R. Bürgi
X. Daura
A. Mark
M. Bellanda
S. Mammi
E. Peggion
W. van Gunsteren

Folding study of an Aib-rich peptide in DMSO by molecular dynamics simulations

Authors' affiliations:
R. Bürgi, X. Daura, A. Mark and W. van Gunsteren*, Laboratorium für Physikalische Chemie, ETH Zürich, CH-8092 Zürich, Switzerland. *Present address: Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4, NL-9747 AG, Groningen, The Netherlands.

M. Bellanda, S. Mammi and E. Peggion
University of Padova, Department of Organic Chemistry, Biopolymer Research Center, C.N.R Via Marzolo 1, Padova, Italy, I-35131.

Correspondence to:
Professor *Wilfred van Gunsteren*
Laboratorium für Physikalische Chemie
ETH Zürich
CH-8092 Zürich
Switzerland
E-mail: wfvgn@igc.phys.chem.ethz.ch

Dates:
Received 13 April 2000
Revised 08 June 2000
Accepted 11 July 2000

To cite this article:
Bürgi, R., Daura, X., Mark, A., Bellanda, M., Mammi, S., Peggion, E., van Gunsteren, W. Folding study of an Aib-rich peptide in DMSO by molecular dynamics simulations.
J. Peptide Res., 2001, 57, 107-118.

ISSN 1397-002X

Key words: Aib; DMSO; GROMOS; 3_{10} helix; molecular dynamics; peptide folding

Abstract: To evaluate the ability of molecular dynamics (MD) simulations using atomic force-fields to correctly predict stable folded conformations of a peptide in solution, we show results from MD simulations of the reversible folding of an octapeptide rich in α-aminoisobutyric acid (2-amino-2-methyl-propanoic acid, Aib) solvated in di-methyl-sulfoxide (DMSO). This solvent generally prevents the formation of secondary structure, whereas Aib-rich peptides show a high propensity to form secondary structural elements, in particular 3_{10}- and α-helical structures. Aib is, moreover, achiral, so that Aib-rich peptides can form left- or right-handed helices depending on the overall composition of the peptide, the temperature, and the solvation conditions. This makes the system an interesting case to study the ensembles of peptide conformations as a function of temperature by MD simulation. Simulations involving the folding and unfolding of the peptide were performed starting from two initial structures, a right-handed α-helical structure and an extended structure, at three temperatures, 298 K, 340 K, and 380 K, and the results are compared with experimental nuclear magnetic resonance (NMR) data measured at 298 K and 340 K. The simulations generally reproduce the available experimental nuclear Overhauser effect (NOE) data, even when a wide range of conformations is sampled at each temperature. The importance of adequate statistical sampling in order to reliably interpret the experimental data is discussed.

The prediction of whether and how proteins and peptides in solution fold to a relatively compact, stable structure is one of the grand challenges in computational biochemistry. Although expressed proteins generally fold to a unique,

stable 3-dimensional structure, shorter peptides may adopt a variety of structures in solution. This makes the characterization of the ensemble of relatively stable structures of peptides using experimental methods a non-trivial task. An experiment yields ensemble averaged properties from which it is possible to conclude that the ensemble is dominated by one unique conformation only in particular cases.

In this context, molecular dynamics (MD) simulations have proven to be an excellent tool for studying the process of reversible peptide folding. Recently, secondary structures of various peptides have been studied by means of MD simulation using explicit solvents (1–7), and implicit solvation models (8,9). In contrast with other analysis methods, MD simulation generates an ensemble of structures from which the behaviour of a system can be inferred and averages calculated, which can in turn be compared to experimental averages. With present computers, it is possible to simulate hundreds of ns for small systems. This makes it possible to sample a vast variety of conformations. Moreover, the dynamics of folding–unfolding can be studied, as these processes can occur on the ns to microsecond timescale for small peptides (10).

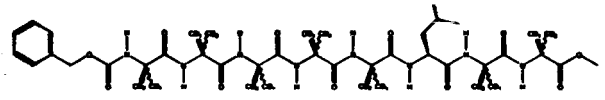
In the current work, we study the secondary structure formation of an octapeptide rich in α-aminoisobutyric acid (2-amino-2-methyl-propanoic acid, Aib) in DMSO. The chemical structure is shown in Fig. 1. As a solvent, DMSO usually reduces or even prevents the formation of stable secondary structure. Aib-rich peptides, however, have a high propensity to adopt secondary structure, particularly 3_{10}- and α-helices (11), even in DMSO. Because the octapeptide contains seven achiral (Aib) residues and only one chiral (L-Leu) residue, it is expected to adopt both left-handed (L) and right-handed (R) helical structures. The presence of the L-Leu residue at position 6 would result in a slight preference for the R-helical form. This makes the present system an interesting test of the interatomic force-field used, provided the MD simulations are sufficiently long so that both L- and R-helical structures are sampled.

Nuclear magnetic resonance (NMR) studies have been performed on the octapeptide studied (12). The lack of α-protons makes NMR experiments on such peptides very difficult to interpret. The NMR results seem to indicate the right-handed 3_{10}-helix as the predominant structure. It is, however, highly desirable to use an independent technique to unambiguously determine the dominant secondary structure of the octapeptide, as the NMR data are a time average over an ensemble of structures, and especially for flexible molecules it is questionable to what extent a single structure can be representative for the ensemble. Therefore, the current study has two aims: (i) to model the conformational equilibria of the octapeptide in solution in order to enable interpretation of the NMR data, and (ii) to test the ability of the force-field to reproduce the available experimental data independent of an initial structural model.

Four MD simulations were performed. To investigate the stability of the right-handed helical fold, two of the simulations were started from a right-handed (R) α-helical structure, one at 298 K ($298_{\alpha R}$) and one at 380 K ($380_{\alpha R}$). To investigate the ability of the simulations to fold the peptide from an arbitrary conformation, two simulations were performed starting from an extended conformation (all backbone dihedral angles equal to 180°), one at 298 K (298_e) and one at 340 K (340_e). Using higher temperatures the equilibrium between folded and unfolded conformations is shifted towards the latter and the number of (un)folding transitions is enhanced, which leads to broader sampling. Table 1 gives an overview of the MD simulations.

Methods

All simulations were performed with the GROMOS96 package (13,14), using the GROMOS force-field 43A1 (14). The peptide was placed in a right-handed α-helical conformation in a periodic truncated octahedron with 768 DMSO molecules (15) for the $298_{\alpha R}$ and the $380_{\alpha R}$ simulations, and in an extended conformation (all backbone torsional angles set to the trans-configuration) with 1119 DMSO molecules for the 298_e and the 340_e simulations.



Z - Aib[1] - Aib[2] - Aib[3] - Aib[4] - Aib[5] - Leu[6] - Aib[7] - Aib[8] - OMe

Figure 1. Chemical formula of the octapeptide. CB_1 and CB_2 indicate the pro-S and pro-R methyl groups, respectively.

Table 1. Overview of the four MD simulations of the octapeptide Z-$(Aib)_5$-L-Leu-$(Aib)_2$-OMe in DMSO

Label	Temperature (K)	Starting configuration	Length (ns)
$298_{\alpha R}$	298	α-helix (right-handed)	50
298_e	298	Extended chain	150
340_e	340	Extended chain	150
$380_{\alpha R}$	380	α-helix (right-handed)	50

Covalent bond lengths were kept rigid using the SHAKE procedure (16) with a geometric tolerance of 10^{-4}. Before starting the simulations, a steepest-descent energy minimization was carried out to relax the solvent molecules around the peptide. Initial velocities were assigned from a Maxwell-Boltzmann distribution at 200 K. During the initial phases of the simulations, the dihedral angles were restrained using a harmonic potential energy function relaxing the force constant from 0.1 kJ mol^{-1} deg^{-2} to 0 within 150 ps.

Peptide and solvent separately were weakly coupled to a temperature bath (17) with a relaxation time of 0.1 ps. In addition, the system was coupled to a pressure bath (17) of 1 atm using the isothermal compressibility $\kappa_t = 4.575 \times 10^{-4}$ (kJ mol^{-1} nm^{-3})$^{-1}$ and a relaxation time of 0.5 ps. The time step for the leap-frog algorithm was set to 0.002 ps. For the nonbonded interactions, a twin-range cutoff was used, evaluating the short-range contributions of the van der Waals and electrostatic interactions within 0.8 nm at every time step and the long-range contributions within 1.4 nm every five time steps. Electrostatic forces outside 1.4 nm were treated using a reaction field with the relative dielectric permittivity $\varepsilon_{RF} = 54$. Trajectory coordinates were saved every 0.5 ps for analysis.

Results and Discussion

Secondary structure

Figure 2 shows the backbone atom positional root mean square deviations (r.m.s.d.) for residues 2–7 from a left-handed (left panels) and a right-handed (right panels) 3_{10}-model helix (upper panels), and from a left-handed and a right-handed α-model helix (lower panels) for all four simulations as a function of time. Figure 3 shows the occurrence per residue of the two predominant secondary structure elements, α-helical (blue) and 3_{10}-helical (red), as defined in the program PROCHECK (18), for the four simulations as a function of time. PROCHECK does not differentiate between left-handed and right-handed structures. A variety of other secondary structure elements such as hydrogen-bonded turns and bent residues were detected. For clarity, these were omitted from Fig. 3.

Both Fig. 2 and Fig. 3 show that the helical structure in the $298_{\alpha R}$ simulation is stable for extended periods on the ns time-scale; during the first 25 ns, only transitions between the α-helical and the 3_{10}-helical structure are observed. Starting from an extended structure (simulation 298_e) R-helical structures are sampled at different times throughout

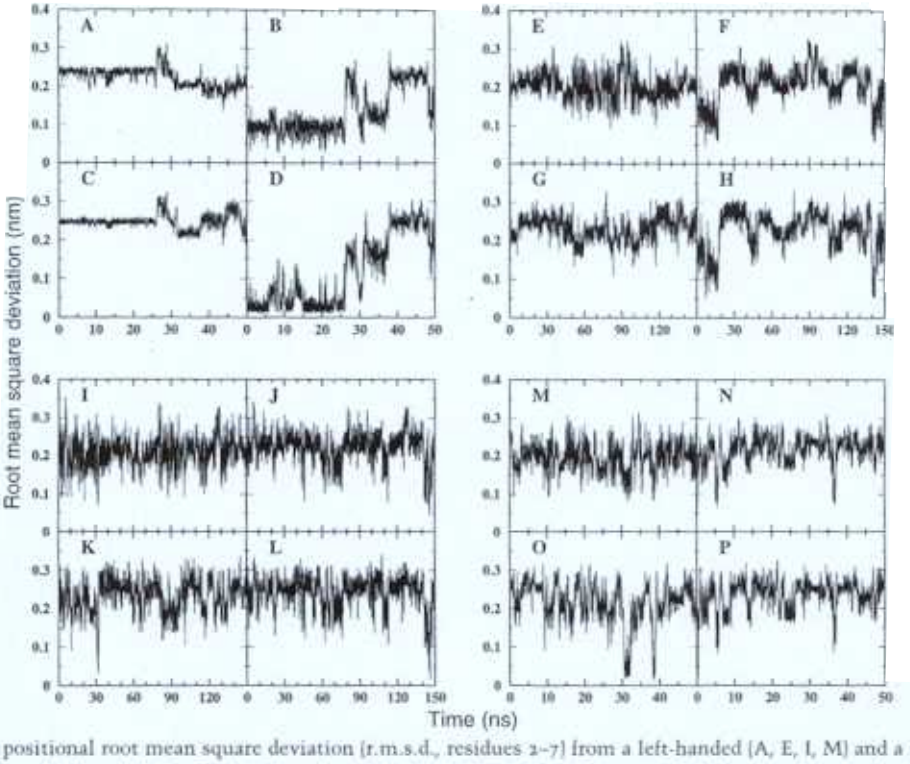


Figure 2. Backbone atom positional root mean square deviation (r.m.s.d., residues 2–7) from a left-handed (A, E, I, M) and a right-handed (B, F, J, N) 3_{10}-model helix and from a left-handed (C, G, K, O) and a right-handed (D, H, L, P) α-model helix for the simulations (see Table 1) $298_{\alpha R}$ (A, B, C, D), 298_e (E, F, G, H), $340_{\alpha L}$ (I, J, K, L), $380_{\alpha R}$ (M, N, O, P).

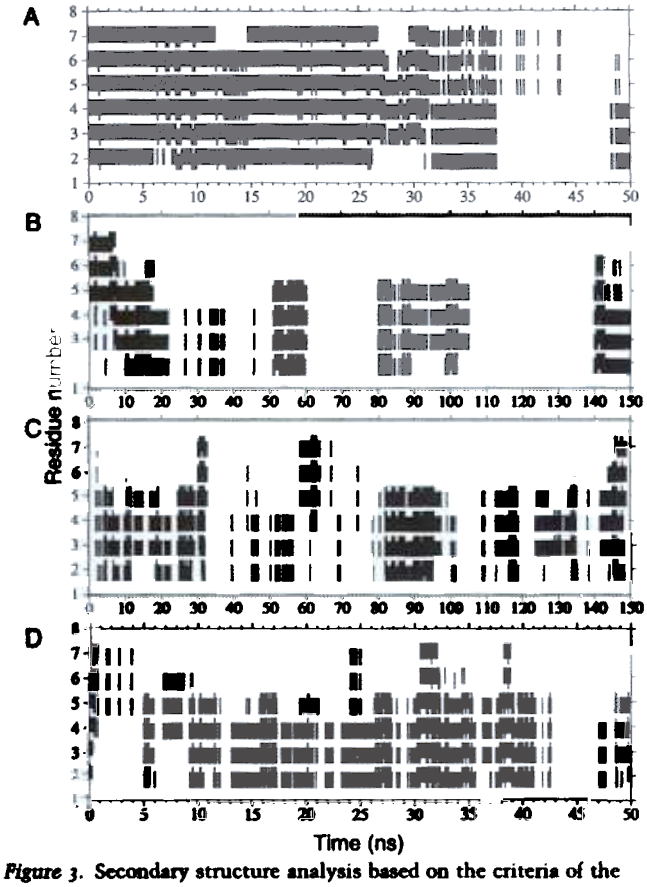


Figure 3. Secondary structure analysis based on the criteria of the program PROCHECK (18) for the simulations $298_{\alpha R}$ (A), 298_e (B), 340_e (C), $380_{\alpha R}$ (D). Residues with 3_{10}-helical conformations are plotted red, those with α-helical conformations in blue.

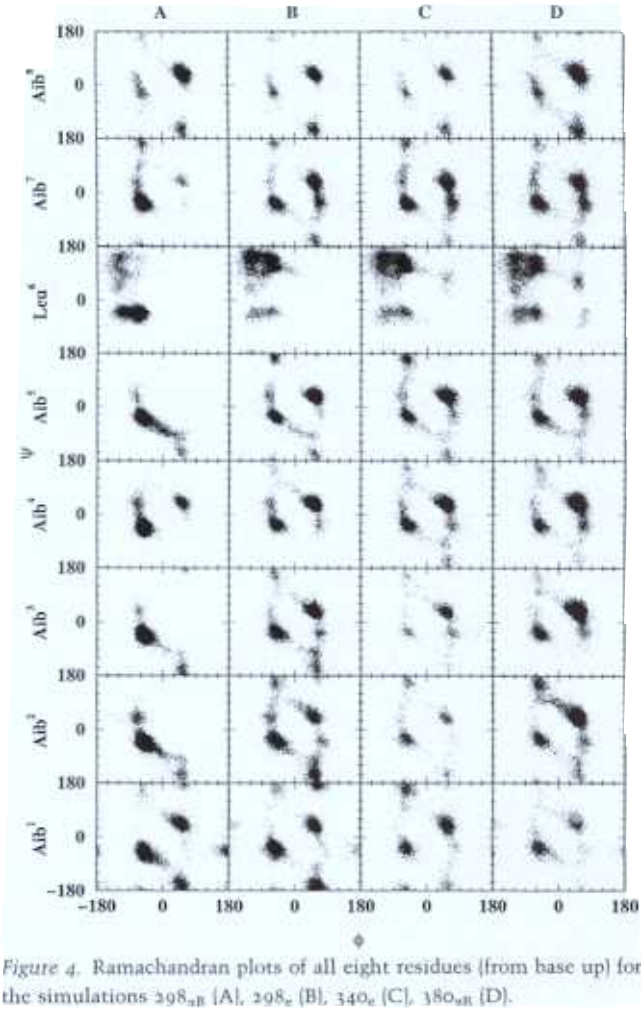


Figure 4. Ramachandran plots of all eight residues (from base up) for the simulations $298_{\alpha R}$ (A), 298_e (B), 340_e (C), $380_{\alpha R}$ (D).

the simulation. Comparing the two room-temperature simulations ($298_{\alpha R}$ and 298_e) it is clear that the former is dominated for tens of ns by its R-helical starting structure. This is not the case at 380 K (simulation $380_{\alpha R}$), where the α-helix is lost after half a ns and the N-terminal part of the peptide then folds repeatedly into 3_{10}-helical structures (Fig. 3). A similar pattern is observed in the 340_e simulation. When raising the temperature from 298 K (298_e simulation) to 340 K (340_e simulation) and to 380 K ($380_{\alpha R}$ simulation), the time-scale for major conformational changes becomes shorter (Fig. 3): from 10–30 ns (298_e) to 4–10 ns (340_e) and to 1–4 ns ($380_{\alpha R}$). Figure 3 also shows how close in conformational space the α- and the 3_{10}-helical structures are. Once a helical conformation is reached the peptide starts swapping between α- and 3_{10}-helical forms.

As for the chirality of the peptide conformations, in Fig. 2 it can be seen that at room temperature only right-handed helical structures are formed. However, at room temperature the time-scale of major conformational changes is of the order of tens of ns. Even simulations of 50–150 ns do not result in a representative sampling of conformational space, and it is therefore not possible to draw definitive conclusions based on the $298_{\alpha R}$ and 298_e simulations. At the higher temperatures, the peptide behaves as though it were essentially achiral. In the $380_{\alpha R}$ simulation, there seems to be a slight preference for left-handed structures over right-handed. These observations are confirmed in the Ramachandran plots shown in Fig. 4; the higher the temperature, the greater the apparent preference for left-handed structures (i.e. more density in the upper-right quadrant). Even at room temperature the last residue, Aib⁸, is predominantly found in left-handed configurations. This is in agreement with the observation of Karle & Balaram (11) that in crystal structures of Aib-containing peptides, the ones terminating with a Leu-Aib-OMe, an Aib-Aib-OMe or an Ala-Aib-OMe sequence generally (14 out of 17) show a left-handed helical conformation at the final Aib residue. The crystal structure of a bromo-substituted analog of the octapeptide simulated here contains two equally populated enantiomeric helical conformations for the terminal residue Aib⁸ (19). Furthermore, the Ramachandran plots show that,

except for the $298_{\alpha R}$ simulation, the Leu residue prefers extended conformations as opposed to helical ones. Thus, L-Leu serves rather as a helix-breaker than as a helix-former in the octapeptide, in agreement with the observation that Leu is more frequently found in β- than in α-structures (20).

Clustering

In order to obtain a better picture of the conformational variety in the ensemble of structures of the octapeptide, the structures have been grouped into clusters with respect to their backbone atom positional r.m.s.d. for residues 2–7. For the precise definition of the clustering algorithm, we refer to Daura *et al.* (3). Each cluster consists of structures that differ by less than 0.1 nm in backbone atom positional r.m.s.d. The clusters are mutually exclusive, i.e. different clusters cannot contain the same structure. For the $298_{\alpha R}$ simulation 26 clusters were found; for the 298_e simulation 77 clusters; for the 340_e simulation 147 clusters; and for the $380_{\alpha R}$ simulation 105 clusters. Figure 5 shows the relative population of all the clusters. For all simulations, only about 20 clusters are populated by more than 1%. The ensemble of peptide structures is dominated by only about 5–10 peptide conformations. This is in agreement with MD simulations of the unfolded state of other peptides showing that the unfolded or denatured state is not of random conformational nature but can be characterized by a low number, 10^1–10^2, of conformations (2,3). Example conformations selected from the trajectories of the octapeptide are shown in Fig. 6. In panel I, the variety of backbone structures within a cluster is indicated: the five superimposed structures shown are members of the first (most populated) cluster of the $298_{\alpha R}$ simulation taken at different times. Because the clustering criterion involves only the backbone atoms of residues 2–7, the terminal residues and the side-chains show a variety of conformations.

A secondary structure analysis, using the program PROCHECK, of the different clusters can be found in Fig. 7. The lengths of the bars indicate the percentage of occurrence of a given secondary structure element. Simulation $298_{\alpha R}$ (panel A) is biased by its α-helical (dark blue) starting structure. In the simulations 340_e and $380_{\alpha R}$ (panels C and D), the most populated cluster (sequence number 1) has its N-terminal part folded into a 3_{10}-helical (red) form. Cluster one in simulation 298_e (panel B) is only slightly 3_{10}-helical in its N-terminal part, whereas cluster two is more helical. However, as noted before, the room-temperature simulations $298_{\alpha R}$ and 298_e are not long enough to provide adequate statistics. In terms of the number of highly populated clusters as well as their relative populations (Fig. 5A,B) and

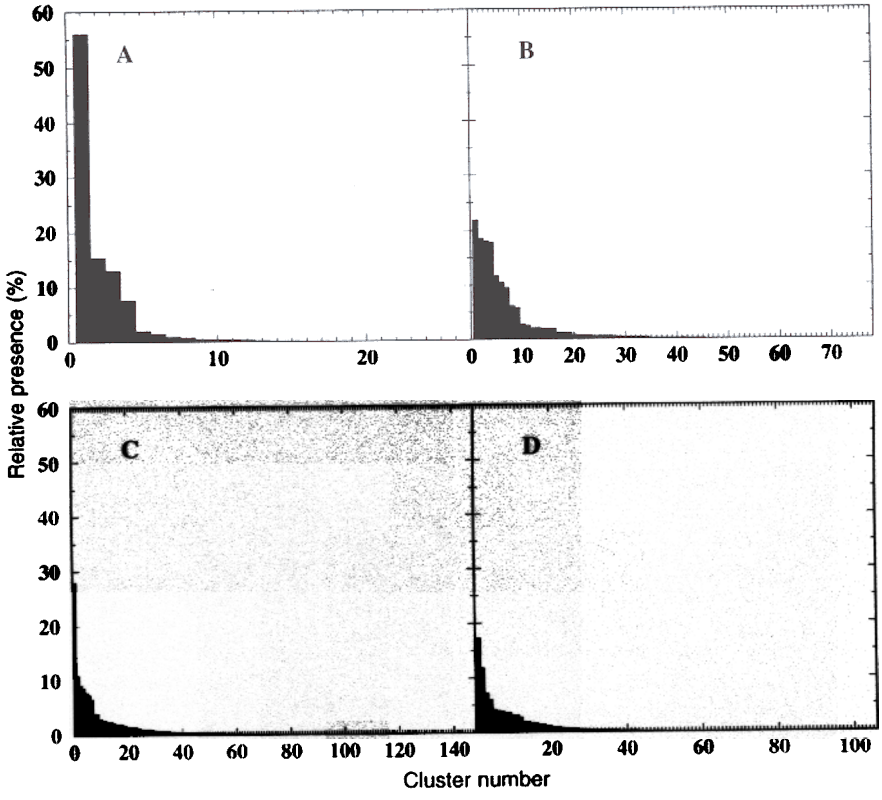


Figure 5. Relative population of all clusters in simulations $298_{\alpha R}$ (A), 298_e (B), 340_e (C), and $380_{\alpha R}$ (D) using a backbone atom positional r.m.s.d. (residues 2–7) of less than 0.1 nm as the clustering criterion.

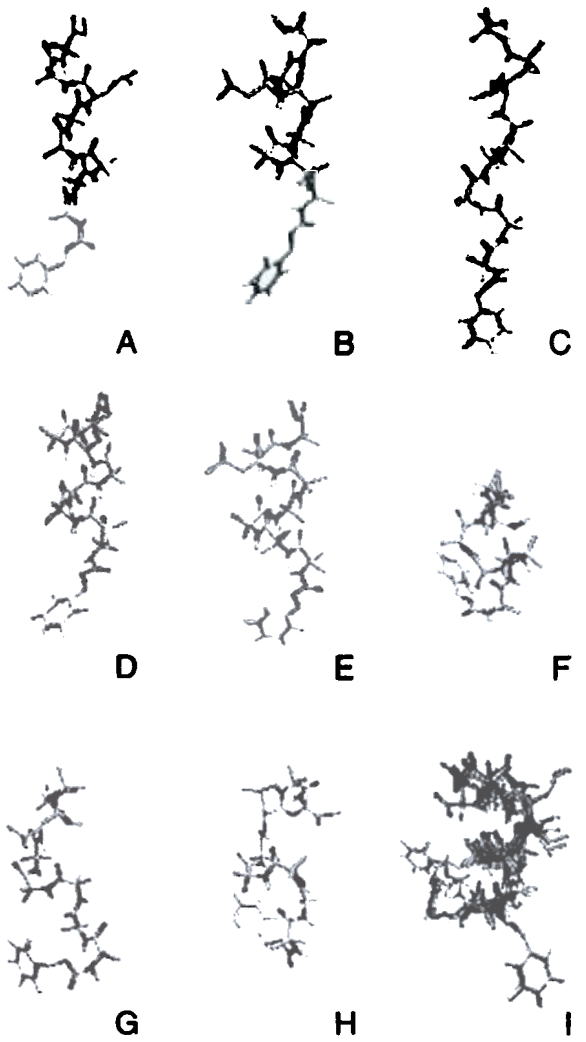


Figure 6. Selected structures of the octapeptide: left-handed 3_{10} helix (A); right-handed 3_{10} helix (B); left-handed α-helix (D); right-handed α-helix (E); structure with maximum radius of gyration in simulation 298_e (C); structure with minimum radius of gyration in simulation 298_e (F); central member of the most populated cluster in simulation 298_e (G); central member of the most populated cluster in simulation 340_e (H); several members of the most populated cluster of simulation $298_{\alpha R}$. (I).

the dominant type of secondary structure present (Fig. 7A,B), the simulations at 298 K are not converged. The sampling difficulties at 298 K are also clear when comparing results at different temperatures. When lowering the temperature from 380 K to 340 K and 298 K, one would normally expect the relative population of the most stable (lowest cluster sequence number) conformation to increase. This is the case comparing simulations $380_{\alpha R}$ and 340_e in Fig. 5, but not when comparing simulations 340_e and 298_e. The structures of clusters one and two of simulation 298_e, (see Fig. 8), are very similar, and their relative populations differ by less than 4%. Yet, their secondary structures are characterized differently by the program PROCHECK (Fig. 7B). Structures in cluster 1 are mainly identified as being comprised of turns, while structures in cluster 2 are predominantly recognized as partial 3_{10}-helices. Longer simulations at 298 K may easily shift the character of the most populated cluster towards 3_{10}-helical as it is found in the simulations at higher temperatures (Fig. 7C,D).

To explore the handedness of the peptide conformations, a detailed analysis of the handedness per residue was made for the first 10 clusters of each simulation. Table 2 shows the results grouped into R (third quadrant of the Ramachandran plot); L (first quadrant of the Ramachandran plot); and – (second or fourth quadrant of the Ramachandran plot). Residues 1 and 8 were not considered as they were excluded from the clustering criterion. Except for the $298_{\alpha R}$ simulation, which is biased by its starting structure, there is an overall preference for L-handed conformations. However, we can see that both L- and R-conformations are often present at the same time for different residues along the peptide chain. Sometimes the handedness changes more than once along the peptide chain. It is also apparent that the Leu[6] residue is hardly found in either R- or L-handed conformations.

Comparison with experiment

For a comparison of the simulated proton–proton distances with experimentally determined nuclear Overhauser effect (NOE) distance bounds, 34 values at 298 K and 32 values at 340 K were available (12). As all simulations were based on a united atom model, distances to methyl groups had to be calculated using appropriate pseudo atoms (14). For comparison with the experimental NOE bounds including methyl groups, the latter had to be recalculated in the following way:

1. All the experimental NOE distance bounds to methyl groups had been obtained as a weighted sum, $\bar{r}_{ij}^{-3} = \sum_{k=1} \bar{r}_{ijk}^{-3}$, where r_{ij} denotes the distance from proton i to the methyl group j and r_{ijk} is the distance from proton i to proton k of the methyl group j (12). The distances calculated with pseudo atoms, however, correspond to a geometric mean position of the three protons of the methyl group $\bar{r}_{ij}^{-3} = {}^1/_3 \sum_{k=1} \bar{r}_{ijk}^{-3}$. Therefore, the experimental NOE bounds involving methyl groups of Bellanda *et al.* (12) were multiplied by a factor of $3^{1/3}$.

2. A pseudo atom correction of 0.03 nm was added to the NOE bounds involving methyl groups (21).

The modified NOE distance bounds are listed in Table 3. The NOE bounds to the Aib methyl groups CB1 (pro-S) and CB2 (pro-R) have been assigned such that they match the

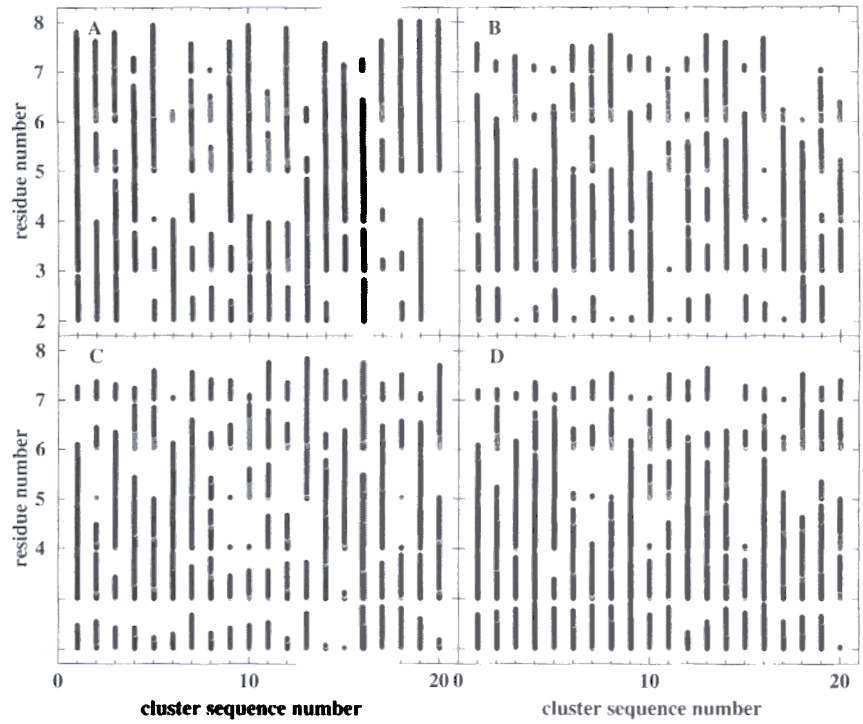


Figure 7. Secondary structure analysis of the 20 most populated clusters for the simulations $298_{\alpha R}$ (A), 298_e (B), 340_e (C), $380_{\alpha R}$ (D) calculated with the program PROCHECK (18). 3_{10}-Helical residues are plotted in red, α-helical in dark blue, bends in brown, turns in light blue, and beta bridges in yellow. The lengths of the bars indicate the percentage of occurrence of a given secondary structure element per residue. The clusters are ranked according to decreasing population.

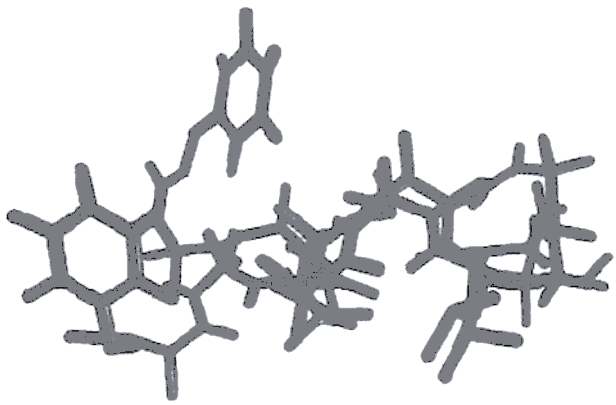


Figure 8. Structures of the central members of clusters one (most populated, blue) and two (second most populated, red) of simulation 298_e.

right-handed 3_{10}-helix. Because we did not want to assume any previous knowledge about the structure of the peptide, we calculated the violation of the NOE distance bounds for both the assignments as given in Table 3 (Fig. 9A) as well as the opposite assignments (CB1 and CB2 interchanged) for the Aib methyls (Fig. 9B). Considering Fig. 9A, all the NOE-bound violations (blue and red lines indicate bound violations of experimental data collected at 298 K and 340 K, respectively) are below 0.1 nm for the $298_{\alpha R}$ simulation. The biggest violation observed is 0.08 nm for the distance 5CB2–7HN (no. 30). This simulation is, however, dominated by the R-helical starting structure. For the simulation starting from the extended structure (298_e), the same NOE shows the biggest violation. It is the only violation greater than 0.1 nm. Comparison with the $298_{\alpha R}$ simulation shows that the NOE violations are reduced as soon as there is a longer period of helical structures. However, in neither simulation is the sampling sufficient. For the 340_e simulation, all the NOE violations are below 0.05 nm. The NOE bounds are thus essentially satisfied at this temperature. Even for the $380_{\alpha R}$ simulation, the NOE data measured at 340 K are basically satisfied.

When choosing the opposite assignment of NOE peaks to the chirally indistinguishable methyl groups CB1 and CB2, e.g. by assigning the NOE bounds to match a left-handed 3_{10}-helix, i.e. by exchanging CB1 and CB2 in Table 3, we notice that all the NOE-bound violations increase in the $298_{\alpha R}$ simulation (Fig. 9B). This is understandable as the $298_{\alpha R}$ simulation is biased by its right-handed starting structure. However, for all the other simulations we see that the NOE bounds 1CB2–4HN (no. 8), 2CB2–5HN (no. 14) and 3CB2–6HN (no. 19) are less violated when the assignments are made according to a left-handed 3_{10}-helix, whereas the NOE bounds 5CB2–7HN (no. 30) and 5CB2–8HN (no. 31) are less violated if the assignments match the right-handed 3_{10}-helix. This indicates that in the 298_e, 340_e and $380_{\alpha R}$ simulations, the first five residues seem to prefer a left-handed rather than a right-handed conformation, and that there is a change in chirality at the Leu[6]. As we can see from Fig. 4, the last two residues prefer a left-handed conformation again. It is also interesting to see that the NOEs

Table 2. Handedness of the central member conformations of the first 10 clusters for the simulations $298_{\alpha R}$, 298_{e}, 340_{e}, and $380_{\alpha R}$ as a function of the residue number. The population (Pop.) of the clusters is given in per cent. The symbols R and L indicate ϕ,ψ angles in the third ($\phi,\psi<0$) and first ($\phi,\psi>0$) quadrant of the Ramachandran plot, respectively; the symbol – indicates ϕ,ψ angles in the second ($\phi>0$, $\psi<0$) or fourth ($\phi<0$, $\psi>0$) quadrant of the Ramachandran plot. Residues 1 and 8 were not considered, as they were excluded from the clustering criterion.

		Residue number					
Nr.	Pop. (%)					6	
Simulation $298_{\alpha R}$							
	56	R	R	R	R	R	R
	15	R	R	L	R	–	–
	13	R	R	–	R	R	R
	8	R	–	R	R	–	R
		–	R	–	R	R	R
		R	R	L	–	–	R
			–	R	–	R	R
		L	R	L	L	–	–
	0.5	R	R	R	R	R	R
10	0.5	R	R	–	R	R	R
Simulation 298_{e}							
	22	–	R	L	L		L
	18	L	L	L	L		R
	18	R	L	L	R		L
	18	–	R	R	L		R
	12	R	R	R	R		R
	10	R	L	L	–		–
	9	L	R	–	R		R
8	6	R	L	L	R		–
9	6	L	–	L	L		R
10	3	–	R	–	–		R
Simulation 340_{e}							
1	28	L	L	L	L		L
2	11	L	–	R	L		–
3	9	R	–	L	L		L
4	9	–	L	L	R		R
5	8	R	L	L	–		L
6	8	R	L	L	L		R
7	7	L	L	R	R		–
8	4	R	L	–	L		L
9	4	R	–	–	L		L
10	3	–	L	R	–		R
Simulation $380_{\alpha R}$							
	17	–	L	L	L		–
2	12	L	L	L	R		R
3	7	R	R	L	L	–	–
4	6	L	L	L	–	L	R
5	4	–	L	R	R	R	L
6	4	L	L	L	–	R	–
7	4	L	L	R	L	–	L
8	4	L	L	L	–	–	–
9	4	L	L	L	L	R	L
10	3	R	R	L	R	–	R

Table 3. Experimentally derived proton–proton and proton–methyl NOE bounds, measured at two temperatures: 298 K and 340 K. Protons of the Aib methyl groups are indicated by the symbols CB1 (pro-S) and CB2 (pro-R). The bounds involving the CB methyl groups include a pseudo-atom bound correction of 0.03 nm, because the geometric mean of the positions of the three methyl hydrogens is used in the NOE distance calculation (21).

			NOE upper bounds at	
NOE sequence number	Atom 1	Atom 2	298 K (nm)	340 K (nm)
1	1HN	2HN	0.351	0.326
2	1CB1	1HN	0.365	
3	1CB2	1HN	0.378	0.415
4	1CB1	2HN	0.404	0.444
5	1CB2	2HN	0.455	0.512
6	1CB1	3HN	0.496	
7	1CB2	3HN	0.502	0.600
8	1CB2	4HN	0.486	
9	2HN	3HN	0.337	0.300
10	2CB1	2HN	0.367	0.406
11	2CB2	2HN	0.405	0.441
12	2CB2	3HN	0.481	0.515
13	2CB2	4HN	0.519	0.666
14	2CB2	5HN	0.471	0.502
15	3CB2	3HN		0.454
16	3HN	4HN	0.299	0.279
17	3CB1	4HN	0.393	
18	3CB2	4HN	0.470	0.568
19	3CB2	6HN	0.461	
20	4HN	5HN	0.286	0.280
21	4CB1	4HN	0.350	0.388
22	4CB2	5HN	0.486	
23	4CB2	6HN	0.492	
24	5CB2	5HN		0.490
25	5HN	6HN	0.281	0.259
26	5CB1	6HN	0.402	
27	5CB2	6HN		0.572
28	5CB1	6HA	0.492	0.574
29	5CB2	6HA		0.657
30	5CB2	7HN	0.477	0.546
31	5CB2	8HN	0.486	0.564
32	6HN	7HN	0.261	0.266
33	6HA	7HN	0.297	0.298
34	6HA	8HN	0.349	0.392
35	7HN	8HN	0.288	0.289
36	7CB1	7HN		0.392
37	7CB2	7HN		0.453
38	7CB1	8HN	0.414	
39	7CB2	8HN		0.582
40	8CB1	8HN	0.380	0.434
41	8CB2	8HN	0.406	0.445

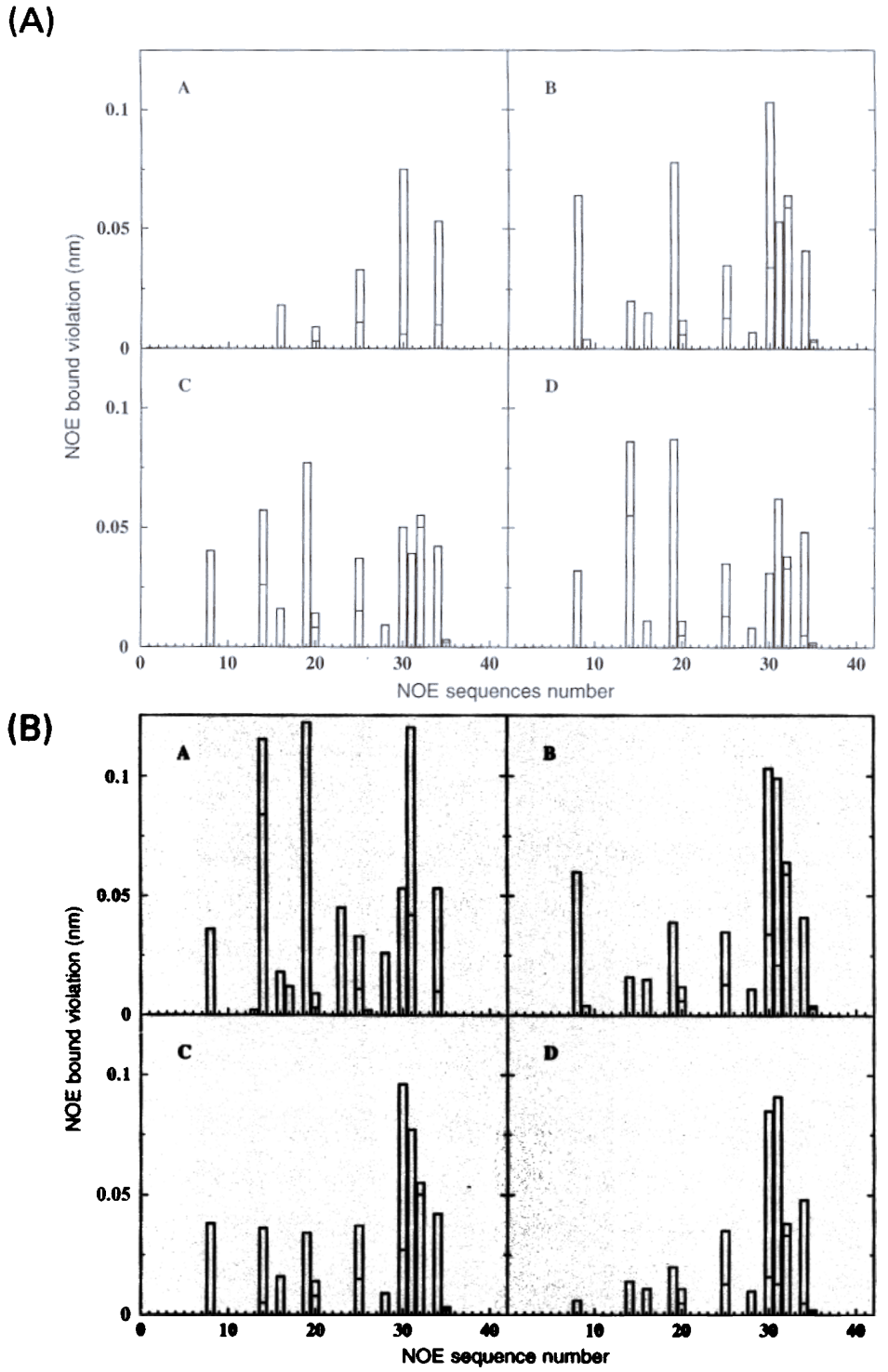


Figure 9. Violations of the experimentally derived NOE distance bounds (see Table 3) at 298 K (blue; 34 NOEs) and 340 K (red; 32 NOEs) for the simulations $298_{\alpha R}$ (A), 298_{ϵ} (B), 340_{ϵ} (C), and $380_{\alpha R}$ (D). The NOE sequence numbers for Fig. 9A are defined in Table 3. In Fig. 9B, the assignments for the pro-S and pro-R methyls have been switched in comparison with Table 3. At 298 K, no bounds are available for NOEs with sequence numbers 15, 24, 27, 29, 36, 37, and 39. At 340 K, no bounds are available for NOEs with sequence numbers 2, 6, 8, 17, 19, 22, 23, 26, and 38.

5CB2–7HN (no. 30), and 5CB2–8HN (no. 31) are moderately violated at 298 K (Fig. 9 A, B panels A, B, blue lines) and not violated at all at 340 K (Fig. 9 A, B panel C, red lines) although the carboxy terminus of the peptide is highly flexible and only adopts helical forms for short times.

Regarding the averaging involved in the NMR experiment, it is of interest to compute the NOE violations of single structures. Figure 10A shows the violations of the NOE bounds as defined in Table 3, for several conformations displayed in Fig. 6 taken from the simulations. Figure 10B

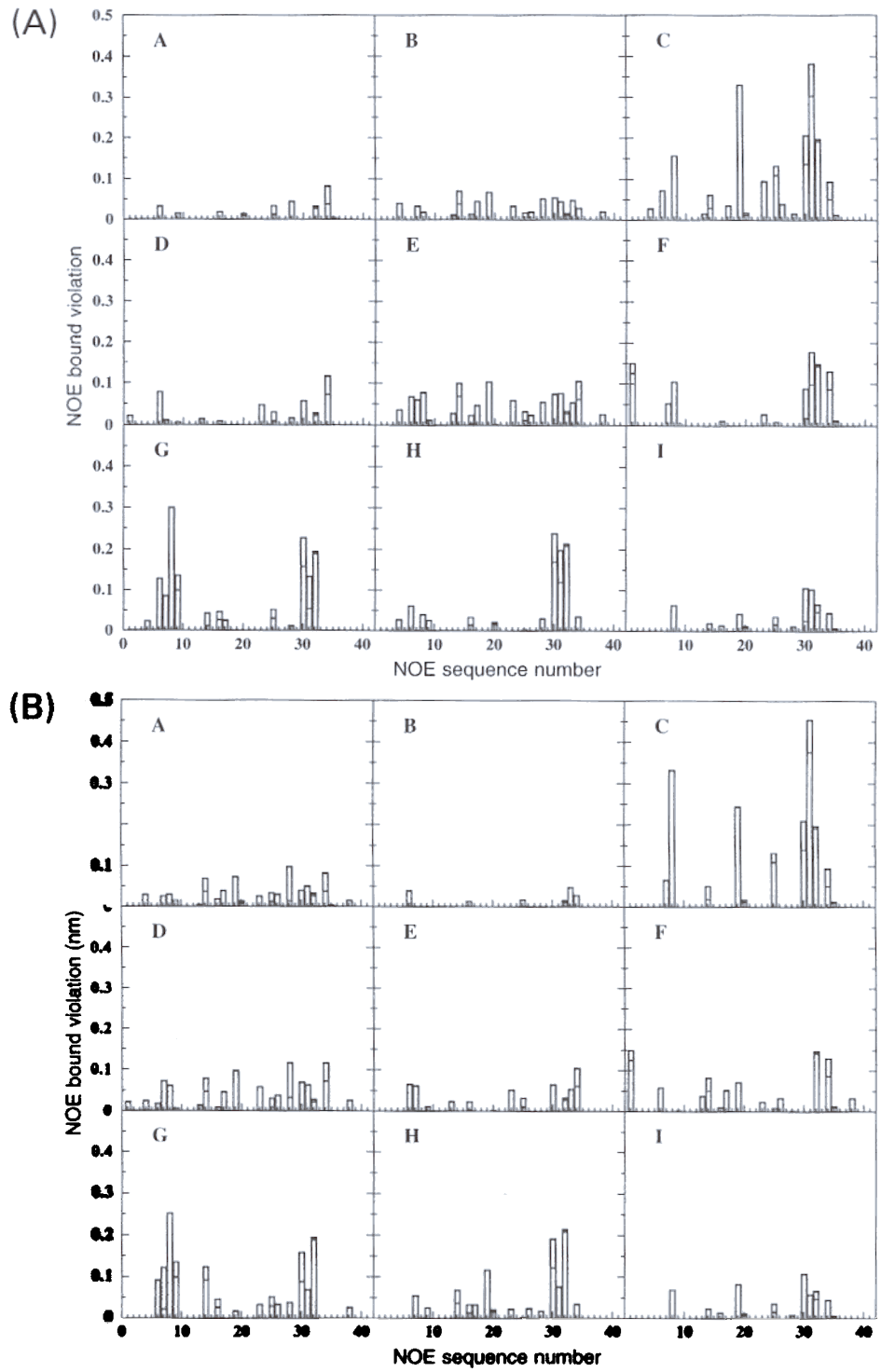


Figure 10. Violations of the experimental NOE distance bounds at 298 K (blue) and 340 K (red) of selected structures (see also Fig. 6): left-handed 3_{10} helix (A); right-handed 3_{10}-helix (B); left-handed α-helix (D); right-handed α-helix (E); structure with maximum radius of gyration in simulation 298_e (C); structure with minimum radius of gyration in simulation 298_e (F); central member of the most populated cluster in simulation 298_e (G); central member of the most populated cluster in simulation 340_e (H); and the violation of the $\langle r^{-6}\rangle^{-1/6}$ averages for simulation 298_e compared with the data at 298 K (blue) and for simulation 340_e compared with the data at 340 K (red) (I). The NOE sequence numbers for Fig. 10A are defined in Table 3. In Fig. 10B, the assignments for the pro-S and pro-R methyls have been switched in comparison with Table 3.

shows the violations of the same structures to the NOE bounds with the opposite assignment of NOE peaks to the chirally indistinguishable methyl groups CB1 and CB2 (CB1 and CB2 interchanged in Table 3). In Fig. 10A, we can see that the right-handed 3_{10}-model helix (Fig. 10A panel B) satisfies the experimental NOE distance bounds quite well, whereas all the other model helices (Fig. 10A panels A, D, E) show violations greater than 0.1 nm for some of the NOE bounds. Not unexpectedly, the most extended conformation observed (Fig. 6C) shows the largest violations (Fig. 10A panel C). Although the simulation at 340 K satisfies the NOE bounds well (red in Fig. 10A panel I), the ensemble contains a considerable number of left-handed helical structures. Furthermore, it is interesting to note that the structures of the most populated clusters taken from simulations 298_e and 340_e (Fig. 10A panels G, H) do not satisfy the experimental NOE bounds very well. Yet the average NOE violations of these simulations are quite low (Fig. 10A panel I). This reflects the nonlinearity of the r^{-6} averaging involved in the NMR experiment.

In Fig. 10B, we make similar observations to those made in Fig. 10A. The model structure satisfying the NOE bounds best for the opposite assignment of the Aib methyl groups is obviously the left-handed 3_{10}-helix (Fig. 10B panel A). Unlike in Fig. 10A, however, the opposite handedness of the 3_{10}-model helix (Fig. 10B panel B) induces no violations greater than 0.1 nm. Furthermore, it is interesting to note that none of the model structures satisfies the NOE bounds at 340 K (Fig. 10B panels A, B, D, E, red lines) as well as the ensemble average of the 340_e simulation (Fig. 10B panel I, red lines).

Conclusions

For all four simulations of the octapeptide in DMSO, a variety of secondary structures is observed. The right-handed α-helical structure is found to be stable for an extended time period at room temperature, but is not necessarily thermodynamically the most stable structure. We could also observe folding into right-handed 3_{10}- and α-helical structures from an arbitrary (extended) conformation at room temperature. At 340 K and 380 K, folding into both left- and right-handed 3_{10}- and α-helical structures is observed. In all the simulations except for $298_{\alpha R}$, which is biased by its R-α-helical starting structure, the 3_{10}-helical structure is predominant, which is in agreement with the NMR data. Most often, however, only the N-terminal part of the peptide (residues 2–5) is folded, whereas residues 6–8 are highly mobile. This is in agreement with X-ray crystallographic studies of Aib-containing peptides, which indicate the occurrence of both R- and L-helical conformations at the C-terminus of these peptides. The apparent folding/unfolding transition rates depend on the temperature: from 0.03–0.1 ns^{-1} at 298 K, 0.1–0.3 ns^{-1} at 340 K, to 0.3–1 ns^{-1} at 380 K.

The experimental NOE distance bounds are well-satisfied in the simulations; all but two violations at room temperature can be explained by the lack of sampling owing to the slow motions at this temperature. At 340 K, the simulated proton–proton and proton–methyl distances match the experimental NOE bounds within experimental error. Comparing the alternative assignments of the NOE bounds to the chirally indistinguishable methyl groups that match right- and left-handed 3_{10}-helices, respectively, suggests a preference for the left-handed over the right-handed structures for all Aib residues, with the Leu being the only residue to prefer right-handed structures.

The results show that MD simulations of peptides using empirical force-fields can be used to investigate the different conformational states accessible to short peptides in solution, provided the length of the simulations is much longer than the time-scale of the folding/unfolding transitions at the chosen temperature.

Acknowledgments: Financial support was obtained from the Schweizerischer Nationalfonds, project number 21-50929.97, which is gratefully acknowledged.

References

1. Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W.F. & Mark, A.E. (1999) Peptide folding: when simulation meets experiment. *Angew. Chemie Intl. Ed.* **38**, 236–240.
2. Daura, X., Jaun, B., Seebach, D., van Gunsteren, W.F. & Mark, A.E. (1998) Reversible peptide folding in solution by molecular dynamics simulation. *J. Mol. Biol.* **280**, 925–932.
3. Daura, X., van Gunsteren, W.F. & Mark, A.E. (1999) Folding–unfolding thermodynamics of a β-heptapeptide from equilibrium solutions. *Proteins* **34**, 269–280.
4. Daura, X., van Gunsteren, W.F., Rigo, D., Jaun, B. & Seebach, D. (1997) Studying the stability of a helical β-heptapeptide by molecular dynamics simulations. *Chemistry – A European Journal* **3**, 1410–1417.
5. Pande, V.S. & Rokhsar, D.S. (1999) Molecular dynamics simulations of unfolding and refolding of a β-hairpin fragment of protein G. *Proc. Natl. Acad. Sci. USA* **96**, 9062–9067.
6. Shirley, W.A. & Brooks, C.L. III (1997) Curious structure in 'canonical' alanine-based peptides. *Proteins: Structure, Function, Genetics* **28**, 59–71.

7. Takano, M., Yamato, T., Higo, J., Suyama, A. & Nagayama, K. (1999) Molecular dynamics of a 15-residue poly(L-alanine) in water: helix formation and energetics. *J. Am. Chem. Soc.* **121**, 605–612.
8. Schaefer, M., Bartels, C. & Karplus, M. (1998) Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit solvation model. *J. Mol. Biol.* **284**, 835–848.
9. Sung, S.-S. & Wu, X.-W. (1996) Molecular dynamics simulations of synthetic peptide folding. *Proteins: Structure, Function, Genetics* **25**, 202–214.
10. Muñoz, V., Thompson, P.A., Hofrichter, J. & Eaton, W.A. (1997) Folding dynamics and mechanism of β-hairpin formation. *Nature* **390**, 196–199.
11. Karle, I.L. & Balaram, P. (1990) Structural characteristics of α-helical peptide molecules containing Aib residues. *Biochemistry* **29**, 6747–6756.
12. Bellanda, M., Peggion, E., Bürgi, R., van Gunsteren, W. & Mammi, S. (2001) Conformational study of an Aib-rich peptide in DMSO by NMR. *J. Peptide Res.* **57**, 97–106.
13. Scott, W.R.P., Hünenberger, P.H., Tironi, I.G., Mark, A.E., Billeter, S.R., Fennen, J., Torda, A.E., Huber, T., Krüger, P. & van Gunsteren, W.F. (1999) The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **103**, 3596–3607.
14. van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P. & Tironi, I.G. (1996). *Biomolecular Simulation: The GROMOS96 Manual and User Guide.* Vdf Hochschulverlag AG an der ETH Zürich, Zürich.
15. Liu, H., Müller-Plathe, F. & van Gunsteren, W.F. (1995) A force field for liquid dimethyl sulfoxide and physical properties of liquid dimethyl sulfoxide calculated using molecular dynamics simulation. *J. Am. Chem. Soc.* **117**, 4363–4366.
16. Ryckaert, J.P., Ciccotti, G. & Berendsen, H.J.C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341.
17. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. & Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
18. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.
19. Bavoso, A., Benedetti, E., Di Blasio, B., Pavone, V., Pedone, C., Toniolo, C., Bonora, G.M., Formaggio, F. & Crisma, M. (1988) Long, chiral polypeptide 3_{10}-helices at atomic resolution. *J. Biomol. Struct. Dyn.* **5**, 803–817.
20. Chou, P.Y. & Fasman, G.D. (1978) Empirical predictions of protein conformation. *Annu. Rev. Biochem.* **47**, 251–276.
21. Koning, T.M.G., Boelens, R. & Kaptein, R. (1990) Calculation of the nuclear Overhauser effect and the determination of proton–proton distances in the presence of internal motions. *J. Magn. Reson.* **90**, 111–123.