PROTEINS: Structure, Function, and Genetics 24:450–466 (1996)

Molecular Dynamics Simulations of N-Terminal Peptides From a Nucleotide Binding Protein

David van der Spoel,[1] Hans J. Vogel,[2] and Herman J.C. Berendsen[1]
[1]*Department of Biophysical Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands; and*
[2]*Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada T2N 1N4*

ABSTRACT **Molecular dynamics (MD) simulations of N-terminal peptides from lactate dehydrogenase (LDH) with increasing length and individual secondary structure elements were used to study their stability in relation to folding. Ten simulations of 1–2 ns of different peptides in water starting from the coordinates of the crystal structure were performed. The stability of the peptides was compared qualitatively by analyzing the root mean square deviation (RMSD) from the crystal structure, radius of gyration, secondary and tertiary structure, and solvent accessible surface area. In agreement with earlier MD studies, relatively short (< 15 amino acids) peptides containing individual secondary structure elements were generally found to be unstable; the hydrophobic α_1-helix of the nucleotide binding fold displayed a significantly higher stability, however. Our simulations further showed that the first βαβ supersecondary unit of the characteristic dinucleotide binding fold (Rossmann fold) of LDH is somewhat more stable than other units of similar length and that the α_2-helix, which unfolds by itself, is stabilized by binding to this unit. This finding suggests that the first βαβ unit could function as an N-terminal folding nucleus, upon which the remainder of the polypeptide chain can be assembled. Indeed, simulations with longer units (βαβα and βαβαββ) showed that all structural elements of these units are rather stable. The outcome of our studies is in line with suggestions that folding of the N-terminal portion of LDH in vivo can be a cotranslational process that takes place during the ribosomal peptide synthesis.**

Key words: lactate dehydrogenase, Rossmann fold, cotranslational folding

INTRODUCTION

The folding of a protein into its correct three-dimensional structure is a topic of continuing interest.[1] Following the original demonstration by Anfinsen[2] that a denatured protein can regain its original structure when the denaturing agent is removed, it is clear that the information required for the folding of the protein is retained in its amino acid sequence. Since this discovery it has been customary to study protein folding by following the renaturation of the denatured polypeptide, which comprises the entire amino acid sequence. These studies have led to the widely held view that so-called molten globules exist as obligatory intermediates along the in vitro folding pathway of many proteins.[3–6] Such intermediates have been characterized in detail primarily by nuclear magnetic resonance (NMR) hydrogen exchange measurements and have been found for various segments in the amino acid sequence.[7–9] This view is no longer generally accepted, however, since recent experiments have suggested that the partially folded forms of some proteins could be trapped unproductive forms rather than obligatory folding intermediates.[10,11] Other studies aimed at solving the protein folding problem have relied on the assumption that the overall folding of a protein is dictated by the intrinsic folding properties of short stretches of the amino acid sequence. Consequently, short linear peptides resembling parts of the amino acid sequence of various proteins have been produced synthetically and have been studied by high-resolution NMR and circular dichroism (CD) spectroscopy for their folding properties.[12–14] While this work has clearly established the transient formation of elements of secondary structure in such peptides, the role of these secondary structure elements in directing the course of folding remains to be determined.[13,14]

The question remains whether protein folding in vivo really occurs from a completely synthesized and unstructured random coil peptide. Several authors have argued that it would be possible in principle to initiate folding in the N-terminal part of a protein because the polypeptide chain emerges in an ordered manner from the ribosome during protein translation.[15–18] This process could give rise to an N-terminal folding nucleus, which can subsequently

Abbreviations: LDH, lactate dehydrogenase; MD, molecular dynamics; RMSD, root mean square deviation; CPU, central processing unit; CD, circular dichroism.

Received July 13, 1995; revision accepted October 18, 1995.
Address reprint requests to Herman J.C. Berendsen, Department of Biophysical Chemistry, University of Groningen, University of Groningen, Nijenborgh 4, 9747 Groningen, the Netherlands.

serve as a site of assembly for the remaining portions of the polypeptide chain. This phenomenon is referred to as cotranslational rather than posttranslational folding,[16] and a number of molecular biology studies with expressed truncated proteins[19–22] have indicated that the folding of certain classes of proteins may well proceed in this fashion. If such an N-terminal folding nucleus were to be present within a group of proteins with identical folds, it should have a recognizable feature near the N-terminal end of the protein. One group of proteins that has a well-characterized overall fold, which can be readily recognized on the basis of amino acid homology, comprises the nucleotide binding proteins.[23] In particular, the well-characterized Rossmann fold, also known as the "classical dinucleotide binding chain" fold, is found in all NAD^+, $NADP^+$, and FAD binding proteins.[23,24] In addition, a related, but structurally dissimilar "mononucleotide binding chain" fold is found in many proteins that bind nucleotides such as ATP and GTP.[23,25,26] That this class of proteins may be suitable for studying the formation of an N-terminal folding nucleus was suggested in a paper by Wierenga et al.[27] They commented that the nucleotide binding βαβ unit could be a possible nucleation site for protein folding: " . . . In the known structures these βαβ-structural entities do always occur near the N terminus . . . . This suggests that these βαβ-folds might function as a nucleation center for the folding of the complete domain . . ." The same region was also indicated as a possible folding unit by an analysis of the ribosomal translation rate.[28] Finally, it is known that N-terminal proteolytic fragments of lactate dehydrogenase (LDH) retain the capacity to bind dinucleotides, indicating that they are properly folded.[29] The βαβ unit, which is an integral part of the Rossmann fold, is a common supersecondary structural element in proteins.[30] In all, at least 40 proteins are presently known or predicted to have this fold, and considering the "age" of the fold, which was estimated to be over 3 billion years,[30] there will probably be more than these.

In this work we have used molecular dynamics (MD) calculations to obtain further information about the stability of peptides derived from the N-terminal portion of LDH, a protein with the dinucleotide binding fold. Over the last few years MD calculations, when used judiciously, have become a reliable means of studying the stability and flexibility of small proteins in water.[31,32] With continuing improvements in force fields and computing capacity, it has become possible to study processes such as protein denaturation.[31,33–36] In addition, simulations of peptides in aqueous solution can now be performed routinely, and such studies have been done to assess the stability of turns and helices in short linear peptides.[37–43] It is not possible to assess the thermodynamic stability of a peptide or protein in solution directly from MD simulations, due to limited sampling of unfolded as well as partially folded states. However, it is possible to investigate the *kinetic stability*, i.e., the resistance against unfolding and compare this for various peptides or proteins in different environments or at different temperatures. This has recently been done in a number of cases, in four for α-helical peptides[44–47] in another case for a β-hairpin.[48] Finally, it has also been noted that MD simulations can assist in an analysis of protein folding pathways.[49]

Tsou[16] has suggested that an N-terminal folding nucleus should have a structure that may not be identical but should resemble the final structure in the completed protein. Therefore we decided to use the folded parts as determined in the crystal structure of the protein as starting structures for our MD calculations. To simulate the effect of the ribosomal protein synthesis, we studied the N-terminal region and systematically increased the length of the simulated polypeptide chain. For comparison, a number of other units derived from the same amino acid sequence were also simulated. While this approach does not assess folding directly, the *kinetic stability* of the various peptides can be compared in this manner, as was done previously in MD studies of α-helices and β-turns.[37–41,44–46,48]

SIMULATION SETUP

Ten MD simulations of peptides from LDH in water, varying in length from 6 to 76 amino acid residues, were performed. One of these was a deprotonated peptide, in which the side chain of a lysine residue (Lys-22) was made neutral, to assess the influence of a salt bridge on the stability. In our simulations we ignored the first 20 residues of the protein, because in the complete protein they form a motif that is sticking out into the solvent. This loop is necessary for oligomerization, in which the loop stabilizes the tetramer[29]; therefore the loop is presumably less important for the structure of the monomeric folded protein. All but four of the peptides start at residue Asn-21. To distinguish the simulations we named them β_1, α_1, βα, β_2, $\beta\alpha\beta_1$, α_2, $\beta\alpha\beta_2$, βαβα, and βαβαββ; the names refer to the structural elements simulated. The simulation of the deprotonated peptide was named $\beta\alpha\beta^-$. An overview of the simulations is given in Figure 1. Simulations α_1, βα, α_2, $\beta\alpha\beta_1$, $\beta\alpha\beta^-$, and $\beta\alpha\beta_2$ were 2.0 ns, and the other simulations were 1.0 ns long. The sequence for the longest βαβαββ peptide along with some characteristics is given in Table I. A plot of the $\beta\alpha\beta_1$ and βαβαββ peptides is given in Figure 2.

The coordinates for the peptides were taken from the crystal structure of dogfish, apo-LDH, which was determined by Abad-Zapatero et al.[50] The structure was retrieved from the Brookhaven pdb database, entry 6LDH. They were solvated in a rectangular box of SPC water[51] using a shell of at least 0.8 nm on

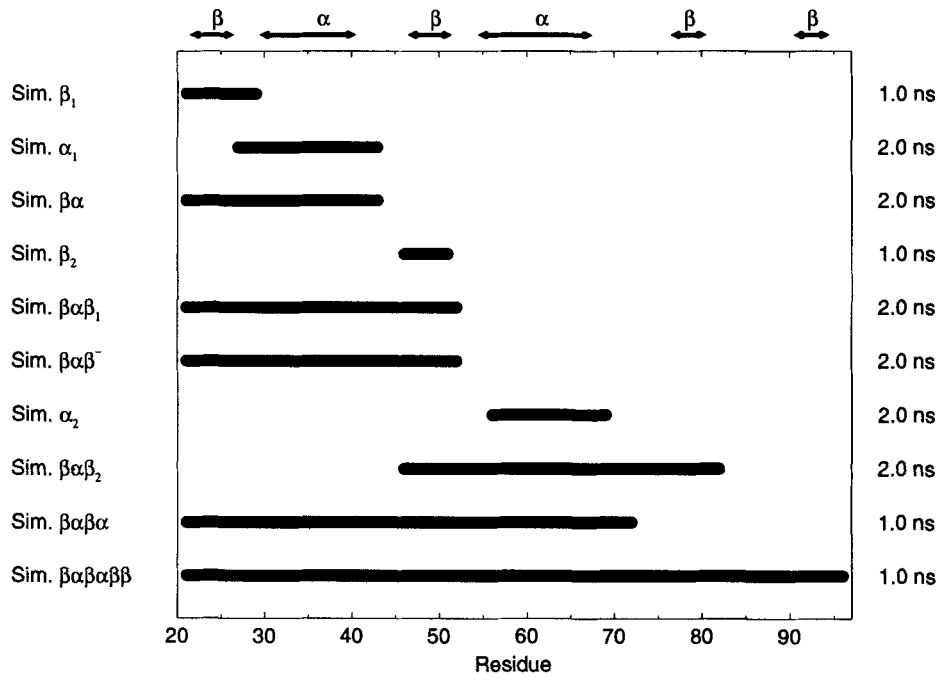

Fig. 1. Overview of the peptide names and sequences used.

every side of the peptide. Although the use of counterions would make the total system electrostatically neutral, we feel that this would not improve the results from the simulation without using a proper long-range electrostatic method such as Ewald summation[52,53] or particle-particle, particle-mesh (PPPM).[54,55] Although these methods are well established, there are some serious drawbacks to using them: 1) a crystalline environment is imposed on a system that is meant to be a peptide in infinite dilution; 2) a correct lattice sum method requires the presence of counterions, which, when not sampled properly, will induce artifacts; 3) the central processing unit (CPU) time is much longer due to different scaling, $O(N^2 \log N)$ for Ewald vs. $O(N)$ for simple cut-offs, although the required CPU time for the PPPM method is comparable to a simple cut-off.[5] Reaction field methods assume a dielectric continuum outside a specified cut-off range around a given particle,[56] which implies that the whole peptide should be within cut-off distance; this is impractical for peptides with a typical length of 2 nm, twice the cut-off distance. A further complication in using, e.g., the PPPM method is, to our knowledge, that there is no implementation of this method on message passing parallel computers, which we have used for our simulations. Some promising methods to deal with long-range electrostatic interactions are currently under development[57,58]; these will be implemented on our parallel computers in the near future.

All simulations were performed using periodic boundary conditions. The solvated peptides were subsequently energy minimized with the steepest descent method for 100 steps. Then the system was run with position restraining on the peptide for 50 ps to allow for further relaxation of the solvent molecules. During these restrained MD runs the temperature was controlled using weak coupling to a bath of constant temperature (T_0 = 300 K, coupling time τ_T = 0.1 ps), and the pressure was controlled using weak coupling to a bath of constant pressure (P_0 = 1 bar, coupling time τ_P = 0.5 ps).[59] The production runs were done with the same pressure- and temperature-coupling constants as the restrained runs. Protein and solvent were coupled to the temperature bath separately in restrained as well as free MD. The center of mass motion of the entire simulation system was removed every step. The Gromos-87 forcefield[60] was used with modifications as suggested by van Buuren et al.[61] The time step was 2.0 fs, SHAKE[62] was used for all covalent bonds, and the cut-off for nonbonded interactions was set at 1.0 nm. Neighborlists were used and updated every 20 fs.

To each peptide we added a $CH_3C=O$ group at the N-terminus and an NH_2 group at the C-terminus; these neutral groups were used to avoid disturbing the peptides by introducing charges on their N- and C-terminal ends. For comparison with future NMR and CD experiments, residue Cys35 was mutated to Ser to eliminate the possibility of disulphide bridge formation in the test tube; we think this substitution is unlikely to have a great impact on the structure of the peptides. The number of atoms and the

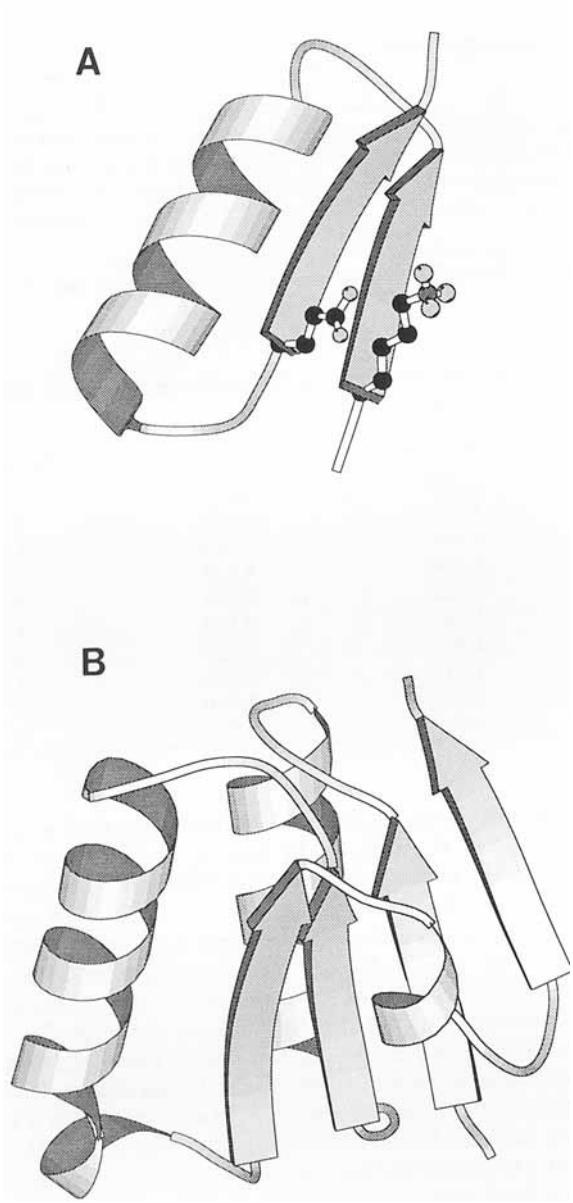

Fig. 2. Plot of the βαβ unit (**A**) with Lys-22 and Glu-47 in all atom model and βαβαββ unit (**B**). Plots were made using Molscript.[84]

number of water molecules for each peptide simulation are given in Table II. The simulations were carried out using the GROMACS software package[63] on custom-built parallel computers[64,65] with Intel *i*860 CPUs and on a Silicon Graphics Power Challenge. The total computer time for all simulations was approximately 2,000 h.

RESULTS

When analyzing long MD simulations of peptides or proteins it is useful to define properties that give a relative estimate of the stability. We have analyzed the root mean square deviation (RMSD) from the crystal structure, the radius of gyration, and the solvent accessible hydrophobic surface area. These values can be plotted as a function of time; in a stable protein these values fluctuate around an equilibrium value; in a changing conformation there will be a drift in one or more of these observables. Since we do not expect all our peptides to be stable in the simulation it is also interesting to look at the dynamic processes taking place during the simulation. To illustrate this we have also looked at secondary structure analysis, distance matrices, and salt bridges involving Lys-22. These items will be addressed in subsequent sections.

RMSD From the Crystal Structure

The RMSD of the backbone atoms (N,Cα,C) with respect to the crystal structure was calculated by least-squares fitting of the backbone atom positions of each peptide to the crystal structure and subsequently calculating the RMSD (Eq. 1).

$$\mathrm{RMSD}(t) = \left[\frac{1}{N_S}\sum_{i\in S}(\mathbf{r}_i(t) - \mathbf{r}_i(0))^2\right]^{\frac{1}{2}} \qquad (1)$$

where $\mathbf{r}_i(t)$ is the position of atom i at time t and S is a given subset of all atoms with size N_S. In Figure 3 the RMSD averaged over all residues is plotted as a function of time for all the simulations. The RMSDs of the β-strands in Sim. β_1 and β_2 are very high; these peptides deviate very much from the crystal structure. β_2 deviates gradually, whereas β_1 is more fluctuating. It should be mentioned that an RMSD of 0.7 nm for a 9 residue peptide (β_1) means that it is not very meaningful to fit the structures on top of each other. After an initial rise of the RMSD to 0.3 nm, the isolated α_1-helix is stable for 800 ps at an RMSD of 0.2 nm (from 350 until 1,150 ps). After this point in the simulation the RMSD gradually rises to more than 0.4 nm due to unfolding of the C-terminal part of the α-helix. The α_2 peptide deviates up to 0.4 nm from the crystal structure within 800 ps and remains at that level for the rest of the 2.0 ns simulation. In the βα simulation the RMSD goes up after 600 ps due to unfolding of the helix starting from the N-terminal side, i.e., from the flexible loop. Around 1.0 ns the βα peptide reaches an RMSD of 0.5 nm, after which point the RMSD decreases to 0.3 nm at the end of the simulation. In the runs of the $\beta\alpha\beta_1$, the $\beta\alpha\beta^-$, and the $\beta\alpha\beta_2$ peptides, the RMS deviation slowly increases to 0.4 nm after 1,000 ps, but after 1,250 ps it decreases again until it stabilizes at 0.3 nm. In all three βαβ units the RMSD is then virtually constant for over 700 ps until the end of the 2.0 ns simulation. In the two longest peptides the RMSD slowly rises to 0.2 (βαβαββ) resp 0.3 (βαβα) nm. In both simulations there is still a small drift at the end of the simulation (1,000 ps), indicating that the peptides are not fully equilibrated.

TABLE I. Sequence Characteristics*

Number	2	3	4	5	6	7	8	9
	12345678901234567890123456789012345678901234567890123456789012345678901234567890123456							
Amino Acid	NKITVVGVGAVGMACAISILMKDLADEVALVDVMEDKLKGEMMDLQHGSLFLHTAKIVSGKDYSVSAGSKLVVITA							
Sec. Struct.	S EEEEE SHHHHHHHHHHHHHHT SEEEEE S HHHHHHHHHHHHHHH GGG SEEEEESSGGGT S SEEEE							
Hydrophobic	oooo o oo ooooo ooo oo oooo oo o oo o ooo oo oo o o ooooo							
Charge	+ +− −− − −−+ + − − + + + +− +							

*H = Helix, G = 3 Helix or 3_{10} Helix, T = Hydrogen Bonded Turn, S = Bend, E = Extended strand participating in Beta Sheet, Space = No defined secondary structure.

TABLE II. Overview of the Peptides Used in the Simulations, the Number of Water Molecules, and the Net Charge on Each Peptide

	Residues	No. of residues	No. of waters	No. of atoms	Charge
β_1	Asn-21–Gly-29	9	1,346	4,120	+1
α_1	Gly-27–Asp-43	17	1,612	4,974	0
$\beta\alpha$	Asn-21–Asp-43	23	1,886	5,854	+1
β_2	Asp-46–Val-51	6	1,394	4,238	−2
$\beta\alpha\beta_1$	Asn-21–Asp-52	32	1,929	6,057	−2
$\beta\alpha\beta^-$	Asn-21–Asp-52	32	1,947	6,110	−3
α_2	Asp-56–Ser-69	14	1,627	5,020	0
$\beta\alpha\beta_2$	Asp-46–Asp-82	37	1,919	6,105	−2
$\beta\alpha\beta\alpha$	Asn-21–Leu-72	52	2,340	7,480	−3
$\beta\alpha\beta\alpha\beta\beta$	Asn-21–Ala-96	76	2,510	8,202	0

The RMSD was also averaged over the backbone atoms per residue over the last 1,000 ps of each simulation (i.e., over the entire trajectory for Sims. β_1, β_2, $\beta\alpha\beta\alpha$, and $\beta\alpha\beta\alpha\beta\beta$) (Fig. 4). Here we can compare how parts of the secondary structure behave in different peptides. In most simulations that contain the α_1-helix ($\beta\alpha$, $\beta\alpha\beta_1$, $\beta\alpha\beta^-$, and, to a lesser extent $\beta\alpha\beta\alpha$) the C-terminal end of the α_1-helix is flexible. The α_1-helix is flexible at the ends only, whereas the α_2-helix is flexible over its total length, especially near Gly-60. In all simulations containing residues 70–72 (Leu-Phe-Leu) ($\beta\alpha\beta_2$, $\beta\alpha\beta\alpha$, and $\beta\alpha\beta\alpha\beta\beta$), this section of the polypeptide chain is very flexible because three hydrophobic side chains are solvent exposed and are trying to move away from the solvent. There is no suitable place to go for these side chains, however, so they remain solvated. Another noticeable feature is the Gly-rich loop (residues 27–32), which shows a peak in β_1, $\beta\alpha$, and $\beta\alpha\beta^-$ although it varies in height.

Radius of Gyration

For each peptide the radius of gyration R_g was calculated according to Eq. 2

$$R_g = \left(\frac{\sum_i \mathbf{r}_i^2 m_i}{\sum_i m_i} \right)^{\frac{1}{2}} \qquad (2)$$

where m_i is the mass of atom i and $\mathbf{r}_i$ the position of atom i with respect to the center of mass of the peptide. The radius of gyration is a rough measure for the compactness of a structure. In all but the longest peptides ($\beta\alpha\beta_2$, $\beta\alpha\beta\alpha$, and $\beta\alpha\beta\alpha\beta\beta$) we find some change in R_g (Fig. 5). In Sim. β_1 the decrease at about 200 ps is due to a collapse of the peptide where a β-turn-like structure is formed, which later unfolds again. β_2 collapses somewhat and remains stable after 700 ps. In α_1 R_g decreases gradually because the loops at both ends of the α_1-helix sometimes curl inward. By contrast, the large fluctuations in R_g in Sim. α_2 are due to unfolding and refolding of parts of the α_2-helix. In the first 300 ps of the $\beta\alpha$ simulation we see R_g decreasing with 0.1 nm because the C-terminal loop at the end of the helix collapses inward, making a short-lived hydrogen bond between the side chain NH_2 of Asn-21 and the carbonyl oxygen of residue Asp-43. Movements in the Gly-rich loop (residues 27–32) also contribute to the change in R_g. After 1.0 ns R_g decreases slowly to 0.75 nm. In Sim. $\beta\alpha\beta^-$, R_g decreases somewhat after 800 ps due to movements in the second loop region; in $\beta\alpha\beta_1$, in contrast, R_g rises slightly after 800 ps although the final value (0.9 nm) is close to the initial one (0.87 nm).

Secondary Structure Analysis

A secondary structure analysis was performed with the DSSP program, which determines the existence of hydrogen bonds as criteria for the presence of secondary structure.[66] In Figure 6 we have plotted the secondary structure elements of each residue in the different peptides as a function of time. Colors were used to distinguish between secondary structure types, the most important colors being red

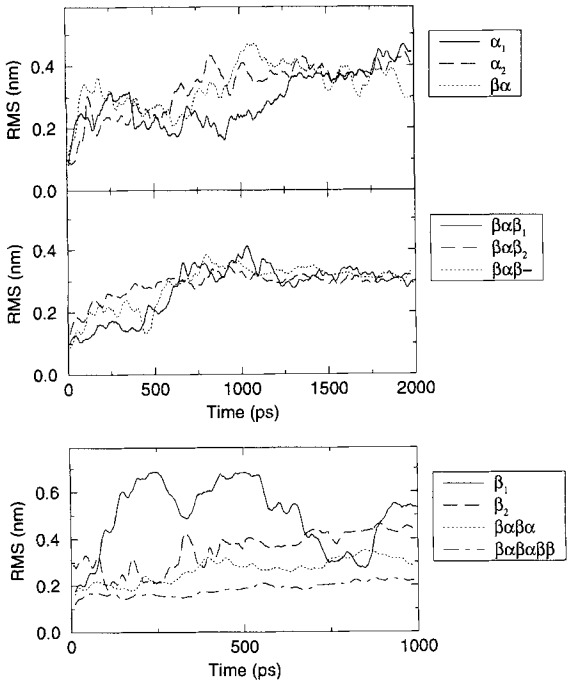

Fig. 3. Root mean square deviation (RMSD) of all backbone atoms from the X-ray structure. A running average over 25 ps is given to improve clarity.

for β-sheet and blue for α-helix. From Figure 6 it can be seen that the isolated β_1 (residues 21–27) and loop (residues 28–29) form a hydrogen bonded turn sometimes centered around Gly-27 and a bend most of the time. In β_2, a hydrogen bonded turn appears to be the most stable form in the simulation; during formation of the turn structure a transient 3_{10} Helix is found, at 320 ps and around 375 ps. Neither of the two β-strands remains even close to their starting structure. The α_1-helix is rather stable; compared with the crystal structure only two residues on each side come off initially. Three residues in the middle of the α_1-helix also lose their α-helicity temporarily, but they rapidly regain their original structure. After 1,500 ps, however, four residues on the C-terminal end lose their helicity although they remain hydrogen bonded in a turn.

As can be seen in the α_1 simulation (around 125 ps), the helix secondary structure type (blue) can be rapidly changed into the turn structure type (yellow) and back, indicating that these secondary structure types can interconvert rapidly within the definitions of the DSSP program. In the βα unit the hydrophobic β_1-strand and the Gly-rich loop apparently are "pulling" on the α_1-helix, thereby unfolding it from the N-terminal side; at 1,000 ps α_1 is completely unfolded in this peptide. The α_1-helix rapidly reforms, however, starting from the N-terminal end at Met-33, which directly follows a Gly residue: a nine residue α-helix is formed in 200 ps. Simulations $\beta\alpha\beta_1$ and $\beta\alpha\beta^-$ can also be compared in detail using this analysis. Both units are altogether stable; the number of hydrogen bonds in the β-sheet varies but is never less than two. In the α-helical region of these peptides there is a difference between the two units; the α_1-helix loses three residues on the C-terminal end in the $\beta\alpha\beta_1$ simulation, whereas it loses seven α-helical residues in the C-terminal end in the $\beta\alpha\beta^-$ simulation. The loss of helicity is due to rearrangements involving the hydrophobic residues in the second loop region and at the end of the α_1-helix. After 2.0 ns eight helical residues remain in simulation $\beta\alpha\beta_1$, four in $\beta\alpha\beta^-$, and a stable β-sheet in both. Considerable structural changes occur in the α_2 peptide simulation, with α-helical residues transforming into π-helix or a hydrogen bonded turn, initially at the C-terminal end only, but after 600 ps the α_2-helix unfolds from the N-terminal side and for several periods of up to 100 ps (1,150–1,250 ps), all α-helix is gone. A central Gly residue (Gly-60) is probably the reason for breaking. During the whole simulation we see α-helical parts in either the N-terminus or the C-terminus of the peptide but rarely simultaneously (except around 1,700 ps). In the $\beta\alpha\beta_2$ peptide we see that the α_2-helix unfolds from the C-terminus. All α-helical residues after Gly-60 are gone after 850 ps. By contrast, the β-sheet is highly stable; it restricts the conformational space of the residues in between. It is also interesting to note that considerable non-native secondary structure is formed after the C-terminal part of the α_2-helix has unrolled. When the α_2-helix is connected to the $\beta\alpha\beta_1$ peptide to form the βαβα peptide, the α_2-helix is more stable than in isolated form in solution (Fig. 6). In the βαβα peptide the most flexible part (apart form the three C-terminal residues) is the Gly-rich loop (residues 27–32), which unfolds the α_1-helix from the N-terminal side, similar to what was observed in the βα simulation. As in the α_2 simulation the α_2-helix breaks at Gly-60, but at either side of Gly-60 the α_2-helix is rather stable. The longest βαβαββ peptide is stable during the whole simulation; only the α_2-helix (which breaks at Gly-60 again) and the three solvent-exposed hydrophobic residues (70–72) change their secondary structure, although almost always ten residues remain in α-helical conformation in the α_2-helix. In both βαβα and βαβαββ, the β-sheet is stable; there are almost always two or more hydrogen bonds connecting the individual β-strands, but nevertheless in the βαβαββ peptide the first and second β-strands are even more stable than in the βαβα, due to the addition of the two extra β-strands on either side of the original β-sheet (β-strands are in order $\beta_4\beta_1\beta_2\beta_3$). Although this increased fixation of β-strands also gives rise to a more stable α_1-helix, the secondary structure pattern of the α_2-helix in the βαβαββ peptide is similar to that in the βαβα peptide.

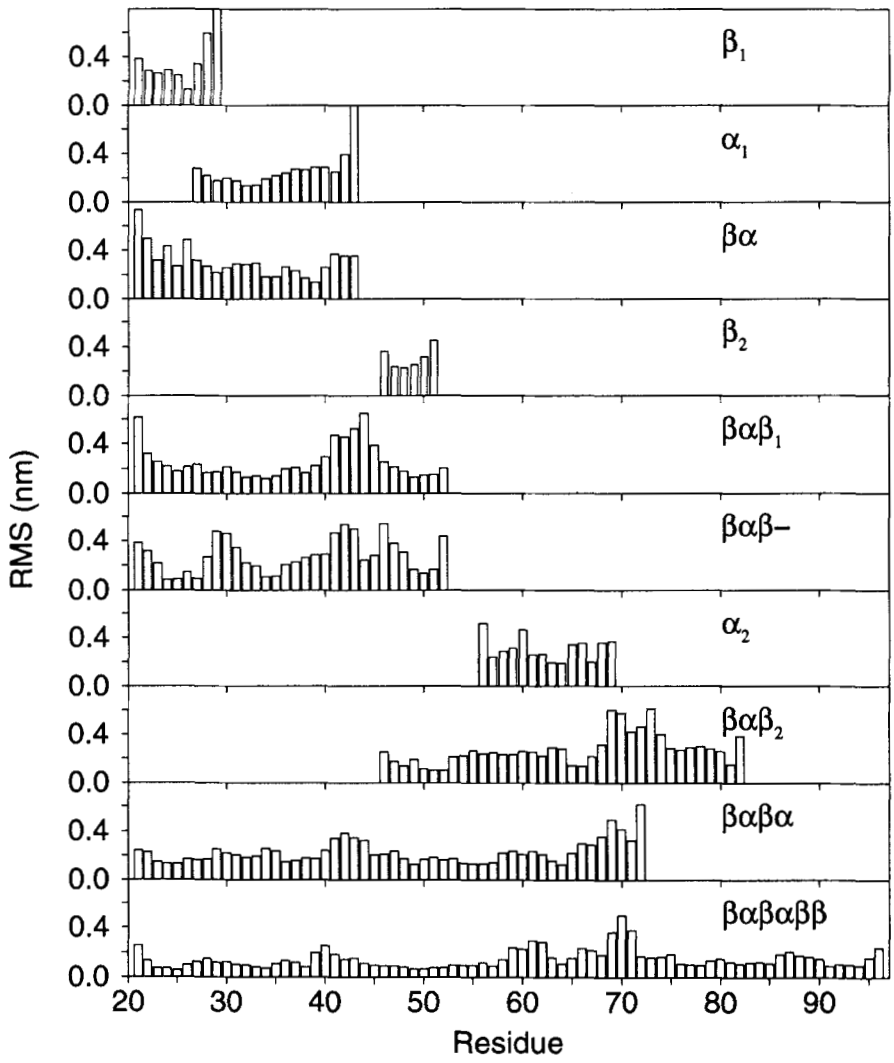

Fig. 4. Root mean square deviation (RMSD) from the X-ray structure per residue, averaged over the last 1,000 ps of each simulation.

Distance Matrices

Although the secondary structure analysis is very important for understanding the simulation, it does not tell all about the tertiary structure of the protein. We have therefore calculated distance matrices, in which a matrix of residue-residue distances is computed as a function of time. We defined the distance between two residues A_i and A_j as the smallest distance between any pair of atoms ($i \in A_i$, $j \in A_j$). The distance matrix is thus symmetric by definition. We have visualized this information using gray scales to represent distances. In Figure 7 the distance matrices for five different time frames in simulations $\beta\alpha\beta_1$, $\beta\alpha\beta^-$, and $\beta\alpha\beta_2$ are plotted. The distances have been discretized in 13 levels from 0 (black) to 1.2 nm (white), as is depicted in the figure legend. All distances greater than 1.2 nm are plotted in white. Although the figures appear to be very similar, a closer look reveals detailed information. The α_1-helix can be seen in the center of the figures, where a subdiagonal (and a superdiagonal) line close to the diagonal shows that residues at position (n+4) in the sequence are close to residues at position (n), where (n) runs from residues 29 to 38. At 800 ps in both simulations part of the α_1-helix has unfolded at the C-terminal end, but more so in the $\beta\alpha\beta^-$ simulation than in the $\beta\alpha\beta_1$ simulation. Parallel β-strands show up as sub (super) diagonal lines further removed than four residues from the diagonal. It can also be seen that residues 21 through 27 are close to residues 47 through 52, which is exactly the β-sheet region in the peptide. The β-sheet remains intact during the entire simulation of these peptides. Finally, it can also be seen from these distance matrices that unfolding of the C-terminal part of the α_1-helix coincides with closer contact of al-

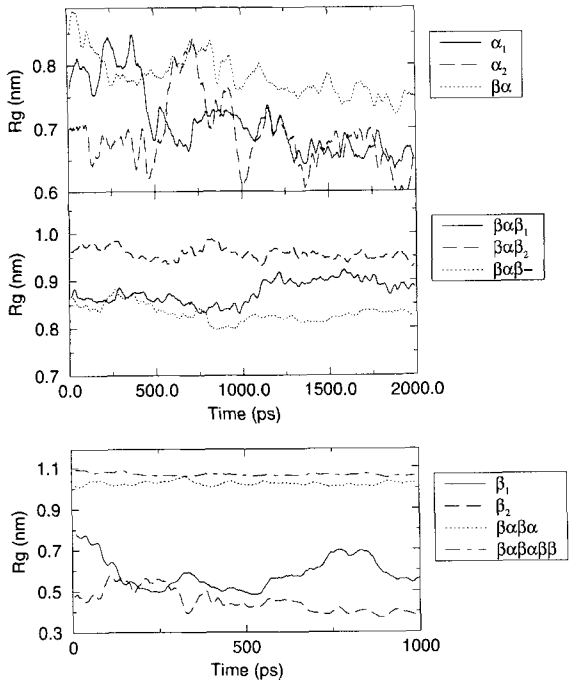

Fig. 5. Radius of gyration (Rg) during the MD simulations. A running average over 25 ps is given to improve clarity.

most all residues in the loop region that connects the α-helix to the second β-strand. The $\beta\alpha\beta_2$ unit behaves rather like the other βαβ unit, the β-sheet is stable, and the end of the α_2-helix is somewhat disordered after 1,000 ps. Also at t = 1,000 ps we see events taking place around residue Gly-60: some local close contacts between residues 59, 60, and 61 have formed.

In Figure 8 we have plotted the distance matrices of the larger βαβα and βαβαββ peptides at 0, 500, and 1,000 ps. As in the smaller peptide, the secondary structure elements such as the two α-helices are easily recognized in the distance matrix. In Sim. βαβαββ around residue 85, the 3_{10}-helix is visible as a dark region and it is also possible to deduce the order in which the four β-strands comprise the β-sheet from this plot ($\beta_4\beta_1\beta_2\beta_3$). From the distance matrices at 500 and 1,000 ps, it can be seen that the α-helices have moved somewhat closer, giving rise to more α-α contacts (roughly residues 30–40 and 55–65) in both βαβα and βαβαββ and closer (n, n+4) contacts at 1,000 ps in βαβαββ only. The most important conclusion from these plots is, however, that the tertiary structure from the crystal structure is still present at the end of each of the simulations that have at least a βαβ unit.

Solvent Accessible Surface Area

To examine the effect of the solvent on hydrophobic residues, the solvent accessible surface area was calculated with the DSSP program.[66] The program calculates the surface area and plots it per residue. We have summed up the total area for hydrophobic side chains and plotted that as a function of time in Figure 9. In some of the peptides we see that the total hydrophobic surface area is reduced over the course of the simulation (β_2, βα, $\beta\alpha\beta^-$, $\beta\alpha\beta_2$, and βαβα). In $\beta\alpha\beta_1$ we note a clear decrease in the solvent accessible hydrophobic area around 900 ps, followed by a sharp increase. Eventually the area is similar to the value at t = 0. Thus, in some peptides there is a clear hydrophobic effect that seems to drive the conformational changes that occur in our simulations. When we average the solvent accessible hydrophobic surface area per residue over the simulation (Fig. 10), we see that some residues are protected from solvent almost completely in the larger peptides, whereas they are exposed in the smaller ones. This applies especially to the sequence Ile-23 through Gly-28, comprising the first β-strand, and Val-48 through Val-51, which largely makes up the second β-strand; these residues are exposed to solvent in the $\beta\alpha\beta_1$ unit, while they are protected in the larger βαβαββ peptide.

Salt Bridges

It is interesting that oppositely charged residues can be found at the start of the first (Lys-22) and second strands (Asp-46, Glu-47) of the parallel β-sheet in $\beta\alpha\beta_1$. In the crystal structure of LDH, Glu-47 forms a salt bridge with Lys-22, while the side chain of Asp-46 ($O\delta_1$) forms a hydrogen bond with the backbone NH of Glu-47. From the beginning of simulation $\beta\alpha\beta_1$ we see salt bridges occurring intermittently between Lys-22 and Asp-46 or Glu-47 (Fig. 11). Although the Lys-22 to Glu-47 salt bridge is present most of the time before 1.0 ns, Glu-47 is replaced by Asp-46 for some time around 300 ps; after 1.0 ns Asp-46 and Glu-47 both form salt bridges with Lys-22, but neither is very persistent. In simulation $\beta\alpha\beta^-$, where we deprotonated the Lys-22 side chain from NH_3^+ to an NH_2 group, hydrogen bonds between Lys-22 and Glu-47 also occur, but less frequently. A hydrogen bond between Asp-46 and Lys-22 is never seen (Fig. 11). The large peak in the distance between Lys-22 and Glu-47 of more than 1 nm coincides with a temporary reduction in the number of hydrogen bonds in the β-sheet. From this figure it can also be seen that the motion of the negative side chains is coupled in this simulation; they move as a couple all the time, whereas in simulation $\beta\alpha\beta_1$ there is no such coupling. This may be because Glu-47 is very much immobilized by the salt bridge whereas the Asp-46 side chain is still free to move. In the βαβα and βαβαββ simulations the salt bridge between Lys-22 and Glu-47 is present almost continuously (Fig. 11), implying that the residues are increasingly fixed in their position due to surrounding residues. In Sim. βαβα the Asp-46 to

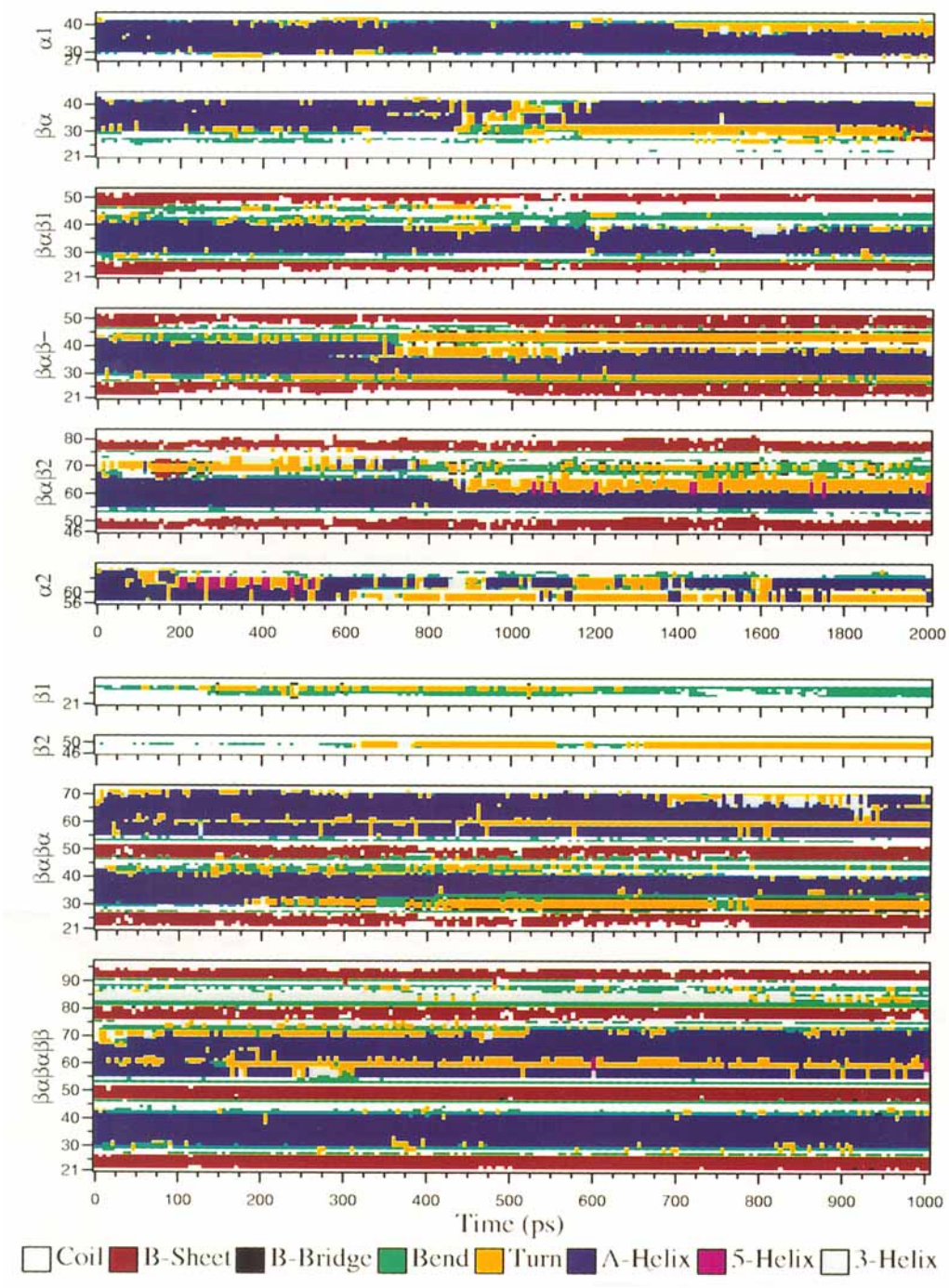

Fig. 6. Secondary structure as a function of time for all simulations.

Lys-22 salt bridge is sometimes present simultaneously with the Glu-47 to Lys-22 salt bridge.

DISCUSSION

It has often been suggested that protein folding is a spontaneous process that starts during protein synthesis on the ribosome.[15] There has been some controversy on the topic, however, because currently not a lot of direct experimental evidence supports this notion. Nevertheless Tsou[16] has strongly argued that protein folding during biosynthesis is at least partially cotranslational, while the extent of folding may vary for different proteins. It has also been found that the fully unfolded form of a complete polypeptide does not exist within the living cell.[18] In recent years some evidence for cotransla-

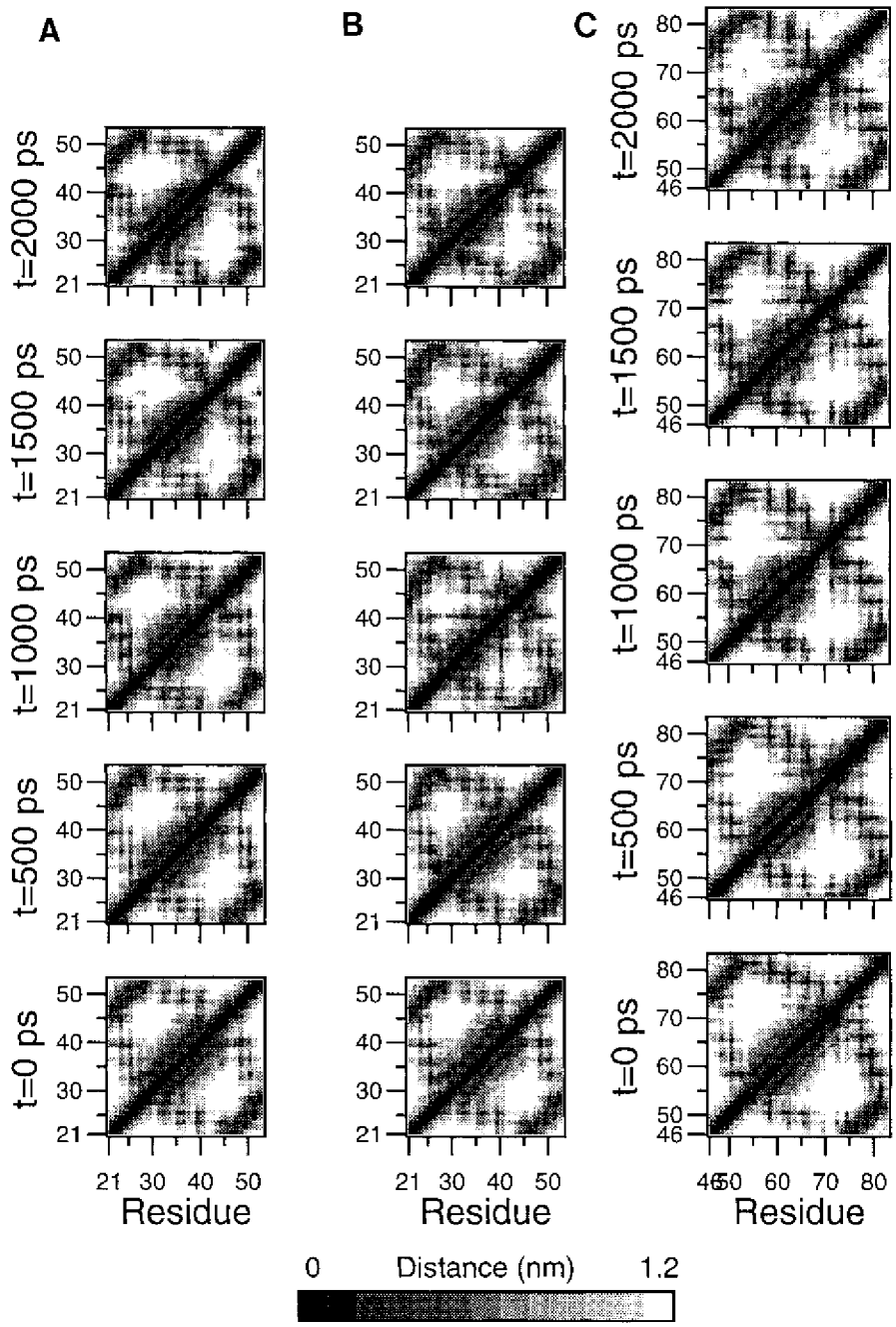

Fig. 7. Residue–residue distance matrices from five time frames of simulations $\beta\alpha\beta_1$ (**A**) and $\beta\alpha\beta^-$ (**B**), and $\beta\alpha\beta_2$ (**C**). The distance between residues A_i and A_j is the smallest distance between any pair (i, j) of atoms ($i \in A_i$, $j \in A_j$).

tional folding has been provided; for example, Fedorov et al.[17] demonstrated the correct folding of the N-terminal end of the β-subunit of tryptophane synthetase during its synthesis on the ribosome. In addition, it has been demonstrated that the ribosome itself can induce folding.[67,68] Proteins with repeating αβ structures with parts of their C-terminal ends truncated, such as ras-protein,[19] phosphoglycerate kinase,[20] glyceraldehyde-3-phosphate dehydrogenase,[22] or phosphoribosyl anthranilate isomerase,[21] all fold into a stable native-like structure. Interestingly, in the case of the latter enzyme, the carboxy-terminal fragment was unstructured by itself, but rapidly acquired its native structure upon binding to the N-terminal domain. Experimental work has shown that N-terminal fragments of por-

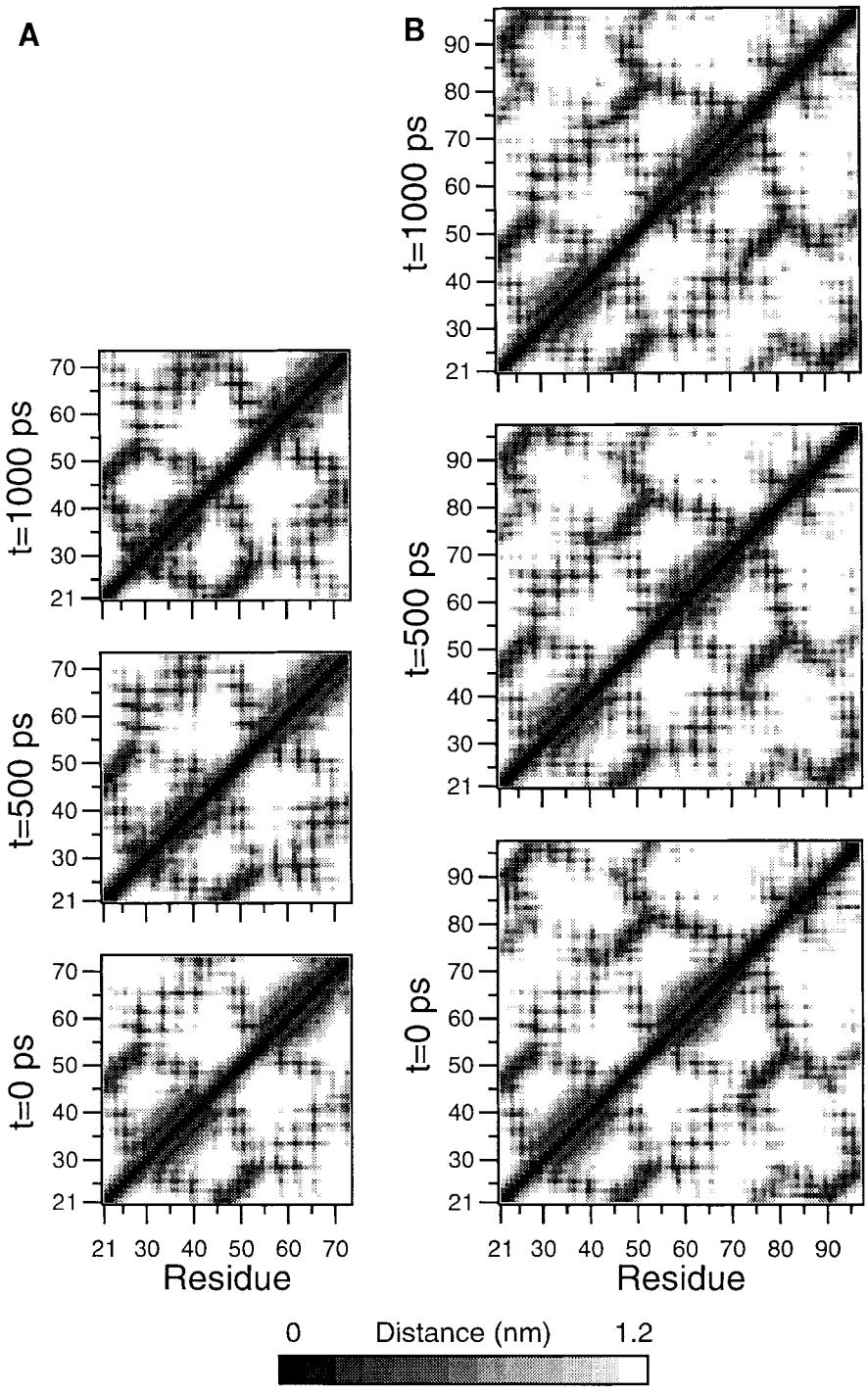

Fig. 8. Residue–residue distance matrices of three time frames from simulation βαβα (**A**) and βαβαββ (**B**). The distance between residues A_i and A_j is the smallest distance between any pair (i, j) of atoms ($i \in A_i$, $j \in A_j$).

cine muscle LDH can bind dinucleotides, indicating that the Rossmann fold is present in these fragments.[29,69] Although C-terminal fragments of LDH were also found to possess the native structure grossly, the stability of these fragments is drastically reduced compared with the native protein.[1] Taken together, these observations suggest that the N-terminal region of these proteins might function as a folding nucleus upon which the remainder of the structure is assembled.

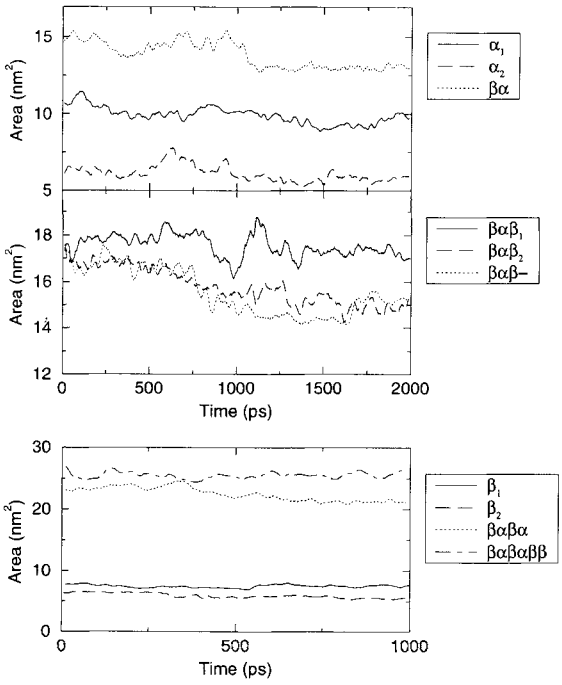

Fig. 9. Solvent accessible hydrophobic surface area during the simulations. A running average over 25 ps is given to improve clarity.

Recent MD simulation work on amphiphilic α-helical peptides in aqueous solution[38–41,43,44] has shown that most isolated α-helices have a tendency to unfold in less than 200 ps under these conditions. In another case, of myoglobin α-helices, this was less clear; only one of these (helix F) unfolds within 1 ns at 298 K; of the other α-helices, helices A, G, and H seem to be most stable.[45] In yet another simulation study of the Helix I/Loop I fragment of Barnase, which is assumed to be an initiation site for folding of the enzyme, the authors found the α-helical part to be rather stable, mainly due to hydrophobic interactions[47]; from their simulation data the authors build a detailed folding pathway for the α-helix. Nevertheless, most peptide α-helices are unstable in aqueous solution; water penetration and insertion into the peptide backbone appear to be the major mechanisms that give rise to helix instability.[37,43] We have found that the stability of the isolated α_1-helix in water is significantly higher than most α-helices in MD simulations; the hydrophobic surface of the α-helical peptide apparently shields the backbone from water insertion.

By contrast, the amphiphilic α_2-helix of LDH simulated in water is relatively unstable and behaves quite like other α-helical peptides in MD simulations. In Figure 6 it can be seen that both ends of the peptide are sometimes α-helical, most notably the C-terminal end. In the center of this α-helix is a Gly residue that facilitates water insertion in the α-helix backbone. The stability of the α_1-helix is likely to be real, because of the hydrophobic character of the peptide. Although simulation α_1 is not equilibrated after 2.0 ns, the same α-helix in simulation βα fully unfolds and subsequently refolds again, which strongly suggests that the α-helical structure is most favorable for this sequence. Dill et al.[70] have suggested that such structures could form and be stabilized through hydrophobic contacts between residues in close proximity in a polypeptide chain. Indeed, it seems that upon refolding the hydrophobic surface area of the βα peptide is reduced (Fig. 9). Recent MD simulations of a highly hydrophobic membrane spanning peptide in water have revealed that such α-helical peptides have a stability comparable to that of the LDH α_1-helix[46]; the A helix of myoglobin, which is highly stable in a simulation, also has a very hydrophobic sequence in the middle of the peptide spanning residues W7-A15.[45]

MD simulations of short peptides in water have demonstrated that turn-like structures can rapidly form, disappear, and reform.[41,71] This notion has been confirmed by NMR studies of short turn-forming peptides.[13,72,73] Indeed, our simulations of the isolated β_1, β_2, and also α_2-helix show similar behavior. In the simulations of longer peptides, the most flexible regions are generally the turn-forming regions and the ends of the peptides. All proteins with a nucleotide binding fold contain a Gly-rich region between the β_1- and α_1-helix that forms a characteristic turn in the final structure.[24,74] In the larger structures, particularly in the βαβαββ peptide, the flexibility of this region is substantially reduced. We believe that the glycines in this region may form a flexible joint that facilitates the overall folding of the $\beta\alpha\beta_1$ unit. Folding studies of turn-forming peptides are in agreement with this, as they show that a central Gly residue allows the two ends of the peptide to approach each other faster.[75] It is interesting to note that the first 45 residues of adenylate kinase, which contain its nucleotide binding fold and its Gly-rich loop, can fold into a structure of two helices and three β-strands, which resembles the folding in the intact protein.[76] Like the intact protein this peptide can bind MgATP, which stabilizes the fold.[77] In our simulation of the βα unit the N-terminal strand is connected to the helix by the Gly-rich loop. Due to this loop we see much more flexibility in the α_1-helix than in the isolated α_1-helix simulation.

Most MD simulations in water performed to date, have focused on well-characterized small globular proteins or isolated peptides with elements of secondary structure. Here we have simulated isolated elements of supersecondary structure derived from LDH. We note that in our simulations the secondary and tertiary structure of the $\beta\alpha\beta_1$ and $\beta\alpha\beta_2$ units is quite well preserved. In general, the structure in these characteristic units of the Rossmann fold

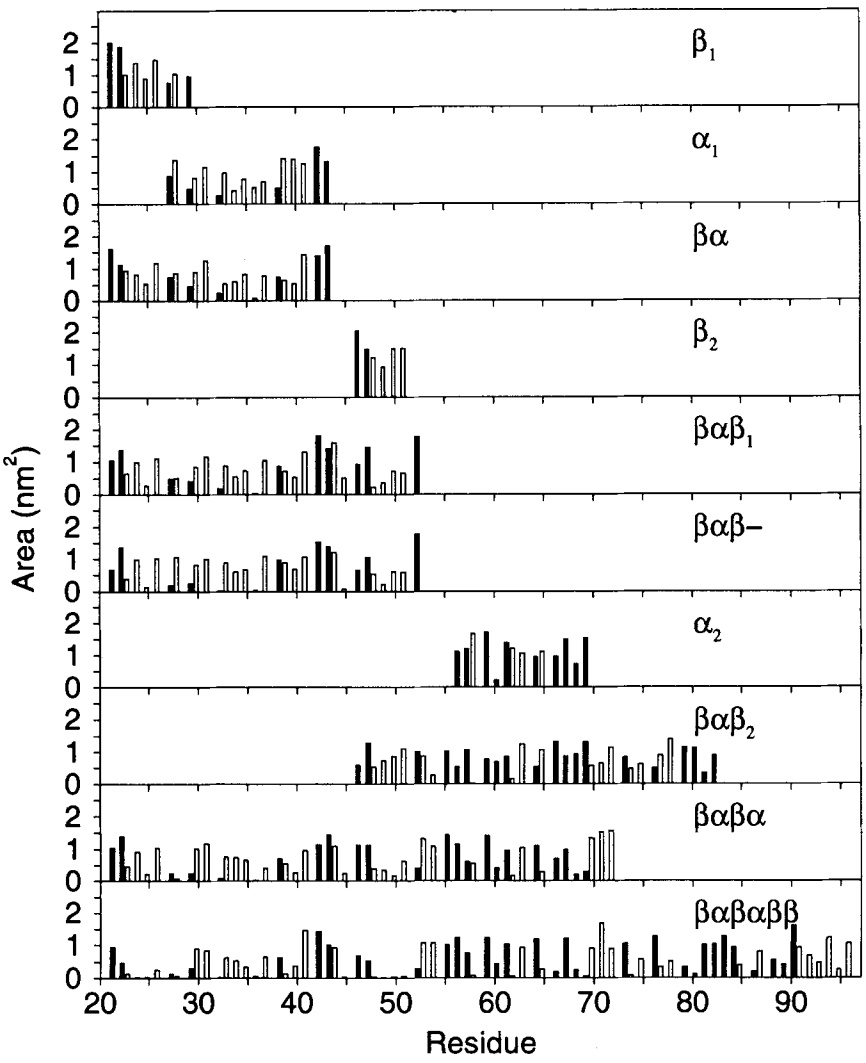

Fig. 10. Average solvent accessible surface area per residue. Solid bars, hydrophilic residues; open bars, hydrophobic residues.

seems less well preserved than in compactly folded proteins, but significantly better than in linear peptides. In none of our simulations does a break occur in the β-sheet; once formed, this unit seems extremely stable. Perhaps this stability is related to the need for cooperative and simultaneous disruption of three or more hydrogen bonds, as was also found from a model study.[78] In addition, hydrophobic contacts between the β-sheet and the α_1-helix contribute significantly to the stability of this supersecondary structure (for contribution of electrostatics, see below). In the absence of MD simulations of isolated β-sheets, it is difficult at this stage to interpret our data in great detail. However, it has been reported that the β-sheet of lysozyme, for example, is highly resistant to denaturation as well.[79] To compare our data for the $\beta\alpha\beta_1$ unit, we also simulated the $\beta\alpha\beta_2$ unit of LDH. This unit is also comparatively stable, although somewhat less so, because the α_2-helix is not fully preserved, unlike the α_1-helix in the $\beta\alpha\beta_1$ unit. However, in the $\beta\alpha\beta_2$ simulation the small β-sheet is also stable and does not come apart. Again, hydrophobic contacts between α-helix and β-sheet stabilize this unit.

A minimal requirement for a folding nucleus is that it can stabilize structure in polypeptide segments that are subsequently added to it. Indeed, the $\beta\alpha\beta_1$ unit of LDH fulfills this requirement (see Results and below). Since the $\beta\alpha\beta_1$ unit is the first N-terminal supersecondary structure element that could fold into a stable structure, it can in principle fulfill the role of an N-terminal folding nucleus in LDH. Even when later parts of the sequence would be more stable thermodynamically in an isolated peptide, during synthesis on the ribosome they are formed later; when our $\beta\alpha\beta_1$ unit does indeed form during translation, this will be the nucleus upon which the remainder of the protein folds. Clearly,

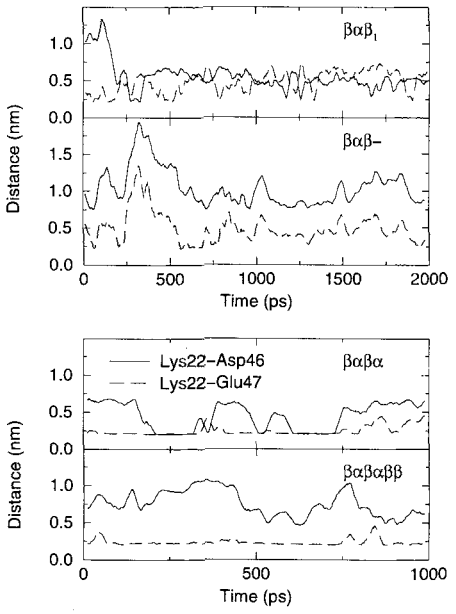

Fig. 11. Distance between charged groups of side chains of Lys-22 and Asp-46 and between Lys-22 and Glu-47 during the MD simulations. A running average over 25 ps is given to improve clarity.

studies with synthetic peptides are required to verify this behavior experimentally. Provided that solubility of these peptides is sufficient, such studies should demonstrate that a "nascent" $\beta\alpha\beta_1$ unit can form in aqeous solution, analogous to the well-known "nascent" α-helix formation.[12,13]

A simulation of the $\beta\alpha\beta^-$ unit was set up with the purpose of investigating the role of the salt bridge between Lys-22 and Glu-47. Removal of this electrostatic interaction gives rise to a destabilization of the α_1-helix in the $\beta\alpha\beta_1$ unit (Fig. 6), underscoring the importance of this interaction for the overall structural stability of the $\beta\alpha\beta_1$ unit. It is possible that the long-range electrostatic interaction between Lys-22 and the combination Asp-46 and Glu-47 can help to drive the formation of the β-sheet of $\beta\alpha\beta_1$. Studies with mutated proteins and synthetic peptides have indeed shown that attractive electrostatic interactions can place secondary structure elements in their correct orientation; moreover, they can even accelerate their folding.[80–82]

We have also simulated regions that extend beyond the first βαβ unit. We note that the α_2-helix is notoriously unstable on its own; however, as part of the βαβα structure it gains considerably in structural stability. When further elements of the structure are added to form the βαβαββ structure, the extra β-strands are also stable in the simulation. The reduction in solvent accessible surface area of the β-sheet and α-helix regions in the βαβαββ unit also contributes to a remarkable stabilization of these units. This is related to the more efficient internal packing of hydrophobic residues than can occur in larger units.

The inaccurate treatment of long-range electrostatic interactions requires further consideration. It has been well established that for uncorrelated dipoles (such as in water) the use of a cut-off of 1.0 nm is acceptable, but if full charges are present artefacts may occur.[53] For distances below the cut-off, the charge interactions may be somewhat exaggerated due to the neglect of screening effects and the omission of electronic polarizability. For distances above the cut-off the interactions are ignored. In aqueous solution screening is such that effective charge interactions are small; this is clearly exemplified by the difference in pK values for the two acid groups in adipic acid, which are at a distance of ~ 0.6 nm and differ by 0.02 pK unit, corresponding to a negligable 0.12 kJ mol^{-1}.

Whereas our MD simulations do not prove in themselves that folding of this part of the protein occurs in a cotranslational manner, our results are consistent with this suggestion, because we establish here that from the βαβ unit onwards, the folded portion is relatively stable. This is a necessary minimum requirement for cotranslational folding. The outcome of our studies hints at the possibility that the folding of most proteins containing a Rossmann fold, which all share amino acid homology, can in principle proceed in a similar cotranslational fashion. This agrees with other work indicating that the first 50 amino acids of many dehydrogenases contain a folding unit.[28] Whether this process can play a role in the folding of other classes of proteins remains to be determined. However, a statistical analysis of the compactness, number of neighboring amino acids, and stability of secondary structure in a large number of proteins suggests that it may be widespread.[83]

CONCLUSIONS

Our MD simulations show that the α_1-helix of LDH is rather stable, when compared with similar simulations of a range of α-helices. The hydrophobic character of this helix makes it an ideal candidate for spontaneous formation through hydrophobic collapse[70]; its hydrophobic surface can subsequently interact with other hydrophobic residues of the β_1 and β_2 strands, as noted throughout our simulations. The $\beta\alpha\beta_1$ unit is the first supersecondary structure element in the sequence of LDH; this unit is stable in our simulations; in particular, its β-sheet never falls apart and the α_1-helix is not destabilized. Its formation may be facilitated by the presence of the Gly-rich loop; indeed, proper refolding of the α_1-helix in the βα unit could be observed in our simulations. The $\beta\alpha\beta_1$ unit is further stabilized by the formation of the β-sheet and the cross-strand ion pair. If it is stable, it can function as a nucleus upon which the remainder of the structure can be assem-

bled[27]: this is demonstrated by the markedly increased structural stability of the α_2-helix in the $\beta\alpha\beta\alpha$ unit compared with the simulation of the isolated α_2-helix. Also, the strands β_3 and β_4 are extremely stable when simulated as part of the $\beta\alpha\beta\alpha\beta\beta$ unit; as isolated units, they would probably behave like the structurally flexible β_1 and β_2 strands. These MD results suggest a potential pathway for cotranslational folding of the nucleotide binding fold. Future high-resolution NMR and CD experiments with N-terminal peptides of LDH with increasing length should test the validity of this proposal.

ACKNOWLEDGMENTS

The authors thank Dr. A.R. van Buuren and Dr. R.M. Scheek for useful comments. D.v.d.S. acknowledges support from the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO). H.J.V. acknowledges support from the Medical Research Council of Canada and from the Netherlands Organization for Scientific Research (NWO) for a visiting professorship.

REFERENCES

1. Jaenicke, R. What does protein refolding in vitro tell us about protein folding in the cell? Philos. Trans. R. Soc. Lond. 339:287–295, 1993.
2. Anfinsen, C.B. Principles that govern the folding of protein chains. Science 181:223–230, 1973.
3. Ohgushi, M., Wada, A. 'Molten-globule state': A compact form of globular proteins with mobile side-chains. FEBS Lett. 164:21–24, 1983.
4. Kuwajima, K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. Proteins 6:87–103, 1989.
5. Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E., Razgulyaev, O.I. Evidence for a molten globule state as a general intermediate in protein folding. FEBS Lett. 262: 20–24, 1990.
6. Kim, P.S., Baldwin, R.L. Intermediates in the folding reactions of small proteins. Annu. Rev. Biochem. 60:631–660, 1990.
7. Roder, H., Elöve, G.A., Englander, S.W. Structural characterization of folding intermediates in cytochrome c by H-exchange labeling and proton NMR. Nature 335:700–704, 1988.
8. Udgaonkar, J.B., Baldwin, R.L. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. Nature 335:694–699, 1988.
9. Englander, S.W., Mayne, L. Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. Annu. Rev. Biophys. 21:243–265, 1992.
10. Sosnick, T.R., Mayne, L., Hiller, R., Englander, S.W. The barriers in protein folding. Nature Struct. Biol. 1:149–156, 1994.
11. Creighton, T.E. The energetic ups and downs of protein folding. Nature Struct. Biol. 1:135–138, 1994.
12. Wright, P.E., Dyson, H.J., Lerner, R.A. Conformation of peptide fragments of proteins in aqueous solution: Implications for initiation of protein folding. Biochemistry 27: 7167–7175, 1988.
13. Dyson, H.J., Wright, P.E. Peptide conformation and protein folding. Curr. Opin. Struct. Biol. 3:60–65, 1993.
14. Shin, H.-C., Merutka, G., Waltho, J.P., Tennant, L.L., Dyson, H.J., Wright, P.E. Peptide models of protein folding initiation sites. 3. The G-H helical hairpin of myoglobin. Biochemistry 32:6356–6364, 1993.
15. Schulz, G.E., Schirmer, R.H. "Principles of Protein Structure." New York: Springer Verlag, 1978.
16. Tsou, C.-L. Folding of the nascent peptide chain into a biologically active protein. Biochemistry 27:1807–1812, 1988.
17. Fedorov, A.N., Friguet, B., Djavadi-Ohaniance, L., Alakhov, Y.B., Goldberg, M.E. Folding on the ribosome of *Escherichia coli* tryptophan synthase β subunit nascent chains probed with a conformation-dependent monoclonal antibody. J. Mol. Biol. 228:351–358, 1992.
18. Fedorov, A.N., Baldwin, T.O. Contribution of cotranslational folding to the rate of formation of native protein structure. Proc. Natl. Acad. Sci. USA 92:1227–1231, 1995.
19. Hata-Tamake, A., Kawai, G., Yamasaki, K., Ito, Y., Kajiura, H., Ha, J.-M., Miyazawa, T., Yokoyama, S., Nishimura, S. Spin-labeling proton NMR study on aromatic amino acid residues in the guanine nucleotide binding site of human c-Ha-ras(1–171) protein. Biochemistry 28:9550–9556, 1989.
20. Minara, P., Hall, L., Betton, J.-M., Missiakis, D., Yon, J.M. Efficient expression and characterization of isolated structural domains of yeast phosphateglycerate kinase generated by site-directed mutagenesis. Protein Eng. 3:55–60, 1989.
21. Eder, J., Kirschner, K. Stable substructures of eightfold βα-barrel proteins: Fragment complementation of phosphoribosylanthranilate isomerase. Biochemistry 31:3617–3635, 1992.
22. Jecht, M., Tomschy, A., Kirschner, K., Jaenicke, R. Autonomous folding of the excised coenzyme-binding domain of an D-glyceraldehyde 3-phosphate dehydrogenase from *Thermotoga maritima*. Protein Sci. 3:411–418, 1994.
23. Schulz, G.E. Binding of nucleotides by proteins. Curr. Opin. Struct. Biol. 2:61–67, 1992.
24. Baker, P.J., Britton, K.L., Rice, D.W., Stillman, A.R.T.J. Structural consequences of sequence patterns in the fingerprint region of the nucleotide binding fold. J. Mol. Biol. 228:662–671, 1992.
25. Saraste, M., Sibbald, P.R., Wittinghofer, A. The P-loop—a common motif in ATP and GTP binding proteins. Trends Biochem. Sci. 15:430–434, 1990.
26. Bossemeyer, D. The glycine-rich sequence of protein kinases: A multifunctional element. Trends Biochem. Sci. 19:201–205, 1994.
27. Wierenga, R.K., Terpstra, P., Hol, W.G.J. Prediction of the occurence of the ADP-binding βαβ-fold in proteins, using an amino acid sequence fingerprint. J. Mol. Biol. 187:101–107, 1986.
28. Krasheninnikov, I.A., Komar, A.A., Adzhubei, I.A. Frequencies of utilization of codons in mRNA and coding of the domain structure of proteins. Dokl. Akad. Nauk. SSSR 305:1006–1012, 1989.
29. Opitz, U., Rudolph, R., Jaenicke, R., Ericsson, L., Neurath, H. Proteolytic dimers of porcine muscle lactate dehydrogenase: Characterisation, folding, and reconstitution of the truncated and nicked polypeptide chain. Biochemistry 26: 1399–1406, 1987.
30. Rossmann, M.G., Moras, D., Olsen, K.W. Chemical and biological evolution of a nucleotide-binding protein. Nature 250:194–199, 1974.
31. Daggett, V., Levitt, M. Realistic simulations of native-protein dynamics. Annu. Rev. Biophys. Biomol. Struct. 22: 353–380, 1993.
32. Caflisch, A., Karplus, M. Molecular dynamics studies of protein and peptide folding and unfolding. In: "The Protein Folding Problem and Tertiary Structure Prediction." Merz Jr., K.M., Le Grand, S.M. (eds.). Boston: Birkhäuser, 1994: 193–230.
33. Mark, A.E., van Gunsteren, W.F. Simulation of the thermal denaturation of hen egg white lysozym: Trapping the molten globule state. Biochemistry 31:7745–7748, 1992.
34. Daggett, V., Levitt, M. A model of the molten globule state from molecular dynamics simulations. Proc. Natl. Acad. Sci. USA 89:5142–5146, 1992.
35. Brooks III, C.L. Molecular simulations of peptide and protein unfolding: In quest of a molten globule. Curr. Opin. Struct. Biol. 3:92–98, 1993.
36. Tirado-Rives, J., Jorgensen, W.L. Molecular dynamics simulations of the unfolding of apomyoglobin in water. Biochemistry 33:4175–4184, 1993.
37. DiCapua, F.M., Swaminathan, S., Beveridge, D.L. Theoretical evidence for destabilization of an α helix by water

insertion: Molecular dynamics of hydrated decaalanine. J. Am. Chem. Soc. 112:6768–6771, 1990.
38. Tirado-Rives, J., Jorgensen, W.L. Molecular dynamics simulations of the unfolding of an α-helical analogue of ribonuclease A S-peptide in water. Biochemistry 30:3864–3871, 1991.
39. deLoof, H., Nilsson, L., Rigler, R. Molecular dynamics simulations of galanin in aqueous and nonaqueous solution. J. Am. Chem. Soc. 114:4028–4035, 1992.
40. van Buuren, A.R., Berendsen, H.J.C. Molecular dynamics simulation of the stability of a 22 residue alpha-helix in water and 30% trifluoroethanol. Biopolymers 33:1159–1166, 1993.
41. Hermans, J. Molecular dynamics simulations of helix and turn propensities in model peptides. Curr. Opin. Struct. Biol. 3:270–276, 1993.
42. Mierke, D.F., Kessler, H. Improved molecular dynamics simulations for the determination of peptide structures. Biopolymers 33:1003–1017, 1993.
43. Soman, K.V., Karimi, A., Case, D.A. Molecular dynamics analysis of a ribonuclease c-peptide analogue. Biopolymers 33:1567–1580, 1993.
44. Kazmirski, S.L., Alonso, D.O.V., Cohen, F.E., Prusiner, S.B., Daggett, V. Theoretical studies of sequence effects on the conformational properties of a fragment of the prion protein: Implications for scrapie formation. Curr. Biol. 2:305–315, 1995.
45. Hirst, J.D., Brooks III, C.L. Molecular dynamics simulations of isolated helices of myoglobin. Biochemistry 34: 7614–7621, 1995.
46. Kovacs, H., Mark, A.E., Johansson, J., van Gunsteren, W.F. The effect of environment on the stability of an integral membrane helix: Molecular dynamics simulations of surfactant protein c in chloroform, methanol and water. J. Mol. Biol. 247:808–822, 1995.
47. Braxenthaler, M., Avbelj, F., Moult, J. Structure, dynamics and energetics of initiation sites in protein folding: I. Analysis of a 1 ns molecular dynamics trajectory of an early folding unit in water: The helix I/loop I-fragment of barnase. J. Mol. Biol. 250:239–257, 1995.
48. Pugliese, L., Prévost, M., Wodak, S.J. Unfolding simulations of the 85-102 β-hairpin of barnase. J. Mol. Biol. 251: 432–447, 1995.
49. Hinds, D.A., Levitt, M. Simulation of protein-folding pathways: Lost in (conformational) space. Trends Biotech. 13: 23–27, 1995.
50. Abad-Zapatero, C., Griffith, J.P., Sussman, J.L., Rossmann, M.G. Refined crystal structure of M4 dogfish apo-lactate dehydrogenase. J. Mol. Biol. 198:445–467, 1987.
51. 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." Dordrecht: D. Reidel Publishing Company, 1981:331–342.
52. Allen, M.P., Tildesley, D.J. "Computer Simulations of Liquids." Oxford: Oxford Science Publications, 1987.
53. Smith, P.E., Pettitt, B.M. Peptides in ionic solutions: A comparison of the Ewald and switching function techniques. J. Chem. Phys. 95:8430–8441, 1991.
54. Hockney, R.W., Eastwood, J.W. "Computer Simulation Using Particles." New York: McGraw-Hill, 1981.
55. Luty, B.A., Davies, M.E., Tironi, I.G., van Gunsteren, W.F. A comparison of particle-particle, particle-mesh and Ewald methods for calculating electrostatic interactions in periodic molecular systems. Mol. Sim. 14:11–20, 1994.
56. Tironi, I.G., Sperb, R., Smith, P.E., van Gunsteren, W.F. A generalized reaction field method for molecular dynamics simulations. J. Chem. Phys. 102:5451–5459, 1995.
57. Berendsen, H.J.C. Electrostatic interactions. In: "Computer Simulation of Biomolecular Systems." van Gunsteren, W.F., Wilkinson, P.K.W. (eds.) Leiden: ESCOM, 1993:161–181.
58. Luty, B.A., van Gunsteren, W.F. Calculating electrostatic interactions using the particle-particle, particle-mesh method with non-periodic long-range interactions. Submitted.
59. Berendsen, H.J.C., Postma, J.P.M., DiNola, A., Haak, J.R. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81:8, 1984.
60. van Gunsteren, W.F., Berendsen, H.J.C. "Gromos-87 Manual." Biomos BV Nijenborgh 4, 9747 AG Groningen, The Netherlands 1987.
61. van Buuren, A.R., Marrink, S.J., Berendsen, H.J.C. A molecular dynamics study of the decane/water interface. J. Phys. Chem. 97:9206–9212, 1993.
62. 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. Comp. Phys. 23:327–341, 1977.
63. van der Spoel, D., Berendsen, H.J.C., van Buuren, A.R., Apol, E., Meulenhoff, P.J., Sijbers, A.L.T.M., van Drunen, R. "Gromacs User Manual." Nijenborgh 4, 9747 AG Groningen, The Netherlands. Internet: http://rugmd0.chem.rug.nl/~gmx 1995.
64. Bekker, H., Berendsen, H.J.C., Dijkstra, E.J., Achterop, S., van Drunen, R., van der Spoel, D., Sijbers, A., Keegstra, H., Reitsma, B., Renardus, M.K.R. Gromacs: A parallel computer for molecular dynamics simulations. In: "Physics Computing 92." de Groot, R.A., Nadrchal, J. (eds.). Singapore: World Scientific, 1993.
65. Berendsen, H.J.C., van der Spoel, D., van Drunen, R. Gromacs: A message-passing parallel molecular dynamics implementation. Comp. Phys. Commun. 91:43–56, 1995.
66. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637, 1983.
67. Das, B., Chattopadhyay, S., Gupta, C.D. Reactivation of denatured fungal glucose 6-phosphate dehydrogenase and *E. coli* alkaline phosphatase with *E. coli* ribosome. Biochem. Biophys. Res. Commun. 183:774–780, 1992.
68. Bera, A.K., Das, B., Chattapahyay, S., Dasgupta, C. Refolding of denatured restriction endonucleases with ribosomal preparations from *Methanoscarcina barkeri*. Biochem. Biophys. Res. Commun. 32:315–323, 1994.
69. Girg, R., Jaenicke, R., Rudolph, R. Dimers of porcine skeletal muscle lactate dehydrogenase produced by limited proteolysis during reassociation are enzymatically active in the presence of stabilizing salt. Bioch. Int. 7:433–441, 1983.
70. Dill, K.A., Fiebig, K.M., Chan, H.S. Cooperativity in protein-folding kinetics. Proc. Natl. Acad. Sci. USA 90:1942–1946, 1993.
71. Tobias, D.J., Mertz, J.E., Brooks III, C.L. Nanosecond timescale folding dynamics of a pentapeptide in water. Biochemistry 30:6054–6058, 1991.
72. Blanco, F.J., Jimenez, M.A., Herranz, J., Rico, M., Santoro, J., Nieto, J.L. NMR evidence of a short linear peptide that folds into a β-hairpin in aqueous solution. J. Am. Chem. Soc. 115:5887–5888, 1993.
73. Blanco, F.J., Jimenez, M.A., Pineda, A., Rico, M., Santoro, J., Nieto, J.L. NMR solution structure of the isolated N-terminal fragment of protein-G B_1 domain: evidence of trifluoroethanol induced native-like β-hairpin formation. Biochemistry 33:6004–6014, 1994.
74. Wierenga, R.K., de Maeyer, M.C.H., Hol, W.G.J. Interaction of phosphate moieties with α-helices in dinucleotide binding proteins. Biochemistry 24:1346–1357, 1985.
75. Stroup, A.N., Gierasch, L.M. Reduced tendency to form a β turn in peptides from the p22 tailspike protein correlates with a temperature-sensitive folding defect. Biochemistry 29:9765–9771, 1990.
76. Fry, D.C., Kuby, S.A., Mildvan, A.S. NMR studies of the MgATP binding site of adenylate kinase and of a 45-residue peptide fragment of the enzyme. Biochemistry 24: 4630–4694, 1985.
77. Fry, D.C., Byler, D.M., Susi, H., Brown, E.M., Kuby, S.A., Mildvan, A.S. Solution structure of the 45-residue MgATP-binding peptide of adenylate kinase as examined by 2D-NMR. FTIR, and CD spectroscopy. Biochemistry 27: 3588–3598, 1988.
78. Tobias, D.J., Sneddon, S.F., Brooks III, C.L. Stability of a model β-sheet in water. J. Mol. Biol. 227:1244–1252, 1992.
79. Dobson, C.M., Evans, P.A., Radford, S.E. Understanding how proteins fold: The lysozyme story so far. Trends Biochem. Sci. 19:31–37, 1994.
80. Tweedy, N.B., Hurle, M.R., Chrunyk, B.A., Matthews, C.R. Multiple replacements at position 211 in the α subunit of tryptophan synthase as a probe of the folding unit association reaction. Biochemistry 29:1539–1545, 1990.
81. Schafmeister, C.E., Miercke, L.J.W., Stroud, R.M. Struc-

ture at 2.5 Å of a designed peptide that maintains solubility of membrane proteins. Science 262:734–738, 1993.
82. Monera, O.D., Kay, C.M., Hodges, R.S. Electrostatic interactions control the parallel and antiparallel orientation of α-helical chains in two stranded α-helical coiled coils. Biochemistry 33:3862–3871, 1994.
83. Alexandrov, N. Structural arguments for N-terminal initiation of protein folding. Protein Sci. 2:1989–1991, 1993.
84. Kraulis, P.J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. 24:946–950, 1991.