Theoretical Evidence for Water Insertion in α-Helix Bending: Molecular Dynamics of Gly_{30} and Ala_{30} in Vacuo and in Solution

Frank M. DiCapua, S. Swaminathan, and D. L. Beveridge*

Contribution from the Department of Molecular Biology and Biochemistry and Department of Chemistry, Hall-Atwater Laboratories, Wesleyan University, Middletown, Connecticut 06457.
Received December 7, 1990

Abstract: In protein crystal structures, ordered water molecules have been frequently observed at instances where α-helices bend or fold. A series of molecular dynamics (MD) simulations have been carried out in order to investigate the possibility that a water-inserted structure may be implicated in the helix-folding process. Both in vacuo and aqueous solution simulations were performed on a variety of canonical right-handed α-helices; these are of various lengths and are comprised entirely of either alanine or glycine. Analysis of the calculations reveals that the polyglycine helices bend due to the intrinsic nature of the constituent amino acids. For alanine helices, bending occurs in a more ordered manner and involves the insertion of a water molecule within the helical hydrogen bond. The calculations indicate a mechanism by which this might occur: in the course of the vibrational motions of the stabilizing N—H···O═C hydrogen bonds, several water molecules hydrogen bond with both the carboxy and amino groups and serve to weaken the intramolecular hydrogen bonds. When the N—H···O═C hydrogen bond is sufficiently extended, one of the hydrogen-bonded water molecules inserts so as to form a bridge; this insertion must occur according to precise orientation requirements. The resulting three-centered hydrogen bond system serves as an intermediate in the breakup (destabilization) of the α-helix at that position, the onset of microfolding. These studies further support the potential role of water insertion in the protein-folding process.

I. Introduction

Recently, Sundaralingam and Sekharudu[1] surveyed the crystal structures of a number of proteins containing hydrated α-helices. They found a number of instances in which a water molecule is inserted into the N—H···O═C hydrogen bond and correlated this observation with a local destabilization of the helix in the direction of reverse turns and open structures in (Ψ,ϕ) space. They further suggested that the observed water-inserted structures may be intermediates in protein-folding processes. Similar observations have been reported in the crystallography of oligopeptide units by Karle and co-workers[2,3] and by Thornton.[4] We recently reported an extensive molecular dynamics simulation on a polyalanine decamer, Ala_{10}, together with 550 water molecules.[5] This simulation revealed one example of a sequence of residues in which the helix is destabilized by water insertion and thus provided some theoretical corroboration of the idea. To investigate this matter further, we have now carried out a series of molecular dynamics simulations on several different canonical α-helices, both in vacuo and with the explicit inclusion of water. The effect of including water in the simulation and the extent to which water insertion and helix destabilization is observed in the simulation are explored in more detail. The implications of our results in the development of models for protein-folding mechanism are discussed.

II. Background

The influence of water on protein structure[6] and the role of water in protein folding are clearly an important subject in molecular biophysics, but they have been difficult to clarify. The principal experimental data in this area come from the position of ordered water molecules in protein crystal structures. This has provided leading but ultimately fragmentary information, since only a fraction of the surface waters turn out to be crystallographically ordered. NMR evidence for the ordering of water at the protein surface has proven to be interesting, but controversial.[7]

Molecular dynamics (MD) simulation now affords the possibility of theoretical studies including all waters of hydration, but the results are sensitive to the choice of force fields, the simulation protocol, and the convergence characteristics. Nevertheless, MD simulation is presently suitable for conducting computer experiments on components of proteins, and the results can reveal possibilities for the role of water in structure and process. In the study described herein, we construct a molecular dynamics experiment in which we examine the role of water structure on microfolding phenomenon. The calculated results are compared and correlated with those found in crystal structures and suggest a specific mechanism for water-mediated protein-folding processes.

III. Preliminary Results

We review here the results from a previous study from this laboratory pertinent to the present investigation. A 130-ps molecular dynamics simulation was carried out on an alanine decapeptide surrounded by 550 water molecules in a hexagonal prism cell.[5] The results were analyzed in a form which corresponds as closely as possible to that adopted by Sundaralingam and Sekharudu, using the geometrical conventions described in Figure 1, which is also applicable to the present study. This allows the results of the simulation to be directly compared to crystallographic observations, wherein the location of the hydrogens is not known. The primary diagnostic index of α-helix formation is the distance $d2$, the separation between the carbonyl oxygen and the amide nitrogen in the $(O—C)_i\cdots(N—H)_{i+4}$ hydrogen bond. The distances $d1$ and $d3$ respectively represent those distances between the carbonyl oxygen and the amide nitrogen to the center of the water oxygen. The value of $d2$ in the canonical α-helix is 3.2 Å. Analysis of the MD simulation revealed that most all of the distances, including $d2$, were seen to be oscillatory and stable. Occasional displacements of ca. 0.5 Å were seen, but these were short-lived. When the structures were reordered according to monotonically increasing $d3$, henceforth called the $d3$ structural index ($d3$si),

(1) Sundaralingam, M.; Sekharudu, Y. C. *Science* **1989**, *244*, 1333.
(2) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Biopolymers* **1989**, *28*, 773.
(3) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 299.
(4) Blundell, T.; Barlow, D.; Borkakoti, N.; Thornton, J. *Nature* **1983**, *306*, 281.
(5) DiCapua, F. M.; Swaminathan, S.; Beveridge, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 6768.
(6) Edsall, J. T.; McKenzie, H. A. *Adv. Biophys.* **1983**, *16*, 53.
(7) Otting, G.; Wütrich, K. *J. Am. Chem. Soc.* **1989**, *111*, 1871.

0002-7863/91/1513-6145$02.50/0

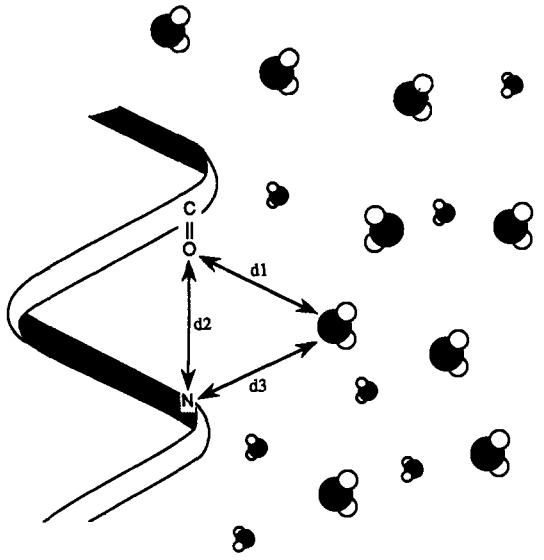

Figure 1. Schematic drawing illustrating the diagnostic parameters $d1$, $d2$, and $d3$.

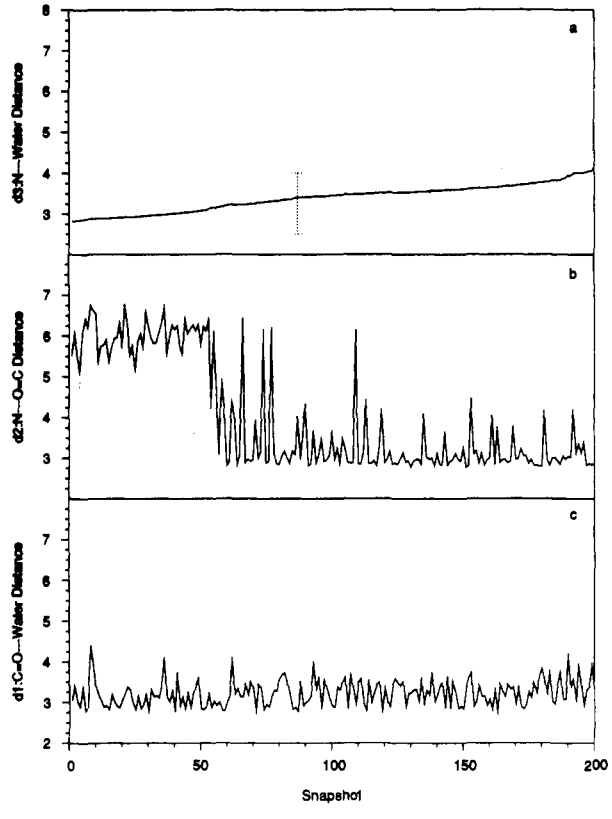

Figure 2. Structural behavior of $d1$, $d2$, and $d3$ for the ^{1}Ala···^{5}Ala hydrogen bond. Four thousand snapshots were taken over the 100 ps of trajectory; in each snapshot, the closest water was used in the calculation of distances. Distances were sorted in the ascending order of the $d3$ parameter, and every 20 points were averaged; the location of 3.4 Å is noted by the dotted line.

$d2$ and $d1$ were also seen to be stable and oscillatory, indicating the helix is dynamically stable in this region. The range of values seen for $d1$ indicates that the acceptor site for hydrogen bonding on the carbonyl oxygen was always saturated.

All of the α-helix hydrogen bonds behaved dynamically in this manner with the exception of the terminal hydrogen bond, which separated over the course of the trajectory. The behavior of $d1$, $d2$, and $d3$ vs $d3$si for this case is shown in Figure 2. The value of $d3$ = 3.4 Å was chosen by Sundaralingam and Sekarudu as

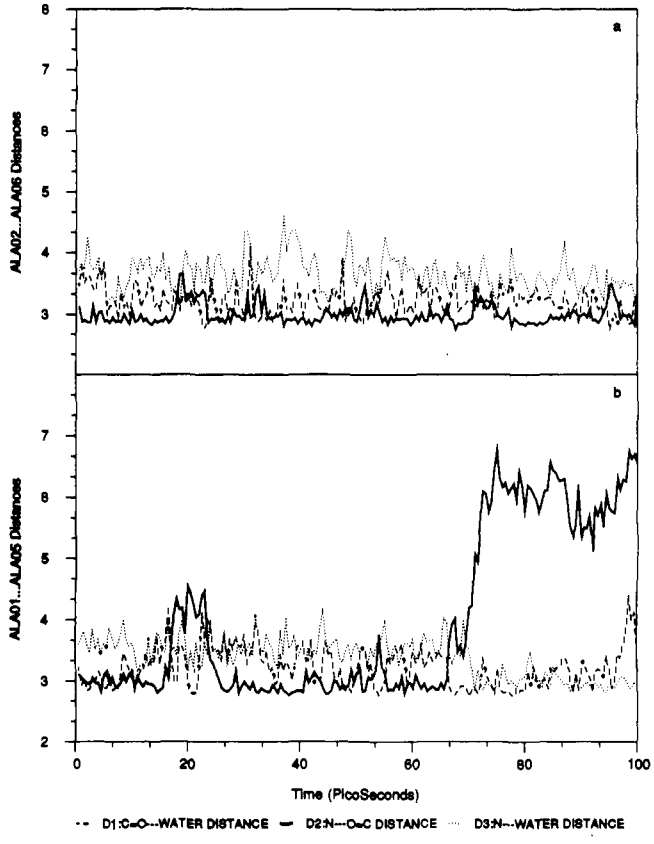

Figure 3. Dynamical behavior of $d1$, $d2$, and $d3$ distances for both ^{2}Ala···^{6}Ala and ^{1}Ala···^{5}Ala.

the threshold value for a water entering the first hydration shell.[1] Below this value of $d3$, one has "water-inserted" structures. We observe that when $d3 < 3.4$ Å, the corresponding distance $d2$ moves into a dynamical regime of large values, indicating helix destabilization. Destabilization of the α-helix in a single dynamic trajectory is thus marked by similarities to structures observed as "water-inserted" in the ensemble of crystallographic structures by Sundaralingam and Sekharudu, c.f., Figure 2B of ref 1. The behavior of $d2$ above 3.4 Å in Figure 2b shows that some structures in this range also correspond to destabilized helices. We determined that this arises from a destabilization that persists after an inserted water has departed.

Sundaralingam and Sekharudu proposed that helix destabilization occurs only when a water is inserted such that $d1$, $d2$, and $d3$ achieve threshold values necessary for forming a "three-centered case". The calculated time evolution of these distances in the simulation (Figure 3) support this proposal, as it is only when $d1$, $d2$, and $d3$ intersect that either the transient (~15 ps) or the persistent (~70 ps) destabilization is observed. Examination of the orientation of the water during the transient destabilization reveals that although a water is "inserted", a bridge structure is not actually formed, i.e. the orientation of the water is not correct for supporting a bridge between the C=O and N--H. In the 70-ps destabilization, the orientation of the inserted water is such that a bridge of the form C=O···H—OW···H--N is clearly present, and the destabilization still persists at the termination of the trajectory.

In the above MD simulation, the observed destabilization of an α-helix hydrogen bond indicates that such a phenomenon can happen, but the possibility that an end effect causes the terminal hydrogen bonds to be significantly weaker than the central ones could not be eliminated. Also, we wish to differentiate a water-mediated microfolding event from one in which the helix folds intrinsically and is followed by a migration of the surrounding water molecules into the resulting cavity. A series of additional MD simulations on several different canonical α-helices were performed and analyzed in order to explore in more detail the

participation of water in the microfolding process.

The two prototype examples chosen for this study are α-helixes consisting of polyglycine and polyalanine, which present extreme cases of helix stability. Glycine, due to its intrinsic conformational flexibility in (ψ,ϕ) space, is a helix breaker and will degrade to a lower energy secondary structure. In such a case, the solvent need not initiate the helical bending; rather, the individual waters will simply readjust their position around the helix as it changes its conformation. Alanine is well-known to be a helix former, and thus tends to stabilize helical secondary structure. Moreover, recent evidence has shown that helices comprised of alanine are stable in an aqueous environment.[8] Thus for polyalanine helices, the solvent may take on a more active role, and initiate the bending event. Computer experiments based on MD simulation can provide new knowledge about these possibilities. In order to distinguish intrinsic folding processes from ones that are water mediated, all simulations are performed in vacuo as well as in an aqueous environment.

IV. Calculations

All simulations are performed with WESDYN 1.0, a locally developed molecular simulations program that has been optimized for performance on the CRAY YMP supercomputer. All simulations begin with a right-handed α-helix that has been constructed on the basis of canonical values for the (ψ,ϕ) angles. The molecular topology of our system is determined by using PROGMT from GROMOS86[9] and converted into WESDYN format. For the simulations involving water, we use WESDYN's PROSOL in order to hydrate the solute with a 9.0-Å layer of water molecules in a hexagonal prism cell at 1-g/cm^3 density. This unit cell is treated under periodic boundary conditions throughout the simulations. The theoretical object of study is thus a representation of the dilute aqueous solution of the helix.

All simulations use the GROMOS86 force field for the oligopeptides[9] and the SPC model for water.[10] SPC water gives good agreement with experiment for the water–water radial distributions in the pure liquid[11] and a dielectric constant of 68[12] in reasonable accord with experiment. Switching functions are used to make the nonbonded interactions of all types go smoothly to zero between 7.5 and 8.5 Å; this is consistent with the parametrization of SPC water. Vacuum simulations use a dielectric constant of 1. In order to avoid artificially splitting dipoles, the nonbonded pair list is formed on a group by group basis, where the groups correspond to sets of atoms comprising well-defined chemical subunits.

The preparatory phase involves equilibrating the system to an energy minima corresponding to room temperature. As a first step to this, a Monte Carlo calculation of 30 000 passes is performed on the water in the field of the fixed helix; this equilibrates the water into a favorable configuration around the protein. Once the solvent–solvent and solvent–solute energies are stabilized, the entire system is briefly minimized with 50 steps of the conjugate gradient technique. This step serves to release any strain that might be present in the bonds and bond angles of the solute, as well as to remove any unfavorable van der Waals contacts that might exist within the system. In order to perform the simulation at room temperature, kinetic energy must now be added to the system through a process that is referred to as heating. Thus, the molecular dynamics calculation is begun by slowly heating the entire system to an energy corresponding to 300 deg K. This method involves a slow input of kinetic energy by randomly assigning the velocities of each of the atoms in the system. These velocities are constantly reassigned so as to remove any bias that might be introduced by an unfavorable random number; this has the effect of distributing energy equally among all of the atoms while assuring that the initial geometry of the system does not drastically change due to the increased heat input. This process, Gaussian heating and equilibration, is performed over 5 ps. Once the system achieves an energy corresponding of 300 K, the randomization of velocities is terminated, and the energy is controlled by scaling the system temperature back to 300 K whenever its average fluctuates more than 5 deg. This equilibration process serves eventually to fine tune the system into a thermally bound energy minimum. The amount of time over which this scaling must be performed is system dependent, but it usually lasts for a period of 5–25 ps.

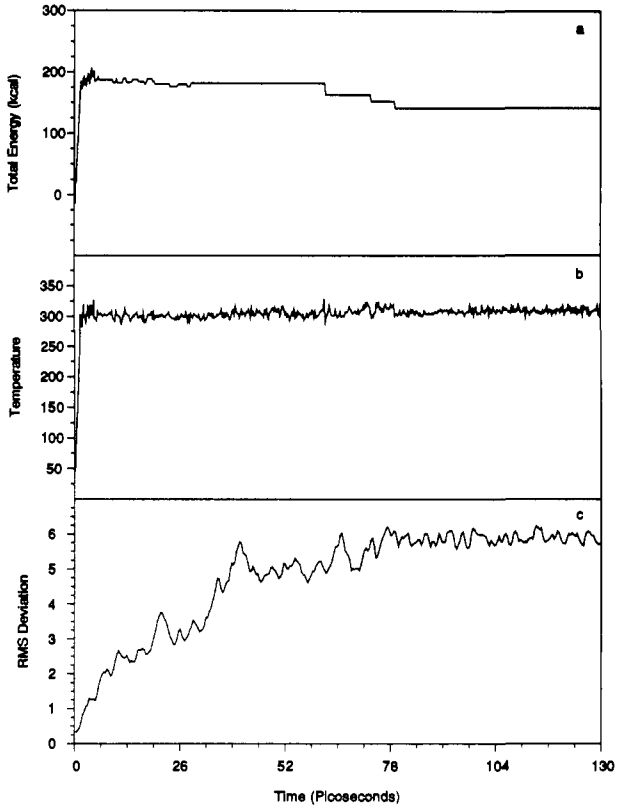

Figure 4. Convergence profiles for the Gly_{30} vacuum simulation: (a) total energy of the system remains constant for relatively long stretches; (b) system temperature; (c) root-mean-square deviation.

In the trajectory phase, the temperature window for velocity rescalings is increased to 10 deg, and the system is allowed to move under free dynamics; this portion of the calculation produces the results for detailed analysis. The stability of the calculation can be monitored by examining the convergence profiles of total energy and temperature. Constant energy is maintained throughout the entire trajectory portion of the simulation, and low-temperature fluctuations are observed. All of the simulations presented here conserved energy and remained naturally within the 10-deg temperature window.

In addition to monitoring the time evolution of the critical distance parameters for helix stability and water insertion, an analysis of hydrogen bond strength was performed on all sequences and is monitored over the course of the trajectory. The relative hydrogen bond strength is determined on the basis of the angle of the hydrogen bonding vectors of both N--H and C=O, as well as the distance between them. In each case, we have searched for all possible hydrogen bonding modes: $i\cdots(i+4)$, $i\cdots(i+3)$, and $i\cdots(i+5)$.

V. Results and Discussion

Glycine$_{30}$. The convergence profiles for the in vacuo simulation on Gly_{30} are shown in Figure 4. Figure 4a shows that after the initial heating and equilibration steps, the total energy of the simulation undergoes a number of rescalings, but ultimately remains constant for 50 ps with no rescaling of velocities. These constant stretches of energy translate into stable temperature segments (Figure 4b) in which the temperature never fluctuates more than 10 deg from the 300 K target temperature. The corresponding changes in energy correspond to transitional alterations within the structure. The root-mean-square deviation of the dynamical structure (Figure 4c) changes quite a bit as the calculation proceeds, but ultimately stabilizes.

The corresponding changes in the helix can be seen in the sequence of snapshots from the simulation depicted in Figure 5, which show ribbon diagrams of the polyglycine structure over the

(8) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5286.

(9) van Gunsteren, W. F.; Berendsen, H. J. C. GROMOS86: Groningen Molecular Simulation Program, University of Groningen, 1986.

(10) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, H. J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, Holland, 1981; p 331.

(11) Beveridge, D. L.; Mezei, M.; Mehrotra, P. K.; Marchese, F. T.; Ravishanker, G.; Vasu, T. R.; Swaminathan, S. In Molecular Based Study and Prediction of Fluid Properties. Haile, J. M., Mansoori, G., Eds. *Adv. Chem.* **1983**, *204*, 297.

(12) Alper, H. E.; Levy, R. M. *J. Chem. Phys.* **1989**, *91*, 1242.

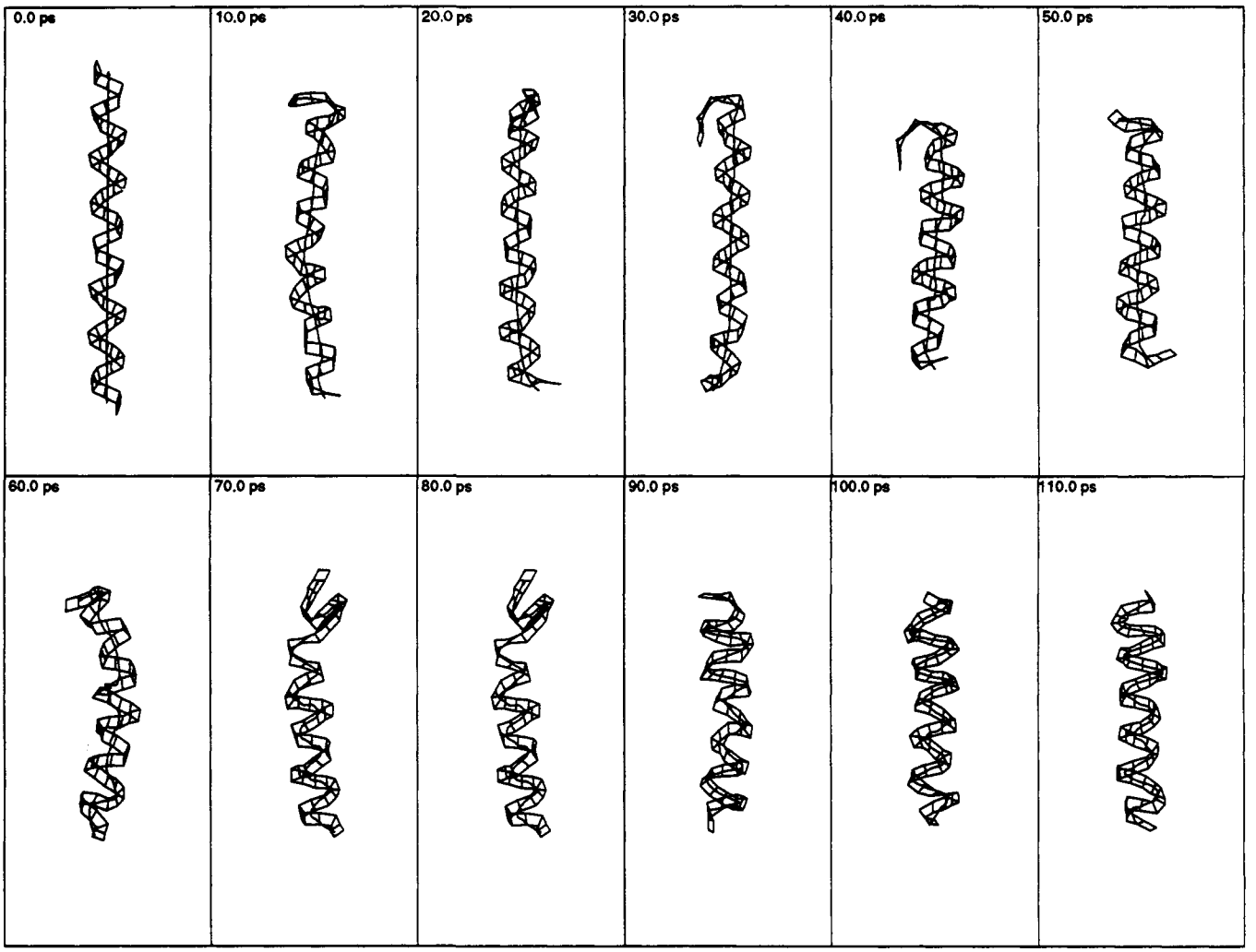

Figure 5. Ribbon snapshots representing equally spaced intervals over the 130 ps of the Gly_{30} vacuum simulation; the ribbon represents the secondary structure of the helix, and the inset line represents the helical axis.

130 ps of the simulation. This ribbon of consecutive peptide planes describes the secondary structure of the helix, and the inset solid line represents the helical axis as obtained from Curves analysis.[13] Soon after the dynamics has begun, and even before the equilibration is finished, we see that the helix begins to contract somewhat in a manner typical of in vacuo simulations. There is a certain floppiness at the amino terminal end, and small curvatures at various points along the helical axis. As the trajectory continues on, the axis deformation continues and become more prevalent. Between 50 and 70 ps into the simulation, a transition to a supercoiled regime is observed. After rescaling, the simulation becomes dynamically stable at ~80 ps and remains so up to the termination of the simulation. The behavior of the dynamical structure can be seen more clearly in Figure 6a. Here 26 structures, taken at equal intervals over the course of the simulation, are superimposed onto one another; the dark bold line running down the center corresponds to the starting canonical structure. The earlier structures are seen as the extended fluctuating forms that align more closely with the axis of the reference structure, and the later contracted helices are seen as shortened supertwisted forms. The dynamical behavior of the in vacuo helical axis is paralleled by that of the hydrogen bonds, which quickly convert from those of an α-helix to those of a 3_{10} helix soon after the trajectory phase of the simulation has begun.

The corresponding convergence profiles for the simulation of solvated polyglycine (Gly_{30} + 889 H_2O) are shown in Figure 7. The root-mean-square deviation from the reference structure (Figure 7c) is not as large as that of the vacuum simulation. This is likely a result of a frictional force provided by the surrounding water environment; the helix will not be able to move as readily as in the in vacuo situation because the waters act to dampen many of the fluctuations and motions attempted by the helix. For similar reasons, the total energy of the simulation (Figure 7a) remains constant after the initial heating and equilibration steps. The decreased motion of the helix results in a corresponding decrease in the production of excess kinetic energy, and the small amounts that are produced are readily absorbed and redistributed by the solvent. This behavior translates to a constant system temperature which never fluctuates more than 5 deg from 300 K. Note, however, that while the system temperature remains constant, that of the solute oscillates around the 300 K system temperature due to the transfer of energy between the solute and the solvent.

The dynamical structure of the hydrated polyglycine helix (Figure 6b) remains in an α-helical form except for the amino end and a four-residue stretch between ^{18}Gly and ^{21}Gly, which is characterized by both a helical curvature and a breaking of the hydrogen bonds. This portion is bordered by small stretches in which the helix tends toward the 3_{10} form. Figure 8 shows the $d3$si plots resulting from the analysis of one hydrogen bond within the curved stretch (^{18}Gly···^{22}Gly), and a recalculation in order to consider 3_{10} hydrogen bonding (^{18}Gly···^{21}Gly) is shown in Figure 9. Since $d1$ and $d2$ fluctuate extensively on either side of the $d3$ threshold value in either case, there is no correlation between the approach of a water molecule and the behavior of the hydrogen bond. The bond breakage is thus due to the intrinsic tendency of polyglycine to bend: it first forms a 3_{10} helix, and then bends

(13) Sklenar, H.; Etchebest, C.; Lavery, R. *Proteins: Structure, Function Genet.* **1989**, *6*, 46.

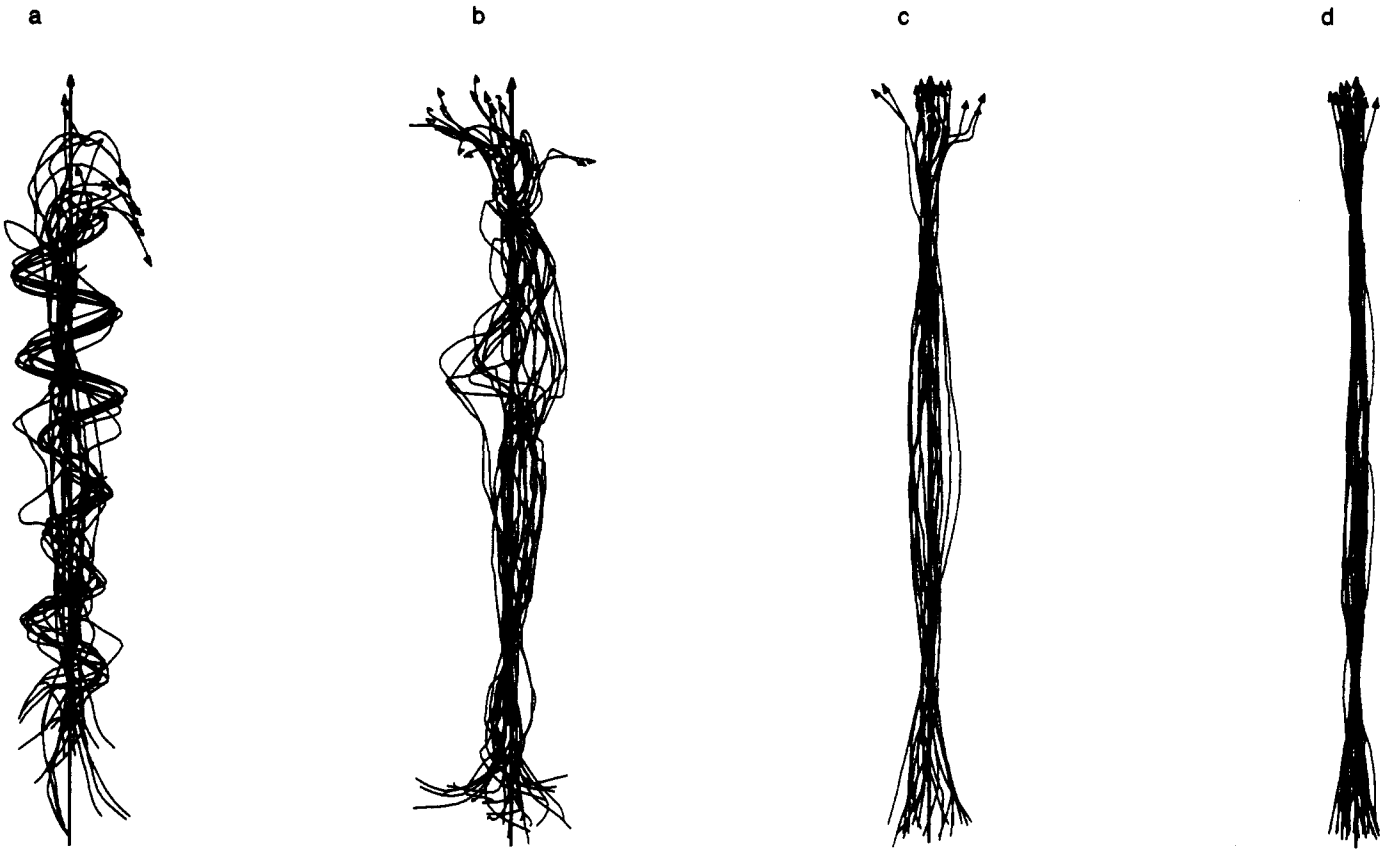

Figure 6. Helical axes for the various simulations that are performed. In all cases, 26 structures are taken at equally spaced intervals over the course of the simulation and are superimposed on one another; the dark bold line running down the center corresponds to the reference structure. (a) 130-ps simulation of Gly_{30} under vacuum; (b) 130-ps simulation of solvated Gly_{30}; (c) 130-ps simulation of Ala_{30} under vacuum; (d) 280-ps simulation of solvated Ala_{30}.

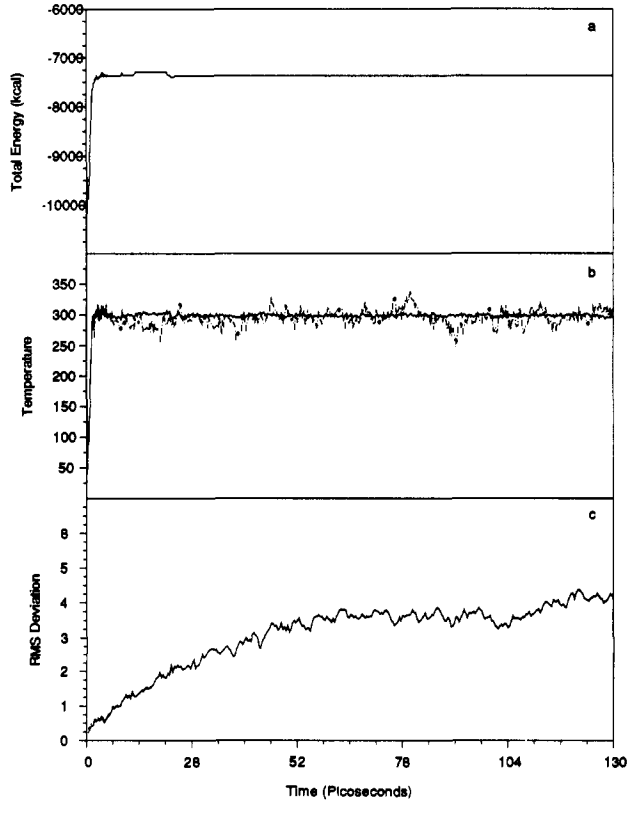

Figure 7. Convergence profiles for the Gly_{30} solvent simulation: (a) total energy of the system remains constant; (b) system temperature is represented by the solid line, and the solute temperature is denoted by the dashed line; (c) root-mean-square deviation.

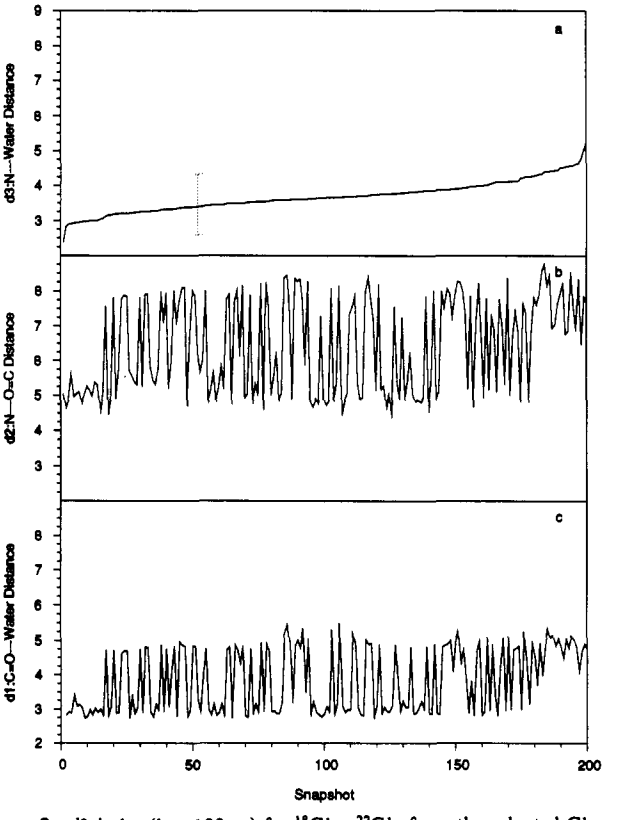

Figure 8. *d*3si plot (last 100 ps) for $^{18}Gly\cdots^{22}Gly$ from the solvated Gly_{30}.

due to an intrinsic effect rather than as a consequence of an interaction with water. This result is in sharp contrast with the

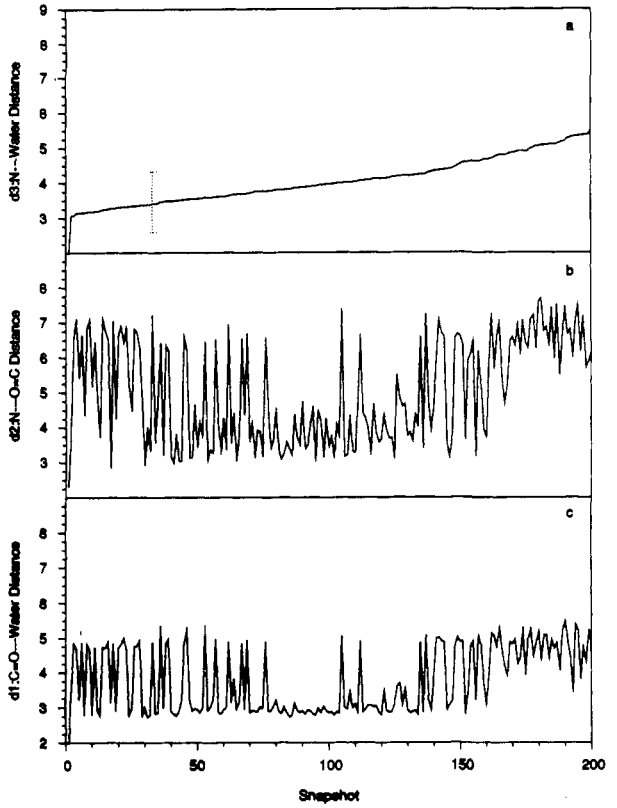

Figure 9. A recalculation of the *d*3si plot (last 100 ps) of the solvated Gly_{30}. This plot represents the $^{18}Gly\cdots^{21}Gly$ hydrogen bond when taking a 3_{10} helix into consideration for the analysis.

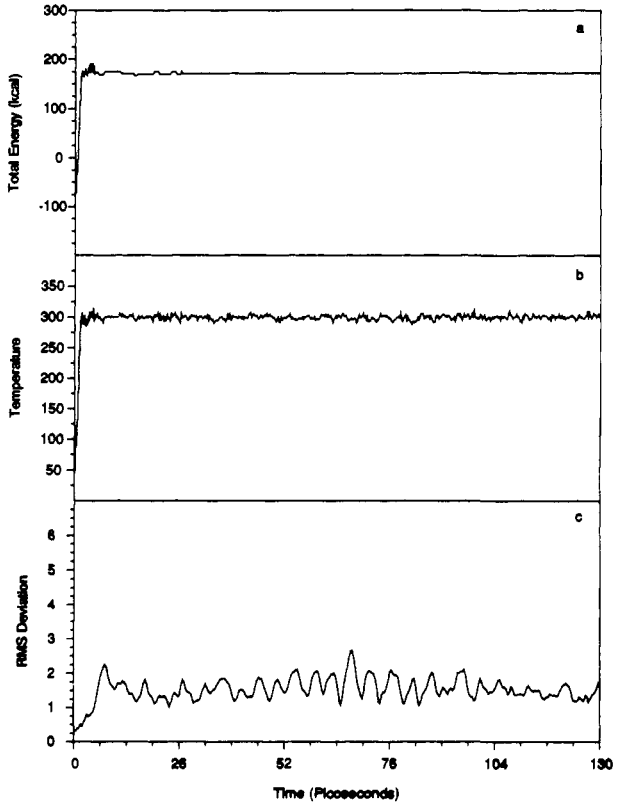

Figure 10. Convergence profiles for the Ala_{30} vacuum simulation (a) total energy of the system remains constant; (b) temperature; (c) root-mean-square deviation.

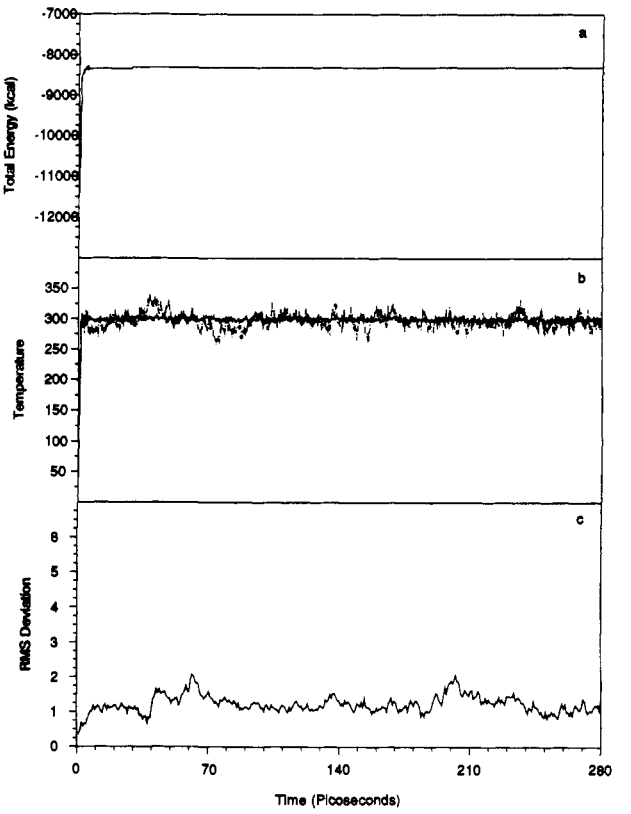

Figure 11. Convergence profiles for the Ala_{30} solvent simulation: (a) total energy of the system remains constant; (b) temperature of the solute (dashed line) oscillates around the temperature of the system (solid line); (c) root-mean-square deviation.

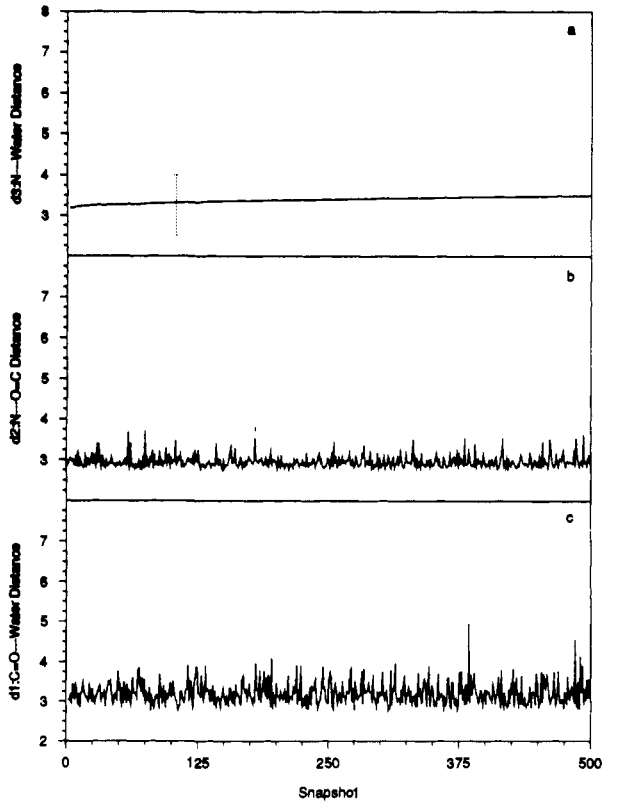

Figure 12. *d*3si plots for $^{22}Ala\cdots^{26}Ala$ from the solvated alanine 30-mer that is run at 300 K for 250 ps. Note that the *d*2 hydrogen bonding distance does not change appreciably as the water approaches.

decaalanine case, in which the corresponding *d*1 and *d*2 behavior is sensitive to the threshold and the curvature arises from an interaction with a water molecule. If the calculation were allowed to proceed for a much longer period of time, it is likely that the

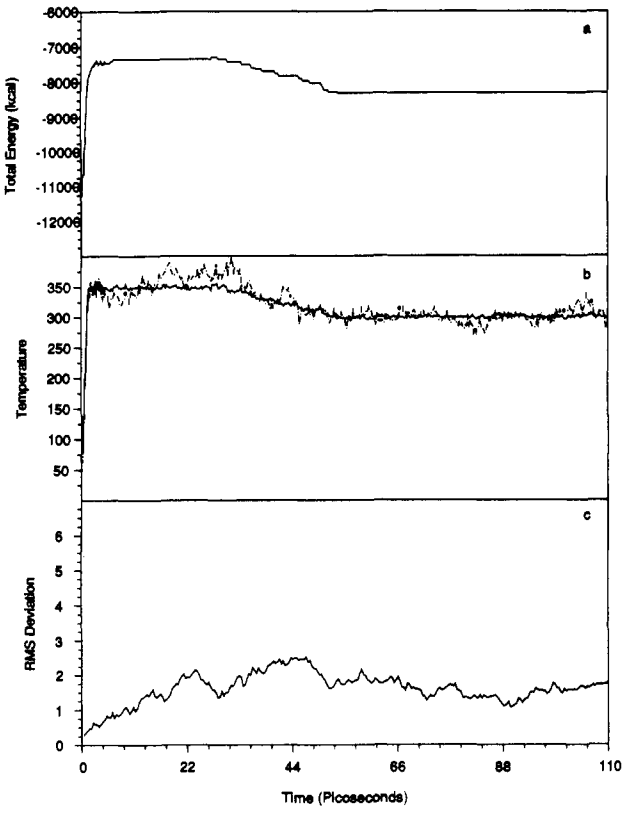

Figure 13. Convergence profiles for the Ala_{30} solvent simulation that is heated to 350 K and allowed to remain there for 20 ps: (a) total energy of the system remains constant once it has been cooled back down; (b) temperature—system temperature is shown as a solid line and the solute temperature is represented by a dashed line; (c) root-mean-square deviation.

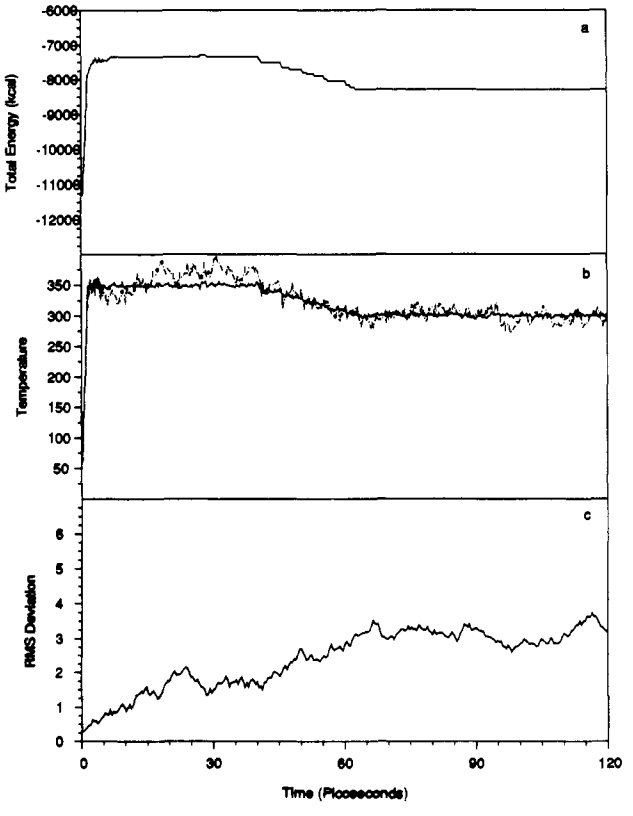

Figure 14. Convergence profiles for the Ala_{30} solvent simulation that is heated to 350 K and allowed to remain there for 30 ps: (a) total energy of the system remains constant once it has been cooled back down; (b) temperature; (c) root-mean-square deviation.

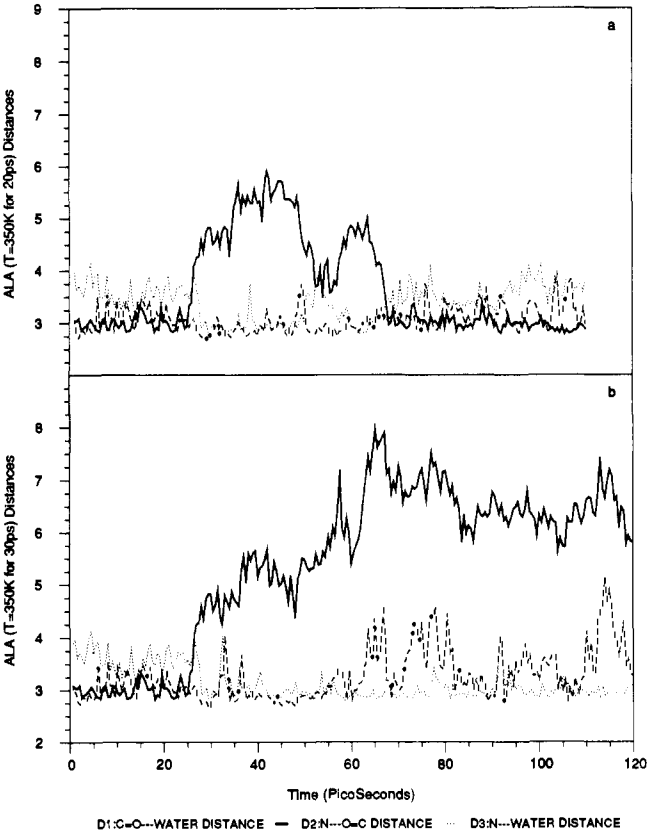

Figure 15. Time course of the distances for the $^{22}Ala\cdots^{26}Ala$ hydrogen bond during the solvated Ala_{30} simulation at 350 K. Plot (a) represents the simulation that is heated to 350 K and left there for 20 ps. Note that a transient destabilization occurs at ~25 ps, but later reforms. Plot (b) represents the simulation that is heated to 350 K and left there for 30 ps; in this case, the hydrogen bond destabilizes and remains that way even as the system is cooled back down to 300 K.

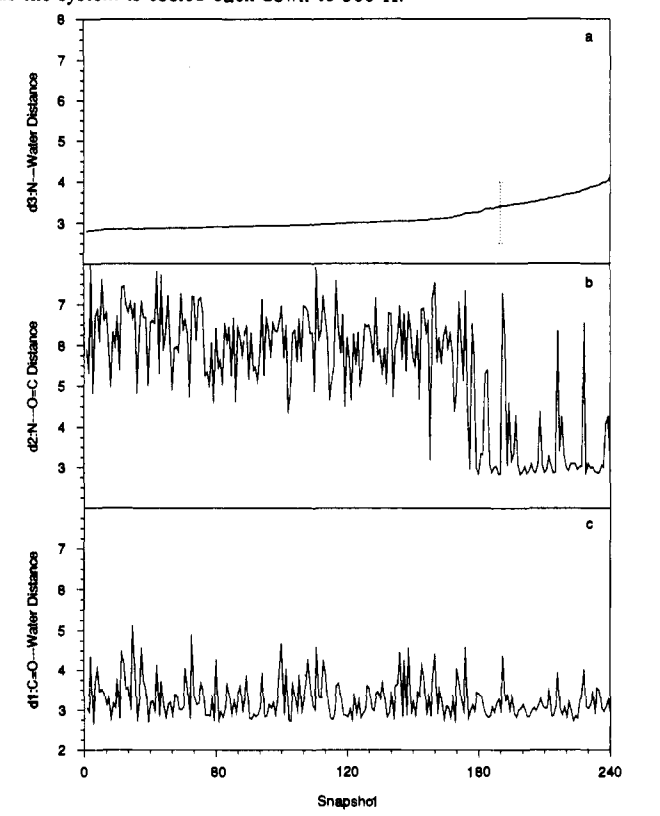

Figure 16. *d*3si plots for $^{22}Ala\cdots^{26}Ala$ from the solvated alanine 30-mer that is heated to 350 K and left there for 30 ps. In this case, the hydrogen bond destabilizes when the water molecule approaches and remains that way even as the system is cooled back down to 300 K.

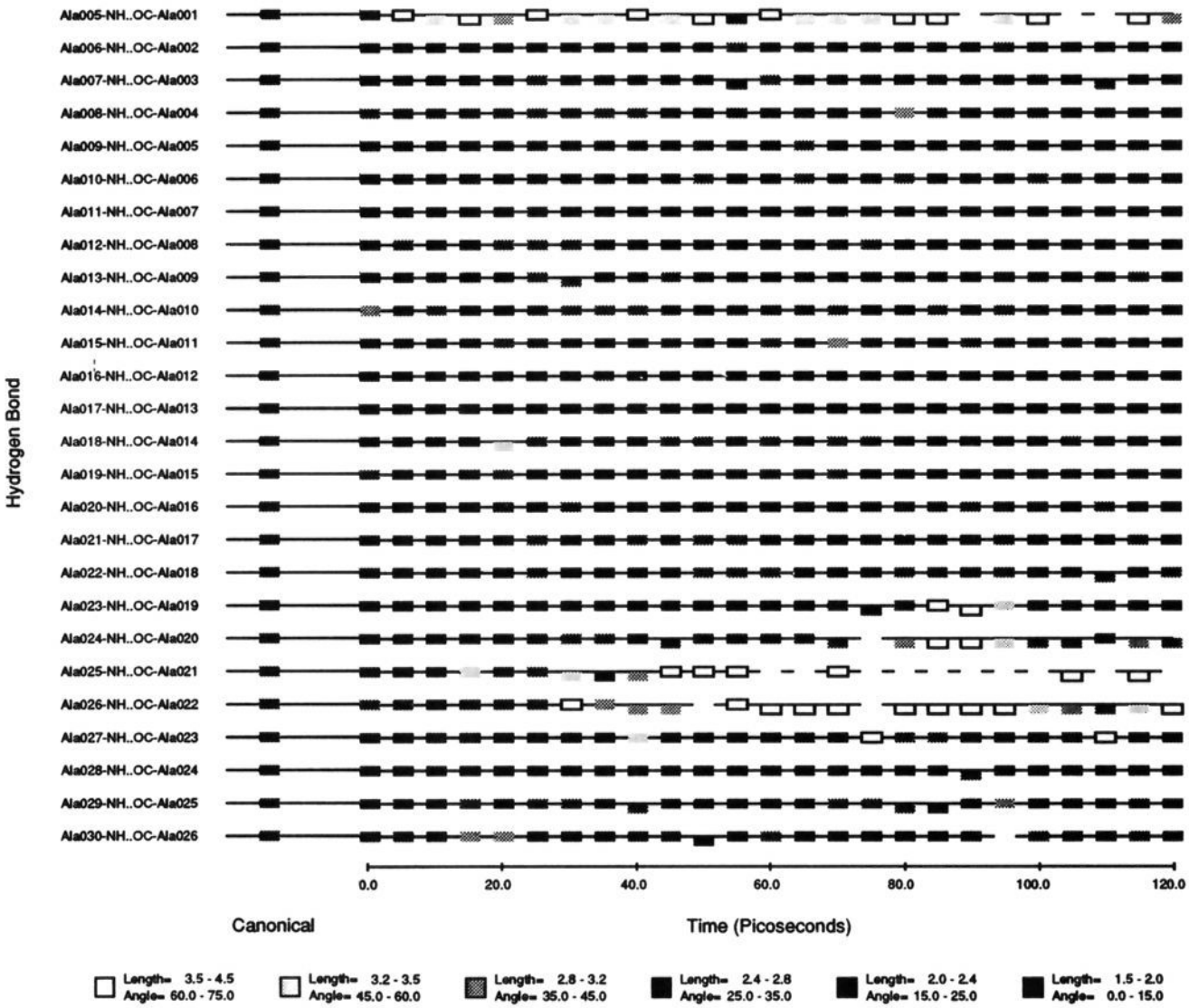

Figure 17. Plot of hydrogen bond strength for the Ala_{30} solvent simulation in which the system was kept at 350 K for 30 ps and then cooled down: The 26 different hydrogen bonds are represented on the y axis, and the passage of time is shown on the x axis; the entire 120 ps of the simulation (including heating and equilibration) is represented. The relative hydrogen bond strength is determined on the basis of both the angle of the hydrogen bonding vectors (N--H and C=O) and the distance between them; these are represented by different levels of gray. Each line shows the behavior of the hydrogen bond over the 120-ps time course, and each small rectangle represents the bond strength at that interval within the trajectory. In each case, all possible hydrogen-bonding modes are searched for and displayed. If the rectangle is centered on the line, the hydrogen bond is an $i\cdots(i+4)$ hydrogen bond; an $i\cdots(i+5)$ bond is shown by drawing the box above the center line, and an $i(i+3)$ hydrogen bond is represented by having the rectangle appear below the line. Note that the local hydrogen-bonding structure changes from $1 \rightarrow 4$ to $1 \rightarrow 3$ type just before the rupturing of the helix.

polyglycine helix would continue to lose its α-helical nature until only a very few small stretches remained. The resulting conformation would be chiefly random structure, consistent with the helix breaking tendency of the glycine residue.

Alanine$_{30}$. The convergence profiles for the vacuum simulation of alanine$_{30}$ are shown in Figure 10. As in the previous simulations, the overall energy remains constant during the trajectory portion of the calculation (30–130 ps); during this time, the temperature never fluctuates more than 10 deg from the 300 K simulation temperature, and the root-mean-square deviation very clearly oscillates with a periodic motion. The periodic motion corresponds to a very stable structure in which the various degrees of freedom oscillate about a mean value. This observation is confirmed by the superimposition of helical axes shown in Figure 6c. Unlike the vacuum simulation of glycine, all of the polyalanine axes remain very close to the initial structure, and no superhelical twist arises. Although there is some bending in the helix, all of the hydrogen bonds remain strong and do not break. Hydrogen bond analysis on this simulation found all of the bonds to remain stable as $i\cdots(i+4)$ hydrogen bonds; the exception to this is the amino termina! hydrogen bond, which is probably an end effect.

In the MD simulation of the solvated form (Ala_{30} + 997 H_2O), we see from the convergence profiles (Figure 11) that the frequency of the fluctuations of the helix is reduced compared to that in the in vacuo case. This is observed in the plot of overall root-mean-square deviation (Figure 11c) as well as in the superposition of the helical axes (Figure 6d). This simulation was run for a period of 250 ps, but produces no helical bending; this must be due to the synergistic effect of the greater number of hydrogen bonds. None of the hydrogen bonds ever separate very much from the canonical values, and $d3si$ plots for all of the hydrogen bonds are unremarkable. One example of such behavior is shown in Figure 12, but all remaining hydrogen bonds behave in the same manner. There is no separation of the $d2$ distance corresponding to the approach of a water molecule. In fact, the effect of the solvent seems to be such that the helix is stabilized even more; hydrogen bond analysis shows that the bonds increase in strength as the simulation proceeds to the 250-ps mark, and a superimposition of the helical axes shows that the tendency for the fluctuations is dampened by the surrounding waters.

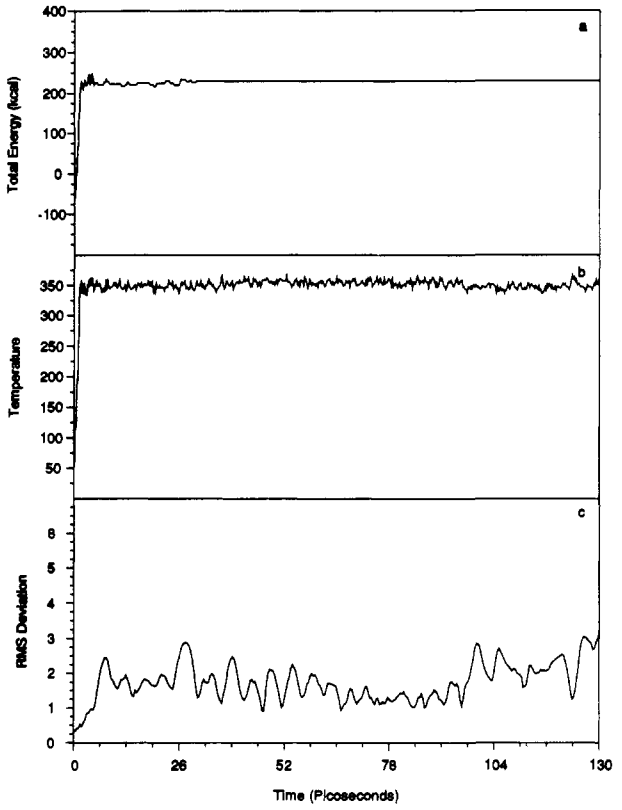

Figure 18. Convergence profiles for the Ala_{30} vacuum simulation at 350 K: (a) total energy of the system remains constant; (b) temperature; and (c) root-mean-square deviation.

Alanine$_{30}$ (T = 350 K). Even though the α-helix of Ala_{30} was stable at 300 K, it is of interest to explore if breakage will eventually occur. However, the length of time necessary for us to observe such an event will be very long and would require an extensive amount of computer cycles. Rather than run the simulation for sustained periods, we chose to provide some smaller amount of energy into the system through the use of simulated annealing. The procedure for simulated annealing is very similar to that of the heating process, and its objective is to heat the system temperature to a value higher than that which would normally be used as the simulation temperature. This higher temperature simulation allows the system to explore the energy surface more quickly because it has sufficient energy to surmount any small energy barriers that might exist at 300 K. Although this elevated temperature could be assigned any value, we did not increase it too much; rather, the system temperature was increased to a quantity adequate for the components to overcome the small barriers that would be normally transversed if it were given a sufficient amount of time.

In the particular calculation carried out here, Gaussian heating and equilibration methods were applied in order to slowly heat the system up to 350 K rather than the 300 K. The system was left at this temperature with a 5 deg temperature window for a length of time adequate for forming a three-centered water bridge and is then slowly cooled back down to 300 K over a series of five small steps. In each of these steps, the system temperature is cooled by 10 deg and allowed to equilibrate at that temperature for a period of 5 ps. In this way, there is no sudden influx of energy into or out of the system, and the solute is hopefully allowed to experience the same motions that it customarily would, but in a much shorter time scale than would be normally required.

Two such simulations were performed, and the convergence profiles are similar to the previous ones in that there is low fluctuation of the temperature and the energy is conserved. In one (Figure 13), the system is slowly heated up to 350 K and allowed to remain there for a period of 20 ps. Here, there was a temporary destabilization in the region of ^{21}Ala···^{25}Ala and ^{22}Ala···^{26}Ala which may be traced to the formation of a three-centered water bridge: however, this bond reformed as soon as the system was cooled back down to room temperature. In the other case (Figure 14), the system was left at 350 K for a period of 30 ps; a three-centered water bridge also formed, but the hydrogen bond did not reform once the system was cooled down. Figure 15 demonstrates the behavior of the ^{22}Ala···^{26}Ala hydrogen bond in both cases. When the helix is left at 350 K for 20 ps, this bond is transiently destabilized as a result of the bonds interaction with a water molecule; however, this bond quickly reforms as the system is cooled back down to 300 K. When the system is left at 350 K for 30 ps, however, the helix continues to bend even after the system is cooled down to 300 K. The behavior of ^{22}Ala···^{26}Ala (for the 30 ps, T = 350 K, simulation) is shown in Figure 16; ^{21}Ala···^{25}Ala behaves in a similar fashion (not shown). While the bond reforms as the system is cooled down in the 20-ps case, in the 30-ps case it does not because the N--H and C=O groups have moved too far apart and water molecules have already come in to fill the gap caused by the bend.

Figure 17 illustrates the hydrogen bond analysis from this simulation. This figure represents the 26 different hydrogen bonds on the y axis and the time passage on the x axis. The relative hydrogen bond strength is determined on the basis of both the angle of the hydrogen bonding vectors (N--H and C=O), and the distance between them. Each line shows the behavior of the hydrogen bond over the 120 ps of trajectory, and each small rectangle represents the bond strength at that interval within the trajectory. In each case, all possible hydrogen bonding modes are searched for and displayed. If the rectangle is centered on the line, the hydrogen bond is an i···(i + 4) hydrogen bond; an i···(i + 5) bond is shown by drawing the box above the center line, and an i···(i + 3) hydrogen bond is represented by having the rectangle appear below the line.

Hydrogen bond analysis on both of these ruptured helices demonstrates that both helixes change their local hydrogen-bonding structure from a 1 → 4 to a 1 → 3 (or, 3_{10}) type just before the bond breaks. This local structure change may or may not be due to the effect of the water molecule because the vacuum simulation of Ala_{30} at 350 K (convergence profiles are seen in Figure 18) likewise shows this same tendency to alter structure toward 1 → 3 and 1 → 5 hydrogen-bonding patterns for short periods of time. This action is not localized to any specific point along the helix, but rather it occurs randomly throughout the structure, a manifestation of failed attempts at bending. Because α-helixes are rarely found to be very long, this type of behavior is probably due to a persistence length effect; however, no discriminatory reason exists for the helix to fold at a specific point, and so short intervals of weakness are exhibited throughout the entire structure. This weakness is not seen in the 300 K simulation because the higher temperature of the 350 K simulation allows this helix to explore its local minimum more quickly. Although this local impermanent behavior exists, the tendency of the helix is for it to remain stable throughout the 350 K vacuum simulation, and it shows no signs of curvature on its own. Thus, the bending of the polyalanine is not an intrinsic phenomenon, and there is only a very small weakness that arises as a result of its persistence length. Moreover, the water is not able to drive the bending on its own; this is evidenced by the fact that the 300 K solvent simulation does not result in bending. Therefore, it is only through the cooperation of both the solvent and the helix that the bending occurs; the water molecule functions by taking advantage of a prevailing weakness within the helix and serves to facilitate the bending process.

While the carboxy end of the solvated helix in Ala_{30} (T = 350 K) ruptures; a d3si plot of this bond (not shown) is very similar to that of the polyglycine hydrogen bond breakage. Thus this destabilization is likely due to an intrinsic end effect rather than to a water insertion and can be ignored.

VI. Summary and Conclusions

Molecular dynamics simulations have been successfully performed on α-helices of glycine and alanine oligopeptides, both in

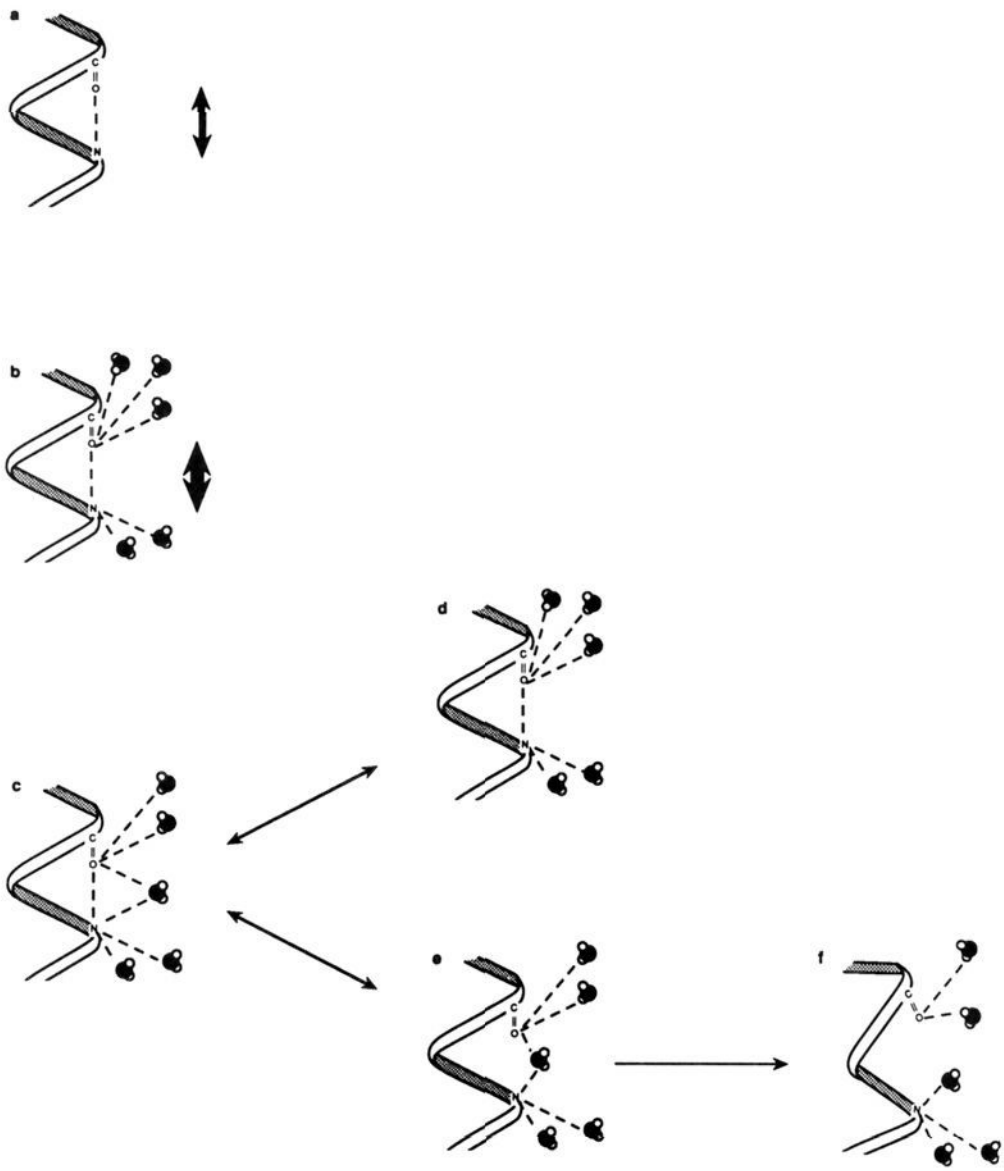

Figure 19. Model for Helix Destabilization. While the helix undergoes a range of vibrational motions, the stabilizing hydrogen bonds undergo a breathing motion such that the length is periodic and oscillatory about a mean bond length (a). Some portion of the helix will contain an inherent weakness that is due to the persistence length, local geometry, or electrostatic field. As this is happening, water molecules are hydrogen bonding to both the amino and the carboxy groups along the helix (b); this serves to weaken the bonds further. As the breathing arrives to a point wherein the bond is very large, one of the attached water molecules realigns itself with the hydrogen so as to form a bridge between the carboxy and amino group of one hydrogen bond (c). This three-centered hydrogen bond serves as an intermediate which can revert back to an unbridged structure (d), or it can continue on to cause a transiently destabilized hydrogen bond (e). This transiently destabilized form can then result in a bending of the helix (f), or it can eventually revert back to an unbridged structure (d). Structures b and d are identical in this schematic.

vacuo and with the explicit inclusion of water. The results on polyglycine show that the helix bends soon after the trajectory commences, regardless of whether the water is present in the simulation. Thus, the results from the theoretical model system are consistent with the observed helix-breaking tendencies of the glycine residue. For the polyalanine case, the α-helix remains intact at room temperature, for both the in vacuo and solvated simulations. This result is consistent with the helix-forming tendency of the alanine residue.

The simulations on the polyalanine were subsequently carried out at 350 K to see if helix bending could be induced and, if so, to study the mechanism. At this slightly higher temperature, some points of instability in the helix are found. Detailed analysis reveals that the bending occurs with a cooperation of both intrinsic fluctuations in the helix hydrogen bonds and orientation specific water insertions.

On the basis of the inspection of the trajectory snapshots from these series of experiments, we have constructed a mechanism (Figure 19) by which this folding might occur: The α-helical structure undergoes a range of vibrational motions; among these are those of the stabilizing hydrogen bonds which undergo a breathing motion such that the length is periodic and oscillatory about a mean bond length. Some point along the helix will contain an inherent weakness such that the hydrogen bonds associated with this region will be less strong and will oscillate with greater frequency and amplitude. This weakness will be sequence dependent and will be manifested in either the persistence length, local geometry, or the local electrostatic environment. In our case, this weakness probably arises as a result of the persistence length of alanine, but different types of residues will result in different persistence lengths. It might also be possible that a certain length of the helix will be less stable due to the constituent amino acids (based upon both the electrostatics and geometry, some amino acid sequences are better helix formers than others) in that region. As this is happening, water molecules hydrogen bond all along the helix to either the carboxy or amino groups of the substituent

amino acids (in all of our simulations, the waters show a greater tendency to bind at the carboxy groups); this serves to weaken the bonds even further. As the breathing arrives to a point wherein the hydrogen-bonding distance is very large, one of the attached water molecules realigns itself with the hydrogen so as to form a bridge between the carboxy and amino group of one hydrogen bond. This three-centered hydrogen bond serves as an intermediate which then causes the α-helix to break at that position and begin to fold up. This intermediate structure can then move in either direction: either back to the intact helix or toward a further separation. As it continues to separate, a local perturbation will occur such that the bent structure will be better accommodated; this seems to be a tendency toward a 3_{10} helix.

Alanine and glycine provide examples of two extremes that are involved in the α-helical microfolding process, and we have been able to explain the role that water plays in their helical bending. Other sequences will want to bend due to a combination of these two extremes, and so the role played by the water will demonstrate a corresponding dependence: Water will act in a passive fashion, providing a frictional force that dampens the motions of the helix (as in the case of polyglycine); it will also serve a more active role as a tool for guiding the overall bending pathway (as in the case of polyalanine). How much either of these two extremes contribute to the overall role of the bending will be dependent on the sequence of the helix and on its local environment.

It is worthwhile to mention the fact that proteins do not fold by forming one long chain that later goes on to bend at certain areas; rather, they will form some type of intermediate that consists of some initial secondary and tertiary structure. The role that will be played by the solvent must be balanced with the accessibility of the helix within this intermediate form as well as the local environment that is bestowed upon the helix by the surrounding transitionary structure. This environment will further weaken or strengthen various points along the helix and affect the role of the water within the binding process. Thus, while the internal geometry will play a crucial role in the microfolding process, so to must the local nonbonded environment that is projected by the surrounding protein, water, and any other constituents of the system.

Acknowledgment. This research is supported by NIH Grant No. GM-37909 from the National Institutes of Health. The calculations were made possible by a generous allocation of CRAY YMP time at the Pittsburgh Supercomputer Center. F.M.D. is supported by a Traineeship in Molecular Biophysics via NIH Grant No. GM-08271.

Registry No. Gly30, 134110-05-7; Ala30, 109376-45-6; H_2O, 7732-18-5.

The Dehydration of Glyoxylate Hydrate: General-Acid, General-Base, Metal Ion, and Enzymatic Catalysis[1a]

J. E. Meany[1b] and Y. Pocker*

Contribution from the Department of Chemistry, BG-10, University of Washington, Seattle, Washington 98195. *Received November 9, 1990*

Abstract: The lactate dehydrogenase catalyzed reduction of glyoxylate by NADH was studied at 25.0 °C and an ionic strength of 0.15 by using a spectrophotometric method. We have demonstrated that this reaction is a sequential one that requires prior dehydration of glyoxylate hydrate. Accordingly, reaction rates extrapolated to infinite LDH concentrations were used to determine the rate of hydration of glyoxylate hydrate. The dehydration reaction was shown to be susceptible to general-acid and general-base catalysis in phosphate buffers and to catalysis by transition-metal ions and by the high-activity zinc metalloenzyme carbonic anhydrase (CA II) from bovine erythrocytes. The Brønsted constant, β, was estimated from the catalytic rate coefficients of H_2O, HPO_4^{2-}, and OH^- and compared with those for the reversible hydration reactions of other substrates of varying electrophilicity. The magnitude and relative order of catalytic efficiency of divalent transition-metal ions, $Zn^{2+} > Cu^{2+} > Co^{2+} > Ni^{2+} > Cd^{2+} > Mn^{2+}$, were compared with other reactions in which the substrate has the capacity to bind these cations. For the dehydration of glyoxylate a mechanism involving the metal ion in the direct transfer of a water molecule and at the same time acting as a general acid is proposed. The CA II catalyzed dehydration of glyoxylate hydrate was determined as a function of pH. The data show that the dependency of enzymatic rate on pH is similar to that for the dehydration of bicarbonate, and accordingly suggest certain parallels in mechanisms for their bovine carbonic anhydrase catalyzed dehydrations.

Introduction

Glyoxylate is an anion of biochemical significance in both plants and animals. In plants and in bacteria the glyoxylate cycle provides a means of utilizing acetyl coenzyme A for the ultimate production of carbohydrates and amino acids. In mammalian brains, glyoxylate is produced from glycine by D-amino acid oxidase.[2] It has been proposed that glyoxylate may have a controlling effect on various metabolic processes in mammals.[3] It is known to inhibit oxidative metabolism, and adducts of glyoxylate with various nucleophiles appear to serve as physiological substrates for a number of oxidases. For example, thioacetates serve as substrates for L-hydroxy acid oxidase to form oxalyl thioester compounds, which may serve as intracellular messengers for insulin and other hormones.[3a-c]

In aqueous solution, glyoxylate exits almost entirely as its *gem*-diol, glyoxylate hydrate:[4]

$$\begin{matrix}\mathrm{CHO}\\ |\\ \mathrm{CO_2^-}\end{matrix} + \mathrm{H_2O} \rightleftharpoons \begin{matrix}\mathrm{CH(OH)_2}\\ |\\ \mathrm{CO_2^-}\end{matrix} \qquad K_{eq} = 163 \qquad (1)$$

It has been shown that lactate dehydrogenase (EC 1.1.1.27, L-lactate; NAD oxidoreductase) catalyzes the oxidation of glyoxylate to oxalic acid:

(1) (a) Support of this research by grants from the National Science Foundation and the Muscular Dystrophy Association is gratefully acknowledged. (b) Visiting scientist under the National Science Foundation Research Opportunity Award Program, Summer 1988. Permanent address: Department of Chemistry, Seattle University, Seattle, WA 98122.

(2) de Marchi, W. J.; Johnston, G. A. R. *J. Neurochem.* **1969**, *16*, 355.

(3) (a) Gunshore, S.; Brush, E. J.; Hamilton, G. A. *Bioorg. Chem.* **1985**, *13*, 1. (b) Hamilton, G. A.; Beatty, S. M. *Bioorg. Chem.* **1985**, *13*, 14. (c) Harris, R. K.; Hamilton, G. A. *Biochemistry* **1987**, *26*, 1.

(4) (a) Debus, H. *J. Chem. Soc.* **1904**, *85*, 1382. (b) Rendina, A. R.; Hermes, J. D.; Cleland, W. W. *Biochemistry* **1984**, *23*, 5148. (c) Craig, M. M.; Baff, M.; Gresser, M. J. *J. Am. Chem. Soc.* **1988**, *110*, 3302.

0002-7863/91/1513-6155$02.50/0