T. Structure of myoglobin refined at 2.0 Å resolution. Crystallographic refinement of metmyoglobin from sperm whale. J. Mol. Biol. 110:537–568, 1977.
26. Dreusicke, D., Karplus, P.A., Schulz, G.E. Refined structure of porcine cytosolic adenylate kinase at 2.1 Å resolution. J. Mol. Biol. 199:359–371, 1988.
27. A very strongly kinked proline-containing helix is found in melittin; in this helix an adjacent glycine residue having a nonhelical conformation serves as a "hinge" which interrupts the hydrogen bonding pattern, but without intervention of a water molecule.[28]
28. Terwilliger, T.C., Eisenberg, D. The structure of melittin. I. Structure determination and partial refinement. J. Biol. Chem. 257:6010–6015, 1982.
29. Piela, L., Némethy, G., Scheraga, H.A. Proline-induced constraints in α-helices. Biopolymers 26:1587–1600, 1987.
30. Karle, I.L., Flippen-Anderson, J., Sukumar, M., Balaram, P. Conformation of a 16-residue zervamicin IIA analog peptide containing three different structural features: 3_{10} helix, α-helix and β-bend ribbon. Proc. Natl. Acad. Sci. U.S.A. 84:5087–5091, 1987.
31. Brandl, C.J., Deber, C.M. Hypothesis about the function of membrane-buried proline residues in transport proteins. Proc. Natl. Acad. Sci. U.S.A. 83:917–921, 1986.
32. Mogi, T., Stern, L.S., Chao, B.H., Khorana, H.G. Structure-function studies of bacteriorhodopsin VIII. Substitutions of the membrane-embedded prolines 50, 91, 186: The effects are determined by the substituting amino acids. J. Biol. Chem. 264:14192–14196, 1989.
33. Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., Downing, K.H. Model for the structure of bacteriorhodopsin based on high-resolution electron cryomicroscopy. J. Mol. Biol. 213:899–929, 1990.