Influence of Variation of a Side Chain on the Folding Equilibrium of a β-Peptide

by **Zhixiong Lin**[a])[b]), **Florian H. Hodel**[a]), and **Wilfred F. van Gunsteren***[a])

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

[b]) School of Life Sciences and Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China (USTC), Hefei 230027, P. R. China

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 the peptide. Here, we present a molecular-dynamics (MD) study of the influence of a variation of a side chain on the folding equilibrium of a β-heptapeptide that folds into a 3_{14}-helical structure. The side chain of the 5th residue, a valine, was changed into five differently branched side chains of different lengths and polarity, Ala, Leu, Ile, Ser, and Thr. Two computational techniques, long-time MD simulations and the one-step perturbation method, were used to obtain free enthalpies of folding. The simulations show that all six peptides exhibit similar folding behavior, and that their dominant fold is the same, *i.e.*, a 3_{14}-helix. Despite the similarities of their structural properties, a small stabilization effect of *ca.* 2 kJ mol^{-1} on the folding equilibrium of the 3_{14}-helical structure due to a branching C_γ-atom in the β^3-side chain is observed. These results confirm those of previous circular dichroism (CD) studies. The length of side chain and its polarity seem to have no apparent (de)stabilization effect. Application of the cost-effective one-step perturbation method to predict free-enthalpy differences appeared to yield an overall accuracy of about k_BT, which is not sufficient to detect the small stabilization effect.

1. Introduction. – Foldamers [1][2] are a class of non-natural polymers which exhibit a strong tendency to form stable, well-defined secondary or even tertiary structures. Among them, β-peptides [3][4] have attracted particular interest because of their resistance to proteases [5] and their ability to permeate cell membranes [6], which make them promising candidates for pharmaceutical applications [7]. Moreover, β-peptides of relatively short lengths solvated in MeOH can fold into stable secondary structures. This solvent is a computationally efficient one because the big Me group, modelled as a united atom, gives it a lower density of atomic interaction sites than H_2O. Therefore, β-peptides are ideal systems to investigate the folding process and folding/unfolding equilibrium.

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. Having two sp^3-C-atoms, C_α and C_β, in the backbone of each residue, its folded conformation is not only influenced by the side-chain sequence but also by its substitution pattern [8]. Experiments [9], quantum calculations [10–12], and molecular-dynamics (MD) simulations [13–15] have shown that peptides with substitution exclusively at the C_β-atom (β^3-peptides) preferentially

adopt a 3_{14}-helical conformation, while an alternating substitution pattern at the C_α- and C_β-atoms (β^2/β^3-peptides) prefer a $2.7_{10/12}$-helix.

Circular dichroism (CD) studies [16–18] have suggested that branched side chains, *i.e.*, a branch at the C_γ-atom, *e.g.*, in Val, Ile, and Thr side chains, substituted at the C_β-atom, stabilize the 3_{14}-helical conformation of β^3-peptides. However, these conclusions were not based on high-resolution NMR data, but on CD spectra only, which have been shown to be a less reliable secondary-structure indicator for small and highly flexible β-peptides [19]. The effect of branched side chains has been investigated using MD simulations by simulating four β-peptides with differently branched side chains [15], and it was concluded that β^3-peptides show a structural preference for the 3_{14}-helix independent of the branching nature of some side chains, in contrast to what was previously proposed based on CD measurements.

Here, we address this question systematically by taking a well-folding β-heptapeptide (*Fig. 1*) and changing the 5th residue, a Val, to Ala, Leu, Ile, Ser, and Thr. In this way, we can investigate the effect of *a*) branching of the side chain, *b*) the length of the side chain, and *c*) the presence of a polar OH group in the side chain on the folding equilibrium of the 3_{14}-helical conformation using MD simulations. This is expensive though because of the large number of long-time MD simulations required. As a cost-effective alternative, we applied the one-step perturbation technique [20–25] to predict the folding equilibria or folding free energies of the six β-peptides with different side-chain substitutions from a single unphysical reference simulation, and investigated the accuracy of the one-step perturbation method.

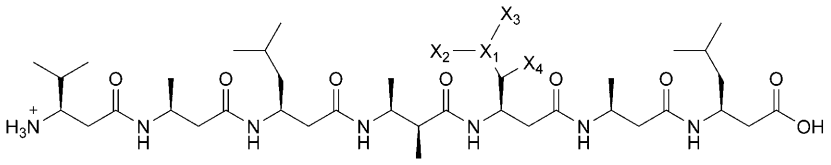

Fig. 1. *Chemical formula of the hepta-β-peptides studied, where X denotes soft-core atoms in the reference state and different atom types in the perturbed-state real peptides.* Sequence: H_2^+-β-HVal-β-HAla-β-HLeu-(*S*,*S*)-β-HAla(αMe)-β-HY-β-HAla-β-HLeu-OH, in which the 5th residue Y is an artificial, designed residue in the reference state, and is Ala, Val, Leu, Ile, Ser, or Thr in the six perturbed-state real peptides, respectively.

2. Methods. – 2.1. *One-Step Perturbation.* Assume that we have a conformational ensemble representing a peptide (un)folding equilibrium that was generated in one long MD simulation using a reference state Hamiltonian H_{ref}. If we now divide the conformational ensemble into two subensembles representing folded *vs.* unfolded conformations, C_F and C_U, we can calculate the free-enthalpy difference between them, $\Delta G^{\text{ref}}_{F,U}$, using the expression

$$\Delta G^{\text{ref}}_{F,U} = -k_B T \ln\left(N^{\text{ref}}_F / N^{\text{ref}}_U\right) \tag{1}$$

where the number of configurations belonging to subensembles C_F or C_U is denoted by N_F or N_U, respectively, and $k_B T$ is the *Boltzmann* constant multiplied by the

temperature. According to statistical mechanics the free-enthalpy change of a (sub)ensemble C due to changing the Hamiltonian from H_{ref} to a perturbed state Hamiltonian H_{pert} is given by

$$\Delta G^{\text{C}}_{\text{pert,ref}} = -k_B T \ln\left\langle e^{-(H_{\text{pert}} - H_{\text{ref}})/k_B T} \right\rangle_{\text{ref,C}} \tag{2}$$

where the (sub)ensemble averaging is denoted by $\langle \ldots \rangle$ and is carried out over all configurations that were generated using H_{ref}, and that belong to the conformational (sub)ensemble C. Using the thermodynamic cycle shown in *Fig. 2* [22–24], we have

$$\Delta G^{\text{pert}}_{\text{F,U}} = \Delta G^{\text{ref}}_{\text{F,U}} + \Delta G^{\text{F}}_{\text{pert,ref}} - \Delta G^{\text{U}}_{\text{pert,ref}} \ . \tag{3}$$

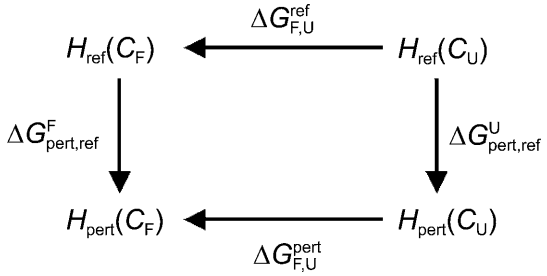

Fig. 2. *Schematic representation of a thermodynamic cycle used in the one-step perturbation method.* H_{ref} and H_{pert} refer to a reference-state Hamiltonian and a perturbed-state Hamiltonian, respectively, in which the conformational ensembles belonging to these Hamiltonians are separated into two conformational states, folded C_{F} and unfolded C_{U}.

From a single equilibrium simulation using a possibly unphysical reference state Hamiltonian H_{ref}, we may thus derive the free-enthalpy differences $\Delta G^{\text{pert}}_{\text{F,U}}$ for many different perturbed-state Hamiltonians H_{pert}.

2.2. *Molecular Model and Simulation Setup.* The simulations of the β-heptapeptides [26] (*Fig. 1*) were carried out in explicit solvent MeOH using the GROMOS05 [27][28] simulation package and GROMOS force-field parameter set 45A3 [29]. The MeOH molecules were represented using a rigid three-site model belonging to the standard GROMOS96 set of solvents [27]. Aliphatic CH_n groups were treated as united atoms, both in the solute and solvent. The N- and C-termini of the peptides were protonated. No counterions were used.

Two types of simulations were performed (*Table 1*). In the two reference-state (R) simulations, the atoms $X_1 - X_4$ in *Fig. 1* are modelled as so-called soft-core atoms, which allow the use of the perturbation *Eqn. 2* to obtain free-enthalpy differences between the perturbed states H and the reference state R. In the six perturbed- or end-state simulations H_{Ala} to H_{Thr}, the folding equilibria of the six real peptides are obtained.

For each simulation listed in *Table 1*, the 3_{14}-helical folded conformation served as initial structure. Periodic boundary conditions were applied based on a rectangular box.

A minimum distance of 1.4 nm between any peptide atom and the closest box wall was enforced while solvating the peptide. The resulting numbers of solvent molecules are listed in *Table 1*. After a steepest-descent energy minimization to remove close contacts between solute and solvent atoms, an equilibration scheme was carried out which included sampling of the atom velocities from a *Maxwell* distribution at 60 K and gradually raising the simulation temperature, while decreasing the strength of the position-restraining potential-energy term for the solute atoms.

Table 1. *Overview of the MD Simulations*

Simulation	Substituent at the 5th residue	No. of solvent molecules	T [K]	Branched at C_γ
Reference States				
$R_{\alpha 1.51}^{340K}$	as in *Fig. 1*	1091	340	–
$R_{\alpha 1.51}^{400K}$	as in *Fig. 1*	1091	400	–
Perturbed States				
H_{Ala}	Ala	1095	340	No
H_{Val}	Val	1090	340	Yes
H_{Leu}	Leu	1095	340	No
H_{Ile}	Ile	1091	340	Yes
H_{Ser}	Ser	1095	340	No
H_{Thr}	Thr	1091	340	Yes

The simulations were carried out for 201 ns at constant temperature, 340 K or 400 K (*Table 1*), and constant pressure, 1 atm. The temperature of 340 K was chosen to obtain sufficient folded and unfolded conformations within each simulation. The reference state (H_{ref}) was also simulated at a higher temperature, 400 K, in order to sample more unfolded conformations. The solute molecules and the MeOH solvent were separately coupled to a temperature bath by means of weak coupling [30], using a coupling time of 0.1 ps. The pressure was calculated with an atomic virial and held constant by weak coupling [30] to an external pressure bath with a coupling time of 0.5 ps, using an isothermal compressibility of 4.575×10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$. All bond lengths and the geometry of the MeOH molecules were constrained using the SHAKE algorithm [31] with a relative geometric accuracy of 10^{-4}, allowing a time step of 2 fs in the leap-frog algorithm to integrate the equations of motion. For the treatment of the nonbonded interactions, twin-range cutoff radii of 0.8/1.4 nm were used. 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 [32] according to a dielectric continuum with a dielectric permittivity of 17.7, the value of the dielectric permittivity of the MeOH model.

In the reference-state simulations, we used soft-core *Lennard–Jones* interactions [33] instead of the normal *Lennard–Jones* functions for the interactions between atoms X_i and other atoms (*Fig. 1*),

$$V(r_{ij}) = \frac{C_{12}^{ij}}{\left[\alpha_{LJ}\lambda^2 \frac{C_{12}^{ij}}{C_6^{ij}} + r_{ij}^6\right]^2} + \frac{C_6^{ij}}{\left[\alpha_{LJ}\lambda^2 \frac{C_{12}^{ij}}{C_6^{ij}} + r_{ij}^6\right]} \quad (4)$$

where r_{ij} is the distance between atoms i and j, C_{12}^{ij} and C_6^{ij} are the *Lennard–Jones* parameters for atom pair (i, j), $\lambda = 0.5$, and $\alpha_{LJ} = 1.51$ is a positive constant controlling the softness of the soft-core atoms [24]. The *Lennard–Jones* parameters of atom type CH1 were used for atom X_1 and type CH_3 for atoms X_2, X_3, and X_4.

2.4. *Analysis.* Trajectory coordinates and energies were stored every 0.5 ps. The 201-ns trajectories were used for structural analysis, while the first 1 ns was treated as equilibration period and thus excluded from calculating free enthalpies.

Atom-positional root-mean-square deviations (RMSD) were calculated after translational superposition of solute centers of mass and rotational least-squares fitting of the atomic coordinates of all backbone atoms (N,CB,CA,C) except for those in the N- and C-terminal residues of the β-peptides. An RMSD of less than 0.1 nm with respect to an ideal 3_{14}-helix was used as the criterion to define the folded conformation (C_F), all other configurations being assigned to the ensemble of unfolded conformations (C_U). Hydrogen bonds were defined by a maximum H-atom-acceptor distance of 0.25 nm and a minimum donor–H-atom–acceptor angle of 135°. A conformational clustering analysis was performed as described by *Daura et al.* [34] on the set of 20100 peptide structures taken at 10-ps intervals from the simulations. An RMSD of 0.1 nm was used as conformational similarity criterion for the β-heptapeptide. We performed clustering using the 201-ns simulation, chose those clusters which made up 95% of the trajectory, and then counted how many of these ones occurred as a function of time. Conformational clustering analyses were also carried out with combined trajectories of equal length for pairs of different simulations. Portions of structures per cluster that belonged to either of the trajectories were calculated afterwards.

For the perturbation free-enthalpy calculations, we did perturbations from the two reference states to the six peptides with the side chain at the C_β-atom of the 5th residue being Ala, Val, Leu, Ile, Ser, or Thr, respectively (*Fig. 1*).

3. Results and Discussion. – 3.1. *Simulations of the Six Peptides.* The atom-positional root-mean-square deviations (RMSD) of the backbone atoms of residues 2 to 6 with respect to a 3_{14}-helical structure are shown in *Fig. 3* for simulations of the six peptides with different side-chain substitutions as a function of simulation time together with their distributions. The dominant fold of all six peptides is a 3_{14}-helix. Many folding/unfolding events occurred during all the simulations, promising an accurate estimation of the folding equilibria. The number of conformational clusters is also shown in *Fig. 3* as a function of time. Few new clusters appear after 150 ns of simulation, indicating that the simulations are converging on a particular set of conformations.

The RMSD distributions of the simulations of the six peptides show the same pattern (*Fig. 3*, and Fig. S1 in *Supplementary Material*[1])), indicating that they access similar conformational spaces. Indeed, a conformational clustering analysis of the

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

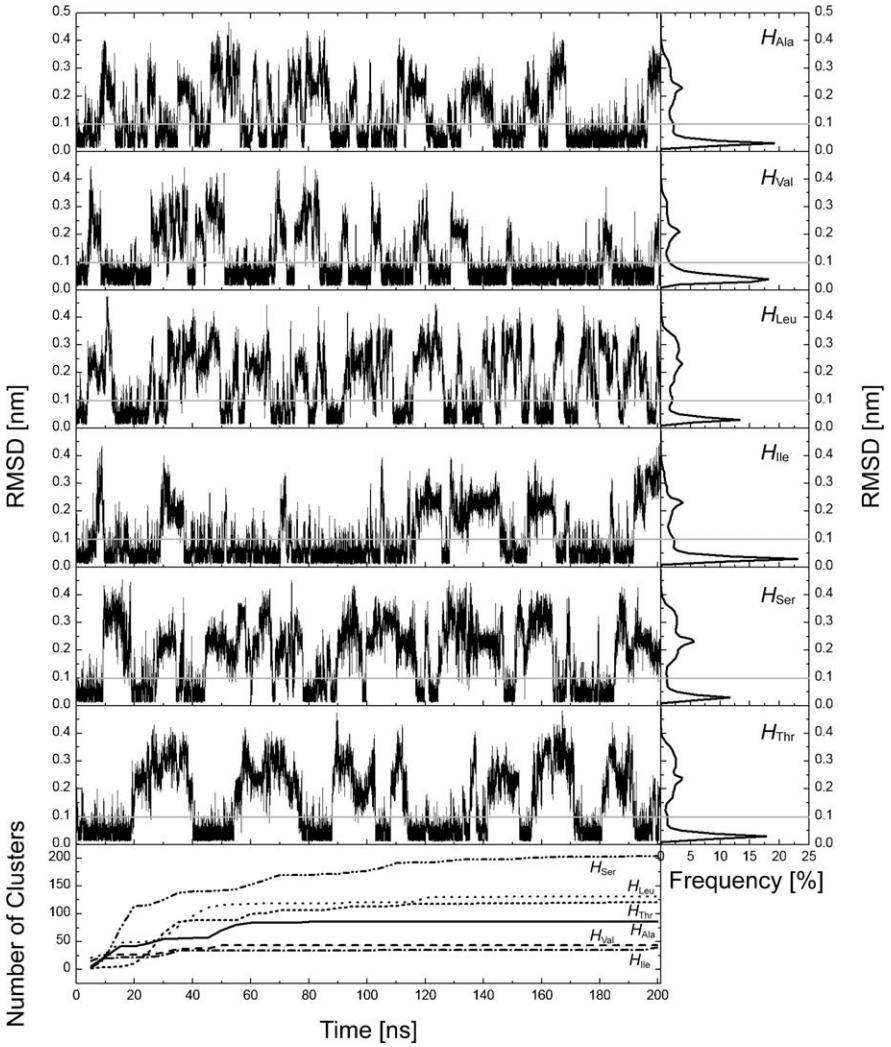

Fig. 3. *Time evolution of the backbone atom-positional RMSD of residues 2–6 with respect to the* 3_{14}*-helix for the six simulations of the β-heptapeptide and the corresponding distribution* (upper panels). Lower panel: *Number of conformational clusters* –: H_{Ala}; – – –: H_{Val}; ⋯ : H_{Leu}; – · –: H_{Ile}; – ⋯ –: H_{Ser}; - - -: H_{Thr}. The grey lines represent the criterion used to distinguish folded from unfolded conformations.

combined trajectories of the H_{Val} simulation and the simulations of the other five peptides shows that the conformational spaces sampled in the simulations show considerable overlap (Fig. S2 in *Supplementary Material*[1])), meaning that they not only have the same folded conformation, but also they share similar unfolded conformations.

The intramolecular H-bond populations of the simulations of the six peptides are listed in *Table 2*. Only 3_{14}-helical H-bonds are populated more than 10% during the

Table 2. *Intramolecular H-Bond Populations* [%] *in 201-ns Simulations of the Six Peptides with Different Side-Chain Substitutions*[a])

Donor···Acceptor	H_{Ala}	H_{Val}	H_{Leu}	H_{Ile}	H_{Ser}	H_{Thr}
NH(i)···O($i+2$) [HB_{14}]						
NH(1)···O(3)	19	20	14	19	16	17
NH(2)···O(4)	46	57	35	54	28	39
NH(3)···O(5)	51	60	38	60	31	43
NH(4)···O(6)	47	56	33	52	32	44
NH(5)···O(7)	18	15	11	16	16	20

[a]) Only H-bonds with a population larger than 10% in at least one of the six simulations are reported.

simulations. No other dominant fold was observed in any of the simulations. The side-chain OH group of the 5th residue of H_{Ser} or H_{Thr} did not form any significant intra-solute H-bonds, *i.e.*, more than 10%, with other atoms in the peptide, neither as an acceptor nor as a donor. Only one such H-bond is populated more than 5%: the one between the OH of the C terminus as a donor and the O of the 5th side chain as an acceptor. The populations are 6 and 9% for the H_{Ser} and H_{Thr} simulations, respectively. In other words, the side-chain OH group of the 5th residue does not compete with H-bonds between backbone atoms which are necessary to maintain a helical structure. The simulations of the different peptides also have very similar distributions for the backbone dihedral angles of the 5th residue (see Fig. S3 in *Supplementary Material*[1])). However, the distributions of the side-chain dihedral angles $\chi1$ and $\chi2$ shown in *Fig. 4* show peaks of different height, while their positions remain largely the same. The solute dipole moments of the six peptides show strong similarities as well (Fig. S4 in *Supplementary Material*[1])).

Despite the overall similarities between the structural properties of the six peptides, there are small differences in their folding equilibria, or folding free enthalpies (shown in *Table 3*), as is also indicated by the RMSD distributions (*Fig. 3*) and H-bond populations (*Table 2*). Indeed, the peptides with side chains branched at the C_γ-atom having folding free enthalpies of -1.8, -1.2, and 0.6 kJ mol^{-1} slightly stabilize the 3_{14}-helical structure, compared to those peptides without a branching C_γ-atom having folding free enthalpies of -0.2, 1.6, and 2.2 kJ mol^{-1}; but with 2 kJ mol^{-1}, the effect is rather small. Although the statistical uncertainties of 0.3 to 0.4 kJ mol^{-1} calculated using block averaging [35] shown in *Table 3* are quite small, *Boned et al.* [36] pointed out that the error bar of calculating free enthalpies through *Eqn. 1* using a long-time MD simulation is *ca.* 1 kJ mol^{-1}. So the differences in the folding free enthalpies between peptides with branched and non-branched side chains of *ca.* 2 kJ mol^{-1} are slightly greater than the statistical uncertainties. When comparing the folding free enthalpies of peptides differing in the length or polarity of the side chain, no definite conclusion can be drawn (*Table 3*).

3.2. *One-Step Perturbation Results.* The atom-positional RMSDs of the backbone atoms of residues 2 to 6 with respect to a 3_{14}-helical structure are shown in *Fig. 5* for the two reference simulations, $R_{\alpha1.51}^{340K}$ and $R_{\alpha1.51}^{400K}$, as a function of simulation time, together with their distributions. Although $R_{\alpha1.51}^{400K}$ sampled less-folded conformations, the

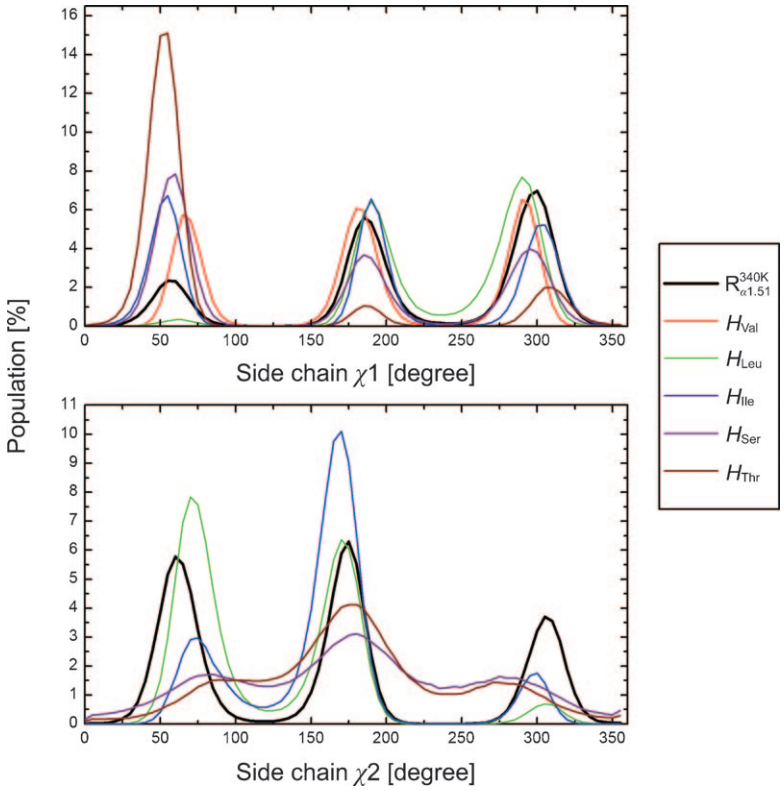

Fig. 4. *Distributions of the side-chain dihedral angles χ1 and χ2 of the 5th residue of the peptide in the simulations of the reference state* $R^{340K}_{\alpha 1.51}$ *and of the real peptides*

Table 3. *Folding Free Enthalpies and Perturbation Free Enthalpies* [kJ mol^{-1}] *Associated with Perturbations from the Two Reference States to the Six Peptides with Different Side-Chain Substitutions*[a])

	H_{Ala}	H_{Val}	H_{Leu}	H_{Ile}	H_{Ser}	H_{Thr}
$\Delta G^{sim}_{F,U}$	-0.2 ± 0.3	-1.8 ± 0.4	1.6 ± 0.3	-1.2 ± 0.4	2.2 ± 0.4	0.6 ± 0.4
$R^{340K}_{\alpha 1.51}$						
$\Delta G^{ref}_{F,U}$			-0.3 ± 0.4			
$\Delta G^{F}_{pert,ref}$	20.5 ± 0.7	14.6 ± 0.6	22.2 ± 0.5	26.4 ± 0.3	23.5 ± 1.6	20.4 ± 1.4
$\Delta G^{U}_{pert,ref}$	19.3 ± 1.5	15.0 ± 0.4	23.1 ± 0.2	26.3 ± 0.3	20.4 ± 2.2	15.2 ± 2.5
$\Delta G^{pert}_{F,U}$	0.9 ± 1.7	-0.7 ± 0.8	-1.2 ± 0.7	-0.2 ± 0.6	2.8 ± 2.7	4.9 ± 2.9
$R^{400K}_{\alpha 1.51}$						
$\Delta G^{ref}_{F,U}$			2.9 ± 0.3			
$\Delta G^{F}_{pert,ref}$	9.8 ± 1.9	9.9 ± 0.5	17.2 ± 0.8	18.9 ± 1.6	4.7 ± 2.6	4.7 ± 2.5
$\Delta G^{U}_{pert,ref}$	7.4 ± 1.3	6.2 ± 1.2	13.4 ± 2.1	20.0 ± 0.9	9.5 ± 1.8	4.1 ± 1.8
$\Delta G^{pert}_{F,U}$	5.3 ± 2.3	6.6 ± 1.3	6.7 ± 2.3	1.8 ± 1.9	-1.9 ± 3.2	3.5 ± 3.1

[a]) The folding free enthalpies $\Delta G^{sim}_{F,U}$ and $\Delta G^{ref}_{F,U}$ were obtained using *Eqn. 1*, the perturbation free enthalpies $\Delta G^{F}_{pert,ref}$ and $\Delta G^{U}_{pert,ref}$ from *Eqn. 2*, and $\Delta G^{pert}_{F,U}$ from *Eqn. 3*. The statistical uncertainties were estimated using block averaging [35].

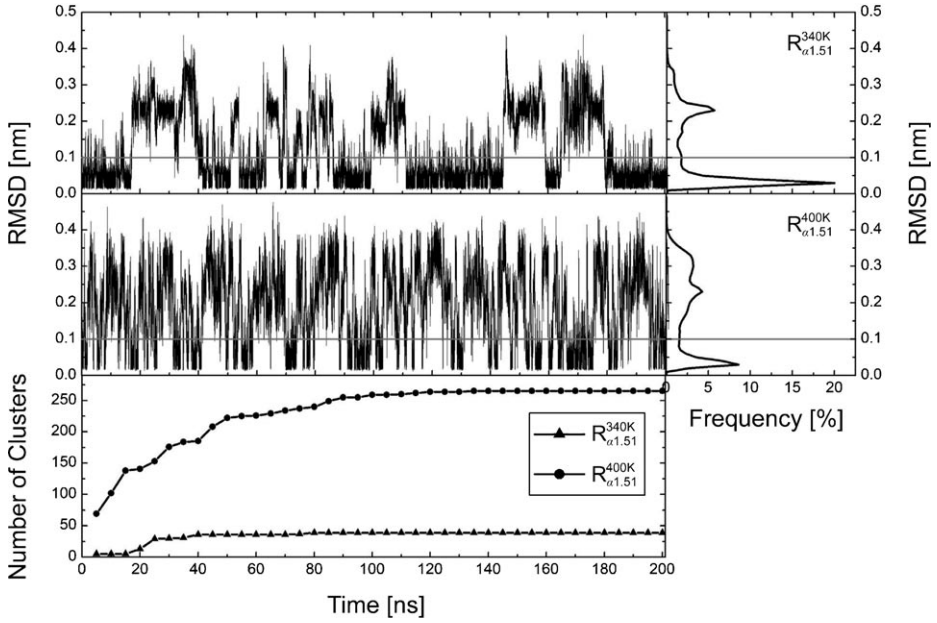

Fig. 5. *Time evolution of the backbone atom-positional RMSD of residues 2–6 with respect to the* 3_{14}*-helix for the two reference simulations* $R^{340K}_{\alpha1.51}$ *and* $R^{400K}_{\alpha1.51}$ *of the β-heptapeptide and the corresponding distribution* (upper panels). Lower panel: *Number of conformational clusters.* The grey lines represent the criterion used to distinguish folded from unfolded conformations.

reference simulations show the same pattern of RMSD distribution as the simulations of the peptides (Fig. S1 in *Supplementary Material*[1])).

Sufficient overlap of the sampled phase space between reference and perturbed state simulations is crucial for the application of one-step perturbation, or more general free energy perturbation [37][38], *i.e.*, the reference simulation should sample the important phase space for all perturbed states in order to obtain accurate free enthalpies. To examine the conformational space overlap between the reference- and the perturbed-state simulation trajectories, a conformational clustering analysis of the combined trajectories of the $R^{340K}_{\alpha1.51}$ and the simulations of the six peptides was performed. The results are shown in *Fig. 6*. The major clusters comprise configurations from both simulations, of the reference and the perturbed states, no cluster exclusively containing configurations from only one simulation. The results indicate that the conformational spaces sampled in the simulations of the reference and the perturbed states largely overlap, not only in the folded conformation, but also in regard to the unfolded ones. Similar results were found for the reference simulation $R^{400K}_{\alpha1.51}$ (Fig. S5 in *Supplementary Material*[1])). In addition, simulation $R^{340K}_{\alpha1.51}$ visited the same range of backbone and the side-chain dihedral angle values for the 5th residue as the perturbed-state simulations (*Fig. 4*, and Fig. S3 in *Supplementary Material*[1])), another indication that this reference simulation sampled a similar conformational ensemble as the perturbed state ones. The overlap of the sampled phase spaces was confirmed by a

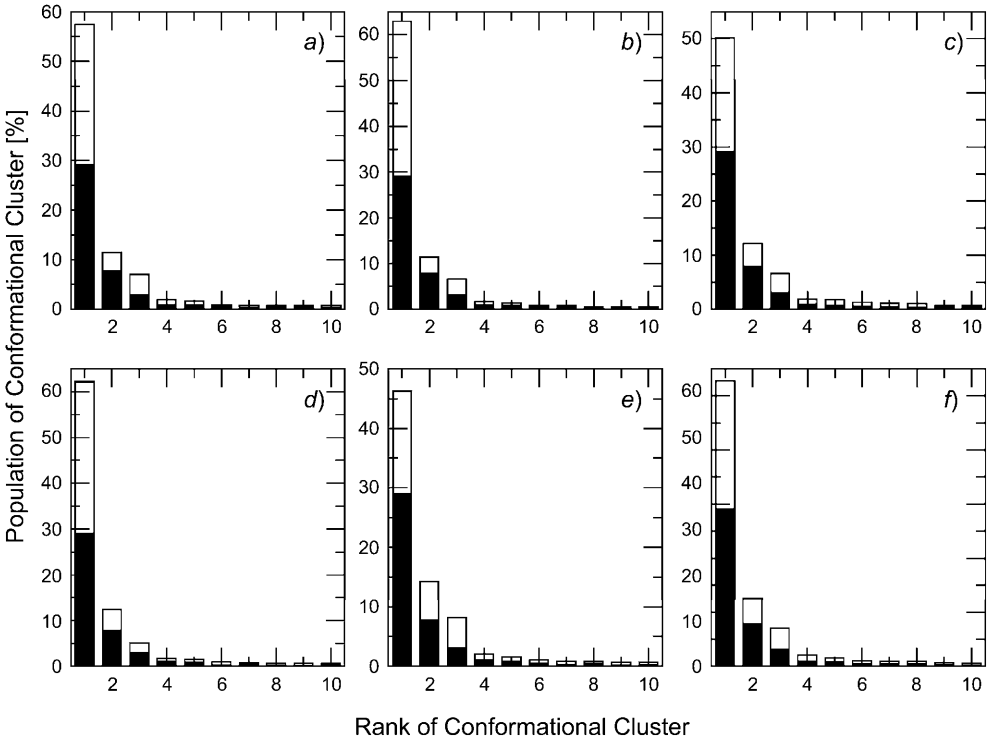

Fig. 6. *Conformational clustering analysis of the combined trajectories of equal lengths of the reference simulation* $R^{340K}_{\alpha 1.51}$ *and the six peptide simulations.* The population in percentage per cluster and the portion of structures per cluster that belong to the trajectories of either $R^{340K}_{\alpha 1.51}$ (black) or the peptide simulations (white) is shown in decreasing order. *a*) $R^{340K}_{\alpha 1.51}$ and H_{Ala}; *b*) $R^{340K}_{\alpha 1.51}$ and H_{Val}; *c*) $R^{340K}_{\alpha 1.51}$ and H_{Leu}; *d*) $R^{340K}_{\alpha 1.51}$ and H_{Ile}; *e*) $R^{340K}_{\alpha 1.51}$ and H_{Ser}; *f*) $R^{340K}_{\alpha 1.51}$ and H_{Thr}.

comparison of the energy distributions of the solute–solute and solute–solvent energies shown in *Fig. 7.*

Free-enthalpy differences of perturbing from the two reference states to the six peptides with different side-chain substitutions (*Eqn. 2*) are listed in *Table 3* together with the predicted folding free enthalpies (*Eqn. 3*). The statistical uncertainties were estimated using block averaging [35]. Among the peptides with apolar side chains, the folding free enthalpies of H_{Ala}, H_{Val}, and H_{Ile} predicted by one-step perturbation (*Eqn. 3*) using the reference simulation $R^{340K}_{\alpha 1.51}$ differ from the simulation results (*Eqn. 1*) by *ca.* 1 kJ mol^{-1}, while, for H_{Leu}, the deviation is 2.8 kJ mol^{-1}, which is about k_BT at 340 K. So one-step perturbation gives an overall accuracy of about k_BT, and the average absolute deviation is 1.5 kJ mol^{-1} for the folding free enthalpies of the peptides with apolar side chains. This is very similar to the results obtained from a previous one-step perturbation study of a similar case [24]. For peptides with polar side chains, the differences in the folding free enthalpies between one-step perturbation and the simulation results are not consistent. For H_{Ser}, the deviation is only 0.6 kJ mol^{-1}, while, for H_{Thr}, it is 4.3 kJ mol^{-1}. Overall, according to one-step perturbation based on

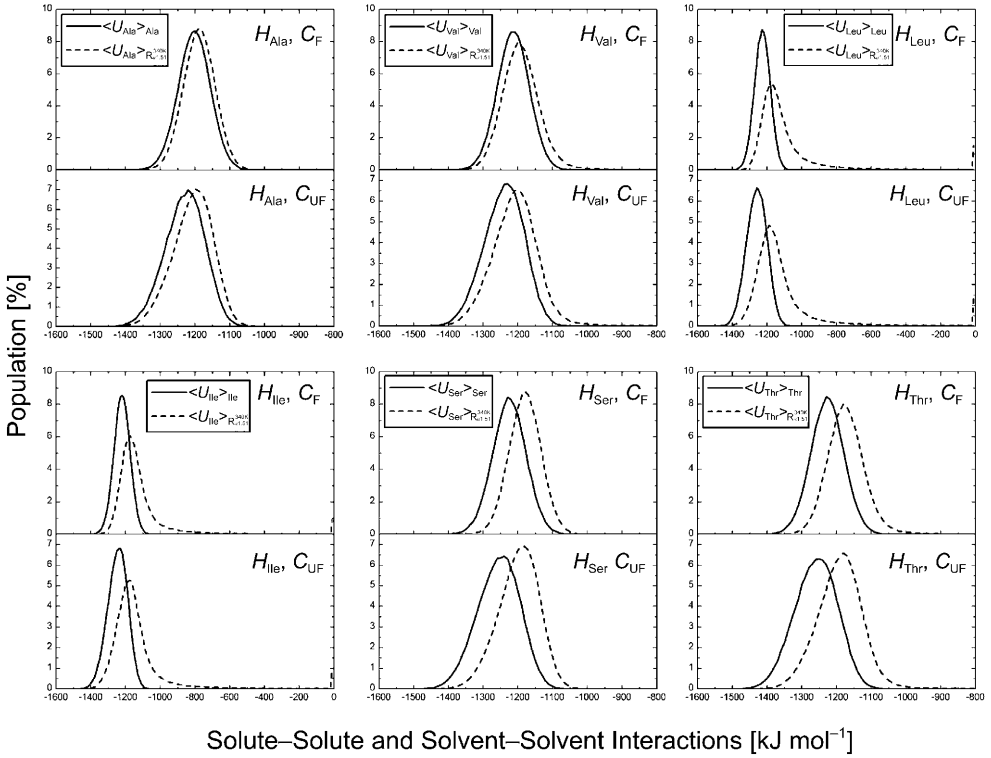

Fig. 7. *The distributions of solute–solute and solute–solvent interaction energies of the six peptides* (H_{Ala}, H_{Val}, H_{Leu}, H_{Ile}, H_{Ser}, and H_{Thr}) *in the simulations of themselves* (solid lines) *or in the reference simulation* $R_{\alpha 1.51}^{340K}$ (dashed lines).

simulation $R_{\alpha 1.51}^{340K}$, side chains branched at the C_γ-atom have no 3_{14}-helix-stabilizing effect.

The reference simulation $R_{\alpha 1.51}^{400K}$ sampled more folding/unfolding events, and also more unfolded conformations (*Fig. 5*, and Fig. S1 in *Supplementary Material*[1])), which make it a better reference state simulation [24]. Except for H_{Ser}, all the folding free enthalpies predicted by $R_{\alpha 1.51}^{400K}$ at 400 K are larger than the explicitly simulated ones, and, in principle, should be higher than those obtained from the simulations at 340 K. Reweighting the ensemble averages to 340 K did not give any useful result (see Table S1 in *Supplementary Material*[1]) and the corresponding texts). Again, there is no stabilizing effect by side chains branched at the C_γ-atom.

For the peptides with polar side chains, $R_{\alpha 1.51}^{340K}$ gives a better folding free enthalpy for H_{Ser}, whereas $R_{\alpha 1.51}^{400K}$ gives a better result for H_{Thr}. In addition, the statistical uncertainties of the results of both H_{Ser} and H_{Thr} are larger than those for the peptides with apolar side chains. In fact, when doing a perturbation from an apolar reference state to a polar perturbed state, the orientations of the surrounding polar solvent molecules pose a problem [39]. That is, in the apolar reference simulation, those orientations are random, whereas in the polar perturbed state, due to the electrostatic

interactions, the surrounding polar solvent molecules have preferences for specific orientations. Therefore, this kind of perturbation requires much longer sampling or, in the present case, an additional sampling of the C–C–O–H torsional angle degree of freedom. Only if the reference simulation happens to sample enough configurations of the right orientations of this torsional angle and the solvent molecules for the perturbed state, the results may be correct. Increasing the simulation time might help to improve the accuracy and reduce statistical uncertainties as well. A combined linear interaction energy and one-step perturbation approach as proposed by *Oostenbrink* [40] offers a possible solution when there are partial charges in perturbed states.

Considering the peptides with apolar side chains, the overall accuracy of one-step perturbation is of the order of k_BT. In this particular case, however, to correctly predict the small stabilization effect, an accuracy higher than k_BT is required. Hence, no effect on the folding could be deduced from the results of one-step perturbation.

4. Conclusions. – The effect of varying a side chain on the folding equilibrium of a 3_{14}-helical fold of a β-peptide was systematically investigated by means of long-time MD simulations and the one-step perturbation method. A well-folding β-heptapeptide was selected, and the 5th residue Val was changed into different branched or non-branched side chains, *i.e.*, Ala, Leu, Ile, Ser, and Thr.

The simulation results show that, for all six peptides with different side chains, the dominant fold is a 3_{14}-helix, while no other dominant secondary structure was observed in any of the simulations. Despite all the similarities between the structural properties of these six peptides, *e.g.*, their conformational space accessible at a given thermodynamic state point, RMSD distribution with respect to the 3_{14}-helical structure, H-bonds, dihedral angles, and solute dipole moments, there is indeed a small stabilization effect on the folding equilibrium of the 3_{14}-helical structure due to the presence of a branching C_γ-atom in the side chain at the β^3-C-atom of the backbone. The stabilization effect is *ca.* 2 kJ mol^{-1} for varying the side chain of one residue, which is slightly larger than the statistical uncertainty. No (de)stabilization effect was observed when varying the length or the polarity of the side chain.

The folding equilibria of these six peptides were also predicted using the more cost-effective one-step perturbation method. This could reduce the number of separate simulations required by a factor of six. The one-step perturbation results showed an overall accuracy of about k_BT. For the folding free enthalpies of the peptides with apolar side chains, the average absolute deviation from the simulation results is with 1.5 kJ mol^{-1}, better than k_BT. This accuracy was, however, not sufficient to detect a stabilization effect. For the two peptides with a polar side chain, the differences in the folding free enthalpies between one-step perturbation and simulations results are not consistent. Longer simulation time or better sampling methods are needed.

In summary, a slight stabilizing effect on the 3_{14}-helical structure due to a branching at the C_γ-atom of the side chain at the β^3-C-atom of the 5th residue is observed. This confirms indications from previous CD studies. No obvious (de)stabilization effect due to the length or the polarity of the 5th side chain was observed.

This work was financially supported by the *National Center of Competence in Research* (*NCCR*) *in Structural Biology* and by grant No. 200020-121913 of the *Swiss National Science Foundation*, by grant

No. 228076 of the *European Research Council* (*ERC*), and by grant No. IZLCZ2-123884 of the *Sino-Swiss Science and Technology Cooperation Program*, which are gratefully acknowledged.

REFERENCES

[1] S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173.
[2] D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893.
[3] R. P. Cheng, S. H. Gellman, W. F. DeGrado, *Chem. Rev.* **2001**, *101*, 3219.
[4] D. Seebach, J. L. Matthews, *Chem. Commun.* **1997**, *21*, 2015.
[5] J. Frackenpohl, P. I. Arvidsson, J. V. Schreiber, D. Seebach, *ChemBioChem* **2001**, *2*, 445.
[6] N. Umezawa, M. A. Gelman, M. C. Haigis, R. T. Raines, S. H. Gellman, *J. Am. Chem. Soc.* **2002**, *124*, 368.
[7] M. Werder, H. Hauser, S. Abele, D. Seebach, *Helv. Chim. Acta* **1999**, *82*, 1774.
[8] T. A. Martinek, F. Fulop, *Eur. J. Biochem.* **2003**, *270*, 3657.
[9] D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L. Matthews, J. V. Schreiber, *Helv. Chim. Acta* **1998**, *81*, 932.
[10] Y. D. Wu, D. P. Wang, *J. Am. Chem. Soc.* **1998**, *120*, 13485.
[11] Y. D. Wu, D. P. Wang, *J. Am. Chem. Soc.* **1999**, *121*, 9352.
[12] R. Gunther, H. J. Hofmann, *Helv. Chim. Acta* **2002**, *85*, 2149.
[13] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem., Int. Ed.* **1999**, *38*, 236.
[14] X. Daura, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *J. Mol. Biol.* **1998**, *280*, 925.
[15] A. Glättli, D. Seebach, W. F. van Gunsteren, *Helv. Chim. Acta* **2004**, *87*, 2487.
[16] B. W. Gung, D. Zou, A. M. Stalcup, C. E. Cottrell, *J. Org. Chem.* **1999**, *64*, 2176.
[17] Y. Hamuro, J. P. Schneider, W. F. DeGrado, *J. Am. Chem. Soc.* **1999**, *121*, 12200.
[18] T. L. Raguse, J. R. Lai, S. H. Gellman, *Helv. Chim. Acta* **2002**, *85*, 4154.
[19] A. Glättli, X. Daura, D. Seebach, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2002**, *124*, 12972.
[20] H. Y. Liu, A. E. Mark, W. F. van Gunsteren, *J. Phys. Chem.* **1996**, *100*, 9485.
[21] C. Oostenbrink, W. F. van Gunsteren, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6750.
[22] Z. X. Cao, H. Y. Liu, *J. Chem. Phys.* **2008**, *129*, 15101.
[23] Z. X. Lin, H. Y. Liu, W. F. van Gunsteren, *J. Comput. Chem.* **2010**, *31*, 2419.
[24] Z. X. Lin, J. Kornfeld, M. Machler, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2010**, *132*, 7276.
[25] Z. X. Lin, W. F. van Gunsteren, *Phys. Chem. Chem. Phys.* **2010**, *12*, 15442.
[26] D. Seebach, P. E. Ciceri, M. Overhand, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amstutz, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 2043.
[27] 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.
[28] 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.
[29] L. D. Schuler, X. Daura, W. F. van Gunsteren, *J. Comput. Chem.* **2001**, *22*, 1205.
[30] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
[31] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
[32] I. G. Tironi, R. Sperb, P. E. Smith, W. F. van Gunsteren, *J. Chem. Phys.* **1995**, *102*, 5451.
[33] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, W. F. van Gunsteren, *Chem. Phys. Lett.* **1994**, *222*, 529.
[34] X. Daura, W. F. van Gunsteren, A. E. Mark, *Proteins: Struct., Funct., Genet.* **1999**, *34*, 269.
[35] M. P. Allen, D. J. Tildesley, 'Computer Simulation of Liquids', Oxford University Press, New York, 1987.
[36] R. Boned, W. F. van Gunsteren, X. Daura, *Chem.–Eur. J.* **2008**, *14*, 5039.

[37] D. Wu, D. A. Kofke, *J. Chem. Phys.* **2005**, *123*, 54103.
[38] D. Wu, D. A. Kofke, *J. Chem. Phys.* **2005**, *123*, 84109.
[39] J. W. Pitera, W. F. van Gunsteren, *J. Phys. Chem. B* **2001**, *105*, 11264.
[40] C. Oostenbrink, *J. Comput. Chem.* **2009**, *30*, 212.

Received January 6, 2011