Molecular Simulations of Thermodynamic Properties for the System α-Cyclodextrin/Alcohol in Aqueous Solution

Daniel Markthaler[1], Julia Gebhardt[1], Sven Jakobtorweihen[2], and Niels Hansen[1,*]

DOI: 10.1002/cite.201700057

Dedicated to Prof. Dr. Frerich J. Keil on the occasion of his 70th birthday

Free-energy calculations based on molecular simulations provide access to a wide range of thermodynamic properties such as the solubility and partitioning of a molecule in and between various phases. It is demonstrated how molecular dynamics free-energy simulations may be used to obtain the solubilities of primary alcohols and *n*-alkanes in water and binding affinities of primary alcohols to α-cyclodextrin. The equivalence of two distinct routes to calculate binding free energies is shown leading to the conclusion that host-guest binding affinities may be used to probe the underlying molecular force field.

Keywords: Force field, Free energy, Host-guest complex, Molecular dynamics, Statistical mechanics

Received: May 28, 2017; *accepted:* July 27, 2017

1 Introduction

Methods to predict physicochemical properties of complex systems based on statistical mechanics have gained an ever increasing importance in chemical engineering research [1]. They have been percolating a diverse field of applications ranging from adsorption in porous media over protein engineering to materials design [2]. Atomistic simulation has seen a particular increase in relevance due to its permanently expanding scope that is directly linked to the increase in computational speed, the use of massively parallel hardware architecture [3, 4] as well as acceleration through the use of graphics processors [5]. Molecular simulation techniques using classical force fields is probably the most direct way to theoretically explore molecular behavior in the condensed phase that may not be accessible experimentally at all, or may be impeded by high cost or the difficulty of detangling multivariate relationships through indirect experimental measurements, making this tool complementary to experimental techniques. Molecular simulations are used in a discovery driven mode (i.e., to explain or search for new phenomena) but also more and more in a data-driven manner (i.e., to predict properties, test theories and models, or to validate difficult experiments) [6]. Property prediction is directly linked to two major challenges in molecular simulation, referred to as the force-field problem and the sampling problem [7]. These two problems are interrelated. Only if a calculated property is sufficiently converged can an assessment of the quality of a force field be carried out.

With regard to chemical engineering applications, molecular simulations may give access to a range of properties needed in practice [8]. Important physicochemical parameters in for example purification and separation technology, drug design and environmental chemistry are the solubility and partitioning of molecules in and between various media. Therefore, there is great need for accurate predictions of solubilities and partition or distribution coefficients of known compounds but also of hypothetical ones that have not yet been identified. Molecular dynamics simulations are usually not resorted to when predicting solubility because other methods often require less computational effort [9 – 13]. Although these methods may offer very accurate

[1]Daniel Markthaler, Julia Gebhardt, Dr.-Ing. Niels Hansen
hansen@itt.uni-stuttgart.de
University of Stuttgart, Institute of Thermodynamics and Thermal Process Engineering, Pfaffenwaldring 9, 70569 Stuttgart, Germany.
[2]Dr.-Ing. Sven Jakobtorweihen
Hamburg University of Technology, Institute of Thermal Separation Processes, Eißendorfer Straße 38, 21073 Hamburg, Germany.

and fast prediction in many cases, their predictive power depends on the chemical structure and definite methods that are best to use are not easy to identify [14, 15]. Atomistic simulations have the benefit that they may provide additional conceptual understanding of the underlying molecular level forces driving the solvation process and offer the possibility to obtain a diverse set of physical properties of a complex system based on the same underlying potential energy function (i.e., the molecular force field) which enhances the interpretability of molecular simulation results. For drug-like or polyfunctional molecules which are often of interest for simulation studies few experimental data exist on the solvation free energy, a quantity, which is often used as target property in force field parametrization and validation [16 – 19]. In contrast data on solubility are often more abundant, making the calculation of solubilities – even relative solubilities – particularly important [20].

Solubilities and other physical properties of high practical interest such as activity coefficients, Henry's law constants, and distribution of chemical species between different phases are related to the free energy of solvation [21]. As a result such calculations have seen a tremendous progress over the last decades and standards are beginning to evolve that allow for reproduction of published results and reduce the barriers to perform such calculations routinely [22 – 24]. However, partitioning is not only relevant between homogeneous phases but also between a bulk phase and a heterogeneous phase such as a membrane [25], a micelle [26], a stationary phase in liquid chromatography [27], or a larger biomolecule such as a protein [28]. Therefore, a similar active area of research is the prediction of binding affinities for biomolecular complexes from molecular dynamics simulations due to the expected impact such calculations will have in early stage drug discovery. Yet these calculations have a higher level of complexity and are, thus, less mature than those devoted to solvation free energies [29]. Two types of approaches are principally used to compute standard binding free energies using molecular dynamics simulations, namely the alchemical double decoupling method and the potential of mean force method, respectively [30 – 33]. In the present work the two approaches are compared using the example of complexation between α-cyclodextrin and primary alcohols.

The purpose of the present article is twofold. First, the possibility of making predictions of aqueous solubilities of *n*-alkanes and primary alcohols from free-energy calculations by utilizing a relationship between solubility, free energy or rather free enthalpy [34] or Gibbs energy [35] of solvation, and solute vapor pressure that is strictly valid only when all activity and fugacity coefficients are unity is investigated. Second it is demonstrated that binding free enthalpies of host-guest systems can be obtained that are method-independent and thus, solely a function of the underlying potential energy function (i.e., the molecular force field).

2 Aqueous Solubility of *n*-Alkanes and Primary Alcohols

Consider a solute that is liquid in its pure state and saturated in the solvent at such dilute concentrations that Henry's law is obeyed, that is the saturated solution is infinitely diluted such that the rational activity coefficient is equal to one. Then, (aqueous) solubility expressed in molar concentration of a solute can be estimated given the free enthalpy of solvation of a single solute molecule A and its pure-substance vapor pressure from [36, 37]

$$c_{\mathrm{A}}^{\mathrm{s}} = \left(\frac{P_{\mathrm{A}}^{\mathrm{sat}}}{P^{\mathrm{o}}}\right) \exp\left[\frac{-\Delta G_{\mathrm{A,solv}}^{\mathrm{s}}}{RT}\right] \tag{1}$$

where R is the universal gas constant, T is the temperature, $P_{\mathrm{A}}^{\mathrm{sat}}$ is the equilibrium vapor pressure of A over pure A, P^{o} is the pressure (24.77 bar) of an ideal gas at 1 molar concentration and 298.15 K. Alternatively the solubility can be expressed in units of mole fraction according to [38]

$$x_{\mathrm{A}}^{\mathrm{s}} = \frac{P_{\mathrm{A}}^{\mathrm{sat}}}{RT\rho_{\mathrm{s}}} \exp\left[\frac{-\Delta G_{\mathrm{A,solv}}^{\mathrm{s}}}{RT}\right] \tag{2}$$

where ρ_{s} denotes the molar density of the solvent and the right hand side can be identified as the ratio between the vapor pressure and Henry's law constant of the solute [39]. Under the assumption of small values for x the relation between both definitions is [40]

$$c_{\mathrm{A}}^{\mathrm{s}} = x_{\mathrm{A}}^{\mathrm{s}}\rho_{\mathrm{s}} \tag{3}$$

These relations were applied frequently to calculate the solubility of light gases in ionic liquids [41], in water, or aqueous solutions [42, 43], or the solubility of *n*-alkanes in water [44], respectively. For the prediction of the solubility of solutes that are solid in their pure state see [45, 46].

In the present work, Eq. (2) is examined to estimate the solubility of *n*-alkanes and primary alcohols in water. Fig. 1 shows the results of calculated solubilities up to *n*-nonane which are in very good agreement with experimental data [47]. The required vapor pressures for *n*-alkanes lighter than *n*-pentane, which are in the gas state at 298.15 K, was taken to be equal to the total pressure 1 bar. For *n*-pentane to *n*-nonane, which are liquids under these conditions the vapor pressures were taken from the literature [47]. The hydration free enthalpies were obtained from reported molecular dynamics simulations using the GROMOS 45A3 force field [49] except for methane for which the hydration free enthalpy was recalculated in the present work due to conflicting values reported in [49, 50]. Here the value of $8\,\mathrm{kJ\,mol^{-1}}$ as reported in [50] was confirmed. Note that there is no difference between the GROMOS force fields 45A3 and 53A6 [49] for the alkanes considered in this work. For alkanes longer than those considered in this work, experimental data from different sources become inconsistent, often showing a sudden change in the depend-

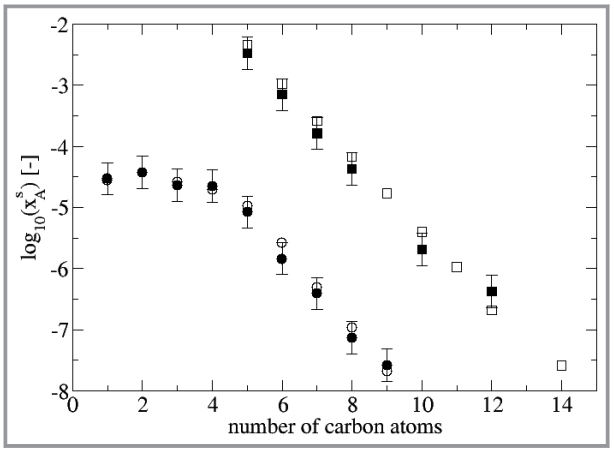

Figure 1. Aqueous solubility of *n*-alkanes (circles) and linear alcohols (squares) at 298.15 K and 1 bar. The simulation results (filled symbols) were obtained with the GROMOS 53A6 force field assuming an error of ±1.5 kJ mol^{-1} in the free enthalpies of hydration. The experimental data (open symbols) were taken from [47, 48], respectively.

ence of solubility upon carbon number, while simulation results do not support this observation [44].

Compared to *n*-alkanes, primary alcohols interact more strongly with the solvent such that the validity of Eqs. (1) and (2) is more in doubt [51]. However, Fig. 1 shows very good agreement between calculated and experimental [48] data also for this system. The required vapor pressures were taken from the literature (C5 to C8 [52], C10 [53], C12 [54]). The hydration free enthalpies were obtained from molecular dynamics simulations with the GROMOS 53A6 force field. For 1-butanol to 1-octanol the values reported in [55] were used while the increments in hydration free enthalpy with carbon number relative to 1-octanol calculated in a recent work [56] were used to obtain values for 1-decanol and 1-dodecanol. Finally it is noted that if the assumption that all activity coefficients are unity proves to be unrealistic, finite concentration activity coefficients may also be obtained from molecular dynamics free-energy calculations [57]. For saturated solutions of solid solutes the need to calculate the residual chemical potential of the solid hampers the use of molecular simulations [58]. However, molecular dynamics simulations can straightforwardly be used to calculate relative solubilities for the same solute in different solvents [20]. Therefore, we conclude that molecular dynamics simulations with recent force fields will gain an increasing importance in predicting solubilities in complex systems. The same holds true for binding affinities which are considered in the next section.

3 Standard Binding Free Enthalpy of Primary Alcohols to α-Cyclodextrin

Cyclodextrins are often used as model systems to study binding mechanisms [59 – 61] but are also of much practical interest themselves, as chiral selectors [62 – 67], molecular reactors [68], drug solubility agents [69], and in many other applications [70 – 74]. Complexation thermodynamics of cyclodextrins was widely studied experimentally [75, 76]. Computationally the standard binding free enthalpy can be obtained through two distinct routes, one relying on alchemical transformations (decoupling) along a non-physical path and the other on probabilities of being at a (predefined) physical path that may (but does not have to) resemble the true reaction coordinate. The probabilities are related to the potential of mean force (PMF) from which the binding free enthalpy can be calculated. Both routes may have advantages and drawbacks for the particular system to be studied [32]. However, from the statistical mechanics point of view they have to result in the same number, a condition that may be used for validation and error detection [77]. If the two routes provide the same results within the statistical errors, the calculated binding free enthalpy is most likely the one prescribed by the underlying molecular force field. In a previous work the standard binding free enthalpy of primary alcohols from 1-butanol to 1-dodecanol to α-cyclodextrin for different GROMOS force fields using the alchemical route was calculated [56]. In the present work these simulations were extended for alcohols up to 1-eicosanol (C20) using the GROMOS 53A6$_{GLYC}$ [78] force field, applying the same methodology as in the previous work [56]. Additionally the PMF route was examined for 1-butanol to 1-dodecanol using the GROMOS 53A6$_{GLYC}$ and CHARMM36 [79] force fields. To calculate the standard binding free enthalpy via the PMF route a one-dimensional setup [80] was chosen:

- The six glycosidic oxygens (O1) were positionally restrained using harmonic potentials with a force constant of 1000 kJ mol^{-1}nm^{-2} such that the central axis of the α-cyclodextrin was aligned with the *z*-axis (see Fig. 2 for the definition of the coordinate system and the nomenclature used to specify the atoms in α-cyclodextrin).
- A position restraint acting on the center of mass of the guest molecule was used to prevent it from moving too far orthogonal to the *z*-axis in order to restrict the sampled area. This restraint was applied in the form of a flat-bottom potential depending on the distance between the center of mass of the guest and the *z*-axis:

$$U_R = \begin{cases} K_{xy}(r_{com} - r_{res})^4_{xy} & \text{if } r_{com,xy} > r_{res,xy} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

- where the force constant K_{xy} was set to 7500 kJ mol^{-1}nm^{-4}, and the switching distance $r_{res,xy}$ was set to 0.6 nm. The quantity $r_{com,xy}$ denotes the center of mass coordinate of the guest molecule in the *x*-*y*-plane. With this choice of $r_{com,xy}$ the restraint has no contribution in the bound state.
- An angle restraint was applied, defined in terms of two vectors, the *z*-axis and the molecular axis of the alcohol molecule. The latter was determined by the oxygen atom and the carbon atom of the methyl group at the opposite

end of the alcohol molecule. This restraint was used to avoid a flip of the alcohol molecule and was applied in the form of a harmonic potential in the angle. The restraint energy is, thus,

$$U_\theta = \frac{1}{2} K_\theta (\theta - \theta_0)^2 \tag{5}$$

where the force constant K_θ was set to 500 kJ mol^{-1}rad^{-2} and the reference angle θ_0 to 0°.

- Umbrella sampling [81] with windows placed at successive positions each 0.1 nm apart along the z-axis was performed using the z-component of the distance between the center of mass of the alcohol and the center of mass of α-cyclodextrin as order parameter. A harmonic umbrella potential with a force constant of 500 kJ mol^{-1}nm^{-2} that acts on the deviation of the z-coordinate of r_{com} from the reference position of the respective umbrella window was applied. Molecular dynamics simulations of 20 ns were carried out per window using the GROMACS program package [82] (version 5.1.4 for simulations with the GROMOS force field and version 2016.x for those with the CHARMM force field) patched to the free-energy library PLUMED 2.3.0 [83] that handles the definition of the restraints.
- The free-energy profiles were obtained applying the umbrella integration (UI) [84, 85] method using a freely available python implementation [86], from the weighted histogram analysis (WHAM) [87] implementation offered by the GROMACS program package [88] and additionally by the multistate Bennett acceptance ratio approach (MBAR) [89] for which a freely available python implementation [90] was employed.

Many other different setups are of course possible. However, below it will be shown that this particular setup leads to reliable results. Furthermore, the rationale for using this particular setup will be discussed in a future publication. From this 1D setting, the standard binding free enthalpy can be calculated by [80]

$$\Delta G^{o}_{bind} = -RT \ln \frac{l_b A_{u,R}}{V^o} - RT \ln \frac{\Omega}{8\pi^2} + \Delta W_R + RT \ln \left\langle \exp\left[\frac{-K_\theta(\theta-\theta_0)^2}{2RT}\right] \right\rangle_{K_\theta=0} \tag{6}$$

where V^o is the standard state volume of 1.661 nm^3 and l_b denotes the bound length defined as Boltzmann-weighted physical length integrated over the bound region b,

$$l_b = \int_b \exp\left[\frac{-W_R(z)}{RT}\right] dz \tag{7}$$

where $W_R(z)$ is the potential of mean force as a function of z and defined to be zero at its lowest point where the guest molecule is bound. The bound region was defined to span the range between the minimum and the flat part of the PMF and was set to encompass a distance of 1.5 nm in all cases. Due to the small weights of larger distances this choice has no significant effect on the standard binding free enthalpy. The cross-sectional area $A_{u,R}$ available for the unbound ligand in x-y directions in the presence of the applied position restraint is obtained from the partition function of the restraining potential by

$$A_{u,R} = \int_0^\infty 2\pi r \exp\left[\frac{-U_R}{RT}\right] dr = \pi r_{res,xy}^2 + \frac{\pi r_{res,xy} \Gamma(0.25)}{2k_R^{0.25}} + \frac{\pi\sqrt{\pi}}{2\sqrt{k_R}} \tag{8}$$

where Γ denotes the Gamma function and k_R the dimensionless restraining force constant K_{xy}/RT. With the numerical values given above $A_{u,R}$ evaluates to 1.64 nm^2.

The contribution to the free enthalpy in the unbound state due to the angle restraint arises from the reorientation permitted about the three Euler angles of a rigid body in a homogeneous bulk phase. For the restraining potential given by Eq. (5) the available rotational volume Ω under the influence of the restraint is [91, 92]

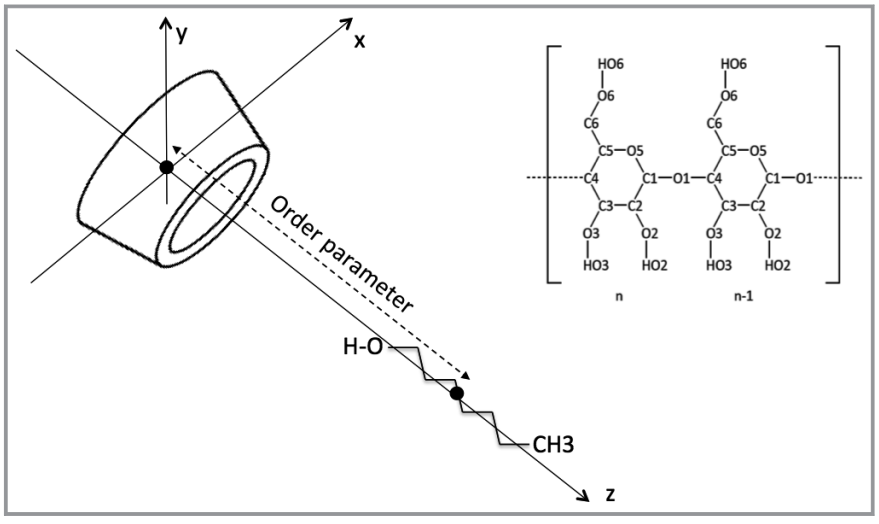

Figure 2. Schematic representation of the host-guest arrangement in configuration 2 (conf2) in which the hydroxyl group of the alcohol in the bound state is located at the secondary rim of the α-cyclodextrin. On the right-hand side the nomenclature used to specify the atoms in the monomeric cyclodextrin units are shown.

$$\Omega = \int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi} \exp\left[\frac{-U_\theta}{RT}\right] \sin\theta \mathrm{d}\theta\mathrm{d}\phi\mathrm{d}\psi = 4\pi^2 \int_0^{\pi} \exp\left[\frac{-U_\theta}{RT}\right] \sin\theta \mathrm{d}\theta \tag{9}$$

and evaluates to 0.1956 using the numerical values given above. The quantity ΔW_R is the depth of the potential of mean force and the last term in Eq. (6) accounts for the contribution of releasing the angle restraint in the bound state. This contribution may be evaluated by free energy perturbation [93, 94] (as indicated in Eq. (6)) or thermodynamic integration [95] over the derivative of the restraining potential with respect to a coupling parameter. Here the first approach is used. The results depend slightly on the specific case and are in the range –5.0 to –7.6 kJ mol^{-1}.

Figs. 3 and 4 show the potentials of mean force for all alcohol molecules studied. Because the free enthalpy is a state function the left and the right branch of the PMFs should result in the same standard binding free enthalpy when evaluated with Eq. (6). Therefore, the depths of the PMF when evaluated from the left or from the right PMF branch, respectively, are related by $\Delta W_{R,left} - \Delta W_{R,right} = RT\ln(l_{b,left}/l_{b,right})$. For the simulations conducted in the present work the difference $\Delta W_{R,left} - \Delta W_{R,right}$ was usually larger in magnitude than the expression $RT\ln(l_{b,left}/l_{b,right})$ but in most cases within the thermal energy of *RT*, and thus, within the uncertainties associated with ΔW_R. For the final binding free enthalpy of a particular configuration (i.e., conf1 or conf2) the arithmetic mean value of the two branches were taken. The results show that the PMF depth for shorter alcohols is larger in configuration 2 (Fig. 3) compared to configuration 1 (Fig. 4), in agreement with an earlier work [56]. For longer chains the difference between the two configurations vanishes. Compared to the PMFs ob-

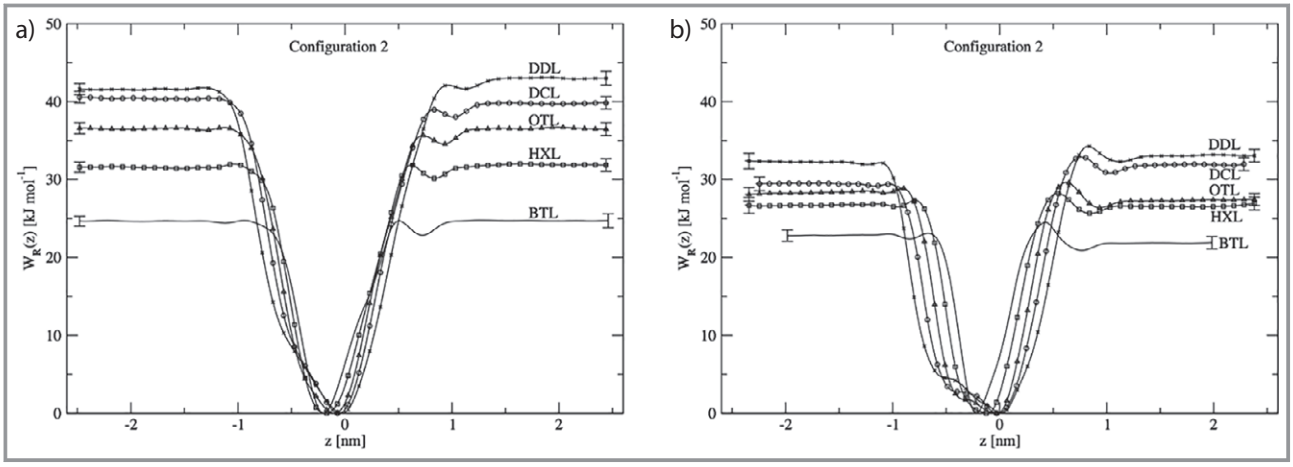

Figure 3. 1D potentials of mean force from umbrella sampling simulations, evaluated with the umbrella integration method, for α-cyclodextrin/alcohol complexes in configuration 2 using the *z*-component of the distance between the center of mass of the alcohol molecule and that of α-cyclodextrin as order parameter. Error estimates were performed according to [96]. a) GROMOS 53A6$_{GLYC}$ force field and b) CHARMM36 force field.

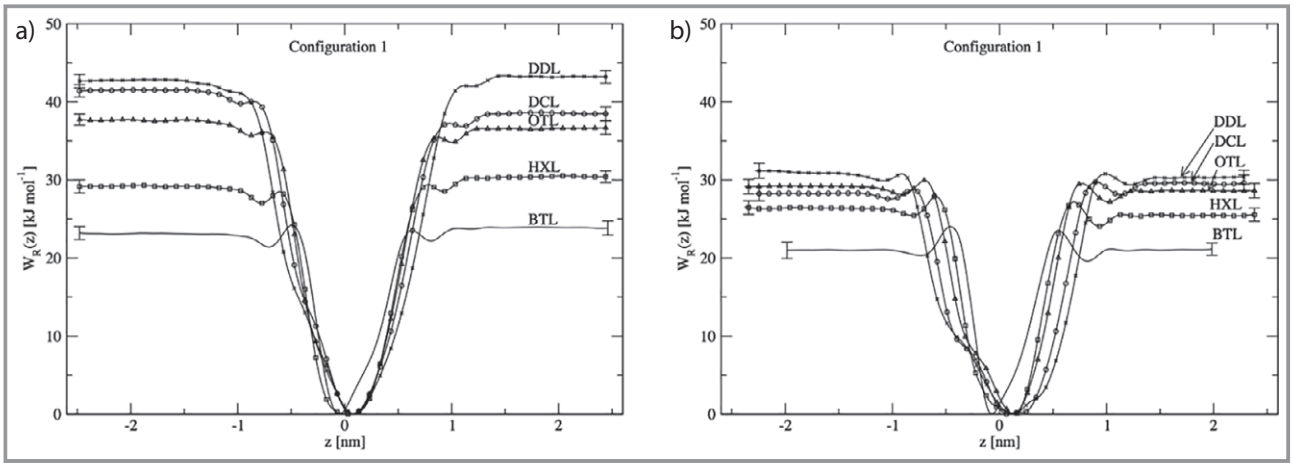

Figure 4. 1D potentials of mean force from umbrella sampling simulations, evaluated with the umbrella integration method, for α-cyclodextrin/alcohol complexes in configuration 1 using the *z*-component of the distance between the center of mass of the alcohol molecule and that of α-cyclodextrin as order parameter. Error estimates were performed according to [96]. a) GROMOS 53A6$_{GLYC}$ force field and b) CHARMM36 force field.

tained with the CHARMM36 force field the GROMOS 53A6 ones show deeper well depths ΔW_R, while the overall shape is very similar. It is noted that the barriers observed in some PMF curves between the flat region and the borders of the well do not necessarily have a physical significance but are most likely a consequence of the 1D setup, which does not resemble the true reaction coordinate [97]. Standard binding free enthalpies derived from experimental measurements do not distinguish between two different binding modes. Therefore, the logarithmic average [98] of the two values corresponding to the two configurations was taken to arrive at a final value for each alcohol molecule which can be compared to experimental data. Fig. 5 shows that the ΔG^o_{bind}-results for the GROMOS force field obtained by double decoupling and the PMF route are in very good agreement, leading to the conclusion that they reflect the properties of the underlying potential energy function. Therefore, the earlier conclusion that the GROMOS 53A6$_{GLYC}$ force field slightly overestimates the binding strength whereas the CHARMM36 force field underestimates it is confirmed. Note that the method used to analyze the umbrella sampling simulations (UI vs. WHAM vs. MBAR) had no significant effect on the results.

The simulation results clearly suggest an increase in magnitude of the binding free enthalpy as function of chain length and no sudden change in this trend even for the longest alcohol, 1-eicosanol. This is in contrast to some of the experimental studies but in agreement with the recent measurements of Linden et al. [100] who reported binding constants for the entire homologous series from 1-butanol to 1-dodecanol. To rationalize the increase in magnitude of ΔG^o_{bind} the binding modes were studied in more detail. Fig. 6 shows that the bulk water phase starts at ±0.8 nm (measured from the O1-atoms). Even for the C20 alcohol the complete chain is within this range (Fig. 7). That is, all parts of the guest molecule are interacting with the α-cyclodextrin host and contribute to the binding free enthalpy.

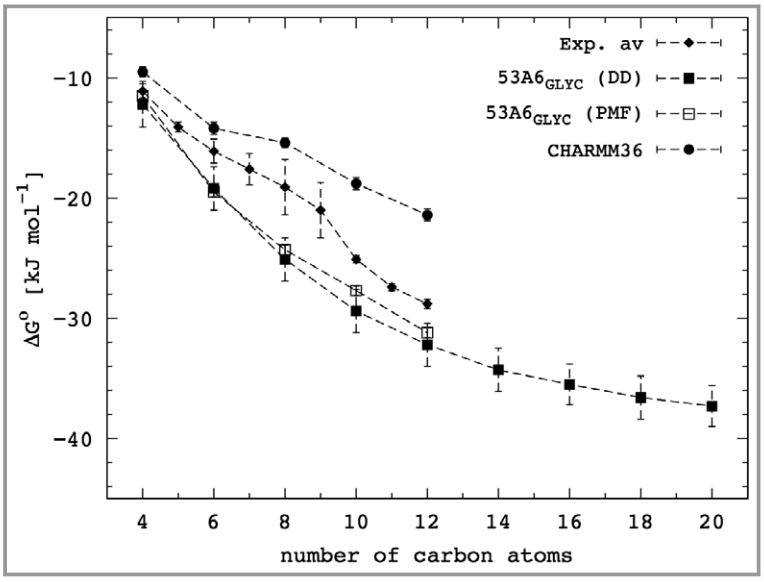

Figure 5. Experimental and calculated standard binding free enthalpies versus the number of carbon atoms. Experimental data (exp. av.) are based on our compilation reported earlier [56] supplemented by additional data [99]. Simulation results are reported for the GROMOS force field 53A6$_{GLYC}$ [78] calculated either by alchemical double decoupling (DD) [56] or via the potential or mean force (PMF) and for the force field CHARMM36 [79] calculated via the potential of mean force. Dashed lines are used as guide to the eye.

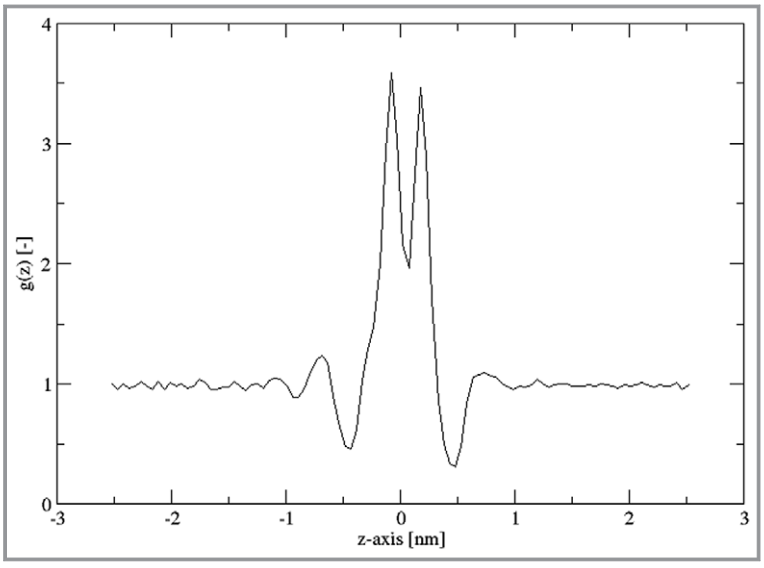

Figure 6. One-dimensional distribution function of water molecules along the central axis of α-cyclodextrin. The origin lies in the center of mass of the glycosidic oxygen atoms. The positive direction points towards the primary hydroxyl groups (primary rim), the negative direction points towards the secondary hydroxyl groups (secondary rim). All water molecules that are located within a cylindrical volume of radius 0.1 nm around the central axis were considered. Simulations were conducted with the GROMOS 53A6$_{GLYC}$ force field [78].

4 Conclusion and Outlook

The calculation of aqueous solubilities from molecular dynamics free-energy simulation was demonstrated for *n*-alkanes and primary alcohols. For the latter binding affinities to α-cyclodextrin were calculated via the potential of mean force. The equivalence of the results with earlier double decoupling simulations shows that the binding affinity is merely a function of the underlying potential energy function. This opens up new routes for force field optimization as a more diverse set of experimental target data can be incorporated. Indeed a growing body of literature deals with identifying inaccuracies in current force fields, which remain undetected when considering

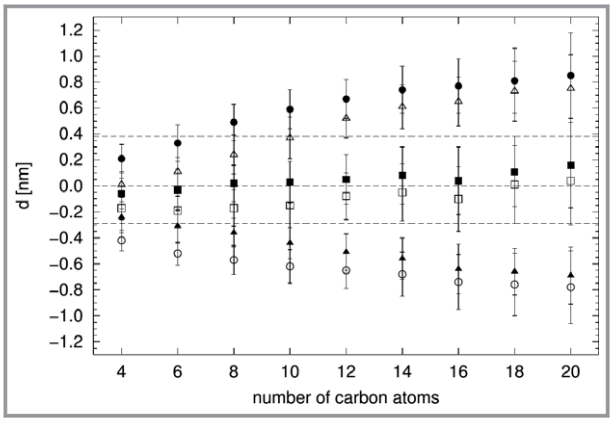

Figure 7. Characterization of the binding modes of primary alcohols in α-cyclodextrin obtained from simulations with the GROMOS 53A6$_{GLYC}$ [78] force field. The coordinate d measures the z-component of the distance between a reference point in α-cyclodextrin (center of mass of C1 atoms) and a reference point of the alcohol molecule. For the latter three choices were considered: (i) the CH_3-end of the alcohol (triangles), (ii) the center of mass of the two central C-atoms of the aliphatic chain (squares), and (iii) the oxygen atom of the hydroxyl group (circle). Filled symbols correspond to conf1 and open symbols to conf2. The horizontal dashed lines illustrate the size of the α-cyclodextrin cavity. The line at $d = 0$ represents the position of the α-cyclodextrin reference point, while the lines at $d = 0.38$ nm and $d = -0.29$ nm depict the average distance of the oxygen atoms of the primary hydroxyl groups and the secondary hydroxyl groups, respectively, from the α-cyclodextrin reference point.

only the thermodynamic properties used typically in the original parametrizations [101]. By considering both solubility and binding affinity on equal footing a thermodynamical consistent description of complex host-guest systems in complex solvent environments will become possible in the future.

This work was financially supported by the German Research Foundation (DFG) through the Cluster of Excellence in Simulation Technology (EXC 310/2) and the collaborative research center SFB 716, both at the University Stuttgart, and by the Baden-Württemberg Stiftung. Simulations were performed on the computational facilities bwUniCluster and BinAC funded by the Ministry of Science, Research, and Arts and the Universities of the State of Baden-Württemberg, Germany, within the framework program bwHPC and the DFG through grant no INST 39/963-1 FUGG and on resources provided by The North-German Supercomputer Alliance (HLRN). This contribution is dedicated to Frerich J. Keil on the occasion of his 70th birthday. SJ and NH express him their gratitude for long-standing advice and support.

List of Symbols

$A_{u,R}$	[nm^2]	cross-sectional area available to the guest molecule under the influence of the position restraint orthogonal to the z-axis
c_A^s	[mol L^{-1}]	solubility of compound A in solvent s
$\Delta G_{A,solv}^s$	[kJ mol^{-1}]	free enthalpy of solvation of compound A in solvent s
ΔG_{bind}^0	[kJ mol^{-1}]	standard binding free enthalpy
K_{xy}	[kJ mol^{-1}nm^{-4}]	restraining force constant of the distance restraint acting orthogonal to the z-axis
K_θ	[kJ mol^{-1}rad^{-2}]	restraining force constant of angle restraint
l_b	[nm]	bound length
p_A^{sat}	[bar]	vapor pressure of compound A
p^0	[bar]	pressure of ideal gas at 1 molar concentration and 298.15 K
r	[nm]	distance
R	[J mol^{-1}K^{-1}]	universal gas constant
T	[K]	temperature
U_R	[kJ mol^{-1}]	potential energy of the distance restraint
U_θ	[kJ mol^{-1}]	potential energy of the angle restraint
V^0	[nm^3]	standard state volume
ΔW_R	[kJ mol^{-1}]	PMF depth, defined as the lowest point, zero, minus the exponential average of the PMF over the entire unbound region
x_A^s	[–]	solubility of compound A in solvent s expressed as mole fraction

Greek symbols

θ	[rad]	angle between molecular axis of the guest molecule and the z-axis
ρ_S	[mol L^{-3}]	molar density of the solvent
Ω	[–]	rotational volume accessible to the guest molecule under the influence of the angle restraint

Abbreviations

MBAR	multistate Bennett acceptance ratio approach
PMF	potential of mean force
UI	umbrella integration
WHAM	weighted histogram analysis

References

[1] M. W. Deem, *AIChE J.* **1998**, *44 (12)*, 2569. DOI: 10.1002/aic.690441202

[2] K. E. Gubbins, J. D. Moore, *Ind. Eng. Chem. Res.* **2010**, *49 (7)*, 3026. DOI: 10.1021/ie901909c

[3] G. S. Heffelfinger, *Comput. Phys. Commun.* **2000**, *128 (1 – 2)*, 219. DOI: 10.1016/S0010-4655(00)00050-3

[4] J. C. Palmer, P. G. Debenedetti, *AIChE J.* **2015**, *61 (2)*, 370. DOI: 10.1002/aic.14706

[5] J. E. Stone, J. C. Phillips, P. L. Freddolino, D. J. Hardy, L. G. Trabuco, K. Schulten, *J. Comput. Chem.* **2007**, *28 (16)*, 2618. DOI: 10.1002/jcc.20829

[6] E. J. Maginn, *AIChE J.* **2009**, *55 (6)*, 1304. DOI: 10.1002/aic.11932

[7] W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. A. van der Vegt, H. B. Yu, *Angew. Chem. Int. Ed.* **2006**, *45 (25)*, 4064. DOI: 10.1002/anie.200502655

[8] E. Hendriks, G. M. Kontogeorgis, R. Dohrn, J.-C. de Hemptinne, I. G. Economou, L. F. Žilnik, V. Vesovic, *Ind. Eng. Chem. Res.* **2010**, *49 (22)*, 11131. DOI: 10.1021/ie101231b

[9] K. Göke, T. Lorenz, A. Repanas, F. Schneider, D. Steiner, K. Baumann, H. Bunjes, A. Dietzel, J. H. Finke, B. Glasmacher, A. Kwade, *Eur. J. Pharm. Biopharm.* in press. DOI: 10.1016/j.ejpb.2017.05.008

[10] A. Lusci, G. Pollastri, P. Baldi, *J. Chem. Inf. Model.* **2013**, *53 (7)*, 1563. DOI: 10.1021/ci400187y

[11] Y. Zhao, W. Liu, X. Pei, D. Yao, *Fluid Phase Equilib.* **2017**, *437*, 43. DOI: 10.1016/j.fluid.2017.01.006

[12] B. Bouillot, T. Spyriouni, S. Teychené, B. Biscans, *Eur. Phys. J.: Spec. Top.* **2017**, *226 (5)*, 913. DOI: 10.1140/epjst/e2016-60225-5

[13] A. Klamt, F. Eckert, M. Hornig, M. E. Beck, T. Bürger, *J. Comput. Chem.* **2002**, *23 (2)*, 275. DOI: 10.1002/jcc.1168

[14] R. E. Skyner, J. L. McDonagh, C. R. Groom, T. van Mourik, J. B. O. Mitchell, *Phys. Chem. Chem. Phys.* **2015**, *17 (9)*, 6174. DOI: 10.1039/c5cp00288e

[15] A. J. Hopfinger, E. X. Esposito, A. Llinas, R. C. Glen, J. M. Goodman, *J. Chem. Inf. Model.* **2009**, *49 (1)*, 1. DOI: 10.1021/ci800436c

[16] J. Fu, J. Wu, *Fluid Phase Equilib.* **2016**, *407*, 304. DOI: 10.1016/j.fluid.2015.05.042

[17] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25 (13)*, 1656. DOI: 10.1002/jcc.20090

[18] J. Zhang, B. Tuguldur, D. van der Spoel, *J. Chem. Inf. Model.* **2015**, *55 (6)*, 1192. DOI: 10.1021/acs.jcim.5b00106

[19] J. Zhang, B. Tuguldur, D. van der Spoel, *J. Chem. Inf. Model.* **2016**, *56 (4)*, 819. DOI: 10.1021/acs.jcim.6b00081

[20] S. Liu, S. Cao, K. Hoang, K. L. Young, A. S. Paluch, D. L. Mobley, *J. Chem. Theory Comput.* **2016**, *12 (4)*, 1930. DOI: 10.1021/acs.jctc.5b00934

[21] G. Raabe, *Molecular Simulation Studies on Thermophysical Properties*, Springer Nature, Singapore **2017**. DOI: 10.1007/978-981-10-3545-6

[22] G. D. R. Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts, D. L. Mobley, *J. Chem. Eng. Data* **2017**, *62 (5)*, 1559. DOI: 10.1021/acs.jced.7b00104

[23] A. Pohorille, C. Jarzynski, C. Chipot, *J. Phys. Chem. B* **2010**, *114 (32)*, 10235. DOI: 10.1021/jp102971x

[24] P. V. Klimovich, M. R. Shirts, D. L. Mobley, *J. Comput.-Aided Mol. Des.* **2015**, *29 (5)*, 397. DOI: 10.1007/s10822-015-9840-9

[25] D. Lopes, S. Jakobtorweihen, C. Nunes, B. Sarmento, S. Reis, *Prog. Lipid Res.* **2017**, *65*, 24. DOI: 10.1016/j.plipres.2016.12.001

[26] D. Yordanova, E. Ritter, T. Gerlach, J.-H. Jensen, I. Smirnova, S. Jakobtorweihen, *J. Phys. Chem. B* **2017**, *121 (23)*, 5794. DOI: 10.1021/acs.jpcb.7b03147

[27] S. M. Melnikov, A. Höltzel, A. Seidel-Morgenstern, U. Tallarek, *Angew. Chem. Int. Ed.* **2012**, *51 (25)*, 6251. DOI: 10.1002/anie.201201096

[28] R. Wedberg, J. Abildskov, G. H. Peters, *J. Phys. Chem. B* **2012**, *116 (8)*, 2575. DOI: 10.1021/jp211054u

[29] D. L. Mobley, M. K. Gilson, *Annu. Rev. Biophys.* **2017**, *46*, 531. DOI: 10.1146/annurev-biophys-070816-033654

[30] P. Procacci, R. Chelli, *J. Chem. Theory Comput.* **2017**, *13 (5)*, 1924. DOI: 10.1021/acs.jctc.6b01192

[31] Y. Deng, B. Roux, *J. Phys. Chem. B* **2009**, *113 (8)*, 2234. DOI: 10.1021/jp807701h

[32] J. C. Gumbart, B. Roux, C. Chipot, *J. Chem. Theory Comput.* **2013**, *9 (1)*, 794. DOI: 10.1021/ct3008099

[33] M. K. Gilson, J. A. Given, B. L. Bush, J. A. McCammon, *Biophys. J.* **1997**, *72*, 1047. DOI: 10.1016/S0006-3495(97)78756-3

[34] IUPAP, *Phys. A* **1978**, *93 (1 – 2)*, 1. DOI: 10.1016/0378-4371(78)90209-1

[35] IUPAC, *Quantities, Units and Symbols in Physical Chemistry*, Blackwell Scientific, Oxford **1988**.

[36] J. D. Thompson, C. J. Cramer, D. G. Truhlar, *J. Chem. Phys.* **2003**, *119*, 1661. DOI: 10.1063/1.1579474

[37] C. J. Cramer, D. G. Truhlar, in *Free Energy Calculations in Rational Drug Design* (Eds: M. R. Reddy, M. D. Erion), Springer, New York **2001**, 63 – 95.

[38] A. Ben-Naim, *Molecular Theory of Solutions*, Oxford University Press, Oxford **2006**.

[39] K. S. Shing, K. E. Gubbins, K. Lucas, *Mol. Phys.* **1988**, *65 (5)*, 1235. DOI: 10.1080/00268978800101731

[40] H. Gamsjäger, J. W. Lorimer, M. Salomon, D. G. Shaw, R. P. T. Tomkins, *Pure Appl. Chem.* **2010**, *82 (5)*, 1137. DOI: 10.1351/PAC-REP-09-10-33

[41] D. Kerlé, M. Namayandeh, R. Ludwig, S. Wohlrab, D. Paschek, *Phys. Chem. Chem. Phys.* **2017**, *19 (3)*, 1770. DOI: 10.1039/C6CP06792A

[42] Q. Chen, S. P. Balaji, M. Ramdin, J. J. Gutiérrez-Sevillano, A. Bardow, E. Goetheer, T. J. H. Vlugt, *Ind. Eng. Chem. Res.* **2014**, *53 (46)*, 18081. DOI: 10.1021/ie503488n

[43] M. Kohns, S. Werth, M. Horsch, E. von Harbou, H. Hasse, *Fluid Phase Equilib.* **2017**, *442*, 44. DOI: 10.1016/j.fluid.2017.03.007

[44] A. L. Ferguson, P. G. Debenedetti, A. Z. Panagiotopoulos, *J. Phys. Chem. B* **2009**, *113 (18)*, 6405. DOI: 10.1021/jp811229q

[45] A. S. Paluch, D. D. Cryan, E. J. Maginn, *J. Chem. Eng. Data* **2011**, *56 (4)*, 1587. DOI: 10.1021/je101251n

[46] A. S. Paluch, E. J. Maginn, *AIChE J.* **2013**, *59 (7)*, 2647. DOI: 10.1002/aic.14020

[47] D. Mackay, W. Y. Shiu, *J. Phys. Chem. Ref. Data* **1981**, *10 (4)*, 1175. DOI: 10.1063/1.555654

[49] L. D. Schuler, X. Daura, W. F. van Gunsteren, *J. Comput. Chem.* **2001**, *22 (11)*, 1205. DOI: 10.1002/jcc.1078

[50] X. Daura, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **1998**, *19 (5)*, 535. DOI: 10.1002/(SICI)1096-987X(19980415)19:5<535::AID-JCC6>3.0.CO;2-N

[51] H. H. Uhlig, *J. Phys. Chem.* **1937**, *41 (9)*, 1215. DOI: 10.1021/j150387a007

[48] M. Goral, B. Wisniewska-Goclowska, A. Maczynski, *J. Phys. Chem. Ref. Data* **2006**, *35 (3)*, 1391. DOI: 10.1063/1.2203354

[52] K. Nasirzadeh, R. Neueder, W. Kunz, *J. Chem. Eng. Data* **2006**, *51 (1)*, 7. DOI: 10.1021/je049600u

[53] E. Olsen, *Molecules* **2001**, *6 (4)*, 370. DOI: 10.3390/60400370

[54] E. J. Davis, in *Aerosol Measurement: Principles, Techniques, and Applications*, 3rd Ed. (Eds: P. Kulkarni, P. A. Baron, K. Willeke),

Wiley, Hoboken **2011**, 417 – 434. DOI: 10.1002/9781118001684.ch19
[55] B. A. C. Horta, P. F. J. Fuchs, W. F. van Gunsteren, P. H. Hünenberger, *J. Chem. Theory Comput.* **2011**, *7 (4)*, 1016. DOI: 10.1021/ct1006407
[56] J. Gebhardt, N. Hansen, *Fluid Phase Equilib.* **2016**, *422*, 1. DOI: 10.1016/j.fluid.2016.02.001
[57] S. Hempel, J. Fischer, D. Paschek, G. Sadowski, *Soft Mater.* **2012**, *10 (1 – 3)*, 26. DOI: 10.1080/1539445X.2011.599698
[58] C. Červinka, M. Fulem, *J. Chem. Theory Comput.* **2017**, *13 (6)*, 2840. DOI: 10.1021/acs.jctc.7b00164
[59] H.-J. Schneider, *Angew. Chem. Int. Ed.* **2009**, *48 (22)*, 3924. DOI: 10.1002/anie.200802947
[60] F. Biedermann, W. M. Nau, H.-J. Schneider, *Angew. Chem. Int. Ed.* **2014**, *53 (42)*, 11158. *Angew. Chem.* **2014**, *126*, 11338. DOI: 10.1002/anie.201310958
[61] K. N. Houk, A. G. Leach, S. P. Kim, X. Zhang, *Angew. Chem. Int. Ed.* **2003**, *42 (40)*, 4872. DOI: 10.1002/anie.200200565
[62] N. M. Maier, P. Franco, W. Lindner, *J. Chromatogr. A* **2001**, *906 (1 – 2)*, 3. DOI: 10.1016/S0021-9673(00)00532-X
[63] S. Wren, T. Berger, K.-S. Boos, H. Engelhardt, E. Adlard, I. Davies, K. Altria, R. Stock, in *The Separation of Enantiomers by Capillary Electrophoresis* (Eds: T. Berger, K.-S. Boos, H. Engelhardt, E. Adlard, I. Davies, K. Altria, R. Stock), Vieweg+Teubner, Wiesbaden **2001**, 59 – 77. DOI: 10.1007/978-3-322-83141-5_5
[64] I. Ilisz, G. Fodor, R. Berkecz, R. Iványi, L. Szente, A. Péter, *J. Chromatogr. A* **2009**, *1216 (15)*, 3360. DOI: 10.1016/j.chroma.2009.01.083
[65] H. Lorenz, A. Seidel-Morgenstern, *Angew. Chem. Int. Ed.* **2014**, *53 (5)*, 1218. DOI: 10.1002/anie.201302823
[66] J. Wang, H. Yang, J. Yu, X. Chen, F. Jiao, *Chem. Eng. Sci.* **2016**, *143*, 1. DOI: 10.1016/j.ces.2015.12.019
[67] M. V. Rekharsky, Y. Inoue, *J. Am. Chem. Soc.* **2000**, *122 (18)*, 4418. DOI: 10.1021/ja9921118
[68] C. J. Easton, *Pure Appl. Chem.* **2005**, *77 (11)*, 1865. DOI: 10.1351/pac200577111865
[69] S. S. Thakur, H. S. Parekh, C. H. Schwable, Y. Gan, D. Ouyang, in *Computational Pharmaceutics* (Eds: D. Ouyand, S. C. Smith), Wiley, Hoboken **2015**, 31 – 51. DOI: 10.1002/9781118573983.ch3
[70] E. M. M. Del Valle, *Process Biochem.* **2004**, *39 (9)*, 1033. DOI: 10.1016/S0032-9592(03)00258-9
[71] G. Crini, *Chem. Rev.* **2014**, *114 (21)*, 10940. DOI: 10.1021/cr500081p
[72] X. Ma, Y. Zhao, *Chem. Rev.* **2015**, *115 (15)*, 7794. DOI: 10.1021/cr500392w
[73] G. Yu, K. Jie, F. Huang, *Chem. Rev.* **2015**, *115 (15)*, 7240. DOI: 10.1021/cr5005315
[74] B.-W. Liu, H. Zhou, S.-T. Zhou, J.-Y. Yuan, *Eur. Polym. J.* **2015**, *65*, 63. DOI: 10.1016/j.eurpolymj.2015.01.017
[75] K. A. Connors, *Chem. Rev.* **1997**, *97 (5)*, 1325. DOI: 10.1021/cr960371r
[76] M. V. Rekharsky, Y. Inoue, *Chem. Rev.* **1998**, *98 (5)*, 1875. DOI: 10.1021/cr970015o
[77] W. F. van Gunsteren, A. E. Mark, *J. Chem. Phys.* **1998**, *108 (15)*, 6109. DOI: 10.1063/1.476021
[78] L. Pol-Fachin, V. H. Rusu, H. Verli, R. D. Lins, *J. Chem. Theory Comput.* **2012**, *8 (11)*, 4681. DOI: 10.1021/ct300479h
[79] O. Guvench, E. Hatcher, R. M. Venable, R. W. Pastor, A. D. MacKerell, *J. Chem. Theory Comput.* **2009**, *5 (9)*, 2353. DOI: 10.1021/ct900242e
[80] S. Doudou, N. A. Burton, R. H. Henchman, *J. Chem. Theory Comput.* **2009**, *5 (4)*, 909. DOI: 10.1021/ct8002354
[81] G. M. Torrie, J. P. Valleau, *J. Comput. Phys.* **1977**, *23 (2)*, 187. DOI: 10.1016/0021-9991(77)90121-8
[82] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1 – 2*, 19. DOI: 10.1016/j.softx.2015.06.001
[83] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, *Comput. Phys. Commun.* **2014**, *185 (2)*, 604. DOI: 10.1016/j.cpc.2013.09.018
[84] S. R. Billeter, W. F. van Gunsteren, *J. Phys. Chem. A* **2000**, *104 (15)*, 3276. DOI: 10.1021/jp994127q
[85] J. Kästner, W. Thiel, *J. Chem. Phys.* **2005**, *123 (14)*, 144104. DOI: 10.1063/1.2052648
[86] https://github.com/martinstroet/umbrella_integration (accessed on February 8, 2017)
[87] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, J. M. Rosenberg, *J. Comput. Chem.* **1992**, *13 (8)*, 1011. DOI: 10.1002/jcc.540130812
[88] J. S. Hub, B. L. de Groot, D. van der Spoel, *J. Chem. Theory Comput.* **2010**, *6 (12)*, 3713. DOI: 10.1021/ct100494z
[89] M. R. Shirts, J. D. Chodera, *J. Chem. Phys.* **2008**, *129 (12)*, 124105. DOI: 10.1063/1.2978177
[90] https://github.com/choderalab/pymbar (accessed on February 8, 2017)
[91] J. Hermans, L. Wang, *J. Am. Chem. Soc.* **1997**, *119 (11)*, 2707. DOI: 10.1021/ja963568+
[92] I. Z. Steinberg, H. A. Scheraga, *J. Biol. Chem.* **1963**, *238 (1)*, 172.
[93] L. D. Landau, E. M. Lifshitz, *Statistical Physics*, Clarendon, Oxford **1938**.
[94] R. W. Zwanzig, *J. Chem. Phys.* **1954**, *22 (8)*, 1420. DOI: 10.1063/1.1740409
[95] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3 (5)*, 300. DOI: 10.1063/1.1749657
[96] J. Kästner, W. Thiel, *J. Chem. Phys.* **2006**, *124 (23)*, 234106. DOI: 10.1063/1.2206775
[97] D. I. Kopelevich, *J. Chem. Phys.* **2013**, *139 (13)*, 134906. DOI: 10.1063/1.4823500
[98] D. L. Mobley, J. D. Chodera, K. A. Dill, *J. Chem. Phys.* **2006**, *125 (8)*, 084902. DOI: 10.1063/1.2221683
[99] J.-S. Wu, J.-Z. Zheng, K. Toda, I. Sanemasa, *Anal. Sci.* **1999**, *15 (7)*, 701. DOI: 10.2116/analsci.15.701
[100] L. Linden, K.-U. Goss, S. Endo, *J. Colloid Interface Sci.* **2016**, *468*, 42. DOI: 10.1016/j.jcis.2016.01.032
[101] M. S. Miller, W. K. Lay, S. Li, W. C. Hacker, J. An, J. Ren, A. H. Elcock, *J. Chem. Theory Comput.* **2017**, *13 (4)*, 1812. DOI: 10.1021/acs.jctc.6b01059

