Nanosecond Time Scale Folding Dynamics of a Pentapeptide in Water†

Douglas J. Tobias,‡,§ John E. Mertz,∥ and Charles L. Brooks III*,‡

Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, and Cray Research, Inc., 655-E Lone Oak Drive, Eagan, Minnesota 55121

Received February 11, 1991; Revised Manuscript Received April 15, 1991

ABSTRACT: Reverse turns, four-residue sections of polypeptides where the chain changes direction by about 180°, are thought to be important protein folding initiation structures. However, the time scale and mechanism for their formation have yet to be determined experimentally. To develop a microscopic picture of the formation of protein folding initiation structures, we have carried out a pair of 2.2-ns molecular dynamics simulations of Tyr-Pro-Gly-Asp-Val, a peptide which is known to form a high population of reverse turns in water. In the first simulation, which was started with the peptide in an ideal type II reverse turn involving the first four residues, the turn unfolded after about 1.4 ns. After about 0.6 ns in the second simulation, which was started with the peptide in a fully extended conformation, the peptide folded into a type II turn which had a transient existence before unfolding. The peptide remained unfolded for another 0.9 ns before folding into a type I turn involving the last four residues. The type I turn lasted for about 0.2 ns before unfolding. Thus, these simulations showed that protein folding initiation structures can form and dissolve on the nanosecond time scale. Furthermore, the atomic-level detail of the simulations allowed us to identify some of the interactions which can stabilize the folded structures. The type II turns were stabilized by either a salt bridge between the terminal groups or a backbone–C-terminal group hydrogen bond, and the type I turns were stabilized by a hydrophobic interaction between the proline and valine side chains.

The molecular mechanism by which proteins fold remains one of the greatest unsolved mysteries of molecular biology. Many current popular theories of protein folding are based on a simple sequential model (Kim & Baldwin, 1982), according to which folding occurs through a series of well-defined intermediates. Several experimental studies have demonstrated the existence of such intermediates (Kim & Baldwin, 1990). The intermediates for which detailed structural information is available contain extensive, nativelike secondary structures (Udgaonkar & Baldwin, 1988; Roder et al., 1988). However, since these intermediates are fully formed on the fastest time scale of the experiments, the mechanism and time scale for their formation are unknown. One possibility is that marginally stable structures—such as secondary structural elements [e.g., nascent helical structures (Dyson et al., 1988a) or reverse turns (Dyson et al., 1988b) and/or hydrophobic clusters (Bundi et al., 1978)]—form first and then they coalesce and grow, resulting in the first observable intermediate (Wright et al., 1988). This hypothesis was suggested by the observation that potential folding initiation structures (e.g., reverse turns) in short linear peptides are sufficiently stable under folding conditions to be detected by nuclear magnetic resonance (NMR) spectroscopy (Montelione et al., 1984; Williamson et al., 1986; Dyson et al., 1988b). However, the mechanism and time scale for the formation of these transient folding initiation structures have also yet to be deduced experimentally.

In principle, molecular dynamics (MD) simulations can complement experimental studies of protein folding by providing detailed, atomic-level pictures of motions on time scales which are too short to be probed with present experimental techniques. However, because of the high computational cost of simulating biological molecules in solution, to date, MD simulations of proteins or peptide fragments in water have been limited to a few hundred picoseconds of simulated time (Brooks et al., 1988). Although the unfolding and refolding of a reverse turn (Wright et al., 1990) and the unfolding of a helix (Tirado-Rives & Jorgensen, 1991) in solution have been directly observed in simulations that are a few hundred picoseconds long, evidently that is not long enough to observe folding initiation events in unfolded peptides. Clearly, it will be some time before the gap between the experimental (milliseconds) and simulation (hundreds of picoseconds) time scales is closed, and a complete picture of protein folding dynamics is obtained. Nonetheless, as computing power increases, the gap between the time scales narrows.

In this paper, we report our preliminary findings from two MD simulations of a small peptide in water, each performed for more than 2 ns (1 ns = 1000 ps). It turns out that the order of magnitude extension of simulated time is sufficient to enable us to see both the formation and dissolution of possible folding initiation structures (e.g., reverse turns) in a small peptide in water. Thus, our results provide insight regarding the mechanism and time scale of initial folding events. In addition, the atomic-level detail of the simulations allows us to identify some of the sequence-specific, intrapeptide interactions, such as salt bridges and hydrogen bonds, which might stabilize folding initiation structures in solution.

We carried out a pair of 2.2-ns MD simulations of the pentapeptide, Tyr-Pro-Gly-Asp-Val in water at 300 K. We chose this peptide for study because it is known to form a high population (roughly 50%) of reverse turns in water at 278 K (Dyson et al., 1988b). The first simulation, which we will refer to as the "turn" simulation, was started with the peptide in an ideal type II reverse turn (Venkatachalam, 1968; Chou & Fasman, 1977) conformation, with the chain reversal formed by residues 1–4 (the numbering begins at Tyr, the N-terminal residue). This turn conformation is the type observed in the

† This work was partially supported by an NIH grant to C.L.B. (GM37554). C.L.B. is an A. P. Sloan Foundation Fellow, 1990–1992. D.J.T. was supported by an NIH predoctoral training grant and an internship at Cray Research, Inc.

‡ Carnegie Mellon University.

§ Present address: Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104.

∥ Cray Research, Inc.

0006-2960/91/0430-6054$02.50/0

experiments of Dyson et al. (1988b). The second, or "extd" simulation, was started with the peptide in an extended, parallel β-sheet conformation (Creighton, 1983). We hoped to see the reverse turn "unfold" during the "turn" simulation and the extended structure "fold" into a reverse turn during the "extd" simulation (note: throughout this paper, we will refer to reverse turns as "folded" species, extended conformations as "unfolded" species, and their interconversion as "folding/unfolding"). We thought it was important to run both simulations so that, if the peptide only unfolded once in the "turn" simulation and/or only folded once in the "extd" simulation, we could discuss the folding and unfolding events separately rather than naively assuming one was simply the reverse of the other. More importantly, we wanted to see if our model peptides were capable of both folding from a fully extended conformation, and unfolding from a well-defined folded conformation, on the time scale of our simulations.

Materials and Methods

The "turn" and "extd" trajectories were generated by using similar simulation protocols. The peptide molecules were prepared for the simulations by building them in either a reverse turn or a extended conformation, and then minimizing their energy with harmonic constraint potentials on the backbone dihedral angles to keep them near their initial conformations. Next, we simulated each peptide, with the constraints intact, for 150 ps [the Verlet (1967) algorithm with a time step of 0.0015 ps was used in all the simulations described herein] in a dielectric continuum ($\epsilon = 50$) at a temperature of 300 K to allow the side chains and terminal groups to equilibrate. Finally, each peptide was placed in the center of a 29-Å cubic box containing approximately 750 water molecules and was simulated with periodic boundary conditions for 2.2 ns at 300 K. The CHARMM peptide parameters (Brooks et al., 1983) and the TIP3P water model (Jorgensen et al., 1983) were used. The SHAKE constraint algorithm (Ryckaert et al., 1977) was used to keep the water molecules rigid and to maintain rigid N–H bonds in the peptide molecules. The nonbonded interactions were processed by using a list-based algorithm (Allen & Tildesley, 1989), and the lists were updated every 10 time steps. The nonbonded energies and forces were smoothly truncated at 10 Å, based on atomic centers, according to the minimum image convention (Allen & Tildesley, 1989). The coordinates of the entire peptide/solvent system were stored every 100 time steps during each simulation. All of the computations described in this paper were done with a version of the CHARMM program which we optimized for vector/parallel execution on Cray supercomputers. Each peptide/solvent simulation required about 400 h of cpu time on a Cray YMP supercomputer.

Results and Discussion

The pentapeptide existed in ideal type II reverse turn conformations, with the turn involving residues 1–4, for more than 1300 ps during the "turn" simulation (see Figure 1). The time evolution of the reverse turn is illustrated by the time series of the C_1^α–C_4^α and O_1–H_4 distances in Figure 2 [reverse turns can be identified by using the criterion that the C_i^α–C_{i+3}^α distance is less than 7 Å (Lewis et al., 1973); hydrogen-bonded turns typically have O_i–H_{i+3} distances less than 3.5 Å (Chou & Fasman, 1977)]. The portion of the backbone running from the α-carbon of Gly-3 to the C-terminus (the "exiting" chain) was initially built parallel to the portion of the backbone running from the N-terminus to the α-carbon of Pro-2 (the "entering" chain), and it remained that way for about the first 250 ps of the "turn" simulation. Then it moved so that it was

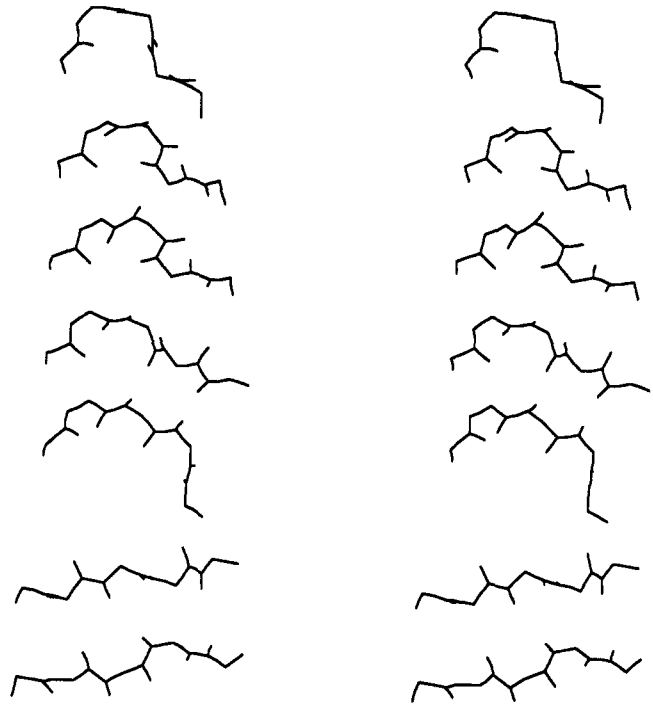

FIGURE 1: Stereoview of a sequence of structures (only the peptide backbone atoms are shown) from the "turn" simulation illustrating the unfolding of a type II reverse turn. The structures correspond to the following simulation times (in picoseconds): top, 1350; 2nd from top, 1386; 3rd from top, 1392; middle, 1403; 3rd from bottom, 1488; 2nd from bottom, 1650; bottom, 1670. The total simulation time covered in this figure is 320 ps. All of the molecular structures shown in this paper were drawn by using the program MOLX (Sneddon, 1990).

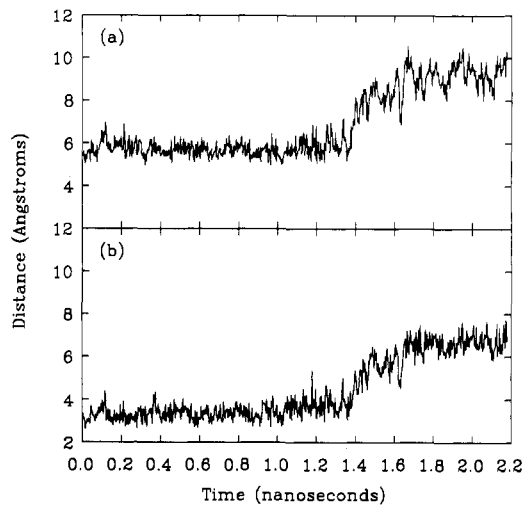

FIGURE 2: Time series from the "turn" simulation: (a) distance between the Tyr C^α and Asp C^α atoms; (b) distance between the Tyr carbonyl O and Asp amide N atoms.

perpendicular to the entering chain, and shortly thereafter the side chain of Asp-4 replaced it as the exiting chain in the reverse turn. This type of structure persisted for most of the remainder of the lifetime of the chain reversal. After about 1350 ps of simulation time, the chain reversal began to significantly distort and become more extended until around 1490 ps, where it gave way to "U-shaped" structures containing a chain reversal involving residues 1–5. These "U-shaped" structures were in rapid equilibrium with mostly extended structures for approximately 150 ps. Finally, after about 1650 ps, the peptide backbone became fully extended and remained that way for the remainder of the 2.2-ns simulation. To

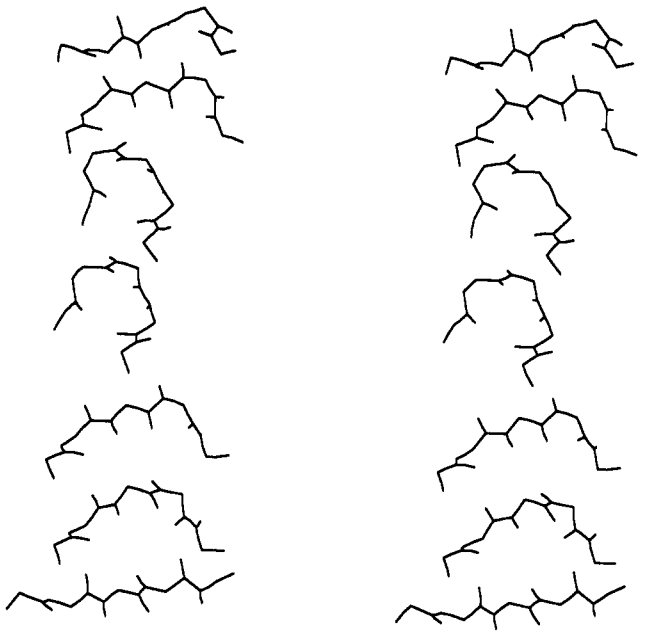

FIGURE 3: Stereoview of a sequence of structures from the "extd" simulation illustrating the peptide folding into a type II reverse turn involving residues 1–4. The turn has a transient existence before unfolding. The structures correspond to the following simulation times (in picoseconds): top, 534; 2nd from top, 548; 3rd from top, 569; middle, 570; 3rd from bottom, 584; 2nd from bottom, 626; bottom, 672. The total simulation time covered in this figure is 138 ps.

summarize the "turn" simulation, we say that the peptide began to unfold after about 1350 ps and that the unfolding was complete after about 1650 ps.

In the "extd" simulation, the peptide remained in fully extended conformations, like sections of β-strand, for the first 500 ps (see Figure 3). At around 530 ps, a kink formed in the chain at the α-carbon of Asp-4, creating "L-shaped" structures, which were in rapid equilibrium with extended structures for about 15 ps. The "L-shaped" structures subsequently turned into "U-shaped" chain reversal structures similar to those observed in the "turn" simulation. The "U-shaped" structures lasted about 20 ps before suddenly changing into a type II reverse turn, like the starting structure for the "turn" simulation, after approximately 570-ps total simulation time. The type II reverse turn structures were transient, lasting only a few picoseconds, before changing back into "U-shaped" structures. This transient folding event is barely discernible in the time series of the C_1^α–C_4^α distance (not shown) as a small "spike" reaching down to ~7.4 Å from a base line of values in the range 8–10 Å. During the next 50 ps, the "U-shaped" structures were in rapid equilibrium with mostly extended structures. The peptide was fully extended again after roughly 670 ps from the start of the simulation, and remained that way for another 700 ps. Before describing the latter part of the "extd" simulation, we summarize the first part: after 570 ps, the extended peptide folded into a type II reverse turn, which had a transient existence before completely unfolding in another 100 ps. In the language of reaction rate theory, the first half of the "extd" trajectory is an example of a "multiple crossing". The peptide free energy was sufficient to allow the peptide to cross the barrier separating the folded and unfolded state wells. However, rather than being quenched in the folded state well, the peptide was deflected back across the barrier and subsequently quenched in the unfolded state well. Finally, comparing Figures 1 and 3, we can see that similar "intermediates", the "U-shaped" structures, were observed in both the folding and unfolding of the peptide

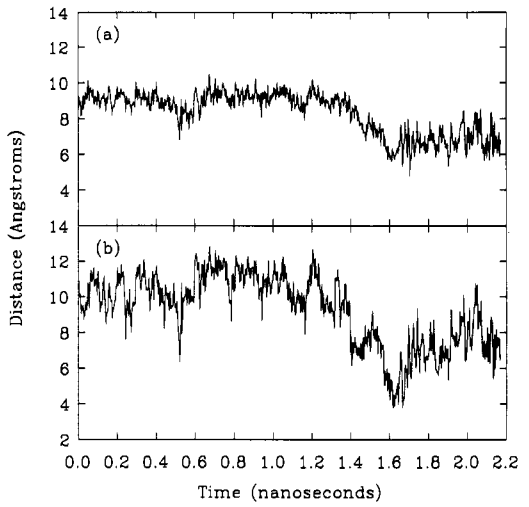

FIGURE 4: Time series from the "extd" simulation: (a) distance between the Pro C^α and Val C^α atoms; (b) distance between the Pro side chain C^γ and Val side chain C^β atoms.

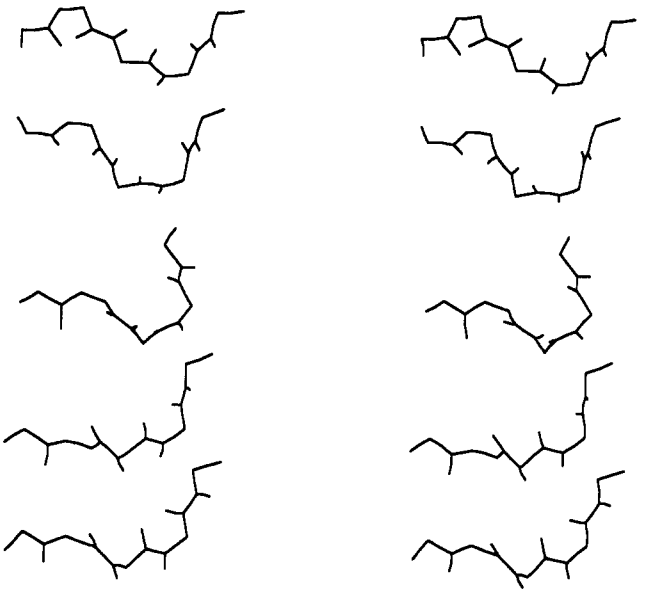

FIGURE 5: Stereoview of a sequence of structures from the "extd" simulation illustrating the peptide folding into a type I reverse turn involving residues 2–5. The turn persists for about 100 ps before unfolding. The structures correspond to the following simulation times (in picoseconds): top, 1503; 2nd from top, 1532; middle, 1629; 2nd from bottom, 1804; bottom, 1917. The total simulation time covered in this figure is 414 ps.

during the "extd" simulation and in the unfolding during the "turn" simulation.

After approximately 1500 ps during the "extd" simulation, another chain reversal began to form, this time involving residues 2–5 (see Figure 5; also see Figure 4a, which shows the time evolution of the C_2^α–C_5^α distance). About 30 ps later, the peptide conformation developed into a distorted type I reverse turn (Venkatachalam, 1968; Chou & Fasman, 1977). During the subsequent 200 ps, there was a rapid interconversion of distorted and well-defined type I turns. The turns disappeared altogether at around 1800-ps total simulation time, giving way to "L-shaped" structures similar to those observed in the "turn" simulation. The "L-shaped" structures were in rapid equilibrium with mostly extended conformations for another 100 ps, before the peptide became fully extended and remained that way for the remainder of the "extd" simulation. In summary, during the second half of the "extd" simulation,

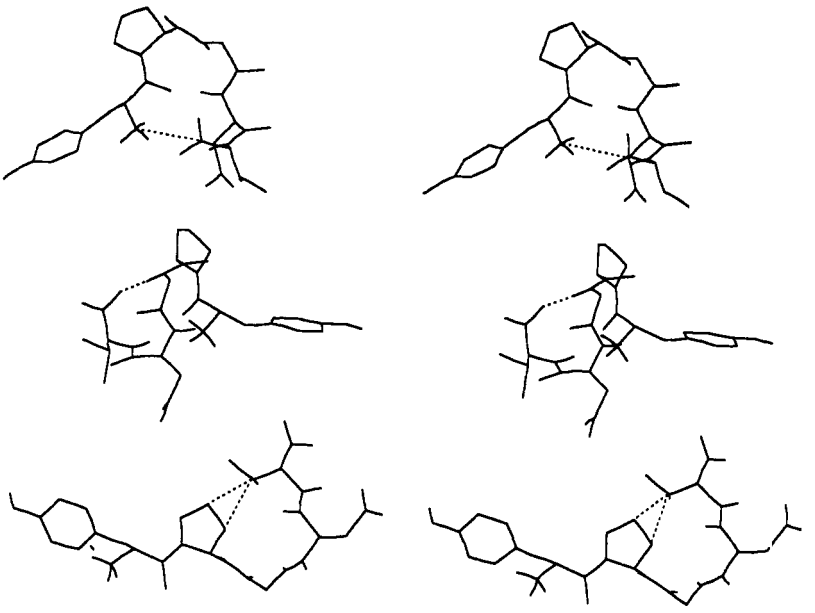

FIGURE 6: Stereoview of preferred intrapeptide interactions (indicated by dotted lines) in folded structures of Tyr-Pro-Gly-Asp-Val: (top) a salt bridge between the charged terminal groups in a type II reverse turn; (middle) a hydrogen bond between the backbone NH group of Gly-3 and a terminal CO group of Val-5; (bottom) a hydrophobic interaction between the pyrrolidine ring of Pro-2 and the side chain of Val-5. The top and middle structures are from the "turn" simulation, and the bottom structure is from the "extd" simulation.

the peptide folded into a novel type I reverse turn which lasted several hundred picoseconds before unfolding again. These type I turns were novel in the sense that they were not observed in the experiments of Dyson et al. (1988b).

The results of several experimental studies (Montelione et al., 1984; Williamson et al., 1986; Dyson et al., 1988b) and our simulations show that reverse turns can form in short peptides in solution. However, Tobias et al. (1991) recently predicted that reverse turns are intrinsically unstable secondary structures in water. That is, in the absence of sequence-specific side chain and/or terminal group interactions, extended conformations are strongly favored over reverse turns in solution. The atomic-level detail of MD simulations allows us to identify the specific interactions which could enable small peptides to overcome the intrinsic instability and form stable reverse turns. Next, we describe some of the interactions which appeared to stabilize reverse turn conformations of Tyr-Pro-Gly-Asp-Val in our simulations. Our criterion for identifying important interactions was that the interacting groups were near contact separation and the interaction persisted for at least 50 ps.

The preparation of the peptide for the "turn" simulation resulted in the formation of a strong salt bridge with the charged NH_3^+- and CO_2^--terminal groups nearly at contact (see Figure 6, top). During the simulation, the salt bridge lasted only 250 ps, breaking up when the exiting chain and Asp side chain switched positions in the reverse turn. Since this contact salt bridge between the terminal groups never appeared again in any of the simulations, we are not certain that it was an important interaction; it could merely have been an artifact of using a minimum energy structure to start the simulation. The most important interaction stabilizing reverse turns during the "turn" simulation seemed to be a hydrogen bond formed between the backbone of Gly-3 and the charged C-terminal group (see Figure 6, middle). This hydrogen bond, which appears as a close approach (~3.5 Å) of the Gly N and one of the Val terminal carboxylate O atoms, formed soon after the breakup of the salt bridge between the terminal groups and lasted for most of the lifetime of the turn (see Figure 7a). The backbone–terminal group hydrogen bond was accompanied by a solvent-separated salt bridge (ion–water–ion) between the N-terminal group and the CO_2^- of the Asp side

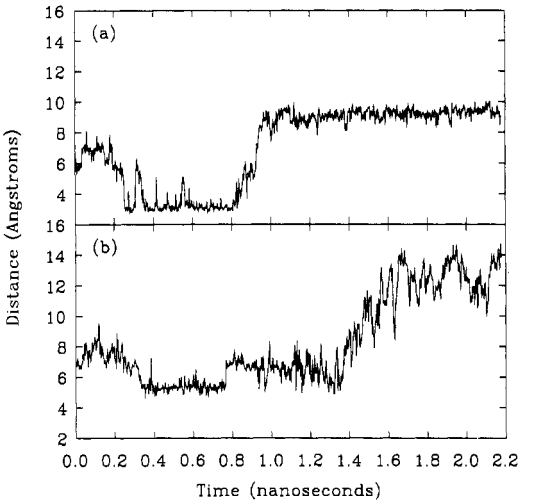

FIGURE 7: Time series from the "turn" simulation: (a) distance between the Gly amide N and one of the Val terminal carboxylate O atoms; (b) distance between the Tyr amide N and one of the Asp side chain O^δ atoms.

chain, which appeared to provide a slight additional stabilization. The latter interaction appears as an ~5-Å approach of the Tyr N and one of the Asp chain O^δ atoms (see Figure 7b). None of the interactions just described were present during the formation of the type II turn in the "extd" simulation. The absence of any strongly stabilizing interactions in that turn is a plausible explanation for its transient existence. This suggests that the formation of a stabilizing sequence-specific interaction is required to complete the folding transition in a short peptide. The novel type I reverse turn observed in the "extd" simulation was apparently stabilized by a hydrophobic interaction involving the pyrrolidine ring of the proline and the Val side chain (see Figure 4b). These groups were approximately at van der Waals contact (4–4.5 Å; see Figure 6, bottom), and the solvent was completely expelled from the region between the entering and exiting chains, forming a tiny "apolar core" in the type I turn. Hydrogen bonds between the backbone amide groups were mostly exposed to the solvent, were rapidly formed and broken (e.g.,

see Figure 2b), and did not appear to strongly stabilize any of the folded structures in either of the simulations. This is not too surprising since the free energy of forming amide hydrogen bonds in water is close to zero, and the free energy barrier to breaking them is small (Sneddon et al., 1989).

We can make a limited comparison of our simulation results with the experimental results of Dyson et al. (1988b). Unfortunately, we cannot compare results for the relative populations of folded and unfolded species because the folding/unfolding events were rare in our simulations and therefore do not constitute a statistically valid sample. However, we can compare our observations regarding specific interactions which stabilize reverse turns in Tyr-Pro-Gly-Asp-Val. The NMR data of Dyson et al., specifically the low temperature coefficient of the Asp amide proton resonance, suggested that a high population of folded structures had a 1–4 hydrogen bond (between the Tyr CO and Asp NH groups). Our results agree with the experimental results in the sense that a large number of 1–4 hydrogen-bonded turn conformations were sampled in our "turn" simulation. The experiments indicated that an (unspecified) electrostatic interaction involving the negatively charged carboxylate group of the Asp-4 side chain may have slightly stabilized the turn. In our "turn" simulation simulation, electrostatic interactions between the Asp-4 side chain and the positively charged N-terminus appeared to stabilize the type II turn. In contrast to our observations, the large temperature coefficient of the Gly amide proton resonance strongly indicated that a C-terminal group–Gly NH hydrogen bond was not present in solution during the experiments. Finally, the hydrophobic interaction we see stabilizing the type I turn is not discernible in the NMR data because the resonances of the aliphatic Pro and Val side chain protons overlap (H. J. Dyson, personal communication).

Overall, it is difficult to verify or dispute our results by comparing them to those of Dyson et al. Some of our observations are consistent with the experimental data, while others are not. The differences could arise because the two studies were carried out at different temperatures. Moreover, we have observed transient conformations which are not necessarily the same as the ensemble-averaged, equilibrium structures observed in the experiments. Finally, the discrepancies could be due to deficiencies in the peptide and/or solvent models used in our simulations. While it is fairly well established that the water and protein models used in this study work well independently (Jorgensen et al., 1983; Brooks et al., 1988), it has not yet been adequately demonstrated that when used together they give reasonable results for the structure and dynamics of small peptides in solution. Nonetheless, the simulations described in this paper show what can happen, namely, that folding initiation structures can form and unfold on the nanosecond time scale at room temperature, given a particular working model. The verification of the results of the present study, as well as the results of past (Wright et al., 1990; Tirado-Rives & Jorgensen, 1991) and future simulations of the formation and dissolution of secondary structures in small peptides in solution, awaits the development of appropriate faster time scale experimental techniques.

ACKNOWLEDGMENTS

The Corporate Computer Network of Cray Research, Inc., generously provided the computer time. We are grateful to Dr. Scott Sneddon and Dr. Erich Wimmer for helpful discussions.

REFERENCES

Allen, M. P., & Tildesley, D. J. (1989) *Computer Simulation of Liquids*, Oxford University Press, Oxford.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187.

Brooks, C. L., III, Karplus, M., & Pettitt, B. M. (1988) *Adv. Chem. Phys. 71*, 1.

Bundi, A., Andreatta, R. H., & Wuthrich, K. (1978) *Eur. J. Biochem. 91*, 201.

Chou, P. Y., & Fasman, G. D. (1977) *J. Mol. Biol. 115*, 135.

Creighton, T. E. (1983) *Proteins*, W. H. Freeman, New York.

Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988a) *J. Mol. Biol. 201*, 201.

Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988b) *J. Mol. Biol. 201*, 161.

Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W., & Klein, M. L. (1983) *J. Chem. Phys. 79*, 926.

Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631.

Lewis, P. N., Momany, F. A., & Scheraga, H. A. (1973) *Biochim. Biophys. Acta 303*, 211.

Montelione, G. T., Arnold, E., Meinwald, Y. C., Stimson, E. R., Denton, J. B., Huang, S.-G., Clardy, J., & Scheraga, H. A. (1984) *J. Am. Chem. Soc. 106*, 7946.

Roder, H., Elove, G. A., & Englander, S. W. (1988) *Nature 335*, 700.

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327.

Sneddon, S. F., Tobias, D. J., & Brooks, C. L., III (1989) *J. Mol. Biol. 209*, 817.

Sneddon, S. F. (1990) *MolX Series of Molecular Graphics Programs*, Department of Chemistry, Carnegie Mellon University.

Tirado-Rives, J., & Jorgensen, W. L. (1991) *Biochemistry 30*, 3864–3871.

Tobias, D. J., Sneddon, S. F., & Brooks, C. L., III (1991) *J. Mol. Biol. 216*, 783.

Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694.

Venkatachalam, C. M. (1968) *Biopolymers 6*, 1425.

Verlet, L. (1967) *Phys. Rev. 159*, 98.

Williamson, M. P., Hall, M. J., & Handa, B. K. (1986) *Eur. J. Biochem. 158*, 527.

Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167.

Wright, P. E., Dyson, H. J., Feher, V. A., Tennant, L. L., Waltho, J. P., Lerner, R. A., & Case, D. A. (1990) in *Frontiers of NMR in Molecular Biology* (Live, D., Armitage, I. M., & Patel, D., Eds.) Wiley-Liss, New York.