JCTC Journal of Chemical Theory and Computation

Article

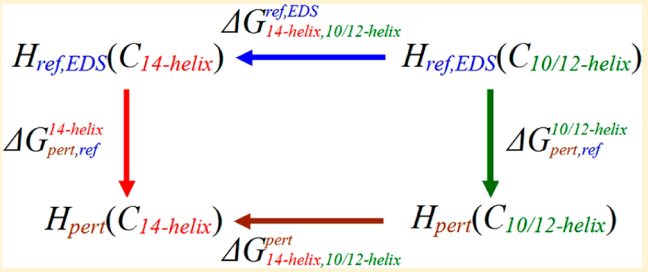
pubs.acs.org/JCTC

Combination of Enveloping Distribution Sampling (EDS) of a Soft-Core Reference-State Hamiltonian with One-Step Perturbation to Predict the Effect of Side Chain Substitution on the Relative Stability of Right- and Left-Helical Folds of β-Peptides

Zhixiong Lin and Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zürich, Switzerland

ABSTRACT: Folding free enthalpies of many not too different polypeptides can be efficiently and accurately predicted with the one-step perturbation (OSP) method using only one or a few molecular dynamics (MD) simulations. In this article, we introduce a combination of enveloping distribution sampling (EDS) and the OSP method (EDS-OSP) and apply it to predict the free enthalpy differences between a right-handed $2.7_{10/12}$-helix and a left-handed 3_{14}-helix for 16 β-peptides with slightly different side-chain substitution patterns. An EDS simulation of a designed soft-core reference-state peptide was carried out in which both helices were sampled. Then, the soft-core atoms were perturbed into physical atoms. Thus, free enthalpy differences between the two helices for the 16 β-peptides can be predicted from only one simulation. The results predicted by EDS-OSP and a previous OSP study are very similar, i.e., the deviations between the results of the 16 peptides are mostly within the order of k_BT, and the average absolute deviation is 1.2 kJ mol^{-1}. Together with the EDS parameter update simulation, about 128 ns of MD simulations needed to be carried out using the EDS-OSP method, while 700 ns of MD simulations were required in the previous OSP study where two separate reference-state simulations and an additional long time MD simulation of one of the 16 β-peptides were carried out. Thus, the computational effort was significantly reduced, i.e., by more than a factor of 5, using the EDS-OSP method. Hence, we consider this method an efficient tool to predict conformational free enthalpy differences from MD simulations.



1. INTRODUCTION

Despite significant advances in the understanding of the protein folding problem due to the development of experimental and computational methods,[1−3] the link between a sequence and its folding equilibrium or process is only partially understood. Elucidation of the effect of different side chains on the folding equilibrium is a prerequisite in the development of accurate protein structure prediction algorithms. This in turn would open the door to the rational design of novel proteins and enzymes with particular folds and biological activities. Moreover, understanding the connection between the side chains and the folding processes would help us to understand processes such as peptide aggregation and may thus enable us to battle diseases such as amyloidosis.

Therefore, studying the change of folding free energies of polypeptides or proteins upon side-chain variations has attracted attention already for more than a decade.[4,5] A great variety of methods have been developed, including machine learning methods[6] which generate predictions based on learned relations delivered from the training set, statistical potential energy methods[7] which use potential energy functions derived from statistical analyses, empirical potential energy methods[8,9] which utilize empirical potential energy combining physical force fields and parameter fitting with respect to particular experimental data, and physical potential energy methods[10−14] which calculate the folding free energy changes based on detailed atomic models, and thus are computationally expensive. Especially when aiming for values of folding free energies, very many folding and unfolding events and unfolded conformations need to be sampled. Since the number of side-chain combinations grows exponentially with the length of a polypeptide chain, calculating the folding free energies of each combination becomes rapidly impossible.

Since 2010, we applied the one-step perturbation (OSP) method[15] using molecular dynamics (MD) simulations to predict the folding free enthalpies of many β-peptides with not too different side chains.[14,16−18] β-peptides[19,20] belong to foldamers,[21,22] which are a class of non-natural polymers which exhibit a strong tendency to form stable, well-defined secondary or tertiary structures. They can fold into stable secondary structures for relatively short chain lengths when solvated in methanol. Also, methanol has a lower density than water, making it a computationally less expensive solvent. Moreover, the folding equilibrium of β-peptides can be simulated within the time scale of hundreds of nanoseconds. These considerations make β-

Received: October 26, 2012
Published: December 12, 2012

dx.doi.org/10.1021/ct300929q | J. Chem. Theory Comput. 2013, 9, 126−134

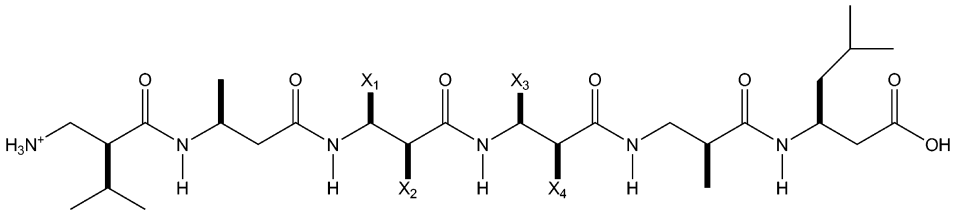


Figure 1. Chemical formula of the β-peptide studied: H_2^+-β^2-HVal-β^3-HAla-(S,S)-β^3-HX$_1$(αX$_2$)-(S,S)-β^3-HX$_3$(αX$_4$)-β^2-HAla-β^3-HLeu-OH, where X$_{1-4}$ denotes soft-core atoms in the soft-core EDS reference-state Hamiltonian and CH_3 or H in the perturbed-state real peptides.

peptides ideal to investigate the folding problem. Having two sp^3 carbon atoms in the backbone of each residue, their folding equilibrium is not only influenced by the side-chain sequence but also by the substitution pattern. The results showed that OSP can accurately, with an overall accuracy of k_BT, predict the folding free enthalpies of many β-peptides, i.e. up to 16, with slightly different side-chain substitutions using one or a few simulations,[14,16] and thus OSP reduces the number of required separate simulations by an order of magnitude.

Recently, we studied the free enthalpy differences between a right-handed $2.7_{10/12}$-helix and a left-handed 3_{14}-helix for a set of 16 β-peptides using the OSP method.[18] Two 100 ns reference-state simulations were carried out with either $2.7_{10/12}$-helical or 3_{14}-helical restraining, respectively, followed by OSP to perturb the soft-core atoms of the reference state into real atoms. Because the reference simulations for the two conformations were carried out separately, the free enthalpy difference between the two helices of the reference-state Hamiltonian could not be calculated. To resolve this, an additional 500 ns MD simulation of one of the 16 β-peptides was required to compute the free enthalpy differences between the two helices for the remaining 15 β-peptides.[18]

Enveloping distribution sampling (EDS)[23−28] is a powerful method to compute relative free enthalpies from a single simulation of an unphysical reference-state Hamiltonian, for which the parameters are first iteratively optimized such that the different end states can be evenly sampled in a subsequent free-enthalpy evaluation simulation. Recently, the EDS formulism has been generalized to compute conformational free enthalpy differences.[29,30] Right-handed α-, π-, and 3_{10}-helical structures of an alanine deca-peptide[29] as well as a right-handed $2.7_{10/12}$-helix and a left-handed 3_{14}-helix of a hexa-β-peptide[30] were studied. In a single simulation of a two-state EDS reference-state Hamiltonian, a pair of helical conformations can be sampled, and thus the free enthalpy difference between these can be calculated.

In this short article, we will introduce a combination (EDS-OSP) of the EDS and OSP methods aimed at much enhanced sampling compared to the use of only one of these methods and apply it to predict the free enthalpy differences between a right-handed $2.7_{10/12}$-helix and a left-handed 3_{14}-helix for a set of 16 β-peptides (Figure 1). The results will be compared with the ones obtained in the previous study using the original OSP method.[18]

2. THEORY

Assuming we wish to calculate the free enthalpy differences between two conformations, α and β, of a set of not too different molecules characterized by Hamiltonians H_{pert}, we may define an EDS reference-state Hamiltonian $H_{ref,EDS}$ such that it envelops conformations α and β and use the one-step perturbation method for the perturbation to H_{pert}.

Two restraining energy function terms are defined which restrain the molecular conformations to conformation α or to conformation β, i.e., $V_X^{rest}(\vec{r}^N;K_X^{rest},\vec{r}_{0\xi}^N)$ where X = A or B, $\vec{r}^N(\vec{r}^1,\vec{r}^2,...\vec{r}^N)$ denotes a configuration of the N atoms of the solute, $\vec{r}_{0\xi}^N$ is the set of parameters which characterizes the conformation ξ, $\xi = \alpha$ or β, and K_X^{rest} is the restraining force constant.

Then, we may construct a soft-core EDS reference-state Hamiltonian

$$\begin{aligned} H_{ref,EDS}(\vec{r}^N;\, s,\, E_{BA}^R) &= -k_BTs^{-1}\ln\{e^{-s(V_A^{rest}(\vec{r}^N)-E_A^R)/k_BT} \\ &+ e^{-s(V_B^{rest}(\vec{r}^N)-E_B^R)/k_BT}\} + H^{phys\_soft}(\vec{r}^N) \\ &= V^{EDS,rest}(\vec{r}^N;\, s,\, E_{BA}^R) + H^{phys\_soft}(\vec{r}^N) \end{aligned} \quad (1)$$

where s is a smoothness parameter and $E_B{}^R - E_A{}^R = E_{BA}{}^R$ is an energy offset parameter difference, which is chosen to optimize the sampling of both end states A and B. k_B are Boltzmann's constant; T is the temperature. $H^{phys\_soft}(\vec{r}^N)$ denotes a molecule with soft-core atoms (see section 3.3), which are then to be perturbed to perturbed-state real atoms of molecules H_{pert}.

With a single simulation of this soft-core EDS reference-state Hamiltonian $H_{ref,EDS}$, both conformations α and β can be sampled. Then, the soft-core atoms can be perturbed into physical atoms using OSP. Thus, the free enthalpy differences between β and α for the perturbed-state molecules H_{pert} can be calculated with only one simulation of the soft-core EDS reference-state Hamiltonian $H_{ref,EDS}$ through

$$\Delta G_{\beta\alpha}^{pert} = -k_BT\ln\left\{\frac{\left\langle e^{-(H_{pert}-H_{ref,EDS})/k_BT}\right\rangle_{ref,set\beta}}{\left\langle e^{-(H_{pert}-H_{ref,EDS})/k_BT}\right\rangle_{ref,set\alpha}}\cdot \frac{N_\beta(H_{ref,EDS})}{N_\alpha(H_{ref,EDS})}\right\} \quad (2)$$

Equation 2 can be simply rewritten as

$$\begin{aligned} \Delta G_{\beta\alpha}^{pert} &= -k_BT\ln\frac{N_\beta(H_{ref,EDS})}{N_\alpha(H_{ref,EDS})} - k_BT \\ &\ln\left\langle e^{-(H_{pert}-H_{ref,EDS})/k_BT}\right\rangle_{ref,set\beta} \\ &+k_BT\ln\left\langle e^{-(H_{pert}-H_{ref,EDS})/k_BT}\right\rangle_{ref,set\alpha} \\ &= \Delta G_{\beta\alpha}^{ref,EDS} + \Delta G_{pert,ref}^{\beta} - \Delta G_{pert,ref}^{\alpha} \end{aligned} \quad (3)$$

Equation 3 is equivalent to the thermodynamic cycle shown in Figure 2 with α a $2.7_{10/12}$-helical conformation and β a 3_{14}-helical

dx.doi.org/10.1021/ct300929q | J. Chem. Theory Comput. 2013, 9, 126−134

one. Only a two-state problem is considered here, but the method can be generalized to multiple-state problems.

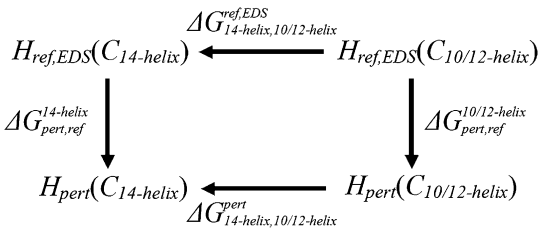


Figure 2. Schematic representation of the thermodynamic cycle used in the combination (EDS-OSP) of the EDS and OSP methods to obtain the free enthalpy differences $\Delta G^{pert}_{14\text{-helix},10/12\text{-helix}}$ between a left-handed 3_{14}-helix and a right-handed $2.7_{10/12}$-helix for a set of 16 real molecules characterized by the Hamiltonians H_{pert}.

The advantage of the EDS-OSP method over the original OSP method is the following. If it is not possible to sample both conformations α and β in a standard MD simulation of the reference-state Hamiltonian, e.g., because one or both of the conformations are not stable, the simulations of the two conformations have to be carried out separately in order to obtain enough statistics. In this case, $\Delta G^{ref}_{\beta\alpha}$ cannot be calculated. Thus, additional simulations are needed, e.g., a simulation of one of the perturbed-state molecules, in order to calculate $\Delta G^{pert}_{\beta\alpha}$ for all other perturbed states. In contrast, in an EDS simulation of the reference-state Hamiltonian, both helices can be sampled in a single simulation, so $\Delta G^{ref,EDS}_{\beta\alpha}$ can be calculated. Thus, $\Delta G^{pert}_{\beta\alpha}$ for all the perturbed states can be calculated with only one simulation of the soft-core EDS reference-state Hamiltonian. Therefore, the EDS-OSP method is much more efficient than the original OSP method in this case.

3. METHOD

3.1. Molecular Model. The β-peptide, which was also studied in ref 18, was modified from the β-peptide in ref 31. The side chains of the third and fourth residues were changed (Figure 1) in order to study the effect of different substitution patterns of these residues on the secondary structure preference of the peptide. The original peptide with the side-chain substitution X_1 = H, X_2 = $CH_2CH(CH_3)_2$, X_3 = $CH(CH_3)_2$, and X_4 = H folds into a right-handed $2.7_{10/12}$-helix and a left-handed 3_{14}-helix.[32,33]

The simulations of the β-peptide were carried out in explicit methanol solvent using the GROMOS11 simulation package[34−38] and the GROMOS force-field parameter set 45A3.[39] The methanol solvent molecules were represented using a rigid three-site model belonging to the standard GROMOS set of solvents.[40] Aliphatic CH_n groups were treated as united atoms, both in the solute and solvent. Both termini were protonated. No counterion was used.

The X_{1-4} atoms in the perturbed-state peptides are either a CH_3 or a H atom (Figure 1), i.e., representing 2^4 = 16 real peptides. The H atom was modeled as a dummy atom without electrostatic or Lennard-Jones interactions while keeping the bonded interactions as for a CH3 united atom, and CH_3 was modeled as a united atom as in the GROMOS force field, keeping its bonded interactions.

3.2. Definition of End-State Hamiltonians and Conformational Sets. Different end-state Hamiltonians A and B, which restrain the peptide into the $2.7_{10/12}$-helix and the 3_{14}-helix, respectively, were characterized using distance restraining potential energy terms. These are defined as an attractive harmonic function applied to the hydrogen-bonding pairs of H and O atoms. Hydrogen bonds NH(3)···O(4), NH(4)···O(1), and NH(6)···O(3) that characterize a $2.7_{10/12}$-helix are restrained in the end-state A, and the pairs NH(2)···O(4) and NH(3)···O(5) that characterize a 3_{14}-helix are restrained in the end state B. The force constants and the reference distances were set to 100 kJ mol^{-1} nm^{-2} and 0.25 nm and 50 kJ mol^{-1} nm^{-2} and 0.25 nm for the end states A and B, respectively. The parameters were chosen such that in the end-state simulations, the average occurrences of the corresponding helical hydrogen-bond populations are about 60−70% (data not shown).

The conformational sets *$2.7_{10/12}$-helix* and *3_{14}-helix*, the free enthalpy between which is to be calculated, were defined through the atom-positional root-mean-square deviation (RMSD) of all backbone atoms (N, CB, CA, C) except for those in the N- and C-terminal residues of the peptides from the ideal helical conformations. The ideal $2.7_{10/12}$-helix was defined as the energy-minimized structure derived from NMR experiments,[18,31,32] and the ideal 3_{14}-helix was defined through the backbone torsional-angle values −180.0°, −121.0°, 58.0°, and −144.0° for ω (CA−C−N−CB), φ (C−N−CB−CA), θ (N−CB−CA−C), and ψ (CB−CA−C−N), respectively, followed by an energy minimization in a vacuum.[18] Configurations belong to set *$2.7_{10/12}$-helix* if

$$\begin{aligned}&\mathrm{RMSD}(\vec{r}^{\,N_{bb}}, \vec{r}^{\,N_{bb}}_{10/12\text{-helix}})\\ &\quad \le \mathrm{RMSD}^{thres}_{10/12\text{-helix}} \text{ and } \mathrm{RMSD}(\vec{r}^{\,N_{bb}}, \vec{r}^{\,N_{bb}}_{14\text{-helix}})\\ &\quad > \mathrm{RMSD}^{thres}_{14\text{-helix}}\end{aligned} \tag{4}$$

they belong to set *3_{14}-helix* if

$$\begin{aligned}&\mathrm{RMSD}(\vec{r}^{\,N_{bb}}, \vec{r}^{\,N_{bb}}_{10/12\text{-helix}})\\ &\quad > \mathrm{RMSD}^{thres}_{10/12\text{-helix}} \text{ and } \mathrm{RMSD}(\vec{r}^{\,N_{bb}}, \vec{r}^{\,N_{bb}}_{14\text{-helix}})\\ &\quad \le \mathrm{RMSD}^{thres}_{14\text{-helix}}\end{aligned} \tag{5}$$

The RMSD threshold value $\mathrm{RMSD}^{thres}_{\xi}$ was set to 0.08 nm for both helices.[32]

We note that these conformational sets *$2.7_{10/12}$-helix* and *3_{14}-helix*, which are defined through atom-positional RMSD, differ

Table 1. Overview of the EDS Simulations

simulation name	description	time step [fs]	simulation time [ns]	EDS parameters s/E_B^R [kJ mol^{-1}] obtained or used
update$1_{1/10\eta}$	EDS update with 1/10 solvent shear viscosity, $\alpha_{LJ}(X_1,X_2,X_3,X_4)$ = 1.51	0.5	256 × 0.15	0.13/−17.3
update$2_{1/10\eta}$	EDS update with 1/10 solvent shear viscosity, $\alpha_{LJ}(X_1,X_4)$ = 2.0, $\alpha_{LJ}(X_2,X_3)$ = 1.51	0.5	256 × 0.15	0.24/−6.4
update$3_{1/10\eta}$	EDS update with 1/10 solvent shear viscosity, $\alpha_{LJ}(X_1,X_4)$ = 2.0, $\alpha_{LJ}(X_2,X_3)$ = 1.51, additional convergence criterion applied	0.5	45 × 0.15	0.45/−1.9
EDS3	EDS ΔG evaluation simulation, $\alpha_{LJ}(X_1,X_4)$ = 2.0, $\alpha_{LJ}(X_2,X_3)$ = 1.51, EDS parameters taken from update$3_{1/10\eta}$	2	101	0.45/−1.9

dx.doi.org/10.1021/ct300929q | J. Chem. Theory Comput. 2013, 9, 126−134

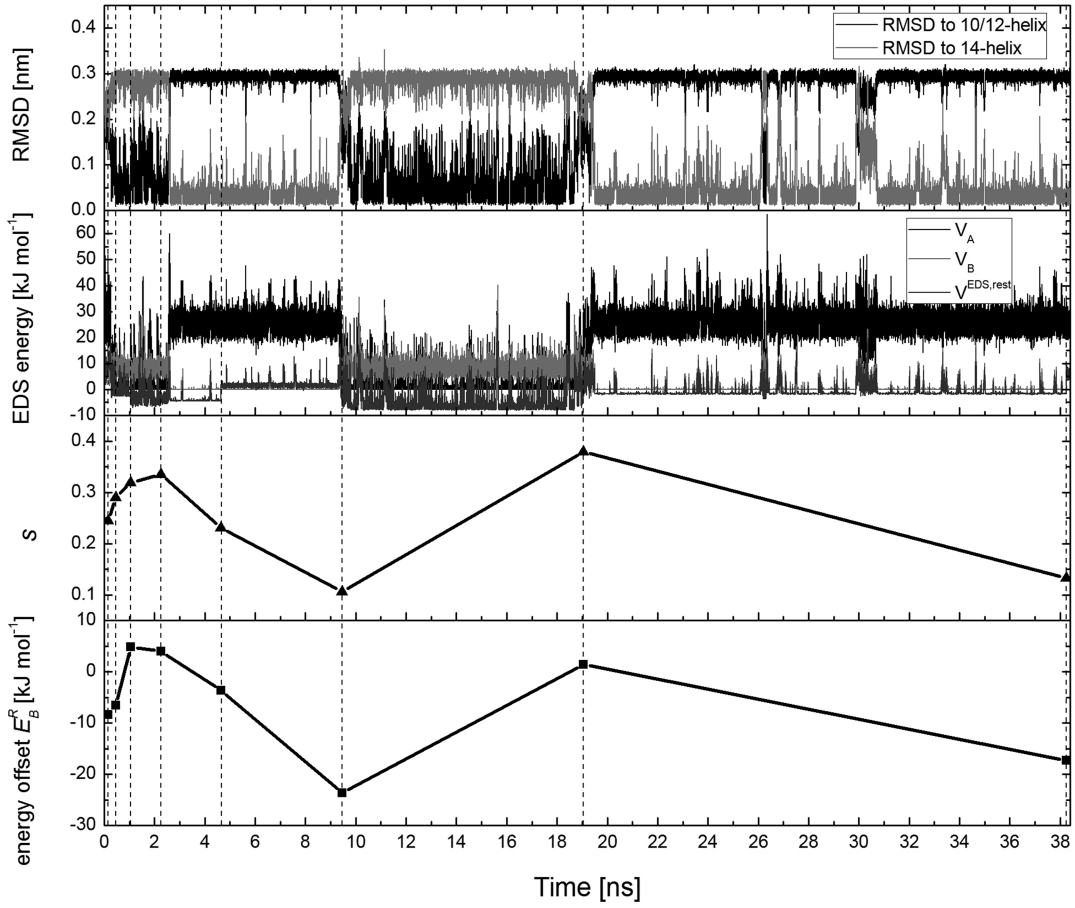


Figure 3. Time evolution of different properties in the EDS parameter update simulation update$1_{1/10\eta}$ with 1/10 solvent shear viscosity. From top to bottom: backbone atom-positional RMSD of the peptide with respect to the right-handed $2.7_{10/12}$-helix (black) and to the left-handed 3_{14}-helix (gray), the restraining (V_A, V_B) and the EDS reference potential energies ($V^{EDS,rest}$), smoothness parameter s, and energy offset $E_B{}^R$. The vertical dashed lines show when the updates of s and $E_B{}^R$ were carried out.

from the conformational ensembles A and B, which are defined by the end-state Hamiltonians and used in the EDS simulations.

3.3. Simulation Setup. The folded structure of the $2.7_{10/12}$-helix served as the initial structure. Periodic boundary conditions were applied on the basis of 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, resulting in 1131 solvent methanol molecules. After a steepest descent energy minimization to remove close contacts between solute and solvent atoms, an equilibration scheme was carried out which included sampling the atom velocities from a Maxwell distribution at 60 K and gradually raising the simulation temperature to 340 K, while decreasing the force constant of the position-restraining potential energy term for the solute atoms from 2.5×10^4 kJ mol^{-1} nm^{-2} to zero.

The simulations were carried out at a constant temperature, 340 K, and a constant pressure, 1 atm. The simulation time is shown in Table 1. The solute molecules and the methanol solvent were separately coupled to a temperature bath by means of weak coupling,[41] using a coupling time of 0.1 ps. The pressure was calculated with a molecular virial and held constant by weak coupling[41] 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 methanol molecules were constrained using the SHAKE algorithm[42] with a relative geometric accuracy of 10^{-4}, allowing a time step of 2 fs in the leapfrog algorithm to integrate the equations of motion. For the treatment of the nonbonded interactions, triple-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[43] according to a dielectric continuum with a dielectric permittivity of 17.7, the value of the dielectric permittivity of the methanol model.[40]

EDS parameter update simulations were carried out with a lower solvent shear viscosity;[30,44,45] i.e., the mass of the solvent was scaled by a factor of 1/100, meaning that the shear viscosity of the solvent was reduced by a factor of 10. This reduction of the solvent mass required a time step of 0.5 fs instead of 2 fs.[30] The parameters s and $E_B{}^R = E_{BA}{}^R$ ($E_A{}^R$ is standardly set to zero in two-state EDS) were updated at fixed time points: after the first, third, seventh, and 15th etc. of a sequence of 0.15 ns simulation periods. We refer to refs 26 and 30 for details of the update scheme. The following additional criterion was added in the parameter update simulations:[30] after each simulation of 0.15 ns (600 configurations), the number of configurations belonging to each conformational set, *$2.7_{10/12}$-helix* or *3_{14}-helix*, during this update time period was counted. If $N_{10/12\text{-helix}}$ and $N_{14\text{-helix}}$ were both above 300, i.e., enough statistics was obtained, and $N_{10/12\text{-helix}}/N_{14\text{-helix}}$ was within the range of 0.8 to 1.2, i.e., roughly equal sampling of the two helices was reached, the parameter

dx.doi.org/10.1021/ct300929q | J. Chem. Theory Comput. 2013, 9, 126−134

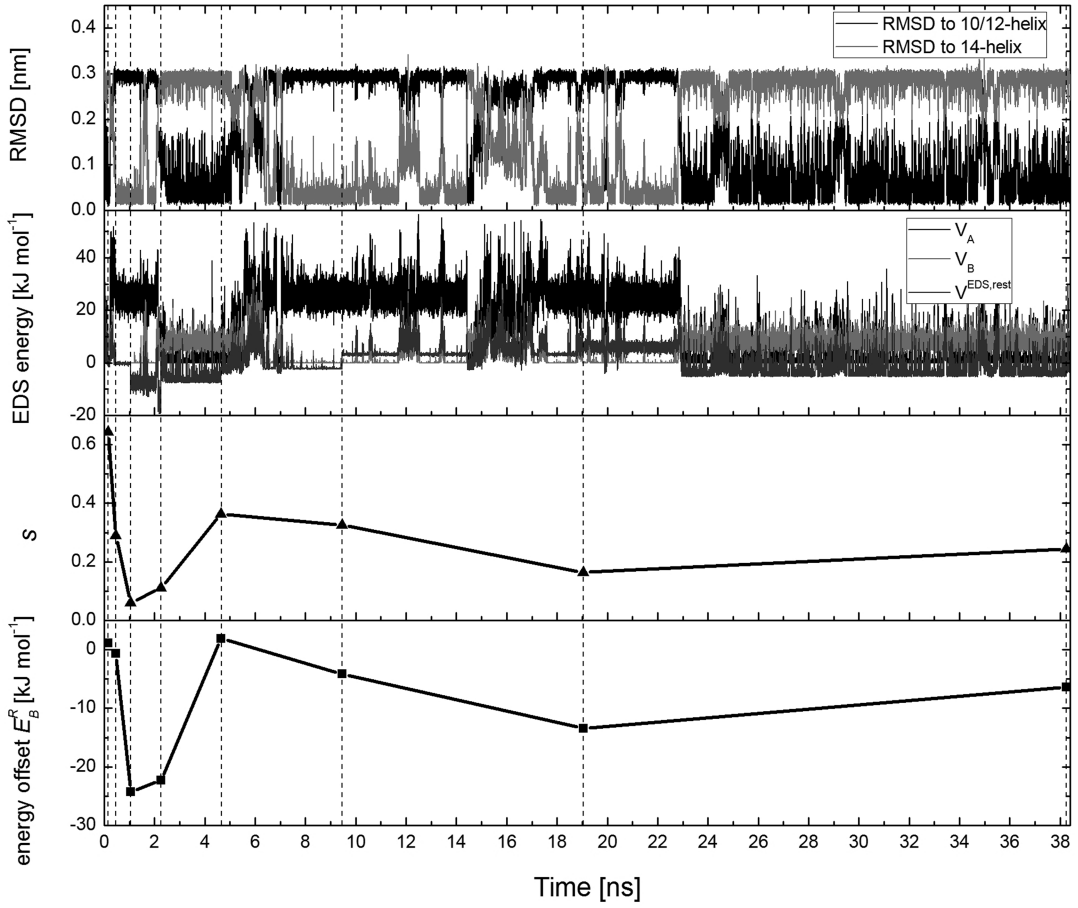


Figure 4. Time evolution of different properties in the EDS parameter update simulation update$2_{1/10\eta}$ with 1/10 solvent shear viscosity. From top to bottom: backbone atom-positional RMSD of the peptide with respect to the right-handed $2.7_{10/12}$-helix (black) and to the left-handed 3_{14}-helix (gray), the restraining (V_A, V_B) and the EDS reference potential energies ($V^{EDS,rest}$), smoothness parameter s, and energy offset $E_B{}^R$. The vertical dashed lines show when the updates of s and $E_B{}^R$ were carried out.

update simulation was terminated, and the final s and $E_B{}^R$ parameters were calculated and used for the free-enthalpy evaluation simulation.[30]

In the EDS reference simulations, we used soft-core Lennard-Jones interactions[46] instead of the normal Lennard-Jones ones for the interactions between the atoms X_{1-4} and other atoms (Figure 1)

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

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 α_{LJ} is a positive constant controlling the softness of the soft-core atoms. In this work, α_{LJ} was first chosen as 1.51 for the atoms X_{1-4} and later changed to 2.0 for the atoms X_1 and X_4 (Table 1) in order to make these atoms softer, thereby allowing more transitions between the R and L helices. The Lennard-Jones parameters of atom type CH3 were used for the atoms X_{1-4}.

An overview of the simulations is shown in Table 1.

In all simulations, trajectory coordinates and energies were stored every 500 steps for analysis. Atom-positional root-mean-square deviations (RMSD) were calculated after translational superposition of the 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-peptide.

4. RESULTS AND DISCUSSION

4.1. EDS Parameter Update Simulations with the Lower Solvent Shear Viscosity. Time evolutions of the backbone atom-positional RMSD of the peptide with respect to the right-handed $2.7_{10/12}$- and left-handed 3_{14}-helical structures, end-state and EDS reference potential energies, and s and $E_B{}^R$ parameters in the EDS parameter update simulation update$1_{1/10\eta}$ with 1/10 solvent shear viscosity are shown in Figure 3. Only 2−3 transitions between the right-handed $2.7_{10/12}$-helix and the left-handed 3_{14}-helix occurred during the 38 ns simulation period. This is due to a high energetic barrier between the two helices and the large distance between the parts of conformational space relevant to these two helices. A large conformational transition is required to go from a right-handed to a left-handed helix. Lowering the solvent shear viscosity can speed up the conformational diffusion of the solute,[30] but due to the high energy barrier between the two helices, the rate at which the peptide moves from one helix to the other is still quite low.

In order to reduce the energetic barrier, the softness parameter α_{LJ} for the atoms X_1 and X_4 of the reference state was changed from 1.51 to 2.0[30] (Table 1). The results of the EDS parameter update simulation update$2_{1/10\eta}$ of the new soft-core reference state are shown in Figure 4. About five transitions between the

dx.doi.org/10.1021/ct300929q | J. Chem. Theory Comput. 2013, 9, 126−134

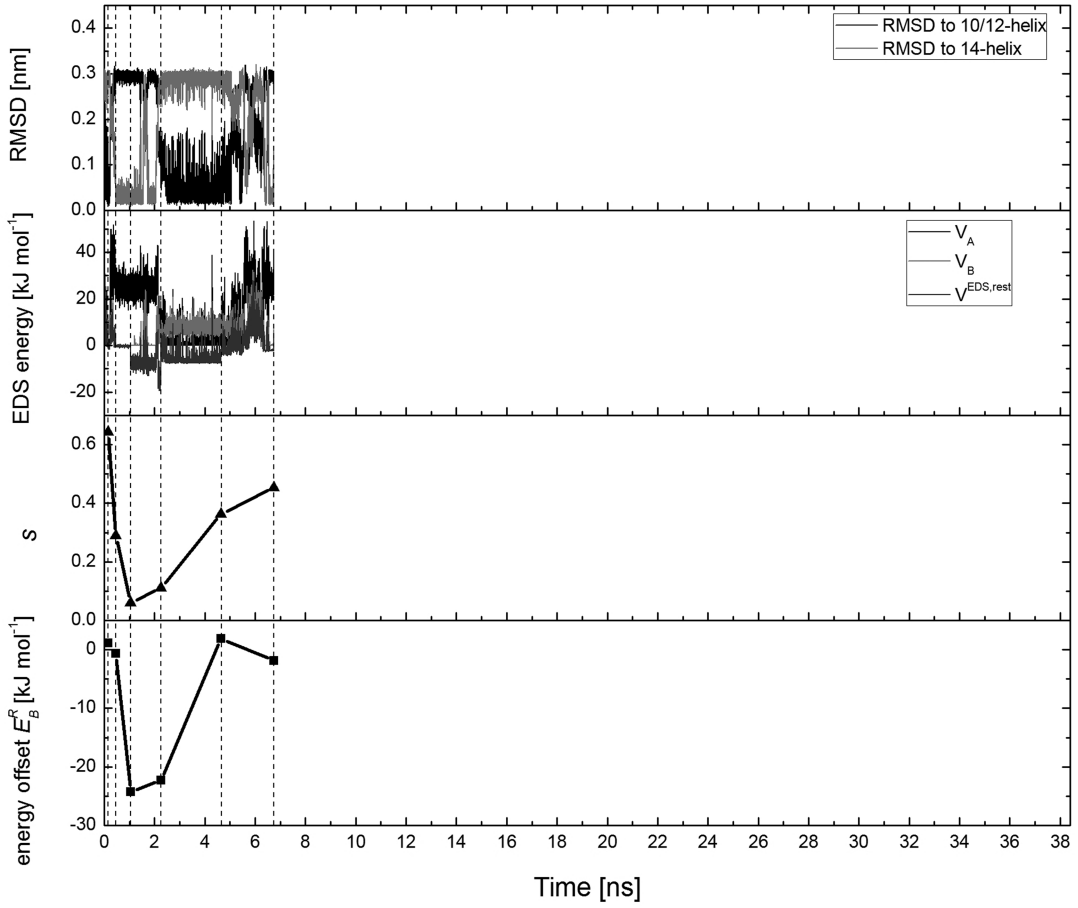


Figure 5. Time evolution of different properties in the EDS parameter update simulation update$3_{1/10\eta}$ with 1/10 solvent shear viscosity. From top to bottom: backbone atom-positional RMSD of the peptide with respect to the right-handed $2.7_{10/12}$-helix (black) and to the left-handed 3_{14}-helix (gray), the restraining (V_A, V_B) and the EDS reference potential energies ($V^{EDS,rest}$), smoothness parameter s, and energy offset $E_B{}^R$. The vertical dashed lines show when the updates of s and $E_B{}^R$ were carried out. The update simulation was terminated after $N_{10/12\text{-helix}}$ and $N_{14\text{-helix}}$ were both above 300 where 4000 configurations per nanosecond were analyzed, and $N_{10/12\text{-helix}}/N_{14\text{-helix}}$ was within the range of 0.8 and 1.2.

right-handed $2.7_{10/12}$-helix and the left-handed 3_{14}-helix occurred during the 38 ns simulation period, i.e. the transitions between the two helices are slightly enhanced compared with the ones in update$1_{1/10\eta}$. Using the mentioned additional criterion[30] in the EDS parameter update simulation update$3_{1/10\eta}$, after 45 of the 0.15 ns simulations, the configurations sampled are 1531 and 1311 for *$2.7_{10/12}$-helix* and *3_{14}-helix*, respectively; i.e. the criterion was fulfilled, so the update simulation was terminated (Figure 5). The obtained s and $E_B{}^R$ parameters were used in the EDS free-enthalpy evaluation simulation.

4.2. EDS Free-Enthalpy Evaluation Simulation. The time evolutions of the backbone atom-positional RMSD of the peptide from the two helical conformations in the EDS free-enthalpy evaluation simulation EDS3, i.e., the soft-core EDS reference simulation with the normal solvent shear viscosity,[30] are shown in Figure 6. The typical residence time of the two helices between transitions was on the time scale of tens of nanoseconds, but both helices were sampled in this single simulation, although an extension of it would improve the statistics. The values for the ensemble averages of the fraction of folded conformations are 0.57 ± 0.07 and 0.12 ± 0.04 for the 3_{14}-helix and the $2.7_{10/12}$-helix, respectively, which were used to calculate the free enthalpy difference $\Delta G^{ref,EDS}_{14\text{-helix},10/12\text{-helix}} = -4.4 \pm 1.0$ kJ mol^{-1} between the two helices for the reference-state Hamiltonian (caption of Table 2). This was not possible in the

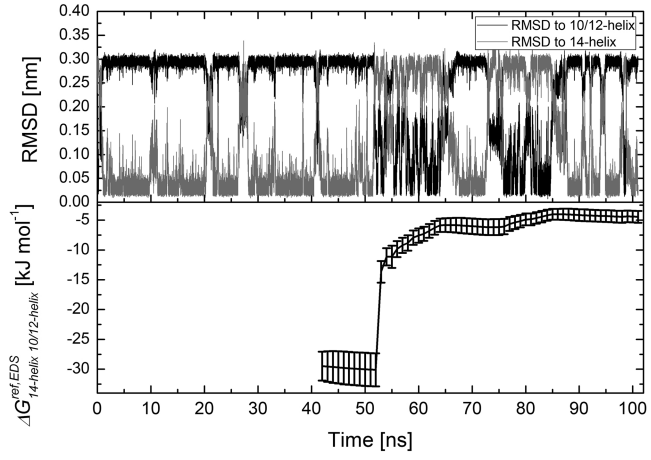


Figure 6. Time evolutions of the backbone atom-positional RMSD of the peptide with respect to the right-handed $2.7_{10/12}$-helix (black) and to the left-handed 3_{14}-helix (gray) and of $\Delta G^{ref,EDS}_{14\text{-helix},10/12\text{-helix}}$ in the EDS free-enthalpy evaluation simulation EDS3.

previous OSP study,[18] where an extra long MD simulation of a perturbed-state peptide had to be carried out.

The distributions of the RMSD are shown in Figure 7. The 3_{14}-helix was more populated than the $2.7_{10/12}$-helix during the EDS free-enthalpy simulation. Not many nonrelevant configurations,

dx.doi.org/10.1021/ct300929q | J. Chem. Theory Comput. 2013, 9, 126−134

Table 2. Free Enthalpy Differences (kJ mol^{-1}) between the Left-Handed 3_{14}-helix and the Right-Handed $2.7_{10/12}$-helix for the 16 Peptides with Different Side-Chain Substitutions Predicted by the EDS-OSP Method Using eq 3[a] and by the Original OSP Method[b]

	third-(β,α), fourth-(β,α)	$\Delta G^{14\text{-helix}}_{pert,ref}$	$\Delta G^{1012\text{-helix}}_{pert,ref}$	$\Delta G^{pert}_{14\text{-helix},10/12\text{-helix}}$ EDS-OSP	$\Delta G^{pert}_{14\text{-helix},10/12\text{-helix}}$ OSP[b]
1	(CH_3,CH_3),(CH_3,CH_3)	21.0 ± 0.2	38.3 ± 1.1	−21.8 ± 1.5	−21.3 ± 0.9
2	(CH_3,CH_3),(CH_3,H)	30.7 ± 0.1	42.8 ± 0.6	−16.6 ± 1.2	−16.8 ± 0.8
3	(CH_3,CH_3),(H,CH_3)	29.2 ± 0.2	47.6 ± 0.8	−22.9 ± 1.3	−21.2 ± 0.9
4	(CH_3,CH_3),(H,H)	42.5 ± 0.2	55.7 ± 0.8	−17.6 ± 1.3	−17.0 ± 1.0
5	(CH_3,H),(CH_3,CH_3)	31.0 ± 0.2	43.6 ± 0.4	−17.1 ± 1.1	−14.8 ± 0.9
6	(CH_3,H),(CH_3,H)	40.4 ± 0.2	46.7 ± 0.6	−10.8 ± 1.2	−10.0 ± 0.9
7	(CH_3,H),(H,CH_3)	38.7 ± 0.3	51.5 ± 0.8	−17.2 ± 1.3	−15.8 ± 0.8
8	(CH_3,H),(H,H)	52.1 ± 0.2	60.8 ± 0.4	−13.1 ± 1.1	−11.2 ± 0.8
9	(H,CH_3),(CH_3,CH_3)	30.2 ± 0.2	33.5 ± 0.4	−7.8 ± 1.1	−4.2 ± 2.0
10	(H,CH_3),(CH_3,H)	40.0 ± 0.1	37.0 ± 0.3	−1.5 ± 1.0	−0.2 ± 0.5
11	(H,CH_3),(H,CH_3)	38.6 ± 0.2	40.9 ± 0.7	−6.7 ± 1.2	−6.2 ± 0.9
12	(H,CH_3),(H,H)	52.0 ± 0.1	49.4 ± 0.3	−1.9 ± 1.0	−1.3 ± 0.9
13	(H,H),(CH_3,CH_3)	43.5 ± 0.2	45.6 ± 0.3	−6.5 ± 1.0	−5.4 ± 0.9
14	(H,H),(CH_3,H)	52.8 ± 0.2	49.0 ± 0.2	−0.6 ± 1.0	1.2 ± 0.9
15	(H,H),(H,CH_3)	51.5 ± 0.1	52.5 ± 0.6	−5.5 ± 1.2	−5.2 ± 0.8
16	(H,H),(H,H)	64.3 ± 0.2	61.4 ± 0.2	−1.5 ± 1.0	−0.2 ± 0.8

[a] $\Delta G^{ref,EDS}_{14\text{-helix},10/12\text{-helix}}$ = −4.4 ± 1.0 kJ mol^{-1}. The statistical uncertainties were estimated by block averaging.[47] [b] Ref 18.

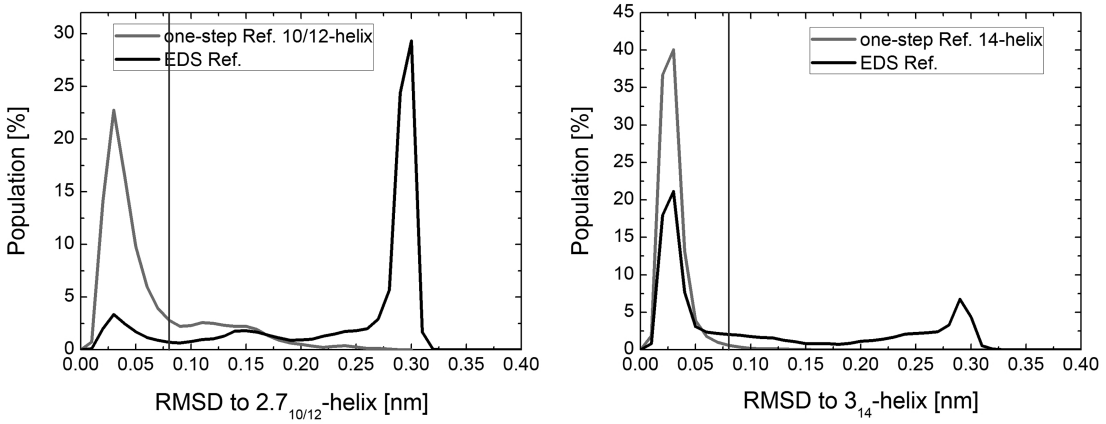


Figure 7. The distributions of the backbone atom-positional RMSD of the peptide with respect to the right-handed $2.7_{10/12}$-helix (left panel) and to the left-handed 3_{14}-helix (right panel) in the reference-state simulations of the previous one-step perturbation study[18] (gray) and in the soft-core EDS reference-state simulation EDS3 (black).

i.e., those not belonging to either *$2.7_{10/12}$-helix* or *3_{14}-helix*, were sampled. Comparing these RMSD distributions with the ones of the two reference-state simulations in ref 18, each helix was less populated in the soft-core EDS reference simulation EDS3. This is not surprising, because in ref 18, the reference simulations were carried out with either $2.7_{10/12}$-helical or 3_{14}-helical restraining, so each reference simulation was meant to sample only one of the two helices. The shape of the RMSD distribution for each helical conformation is similar for the EDS reference simulation and the reference simulations in ref 18 though (Figure 7).

4.3. Free Enthalpy Differences: EDS-OSP versus OSP. The perturbation free enthalpies and free enthalpy differences between the left-handed 3_{14}-helix and the right-handed $2.7_{10/12}$-helix for the 16 perturbed-state real β-peptides, i.e., $\Delta G^{pert}_{14\text{-helix},10/12\text{-helix}}$, predicted by the EDS-OSP method are listed in Table 2, together with the same free enthalpy results from ref 18 predicted in the previous OSP study. The statistical uncertainties were estimated using block averaging.[47] The statistical uncertainties of the $\Delta G^{pert}_{14\text{-helix},10/12\text{-helix}}$ results predicted by EDS-OSP range from 1.0 to 1.5 kJ mol^{-1}, these values being slightly larger than the ones from the previous OSP study which mostly range from 0.8 to 1.0 kJ mol^{-1}.

The resulting free enthalpy differences for the different peptides differ up to more than 20 kJ mol^{-1} (Table 2). For such a large range of free enthalpy values, the results predicted by EDS-OSP and the previous OSP study do not differ too much, i.e., the deviations between them are mostly within the order of k_BT, 2.8 kJ mol^{-1} at 340 K, except for one peptide, and the average absolute deviation is 1.2 kJ mol^{-1}. The only exception is peptide 9 with third-(H,CH_3) and fourth-(CH_3, CH_3), for which the results predicted by EDS-OSP and the previous OSP study differ by 3.6 kJ mol^{-1} (Table 2). We note that the result for this peptide predicted in the previous OSP study has a particularly large statistical uncertainty compared to the other ones, i.e., 2 kJ mol^{-1} instead of 0.8−1.0 kJ mol^{-1} (Table 2). The free enthalpy result of this peptide predicted by the current EDS-OSP method is probably more reliable.

The proposed combination of the EDS and OSP methods is much more efficient than using only the original OSP method. Only a 7 ns EDS parameter update simulation with a 4 times smaller time step and a 100 ns EDS free-enthalpy evaluation

simulation needed to be carried out when using the EDS-OSP method; i.e., the total computational effort is equivalent to about $4 \times 7 + 100 = 128$ ns. Using the original OSP method, two 100 ns reference simulations and a 500 ns perturbed-state simulation needed to be carried out, i.e., the total computational effort was equivalent to $2 \times 100 + 500 = 700$ ns.[18] Thus, similar free enthalpy results were obtained at the cost of more than 5 times less computational effort in the present study compared with the previous one,[18] which shows the efficiency of the proposed EDS-OSP method.

5. CONCLUSION

One-step perturbation (OSP) is a free enthalpy calculation method that uses one or a few simulations of a possibly unphysical reference-state Hamiltonian to predict many free enthalpies between real-state Hamiltonians. Previous studies showed that with a carefully chosen reference state, OSP can be used to predict the folding free enthalpies of many not too different peptides. Enveloping distribution sampling (EDS) is another efficient and powerful free enthalpy calculation method in which a reference-state Hamiltonian enveloping all end states is automatically derived and simulated. In a single simulation of this EDS reference-state Hamiltonian, all the end states can be sampled.

In this article, we proposed a combination (EDS-OSP) of the EDS and OSP methods and applied it to predict the free enthalpy differences between a right-handed $2.7_{10/12}$-helix and a left-handed 3_{14}-helix for 16 β-peptides with slightly different side-chain substitution patterns. An EDS simulation of a soft-core reference-state peptide was first carried out, in which both helices were sampled. Thus, the free enthalpy difference between the two helices for the reference-state Hamiltonian can be calculated. Then, the soft-core atoms were perturbed into physical atoms to calculate the free enthalpy differences between the two helices for all the perturbed-state peptides.

EDS-OSP and the original OSP method give very similar results, i.e., the deviations between the results predicted by the two methods are within the order of k_BT for 15 out of the 16 peptides, and the average absolute deviation is 1.2 kJ mol^{-1}. In regard to computational effort, 700 ns of MD simulation had been carried out previously using the OSP method, while only about 128 ns of MD simulation needed to be carried out using the EDS-OSP method, a 4×7 ns EDS parameter update simulation with a lower solvent shear viscosity and a 100 ns EDS free-enthalpy evaluation simulation. That is, the computational effort was reduced by more than a factor of 5 using EDS-OSP compared with the original OSP method, which shows the advantage of the presented EDS-OSP method.

Hence, we consider the evaluated EDS-OSP method, which combines the advantages of the EDS and the OSP methods, an efficient tool to predict conformational free enthalpy differences. Its application to even larger conformational transitions, e.g., in proteins, would require, however, much longer simulations.

AUTHOR INFORMATION

Corresponding Author

*E-mail: wfvgn@igc.phys.chem.ethz.ch.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was financially supported by the National Center of Competence in Research (NCCR) in Structural Biology and by grant number 200020-137827 of the Swiss National Science Foundation and by grant number 228076 of the European Research Council (ERC), which is gratefully acknowledged.

REFERENCES

(1) Brooks, C. L., 3rd *Acc. Chem. Res.* **2002**, *35*, 447−454.
(2) Daggett, V.; Fersht, A. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 497−502.
(3) Shakhnovich, E. *Chem. Rev.* **2006**, *106*, 1559−1588.
(4) Gromiha, M. M. *Biochem. Soc. Trans.* **2007**, *35*, 1569−1573.
(5) Khan, S.; Vihinen, M. *Hum. Mutat.* **2010**, *31*, 675−684.
(6) Masso, M.; Vaisman, I. I. *Bioinformatics* **2008**, *24*, 2002−2009.
(7) Gilis, D.; Rooman, M. *J. Mol. Biol.* **1996**, *257*, 1112−1126.
(8) Bordner, A. J.; Abagyan, R. A. *Proteins: Struct. Funct. Bioinf.* **2004**, *57*, 400−413.
(9) Zhang, Z.; Wang, L.; Gao, Y.; Zhang, J.; Zhenirovskyy, M.; Alexov, E. *Bioinformatics* **2012**, *28*, 664−671.
(10) Muff, S.; Caflisch, A. *Proteins: Struct. Funct. Bioinf.* **2007**, *70*, 1185−1195.
(11) Baumketner, A.; Krone, M. G.; Shea, J. E. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6027−6032.
(12) Roychaudhuri, R.; Yang, M.; Hoshi, M. M.; Teplow, D. B. *J. Biol. Chem.* **2008**, *284*, 4749−4753.
(13) Shao, Q.; Wei, H.; Gao, Y. Q. *J. Mol. Biol.* **2010**, *402*, 595−609.
(14) Lin, Z. X.; Kornfeld, J.; Machler, M.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **2010**, *132*, 7276−7278.
(15) Liu, H. Y.; Mark, A. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1996**, *100*, 9485−9494.
(16) Lin, Z. X.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2010**, *12*, 15442−15447.
(17) Lin, Z. X.; Hodel, F. H.; van Gunsteren, W. F. *Helv. Chim. Acta* **2011**, *94*, 597−610.
(18) Lin, Z. X.; van Gunsteren, W. F. *J. Phys. Chem. B* **2011**, *115*, 12984−12992.
(19) Seebach, D.; Matthews, J. L. *Chem. Commun.* **1997**, *21*, 2015−2022.
(20) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219−3232.
(21) Gellman, S. H. *Acc. Chem. Res.* **1998**, *31*, 173−180.
(22) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, *101*, 3893−4011.
(23) Christ, C. D.; van Gunsteren, W. F. *J. Chem. Phys.* **2007**, *126*, 184110.
(24) Christ, C. D.; van Gunsteren, W. F. *J. Chem. Phys.* **2008**, *128*, 174112.
(25) Christ, C. D.; van Gunsteren, W. F. *J. Comput. Chem.* **2009**, *30*, 1664−1679.
(26) Christ, C. D.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2009**, *5*, 276−286.
(27) Riniker, S.; Christ, C. D.; Hansen, N.; Mark, A. E.; Nair, P. C.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *135*, 024105.
(28) Hansen, N.; Dolenc, J.; Knecht, M.; Riniker, S.; van Gunsteren, W. F. *J. Comput. Chem.* **2012**, *33*, 640−651.
(29) Lin, Z. X.; Liu, H. Y.; Riniker, S.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2011**, *7*, 3884−3897.
(30) Lin, Z. X.; Timmerscheidt, T. A.; van Gunsteren, W. F. *J. Chem. Phys.* **2012**, *137*, 064108.
(31) Seebach, D.; Abele, S.; Gademann, K.; Guichard, G.; Hintermann, T.; Jaun, B.; Matthews, J. L.; Schreiber, J. V. *Helv. Chim. Acta* **1998**, *81*, 932−982.
(32) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.
(33) Lin, Z. X.; Schmid, N.; van Gunsteren, W. F. *Mol. Phys.* **2011**, *109*, 493−506.

(34) Eichenberger, A. P.; Allison, J. R.; Dolenc, J.; Geerke, D. P.; Horta, B. A. C.; Meier, K.; Oostenbrink, C.; Schmid, N.; Steiner, D.; Wang, D. Q.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2011**, *7*, 3379−3390.
(35) Riniker, S.; Christ, C. D.; Hansen, H. S.; Hünenberger, P. H.; Oostenbrink, C.; Steiner, D.; van Gunsteren, W. F. *J. Phys. Chem. B* **2011**, *115*, 13570−13577.
(36) Kunz, A. P. E.; Allison, J. R.; Geerke, D. P.; Horta, B. A. C.; Hünenberger, P. H.; Riniker, S.; Schmid, N.; van Gunsteren, W. F. *J. Comput. Chem.* **2012**, *33*, 340−353.
(37) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F. *Comput. Phys. Commun.* **2012**, *183*, 890−903.
(38) GROMOS. http://www.gromos.net (accessed Dec. 2012).
(39) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. *J. Comput. Chem.* **2001**, *22*, 1205−1218.
(40) Walser, R.; Mark, A. E.; van Gunsteren, W. F.; Lauterbach, M.; Wipff, G. *J. Chem. Phys.* **2000**, *112*, 10450−10459.
(41) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(42) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(43) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(44) Zagrovic, B.; Pande, V. *J. Comput. Chem.* **2003**, *24*, 1432−1436.
(45) Gee, P. J.; van Gunsteren, W. F. *Chem.—Eur. J.* **2006**, *12*, 72−75.
(46) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(47) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987; pp 191−195.