β-Peptides with Different Secondary-Structure Preferences: How Different Are Their Conformational Spaces?

by **Riccardo Baron**[a]), **Dirk Bakowies**[a]), **Wilfred F. van Gunsteren**[a]), and **Xavier Daura***[b])

[a]) Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zurich, ETH Hönggerberg, CH-8093 Zürich

[b]) Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, E-08193 Bellaterra (e-mail: xavier.daura@uab.es)

The conformational spaces accessible to two β-hexapeptides in MeOH at 298 K and 340 K are investigated by molecular-dynamics simulation with an atomistic model of both solute and solvent. The structural properties of these peptides have been previously studied by NMR in MeOH at room temperature. The experimental data could be fitted to a model (*P*)-*12*/*10*-helix for one of the peptides and a model hairpin with a ten-membered H-bonded turn for the other. The goal of the present work is to determine whether the conformational spaces accessible to these two peptides of seemingly different conformational properties contain any common regions. In other words, to what extent are the evident differences found at the macroscopic level also present at the microscopic structural level? It is found that, for the two peptides studied, the conformational spaces sampled in the respective simulations show significant overlap.

Introduction. – The conformational space accessible to a peptide is defined by those regions of the 3*N*-dimensional configurational space – *N* being the number of atoms – that have a significant probability of being populated. The size and characteristics of this subspace are determined by the length and amino acid sequence of the peptide, as well as by the environment. For an adequate understanding of the macroscopic properties of a peptide, characterization of its accessible conformational space – rather than just of its most stable conformation – is required, since these properties can only be interpreted as weighted averages over the entire ensemble of accessible conformers. Such characterization is presently possible only by computer simulation and has been already used to complement experimental data for a number of peptides [1–6].

A question that remains unexplored is whether the accessible conformational spaces of peptides, which have the same number of residues but differ in their most stable fold, may overlap. This is investigated here for two β-hexapeptides in MeOH. β-Peptides have been a subject of intense research since the first structures were elucidated by the groups of *Seebach* and *Gellman* (see reviews [7–10]). Small β-peptides of as few as six amino acid residues fold into turns [11–16], helices [11][13][17–23], and sheet-like structures [12][15–17], analogous to the secondary structures of proteins. In addition, these compounds are resistant to degradation by most common peptidases and proteases [24][25], which makes them potentially suitable for various biomedical applications [26–31].

The structural formulae and NMR model structures of the two β-hexapeptides studied are shown in *Figs. 1* and *2*, respectively. In MeOH, peptide **1** adopts

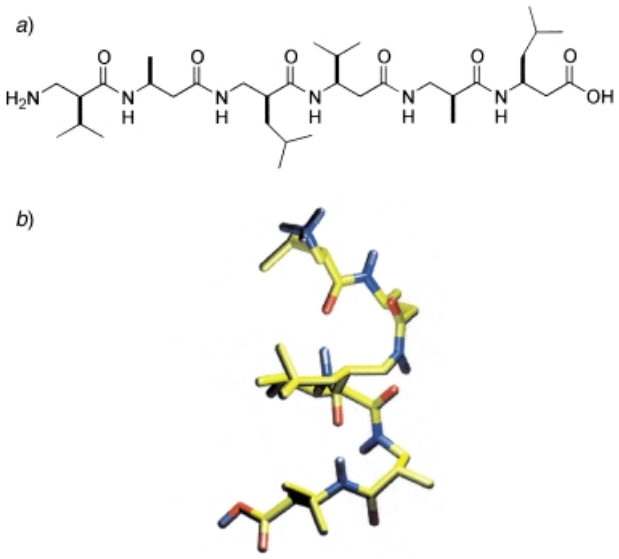

Fig. 1. a) *Structural formula and* b) *NMR model structure of peptide* **1**

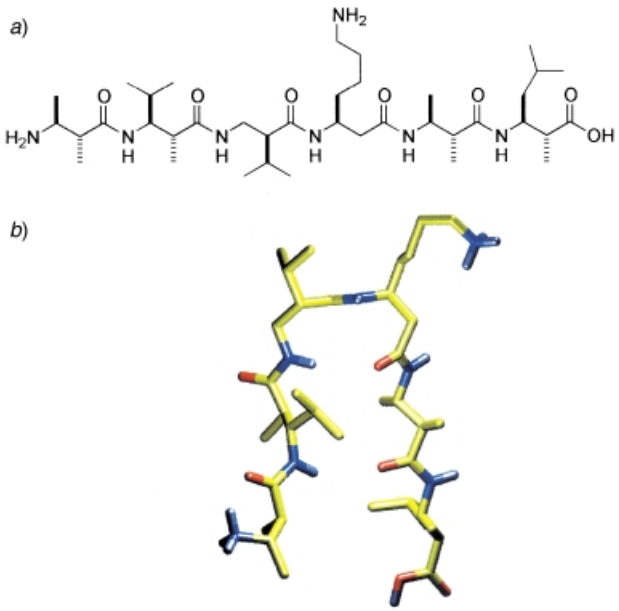

Fig. 2. a) *Structural formula and* b) *NMR model structure of peptide* **2**

predominantly a (*P*)-*12*/*10*-helical structure [2][11][13], while peptide **2** adopts a hairpin structure with a ten-membered H-bonded turn [16][32]. To sample a relevant part of the accessible conformational space of the two peptides in MeOH at two

temperatures, 298 K and 340 K, four 100-ns molecular-dynamics simulations were performed by means of an atomistic model for both solute and solvent. The structures sampled in the simulations were grouped into clusters representing the different backbone conformations adopted by the peptides. On the basis of these clusters, the conformational spaces sampled in the four simulations are compared. First, the degree of overlap between the conformational spaces sampled at 298 K and 340 K by the same peptide is analysed. Second, the question of whether there exists overlap between the conformational spaces sampled by the two peptides at the same temperature is addressed.

Results and Discussion. – The evolution of the structures of peptides **1** and **2** along the 100-ns molecular-dynamics simulations is given in *Figs. 3* and *4* in terms of atom-positional root-mean-square deviations (RMSD) from the model structures shown in *Figs. 1* and *2*, respectively. An RMSD cut-off of 0.08 nm (see *Exper. Part*) has been chosen to distinguish between structures that are representative of the experimentally determined conformer – for convenience referred to as the folded conformer – and other structures – referred to as unfolded conformers. *Figs. 3* and *4* serve to illustrate that, at either temperature, peptides **1** and **2** exist in a conformational equilibrium between the folded conformer – *12/10*-helix and hairpin, respectively – and a number of unfolded conformers.

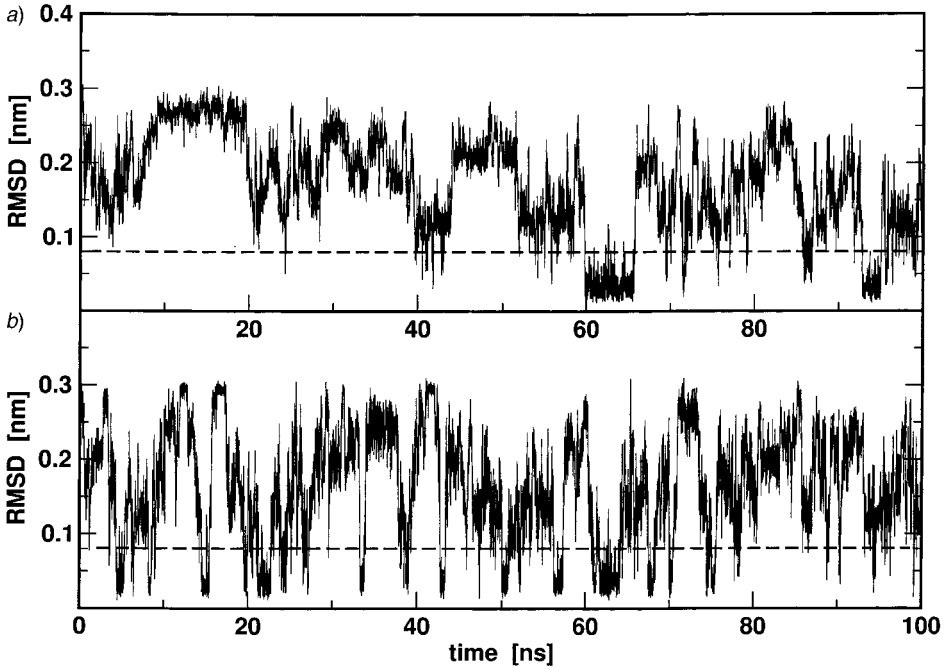

Fig. 3. *Backbone* (residues 2 to 5) *atom-positional root-mean-square deviation* (RMSD) *of peptide* **1** *from its NMR model structure as a function of simulation time. a*) Simulation at 298 K; *b*) simulation at 340 K. The dashed line at 0.08 nm serves to distinguish folded (*12/10*-helix) from unfolded (non-*12/10*-helix) structures.

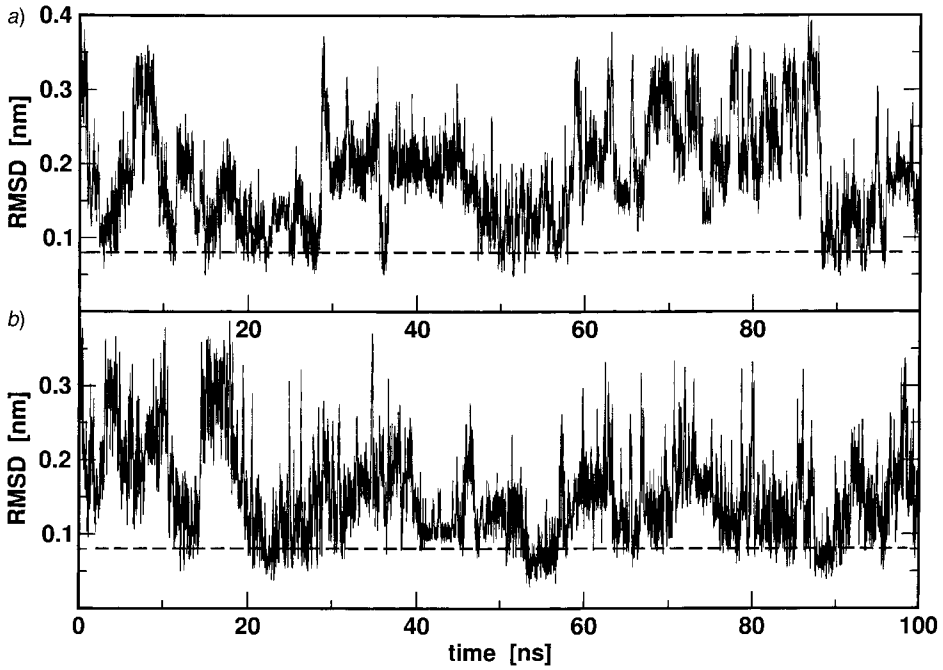

Fig. 4. *Backbone* (residues 2 to 5) *atom-positional root-mean-square deviation* (RMSD) *of peptide* **2** *from its NMR model structure as a function of simulation time. a*) Simulation at 298 K; *b*) simulation at 340 K. The dashed line at 0.08 nm serves to distinguish folded (hairpin) from unfolded (non-hairpin) structures.

To map the structures sampled during the simulations onto a set of generic conformations, a clustering analysis was performed on the individual molecular dynamics trajectories (see *Exper. Part*). The results from this analysis can be summarized as follows. In the simulation of peptide **1** at 298 K a total of 125 conformers (clusters) were visited within 100 ns. The first cluster including the (*P*)-*12*/*10*-helical pattern – clusters are ordered from largest (cluster 1) to smallest (cluster 125) number of elements – is cluster 2, which contains 21% of the ensemble of structures analysed (10^4). At 340 K, the same peptide adopted a total of 212 different conformations. The first cluster including the (*P*)-*12*/*10*-helical pattern is at this temperature cluster 1, which, nevertheless, is also representative for 21% of the ensemble of structures analysed. In the simulation of peptide **2** at 298 K a total of 196 conformers were visited within 100 ns. The first cluster including the 10-membered H-bonded turn corresponding to the hairpin is cluster 1, which contains 20% of the ensemble of structures analysed. At 340 K this peptide sampled 294 different conformers. The first cluster including the ten-membered H-bonded turn is also cluster 1, and is as well representative for 20% of the ensemble. For each individual peptide, not only the weight of the folded conformer is the same at 298 and 340 K, but the central-member structures of the corresponding clusters are also very similar. This is illustrated in *Figs. 5* and *6*, where the central-member structures of cluster 2 at 298 K and cluster 1 at 340 K of peptide **1**, and the central-member structures of cluster 1 at

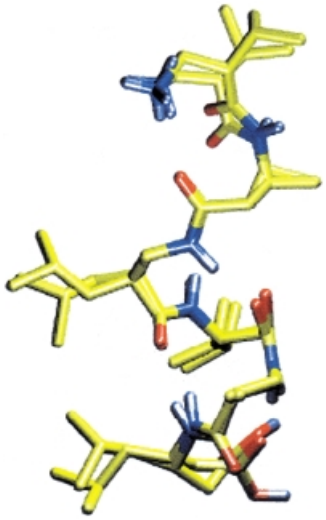

Fig. 5. *Central-member structures of cluster 2 at 298 K and cluster 1 at 340 K from the corresponding simulations of peptide* **1**. The two structures have been fitted with the backbone atoms of residues 2 to 5.

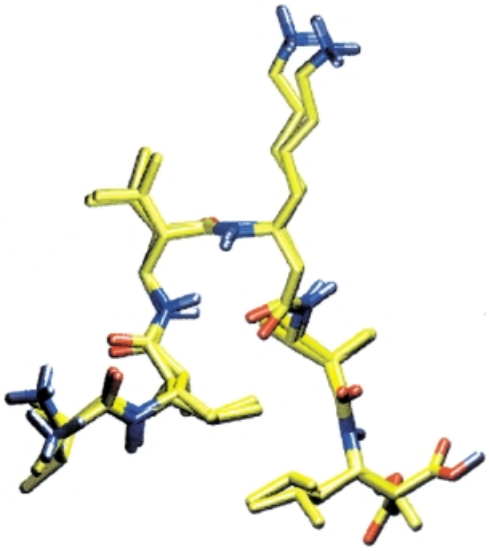

Fig. 6. *Central-member structures of cluster 1 at 298 K and cluster 1 at 340 K from the corresponding simulations of peptide* **2**. The two structures have been fitted with the backbone atoms of residues 2 to 5.

298 K and cluster 1 at 340 K of peptide **2**, respectively, are superimposed. In addition, in spite of the consistently larger number of clusters sampled at 340 K, for either peptide, the distribution of clusters with relatively high populations (*e.g.*, $>1\%$) is similar at both temperatures.

On the basis of this result, it is expected that, for a merged trajectory including the simulations at the two temperatures of any of the two peptides (see *Exper. Part*), the most populated cluster will also have an approximate weight of 20% and about equal contributions from the two trajectories/temperatures. This is confirmed in *Figs. 7* and *8*, which give the percentile populations of clusters 1 to 20 from such merged trajectories

and the contributions of the individual trajectories to each cluster. For peptide **1** (*Fig. 7*), the clustering analysis of the merged trajectory gives 249 clusters, with the first 20 clusters being representative for 85% of the total ensemble of $2 \cdot 10^4$ structures. For peptide **2** (*Fig. 8*), 365 clusters are counted, and the first 20 clusters account for 68% of the ensemble. Overall, for each of the peptides, the conformational spaces sampled at 298 and 340 K significantly overlap.

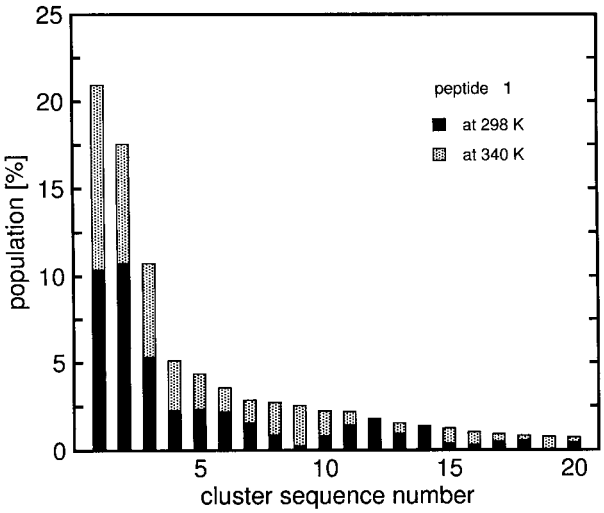

Fig. 7. *Population of clusters 1 to 20 from a merged trajectory including the two simulations of peptide* **1**. 10^4 Structures (at 10-ps intervals) from the simulation at 298 K (solid) plus 10^4 structures from the simulation at 340 K (dotted).

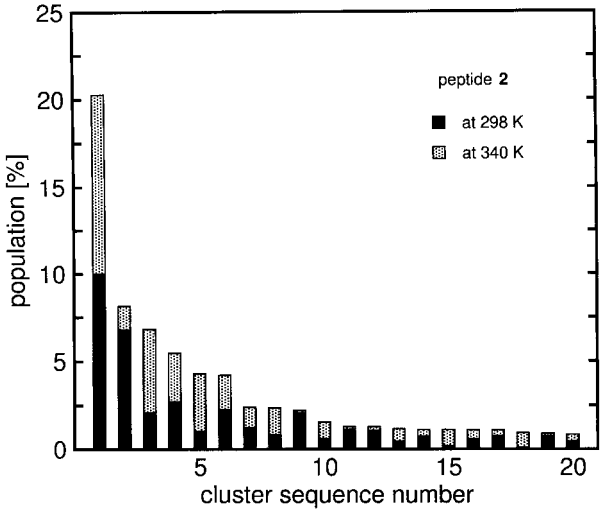

Fig. 8. *Population of clusters 1 to 20 from a merged trajectory including the two simulations of peptide* **2**. 10^4 Structures (at 10-ps intervals) from the simulation at 298 K (solid) plus 10^4 structures from the simulation at 340 K (dotted).

More interestingly, there is also conformational-space overlap in the simulations of peptides **1** and **2** at equal temperature (see *Exper. Part*). This can be observed in *Figs. 9* and *10*, which provide the percentile populations of clusters 1 to 20 from the merged trajectories at 298 and 340 K, respectively, as well as the contributions of each peptide to each cluster. The analysis of the merged trajectory with peptides **1** and **2** at 298 K gives rise to 277 clusters, with the first 20 clusters involving 74% of the total ensemble of $2 \cdot 10^4$ structures. At 340 K the total number of clusters is 407, and the first 20 clusters contain 69% of the ensemble. Of the first 20 clusters, only clusters 10, 13–15, and 17 at 298 K and cluster 20 at 340 K are contributed exclusively by one of the peptides. The latter corresponds to a completely folded (*P*)-*12*/*10*-helix, which, at 298 K, does not appear within the first 20 clusters. The other five clusters correspond to unstructured conformations with a rather large radius of gyration. A common trait of many of the clusters that appear in *Figs. 9* and *10* is that their members feature a central ten-membered H-bonded turn. This is so for clusters 1–3, 5–7, 11, and 16 at 298 K, and clusters 1–5, 7, 8, 11, 14–16 and 18–20 at 340 K. The ten-membered turn closed by the N(3)–H···O=C(4) (residue number in parenthesis) H-bond is actually present in both the *12*/*10* helix and the hairpin. That this turn is the only element of secondary structure present in a high percentage of the structures sampled by peptide **1** suggests that the N(3)–H···O=C(4) H-bond is the center of initiation of not only the hairpin structure of peptide **2** but also of the (*P*)-*12*/*10* helical structure of peptide **1**. In addition, cluster 2 corresponds, at either temperature, to a partially folded (*P*)-*12*/*10* helix in which the first two residues fall out of the helical pattern. In other words, there is a small population of peptide **2** adopting the helical fold. *Figs. 11* and *12* show the central-member structures of cluster 1 at 298 and 340 K, and the central member structures of cluster 2 at 298 and 340 K, respectively.

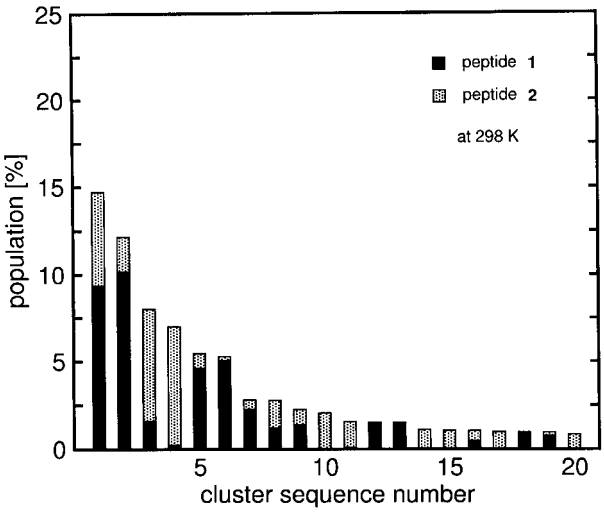

Fig. 9. *Population of clusters 1 to 20 from a merged trajectory including the simulations of both peptides at 298 K.* 10^4 Structures (at 10-ps intervals) from the simulation of peptide **1** (solid) plus 10^4 structures from the simulation of peptide **2** (dotted).

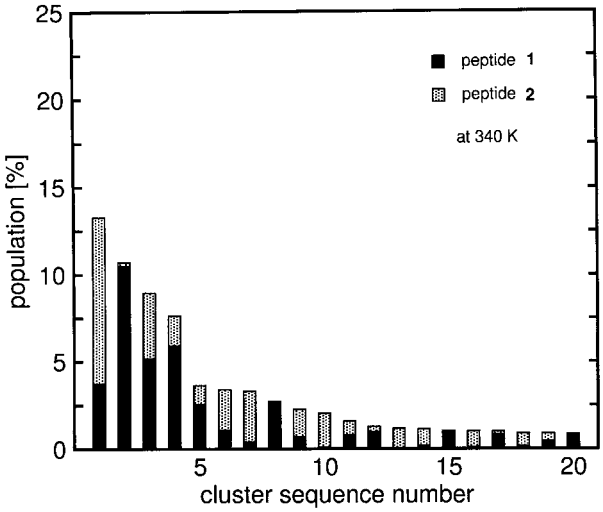

Fig. 10. *Populations of clusters 1 to 20 from a merged trajectory including the simulations of both peptides at 340 K*. 10^4 Structures (at 10-ps intervals) from the simulation of peptide **1** (solid) plus 10^4 structures from the simulation of peptide **2** (dotted).

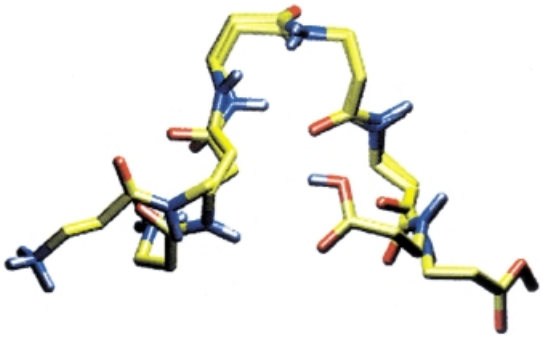

Fig. 11. *Central-member structures of cluster 1 at 298 K and cluster 1 at 340 K from the merged trajectories analysed in* Figs. 9 *and* 10. The two structures have been fitted with the backbone atoms of residues 2 to 5. Only backbone atoms are shown.

Conclusions. – The structures sampled by two β-hexapeptides in MeOH at 298 K and 340 K, in a total of four 100-ns molecular-dynamics simulations, have been mapped to a set of generic backbone conformations that have been used to compare the conformational spaces accessible to the two peptides at the two temperatures. The most stable structure of peptide **1** is a (*P*)-*12*/*10*-helix, while that of peptide **2** is a hairpin with a ten-membered H-bonded turn. For either peptide, the conformational spaces sampled at 298 and 340 K are very similar. Nevertheless, the relative populations of the different conformers are, in general, temperature dependent. More strikingly, a significant part of the conformational spaces sampled by peptides **1** and **2** at a given temperature do also overlap. In particular, the ten-membered turn formed by the N(3)–H$\cdots$O=C(4) H-bond is present in many of the conformers sampled by the two peptides and seems to be the structural feature of initiation of both the (*P*)-*12*/*10*-helical structure and the hairpin structure. Furthermore, the secondary structure motif

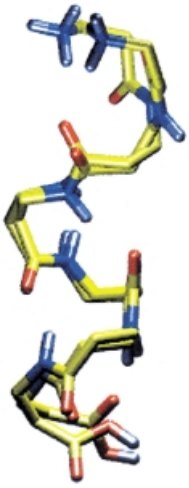

Fig. 12. *Central-member structures of cluster 2 at 298 K and cluster 2 at 340 K from the merged trajectories analysed in* Figs. 9 *and* 10. The two structures have been fitted with the backbone atoms of residues 2 to 5. Only backbone atoms are shown.

adopted preferentially by peptide **1** is also accessible to peptide **2** and *vice versa*. The differences in the relative weights of these shared conformers in the respective ensembles make, nevertheless, the macroscopic structural properties of these two peptides markedly different. Before these conclusions can be generalized, the accessible conformational spaces of more peptides should be studied, especially of peptides that do not share any element of secondary structure in their experimentally observed model structures.

Experimental Part

The simulations and analyses were carried out by means of GROMOS96 [33][34].

β-Hexapeptide **1**. The molecular models for peptide and solvent (MeOH) have been defined previously [2][34]. The initial coordinates of the peptide were taken from an extended structure (all backbone torsional dihedral angles in *trans*). The two ionizable groups (N-terminus and C-terminus) were modelled in their protonated states. The peptide was placed at the center of a periodic truncated-octahedron box, with dimensions such that the minimum distance from any peptide atom to any of the walls was 1.4 nm. The solvent was introduced into the box by replicating a cubic configuration of 216 equilibrated MeOH molecules. All MeOH molecules with the O-atom lying within 0.3 nm of a non-H-atom of the peptide were removed. Thus, the final system consisted of 56 peptide atoms and 4305 solvent atoms (1435 MeOH molecules). Truncated-octahedron periodic boundary conditions were applied from this point onwards. A steepest-descent energy minimization of the system was performed in order to relax the solvent configuration, while the peptide atoms were positionally restrained by a harmonic interaction with a force constant of 250 kJ mol^{-1} nm^{-2}. Next, a steepest-descent energy minimization of the system without restraints was performed to eliminate any residual strain. The energy minimizations were terminated when the energy change per step became smaller than 0.1 kJ mol^{-1}.

Two 100-ns molecular-dynamics simulations at 298 K and 340 K, respectively, and 1 atm were performed. The initial velocities of the atoms were taken from a *Maxwell-Boltzmann* distribution at 100 K. The temp. was brought and maintained at the desired value by means of weak coupling to an external temperature bath [35]. The temperature of the solute and the solvent were independently coupled to the bath with a relaxation time of 0.1 ps. The pressure of the system (calculated *via* a molecular virial) was maintained at the desired value by weak coupling to an external pressure bath [35], with isotropic scaling and a relaxation time of 0.5 ps. A value of $4.575 \cdot 10^{-4}$ kJ^{-1} mol nm^3 was chosen for the isothermal compressibility of the system at the two temperatures. Bond lengths were constrained to ideal values [34] by means of the SHAKE algorithm [36] with a geometric

tolerance of 10^{-4}. In the leap-frog integration scheme a time step of 2 fs was used. The nonbonded interactions were evaluated with a twin-range cut-off of 0.8/1.4 nm and a charge-group pair-list that was updated every 5 time steps [34]. The cut-off radii were applied to the centres of geometry of the solute charge groups and to the O-atoms of the solvent molecules.

β-Hexapeptide **2**. The molecular models for peptide and solvent (MeOH) have also been defined in previous reports [32][34]. The initial coordinates of the peptide were taken from an extended structure. The three ionizable groups (N-terminus, homo-lysine-amino group, and C-terminus) were modelled in their protonated state. The protocol to initialize the simulation described for peptide **1** applies also to peptide **2** and has been described in detail in [32]. The system consisted of 64 peptide atoms and 4359 solvent atoms (1453 MeOH molecules) in a truncated-octahedron periodic box.

Analysis. Trajectory coordinates were stored at 0.5-ps intervals and used for analysis. Least-squares translational and rotational fitting of atomic coordinates for the calculation of atom-positional root-mean-square differences (RMSD) was based on the backbone atoms (N, C^{β}, C^{α}, C) of all but the N- and C-terminal residues of the β-hexapeptides. A conformational clustering analysis was performed on sets of peptide structures taken at 10-ps intervals from the simulations, with the backbone atom-positional RMSD as similarity criterion. The clustering algorithm has been described in previous studies of β-peptide dynamics [37]. A similarity cut-off or maximum cluster radius of 0.08 nm was chosen, which corresponds approximately to the maximum atom-positional RMSD between structures of the set of structures given as model structures satisfying the NMR data. The clustering was performed for four sets of 10^4 structures belonging to the individual 100 ns trajectories as well as for four sets of $2 \cdot 10^4$ structures belonging to the following merged trajectories: *a*) 100-ns trajectory of peptide **1** at 298 K and 100-ns trajectory of peptide **1** at 340 K (*Fig. 7*), *b*) 100-ns trajectory of peptide **2** at 298 K and 100-ns trajectory of peptide **2** at 340 K (*Fig. 8*), *c*) 100-ns trajectory of peptide **1** at 298 K and 100-ns trajectory of peptide **2** at 298 K (*Fig. 9*), and *d*) 100-ns trajectory of peptide **1** at 340 K and 100 ns trajectory of peptide **2** at 340 K (*Fig. 10*).

The authors warmly thank Prof. *Dieter Seebach* for providing the experimental data on the β-peptides and stimulating their study by simulation methods.

REFERENCES

[1] X. Daura, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *J. Mol. Biol.* **1998**, *280*, 925.
[2] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem., Int. Ed.* **1999**, *38*, 236.
[3] D. Seebach, J. V. Schreiber, S. Abele, X. Daura, W. F. van Gunsteren, *Helv. Chim. Acta* **2000**, *83*, 34.
[4] C. Peter, X. Daura, W. F. van Gunsteren, *J. Biomol. NMR* **2001**, *20*, 297.
[5] G. Colombo, D. Roccatano, A. E. Mark, *Proteins: Struct., Funct., Genet.* **2002**, *46*, 380.
[6] A. Glättli, X. Daura, D. Seebach, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2002**, *124*, 12972.
[7] S. Borman, *Chem. Eng. News* **1997**, *75*, 32.
[8] B. L. Iverson, *Nature* **1997**, *385*, 113.
[9] S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173.
[10] K. Gademann, T. Hintermann, J. V. Schreiber, *Curr. Med. Chem.* **1999**, *6*, 905.
[11] D. Seebach, K. Gademann, J. V. Schreiber, J. L. Matthews, T. Hintermann, B. Jaun, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1997**, *80*, 2033.
[12] S. Krauthäuser, L. A. Christianson, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1997**, *119*, 11719.
[13] D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L. Matthews, J. V. Schreiber, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1998**, *81*, 932.
[14] D. Seebach, S. Abele, T. Sifferlen, M. Hänggi, S. Gruner, P. Seiler, *Helv. Chim. Acta* **1998**, *81*, 2218.
[15] Y. J. Chung, L. A. Christianson, H. E. Stanger, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, *120*, 10555.
[16] D. Seebach, S. Abele, K. Gademann, B. Jaun, *Angew. Chem., Int. Ed.* **1999**, *38*, 1595.
[17] D. Seebach, M. Overhand, F. N. M. Kühnle, B. Martinoni, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 913.
[18] D. Seebach, P. E. Ciceri, M. Overhand, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amstutz, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 2043.
[19] D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 13071.

[20] D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. Huang, J. J. Barchi, S. H. Gellman, *Nature* **1997**, *387*, 381.
[21] S. Abele, G. Guichard, D. Seebach, *Helv. Chim. Acta* **1998**, *81*, 2141.
[22] D. H. Appella, J. J. Barchi Jr., S. R. Durell, S. H. Gellman, *J. Am. Chem. Soc.* **1999**, *121*, 2309.
[23] B. W. Gung, D. Zou, A. M. Stalcup, C. E. Cottrell, *J. Org. Chem.* **1999**, *64*, 2176.
[24] D. Seebach, S. Abele, J. V. Schreiber, B. Martinoni, A. K. Nussbaum, H. Schild, H. Schulz, H. Hennecke, R. Woessner, F. Bitsch, *Chimia* **1998**, *52*, 734.
[25] T. Hintermann, D. Seebach, *Chimia* **1997**, *51*, 244.
[26] K. Gademann, M. Ernst, D. Hoyer, D. Seebach, *Angew. Chem., Int. Ed.* **1999**, *38*, 1223.
[27] Y. Hamuro, J. P. Schneider, W. F. DeGrado, *J. Am. Chem. Soc.* **1999**, *121*, 12200.
[28] M. Werder, H. Hauser, S. Abele, D. Seebach, *Helv. Chim. Acta* **1999**, *82*, 1774.
[29] E. A. Porter, X. Wang, H.-S. Lee, B. Weisblum, S. H. Gellman, *Nature* **2000**, *404*, 565.
[30] E. A. Porter, X. Wang, H.-S. Lee, B. Weisblum, S. H. Gellman, *Nature* **2000**, *405*, 298.
[31] K. Gademann, M. Ernst, D. Seebach, D. Hoyer, *Helv. Chim. Acta* **2000**, *83*, 16.
[32] X. Daura, K. Gademann, H. Schäfer, B. Jaun, D. Seebach, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2001**, *123*, 2393.
[33] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, W. F. van Gunsteren, *J. Phys. Chem. A* **1999**, *103*, 3596.
[34] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, 'Biomolecular Simulation: The GROMOS96 Manual and User Guide', vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v., Zürich, Groningen, 1996.
[35] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
[36] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
[37] X. Daura, W. F. van Gunsteren, A. E. Mark, *Proteins: Struct., Funct., Genet.* **1999**, *34*, 269.

Received June 20, 2002