On the Use of a Supramolecular Coarse-Grained Model for the Solvent in Simulations of the Folding Equilibrium of an Octa-β-peptide in MeOH and H_2O

by **Wei Huang**, **Niels Hansen**[1]), and **Wilfred F. van Gunsteren***

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, CH-8093 Zürich
(e-mail: wfvgn@igc.phys.chem.ethz.ch)

Molecular dynamics (MD) simulation can give a detailed picture of conformational equilibria of biomolecules, but it is only reliable if the force field used in the simulation is accurate, and the sampling of the conformational space accessible to the biomolecule shows many (un)folding transitions to allow for precise averages of observable quantities. Here, the use of coarse-grained (CG) solvent MeOH and H_2O models to speed up the sampling of the conformational equilibria of an octa-β-peptide is investigated. This peptide is thought to predominantly adopt a 3_{14}-helical fold when solvated in MeOH, and a hairpin fold when solvated in H_2O on the basis of the NMR data. Various factors such as the chirality of a residue, a force-field modification for the solute, coarse-graining of the solvent model, and an extension of the nonbonded interaction cut-off radius are shown to influence the simulated conformational equilibria and the agreement with the experimental NMR data for the octa-β-peptide.

Introduction. – Molecular dynamics (MD) simulation is an efficient method to study the conformational equilibrium of biomolecules at the atomic level of resolution [1–3]. It yields a detailed picture of the motion of biomolecules, but this is only reliable, if the force field used in the simulation is accurate, and the sampling of the conformational space accessible to the molecules at the given thermodynamic state point is sufficient to obtain precise averages for non-observable or observable quantities characterizing the conformational equilibrium. β-Peptides are non-natural polypeptides which exhibit a marked tendency to form stable, well-defined secondary structure, even at much lower amino-acid sequence lengths than those needed in α-peptides [4–8]. This feature renders them ideal cases to study the folding process and evaluate the force field used based on sufficient (un)folding transitions using MD simulations [9–13]. Although the sampling of conformational space can be reasonably thorough when β-peptides are simulated, any reduction of computational effort will allow increasing the sampling. Here, we report on the use of coarse-grained (CG) supramolecular solvent models to sample the conformational equilibria of an octa-β-peptide [14] (*Fig. 1*) in MeOH and H_2O. This octa-β-peptide was found to predominantly adopt a 3_{14}-helical fold when solvated in MeOH, and hairpin-like conformations when solvated in H_2O on the basis of data obtained in NMR experiments [14]. Using the supramolecular CG solvent, the simulations should correctly reproduce the

[1]) Present address: Institute of Thermodynamics and Thermal Process Engineering, University of Stuttgart, DE-70569 Stuttgart.

properties of the conformational ensembles as determined by NMR spectroscopy, and as obtained from simulations using fine-grained (FG) atomic level of resolution solvent models, provided that the latter agree with experiment. Polarizable supramolecular CG solvent models containing two interaction sites, for H_2O representing five H_2O molecules [15], and for MeOH representing four MeOH molecules [16], have been developed recently. Due to their sizeable reduction of the number of solvent interaction sites, they allow a reduction of the computational effort by more than an order of magnitude. Four proteins have been simulated in CG H_2O [17][18] and three β-peptides in CG MeOH [16]. The results of simulations in pure CG solvent showed that they could not recover all properties of the polypeptide solutes as obtained from fully (solute and solvent) atomistic simulations, because of the lack of H-bonding partners in the supramolecular beads of the CG solvent. To solve this problem, a FG atomic level of resolution solvent layer was put around the solute before solvating the solute plus FG solvent layer into CG solvent. This use of CG solvent did speed up the simulations without perturbing the structural properties of the proteins and peptides significantly [16][18].

The octa-β-peptide of interest (*Fig. 1*) was already simulated before [19], but inadvertently its 5th residue had (*R*) configuration at the C_β-atom instead of (*S*) as in the NMR experiment. Therefore, we repeated the simulation of [19] with 5Lys(*S*) and analyzed the influence of changing the configuration of 5Lys on the conformational ensemble. In [19], the GROMOS force fields 45A3 of 2001 [20], 53A6 of 2004 [21], and 54A7 of 2011 [22] were used. It turned out that the GROMOS 54A7 force field did not reproduce secondary structure preferences well enough for β-peptides, which led to

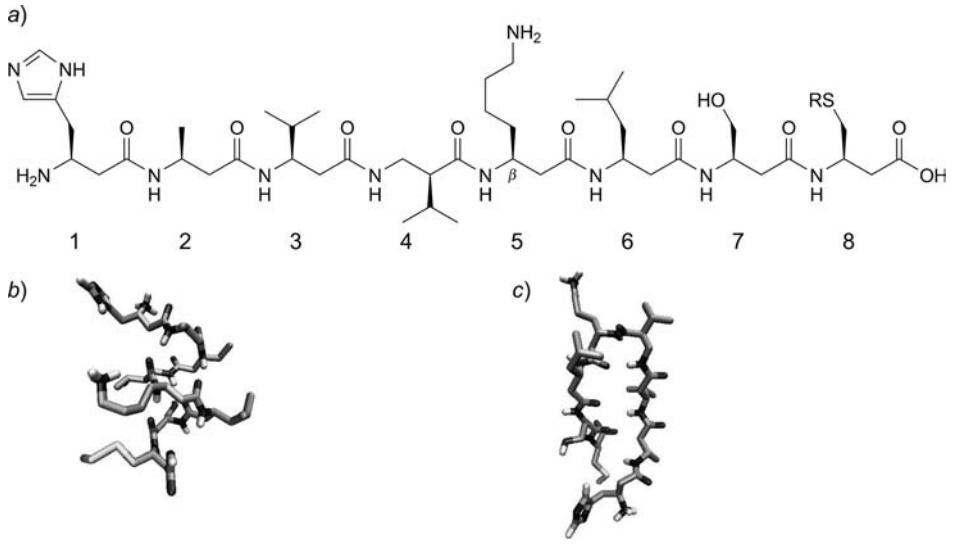

Fig. 1. a) *Structure of octa-β-peptide β³(S)hHis-β³(S)hAla-β³(R)hVal-β²(S)hVal-β³(S)hLys-β³(S)hLeu-β³(R)hSer-β³(R)hCys.* R = MeS in MeOH, R = H in H_2O [14]. b) 3_{14}*-Helical model structure* (MeOH) *derived from NMR data of the peptide in MeOH* [14]. c) *β-Hairpin model structure* (H_2O) *derived from NMR data of the peptide in* H_2O [14].

Table 1. *Overview of the MD Simulations*

Configuration[a])	Solvent	Force field	Simulation name	No. Solv.[b])	Cut-off [nm]	Efficiency[c]) [ns/h]
(*R*)	FG MeOH	54A7	$R^{54A7}_{FG\_MeOH}$	1663	0.8/1.4	0.62
	FG H_2O		$R^{54A7}_{FG\_H_2O}$	1860		0.30
(*S*)	FG MeOH	54A7	$S^{54A7}_{FG\_MeOH}$	1606	0.8/1.4	0.67
	FG H_2O		$S^{54A7}_{FG\_H_2O}$	1866		0.30
	FG MeOH	54A7β	$S^{54A7\beta}_{FG\_MeOH}$	1661	0.8/1.4	0.62
	FG H_2O		$S^{54A7\beta}_{FG\_H_2O}$	1865		0.30
	CG MeOH	54A7β	$S^{54A7\beta}_{CG\_MeOH}$	329	1.4/2.0	5.0
	CG H_2O		$S^{54A7\beta}_{CG\_H_2O}$	496		2.5
	CG MeOH with FG MeOH layer	54A7β	$S^{54A7\beta}_{FG\_layer\_MeOH}$	444/330[d])	1.4/2.0	0.91
	CG H_2O with FG H_2O layer		$S^{54A7\beta}_{FG\_layer\_H_2O}$	740/388[d])		0.36

[a]) Configuration at the C_β-atom of the 5th residue. [b]) Number of solvent molecules (FG) or CG beads. [c]) Run on eight CPUs using MPI parallelization. [d]) FG Molecules/CG beads.

a revision of it, the GROMOS 54A7β force field [23]. Here, we compare the effect of this small force-field change upon the conformational ensembles of the octa-β-peptide. An overview of various simulations is given in *Table 1*.

In [19], the octa-β-peptide was simulated in H_2O without a Zn^{2+} ion present, although such an ion was assumed to be present in the system for which NMR data had been recorded. So, we also performed an MD simulation of the octa-β-peptide in H_2O with a Zn^{2+} ion in order to see whether this ion influences the conformational ensemble, *quod non*.

Results and Discussion. – *Influence of Inverting the Configuration of* ^{5}*Lys.* The atom-positional root-mean-square deviations (RMSDs) of the backbone atoms of residues 2–7 with respect to a hairpin structure and a 3_{14}-helical structure are shown in *Fig. 2* for MD simulations in MeOH and in H_2O using the GROMOS 54A7 force field with the configuration at the C_β-atom of 5th residue being (*R*) or (*S*). From *Panels a–d*, we can see that the simulation of the peptide with (*S*) configuration at the C_β-atom of the 5th residue in MeOH (simulation $S^{54A7}_{FG\_MeOH}$) sampled much more 3_{14}-helical structures than simulation $R^{54A7}_{FG_MeOH}$with (*R*) chirality. This is also reflected in populations of intramolecular H-bonds (see *Table 2*). For simulations in H_2O (*Fig. 2*, *Panels e–h*), no significant difference is found between simulations $R^{54A7}_{FG\_H_2O}$ and $S^{54A7}_{FG\_H_2O}$. Hairpin structures were rarely sampled in either simulation. The hairpin-type H-bonds were only formed in the middle of the hairpin (N(3)·O(6), N(4)·O(5)) for a small fraction (5%) in simulation $R^{54A7}_{FG\_H_2O}$, while no hairpin type H-bonds appear in simulation $S^{54A7}_{FG\_H_2O}$. Instead, 3_{14}-helical structures were sampled in simulations $R^{54A7}_{FG\_H_2O}$ and $S^{54A7}_{FG\_H_2O}$.

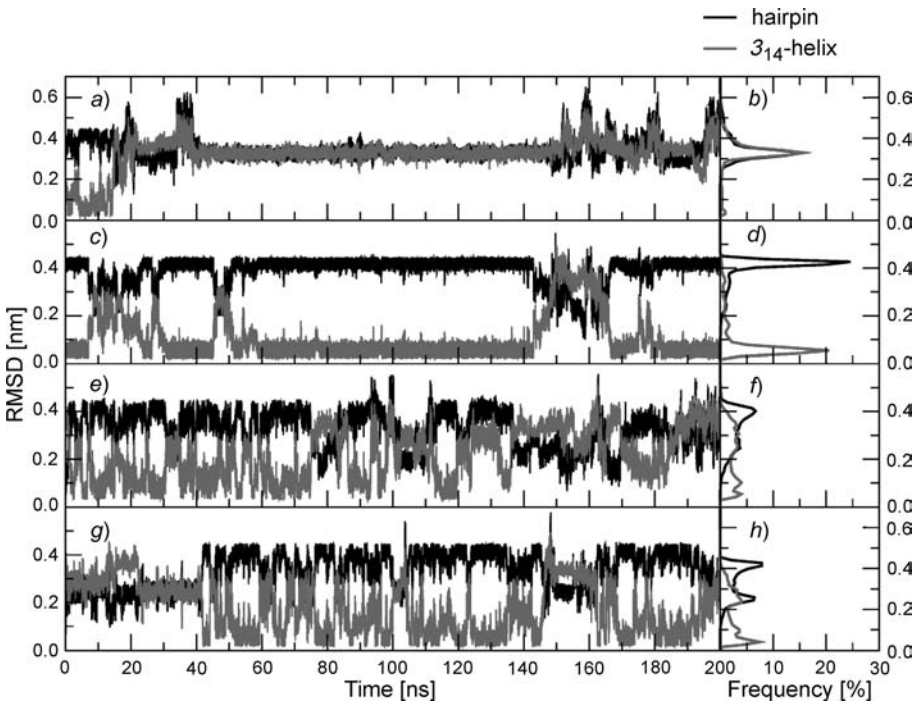

Fig. 2. *Time evolution* (left panels) *and distribution* (right panels) *of the atom-positional RMSD with respect to a hairpin* (black lines) *and with respect to a* 3_{14}*-helix* (gray lines) *for the backbone atoms of residues 2–7 in 200-ns MD simulations in MeOH* (*Panels a–d*) *and in* H_2O (*Panels e–h*) *using the GROMOS 54A7 force field with the configuration at the* C_β*-atom of 5th residue being* (R) or (S). *a*), *b*) $R^{54A7}_{FG\_MeOH}$; *c*), *d*) $S^{54A7}_{FG\_MeOH}$; *e*), *f*) $R^{54A7}_{FG\_H_2O}$; *g*), *h*) $S^{54A7}_{FG\_H_2O}$. The simulation names are specified in *Table 1*.

The H-atom–H-atom NOE distance-bound violations and ^3J values calculated from simulations $R^{54A7}_{FG\_MeOH}$, $S^{54A7}_{FG\_MeOH}$, $R^{54A7}_{FG\_H_2O}$ and $S^{54A7}_{FG\_H_2O}$ are compared with experimental data obtained for the peptide with (*S*) configuration at the 5th residue [14] in *Fig. 3*. The NOE distance-bound violations in MeOH are much smaller in simulation $S^{54A7}_{FG\_MeOH}$ than in simulation $R^{54A7}_{FG\_MeOH}$ (gray lines, upper panels). The (*R*) configuration at residue 5 disfavors 3_{14}-helical structures and lengthens the corresponding NOE H-atom–H-atom distances. The agreement of ^3J coupling constants with experimental data in MeOH is also better for simulation $S^{54A7}_{FG_MeOH}$than for simulation $R^{54A7}_{FG\_MeOH}$. Note that there are two H-atoms bound to the C_β-atom of residue 4Val(*S*), so $^3J(H_N,H_\beta)$ value for this residue was calculated for each H-atom, *i.e.*, using two phase shift values, $-60°$ (gray circles in *Figs. 3*, *5*, *7*, and S2[2])) and 60° (gray diamonds in *Figs. 3*, *5*, *7*, and S2), separately. In H_2O, no significant difference in agreement with experimental NMR data is observed between simulations $R^{54A7}_{FG\_H_2O}$ and $S^{54A7}_{FG\_H_2O}$ (black lines, lower panels).

---

[2]) *Supplementary Material* is available from the corresponding author.

Table 2. *Occurrence, in Percent of Simulation Time, of Backbone–Backbone H-Bonds for the Five Simulations in MeOH.* Only the H-bonds with a population larger than 5% are reported. The H-bonds are separated into three categories depending on whether they belong to a 3_{14}-helix, a hairpin, or to other types of conformations. The residue sequence number is given in parentheses.

Conformation	Donor	Acceptor	$R^{54A7}_{FG\_MeOH}$	$S^{54A7}_{FG\_MeOH}$	$S^{54A7\beta}_{FG\_MeOH}$	$S^{54A7\beta}_{CG\_MeOH}$	$S^{54A7\beta}_{FG\_layer\_MeOH}$
3_{14} Helix	N(1)	O(3)	–	10	11	65	18
	N(2)	O(4)	5	66	18	21	86
	N(3)	O(5)	5	68	18	17	94
	N(4)	O(6)	5	80	24	44	92
	N(5)	O(7)	7	86	30	56	77
	N(6)	O(8)	34	38	18	47	60
Hairpin	N(1)	O(8)	–	–	–	6	–
	N(2)	O(7)	–	–	–	–	–
	N(3)	O(6)	–	–	10	–	–
	N(4)	O(5)	–	–	47	–	–
Other	N(1)	O(2)	–	–	–	51	–
	N(1)	O(4)	–	–	–	25	–
	N(1)	O(6)	26	–	–	–	–
	N(1)	O(7)	–	–	–	24	–
	N(2)	O(5)	50	–	–	35	–
	N(4)	O(2)	7	–	–	–	–
	N(5)	O(2)	61	–	23	–	–

Since the configuration at the C_β-atom of 5th residue of the octa-β-peptide is (*S*) according to the experimental data [14], the subsequent simulations were all run with this topology.

Influence of a Force-Field Modification. Four different torsional dihedral-angle types had been introduced [22] for the backbone φ- and ϕ-torsional angles in the GROMOS 54A7 force field to redress the tendency of the previous GROMOS force field 53A6 [21] to slightly destabilize α-helical structures in proteins. But, it was found that the new backbone φ- and ϕ-torsional-angle energy terms of the 54A7 force field, which were parametrized based on fitting to a large set of high-resolution crystal structures of α-proteins, are not suitable to β-peptides [23]. Thus, the assignment of these 54A7 torsional-angle terms and parameters for β-peptides was reverted to the 53A6 one and is denoted as the 54A7β parameter set [23]. Here, we compare the results of simulations of the octa-β-peptide using the GROMOS 54A7 and 54A7β force fields.

The atom-positional RMSDs of the backbone atoms of residues 2 – 7 with respect to a 3_{14}-helical and a hairpin structure for MD simulations in MeOH and in H_2O using the GROMOS 54A7 and 54A7β force fields are shown in *Fig. 4*. The 3_{14}-helical structure in MeOH is less stable using the 54A7β force field than using the 54A7 force field. This is also reflected in the populations of the H-bonds characteristic for a 3_{14}-helix (*Table 2*). This is because the GROMOS 54A7β force field decreases the H-bond strength between N–H and C=O groups in the polypeptide backbone compared with the 54A7 force field. The simulations in H_2O using either the 54A7 or the 54A7β force field did not sample complete hairpin structures. Hairpin-type H-bonds were only formed in the middle of the hairpin (N(4) · O(5)) in simulation $S^{54A7\beta}_{FG\_H_2O}$, while no hairpin type H-

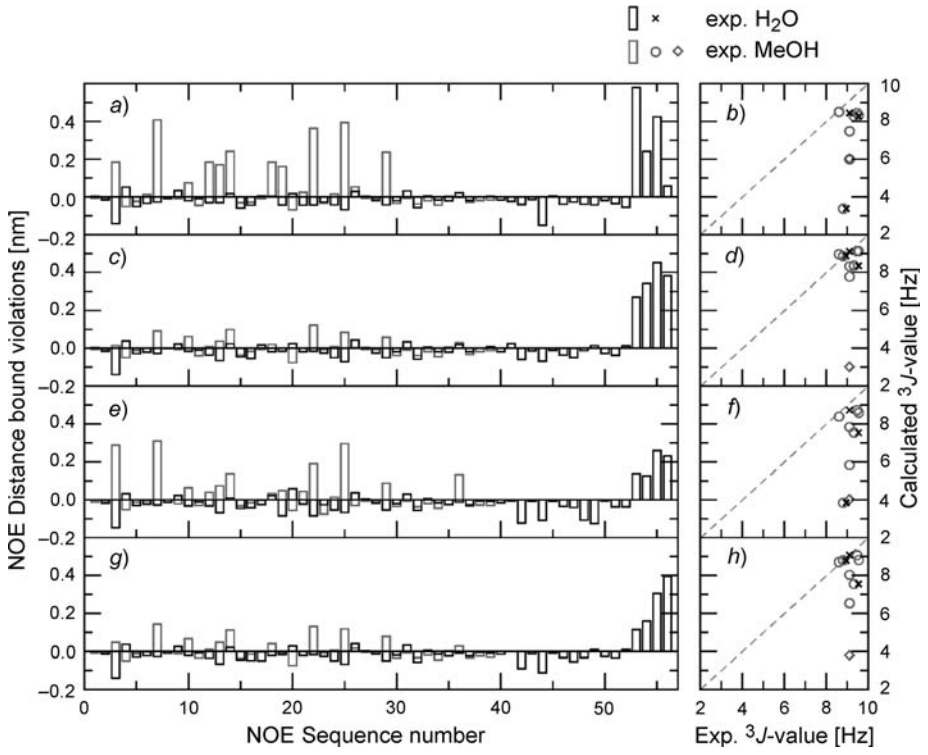

Fig. 3. *Comparison of* $\langle r^{-6}\rangle^{-1/6}$*-averaged NOE distance-bound violations* (left panels) *and average* $^3J(HN,H_\beta)$*-values* (right panels) *as obtained from simulations in MeOH* (*Panels a–d*) *or* H_2O (*Panels e–h*), *and experimental data* [14] *in* H_2O (black bars and crosses) *or in MeOH* (grey bars, circles, and diamonds). The simulations were carried out for 20 ns using the GROMOS 54A7 force field with (*R*)- or (*S*)-configuration at the C_β-atom of the 5th residue. For the specification of the NOE atom pairs and the $^3J(HN,H_\beta)$ values, we refer to Tables S1–S4 in [19]. *a*), *b*) $R^{54A7}_{FG\_MeOH}$; *c*), *d*) $S^{54A7}_{FG\_MeOH}$; *e*), *f*) $R^{54A7}_{FG\_H_2O}$; *g*), *h*) $S^{54A7}_{FG\_H_2O}$. The simulation names are specified in *Table 1*.

bonds were found in simulation $S^{54A7}_{FG\_H_2O}$ (see *Table 3*). Instead, simulation $S^{54A7}_{FG\_H_2O}$ sampled much more 3_{14}-helical structures than simulation $S^{54A7\beta}_{FG\_H_2O}$.

The H-atom–H-atom NOE distance-bound violations and ^3J values calculated from simulations $S^{54A7}_{FG\_MeOH}$, $S^{54A7ß}_{FG\_MeOH}$, $S^{54A7}_{FG\_H_2O}$, and $S^{54A7\beta}_{FG\_H_2O}$ are compared with experimental data [14] in *Fig. 5*. Since the 3_{14}-helical structure is less stable in simulation $S^{54A7\beta}_{FG\_MeOH}$ than in $S^{54A7}_{FG\_MeOH}$, the NOE-bound violations are slightly bigger (*Fig. 5*, *a* and *c*, gray lines). However, ^3J values calculated from simulation $S^{54A7\beta}_{FG\_MeOH}$ agree with experimental data better than those calculated from simulation $S^{54A7}_{FG\_MeOH}$ except the $^3J(H_N,H_\beta)$ value for the 4th residue 4Val(*S*) (*Fig. 5*, *b* and *d*). For simulations in H_2O (lower panels), there are four large violations (larger than 0.1 nm) in simulation $S^{54A7}_{FG\_H_2O}$, *i.e.*, No. 53, $H_N(3)$–$H_\beta(7)$; No. 54, $H_N(2)$–$H_\beta(8)$; No. 55, $H_N(2)$–$H_\gamma(8)$; and No. 56, $H_\gamma(1)$–$H_\beta(8)$ (black lines). Violation No. 54 almost disappears in simulation $S^{54A7\beta}_{FG\_H_2O}$, and violations No. 55 and 56 are smaller in this simulation than in simulation $S^{54A7}_{FG\_H_2O}$. The ^3J values

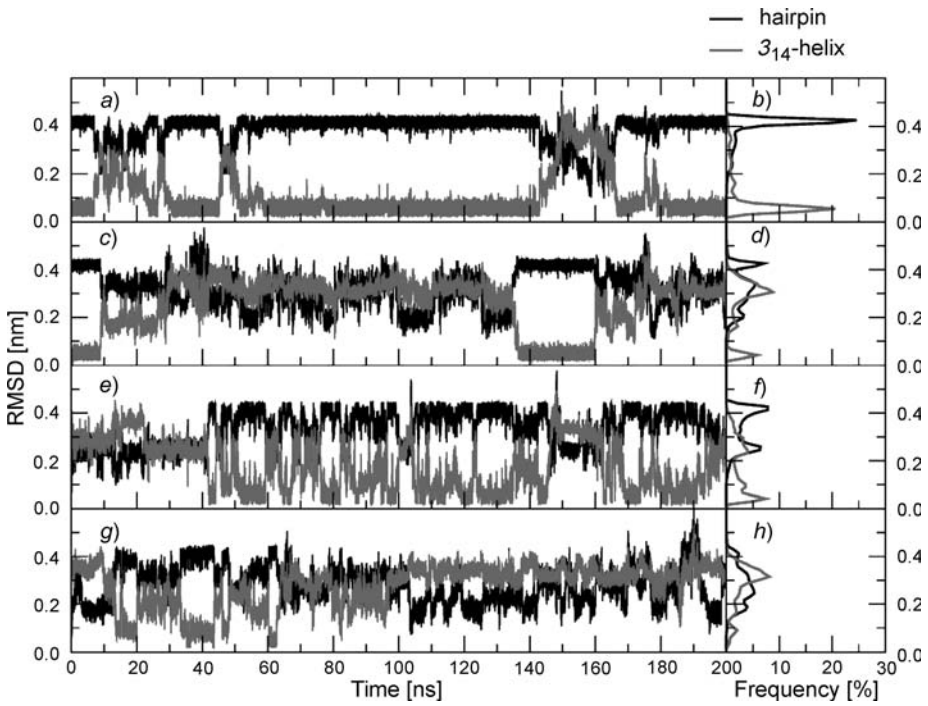

Fig. 4. *Time evolution* (left panels) *and distribution* (right panels) *of the atom-positional RMSD with respect to a hairpin* (black lines) *and with respect to a* 3_{14}*-helix* (gray lines) *for the backbone atoms of residues 2–7 in 200-ns MD simulations in MeOH* (*Panels a–d*) *and in* H_2O (*Panels e–h*) *using the GROMOS 54A7 and 54A7β force fields. a*), *b*) $S^{54A7}_{FG\_MeOH}$; *c*), *d*) $S^{54A7\beta}_{FG\_MeOH}$; *e*), *f*) $S^{54A7}_{FG\_H_2O}$; *g*), *h*) $S^{54A7\beta}_{FG\_H_2O}$. The simulation names are specified in *Table 1*.

calculated from simulation $S^{54A7\beta}_{FG\_H_2O}$ agree with experimental data better than those calculated from simulation $S^{54A7}_{FG\_H_2O}$ (black crosses).

Influence of the Solvent Model. The properties of the octa-β-peptide with different solvent models in the simulations were also compared. The octa-β-peptide was solvated in FG and CG solvent, and CG solvent with a FG solvent layer surrounding the peptide, separately (*Table 1*). The GROMOS force field 54A7β was used in these simulations. The atom-positional RMSDs of the backbone atoms of residues 2–7 with respect to a hairpin and a 3_{14}-helical structure for MD simulations in MeOH and in H_2O are shown in *Fig. 6*. When the octa-β-peptide was solvated in pure CG MeOH solvent (simulation $S^{54A7\beta}_{CG\_MeOH}$), the 3_{14}-helical structure was stable for *ca.* 120 ns, then unfolded (*Fig. 6,c*). But, when a FG MeOH layer was put around the peptide, and the layer together with the peptide was solvated in CG MeOH (simulation $S^{54A7\beta}_{FG\_layer\_MeOH}$), the 3_{14}-helical structure was stable during the whole 200-ns simulation (*Fig. 6,e*). This is also reflected by the populations of 3_{14}-helical H-bonds (*Table 2*) which are larger in simulation $S^{54A7\beta}_{FG\_layer\_MeOH}$ than in simulation $S^{54A7\beta}_{FG\_MeOH}$. For the octa-β-peptide solvated in pure CG H_2O (simulation $S^{54A7\beta}_{CG\_H_2O}$), neither 3_{14}-helical nor hairpin structures were sampled (*Fig. 6,i*). Upon insertion of an FG H_2O layer (simulation $S^{54A7\beta}_{FG\_layer\_H_2O}$), some 3_{14}-

Table 3. *Occurrence, in Percent of Simulation Time, of Backbone–Backbone H-Bonds for the Five Simulations in H_2O.* Only the H-bonds with a population larger than 5% are reported. The H-bonds are separated into three categories depending on whether they belong to a 3_{14}-helix, a hairpin, or to other types of conformations. The residue sequence number is given in parentheses.

Conformation	Donor	Acceptor	$R^{54A7}_{FG\_H_2O}$	$S^{54A7}_{FG\_H_2O}$	$S^{54A7\beta}_{FG\_H_2O}$	$S^{54A7\beta}_{CG\_H_2O}$	$S^{54A7\beta}_{FG\_layer\_H_2O}$
3_{14} Helix	N(1)	O(3)	–	6	10	–	5
	N(2)	O(4)	12	26	–	–	20
	N(3)	O(5)	35	31	12	–	31
	N(4)	O(6)	41	54	29	–	31
	N(5)	O(7)	42	68	30	–	28
	N(6)	O(8)	7	19	24	–	24
Hairpin	N(1)	O(8)	–	–	–	115	–
	N(2)	O(7)	–	–	–	–	–
	N(3)	O(6)	5	–	–	–	9
	N(4)	O(5)	5	–	29	–	19
Other	N(1)	O(2)	–	–	–	18	–
	N(2)	O(6)	5	11	–	–	7
	N(2)	O(8)	–	–	23	16	–
	N(3)	O(7)	–	–	8	–	10
	N(3)	O(8)	–	–	14	–	–
	N(4)	O(8)	6	–	–	–	–
	N(5)	O(2)	–	–	14	–	7
	N(5)	O(5)	–	–	–	17	–
	N(5)	O(6)	–	–	–	11	–
	N(5)	O(8)	–	–	–	–	12
	N(8)	O(6)	–	–	–	12	–

helical and a few hairpin structures are sampled. The RMSD distribution of simulation $S^{54A7\beta}_{FG\_layer\_H_2O}$ is comparable to that of simulation $S^{54A7\beta}_{FG\_H_2O}$. This is also observed for the populations of H-bonds (*Table 3*).

The H-atom–H-atom NOE distance-bound violations and ^3J values calculated from simulations $S^{54A7\beta}_{FG\_MeOH}$, $S^{54A7\beta}_{CG\_MeOH}$, $S^{54A7\beta}_{FG\_layer\_MeOH}$, $S^{54A7\beta}_{FG\_H_2O}$, $S^{54A7\beta}_{CG\_H_2O}$ and $S^{54A7\beta}_{FG\_layer\_H_2O}$, are compared with experimental data [14] in *Fig. 7*. Since the 3_{14}-helical structure is very stable in simulation $S^{54A7\beta}_{FG\_layer\_MeOH}$, the NOE distance-bound violations in this simulation are smaller than in other simulations of the peptide in MeOH (*Fig. 7,e*, upper panels, gray lines), and ^3J values calculated from simulation $S^{54A7\beta}_{FG\_layer\_MeOH}$ also agree best with the experimental data (*Fig. 7,f*). For the simulations in H_2O (lower panels), although simulation $S^{54A7\beta}_{CG\_H_2O}$ sampled neither 3_{14}-helical nor hairpin structure, large NOE-bound violations (larger than 0.1 nm) only appeared for No. 53, $H_N(3)$–$H_\beta(7)$, No. 54, $H_N(2)$–$H_\beta(8)$, and No. 55, $H_N(2)$–$H_\gamma(8)$ (*Fig. 7,i*, black lines). There are four large NOE-bound violations in simulation $S^{54A7\beta}_{FG\_layer\_H_2O}$, No. 53 – 56 (*Fig. 7,k*). The agreement of calculated and experimental ^3J values in H_2O are similar in all simulations, *i.e.*, $S^{54A7\beta}_{FG\_H_2O}$, $S^{54A7\beta}_{CG\_H_2O}$ and $S^{54A7\beta}_{FG\_layer\_H_2O}$ (*Fig. 7,h*, *j*, and *l*; black crosses).

Influence of the Cut-Off Radius for the Non-Bonded Interactions. Since the 3_{14}-helical structure showed much more stability in simulation $S^{54A7\beta}_{FG\_layer\_MeOH}$ than in simulation $S^{54A7\beta}_{FG\_MeOH}$, the influence of the nonbonded interaction cut-off alone rather than in combination with a change in solvent model was investigated too. A simulation

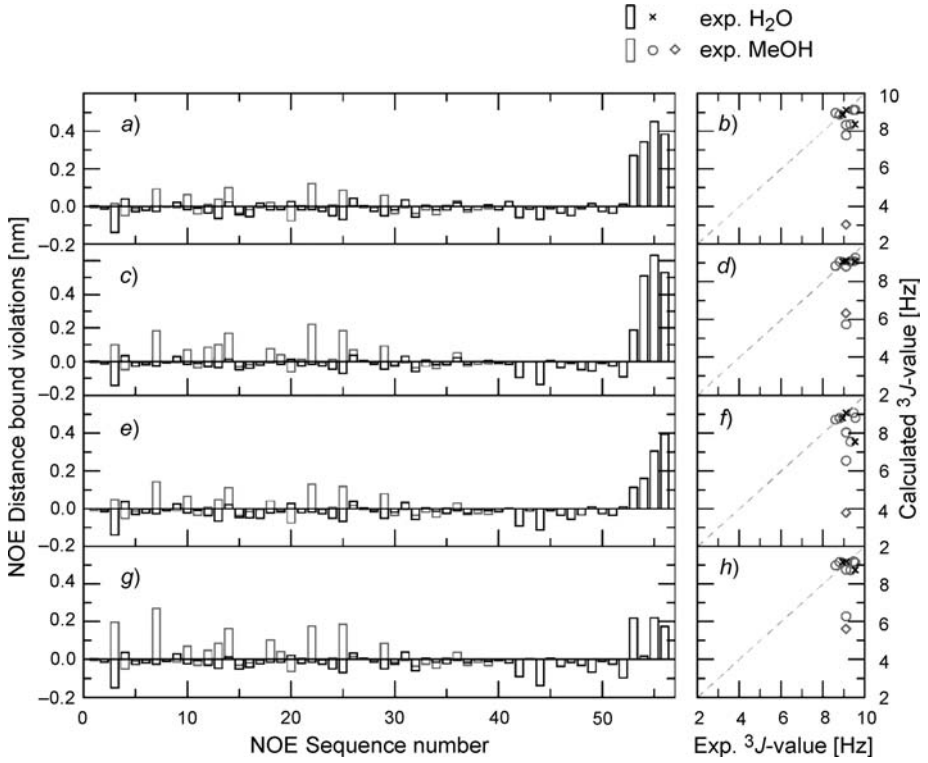

Fig. 5. *Comparison of* $\langle r^{-6}\rangle^{-1/6}$*-averaged NOE distance-bound violations* (left panels) *and average* $^3J(HN,H_\beta)$ *values* (right panels) *as obtained from simulations in MeOH* (*Panels a–d*) *or* H_2O (*Panels e–h*)*, and experimental data* [14] *in* H_2O (black bars and crosses) *or in MeOH* (gray bars, circles, and diamonds). The simulations were carried out for 200-ns using the GROMOS force fields 54A7 and 54A7β. For the specification of the NOE atom pairs and the $^3J(HN,H_\beta)$ values, we refer to Tables S1–S4 in [19]. *a*), *b*) $S^{54A7}_{FG\_MeOH}$; *c*), *d*) $S^{54A7\beta}_{FG\_MeOH}$; *e*), *f*) $S^{54A7}_{FG\_H_2O}$; *g*), *h*) $S^{54A7\beta}_{FG\_H_2O}$. The simulation names are specified in *Table 1*.

of the octa-β-peptide in FG MeOH using the GROMOS force field 54A7β and the nonbonded interaction cut-offs 1.4/2.0 nm was carried out for 200 ns. The comparison of the simulations using different nonbonded interaction cut-offs is presented in *Supplementary Material*. Figs. S1 and S2 show that the larger cut-off stabilizes the 3_{14}-helical structure, and thus the agreement with experimental data is better.

Influence of the Presence of a Zn^{2+} *Ion in Solution.* The comparison of the simulations with and without a single Zn^{2+} ion in aqueous solution, the former corresponding to the NMR experimental set-up, is shown in Figs. S3 and S4 of *Supplementary Material*. No significant change of the behavior of the peptide was detected upon addition of a single Zn^{2+} ion. This can be explained by the observation that the Zn^{2+} ion stays solvated in bulk H_2O, where the H_2O dipoles reduce the *Coulomb* interaction of the ion with the peptide.

Computational Efficiency. The computational efficiency of the simulations is illustrated in *Table 1*. All simulations were run on eight CPUs (central processing units)

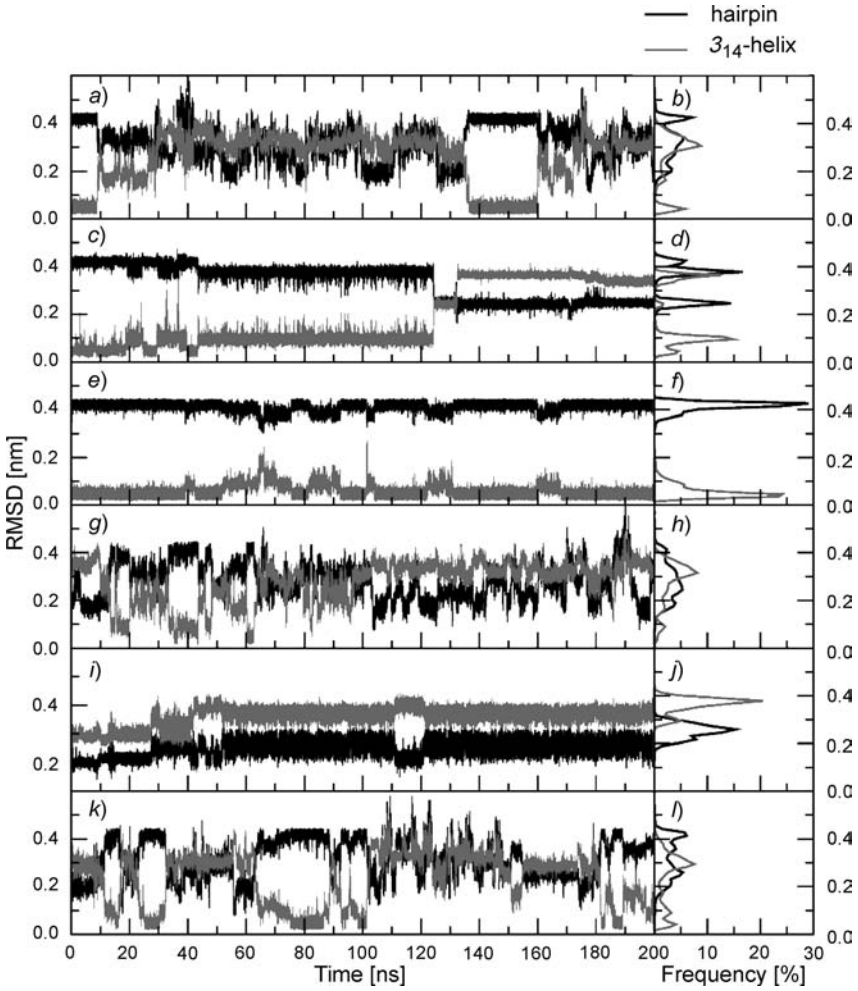

Fig. 6. *Time evolution* (left panels) *and distribution* (right panels) *of the atom-positional RMSD with respect to a hairpin* (black lines) *and with respect to a* 3_{14}*-helix* (gray lines) *for the backbone atoms of residues 2–7 in 200 ns MD simulations in MeOH* (*Panels a–f*) *and in* H_2O (*Panels g–l*) *using the GROMOS 54A7β force field with different solvent models* (FG, CG, or FG_layer). *a*), *b*) $S^{54A7\beta}_{FG\_MeOH}$; *c*), *d*) $S^{54A7\beta}_{CG\_MeOH}$; *e*), *f*) $S^{54A7\beta}_{FG\_layer\_MeOH}$; *g*), *h*) $S^{54A7\beta}_{FG\_H_2O}$; *i*), *j*) $S^{54A7\beta}_{CG\_H_2O}$; *k*), *l*) $S^{54A7\beta}_{FG\_layer\_H_2O}$. The simulation names are specified in *Table 1*.

using MPI (message passing interface) parallelization. In the mix-grained simulations (CG simulation and FG_layer simulation), one CG MeOH bead represents four FG MeOH molecules [16], and one CG H_2O bead represents five FG H_2O molecules [15]. The CG simulation is eight times faster than the FG simulation, and the FG_layer simulation is *ca.* 1.5 times faster than the FG simulation, although a much larger cut-off for the nonbonded interactions is used in the mixed-grained simulations, *i.e.*, 1.4/2.0 nm *vs.* 0.8/1.4 nm in the FG simulations, and much larger box sizes are used in the FG_layer

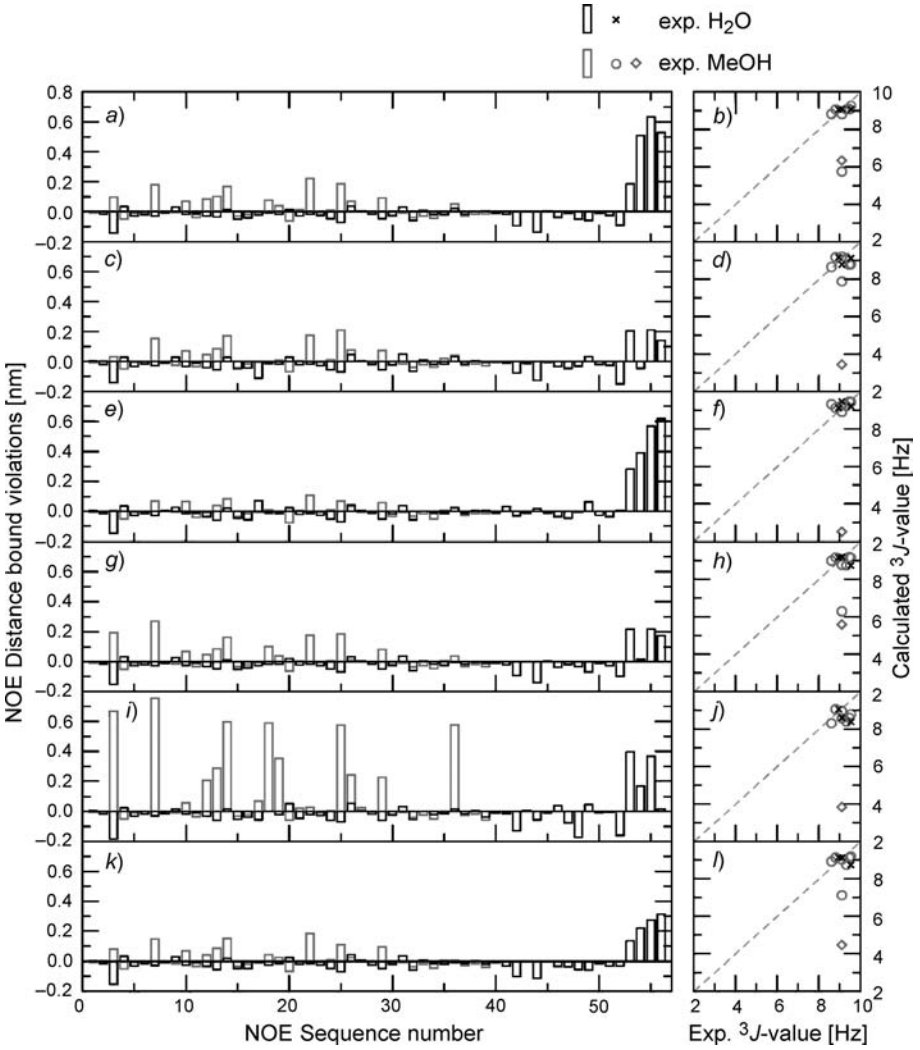

Fig. 7. *Comparison of* $\langle r^{-6}\rangle^{-1/6}$ *averaged NOE distance-bound violations* (left panels) *and average* $^3J(HN,H_\beta)$ *values* (right panels) a*s obtained from simulations in MeOH* (*Panels a–f*) *or* H_2O (*Panels g–l*)*, and experimental data* [14] *in* H_2O (black bars and crosses) *or in MeOH* (gray bars, circles and diamonds). The simulations were carried out for 200 ns using the GROMOS 54A7β force field with different solvent models (FG, CG, and FG_layer). For the specification of the NOE atom pairs and the $^3J(HN,H_\beta)$ values, we refer to Tables S1 – S4 in [19]. *a*), *b*) $S^{54A7\beta}_{FG\_MeOH}$; *c*), *d*) $S^{54A7\beta}_{CG\_MeOH}$; *e*), *f*) $S^{54A7\beta}_{FG\_layer\_MeOH}$; *g*), *h*) $S^{54A7\beta}_{FG\_H_2O}$; *i*), *j*) $S^{54A7\beta}_{CG\_H_2O}$; *k*), *l*) $S^{54A7\beta}_{FG\_layer\_H_2O}$. The simulation names are specified in *Table 1*.

simulations. If the same cut-off radii and box size had been used in FG and mix-grained simulations, the latter would have been faster by more than an order of magnitude.

Comparison with the Experimental NMR Data. When comparing simulated values $\langle Q \rangle_{sim}$ of a quantity Q obtained by averaging over an MD trajectory with experimental values $\langle Q \rangle_{\mathrm{exp}}$ obtained from measurements, one should distinguish between an observable property Q that can be measured directly, and a property Q' that is not an observable so cannot be measured directly, but can be derived from observable, measured $\langle Q \rangle_{\mathrm{exp}}$ values by applying a given or chosen procedure f based on various assumptions and approximations: $Q' = f\left(\langle Q \rangle_{\mathrm{exp}}\right)$ [24]. For example, whereas peak locations and intensities from NMR experiments represent observable, measurable data, molecular structure, NMR order parameters and so on are non-measurable, derived quantities. Comparisons of the latter quantities reflect, at least partly, the approximations and assumptions associated with the conversion of $\langle Q \rangle_{\mathrm{exp}}$ to Q' values and may carry limited experimental information. For this reason, we rather focus on a comparison of NOE distances and ^3J values than on a comparison to helical or hairpin model structures derived from NMR data using single-structure refinement procedures, and a simplified molecular model or force field.

The NOE bounds and ^3J couplings derived from NMR experiments in MeOH are best reproduced in simulation $S^{54A7\beta}_{FG\_layer\_MeOH}$, indicating that a 3_{14}-helical structure dominates the configurational ensemble of the peptide in MeOH.

The NOE bounds and ^3J couplings derived from NMR experiments in H_2O are best reproduced in simulation $S^{54A7\beta}_{FG\_H_2O}$, slightly better than in simulation $S^{54A7\beta}_{FG\_layer\_H_2O}$. *Fig. 6* indicates that a variety of configurations is adopted by the peptide in H_2O, including 3_{14}-helical structures. *Table 3* shows that these comprise a variety of H-bonds.

Conclusions. – The use of coarse-grained (CG) solvent to more efficiently sample the equilibria of an octa-β-peptide in MeOH and H_2O was investigated in the present work. The influence of a change in configuration of a particular residue, of a force-field modification, and of different solvent models was also studied. The comparison of the different simulations led to the following conclusions:

1) Whether the C_β-atom of the 5th residue has (*R*)- or (*S*)-configuration significantly influences the conformational equilibrium in MeOH, while in H_2O no significant differences were observed. Using the correct (*S*)-configuration of the molecule, for which the NMR experiments were performed, the agreement with the NMR data for MeOH is improved compared to the results reported in [19], in which the (*R*)-configuration representing a molecule different from the experimentally studied one was used.

2) The GROMOS force field 54A7β, which was specially adapted for β-peptides, slightly destabilizes the 3_{14}-helical structure in MeOH compared with the force field 54A7. Extending the cut-off radius for nonbonded interactions from 1.4 nm to 2.0 nm restabilizes the 3_{14}-helical structure in MeOH and thus improves the agreement with the NMR data. In H_2O, the 54A7β force field improves the agreement with the NMR data.

3) The supramolecular coarse-grained (CG) solvent model used strongly influences the peptide conformational ensembles, because this CG solvent cannot provide H-bond partners for the solute. The introduction of a fine-grained (FG) solvent layer of molecules around the peptide solved this problem, still leading to a more efficient simulation than when using a pure FG solvent. The FG_layer algorithm proposed by

Riniker et al. [18] restrains the FG solvent layer to the centre of mass (COM) of the solute. This means that, for a non-globular structure such as a hairpin, a thick layer of FG solvent molecules is needed.

An MD simulation of the octa-β-peptide with a Zn^{2+} ion in FG H_2O was also performed using the GROMOS force field 54A7β. The result shows no significant difference from the simulation without the Zn^{2+} ion.

The results of this study illustrate that details of a biomolecular force field, such as torsional-angle energy terms, repulsive *Van der Waals* parameter values, nonbonded cut-off radius, and the solvent model, do matter in regard to obtaining agreement with measured values of observable quantities for polypeptides whose dominant conformation is sensitive to the solvent and a variation of thermodynamic conditions.

Methods. – *Simulation Set-Up.* Eight MD simulations of an octa-β-peptide (*Fig. 1*) in FG/CG MeOH and H_2O, based on the GROMOS force fields 54A7 [22] and 54A7β [23], and two simulations of the same octa-β-peptide, but with different configuration ((*R*) instead of (*S*)) at the C_β-atom of the 5th residue, Lys, in FG MeOH and H_2O, based on the GROMOS force field 54A7 [22] (see *Table 1*) were carried out using the GROMOS [25–28] package of programs. The backbone amino terminus and the Lys side chain were protonated in all simulations; additionally, the carboxy terminus was protonated in the MeOH simulations. The 8th residue, Cys, is slightly different in the MeOH (R = MeS) and H_2O (R = H) simulations (see *Fig. 1*) to match experimental conditions [14]. No counter ions were used.

The coordinates of the first NMR structure from the set of 15 NMR model structures of the peptide in MeOH and ten NMR model structures of the peptide in H_2O, derived from NMR spectroscopic data [14], 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 [14]. In the FG simulations, each peptide was solvated in a periodic, rectangular box with FG MeOH [29] or H_2O [30] as solvent. In the CG simulations, each peptide was solvated in a periodic, rectangular box with CG MeOH [16] or CG H_2O [15] beads as solvent. In the FG_layer simulations, each peptide was first solvated in an FG MeOH or H_2O sphere with a radius of 1.8 nm, then the peptide together with the FG MeOH or H_2O layer was solvated in a periodic, rectangular box with CG MeOH or H_2O beads. The numbers of solvent molecules or beads are compiled in *Table 1*. To keep the FG MeOH or H_2O molecules around the peptide in the FG_layer simulations, attractive harmonic distance restraints beyond a distance r_0 of 1.8 nm were applied between the O-atoms of the FG MeOH or H_2O molecules, and the center of mass (COM) of the peptide approximated by the COM of four selected atoms of the peptide, atoms N(4), C_β(4), C_γ(5), and C_α(5) for the peptide in MeOH, atoms N(3), C_β(3), C(6), and N(7) for the peptide in H_2O. The force constant was 300 kJ mol^{-1} nm^{-2}. The details of the technique to keep a layer of FG solvent around a solute were described in [18].

All simulations were carried out for 200 ns at a constant temp. of 298 K and a constant pressure of 1 atm using the weak coupling algorithm [31]. Solute and solvent were separately coupled to the heat bath. In FG_layer simulations, solute, FG solvent, and CG solvent were separately coupled. The temp. coupling time was set to 0.1 ps, the pressure coupling time to 0.5 ps, and the isothermal compressibility to $4.575 \cdot 10^{-4}$ (kJ mol^{-1} nm^{-3})$^{-1}$. For all simulations, nonbonded interactions were calculated using a triple-range scheme with cut-off radii of 1.4/2.0 nm for mixed-grained simulations and 0.8/1.4 nm for FG simulations. Interactions within the short-range cut-off were evaluated every time step. The intermediate range interactions were updated every fifth time step of 2 fs, and the long-range electrostatic interactions were approximated by a reaction-field force [32] representing a continuum with a relative dielectric permittivity ε_{RF} of 32.63 in mixed-grained MeOH simulations [33], of 19.8 in FG MeOH simulations [29], of 78.5 in mixed-grained H_2O simulations [33], and of 61 in FG H_2O simulations [34]. The bond lengths of the solute and the geometry of the FG solvent molecules were kept rigid using the procedure SHAKE [35] with a precision of 10^{-4}.

Analysis. Trajectory coordinates and energies stored at 0.5-ps intervals were used for analysis. Backbone atom-positional root-mean-square differences (RMSDs) between two solute conformations

were calculated after translational superposition of the solute centers of mass and least-squares rotational fitting of atomic positions, using all backbone atoms (N,C_β,C_α,C) of residues 2–7.

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-atom–H-atom distances extracted from the NOE intensities determined in the NMR experiments were compared with the corresponding average distances in the simulations calculated using $\langle r^{-6}\rangle^{-1/6}$, *i.e.*, r^{-6} averaging, where r is the instantaneous H–H distance. The H-atom–H-atom distances involving aliphatic H-atoms were calculated by defining virtual (CH_1), prochiral (stereospecific CH_2), and *pseudo* (Me and non-stereospecific CH_2) atomic positions, and *pseudo*-atom corrections were added to the distance bounds for the latter, 0.1 nm for non-stereospecific CH_2, 0.15 nm for Me, and 0.29 nm for non-stereospecific rotating Me groups [36].

^3J-Coupling constants were calculated using the *Karplus* relation [37],

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

where $a = 6.4$, $b = -1.4$, and $c = 1.9$ Hz [38]. The dihedral angle $\theta(H_N\text{–}N\text{–}C_\beta\text{–}H_\beta)$ can be related to the dihedral angle $\varphi(C\text{–}N\text{–}C_\beta\text{–}C_\alpha)$ by the relation

$$\theta(H_N - N - C_\beta - H_\beta) = \varphi(C - N - C_\beta - C_\alpha) + \delta \tag{2}$$

where the phase shift δ is $-60°$ for an L-amino acid and 60° for a D-amino acid [38].

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 [39]. The configurations were relaxed by energy minimization *in vacuo* before being used for analysis.

This work was financially supported by grant, No. 200020-137827, of the *Swiss National Science Foundation*, and by grant, No. 228076, of the *European Research Council* (ERC), which are gratefully acknowledged.

REFERENCES

[1] T. Schlick, R. Collepardo-Guevara, L. A. Halvorsen, S. Jung, X. Xiao, *Q. Rev. Biophys.* **2011**, *44*, 191.
[2] R. O. Dror, R. M. Dirks, J. P. Grossman, H. Xu, D. E. Shaw, *Annu. Rev. Biophys.* **2012**, *41*, 429.
[3] S. Piana, J. L. Klepeis, D. E. Shaw, *Curr. Opin. Struct. Biol.* **2014**, *24*, 98.
[4] D. Seebach, J. L. Matthews, *Chem. Commun.* **1997**, 2015.
[5] M. S. Cubberley, B. L. Iverson, *Curr. Opin. Chem. Biol.* **2001**, *5*, 650.
[6] R. P. Cheng, S. H. Gellman, W. F. DeGrado, *Chem. Rev.* **2001**, *101*, 3219.
[7] R. P. Cheng, *Curr. Opin. Struct. Biol.* **2004**, *14*, 512.
[8] E. Hecht, I. Huc, 'Foldamers: Structure, Properties, and Applications', Wiley-VCH, Weinheim, Germany, 2007.
[9] W. F. van Gunsteren, R. Bürgi, C. Peter, X. Daura, *Angew. Chem., Int. Ed.* **2001**, *40*, 351.
[10] X. Daura, A. Glättli, P. Gee, C. Peter, W. F. van Gunsteren, *Adv. Protein. Chem.* **2002**, *62*, 341.
[11] W. F. van Gunsteren, Z. Gattin, 'Simulation of Folding Equilibria in Foldamers: Structure, Properties, and Applications', in 'Simulation of Folding Equilibria in Foldamers: Structure, Properties, and Applications', Eds. S. Hecht, I. Huc, Wiley-VCH, Weinheim, Germany, 2007, p. 173.
[12] Z. Lin, N. Schmid, W. F. van Gunsteren, *Mol. Phys.* **2011**, *109*, 493.
[13] W. Huang, Z. Lin, W. F. van Gunsteren, *J. Chem. Theory Comput.* **2011**, *7*, 1237.
[14] G. Lelais, D. Seebach, B. Jaun, R. I. Mathad, O. Flögel, F. Rossi, M. Campo, A. Wortmann, *Helv. Chim. Acta* **2006**, *89*, 361.
[15] S. Riniker, W. F. van Gunsteren, *J. Chem. Phys.* **2011**, *134*, 084110.
[16] W. Huang, S. Riniker, W. F. van Gunsteren, *J. Chem. Theory Comput.* **2014**, *10*, 2213.
[17] S. Riniker, A. P. Eichenberger, W. F. van Gunsteren, *Eur. Biophys. J.* **2012**, *41*, 647.
[18] S. Riniker, A. P. Eichenberger, W. F. van Gunsteren, *J. Phys. Chem. B* **2012**, *116*, 8873.

[19] A. Choutko, W. F. van Gunsteren, *Helv. Chim. Acta* **2013**, *96*, 189.
[20] L. D. Schuler, X. Daura, W. F. van Gunsteren, *J. Comput. Chem.* **2001**, *22*, 1205.
[21] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656.
[22] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, W. F. van Gunsteren, *Eur. Biophys. J.* **2011**, *40*, 843.
[23] Z. Lin, W. F. van Gunsteren, *J. Comput. Chem.* **2013**, *34*, 2796.
[24] W. F. van Gunsteren, J. Dolenc, A. E. Mark, *Curr. Opin. Struct. Biol.* **2008**, *18*, 149.
[25] W. F. van Gunsteren, GROMOS, http://www.gromos.net (accessed March 21, 2014); the GROMOS software package and force fields can be downloaded from this website.
[26] A. P. Eichenberger, J. R. Allison, J. Dolenc, D. P. Geerke, B. A. C. Horta, K. Meier, C. Oostenbrink, N. Schmid, D. Steiner, D. Wang, W. F. van Gunsteren, *J. Chem. Theory Comput.* **2011**, *7*, 3379.
[27] N. Schmid, C. D. Christ, M. Christen, A. P. Eichenberger, W. F. van Gunsteren, *Comput. Phys. Commun.* **2012**, *183*, 890.
[28] A.-P. E. Kunz, J. R. Allison, D. P. Geerke, B. A. C. Horta, P. H. Hünenberger, S. Riniker, N. Schmid, W. F. van Gunsteren, *J. Comput. Chem.* **2012**, *33*, 340.
[29] R. Walser, A. E. Mark, W. F. van Gunsteren, M. Lauterbach, G. Wipff, *J. Chem. Phys.* **2000**, *112*, 10450.
[30] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, 'Interaction models for H_2O in relation to protein hydration', in 'Intermolecular forces', Ed. B. Pullman, Reidel, Dordrecht, the Netherlands, 1981, p. 331.
[31] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
[32] I. G. Tironi, R. Sperb, P. E. Smith, W. F. van Gunsteren, *J. Chem. Phys.* **1995**, *102*, 5451.
[33] 'CRC Handbook of Chemistry and Physics', 56th Edn., Ed. R. C. West, CRC Press, Boca Raton, FL, 1976.
[34] T. N. Heinz, W. F. van Gunsteren, P. H. Hünenberger, *J. Chem. Phys.* **2001**, *115*, 1125.
[35] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
[36] K. Wüthrich, M. Billeter, W. Braun, *J. Mol. Biol.* **1983**, *169*, 949.
[37] M. Karplus, *J. Chem. Phys.* **1959**, *30*, 11.
[38] A. Pardi, M. Billeter, K. Wüthrich, *J. Mol. Biol.* **1984**, *180*, 741.
[39] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, 'Biomolecular Simulation: The GROMOS96 Manual and User Guide', Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, 1996.

Received June 25, 2014