Conformational Preferences of a β-Octapeptide as Function of Solvent and Force-Field Parameters

by **Alexandra Choutko** and **Wilfred F. van Gunsteren***

Physical Chemistry, Swiss Federal Institute of Technology, ETH, CH-8093 Zürich
(phone: +41 44 632 5501; fax: +41 44 632 1039; e-mail: wfvgn@igc.phys.chem.ethz.ch)

The ability to design properly folded β-peptides with specific biological activities requires detailed insight into the relationship between the amino acid sequence and the secondary and/or tertiary structure of the peptide. One of the most frequently used spectroscopic techniques for resolving the structure of a biomolecule is NMR spectroscopy. Because only signal intensities and frequencies are recorded in the experiment, a conformational interpretation of the recorded data is not straightforward, especially for flexible molecules. The occurrence of conformational and/or time averaging, and the limited amount and accuracy of experimental data hamper the precise conformational determination of a biomolecule. In addition, the relation between experimental observables with the underlying conformational ensemble is often only approximately known, thereby aggravating the difficulty of structure determination of biomolecules. The problematic aspects of structure refinement based on NMR nuclear *Overhauser* effect (NOE) intensities and ^3J-coupling data are illustrated by simulating a β-octapeptide in explicit MeOH and H_2O as solvents using three different force fields. NMR Data indicated that this peptide would fold into a 3_{14}-helix in MeOH and into a hairpin in H_2O. Our analysis focused on the conformational space visited by the peptide, on structural properties of the peptide, and on agreement of the MD trajectories with available NMR data. We conclude that *1*) although the 3_{14}-helical structure is present when the peptide is solvated in MeOH, it is not the only relevant conformation, and that *2*) the NMR data set available for the peptide, when solvated in H_2O, does not provide sufficient information to derive a single secondary structure, but rather a multitude of folds that fulfill the NOE data set.

Introduction. – The functional properties of proteins depend on their three-dimensional structures, which arise, because particular sequences of amino-acid residues in linear polypeptides fold into compact tertiary domains [1]. Although natural proteins typically require more than 50 residues to display stable three-dimensional structures, the careful choice of α-amino acids has permitted the construction of structured peptides with as few as 23 amino acid residues [2]. β-Peptides are nonnatural mimetics of α-peptides, they differ from α-peptides by the insertion of an extra C-atom between the N- and C_α-atoms of the peptide group. This offers the possibility to have a side chain not only at the C_α backbone atom, but alternatively at the C_β backbone atom. This additional variability renders the secondary and even tertiary structures of β-peptides stable even at very short chain lengths [3][4], and these peptides exhibit a large variety of folded structures, including several types of helices. Interestingly, helical structures of β-peptides are typically stable in less polar solvents such as MeOH, whereas the helices of α-peptides are stable in polar solvents such as H_2O [5]. β-Peptides [6] have attracted particular interest because of their ability to form secondary structures similar to those of natural α-peptides and their resistance

to proteases, due to the additional backbone C-atom. β-Peptides are promising candidates for pharmaceutical applications; however, the ability to design well-folding β-peptides with a specific biological activity requires detailed insight into the relationship between the β-amino acid sequence and the three-dimensional structure of a peptide. MeOH is, regarding computations, an efficient solvent because the big Me group, modeled as a united atom, gives it a lower density of atomic interaction sites than H_2O.

The two main experimental techniques, X-ray diffraction and NMR spectroscopy, widely used for structure analysis of biomolecules present some aspects which make the interpretation of experimental data of flexible molecules problematic: the occurrence of conformational- and/or time-averaging, insufficient amount of experimental data compared to molecular degrees of freedom, and insufficient accuracy of the experimental data to restrict the conformational space accessible to the molecule [7][8]. If experimental data such as X-ray reflection intensities [9] and intensities from NMR spectroscopy [10] are used for structure determination, one has to convert recorded values of such a quantity $Q(\mathbf{r}^N)$ which depends on the configuration $\mathbf{r}^N$ of the N atoms of the molecule, and the values of which are averages $< Q >_{\text{exp}}$ over time and space, into a set of configurations $\mathbf{r}^N(Q)$ [11]. Because the number of different quantities that can be measured for a molecular system is generally very much smaller than the number of degrees of freedom of the system or molecule, the problem of determining the conformational distribution $P(\mathbf{r}^N)$ that is compatible with the set of $< Q >_{\text{exp}}$ values is highly underdetermined [12][13].

To address these problems, one may employ unbiased computer simulation that does not use restraints derived from experimental data. If the results of unbiased molecular simulation agree with experimentally measured primary, *i.e.*, directly measured, not derived data, for example, nuclear *Overhauser* effect (NOE) intensities or ^3J couplings in the case of NMR experiments, the interpretation problem is basically solved. If simulated and primary experimental data do not agree, however, further investigation of the reasons for such discrepancy is required. A detailed comparison of the simulated conformational distribution with measured average values of observables that depend on this distribution showed that differences between simulation and experiment can be due to experimental inaccuracies or to deficiencies in the simulation model [12–14]. Indeed, the quality of the simulation results depends on the quality of the force field used, which describes the interactions between particles in the system. Over the years, successive GROMOS force-field parameter sets for biomolecular simulation have been introduced. The most widely used versions of this force field are the 43A1 force field [15][16] of 1996, the 45A3 force field [17] of 2001, the 53A6 force field [18] of 2004, and the 54A7 force field of 2011 [19]. Even using a nearly perfect force field, the ensemble averages of an MD trajectory may be incorrect due to insufficient sampling of particular, important parts of conformational space, which may be due to high energy barriers on the energy hypersurface of the system.

For a particular β-octapeptide [20] (*Fig. 1*), two different folds in two different solvents were proposed on the basis of data obtained in NMR experiments, *i.e.*, upper distance bounds derived from NOEs and dihedral-angle values derived from ^3J-coupling constants. This peptide was determined to adopt a 3_{14}-helical fold when solvated in MeOH and a hairpin fold when solvated in H_2O. In the present study, we

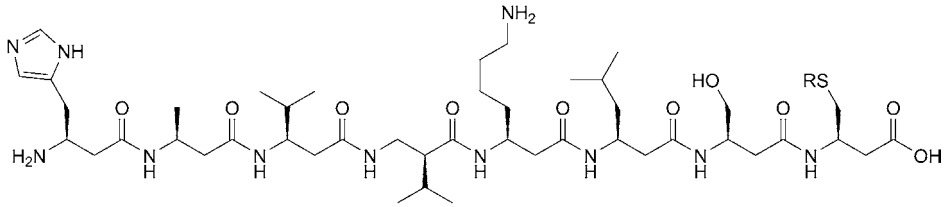

Fig. 1. *Structure of the β-octapeptide β³(S)hHis-β³(S)hAla-β³(R)hVal-β²(S)hVal-β³(S)hLys-β³(S)hLeu-β³(R)hSer-β³(R)hCys.* R = MeS in MeOH, R = H in H_2O.

conformationally interpret the primary experimental NOE data on the peptide using unrestrained MD simulations in explicit solvent, H_2O or MeOH, using the GROMOS05 biomolecular simulation software [21], with three thermodynamically calibrated GROMOS force-field parameter sets, 45A3 [17], 53A6 [18], and 54A7 [19].

Thus, for each solvent we compare the ensembles of structures from three MD simulations with the sets of 15 and 10 NMR model structures in MeOH and H_2O, before and after energy minimization, in terms of conformational space sampled by the peptide, structural properties such as H-bonding, and in terms of the level of agreement with the available experimental NMR data, *i.e.*, NOEs and ^3J-coupling constants.

Results. – We simulated a β-octapeptide (*Fig. 1*) in H_2O and MeOH as solvents with three GROMOS force fields, 45A3, 53A6, and 54A7 (*Table 1*).

Table 1. *Overview of the MD Simulations*

Simulation name	Peptide starting configuration	Substituent R at the 8th residue	Force field	Solvent	Box volume [nm³]	Number of solvent molecules
MeOH45A3	Helix	MeS	45A3	MeOH	42.7	451
MeOH53A6	Helix	MeS	53A6	MeOH	42.7	586
MeOH54A7	Helix	MeS	54A7	MeOH	42.7	613
H_2O45A3	Hairpin	H	45A3	H_2O	61.6	1787
H_2O53A6	Hairpin	H	53A6	H_2O	61.6	1938
H_2O54A7	Hairpin	H	54A7	H_2O	61.6	1952

The atom-positional root-mean-square deviations (RMSDs) of the peptide backbone atoms from the initial helical or hairpin structures are shown in *Fig. 2* as a function of simulation time for the six MD simulations. For the MeOH simulations, where the helical NMR structure was used as starting structure, the RMSD is always below 0.4 nm, shifting back and forth between longer periods at 0.1 and 0.3 nm, indicating a partial folded-unfolded conformational change. The helical fold in MeOH is reproduced by all three force fields. For the H_2O simulations, where the hairpin NMR structure was used as starting structure, the RMSDs are up to 0.5 nm for the 45A3 and 54A7 force fields, and up to 0.4 nm for the 53A6 force field. The three force fields seem to have more difficulty to reproduce the hairpin fold in H_2O that was suggested based on the NMR data. To verify this, we next compare the simulated

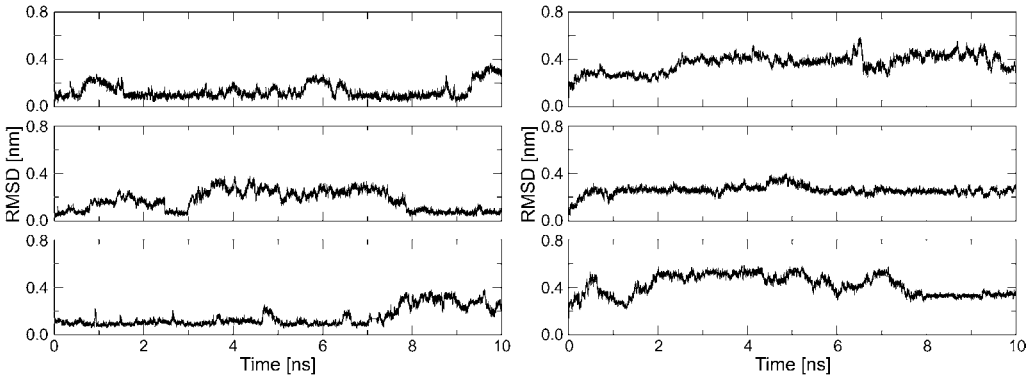

Fig. 2. *Time evolution of atom-positional RMSD with respect to the energy-minimized NMR-model starting structure of the backbone atoms* (N_H, C_O, C_α, C_β) *for the six MD simulations. Left:* for the MeOH simulations with the RMSD calculated to the energy-minimized helical NMR structure. *Right:* for the H_2O simulations with RMSD calculated with respect to the hairpin energy-minimized NMR-model starting structure. From top to bottom: 45A3, 53A6, and 54A7 force fields.

averages for measured quantities or quantities more directly related to recorded data, NOE atom–atom distances, and $^3J(HN,H_\beta)$ couplings.

To check the level of agreement between the experimental and simulated results, the average effective, *i.e.*, r^{-6}-averaged, distances were compared to the proton–proton upper-bound distances and the averaged $^3J(HN,H_\beta)$-coupling constants were calculated, for the simulation ensembles as well as for the sets of NMR model structures before and after energy-minimization (*Figs. 3–5*, and for structure 1 from the set of NMR structures in *Tables S1–S4*[1])). The upper-bound nature of NOE-derived distances implies that only violations with positive values are true violations. For the three MeOH MD simulations (*Fig. 3*), six NOE distance bounds are violated by more than 0.1 nm (up to 0.4 nm). Those six violated distances (3_{HN}–$5_{H\beta}$, $2_{H\gamma}$–$5_{H\beta}$, $4_{H\beta}$–$7_{H\beta}$, $5_{H\beta}$–$2_{H\alpha}$, $6_{H\beta}$–$3_{H\alpha}$, $8_{H\beta}$–$5_{H\alpha}$) are indicated in black in the NMR 3_{14}-helical structure (upper left of *Fig. 7*; see below). Interestingly, the same distance bounds are as well strongly violated in the sets of NMR structures energy-minimized *in vacuo* (see *Fig. 4*). These violations are induced by the energy minimization *in vacuo* of the NMR model structures using GROMOS force fields. Apparently, the simple X-PLOR-like force field used in the single-structure refinement allows structures that are of high energy in the GROMOS force fields. Those six NOE distance-bound violations are all consequences of a too short distance between the backbone atoms of residues 4 and 5, as can be seen in *Fig. 7* (see below). Upon energy minimization, residue 4 moves away from residue 5, and sizable changes in the backbone dihedral angles of residue 5 are observed. The φ, θ, and ψ torsional angles of residue 5 change from $-148°$, $91°$, and $68°$, respectively, in the NMR starting structure to $-59°$, $71°$, and $-55°$ in the energy-minimized NMR starting structure. The three H_2O MD simulations show no violation greater than 0.05 nm, except for the last four long-range NOE distance bounds (3_{HN}–$7_{H\beta}$, 2_{HN}–$8_{H\beta}$, 2_{HN}–$8_{H\gamma}$, $1_{H\gamma}$–$8_{H\beta}$), which are violated by more than 0.1 nm (up to

[1]) *Supplementary Material* may be obtained upon request from the authors.

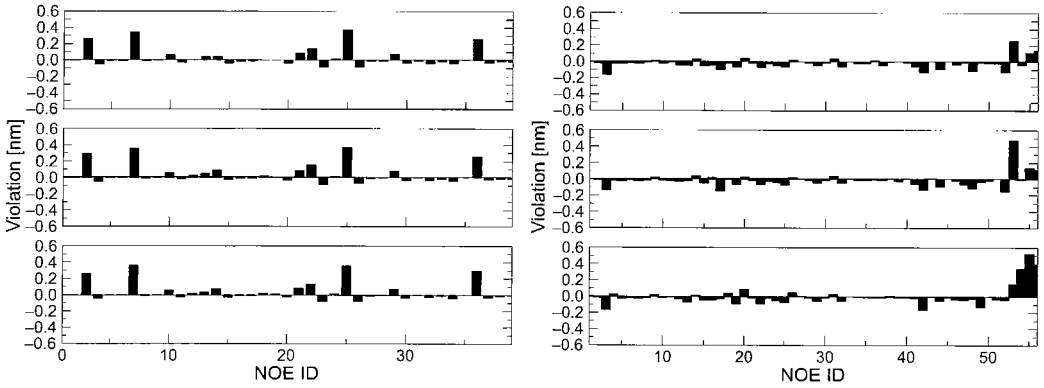

Fig. 3. *Averaged NOE upper distance-bound violations based on* $<r^{-6}>^{-1/6}$ *averages for the bounds derived from NMR experiments for the six MD simulations. Left:* for the MeOH simulations and *right:* for the H_2O simulations. From top to bottom: 45A3, 53A6, and 54A7 force fields. The NOEs can be identified in *Tables S1* and *S2*[1]).

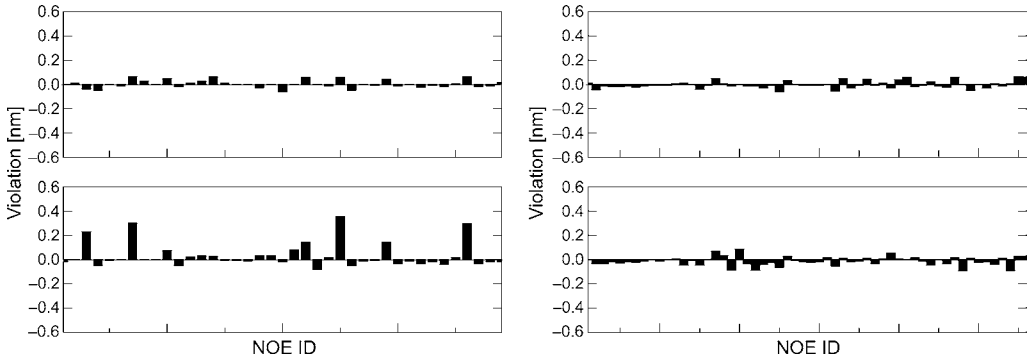

Fig. 4. *Averaged NOE upper distance-bound violations based on* $<r^{-6}>^{-1/6}$ *for the bounds derived from NMR experiments for the sets of NMR structures before* (upper panel) *and after energy minimization* (lower panel) *with force field 54A7. Left:* for the sets of 15 helical NMR structures. *Right:* for the sets of 10 hairpin NMR structures. The NOEs can be identified in *Tables S1* and *S2*[1]).

0.6 nm) in the H_2O54A7 simulation. Simulations H_2O45A3 and H_2O53A6 violate three of these distance bounds. These NOE distance bounds are not violated by more than 0.2 nm in the sets of NMR (hairpin) structures after energy minimization (see *Fig. 4*). These violations are probably due to an opening of the hairpin end, where the Zn^{2+} ion is assumed to reside, during the course of the MD simulations in H_2O. The location of the Zn^{2+} ion is experimentally unknown, and sampling of the spatial distribution of ions in the H_2O surrounding a solute is beyond our computational means. Moreover, force-field parameters for doubly charged ions are rather unreliable. Generally, such ions are strongly solvated by H_2O, staying away from the solute. Therefore, we did not include a Zn^{2+} ion in the MD simulations in H_2O. This made these simulation in addition more readily comparable to the ones in MeOH, which did not contain a Zn^{2+} ion in the experiments [20].

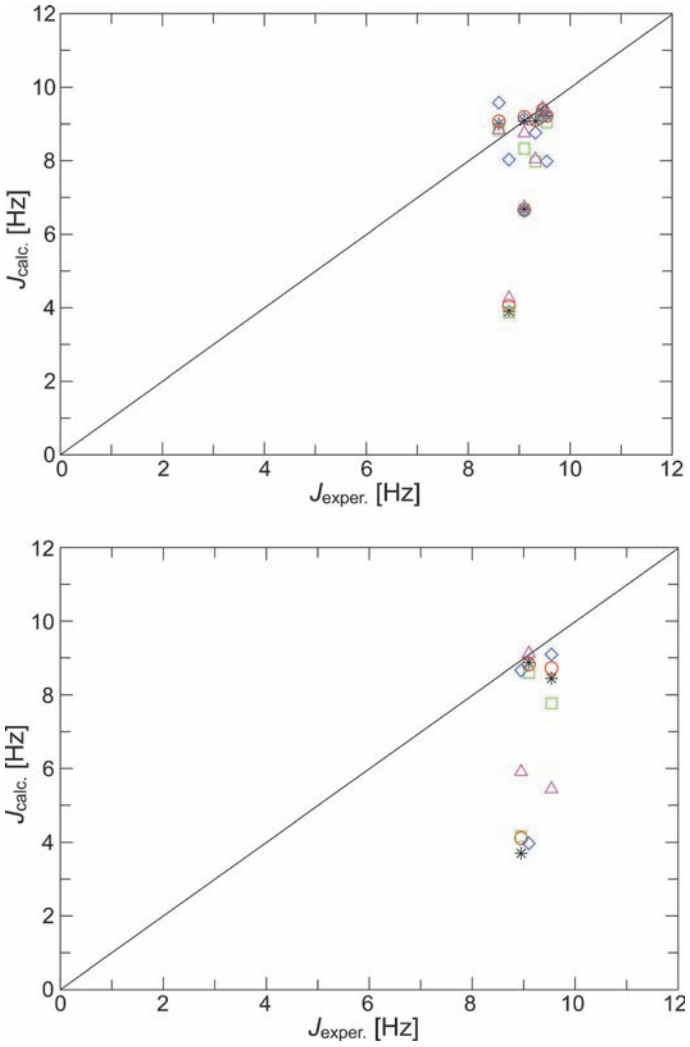

Fig. 5. *Average* ^{3}J*(HN,H$_\beta$) coupling constants as obtained from the six simulations and as calculated from the two sets of NMR model structures. Top:* for the peptide in MeOH. *Bottom:* for the peptide in H_2O. The values obtained from MD simulations using the 45A3 (black star), 53A6 (red circle), and 54A7 (green square) force fields, and from the NMR model structures before (blue diamond) and after energy minimization with force field 54A7 *in vacuo* (pink triangle) are shown. The ^3J(HN,H$_\beta$) couplings can be identified in *Tables S3* and *S4*[1]).

The average ^3J(HN,H$_\beta$)-coupling constants (*Fig. 5*) agree just as well (or as bad) with the experimentally measured values for the simulations as for the sets of NMR structures after energy minimization, except for one value in the MeOH NMR structure set (^3J(HN,H$_\beta$) coupling for residue 4) which is significantly different in the set of NMR model structures from all other structures or averages of them. These

disagreements may also be due to the inaccuracy of the *Karplus* relation and its various parametrizations that relate a ^3J coupling to a corresponding torsional angle [22].

The secondary structures of β-peptides can be characterized in terms of backbone–backbone H-bond patterns. Such H-bond populations with more than 5% occurrence in at least one of the simulations are compiled in *Table 2*. In the MeOH simulations, the observed intramolecular H-bonds are almost all of a type i to $i + 2$, characteristic for a 3_{14}-helix. For the MeOH45A3 simulation, all six H-bonds expected in a 3_{14}-helix (N(1)H–O(3), N(2)H–O(4), N(3)H–O(5), N(4)H–O(6), N(5)H–O(7), and N(6)H–O(8)) are formed. In the two other MeOH simulations, one or both terminal H-bonds are lost. The only H-bond present in simulations MeOH53A6 and MeOH54A7, that does not belong to a 3_{14}-helix but to a hairpin, is the one between the amino H-atom of residue 4 and the O-atom of residue 5. Four H-bonds (N(1)H–O(8), N(2)H–O(7), N(3)H–O(6), and N(4)H–O(5)) are expected to be present when an ideal hairpin is formed by the β-octapeptide. Only one of these H-bonds is formed during the course of the H_2O simulations. In the simulation of the β-peptide in H_2O with the force field 54A7, one observes a H-bond pattern characteristic for a 3_{14}-helix (three out of six H-bonds during more than 27% of the simulation time) and for an α-helix (three out of four H-bonds present during more than 6% of the simulation time). Still other H-bonds are also found in the H_2O simulations.

Table 2. *Occurrence, in Percentage of Simulation Time, of Backbone–Backbone H-Bonds for the Six Simulations.* Only H-bonds with a population larger than 5% are reported. H-Bonds are divided into four categories depending on whether one expects those type of H-bonds in a 3_{14}-helix, a hairpin, an α-helix, or in other types of conformations. The residue sequence number is given between parentheses.

Expected in	Donor	Acceptor	MeOH45A3	MeOH53A6	MeOH54A7	H_2O45A3	H_2O53A6	H_2O54A7
3_{14}-Helix	N(1)	O(3)	6	–	–	–	–	–
	N(2)	O(4)	56	37	55	–	–	–
	N(3)	O(5)	67	73	71	–	–	30
	N(4)	O(6)	69	63	81	–	–	27
	N(5)	O(7)	83	92	96	–	–	32
	N(6)	O(8)	12	14	–	–	–	–
Hairpin	N(1)	O(8)	–	–	–	7	31	–
	N(2)	O(7)	–	–	–	–	–	–
	N(3)	O(6)	–	–	–	–	–	–
	N(4)	O(5)	–	8	6	–	–	7
α-Helix	N(1)	O(5)	–	–	–	–	–	–
	N(2)	O(6)	–	–	–	–	–	6
	N(3)	O(7)	–	–	–	–	–	19
	N(4)	O(8)	–	–	–	–	–	13
Other	N(2)	O(8)	–	–	–	12	41	–
	N(3)	O(1)	–	–	–	–	–	7
	N(5)	O(2)	–	–	–	25	–	–
	N(5)	O(6)	–	–	–	–	–	7
	N(6)	O(2)	–	–	–	–	14	–
	N(6)	O(3)	–	–	–	–	22	–
	N(7)	O(5)	–	–	–	5	–	–

Conformational clustering analysis on the trajectories was used to characterize the conformational ensembles generated during the MD simulations [23]. A conformational cluster was defined as a set of solute structures present in the trajectory that deviate by less than 0.12 nm backbone atom-positional RMSD from each other. *Fig. 6* shows the distribution of the largest conformational clusters in the six simulations. For the MeOH simulations, the largest cluster dominates with a population ranging between 40 and 60% for the three force fields. As can be seen from the bottom part of *Fig. 7*, the conformation of the central member of the main cluster is a 3_{14}-helix in all three force fields. For the H_2O simulations, the largest cluster dominates with a population ranging between 22 and 38% for the three force fields. As can be seen from the middle part of *Fig. 7*, the conformation of the central member of the main cluster has a residual hairpin shape when simulated with force fields 45A3 and 53A6, and a helical structure when simulated with force field 54A7.

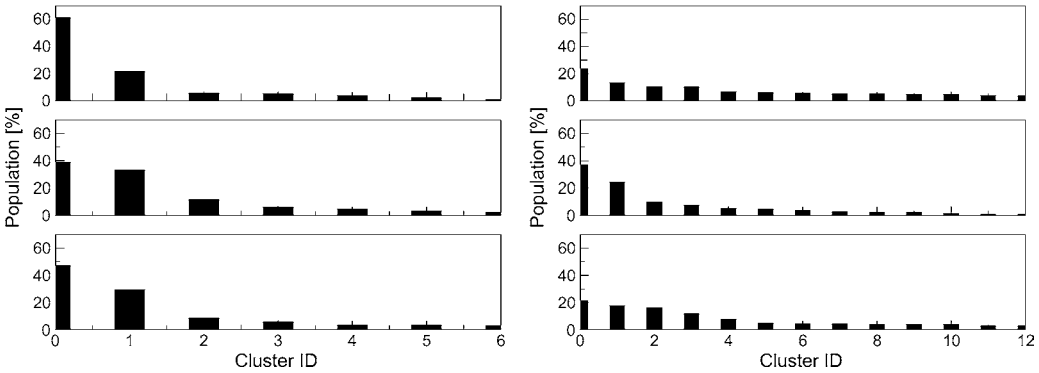

Fig. 6. *Conformational clustering analysis of the 10-ns trajectories using an RMSD cutoff of 0.12 nm for the backbone atoms.* The population in percentage per cluster is shown in decreasing order. Clusters populated by less than 5% of the trajectory are not shown. *Left:* for the MeOH simulations. *Right:* for the H_2O simulations. From top to bottom: 45A3, 53A6, and 54A7 force fields.

Discussion. – The β-peptide simulated in MeOH as a solvent with the three GROMOS force-field parameter sets 45A3, 53A6, and 54A7 starting from the 3_{14}-helical structure derived from NMR data shows a helical conformational ensemble, yet, NOE distance bounds derived from NMR experiments are in part violated by these ensembles generated from MD simulations and by the sets of energy-minimized helical NMR model structures. The information provided by experimental NOE data does not allow us to draw precise conclusions on the structure and stability of the peptide, but only on specific properties of the ensemble. When converting the recorded values of a set of quantities $< Q >_{\text{exp}}$ to a set of molecular configurations $\mathbf{r}^N(Q)$, one should be aware of the scope of the information contained in the quantity $Q(\mathbf{r}^N)$. Different structural ensembles may lead to identical values for the set of quantities $< Q >_{\text{exp}}$. Although a single structure must lead to unique NMR spectra, the inverse is not true [7]. This is because the number of experimental observables is generally too small to determine the underlying set of conformations uniquely. In view of such considerations,

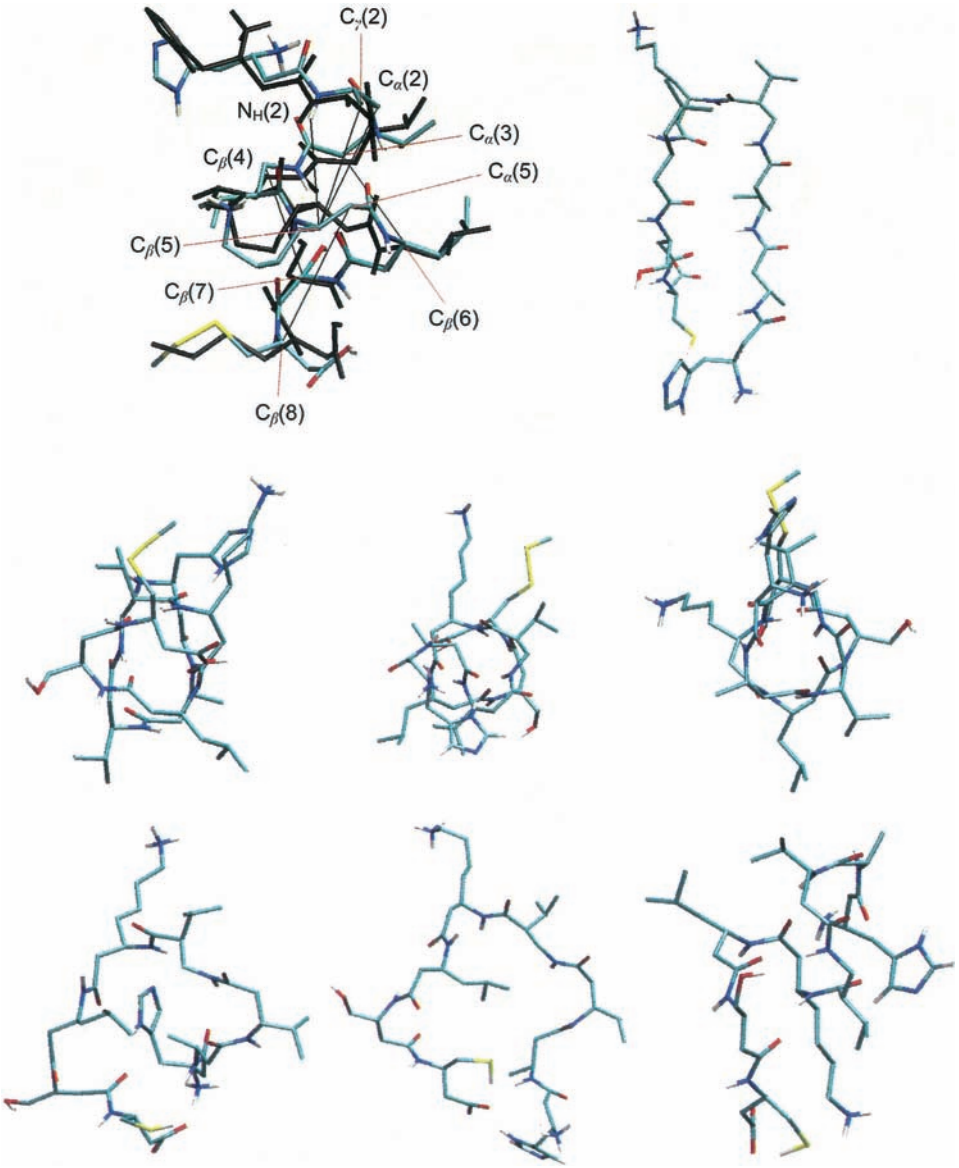

Fig. 7. *NMR-Model starting structures and most dominant conformational folds as obtained from the cluster analysis for the six MD simulations. Upper left:* helical NMR model structure 1 before (black) and after energy minimization using the 45A3 force field *in vacuo* (color). The black lines represent the NOE upper distance bounds violated after the energy minimization and in the course of three MeOH simulations. Atoms involved in those distances are labeled. *Upper right:* hairpin NMR model structure 1. *Middle*, from left to right: central member structure of the most populated cluster for simulations MeOH45A3, MeOH53A6, and MeOH54A7. *Bottom*, from left to right: central member structure of the most populated cluster for simulations H_2O45A3, H_2O53A6, and H_2O54A7.

the simulations of the β-octapeptide in MeOH suggest the presence of a mixed 3_{14}-helical/non-helical conformational ensemble, under the given experimental conditions.

The β-peptide simulated in H_2O as a solvent with the three GROMOS force-field parameter sets 45A3, 53A6, and 54A7 starting from the hairpin structure derived from NMR data shows a helical-conformational ensemble when simulated with force field 54A7, and a partial hairpin conformational ensemble when simulated with force fields 45A3 and 53A6. This observation reflects the modification of the force-field parameters going from 45A3 and 53A6 to the more recent 54A7 set which was meant to increase α-helix stability in α-peptides. The RMSDs from the NMR-derived hairpin structure are large in the three simulations, yet the NOE distance bounds derived from the NMR experiments are not or only slightly violated by these ensembles generated in MD simulations and by the energy-minimized NMR hairpin structures. A hairpin, α-helical, 3_{14}-helical, or partly random coil structure of the simulated β-peptide satisfies the NMR-derived NOE distance bounds in H_2O equally well. The hairpin structure is, therefore, underdetermined, as the NOE data set does not provide a sufficient criterion for distinguishing between these different conformations.

This study demonstrates that unbiased MD simulation using a thermodynamically calibrated force field reproduces experimental NMR data such as NOE upper-bound distances and ^3J-coupling constants just as well as a set of energy-minimized NMR model structures derived from such data, which makes MD simulation a useful tool for the structure refinement of biomolecules and for a conformational interpretation of recorded data. The available experimental data, NOE proton–proton distance bounds and $^3J(\mathrm{HN},\mathrm{H}_\beta)$ couplings, were similarly well-satisfied in the simulations using the three different force fields. So they do not provide sufficient information to distinguish between the three force fields used.

Methods. – *Force-Field Parameter Sets.* The GROMOS force field 45A3 contains 45 individual atom types [16][17] to describe *Van der Waals* interactions. The force field 53A6 reparameterized a set of polar groups against thermodynamic data, including several solvents and added eight new *Van der Waals* atom types [18]. The recently developed force field 54A7 [19] is a modification of the 53A6 force field, in which the torsional-angle energy term for the polypeptide ϕ and ψ dihedral angles was modified, the repulsive *Van der Waals* $C_{12}^{1/2}(I,I)$ parameter for the O–N pair was decreased, a new *Van der Waals* nonbonded atom type for a charged Me group was introduced, and the *Van der Waals* nonbonded interaction parameters for the Na^+ and Cl^- ions were changed.

Molecular Model. Six molecular dynamics simulations of a β-octapeptide (*Fig. 1*) in MeOH and in H_2O based on the 45A3 [17], 53A6 [18], and 54A7 [19] parameter sets (see *Table 1*) were performed using the GROMOS05 software package [21]. The backbone amino terminus and the Lys side chain were protonated in the H_2O and MeOH simulations, additionally the carboxy terminus was protonated in the MeOH simulations. The 8th residue Cys is slightly different in the MeOH (R = Me) and H_2O (R = H) simulations (see *Fig. 1*) to match experimental conditions [20]. No counter ions were used, in contrast to the experiment in H_2O, no Zn^{2+} ion was present in the simulation. The MeOH model included in the 45A3 force field is slightly different from the one used in the 53A6 and 54A7 force fields [18][24].

Simulation Setup. The coordinates of the first NMR structure from the set of 15 NMR model structures for MeOH and 10 NMR model structures for H_2O derived from NMR spectroscopic data [20] were taken as starting structures. The peptide had been determined to adopt a 3_{14}-helical conformation when solvated in MeOH and a hairpin conformation in H_2O. The peptide was solvated in a periodic, rectangular box with MeOH or H_2O as solvent.

All simulations were carried out for 10 ns at a constant temp. of 298 K and a constant pressure of 1 atm using the weak coupling algorithm [25]. Solute and solvent were separately coupled to the heat bath.

The temp. coupling time was set to 0.1 ps and the pressure coupling time to 0.5 ps. An isothermal compressibility of 4.575×10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$ was used [16].

All peptide bond lengths were kept rigid at ideal bond lengths using the SHAKE algorithm [26], as was the H···CH$_3$ distance in MeOH and the H···H distance in H$_2$O, allowing a time step of 2 fs in the leap-frog algorithm to integrate the equations of motion. Nonbonded interactions were calculated using a triple-range cutoff scheme with cutoff radii of 0.8/1.4 nm. Interactions within 0.8 nm were evaluated every time step. The intermediate range interactions were updated every fifth time step, and the long-range electrostatic interactions beyond 1.4 nm were approximated by a reaction field force [27] representing a dielectric continuum with a dielectric permittivity of 61 for the H$_2$O model [28] and either 17.7 for the MeOH model of the 45A3 force field, or 19.8 for that of the 53A6 and 54A7 force fields [24].

Analysis. Trajectory coordinates and energies stored at 0.5-ps intervals were used for analysis. Backbone atom-positional root-mean-square deviations (RMSDs) were calculated after translational superposition of the solute centers of mass and least-squares rotational fitting of atomic positions, using all backbone atoms (N_H, C_β, C_α, and C_O).

Conformational clustering was performed using the approach of *Daura et al.* [23] on the set of peptide structures taken at 10-ps intervals from the complete 10-ns trajectories of the simulations. Only the clusters that make up 95% of a trajectory were selected and counted as a function of time.

The H-Bonds were defined by a maximum H–acceptor distance of 0.25 nm and a minimum donor–H–acceptor angle of 135°. Only H-bonds with a population larger than 5% were reported.

The H···H distance bounds derived from the NOESY cross-peak intensities were compared to the corresponding averaged distances in the sets of NMR model structures before and after energy minimization, and to the average effective H···H distances in the simulations, calculated as $< r^{-6} >^{-1/6}$. The H···H distances involving aliphatic H-atoms were calculated by defining virtual (CH_1), prochiral (stereospecific CH_2), and pseudo (CH_3 and non-stereospecific CH_2) atomic positions, and the distance bounds for the latter were modified to include pseudo-atom distance-bound corrections, *i.e.*, 0.10 nm for non-stereospecific CH_2, 0.15 nm for CH_3, and 0.29 nm for non-stereospecific rotating Me groups [16][29].

^3J-Coupling constants were obtained by averaging over the simulation trajectories and the sets of NMR model structures before and after energy minimization using the *Karplus* relation [30] given in *Eqn. 1*, in which Θ is the dihedral angle between the planes defined by the atoms (H, N, C_β) and the atoms (N, C_β, H_β),

$$^3J(\mathrm{HN,H}_\beta) = a\cos^2\Theta + b\cos\Theta + c. \quad (1)$$

The parameters *a*, *b*, and *c* were chosen to be 6.4, 1.4, and 1.9 Hz, resp. [31].

The coordinates of the NMR model structures were converted into GROMOS format in the following way: the coordinates for H-atoms not present were generated by geometric means according to the topological requirements [16]. The configurations were relaxed by energy minimization *in vacuo* before being used for analysis.

The authors would like to thank *Zrinka Gattin* for many discussions and her help in setting up the project. This work was financially supported by the *National Center of Competence in Research* (NCCR) in *Structural Biology* by the *Swiss National Science Foundation* (grant No. 200020-137827), and by the *European Research Council* (grant No. 228076), which are gratefully acknowledged.

REFERENCES

[1] S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173.
[2] M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342.
[3] R. P. Cheng, *Curr. Opin. Struct. Biol.* **2004**, *14*, 512.
[4] D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893.
[5] J. R. Allison, M. Müller, W. F. van Gunsteren, *Prot. Sci.* **2010**, *19*, 2186.
[6] D. Seebach, J. L. Matthews, *Chem. Commun.* **1997**, 2015.

[7] A. Glättli, W. F. van Gunsteren, *Angew. Chem., Int. Ed.* **2004**, *43*, 6312.
[8] B. Zagrovic, W. F. van Gunsteren, *Proteins* **2006**, *63*, 210.
[9] M. F. C. Ladd, R. A. Palmer, 'Structure Determination by X-Ray Crystallography', Plenum Press, New York, 1985.
[10] K. Wüthrich, 'NMR of Proteins and Nucleic Acids', John Wiley & Sohns, New York 1986.
[11] W. F. van Gunsteren, R. M. Brunne, P. Gros, R. C. van Schaik, C. A. Schiffer, 'Methods in Enzymology. Vol. 239: Nuclear Magnetic Resonance', Academic Press, New York, 1994, p. 619.
[12] W. F. van Gunsteren, J. Dolenc, A. E. Mark, *Curr. Opin. Struct. Biol.* **2008**, *18*, 149.
[13] Z. Gattin, J. Schwartz, R. I. Mathad, B. Jaun, W. F. van Gunsteren, *Chem. – Eur. J.* **2009**, *15*, 6389.
[14] D. Trzesniak, A. Glättli, B. Jaun, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2005**, *127*, 14320.
[15] X. Daura, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **1998**, *19*, 535.
[16] 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, ETH Zürich, 1996.
[17] L. D. Schuler, X. Daura, W. F. van Gunsteren, *J. Comput. Chem.* **2001**, *22*, 1205.
[18] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656.
[19] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, W. F. van Gunsteren, *Eur. Biophys. J.* **2011**, *40*, 843.
[20] G. Lelais, D. Seebach, B. Jaun, R. I. Mathad, O. Flögel, F. Rossi, M. Campo, A. Wortmann, *Helv. Chim. Acta* **2006**, *89*, 361.
[21] M. Christen, P. H. Hünenberger, D. Bakowies, R. Baron, R. Bürgi, D. P. Geerke, T. N. Heinz, M. A. Kastenholz, V. Kräutler, C. Oostenbrink, C. Peter, D. Trzesniak, W. F. van Gunsteren, *J. Comput. Chem.* **2005**, *26*, 1719.
[22] D. Steiner, J. R. Allison, W. F. van Gunsteren, *J. Biomol. NMR* **2012**, *53*, 223.
[23] X. Daura, W. F. van Gunsteren, A. E. Mark, *Proteins* **1999**, *34*, 269.
[24] R. Walser, A. E. Mark, W. F. van Gunsteren, M. Lauterbach, G. Wipff, *J. Chem. Phys.* **2000**, *112*, 10450.
[25] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
[26] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
[27] I. G. Tironi, R. Sperb, P. E. Smith, W. F. van Gunsteren, *J. Chem. Phys.* **1995**, *102*, 5451.
[28] T. N. Heinz, W. F. van Gunsteren, P. H. Hünenberger, *J. Chem. Phys.* **2001**, *115*, 1125.
[29] K. Wüthrich, W. Braun, M. Billeter, *J. Mol. Biol.* **1983**, *169*, 949.
[30] M. Karplus, *J. Chem. Phys.* **1959**, *30*, 11.
[31] A. Pardi, M. Billeter, K. Wüthrich, *J. Mol. Biol.* **1984**, *180*, 741.

Received April 4, 2012