The β-Peptide Hairpin in Solution: Conformational Study of a β-Hexapeptide in Methanol by NMR Spectroscopy and MD Simulation

Xavier Daura,[‡] Karl Gademann,[§,†] Heiko Schäfer,[‡] Bernhard Jaun,[§] Dieter Seebach,*[,§] and Wilfred F. van Gunsteren*[,‡]

Contribution from the Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zurich, ETH-Zentrum, CH-8092 Zurich, Switzerland, and Laboratorium für Organische Chemie, Eidgenössische Technische Hochschule Zürich, ETH-Zentrum, CH-8092 Zürich, Switzerland

Received October 16, 2000. Revised Manuscript Received January 4, 2001

Abstract: The structural and thermodynamic properties of a 6-residue β-peptide that was designed to form a hairpin conformation have been studied by NMR spectroscopy and MD simulation in methanol solution. The predicted hairpin would be characterized by a 10-membered hydrogen-bonded turn involving residues 3 and 4, and two extended antiparallel strands. The interproton distances and backbone torsional dihedral angles derived from the NMR experiments at room temperature are in general terms compatible with the hairpin conformation. Two trajectories of system configurations from 100-ns molecular-dynamics simulations of the peptide in solution at 298 and 340 K have been analyzed. In both simulations reversible folding to the hairpin conformation is observed. Interestingly, there is a significant conformational overlap between the unfolded state of the peptide at each of the temperatures. As already observed in previous studies of peptide folding, the unfolded state is composed of a (relatively) small number of predominant conformers and in this case lacks any type of secondary-structure element. The trajectories provide an excellent ground for the interpretation of the NMR-derived data in terms of ensemble averages and distributions as opposed to single-conformation interpretations. From this perspective, a relative population of the hairpin conformation of 20% to 30% would suffice to explain the NMR-derived data. Surprisingly, however, the ensemble of structures from the simulation at 340 K reproduces more accurately the NMR-derived data than the ensemble from the simulation at 298 K, a question that needs further investigation.

1. Introduction

The unceasing extension of the computationally accessible simulation time scales and improvement of force fields has positioned molecular dynamics (MD) simulation techniques among the basic tools in the study of peptide folding. Previous work from our group[1−5] and from others[6] has shown that the reversible folding of a variety of peptides in various solvents into distinct native folds can be accurately simulated at the atomic level by using a general purpose force field and Newton's equations of motion. The ability to sample the relevant parts of the configurational space of the system as a function of temperature, with a correct distribution of probability densities for the different conformations of the peptide, opens the way to the calculation of a number of properties that remained, until very recently, inaccessible to simulation methods or that could be only calculated by means of severe approximations. Such is the case, for example, of absolute entropies[7] and relative entropies and free energies of different conformational macrostates, e.g., folded and unfolded.[8,9] Knowledge of the relative free energies of different conformational microstates as well as of their enthalpic and entropic components is fundamental to the understanding of the mechanisms of peptide and protein folding. Experimentally these quantities are very difficult to estimate, if not impossible, and in all cases some degree of modeling is required.[10,11] Similarly, the experimental conformational characterization of a peptide in solution is hampered by the mix of states occurring in the volume of the sample on the time scale of the measurement. The interpretation of NOEs and ^3J-coupling constants from NMR experiments in terms of interproton distances and torsional dihedral angles, respectively,

* Address correspondence on the experimental part to Seebach (e-mail: seebach@org.chem.ethz.ch) and on the simulation part to van Gunsteren (e-mail: wfvgn@igc.phys.chem.ethz.ch).

‡ Swiss Federal Institute of Technology Zurich.

§ Eidgenössische Technische Hochschule Zürich.

† Part of the dissertation of K.G., ETH Zurich, No. 13556, 2000. Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138.

(1) Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *J. Mol. Biol.* **1998**, *280*, 925−932.

(2) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem. Int. Ed.* **1999**, *38*, 236−240.

(3) Peter, C.; Daura, X.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **2000**, *122*, 7461−7466.

(4) Bürgi, R.; Daura, X.; Mark, A. E.; Bellanda, M.; Mammi, S.; Peggion, E.; van Gunsteren, W. F. *J. Peptide Res.* In press.

(5) van Gunsteren, W. F.; Bürgi, R.; Peter, C.; Daura, X. *Angew. Chem. Int. Ed.* **2001**, *40*, 351−355.

(6) Takano, M.; Yamato, T.; Higo, J.; Suyama, A.; Nagayama, K. *J. Am. Chem. Soc.* **1999**, *121*, 605−612.

(7) Schäfer, H.; Mark, A. E.; van Gunsteren, W. F. *J. Chem. Phys.* **2000**, *113*, 7809−7817.

(8) Schäfer, H.; Daura, X.; Mark, A. E.; van Gunsteren, W. F. *Proteins: Struct. Funct. Genet.* **2001**, *43*, 45−56.

(9) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins: Struct. Funct. Genet.* **1999**, *34*, 269−280.

(10) DAquino, J. A.; Gómez, J.; Hilser, V. J.; Lee, K. H.; Amzel, L. M.; Freire, E. *Proteins: Struct. Funct. Genet.* **1996**, *25*, 143−156.

(11) Alexandrescu, A. T.; Rathgeb-Szabo, K.; Rumpel, K.; Jahnke, W.; Schulthess, T.; Kammerer, R. A. *Protein Sci.* **1998**, *7*, 389−402.

10.1021/ja003689g CCC: $20.00
Published on Web 02/16/2001

is often complex due to the remarkable flexibility of even secondary-structure-forming peptides and the nonlinearly averaging nature of the NOE experiments.[12−15] The common approach to structure determination relies on the assumption that all NOE signals originate from the same predominant conformer, even when the presence of other conformers is explicitly acknowledged and a percentile population is estimated for the most stable fold. It is frequently assumed that the unfolded state is a conformationally random, unstructured state. Molecular dynamics simulations have suggested that this assumption is not valid, i.e., the unfolded state is neither random nor unstructured,[3−5,9] and that its use can easily lead to misinterpretation of the NMR data, e.g., fitting derived interproton distances originating from two different conformations into one single structure.[2]

While MD simulations have proven successful at reproducing the folding of a number of peptides and have provided new insights on some of its principles, experimental validation is at present still required. For the best understanding of the conformational behavior of a peptide in solution a combined analysis of NMR and MD data is required. Our first computational studies of peptide folding were motivated by the surprising folding capabilities of a β-heptapeptide (seven-residue peptide composed of β-amino acids), which was shown by NMR spectroscopy to fold into a remarkably stable left-handed 3_{14}-helix (3 residues per turn forming 14-membered hydrogen-bonded rings) in methanol solution.[16] Succeeding MD simulations of this peptide at a range of temperatures showed the ability of the computational method to reproduce the experimentally derived data.[1,17] The sensitivity of the calculations to the amino acid sequence was tested by performing MD simulations of a β-hexapeptide containing the same types of side chains (from alanine, valine, and leucine) but in different position. The NMR data collected for this peptide in methanol solution suggested the presence of a new type of helix, a right-handed 12/10 helix (alternating 12-membered and 10-membered hydrogen-bonded rings).[18,19] The 12/10-helical fold is very peculiar in that it has a low backbone dipole moment, as a result of the N−H (and C═O) dipole vectors of consecutive residues pointing in opposite directions. As with the β-heptapeptide, the simulations reproduced the experimentally derived fold with great accuracy and helped interpret some ambiguities present in the NOE data.[2] Now, we extend this series with a combined NMR-MD study of a β-peptide that folds into yet another type of secondary structure, a hairpin-like structure.

β-Peptides have attracted much attention since the first structures were elucidated.[20,21] Small β-peptides of as few as six amino acid residues fold into turns,[18,19,22−25] helices,[16,18,19,26−31] and sheetlike structures,[22−24,26] analogous to the secondary structures of proteins. In addition, these compounds are resistant to degradation by most common peptidases and proteases.[32,33] These two facts suggest the possibility of using β-peptides as nondegradable peptidomimetics for pharmaceutical applications. This possibility was tested by one of our groups and, after the first biologically active β-peptide mimicking a natural α-peptidic hormone was reported,[34] several examples of bioactive β-peptides were described by us[35,36] and others.[37−39]

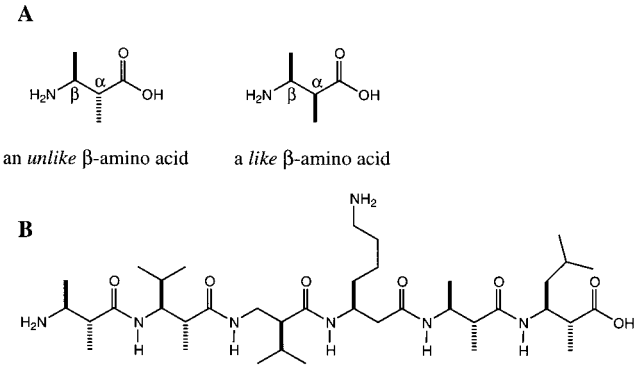

Figure 1. (A) Example of *like* and *unlike* β-amino acids. (B) Structural formula of the β-hexapeptide studied. In the simulations the N- and C-terminal groups were protonated.

The properties of a β-amino acid are not only determined by the nature of the side chain and the chirality of the corresponding backbone center but also by the position of the side chain, i.e., on carbon 2 (or C^{α}), on carbon 3 (or C^{β}), or on both backbone carbon atoms (see Figure 1A), providing an extra degree of freedom for structural design. Seebach and co-workers have predicted that β-peptides consisting of (*R*,*S*)-$\beta^{2,3}$ or *unlike*-$\beta^{2,3}$ residues (see Figure 1A), where the superscripts indicate the positions of the side chains in the β-amino acid, will predomi-

(12) Tropp, J. *J. Chem. Phys.* **1980**, *72*, 6035−6043.
(13) Brüschweiler, R.; Case, D. A. *Prog. Nucl. Magn. Reson. Sp.* **1994**, *26*, 27−58.
(14) Philippopoulos, M.; Lim, C. *J. Phys. Chem.* **1994**, *98*, 8264−8273.
(15) Schneider, T. R.; Brünger, A. T.; Nilges, M. *J. Mol. Biol.* **1999**, *285*, 727−740.
(16) Seebach, D.; Ciceri, P. E.; Overhand, M.; Jaun, B.; Rigo, D.; Oberer, L.; Hommel, U.; Amstutz, R.; Widmer, H. *Helv. Chim. Acta* **1996**, *79*, 2043−2066.
(17) Daura, X.; van Gunsteren, W. F.; Rigo, D.; Jaun, B.; Seebach, D. *Chem.-Eur. J.* **1997**, *3*, 1410−1417.
(18) Seebach, D.; Gademann, K.; Schreiber, J. V.; Matthews, J. L.; Hintermann, T.; Jaun, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1997**, *80*, 2033−2038.
(19) Seebach, D.; Abele, S.; Gademann, K.; Guichard, G.; Hintermann, T.; Jaun, B.; Matthews, J. L.; Schreiber, J. V.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1998**, *81*, 932−982.
(20) Review: Gellman, S. H. *Acc. Chem. Res.* **1998**, *31*, 173−180.
(21) Review: Gademann, K.; Hintermann, T.; Schreiber, J. V. *Curr. Med. Chem.* **1999**, *6*, 905−925.
(22) Seebach, D.; Abele, S.; Gademann, K.; Jaun, B. *Angew. Chem. Int. Ed.* **1999**, *38*, 1595−1597.
(23) Krauthäuser, S.; Christianson, L. A.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119*, 11719−11720.
(24) Chung, Y. J.; Christianson, L. A.; Stanger, H. E.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 10555−10556.
(25) Seebach, D.; Abele, S.; Sifferlen, T.; Hänggi, M.; Gruner, S.; Seiler, P. *Helv. Chim. Acta* **1998**, *81*, 2218−2243.
(26) Seebach, D.; Overhand, M.; Kühnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1996**, *79*, 913−941.
(27) Abele, S.; Guichard, G.; Seebach, D. *Helv. Chim. Acta* **1998**, *81*, 2141−2156.
(28) Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118*, 13071−13072.
(29) Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X.; Barchi, J. J.; Gellman, S. H. *Nature* **1997**, *387*, 381−384.
(30) Appella, D. H.; Barchi, J. J., Jr.; Durell, S. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 2309−2310.
(31) Gung, B. W.; Zou, D.; Stalcup, A. M.; Cottrell, C. E. *J. Org. Chem.* **1999**, *64*, 2176−2177.
(32) Seebach, D.; Abele, S.; Schreiber, J. V.; Martinoni, B.; Nussbaum, A. K.; Schild, H.; Schulz, H.; Hennecke, H.; Woessner, R.; Bitsch, F. *Chimia* **1998**, *52*, 734−739.
(33) Hintermann, T.; Seebach, D. *Chimia* **1997**, *51*, 244−247.
(34) Gademann, K.; Ernst, M.; Hoyer, D.; Seebach, D. *Angew. Chem. Int. Ed.* **1999**, *38*, 1223−1226.
(35) Werder, M.; Hauser, H.; Abele, S.; Seebach, D. *Helv. Chim. Acta* **1999**, *82*, 1774−1783.
(36) Gademann, K.; Ernst, M.; Seebach, D.; Hoyer, D. *Helv. Chim. Acta* **2000**, *83*, 16−33.
(37) Hamuro, Y.; Schneider, J. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1999**, *121*, 12200−12201.
(38) Porter, E. A.; Wang, X.; Lee, H.-S.; Weisblum, B.; Gellman, S. H. *Nature* **2000**, *404*, 565−565.
(39) Porter, E. A.; Wang, X.; Lee, H.-S.; Weisblum, B.; Gellman, S. H. *Nature* **2000**, *405*, 298−298.

nantly adopt extended conformations, and thus should favor the formation of pleated sheets.[26] This is in contrast to β-peptides composed of *like*-β[2,3]-amino acids (see Figure 1A), which have been shown to form predominantly 3_{14}-helical structures.[19] However, the first attempt to synthesize a β-hexapeptide consisting entirely of (*R*,*S*)-β[2,3] residues failed because of insurmountable solubility problems upon chain elongation.[40] The insolubility was presumably caused by the formation of sheets, held by intermolecular hydrogen bonds between the growing β-peptide strands.[22] The design of β-peptides forming hairpin structures was immediately identified as a solution to the aggregation problem, since intramolecular hydrogen bonds between the antiparallel strands would compete with the formation of hydrogen bonds between different molecules. The required turn structure had been, as a matter of fact, already characterized: insertion of (*S*)-β[2]−(*S*)-β[3] dipeptide sequences in β-peptides had been shown to induce the formation of 10-membered hydrogen-bonded turns,[18,19] similar to the β-turns observed in α-peptides.

A β-hexapeptide (Figure 1B) that could fold into a hairpin conformation was thus designed. A (potential) turn composed of a (*S*)-β[2]−(*S*)-β[3] dipeptide sequence, with the side chains of valine and lysine, was chosen for the construction of the antiparallel pleated sheet arrangement. Two dipeptide segments of (*R*,*S*)-β[2,3]-amino acids that (potentially) enforce an extended conformation were attached at the ends of this unit, and were expected to form intramolecular hydrogen bonds. The resulting β-hexapeptide was indeed soluble, even in water. Its conformational behavior in methanol solution has been investigated by NMR spectroscopy and MD simulation. The NMR structure of the β-hexapeptide in methanol has been reported in a preliminary communication.[22]

Here we comment on the NMR data from the perspective given by the MD simulations, and provide additional unpublished NMR data as Supporting Information. An analysis of two 100-ns MD simulations of the β-hexapeptide in methanol solution at 298 and 340 K, respectively, is presented. The simulations were started from an extended structure and reversible folding to the experimentally determined hairpin conformation was observed. The ensembles of structures from the trajectories are analyzed with regard to the conformational space sampled by the peptide, the thermodynamics of folding, and the level of agreement with the NMR-derived data. Implications of the simulation results to the interpretation of the NMR data are discussed.

2. Results and Discussion

2.1. Comparison between NMR-Derived and Simulation Data. This section is organized as follows. First, the lowest energy structure from the simulated annealing runs with restraints (X-PLOR structure number 1, see Figure 2 and section 4.1) will be compared to the structures sampled in the simulations. Second, the interproton distances derived from the NOE data and the experimental ^3J-coupling constants will be compared to the average interproton distances and the average ^3J-coupling constants calculated from the simulations.

The atom-positional root-mean-square deviation (RMSD) from the X-PLOR structure number1 for the backbone atoms of residues 2 to 5 is displayed in the lower panels of Figures 3 and 4 for the simulations at 298 and 340 K, respectively. When the RMSD is lower than 0.08 nm (dashed line) the simulation and experimentally derived conformations are considered equivalent. There is in both simulations an equilibrium between folded (hairpin-like) and unfolded structures. The upper panels of Figures 3 and 4 show the occurrence of backbone hydrogen bonds (those present in at least 5% of the structures) as a function of simulation time. The model (ideal) hairpin structure

(40) Abele, S. Ph.D. Thesis, ETH Zurich, Zurich, 1999.

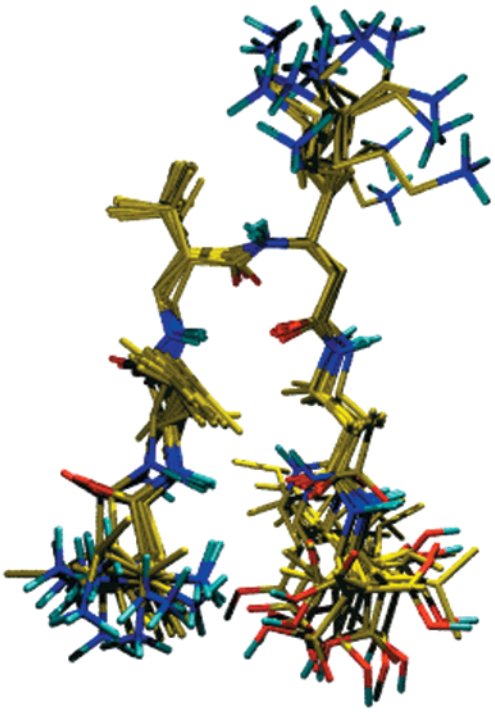

Figure 2. Superposition of the 15 structures with lowest energy from the ab initio simulated-annealing runs with upper-bound distance restraints and torsional-angle restraints derived from the NMR data. The simulated-annealing calculations were performed in vacuo with X-PLOR[54] (see section 4.1).

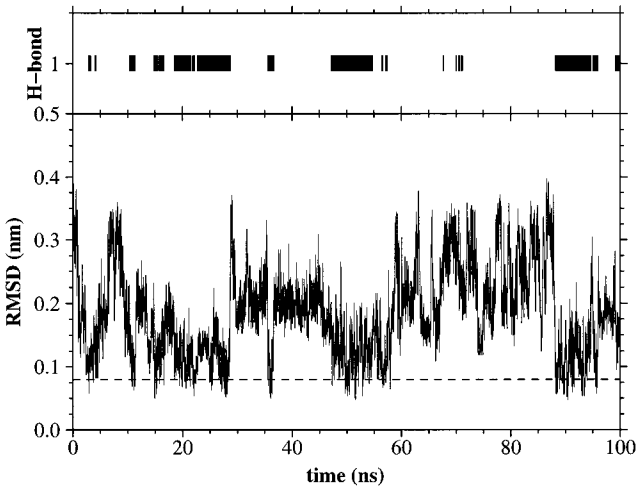

Figure 3. Upper panel: Occurrence of backbone hydrogen bonds at 298 K as a function of simulation time. A hydrogen bond is assumed to exist if the hydrogen−acceptor distance is smaller than 0.27 nm and the donor−hydrogen−acceptor angle is larger than 135°. Only those hydrogen bonds present in at least 5% of the configurations analyzed (2×10^5 at 0.5-ps intervals) are shown. The hydrogen bond NH(3)−CO(4), number 1 in the upper panel, is present in 19% of the configurations. Lower panel: Atom-positional root-mean-square deviation (RMSD) from the X-PLOR structure number 1 for the backbone atoms of residues 2 to 5. The dashed line corresponds to the RMSD (0.08 nm) used as similarity criterion within the clustering algorithm. Simulation structures with RMSD values below this value are considered conformationally equivalent to the reference X-PLOR structure. The initial structure (extended) has an RMSD of 0.40 nm.

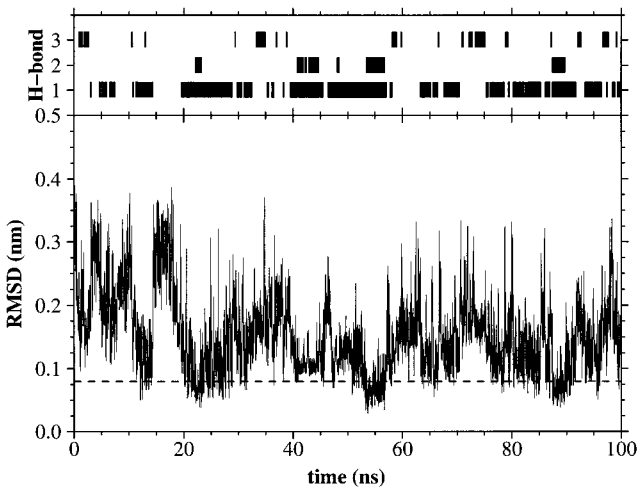

Figure 4. Upper panel: Occurrence of backbone hydrogen bonds at 340 K as a function of simulation time. A hydrogen bond is assumed to exist if the hydrogen−acceptor distance is smaller than 0.27 nm and the donor−hydrogen−acceptor angle is larger than 135°. Only those hydrogen bonds present in at least 5% of the configurations analyzed (2×10^5 at 0.5-ps intervals) are shown. The hydrogen bond NH(3)−CO(4), number 1 in the upper panel, is present in 30% of the configurations; the hydrogen bond NH(2)−CO(5), number 2 in the upper panel, is present in 6% of the configurations; the hydrogen bond NH(5)−CO(2), number 3 in the upper panel, is present in 5% of the configurations. Lower panel: Atom-positional RMSD from the X-PLOR structure number 1 for the backbone atoms of residues 2 to 5. The dashed line corresponds to the RMSD (0.08 nm) used as similarity criterion within the clustering algorithm. Simulation structures with RMSD values below this value are considered conformationally equivalent to the reference X-PLOR structure. The initial structure (extended) has an RMSD of 0.40 nm.

is formed by a turn, stabilized by a hydrogen bond between NH(3) and CO(4), and two antiparallel strands connected by hydrogen bonds between NH(2) and CO(5) and NH(1) and CO(6). The NH(3)−CO(4) hydrogen bond is present in 73% of the (15) X-PLOR structures (see Figure 2), in 19% of the structures from the simulation at 298 K, and in 30% of the structures from the simulation at 340 K. The NH(2)−CO(5) hydrogen bond is present in 7% of the X-PLOR structures and in 6% of the structures from the simulation at 340 K. The occurrence of this hydrogen bond in the simulation at 298 K is less than 1%. The NH(1)−CO(6) hydrogen bond is only present in 7% of the X-PLOR structures and is almost completely absent (<1%) from the simulations. The lack of other important (>5%) backbone hydrogen bonds in Figures 3 and 4 indicates that besides the hairpin turn no other secondary structure elements are significantly populated in the simulations.

The 20 NOEs from the ROESY spectrum and the derived upper-bound distances are listed in Table 1. The average effective violations of these distances by the X-PLOR structures and by the structures from the simulations at 298 and 340 K are also listed. The upper-bound nature of the NOE-derived distances implies that only positive values of $\langle v \rangle_{\mathrm{X-PLOR}}$, $\langle v \rangle_{298}$, and $\langle v \rangle_{340}$ in Table 1 are true violations. Note that in the case of a fast-tumbling molecule like the β-hexapeptide studied here, the $\langle\langle r^{-3} \rangle^2 \rangle^{-1/6}$ and $\langle r^{-6} \rangle^{-1/6}$ distances are essentially equivalent. The only NOEs in Table 1 that carry long-range (in residue sequence) structural information are the last three, i.e., a weak NOE between $HC^{\alpha}(1)$ and $HC^{\beta}(6)$, a medium NOE between $HC^{\alpha}(2)$ and $HC^{\beta}(5)$, and a weak NOE between the methyl groups of the side chains at $C^{\beta}(2)$ and $C^{\alpha}(5)$. The upper-bound distances corresponding to short-range (in residue sequence) NOEs (first 17 in Table 1) are satisfied in both the X-PLOR structures and the simulations. The $HC^{\alpha}(1)-HC^{\beta}(6)$ upper-bound distance is satisfied by the X-PLOR structures and by the ensemble of structures from the simulation at 340 K, but is severely violated in the simulation at 298 K. On the other hand, the $HC^{\alpha}(2)-HC^{\beta}(5)$ upper-bound distance is satisfied by the X-PLOR structures and by the ensemble of structures from the simulation at 340 K, and is only marginally violated in the simulation at 298 K. The upper-bound distance between the side chains of the same two residues is satisfied in all cases. Figure 5 reveals the significant differences in the distribution of interproton distances at the two simulation temperatures for the two NOEs that determine the antiparallel arrangement. The distribution of distances between the $HC^{\alpha}(1)$ and $HC^{\beta}(6)$ hydrogen atoms at 340 K is clearly displaced, with respect to the distribution at 298 K, to lower values (Figure 5A). Most importantly, the distribution at 340 K shows a well-defined peak at around 0.3 nm that is perfectly correlated with the occurrence of ideal hairpin structures (with perfectly antiparallel strands) in the simulation. This peak is not present in the distribution at 298 K, a temperature at which no ideal hairpin structure was sampled indeed. The distribution of $HC^{\alpha}(2)-HC^{\beta}(5)$ distances at 340 K is also shifted, with respect to the distribution at 298 K, to slightly lower values (Figure 5B). At 298 K the distribution has a minimum around 0.6 nm that connects two ranges of distances, most probably corresponding to two distinct groups of conformations, i.e., with and without the turn. Although the peaks in the distributions are at about the same locations at the two temperatures, at 340 K the $HC^{\alpha}(2)-HC^{\beta}(5)$ distances are more uniformly distributed over the entire range. This is not surprising, since at the higher temperature a larger number of conformational microstates become accessible. Two main conclusions arise from this analysis. First, but not unexpectedly, it is clear from Table 1 and Figure 5 that the use of upper-bound interproton distances in NMR structure determination can easily lead to unrealistic hybrid structures if averaging is not taken into account.[2,41] For example, the simulation at 340 K satisfies in an average way all the NOE-derived interproton distances, even if very few ideal hairpin structures have been actually sampled. This contrasts with the more simplified picture of the conformation of the peptide in methanol solution given by the 15 X-PLOR structures shown in Figure 2. In cases where a mix of conformational states may exist in the sample within the time scale of the measurement, the NOE-derived distances should be used as upper-bound restraints for average distances from a physically sound sampling of conformational space rather than as upper-bound restraints for instantaneous distances in some optimization procedure.[42−46] This implies that the properties of an ensemble of structures rather than those of an individual structure need to be studied. To this respect, the sets of structures commonly produced in NMR structure determination are not ensembles in a statistical-mechanical sense but individual solutions to a given set of restraints. The second conclusion is that the last three NOEs in Table 1 are not necessarily indicative of a perfect antiparallel arrangement, i.e.,

(41) Daura, X.; Antes, I.; van Gunsteren, W. F.; Thiel, W.; Mark, A. E. *Proteins: Struct. Funct. Genet.* **1999**, *36*, 542−555.
(42) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1989**, *157*, 289−294.
(43) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. *J. Mol. Biol.* **1990**, *214*, 223−235.
(44) Torda, A. E.; Brunne, R. M.; Huber, T.; Kessler, H.; van Gunsteren, W. F. *J. Biomol. NMR* **1993**, *3*, 55−66.
(45) Nanzer, A. P.; van Gunsteren, W. F.; Torda, A. E. *J. Biomol. NMR* **1995**, *6*, 313−320.
(46) Nanzer, A. P.; Huber, T.; Torda, A. E.; van Gunsteren, W. F. *J. Biomol. NMR* **1996**, *8*, 285−291.

Table 1. NOEs from the 150-ms ROESY Spectrum of the β-Hexapeptide (CD_3OH, 500 MHz) and Average Effective Violations of the Associated Upper-Bound Distances from the X-PLOR Structures and from the Structures in the MD Simulations at 298 and 340 K[a]

H atom	H atom	NOE	$\langle v\rangle_{X-PLOR}$ (nm)	$\langle v\rangle_{298}$ (nm)	$\langle v\rangle_{340}$ (nm)
H−N(2)	H−CMe_2(2)	m	−0.02	−0.07 (−0.08)	−0.07 (−0.08)
H−N(2)	H−C^α(2)	m	−0.07	−0.06 (−0.06)	−0.05 (−0.06)
H−N(2)	H−C^β(1)	w	−0.10	−0.05 (−0.06)	−0.07 (−0.08)
H−N(2)	H−C^α(1)	s	−0.08	−0.08 (−0.08)	−0.08 (−0.08)
H^S−C^β(3)	H−CMe_2(3)	m	−0.08	−0.08 (−0.08)	−0.07 (−0.08)
H^S−C^β(3)	H−C^α(3)	m	−0.11	−0.11 (−0.11)	−0.11 (−0.11)
H−N(3)	H−C^β(2)	m	−0.04	−0.08 (−0.09)	−0.06 (−0.07)
H−N(3)	H−C^α(2)	s	−0.05	−0.07 (−0.07)	−0.08 (−0.08)
H−N(4)	H^S−C^α(4)	w	−0.08	−0.15 (−0.16)	−0.14 (−0.14)
H−N(4)	H^R−C^α(4)	s	−0.04	−0.04 (−0.04)	−0.04 (−0.04)
H−N(5)	Me−C^α(5)	w	−0.20	−0.24 (−0.24)	−0.23 (−0.23)
H−N(5)	Me−C^β(5)	m	−0.15	−0.16 (−0.16)	−0.16 (−0.16)
H−N(5)	H^S−C^α(4)	m	−0.13	−0.10 (−0.10)	−0.11 (−0.11)
H−N(5)	H−C^α(5)	m	−0.09	−0.05 (−0.05)	−0.06 (−0.06)
H−N(5)	H−C^β(4)	m	−0.04	−0.08 (−0.09)	−0.07 (−0.08)
H−N(6)	H−C^α(6)	m	−0.05	−0.06 (−0.07)	−0.07 (−0.08)
H−N(6)	H−CMe_2(6)	m	−0.09	−0.05 (−0.06)	−0.05 (−0.06)
H−C^α(1)	H−C^β(6)	w	−0.02	0.42 (0.38)	0.02 (0.00)
H−C^α(2)	H−C^β(5)	m	−0.10	0.07 (0.04)	0.00 (−0.02)
Me_2−CH(2)	Me−C^α(5)	w	−0.24	−0.05 (−0.08)	−0.13 (−0.15)

[a] The NOEs have been classified in three r_{exp} distance categories: s (strong, ≤0.3 nm), m (medium, ≤0.35 nm), and w (weak, ≤0.45 nm). The average effective distance violations are calculated as $\langle r^{-3}\rangle^{-1/3} - r_{exp}$ for the 15 X-PLOR structures ($\langle v\rangle_{X-PLOR}$), and as $\langle\langle r^{-3}\rangle^2\rangle^{-1/6} - r_{exp}$ and $\langle r^{-6}\rangle^{-1/6} - r_{exp}$ (in parentheses) for 2×10^5 trajectory structures (one per 0.5 ps) from the MD simulations at 298 ($\langle v\rangle_{298}$) and 340 K ($\langle v\rangle_{340}$). H−CMe_2 refers to the proton of the terciary carbon in a valine or leucine side chain (see Figure 1B). Me−C^α and Me−C^β refer to the protons of the methyl group in an alanine side chain in the C^α and C^β positions, respectively. Me_2−CH refers to the protons of the two methyl groups of the terciary carbon in a valine side chain.

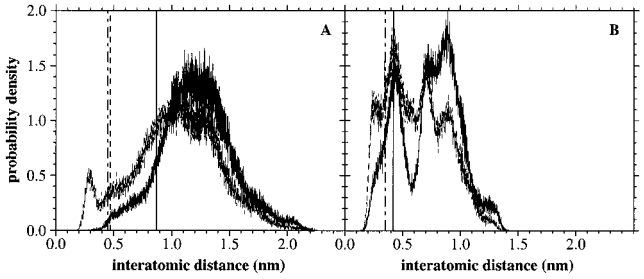

Figure 5. Panel A: Distribution of distances between the hydrogen pair $HC^\alpha(1)/HC^\beta(6)$ (see Table 1) in the simulations at 298 K (solid line) and 340 K (dashed line). The vertical lines correspond to the upper-bound distance derived from the NOE data (0.45 nm, dot−dashed line), and to the $\langle\langle r^{-3}\rangle^2\rangle^{-1/6}$ average distance from the simulations at 298 K (0.87 nm, solid line) and 340 K (0.47 nm, dashed line). Panel B: Distribution of distances between the hydrogen pair $HC^\alpha(2)/HC^\beta(5)$ in the simulations at 298 (solid line) and 340 K (dashed line). The vertical lines correspond to the upper-bound distance derived from the NOE data (0.35 nm, dot−dashed line), and to the $\langle\langle r^{-3}\rangle^2\rangle^{-1/6}$ average distance from the simulations at 298 (0.42 nm, solid line) and 340 K (0.35 nm).

with hydrogen bonds between residues 2 and 5 and 1 and 6. These hydrogen bonds are only marginally present in the 15 X-PLOR structures and in the simulation at 340 K and, yet, the upper-bound distances derived from the NOEs are satisfied in both cases. Finally, the ensemble of structures from the simulation at 340 K reproduces the NOE-derived distances surprisingly better than the ensemble of structures from the simulation at 298 K. This is in contrast to results from a β-heptapeptide in methanol solution studied previously.[41,47] This point will be further discussed in sections 2 and 3.

The 14 ^3J-coupling constants extracted from the one-dimensional ^1H NMR spectrum and the corresponding average coupling constants calculated for the 15 X-PLOR structures and for the trajectory structures from the simulations at 298 and 340 K are listed in Table 2. The distributions of the 11 torsional dihedral angles (3 of the 14 are redundant) in the simulations and the corresponding distributions of ^3J-coupling constants are displayed in Figures 6 (for HC^β−HC^α) and 7 (for HN−HC^β). The dihedral angles corresponding to the experimental ^3J-coupling constants are also listed in Table 2. The following discussion is based on the assumption that the Karplus relation, including the chosen parameters (see section 4.2.3), is valid for the couplings observed in the spectrum. Note also that the torsional dihedral angles calculated with X-PLOR for the final structures from the restrained runs might be slightly different

Table 2. ^3J-Coupling Constants for the β-Hexapeptide (CD_3OH, 500 MHz) and Average Violations from the X-PLOR Structures and from the Structures in the MD Simulations at 298 and 340 K[a]

H atom	H atom	$^3J_{exp}$ (Hz)	angle (deg)	$\langle v\rangle_{X-PLOR}$ (Hz)	$\langle v\rangle_{298}$ (Hz)	$\langle v\rangle_{340}$ (Hz)
H−C^β(1)	H−C^α(1)	7.0	±131/±34	−1.7	3.1	0.6
H−N(2)	H−C^β(2)	10.1	±180	−1.0	−1.0	−1.1
H−C^β(2)	H−C^α(2)	9.7	±146/0	1.6	−3.8	−3.4
H−N(3)	H^S−C^β(3)	7.4	±146	−3.6	−1.6	−1.6
H−N(3)	H^R−C^β(3)	3.8	±117/±48	0.9	2.3	2.2
H^S−C^β(3)	H−C^α(3)	3.8	±113/±56	−0.8	0.1	0.0
H^R−C^β(3)	H−C^α(3)	10.7	±153	2.1	−0.4	−0.8
H−N(4)	H−C^β(4)	9.0	±162	0.6	0.0	0.0
H−C^β(4)	H^S−C^α(4)	4.1	±115/±54	−0.4	2.3	0.7
H−C^β(4)	H^R−C^α(4)	9.6	±146/±6	3.2	−1.8	−0.4
H−N(5)	H−C^β(5)	10.6	±180	−1.0	−1.6	−1.9
H−C^β(5)	H−C^α(5)	9.2	±143/±14	3.6	−4.0	−2.9
H−N(6)	H−C^β(6)	7.0	±142/0	0.9	2.0	1.7
H−C^β(6)	H−C^α(6)	7.0	±131/±34	1.8	−0.8	0.4
⟨|violation|⟩ (Hz)				1.7	1.8	1.3

[a] The average violations are calculated as $\langle^3J\rangle - {}^3J_{exp}$ for the 15 X-PLOR structures ($\langle v\rangle_{X-PLOR}$) and for 2×10^5 trajectory structures (one per 0.5 ps) from the MD simulations at 298 ($\langle v\rangle_{298}$) and 340 K ($\langle v\rangle_{340}$). The torsional dihedral angles in the second numerical column are calculated from the experimental ^3J-coupling constants according to the Karplus relation (eq 1), using $a = 6.4$ Hz, $b = -1.4$ Hz, and $c = 1.9$ Hz for ^3J(HN,HC), and $a = 9.5$ Hz, $b = -1.6$ Hz, and $c = 1.8$ Hz for ^3J(HC,HC).

(47) Gademann, K.; Jaun, B.; Seebach, D.; Perozzo, R.; Scapozza, L.; Folkers, G. *Helv. Chim. Acta* **1999**, *82*, 1−11.

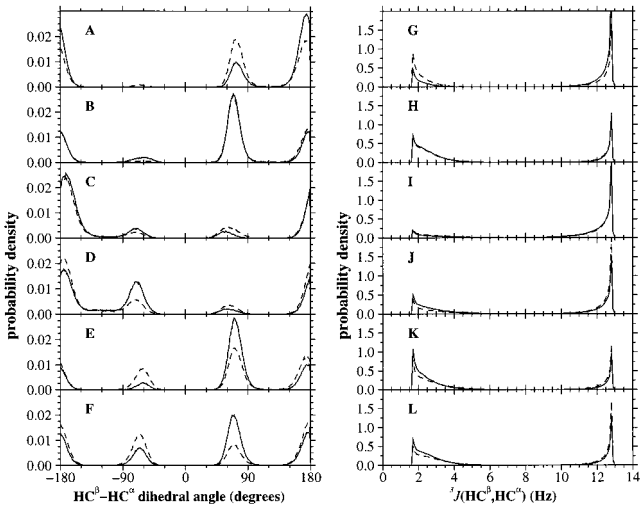

Figure 6. Distribution of values of the torsional dihedral angle HC$^{\beta}$−HC$^{\alpha}$ (panels A to F for residues 1 to 6, respectively) and corresponding distribution of ^{3}J(HC$^{\beta}$,HC$^{\alpha}$) coupling values (panels G to L for residues 1 to 6, respectively) from the simulations at 298 (solid line) and 340 K (dashed line). The ^{3}J(HC$^{\beta}$,HC$^{\alpha}$) coupling values are calculated with the Karplus relation (eq 1), using $a = 9.5$ Hz, $b = -1.6$ Hz, and $c = 1.8$ Hz. Where the same torsion is defined in terms of H^{R}−C and H^{S}−C, only the H^{R}−C case is considered (see Table 2).

from the ones given in column $\langle v \rangle_{\mathrm{X-PLOR}}$, Table 2. In the restrained runs the hydrogen atoms were treated explicitly, and the competition between restraining forces and physical forces might have led to the distortion of bond angles involving hydrogen atoms, which affects all related torsional dihedral angles. In contrast, for the calculation of the ^{3}J values given in column $\langle v \rangle_{\mathrm{X-PLOR}}$ in Table 2, all aliphatic hydrogen atoms were eliminated, and the torsional dihedral angles were recalculated by building virtual atoms at positions based on standard (minimum energy) carbon atom geometries. Overall, the agreement between the experimental ^{3}J values and the average ^{3}J values calculated for the 15 X-PLOR structures and for the ensembles of structures from the simulations at 298 and 340 K is good for the region of the hairpin turn (residues 3 and 4) and poor for residues 1, 2, 5, and 6. The average violation (see Table 2) is 1.3 Hz for the ensemble of structures from the simulation at 340 K, 1.7 Hz for the 15 X-PLOR structures, and 1.8 Hz for the ensemble of structures from the simulation at 298 K. In a model (ideal) hairpin structure the HC$^{\beta}$−HC$^{\alpha}$ torsional dihedral angle would be around 180° in all six residues; in residues 3 and 4 this would be the case for the H^{R}C$^{\beta}$−HC$^{\alpha}$ and HC$^{\beta}$−H^{R}C$^{\alpha}$ torsions, respectively. The HC$^{\beta}$−HC$^{\alpha}$ dihedral angles calculated from the experimental ^{3}J-coupling constants for the six residues (H^{R}C$^{\beta}$−HC$^{\alpha}$ and HC$^{\beta}$−H^{R}C$^{\alpha}$ for residues 3 and 4, respectively, Table 2) are between 131° and 153°, suggesting that the hairpin is not the only populated conformation. In the MD simulations at 298 and 340 K the HC$^{\beta}$−HC$^{\alpha}$ torsional dihedral angles are distributed in three regions (Figure 6), i.e. −60° (or gauche(+)), 60° (or gauche(−)), and 180° (or trans). The actual peaks of the distribution are, nevertheless, displaced from ±60° to slightly larger angles. In this way the system minimizes the repulsive interaction between the side chains at the C$^{\beta}$ and C$^{\alpha}$ carbon atoms, which are spatially close when the HC$^{\beta}$−HC$^{\alpha}$ dihedral angle is ±60°. The main difference between the distributions of dihedral angles at 298 and 340 K is not the position (and not the width) of the peaks but the relative populations of gauche and trans. At the higher temperature the populations of gauche and trans are in general more evenly distributed. The differences between the distributions of torsional dihedral angles at the two temperatures are not entirely reflected in the comparison of the corresponding distributions of ^{3}J values, since the Karplus curve is symmetric for positive and negative angles. In the simulation at 340 K the torsion HC$^{\beta}$(1)−HC$^{\alpha}$(1) populates with similar weights the gauche(−) and trans conformations (Figure 6A,G), and the experimental ^{3}J value is reproduced. Conversely, in the simulation at 298 K the same torsion angle populates predominantly the trans conformation, and the average ^{3}J value is about 3 Hz higher than the experimental one. The very similar distribution of angles of the torsion HC$^{\beta}$(2)−HC$^{\alpha}$(2) in the two simulations and the high population of gauche(−) (Figure 6B,H) have a clear correspondence with the conformation of the central-member structure of cluster number 1 (see section 2.2). For this torsion angle the distribution of angle values is in disagreement with the experimental ^{3}J value (see Table 2). The torsions H^{R}C$^{\beta}$(3)−HC$^{\alpha}$(3) (Figure 6C,I) and HC$^{\beta}$(4)−H^{R}C$^{\alpha}$(4) (Figure 6D,J) show a smooth transition between gauche(+) and trans, i.e., the intermediate angles are also populated, due to the absence of a side chain that would restrict rotation at the C$^{\beta}$ and C$^{\alpha}$ atoms, respectively (see Figure 1B). These two torsions are at the turn of the hairpin and their distribution of torsion angle values in the two simulations is fully compatible with the experimental data: The difference between the experimental and the average calculated ^{3}J values is smaller than 1 Hz (a bit higher for the HC$^{\beta}$(4)−H^{R}C$^{\alpha}$(4) torsion at 298 K, see Table 2). The agreement between experiment and simulation regarding the H^{R}C$^{\beta}$(3)−HC$^{\alpha}$(3) and HC$^{\beta}$(4)−H^{R}C$^{\alpha}$(4) torsions, in combination with the observation that for both torsions there is some population of gauche conformations in the simulations, suggests that there must also be a mix of conformers (with and without the turn) in the "test tube". Thus, these two torsions being predominantly in a trans conformation in the 15 X-PLOR structures, i.e., compatible with the turn structure, result in the corresponding average ^{3}J values being 2 to 3 Hz higher than the experimental ones. The distribution of angles of the torsion HC$^{\beta}$(5)−HC$^{\alpha}$(5) (Figure 6E,K) is in disagreement with the experimental ^{3}J value (see Table 2). As with the same torsion in residue 2, the population of gauche(−) is too large to meet the experimental ^{3}J value of 9.2 Hz, particularly at 298 K. The experimental ^{3}J(HC$^{\beta}$,HC$^{\alpha}$) value for residue 6 is compatible with the distributions from the simulations at 298 and 340 K (Figure 6F,L), even if there are clear differences between the distributions at the two temperatures. The distribution of angles of the HN−HC$^{\beta}$ torsional dihedrals of residues 2 to 5 is very similar in the two simulations (Figure 7). In a model (ideal) hairpin structure these torsional dihedral angles would have values around 180°, with the exception of residue 3, for which the HN−H^{R}C$^{\beta}$ angle would be 120°. The HN−HC$^{\beta}$ dihedral angles calculated from the experimental ^{3}J-coupling constants for residues 2 to 5 (Table 2) are compatible with the hairpin arrangement. On the other hand, the experimental ^{3}J(HN,HC$^{\beta}$) value for residue 6 is substantially lower than expected for a trans conformation, indicating that this torsion might sample a larger range of angles. In the simulations at 298 and 340 K the distributions of angles of the HN−HC$^{\beta}$ torsional dihedrals of residues 2, 4, 5, and 6 (Figure 7) are centered at around 180°, although with relatively large widths. The high experimental ^{3}J(HN,HC$^{\beta}$) values for residues 2 and 5 (see Table 2) give an idea of the uncertainty in the calculated ^{3}J values: The Karplus curve that has been used (see section 4.2.3) has a maximum ^{3}J(HN,HC$^{\beta}$) value of 9.7 Hz at 180°, lower than the experimental 10.1 and 10.6 Hz for residues 2 and 5, respectively.

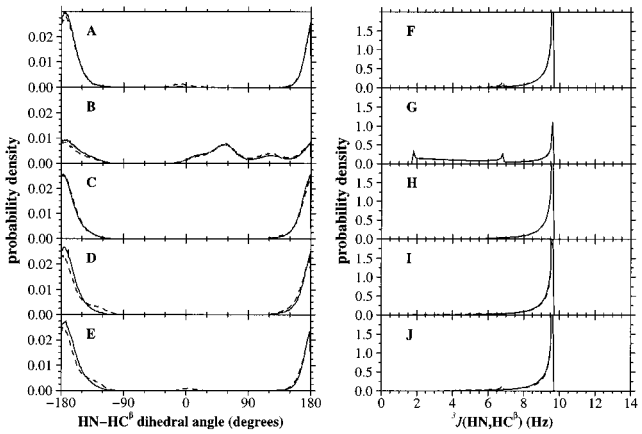

Figure 7. Distribution of values of the torsional dihedral angle HN−HC$^{\beta}$ (panels A to E for residues 2 to 6, respectively) and corresponding distribution of ^3J(HN,HC$^{\beta}$) coupling values (panels F to J for residues 2 to 6, respectively) from the simulations at 298 (solid line) and 340 K (dashed line). The ^3J(HN,HC$^{\beta}$) coupling values are calculated with the Karplus relation (eq 1), using $a = 6.4$ Hz, $b = -1.4$ Hz, and $c = 1.9$ Hz. Where the same torsion is defined in terms of H^{R}−C and H^{S}−C, only the H^{R}−C case is considered (see Table 2).

Taking this into account, it can be concluded that the distribution of angles of the HN(2)−HC$^{\beta}$(2) (Figure 7A,F) and HN(5)−HC$^{\beta}$(5) (Figure 7D,I) torsions in the simulations at 298 and 340 K are compatible with the respective experimental ^3J values. The torsion HN(3)−H^{R}C$^{\beta}$(3), at the beginning of the turn in the hairpin structure, samples a wide range of angles at both temperatures (Figure 7B,G). This is made possible by the absence of a side chain at the C$^{\beta}$ atom that would restrict rotation. For this torsion the distribution of angles is in disagreement with the experimental ^3J value (see Table 2), mostly because of the low population of the 120° conformation. For the torsion HN(4)−HC$^{\beta}$(4), at the middle of the turn, the distribution of angles in the two simulations (Figure 7C,H) is fully compatible with the experimental ^3J value, which suggests an approximately trans conformation (around 160° in average). The distribution of angles of the HN(6)−HC$^{\beta}$(6) torsion in the simulations at 298 and 340 K (Figure 7E,J) translates to ^3J values that are almost 2 Hz higher than the experimental one (see Table 2), indicating that the trans conformation is less populated in the experimental sample than in the simulations.

Overall, the experimental ^3J-coupling constants are less well reproduced by the ensembles of structures from the simulations than the NOE-derived upper-bound distances. Nevertheless, the uncertainty in the conversion of both, experimental NOEs to upper-bound distances and torsional dihedral angles from simulation to ^3J-coupling constants, is high. Furthermore, the poor agreement between the experimental ^3J-coupling constants and the ^3J values calculated for the 15 X-PLOR structures suggests that the derived torsional dihedral angles are partially incompatible with the derived upper-bound distances if they are to be fitted to a single conformation. This is not surprising, since these two properties are subject to two distinct types of averaging. The best agreement with experiment has been obtained for the region of the turn (residues 3 and 4). More importantly, it has been shown that the experimental values reported in Tables 1 and 2 are best explained in terms of ensemble averages. The effect of the averaging can be as striking as in the following example: the average ^3J value calculated from simulation for the hydrogen pair H^{R}C$^{\beta}$−HC$^{\alpha}$ in residue 3 is around 10 Hz at the two temperatures studied, i.e., very close to the experimental value at 298 K (see Table 2), yet the corresponding distributions of ^3J values (see Figure 6I) have very low probability density around 10 Hz.

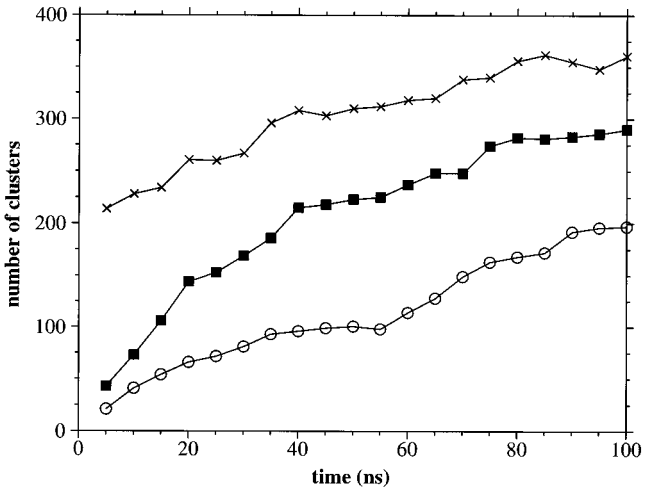

Figure 8. Number of clusters of peptide conformations as a function of simulation time from the trajectories at 298 (circles) and 340 K (squares), and from a 200-ns merged trajectory in which the first 100 ns correspond to the simulation at 298 K (circles) and the second 100 ns correspond to the simulation at 340 K (crosses, shifted to the interval 0−100 ns). The total number of clusters at 100 ns is 197 at 298 K, 290 at 340 K, and 361 for the merged trajectory. The clustering is performed for 10^4 (2×10^4 for the merged trajectory) structures taken at 10-ps intervals.

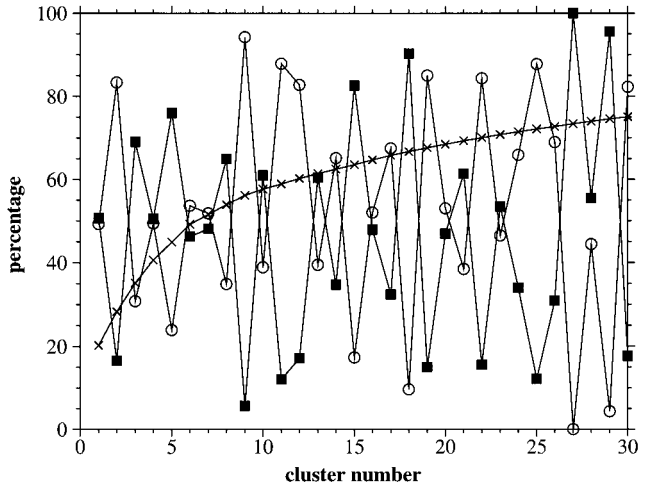

Figure 9. Percentage of members from the simulation at 298 K (circles) and from the simulation at 340 K (squares) in clusters 1 to 30 when the conformational clustering is performed over a merged trajectory in which the first 100 ns correspond to the simulation at 298 K and the second 100 ns correspond to the simulation at 340 K. The crosses indicate the percentage of structures from the total of 2×10^4 structures analyzed (one per 10 ps) that is included in the actual plus all preceding clusters.

2.2. Conformational Analysis. A conformational-clustering algorithm has been used to characterize the configurational space sampled in the simulations (see section 4.2.3). The clustering has been performed for the trajectories at 298 and 340 K as well as for a merged trajectory in which the first 100 ns correspond to the simulation at 298 K and the second 100 ns correspond to the simulation at 340 K. The number of clusters of conformations from these three trajectories is plotted as a function of time in Figure 8. The number of clusters is approximately constant over the last 10 ns of the two independent trajectories, but there is no definite indication of conver-

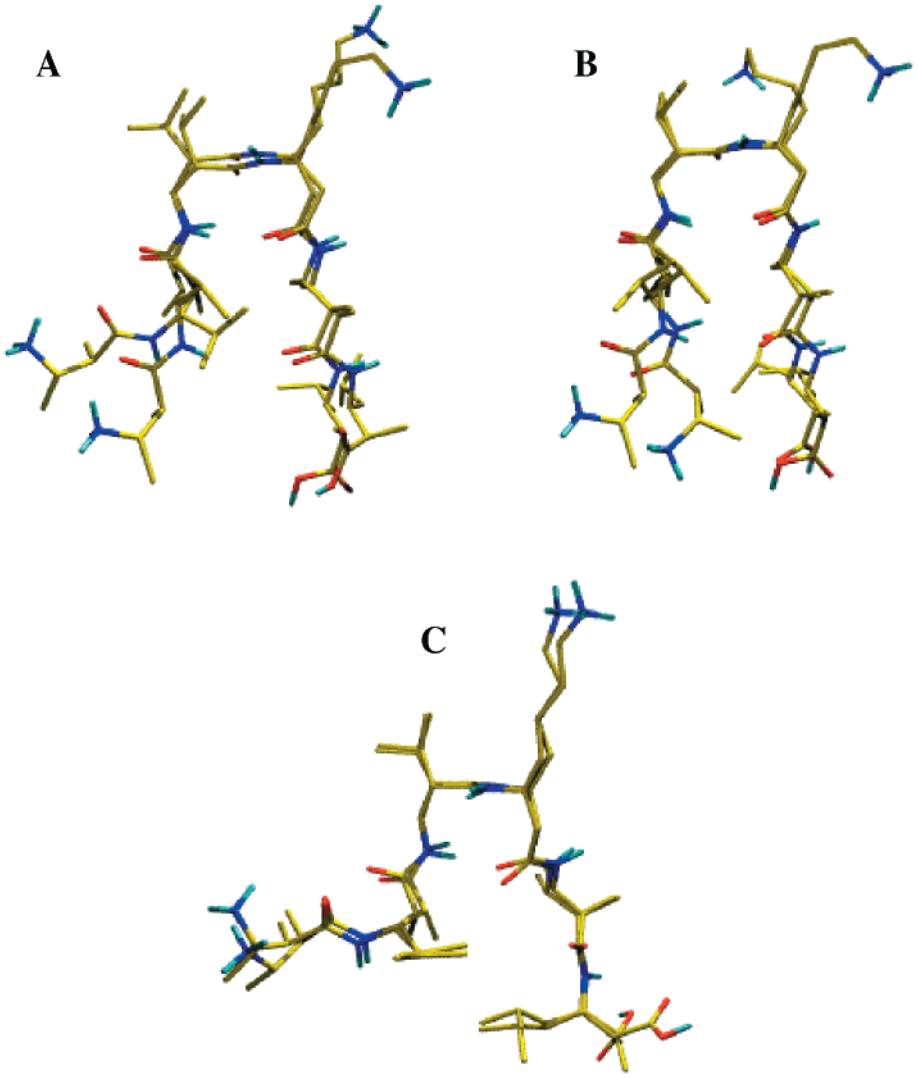

Figure 10. Panel A: Superposition of the reference X-PLOR structure number 1 and the structure with a minimum atom-positional RMSD for the backbone atoms of residues 2 to 5 from this reference structure in the simulation at 298 K (structure at time 90.41 ns, 0.05 nm RMSD). Panel B: Superposition of the X-PLOR structure number 1 and the structure with a minimum atom-positional RMSD from it in the simulation at 340 K (structure at time 53.62 ns, 0.03 nm RMSD). Panel C: Superposition of the central-member structures of cluster number 1 from the simulations at 298 (structure at time 25.35 ns) and 340 K (structure at time 29.96 ns). The atom-positional RMSD between the two structures for the backbone atoms of residues 2 to 5 is 0.02 nm.

gence, i.e., of complete sampling of configurational space. Nevertheless, the most relevant (probable) conformations (clusters), i.e., the conformations with lowest free energy, are likely to be sampled within the 100-ns period (see below). As expected, the number of conformations sampled is bigger in the simulation at 340 K (290 clusters) than in the simulation at 298 K (197 clusters). However, the total number of clusters, even at the higher temperature, is orders of magnitude lower than it could be expected from an exhaustive enumeration of possible (without sizable atom overlap) conformations of a peptide with 18 backbone rotatable bonds. Moreover, there is a substantial overlap between the conformational spaces sampled in the simulations at 298 and 340 K, since the number of clusters in the merged trajectory (361) is much lower than the theoretical maximum of 487 clusters that would result from a zero overlap of the two spaces. This is further illustrated in Figure 9, where the clusters from the merged trajectory are decomposed in terms of percentages of members originating from the simulation at 298 K and members originating from the simulation at 340 K. Approximately 50% of the members of cluster number 1 belong to the 298 K simulation and the other 50% belong to the 340 K simulation. In the rest of the 30 clusters shown in Figure 9 the overlap between the two simulations is also substantial. Indeed, the first cluster with members from only one of the temperatures is cluster number 27. Regarding the weight of the most populated clusters, cluster number 1 contains approximately 20% of the whole (merged) ensemble of structures, and the first 30 clusters (from the total of 361) contain as much as 75% of the ensemble. These numbers suggest that the most relevant parts of the conformational space accessible to the peptide at each of the two temperatures have been already sampled within the 100-ns simulations, even if the complete space has not been sampled.

Figure 10 shows a superposition of the reference X-PLOR structure (number 1 of 15) with a structure from the simulation at 298 K (Figure 10A) and from the simulation at 340 K (Figure 10B). The chosen structures are those with lowest atom-positional RMSD for the backbone atoms of residues 2 to 5 from the reference structure. A superposition of the central member structure of cluster number 1 from each of the two simulations is also shown (Figure 10C). Surprisingly, the central member structure of cluster number 1 is almost identical in the two simulations. It has the 10-membered turn closed by the

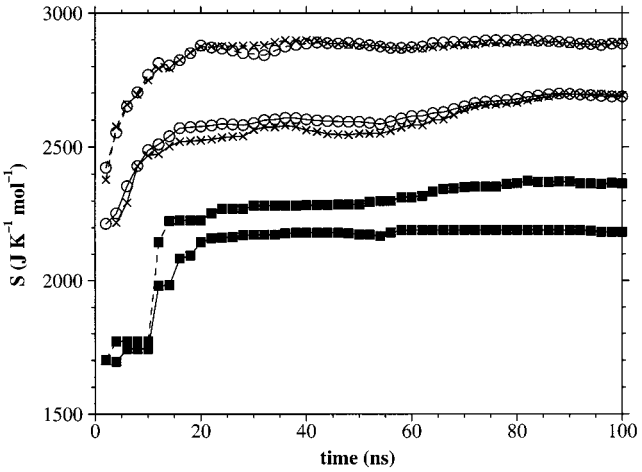

Figure 11. Configurational entropy of the peptide as a function of simulation time in the simulations at 298 (solid lines) and 340 K (dashed lines). Circles: Total peptide entropy; the value at 100 ns is 2686 J K^{-1} mol^{-1} at 298 K and 2886 J K^{-1} mol^{-1} at 340 K. Squares: Entropy of the folded state, constituted by all structures (from the total of 2×10^5 at 0.5-ps intervals) with an atom-positional RMSD for the backbone atoms of residues 2 to 5 from the central-member structure of cluster number 1 smaller than 0.08 nm; the value at 100 ns is 2183 J K^{-1} mol^{-1} at 298 K and 2363 J K^{-1} mol^{-1} at 340 K. Crosses: Entropy of the unfolded state, constituted by all structures with and atom-positional RMSD from the central-member structure of cluster number 1 larger than 0.12 nm; the value at 100 ns is 2692 J K^{-1} mol^{-1} at 298 K and 2892 J K^{-1} mol^{-1} at 340 K.

NH(3)−CO(4) hydrogen bond, but the two strands are not completely extended—especially at residue 2—and there are no further hydrogen bonds between them. The unexpected elbow at residue 2 is responsible for the low $^3J(HC^{\beta},HC^{\alpha})$ value discussed in section 2.1 (see Table 2). As can be observed in Figure 10A,B, there are in the two trajectories some structures which resemble the reference X-PLOR structure, and some which are even closer to a model hairpin. However, the hairpin conformation (including both the turn and properly antiparallel strands) is clearly not the lowest free energy conformation in the simulations. For example, the structure at 90.41 ns from the simulation at 298 K (Figure 10A) belongs to cluster number 4 (from the single-temperature clustering analysis), and has 791 structures (out of a total of 10^4) at an RMSD radius of 0.08 nm. The structure at 25.35 ns (Figure 10C), which is the central member structure of cluster number 1 from the trajectory at 298 K and also of cluster number 1 from the merged trajectory, has an atom-positional RMSD for the backbone atoms of residues 2 to 5 from the X-PLOR structure number 1 of 0.10 nm, and has 1998 structures at an RMSD radius of 0.08 nm. In the simulation at 340 K, the structure at 53.62 ns (Figure 10B) belongs to cluster number 2, and has 698 structures at an RMSD radius of 0.08 nm. The structure at 29.96 ns (Figure 10C), which is the central member structure of cluster number 1 at 340 K, has an atom-positional RMSD for the backbone atoms of residues 2 to 5 from the X-PLOR structure number 1 of 0.10 nm, and has 2038 structures at an RMSD radius of 0.08 nm.

2.3. Peptide Entropy and Free Energy of Folding. The configurational entropy of the peptide is plotted in Figure 11 as a function of time. Translational and rotational entropy due to center of mass motion are removed by least-squares fitting of the structures prior to the entropy calculation (see section 4.2.3). The peptide entropy converges to values around 2686 J K^{-1} mol^{-1} at 298 K and 2886 J K^{-1} mol^{-1} at 340 K within 100 ns of simulation. The configurational entropy of the folded state and that of the unfolded state have also been calculated (see section 4.2.3 for the definition of the folded and unfolded states). Note that the reference structure used in the definition of the folded state is the central member structure of cluster number 1 (see Figure 10C) rather than one of the 15 X-PLOR structures. This choice is validated by the, in general terms, good correspondence between the experimentally derived data and the ensemble averages from simulation (see section 2.1), and is more consistent with the conformational analysis presented in section 2.2. The definition of the boundaries between the folded and unfolded states is, to some extent, arbitrary. However, the conclusions drawn in this section, not the actual numbers, are independent of the spherical cutoffs used to differentiate between the two types of structures. As expected, the entropy of the folded state is lower than that of the unfolded state, and the entropies of the folded and unfolded states are lower in the simulation at 298 K than in the simulation at 340 K (see Figure 11). Not surprisingly either, at both temperatures the configurational entropy of the unfolded state is very close to that of the entire ensemble of structures.

Table 3. Entropy, Enthalpy, and Free Energy of Folding

	T (K)	
	298	340
entropy of all configurations		
S (J K^{-1} mol^{-1})	2686	2886
entropy of folded configurations		
S_{fol} (J K^{-1} mol^{-1})	2183	2363
entropy of unfolded configurations		
S_{unf} (J K^{-1} mol^{-1})	2692	2892
enthalpy of all configurations		
H (kJ mol^{-1})	−229	−218
enthalpy of folded configurations		
H_{fol} (kJ mol^{-1})	−243	−225
enthalpy of unfolded configurations		
H_{unf} (kJ mol^{-1})	−223	−212
entropy of folding		
$\Delta S_{folding} = S_{fol} - S_{unf}$ (J K^{-1} mol^{-1})	−509	−529
enthalpy of folding		
$\Delta H_{folding} = H_{fol} - H_{unf}$ (kJ mol^{-1})	−20	−13
free energy of folding		
$\Delta G_{folding} = \Delta H_{folding} - T\Delta S_{folding}$ (kJ mol^{-1})	132	167
free energy of folding (from populations)		
$\Delta G'_{folding}$ (kJ mol^{-1})	2.8	2.7

[a] $\Delta G'_{folding}$ has been calculated from the populations of folded and unfolded structures with eq 4, as described in section 4.2.3. It includes peptide−solvent and solvent contributions to enthalpy and entropy, whereas $\Delta G_{folding}$ values contain only the peptide contributions to enthalpy and entropy.

With knowledge of the peptide's entropy and enthalpy (taken as the average internal interaction energy) it is possible to calculate the free energy of the peptide—with all the reservations that this concept deserves—using eq 3 (see section 4.2.3). This ill-defined free energy ignores interactions and correlations with the solvent. The calculation can be performed for the whole ensemble of structures as well as for the folded and unfolded states. Therefore, the following analysis serves us to address the suitability of implicit solvation models, which lack any type of solvent entropy, for the study of peptide or protein folding. The enthalpy, entropy, and corresponding free energy for the three sets of structures are listed in Table 3. The free energy of folding and its enthalpic and entropic components are given in the same table. The enthalpy of the various ensembles of structures follows the same lines observed for the entropy (Table 3): The enthalpy of the folded state is lower than that of the unfolded state; the enthalpies of the folded and unfolded states are lower in the simulation at 298 K than in the simulation at

340 K; at both temperatures the enthalpy of the unfolded state is slightly higher than that of the entire ensemble of structures. Upon folding, there is a bigger loss of entropy and a smaller decrease of the enthalpy at 340 K than at 298 K. This implies, independently of the actual values of the differences, that folding would be more favorable at 298 K than at 340 K. The resulting free energies of folding are 132 kJ mol^{-1} at 298 K and 167 kJ mol^{-1} at 340 K. The biggest contribution to these free-energy differences comes from the $T\Delta S$ term. Such high free energies of folding suggest that the structures shown in Figure 10C would never be sampled in a 100-ns simulation. This is, of course, nonsense. The concept of a free energy of the peptide, ignoring interactions and correlations with the rest of the system, is simply wrong, even if we cannot really discard the possibility of ΔS being overestimated by the method used to estimate it. Thus, the free energy of folding has also been calculated from the relative populations of folded and unfolded structures using eq 4 (see section 4.2.3). Again, the reference structure for the definition of the folded state is the central member structure of cluster number 1. The resulting free energies of folding include peptide, peptide solvent, and solvent contributions. The values at 298 and 340 K are very similar, 2.8 and 2.7 kJ mol^{-1}, respectively. Comparison of $\Delta G_{\text{folding}}$ and $\Delta G'_{\text{folding}}$ in Table 3 suggests that the solvent contributions to the free energy of folding must be bigger at 340 K than at 298 K. In particular, the biggest contributions are expected to be of entropic nature and to be due to peptide−solvent correlations.[8,48]

3. Conclusions

The conformational behavior of a β-hexapeptide designed to form a hairpin structure has been studied in methanol solution by NMR spectroscopy and MD simulation. A set of 15 structures from annealing MD simulations in vacuo using the program X-PLOR, with restraints derived from the NMR data, was initially selected as representative of the predominant conformation of the peptide in solution. These structures are characterized by a 10-membered turn closed by a hydrogen bond between the amino group of residue 3 and the carbonyl group of residue 4, and two extended antiparallel strands which are otherwise not within hydrogen-bonding distance. The ensembles of structures from two 100-ns unrestrained MD simulations of the β-hexapeptide in methanol solution at 298 and 340 K, respectively, have been evaluated in terms of relative agreement with the NMR data and are proposed as an alternative to the single-conformation perspective. The ensemble averages of interproton distances and ^3J values accurately reproduce the NMR-derived data involving residues 3 and 4, at the turn of the model hairpin. This is despite a relatively low population of the turn conformation: Inspection of the structures sampled in the simulations reveals that the hydrogen-bonded turn is present in about 20% of the structures at 298 K and in about 30% of the structures at 340 K. These percentages are slightly higher if structures which enclose the basic turn conformation but do not fulfill the hydrogen-bond criterion used in the analysis are also taken into account. However, the initial prediction that the two dipeptide segments of (*R*,*S*)-$\beta^{2,3}$-amino acids at the N- and C-terminal ends of the hexapeptide would adopt an antiparallel extended conformation does not seem to hold in the simulations. This could be due to the very dense packing of aliphatic side chains required for the formation of a perfect hairpin structure, which is clearly not favored by the force field. Indeed, in the X-PLOR structures the two strands are extended and antiparallel, but not within hydrogen-bonding distance either. It is not entirely clear to what extent the observations made in the simulations can be extrapolated to the conformational behavior of the β-hexapeptide in the experimental sample. On one hand, there are some discrepancies between the experimental ^3J values and the ^3J values calculated from the torsional dihedral angles explored by residues 1, 2, 5, and 6 in the simulations. Nevertheless, these discrepancies exist also between the experimental ^3J values and the ^3J values calculated from the torsional dihedral angles of the 15 X-PLOR structures, which are closer in average to those of a model hairpin structure. Furthermore, there is an intrinsic uncertainty in the transformation of torsional dihedral angles to ^3J values which is in this case difficult to estimate. On the other hand, the NOE-derived upper-bound distances between the HC$^\alpha$(2) and HC$^\beta$(5) and between the HC$^\alpha$(1) and HC$^\beta$(6) hydrogen atoms are satisfied by the ensemble of structures from the simulation at 340 K, despite the low number of ideal hairpin structures sampled.

A rather fundamental question that arises at several stages of the analysis of the trajectories from the simulations at 298 and 340 K is why the NMR data (obtained at room temperature) are reproduced more accurately by the ensemble of structures obtained at 340 K than by that obtained at 298 K. This question has two distinct components. The first one refers to the relative discrepancy between simulation and experiment at 298 K. Considering only the simulation results, the discrepancy could be due to (i) deficiencies of the model or force field, (ii) lack of convergence of the probability distribution of conformational states (insuficient sampling), or (iii) the difficulty to exactly reproduce the experimental conditions. The second component of the question refers to the temperature dependence of the probability distribution of conformational states in simulation and experiment. Unfortunately, this point cannot be adequately addressed until two other basic questions are investigated, i.e., (i) the exact relation between the absolute temperatures of experiment and simulation, which depends on the details of force field parametrization, and (ii) the temperature dependence of the NMR data collected for the β-hexapeptide.

The free energy of folding has been estimated, from conformational clustering and population analysis, to be very similar at the two simulation temperatures (2.8 kJ mol^{-1} at 298 K and 2.7 kJ mol^{-1} at 340 K). It has been shown that if the folding process was governed exclusively by the enthalpy and configurational entropy of the peptide, folding would be more favorable at 298 K than at 340 K and practically inviable at both temperatures. Therefore, interactions and correlated motions with the solvent, and within the solvent, play an essential role in the folding/unfolding equilibrium, not only making the process of folding possible, but also modifying the relative probabilities of folding at the two different temperatures. This result puts into question computational models that reduce the folding problem to free-energy landscapes of polymer chains of various levels of complexity,[49] describing the folding process in terms of only chain enthalpy and entropy (at most with an implicit polymer−solvent enthalpic term), which have become so popular in recent years and have led to the funnel-landscape theory of protein folding.[50,51]

4. Methods

4.1. NMR Structure Determination. Sample: 12 mg of β-hexapeptide (see Figure 1B) dissolved in 0.6 mL of CD_3OH. **1D-NMR** (Bruker

(48) Yu, H.-A.; Karplus, M. *J. Chem. Phys.* **1988**, *89*, 2366−2379.

(49) Hao, M.-H.; Scheraga, H. A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 184−188.

(50) Onuchic, J. N.; Wolynes, P. G.; Luthey-Schulten, Z.; Socci, N. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3626−3630.

(51) Chan, H. S.; Dill, K. A. *Proteins: Struct. Funct. Genet.* **1998**, *30*, 2−33.

AMX500): ^{1}H NMR (500 MHz); suppression of the CD_3OH signal by presaturation; 90K data points; 64 scans; 5.6 s acquisition time. ^{1}H-BB-decoupled-^{13}C NMR (125 MHz): 80K data points; 8000 transients; 1.3 s acquisition time; 45° excitation pulse; 1.0 s relaxation delay. Processed with 1.0 Hz exponential line broadening. **2D-NMR**: DQF-COSY (500 MHz, CD_3OH) with pulsed field gradients (PFG) for coherence pathway selection.[52] Acquisition: 2K (t_2) × 512 (t_1) data points; 4 scans per t_1 increment; 0.21 s acquisition time in t_2; 2.0 s relaxation delay; TPPI quadrature detection in ω_1. Processing: Zero filling and FT to 1K × 1K real/real data points after multiplication by sin^2 filter shifted by $\pi/3$ in ω_2 and $\pi/2$ in ω_1. ROESY (500 MHz, CD_3OH).[53] Acquisition: A series of 3 ROESY spectra with mixing times of 50, 100, and 150 ms was acquired; solvent suppression by presaturation; CW spin lock (3.8 kHz) between trim pulses; 4K (t_2) × 768 (t_1) data points; 32 scans per t_1 increment; 0.405 s acquisition time in t_2; other parameters identical to DQF-COSY. Processing: Zero filling and FT to 1K × 1K real/real data points after multiplication by sin^2 filter shifted by $\pi/3$ in ω_2 and cos^2 filter in ω_1. Baseline correction with 3rd degree polynomial in both dimensions.

All resonances in the ^{1}H NMR spectrum were assigned from DQF-COSY experiments (see Supporting Information). The ^3J-coupling constants between the protons of the peptide backbone could be directly obtained from the one-dimensional ^{1}H NMR spectrum (see Table 2). ROESY experiments at three different mixing times were performed to gather information about the three-dimensional structure of the β-hexapeptide. Twenty NOEs were extracted from the 150-ms ROESY spectrum (see Supporting Information and Table 1) and classified in three categories according to the estimated cross-peak volume in the contour plot. Of these 20 NOEs, 11 correspond to intraresidue effects, 6 correspond to effects between i and $(i + 1)$ residues, 2 correspond to effects between i and $(i + 3)$ residues, and 1 corresponds to effects between i and $(i + 5)$ residues. The structure determination was carried out by using X-PLOR 3.851[54] with QUANTA as the front-end. A molecular model of the β-hexapeptide was generated by using QUANTA with neutral NH_2 and COOH groups. The 20 NOEs were classified according to their estimated cross-peak volumes in three distance categories: strong, medium, and weak, with 0.3, 0.35, and 0.45 nm assigned as upper-bound distances and the van der Waals radii as lower-bound distances. These were used together with 9 restraints for backbone torsional dihedral angles (derived from the ^3J values) in simulated-annealing calculations. The simulated-annealing protocol *sa.inp*[55] of X-PLOR 3.851 was used with modified *parallhdg.pro* and *topallhdg.pro* files[54] and the following settings: 10000 steps at 700 K (time step 1.5 fs) and subsequent cooling in 5000 steps to 300 K. The NOE scale was set to 50 and a soft-square potential was used. The DIHE scale was set to 5.0. All other parameters were left unchanged. Thirty structures were generated and the 15 lowest in energy, with no violation larger than 0.01 nm of the experimentally derived distance restraints and no violation larger than 10° of the experimentally derived dihedral-angle restraints, were chosen as representative for the structure in solution (see Figure 2). We refer to them as the X-PLOR structures.

4.2. MD Simulations. The simulations and analysis were carried out by using the GROMOS96 package of programs.[56,57]

4.2.1. Molecular Model. The molecular model was derived from the GROMOS96 43A1 force field.[56] In this force field the aliphatic hydrogen atoms are treated as united atoms together with the carbon atom to which they are attached.[56,58] The molecular-topology building blocks including the force field parameters for the β-amino acid residues were constructed by analogy to the corresponding α-amino acid residues. The initial coordinates of the β-hexapeptide were taken from an extended structure (180° for all backbone torsional dihedral angles). The three ionizable groups (amino-terminal, homo-lysine-amino, and carboxy-terminal) were chosen to be protonated. For solvent methanol the standard model of the force field was taken, i.e., a three-center rigid model.[56]

The extended β-hexapeptide was placed at the center of a periodic truncated-octahedron box. The minimum distance from any peptide atom to the box wall was chosen in this initial configuration as 1.4 nm. The solvent was introduced into the box by using as a building block a cubic configuration of 216 equilibrated methanol molecules. All methanol molecules with the oxygen atom lying within 0.3 nm of a non-hydrogen peptide atom were then removed. Thus, the final system consisted of 64 peptide atoms and 4359 solvent atoms (1453 methanol molecules). Truncated-octahedron periodic boundary conditions were applied from this point onward.

A steepest-descent energy minimization of the system was performed to relax the solvent configuration. The peptide atoms were positionally restrained by using a harmonic interaction with a force constant of 250 kJ mol^{-1} nm^{-2}. Following, 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}.

4.2.2. Simulation Setup. Two 100-ns molecular dynamics simulations at 298 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 temperature was brought and maintained at the desired value by means of weak coupling to an external temperature bath.[59] The temperatures 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,[59] with isotropic scaling and a relaxation time of 0.5 ps. A value of 4.575 10^{-4} kJ^{-1} mol nm^3 was taken for the isothermal compressibility of the system in both simulations. This needs only be an approximate value, since it is used in combination with the pressure relaxation time to determine the strength of the coupling to the pressure bath. Bond lengths were constrained to ideal values[56] by using the SHAKE algorithm[60] with a geometric tolerance of 10^{-4}. A time step for the leapfrog integration scheme of 2 fs was used. The nonbonded interactions were evaluated by means of a twin-range method: The short-range van der Waals and electrostatic interactions were evaluated at every time step by using a charge-group pair list that was generated with a short-range cutoff radius of 0.8 nm. Longer-range van der Waals and electrostatic interactions—between charge groups at a distance longer than the short-range cutoff and shorter than a long-range cutoff of 1.4 nm—were evaluated every 5 time steps, at which point the pair list was also updated, and were kept unchanged between these updates. The cutoff radii were applied to the centers of geometry of the solute charge groups and to the oxygen atoms of the solvent molecules.

4.2.3. Analysis. Trajectory coordinates and energies were stored at 0.5-ps intervals and used for analysis. The energy of the system and the volume reach an equilibrium within tens of picoseconds (data not shown). Given that the total length of the trajectories is 3 to 4 orders of magnitude longer, no initial period of time was discarded as equilibration for the calculation of trajectory averages. 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-hexapeptide. No mass-weighting was used in either translational or rotational fitting, since the RMSD was used exclusively in a configurational space context. A conformational-clustering analysis was performed on a set of 10^4 peptide structures taken at 10-ps intervals from the simulation, using the backbone atom-

(52) Davis, A. L.; Laue, E. D.; Keeler, J.; Moskau, D.; Lohman, J. *J. Magn. Reson.* **1991**, *94*, 637−644.
(53) Griesinger, C.; Ernst, R. R. *J. Magn. Reson.* **1987**, *75*, 261−271.
(54) Brünger, A. T. *X-PLOR. A system for X-ray crystallography and NMR*; Yale University Press: New Haven, 1992.
(55) Nilges, M.; Kuszewski, J.; Brünger, A. T.; Hoch, J. C.; Poulsen, F. M.; Redfield, C., Ed.; Plenum Press: New York, 1991.
(56) 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. *Biomolecular simulation: The GROMOS96 manual and user guide*; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v.: Zürich, Groningen, 1996.
(57) 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. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(58) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535−547.
(59) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(60) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

positional RMSD as similarity criterion. A maximum cluster radius of 0.08 nm was chosen, which corresponds approximately to the maximum atom-positional RMSD between any pair of the 15 X-PLOR structures and to the first minimum in the distribution of atom-positional RMSDs from the central member structure of cluster number 1 (where the clustering was performed with various RMSD criteria). Note that a criterion of 0.09 nm would have been equally adequate on these two grounds and that the more stringent criterion of 0.08 nm was chosen to give the conclusions more significance. The clustering algorithm has been described in previous studies of β-peptide dynamics.[9] Hydrogen bonds were calculated with use of a geometric criterion. A hydrogen bond was thus defined by a minimum donor−hydrogen−acceptor angle of 135° and a maximum hydrogen−acceptor distance of 0.27 nm. Interproton distances derived from the experimental NOE intensities at 298 K (see section 4.1) were compared to the corresponding average effective interproton distances in the simulations. The latter were calculated by means of two different averaging procedures: (i) $\langle r^{-3}\rangle^{-1/3}$ averaging of the instantaneous distances r within the time scale of rotational relaxation of the peptide followed by $\langle r'^{-6}\rangle^{-1/6}$ averaging of the average distances $r' = \langle r^{-3}\rangle^{-1/3}$ within the time scale of the simulation, i.e., $\langle\langle r^{-3}\rangle^2\rangle^{-1/6}$,[12] and (ii) direct $\langle r^{-6}\rangle^{-1/6}$ averaging of the instantaneous distances r. As already mentioned, in the GROMOS96 43A1 force field aliphatic hydrogen atoms are treated within a united-atom model. Interproton distances involving aliphatic hydrogen atoms were thus calculated by defining virtual (for CH_1 and pro-chiral CH_2) and pseudo (for CH_3) atomic positions for these hydrogen atoms at the time of analysis.[56] For consistency, the aliphatic hydrogen atoms of the 15 X-PLOR structures have been treated likewise when comparing the average interproton distances with those in the simulations (see Table 1). Note that the average distances calculated in this way can be only substantially different from the original interproton distances in the all-atom 15 X-PLOR structures if in the latter there were strong deviations from standard geometries due to conflicting attractive restraining interactions and repulsive van der Waals interactions. The average interproton distances for the 15 X-PLOR structures used in Table 1 have been calculated, somewhat arbitrarily, with an $\langle r^{-3}\rangle^{-1/3}$ weighting. ^3J-coupling constants were calculated from simulation by using the Karplus relation,[61]

$$^3J(\mathrm{H,H}) = a\cos^2\theta + b\cos\theta + c \tag{1}$$

where a, b, and c were chosen equal to 6.4 Hz, −1.4 Hz, and 1.9 Hz, respectively, for the calculation of ^3J(HN,HC),[62] and equal to 9.5 Hz, −1.6 Hz, and 1.8 Hz, respectively, for the calculation of ^3J(HC,HC).[63] The rotational relaxation time of the peptide was estimated from the time autocorrelation function

$$C_2(t) = \langle P_2[\mathbf{u}_i(\tau)\cdot\mathbf{u}_i(\tau + t)]\rangle \tag{2}$$

where $\mathbf{u}_i(\tau)$ is the unit vector along one of the three axes of rotation of the peptide at time τ and $P_2(x)$ is the second-order Legendre polynomial. Peptide entropies were calculated from the covariance matrix of atom-positional fluctuations as formulated by Schlitter.[7,64] We refer the reader to Schäfer et al.[8] for a detailed description of the application of Schlitter's method to the calculation of the entropy of a similar β-peptide. Contributions to the entropy from overall translation and rotation of the peptide were eliminated. The translational and rotational motions around the center of mass were removed by mass-weighted translational and rotational least-squares fitting of the configurations analyzed, using the backbone atoms of all but the N- and C-terminal residues of the β-hexapeptide. Note, however, that overall rotation and internal motion are strongly coupled in a flexible molecule and their separation is, therefore, not strictly valid.[7] The free energy (G) of the peptide was estimated from the peptide enthalpy (H, taken as the average internal-interaction energy) and entropy (S) by using the thermodynamic relation

$$G = H - TS \tag{3}$$

where T is the temperature. Note, however, that it is fundamentally wrong to talk about the free energy of a part of the system.

Folding free energies were calculated as

$$\Delta G_{\mathrm{folding}} = -k_{\mathrm{B}}T\ln\frac{P_{\mathrm{folded}}}{P_{\mathrm{unfolded}}} \tag{4}$$

where k_{B} is the Boltzmann constant, T is the temperature, and p_{folded} and p_{unfolded} are the relative probabilities of finding the system in the folded and unfolded states, respectively. p_{folded} and p_{unfolded} can be approximated by the relative number of folded and unfolded structures, respectively, in the simulation. Thus, the folded state has been defined as being constituted by the ensemble of structures (out of the 2×10^5 sampled at 0.5-ps intervals) with an atom-positional RMSD for the backbone atoms of residues 2 to 5 smaller than 0.08 nm from the central member structure of cluster number 1, and the unfolded state as the ensemble of structures with an atom-positional RMSD larger than 0.12 nm from the same structure. This is based on the assumption that the central-member structure of cluster number 1 is the same (or conformationally equivalent) whether the clustering is performed on 10^4 or 2×10^5 equally spaced structures. An indication that this is a reasonable approximation is the fact that the ratio of folded versus unfolded structures defined in this way is approximately the same when calculated over the 10^4 or the 2×10^5 structures.

Acknowledgment. The authors thank B. Brandenberg for measuring NMR spectra. X.D. thanks Christine Peter for helpful discussions on NMR methodology. Financial support was obtained from the Schweizerischer Nationalfonds, project no. 21-50929.97, which is gratefully acknowledged.

Supporting Information Available: ^{1}H chemical shifts of the β-hexapeptide in CD_3OH (500 MHz) and HC^{α}/HC^{β} and CH_3 regions of the 150 ms ROESY spectrum of the β-hexapeptide in CD_3OH (500 MHz) (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA003689G

(61) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11−15.
(62) Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741−751.
(63) de Marco, A.; Llinàs, M.; Wüthrich, K. *Biopolymers* **1978**, *17*, 617−636.
(64) Schlitter, J. *Chem. Phys. Lett.* **1993**, *215*, 617−621.