

DISS. ETH NO. ?????

# Super Title

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH Zurich)

presented by  
Benjamin Joachim Ries

MSc. in Bioinformatics

born on 15.03.1991  
citizen of Germany

accepted on the recommendation of

Prof. Dr. Sereina Riniker, examiner  
Prof. Dr. Philippe Hünenberger, co-examiner  
Prof. Dr. Niels Hansen, co-examiner

2019



*Here a nice dedication*



# Acknowledgements

*“Der Toaster”*

---

Denker

This thesis would not have been possible without many people, who have supported me over the last years.



# Contents



# Summary

Molecular dynamics (MD) simulations offer nowadays a valuable tool for studying phenomena in chemistry and biology. The history of models in chemistry, as well as the methodological and theoretical background of MD is presented in Chapter 1.

Chapter ?? deals with the application of explicit-solvent molecular dynamics (MD) simulations to resorcin[4]arene cavitands, which can adopt a close/contracted (VASE) and an open/expanded (KITE) conformation. The VASE-KITE equilibria of a quinoxaline- and a dinitrobenzene-based resorcin[4]arene are investigated in three solvent environments (vacuum, chloroform and toluene) and at three temperatures (198.15, 248.15 and 298.15 K). The challenge of sampling the millisecond-timescale VASE-KITE transition is addressed by calculating relative free energies using ball-and-stick local elevation umbrella sampling (B&S-LEUS) to promote interconversion transitions. The calculated VASE-to-KITE free-energy changes  $\Delta G$  are in qualitative agreement with the experimental magnitudes and trends.

Chapters 3-?? present the development of the conveyor belt scheme, a method to calculate free-energy differences. The working principle relies on  $K$  coupled replicas of the system that are simulated at different values of a coupling parameter  $\lambda$ . The number  $K$  is taken to be even and the replicas are equally spaced on a forward-turn-backward-turn path, akin to a conveyor belt (CB) between the two end-states. As in  $\lambda$ -dynamics ( $\lambda$ D), the  $\lambda$ -values associated with the individual systems evolve in time

along the simulation. However, they do so in a concerted fashion, determined by the evolution of a single dynamical variable  $\Lambda$  of period  $2\pi$  controlling the advance of the entire CB. Thus, a change of  $\Lambda$  is always associated with one half of the replicas moving forward and the other half moving backward along  $\lambda$ . As a result, the effective free-energy profile of the replica system along  $\Lambda$  is characterized with decreasing barriers upon increasing  $K$ , at least as  $K^{-1}$  in the limit of large  $K$ . When a sufficient number of replicas is used, these variations become small, which enables a complete and quasi-homogeneous coverage of the  $\lambda$ -range by the replica system.

Chapter 3 introduces this scheme with respect to alchemical free-energy calculations. Therefore it is termed *conveyor belt thermodynamic integration*. It provides the mathematical/physical formulation of the scheme, along with an initial application of the method to the calculation of the hydration free energy of methanol.

In Chapter ??, the conveyor belt thermodynamic integration is applied to a Lys-X-Lys tripeptide, involving a side-chain mutation in the central residue, and guanosine triphosphate, involving a hydrogen-to-bromine mutation. With both systems, sampling issues have been encountered, due to the large orthogonal barriers either along the backbone dihedral angles  $\phi$  and  $\psi$  (tripeptide) or along ribose-base dihedral angle  $\chi$  (guanosine). This relative merits of different sampling schemes, orthogonal biasing and estimators to improve the convergence are investigated. The thermodynamic integration scheme is shown to suffer from the constraint of simulations at fixed  $\lambda$ . There is no significant improvement upon changing from the Simpson's quadrature to the MBAR estimator. Both Hamiltonian replica exchange and conveyor belt thermodynamic integration improve the results for the

tripeptide. For the guanosine triphosphate, improvement is only achieved upon application of an orthogonal biasing potential, most efficiently in combination with Hamiltonian replica exchange or conveyor belt thermodynamic integration.

The conveyor belt scheme is extended to the calculation of conformational free-energy differences in Chapter ??, resulting in a so-called conveyor belt umbrella sampling scheme (CBUS). CBUS is here initially applied to the calculation of 45 standard absolute binding free energies of five alkali cations to three crown ethers in three different solvents. Besides introducing and testing the new scheme, it is compared to other methods. Expectedly, the direct counting approach has convergence issues on the accessible simulation timescale, and the corresponding results are rather unreliable. On the other hand, the results obtained using traditional umbrella sampling and CBUS are very consistent. Additionally, comparison of the results to those of previous alchemical calculations *via* an alchemical pathway reveals excellent consistency while the trends of available experimental data are qualitatively reproduced.

Conclusions and an outlook into future developments are given in Chapter ??.



# Zusammenfassung

Molekulardynamikcomputersimulationen (MD-Simulationen) sind heutzutage ein nützliches Mittel, um Phänomene in der Chemie und Biologie zu studieren. Die Geschichte chemischer Modelle sowie der methodische und theoretische Basis von MD-Simulationen wird im 1. Kapitel präsentiert.

Kapitel ?? beschäftigt sich mit der Anwendung von MD-Simulationen auf Resorcin[4]arene in explizitem Lösungsmittel. Diese Käfigmoleküle können geschlossene/kontrahierte (VASE) und offene/ausgestreckte (KITE) Konformationen annehmen. Die Gleichgewichte zwischen VASE und KITE von Resorcin[4]arenen, die auf Quinoxalin und Dinitrobenzen basieren, werden in verschiedenen Umgebungen (Vakuum, Chloroform und Toluol) bei drei verschiedenen Temperaturen (198.15, 248.15 und 298.15 K) untersucht. Die Herausforderung, die Millisekunden-Zeitskala des VASE-KITE Übergangs in der Simulation zu beobachten, wird durch die Berechnung relativer freier Enthalpien mit der *ball-and-stick local-elevation umbrella-sampling* (B&S-LEUS) Methode gemeistert, die die Übergänge zwischen den Konformationen beschleunigt. Die berechneten freien Enthalpieunterschiede  $\Delta G$  stimmen qualitativ mit den experimentellen Werten und Tendenzen überein.

Kapitel 3-?? präsentieren die Entwicklung des *Förderbandschemas* (engl. conveyor belt scheme), welches eine Methode ist, um freie Enthalpie-Differenzen zu berechnen. Die Arbeitsweise beruht auf dem Aneinanderkoppeln von  $K$  Repliken eines Sys-

tems, die an verschiedenen Werten des Kopplungsparameters  $\lambda$  simuliert werden. Die Anzahl  $K$  sollte gerade gewählt werden und die Repliken werden im gleichen Abstand zueinander auf einem vor-und-zurück Pfad angeordnet, was an ein Förderband zwischen zwei Endpunkten erinnert. Wie in der  $\lambda$ -Dynamikmethode ( $\lambda$ D) ändern sich die  $\lambda$ -Werte der einzelnen Systeme im Laufe der Simulation. Dies geschieht allerdings in einer gemeinschaftlichen Art, welche durch die Änderung einer einzigen dynamischen Förderbandvariablen  $\Lambda$  mit Periode  $2\pi$  bestimmt wird. Daher bewirkt eine Änderung von  $\Lambda$  immer eine Vorwärtsbewegung entlang  $\lambda$  von einer Hälfte der Repliken und eine Rückwärtsbewegung entlang  $\lambda$  von der anderen. Daraus resultiert ein effektives freies Enthalpieprofil des Replikensystems entlang  $\Lambda$ , welches sich durch geringere werdende Energiebarrieren auszeichnet, je grösser  $K$  ist. Wenn eine genügende Anzahl  $K$  an Repliken benutzt wird, werden diese Barrieren so klein, dass eine durchgängige und nahezu homogene Erfassung des gesamten  $\lambda$ -Bereichs durch die Repliken erreicht wird.

Kapitel 3 wendet dieses Schema auf alchemische freie Enthalpieberechnungen an. Deshalb wird es als *Förderband thermodynamische Integration* (engl. conveyor belt thermodynamic integration, CBTI) bezeichnet. Eine mathematische/physikalische Formulierung des Schemas wird vorgelegt bevor eine ersten Anwendung der Methode auf die Berechnung der freien Hydratisierungsenthalpie für Methanol beschrieben wird.

In Kapitel ?? wird die CBTI Methode auf zwei andere Systeme angewandt, nämlich auf ein Lys-X-Lys Tripeptid, wobei die zentrale Aminosäure Glu in Gly mutiert wird, und auf Guanosintriphosphat, wobei ein Wasserstoffatom an der C8 Position in ein Bromatom mutiert wird. Mit beiden Systemen wurden Samplingprobleme festgestellt wegen grossen orthogonalen Barrieren

entlang der Torsionswinkel  $\phi$  und  $\psi$  der Peptidbindung (Tripeptid) oder entlang des Ribosentorsionswinkels  $\chi$  (Guanosine). Die jeweiligen Vorzüge verschiedener Samplingmethoden, orthogonalen Biases oder Schätzer, um die Konvergenz zu verbessern, werden untersucht. Es wird gezeigt, dass die thermodynamische Integration nachteilig ist, weil die Simulationen bei konstantem  $\lambda$  stattfinden. Durch eine Änderung des Schätzers von der Simpson's Quadratur zu MBAR lassen sich die Ergebnisse nicht signifikant verbessern. Sowohl Hamiltonian-Austausch zwischen Repliken als auch CBTI verbessern die Resultate für das Tripeptid. Für Guanosintriphosphat wird eine Verbesserung nur durch die Anwendung eines orthogonalen Biases erreicht, wiederum am Effizientesten für den Hamiltonian-Austausch zwischen Repliken und CBTI.

Das Förderbandschema wird im ???. Kapitel auf konformationelle freie Enthalpiedifferenzen ausgeweitet, was zu dem sogenannten *conveyor belt umbrella sampling* (CBUS) Schema führt. CBUS wird hier als Erstes auf die Berechnung von 45 freie Standard-Bindungsenthalpien zwischen 5 verschiedenen Alkaliionen und drei Kronenethern in drei verschiedenen Lösungsmitteln angewandt. Neben dem Vorstellen und Testen des neuen Schemas wird es auch mit anderen Methoden verglichen. Wie erwartet führt das einfache Zählen der Konformationen (direct counting) zu Konvergenzproblemen in der durchgeföhrten Simulationszeit und die erhaltenen Resultate sind daher unzuverlässig. Andererseits sind die Resultate, die durch die traditionelle Umbrella Sampling und die CBUS Methoden erhalten wurden, sehr übereinstimmend. Zusätzlich wurden exzellente Übereinstimmung mit mit früheren alchemischen Berechnungen über einen alchemischen Pfad erreicht und die Tendenzen der experimentellen Daten wurden qualitativ reproduziert.

Eine Schlussfolgerung und ein Ausblick auf zukünftige Entwicklungen

lungen sind im Kapitel ?? aufgeführt.

# Publications

The following publications are included in parts or in an extended version in this thesis. The other chapters are in preparation for publication.

## CHAPTER 2

D.F. Hahn, J.V. Milić, P.H. Hünenberger *Helv. Chim. Acta* **2019**, *submitted*: Vase-Kite Equilibrium of Resorcin[4]arene Cavitands Investigated Using Molecular Dynamics Simulations with Ball-and-Stick Local Elevation Umbrella Sampling.

## CHAPTER 3

D.F. Hahn, P.H. Hünenberger *J. Chem. Theory Comput.* **2019**, *in press*: Alchemical Free-Energy Calculations by Multiple-replica  $\lambda$ -dynamics: The Conveyor Belt Thermodynamic Integration (CBTI) scheme.



# 1

## Introduction

“ *hmm* ”

---

Author

### 1.1 HISTORY OF MOLECULAR MODELING

Since ancient times, humans have strived to understand their environment and the phenomena they observed. They have conceived models describing the composition of the surrounding matter and attempting to describe its behavior. The starting point was century-long disputes between representatives of atomic and continuum theories. The atomic theory was first proposed<sup>?</sup> in the 5<sup>th</sup> century BC by Presocratics Leucippus and Democritus. This theory argued that matter is not infinitely divisible, but that there ultimately particles which are inalterable and indivisible: the atoms. It was Democritus, who claimed<sup>??</sup>

“by convention sweet and by convention bitter, by convention hot, by convention cold, by convention color; but in reality atoms and void”.

This groundbreaking statement was far ahead of its time and was mostly rejected, especially by Aristotle with his work *De caelo*, stating that matter is continuous and consists of five elements.<sup>7</sup> As the church adhered to the continuity theory of Aristotle, it was predominant in Europe, until the theory of Democritus experienced a revival with the mechanical atomism in the 17<sup>th</sup> century, promoted by philosophers like René Descartes, Pierre Gassendi and Robert Boyle.<sup>7</sup>

While Newton led the foundations of classical mechanics in his work *Philosophiae Naturalis Principia Mathematica*,<sup>7</sup> he also shared atomistic views as he stated that<sup>7</sup>

“the least parts of bodies to be - all extended, and hard and impenetrable, and moveable, and endowed with their proper inertia”.

Before, the atomists were uncertain about the laws governing the movements of atoms. With Newton’s three laws of motion, the dynamics of atoms could in principle be determined. The limitation, however, was that the nature of forces between atoms needed to be known.<sup>7</sup>

The early mechanical atomism was then superseded by the atomic theory of John Dalton in the early 19<sup>th</sup> century.<sup>7</sup> For the first time, properties, such as the relative weight could be assigned to atoms. Chemical elements were already known, but now atoms could be ascribed to smallest unit of what was understood as an element. Likewise, chemical compounds were found to be defined by specific combinations of atoms.

The word “chemical structure” was coined in the mid 19<sup>th</sup> century by chemists like Archibald Scott Couper, Friedrich August Kekulé and Alexander Mikhailovich Butlerov. They proposed the first chemical structures, thereby developing important concepts

like valency, chemical bond and substituent.<sup>7</sup> In 1861, Johann Josef Loschmidt published a collection of 384 molecular structures, some examples of which are given in Figs. 1.1a and 1.1b. His graphical representations included the spatial extent of different atoms in surprisingly correct proportions considering the publication date. Additionally, he proposed the first (correct!) benzene structure, although credit is often incorrectly given to Kekulé. However, the latter invented the resonance formulas of benzene, which are showcased in Fig. 1.1c.<sup>7</sup> <sup>8</sup> <sup>9</sup> <sup>10</sup>

The third spatial dimension came into play with August Wilhelm Hofmann, who created, using table croquet balls, three-dimensional models of methane (Fig. 1.1d), chloroform, and other small organic molecules.<sup>7</sup> Although the structures were not all in agreement with today's state of knowledge, he was the one who devised the color scheme for atoms (*e.g.* black for carbon, white for hydrogen) which is still in use today. In the following years such models were progressively refined, following both experimental and theoretical advances. Structural representations of larger and more complex (bio)molecules were created. Important milestones in this respect were the model of penicillin based on X-ray crystallographic data by Hodgkin et al.,<sup>7</sup> the model of protein  $\alpha$ -helices by Pauling et al.,<sup>7</sup> and the famous DNA model by Crick and Watson.<sup>7</sup>

Finally, in 1958, the first (though still coarse) model of a complete protein was proposed by Kendrew et al.<sup>7</sup> (based on the work by Perutz). It was a model of myoglobin made out of wood, plasticine and paint, and is shown in Fig. 1.1e. A model of the structurally similar haemoglobin was created by Perutz et al. a few years later.<sup>7</sup>

The advent of digital computers in the 1940s opened up unprecedented possibilities for building molecular models. Com-

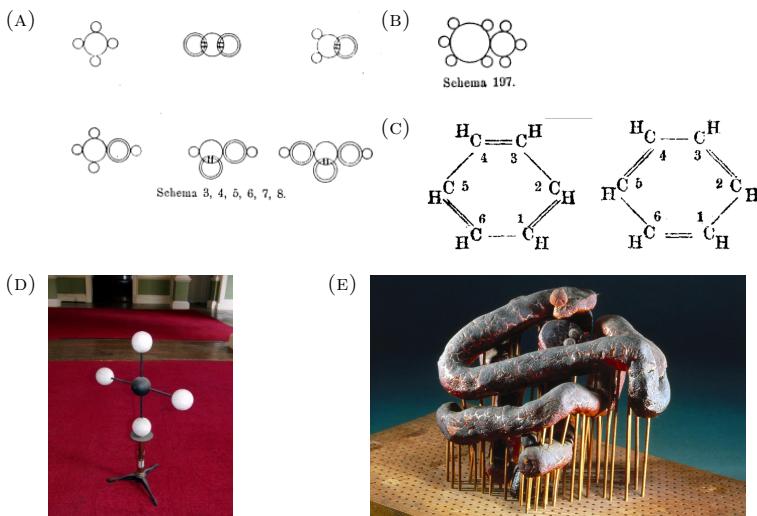


FIGURE 1.1: *Examples of early molecular models.* Graphs (a) and (b) show schemes copied from “Konstitutionsformeln der organischen Chemie in graphischer Darstellung” by J. Loschmidt from 1961, pp. 7 and 62 of Ref.<sup>7</sup> (a): Illustrations of the structures of methane, carbon dioxide, formaldehyde, methanol, formic acid and carbonic acid. (b): Structure of toluene with the first correct prediction of the benzene ring copied from Ref.<sup>7</sup> (c): Drawings of the two resonance formulas of benzene by Kekulé. Copied from (d): Photograph of a methane molecular model, created out of table croquet balls around 1860 by August Wilhelm von Hofmann, which is now part of the collection of the Royal Institution of London. (It is photographed by Henry Rzepa, with the kind permission of the Royal Institution of London, in whose collection the model resides. Wikimedia Commons, the free media repository. Accessed March 10, 2019. [https://commons.wikimedia.org/wiki/File:Molecular\\_Model\\_of\\_Methane\\_Hofmann.jpg](https://commons.wikimedia.org/wiki/File:Molecular_Model_of_Methane_Hofmann.jpg), CC BY-SA 4.0) (e) shows the first model of a complete protein, created and published in 1958 by Kendrew et al.<sup>7</sup> (Science Museum Group. Kendrew’s original model of the myoglobin molecule. 1975-533. Science Museum Group Collection Online. Accessed March 10, 2019. <https://collection.sciencemuseum.org.uk/objects/co13543.1.1>, CC BY-NC-SA)

puters were not only used to resolve the X-ray structures of the proteins mentioned above, but also to solve the equations of the physical models which were otherwise intractable. These included

both quantum-chemical as well as classical-mechanics simulations. After the first Monte Carlo sampling<sup>7</sup> and then molecular dynamics simulation<sup>7</sup> of hard disks in 2 dimensions in the 1950s, the first Lennard-Jones fluid was simulated in 1964.<sup>7</sup> Computer simulations added a new and essential aspect to the molecular models: the alteration of the structure with time, resulting in translations, vibrations, and conformational changes of molecules. The crucial role of the atomic and molecular motions in defining material properties and - ultimately - life was famously phrased by Feynman who stated<sup>7</sup>

“that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms.”

The new field of molecular simulation was born, contributing to the understanding of matter and life in all its details at the atomistic level.

Polyatomic molecules like water<sup>7</sup> and proteins<sup>7</sup> were simulated shortly thereafter. In the following decades, progress was made in developing force fields (interaction functions) and algorithms for reliable, efficient and accurate molecular dynamics (MD) simulations. Since the 1990s, MD has become an indispensable tool in nearly all fields of science, including physics, chemistry, biology, pharmaceutical and material science, and medicine. Due to the ever increasing computational power, larger systems become tractable, which in turn increases the scope of the applications. Apart from visualizing molecules for a better understanding, computations provide a mean for calculating properties, thus complementing and even, sometimes, substituting experiments.

## 1.2 CLASSICAL MECHANICS

The foundations of classical mechanics were laid by Newton,<sup>?</sup> and later generalized by Lagrange and Hamilton. While the laws of classical mechanics are not valid when dealing with high energies (*e.g.* high velocities, light-matter interactions) or low masses (*e.g.* electrons, other elementary particles), where quantum or/and relativistic mechanics is required, they can provide a sufficient accuracy for understanding macroscopic phenomena at a molecular level with comparatively low computational costs.

Classical systems are usually described by  $N$  point particles  $i = 0, 1, \dots, N - 1$  with mass  $m_i$  in Cartesian space, *i.e.* using  $x$ ,  $y$  and  $z$  coordinates. Each particle resides at a certain position  $\mathbf{r}_i = \{x_i, y_i, z_i\}$  with a momentum  $\mathbf{p}_i = m_i \mathbf{v}_i = m_i \{v_{x,i}, v_{y,i}, v_{z,i}\}$  where  $\mathbf{v}$  is the velocity. The vectors  $\mathbf{r}$  and  $\mathbf{p}$  can be united in a position-momentum vector  $\mathbf{x}_i = \{\mathbf{r}_i, \mathbf{p}_i\}$ . The vectors  $\mathbf{x}_i$  of all particles are the elements of the phase-space vector  $\mathbf{x} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1}\}$  which contains the information about the state of a system. The phase space is the associated  $6N$ -dimensional space, which is defined by all possible combinations of positions and momenta of all particles  $i$ .

The associated energy is given by the so-called Hamilonian  $\mathcal{H}$ , which is a function of the phase space vector  $\mathbf{x}$ , and can be separated into a kinetic-energy term  $\mathcal{K}$  and a potential-energy term  $\mathcal{V}$  of the system,

$$\mathcal{H}(\mathbf{x}) = \mathcal{K}(\mathbf{p}) + \mathcal{V}(\mathbf{r}). \quad (1.1)$$

The kinetic energy is given by

$$\mathcal{K}(\mathbf{p}) = \sum_i \frac{\mathbf{p}_i^2}{2m_i}, \quad (1.2)$$

which is the sum over the kinetic energies of all particles. The potential energy term is a mathematical expression for the interaction between the particles, and possibly with the environment. It can be given different functional forms, depending on the nature of the system and the level of approximation, which is typically chosen according to the desired level of accuracy.

The time evolution of the system in Hamiltonian mechanics is given by two equations. The first equation describes how the positions evolve in time,

$$\dot{\mathbf{r}}_i = \frac{d\mathbf{r}_i}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i}. \quad (1.3)$$

By inserting the definition of the kinetic energy, Eq. 1.2, one obtains  $\dot{\mathbf{r}}_i = \partial \mathcal{H} / \partial \mathbf{p}_i = \partial \mathcal{K} / \partial \mathbf{p}_i = \mathbf{p}_i / m_i = \mathbf{v}_i$ , which tells us that the positions change in time according to the velocities of the particles. The second equation describes the time evolution of the momenta,

$$\dot{\mathbf{p}}_i = \frac{d\mathbf{p}_i}{dt} = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}_i}. \quad (1.4)$$

The time derivative of the momentum of a particle  $i$  is equal to its acceleration  $a_i$  multiplied by its mass,  $\dot{\mathbf{p}}_i = m_i \dot{\mathbf{v}}_i = m_i a_i$ . Therefore the second equation is another formulation of Newton's second law  $\mathbf{f}_i = m_i \mathbf{a}_i$ , where  $\mathbf{f}_i$  is the force, given by the negative gradient of the Hamiltonian function with respect to the position  $\mathbf{f}_i = -\partial \mathcal{H} / \partial \mathbf{r}_i$ .

It is important to note that the total energy of a system

governed by Eqs. 1.3 and 1.4 is constant, *i.e.*  $\mathcal{H}$  in Hamiltonian mechanics is a constant of motion.

### 1.3 CLASSICAL STATISTICAL MECHANICS

Statistical mechanics<sup>7 8 9 10 11</sup> is a branch of physics which links the macroscopic properties to the microscopic behavior of a system. The microscopic behavior of the system is governed by the interactions of its particles, *i.e.* by the interaction function  $\mathcal{V}$  in Eq. 1.1. In classical statistical mechanics, the interaction functions are formulated in the framework of classical mechanics and are part of the Hamiltonian of the system. Because of the immense size of macroscopic systems ( $1 \text{ mol} = 6.022 \cdot 10^{23}$  particles) and the non-trivial microscopic interactions between particles in real systems, the treatment even according to classical mechanics is generally not feasible analytically, except for the simplest systems (ideal gas, harmonic crystal). However, by statistically evaluating the microscopic details, we can still calculate the macroscopic properties of a system.

The statistical evaluation is performed on the basis of an ensemble, which is a collection of microscopic states (microstates) of the same system obeying the same microscopic interactions and fulfilling certain macroscopic boundary conditions. A macroscopic property of interest can then be calculated by averaging over the ensemble. According to the ergodic hypothesis,<sup>7</sup> this property can also be calculated by observing the evolution of one system over an infinite period of time. A microstate is defined by a point  $\mathbf{x} = \{\mathbf{r}, \mathbf{p}\}$  in the phase space of the system. Typically, a system

treated by statistical mechanics has many degrees of freedom, *i.e.* a high-dimensional  $\mathbf{x}$ . An ensemble encompasses all microstates which obey some given macroscopic boundary conditions, which are usually three constant thermodynamic observables, among them at least one extensive quantity. Common ensembles are: (*i*) the microcanonical ensemble with constant number  $N$  of particles, volume  $V$  and energy  $E$ ; (*ii*) the canonical ensemble with constant number  $N$  of particles, volume  $V$  and temperature  $T$ ; (*iii*) the isothermal-isobaric (Gibbs) ensemble with constant number  $N$  of particles, pressure  $P$  and temperature  $T$ ; (*iv*) the grand canonical ensemble with constant chemical potential  $\mu$ , volume  $V$  and temperature  $T$ . In the following, we consider only the microcanonical and canonical ensembles, but the discussion can be straightforwardly extended to other systems.

### 1.3.1 THE MICROCANONICAL ENSEMBLE

The microcanonical ensemble is the simplest ensemble, corresponding to an isolated system with a constant number  $N$  of particles in a container of constant volume  $V$  at constant energy  $E$ .

To be part of the ensemble, the microstates defined by the phase space vector  $\mathbf{x}$  have to fulfill the condition

$$\mathcal{H}(\mathbf{x}) = E. \tag{1.5}$$

All accessible microstates  $\mathbf{x}$  of the ensemble lie on a constant-energy hypersurface in phase space and are equally probable. This is known as the assumption of equal a priori probability.

A measure of the amount of phase space available to the system is the partition function. The microcanonical partition function

$\Omega$  is calculated as an integral over phase space

$$\Omega(N, V, E) = \frac{E_0}{N! h^{3N}} \int d\mathbf{x} \delta(\mathcal{H}(\mathbf{x}) - E), \quad (1.6)$$

where the prefactor accounts for particle indistinguishability, unit conversion to a dimensionless quantity, and Heisenberg's uncertainty relation. The partition function is a fundamental quantity in statistical mechanics. All thermodynamic properties of the system can be derived from it. A property related to the number of microstates and, therefore, the partition function is the entropy  $S$ , which is linked to  $\Omega$  via the Boltzmann relation

$$S(N, V, E) = k_B \ln \Omega(N, V, E) , \quad (1.7)$$

where the proportionality factor  $k_B = 8.3146 \text{ J K}^{-1} \text{ mol}^{-1}$  is the Boltzmann constant. For a microcanonical ensemble, the entropy indicates which macrostate is favored. Given two different macrostates, the system will always prefer the state with the higher entropy. A process to this state will always occur spontaneously. However, the entropy difference does not provide any information on the timescale of the corresponding process, which can be short or long due to kinetic hindrance. Other thermodynamic properties can be derived by taking the total differential of the entropy  $S(N, V, E)$ , known from phenomenological thermodynamics,

$$\begin{aligned} dS &= \left( \frac{\partial S}{\partial E} \right)_{N,V} dE + \left( \frac{\partial S}{\partial V} \right)_{N,E} dV + \left( \frac{\partial S}{\partial N} \right)_{V,E} dN \\ &= \frac{1}{T} dE + \frac{P}{T} dV - \frac{\mu}{T} dN , \end{aligned} \quad (1.8)$$

where  $T$  is the temperature,  $P$  the pressure and  $\mu$  the chemical

potential. If one equates the coefficients, *i.e.*

$$\begin{aligned}\frac{1}{T} &= \left( \frac{\partial S}{\partial E} \right)_{N,V}, \\ \frac{P}{T} &= \left( \frac{\partial S}{\partial V} \right)_{N,E}, \\ \frac{\mu}{T} &= \left( \frac{\partial S}{\partial N} \right)_{V,E},\end{aligned}\tag{1.9}$$

one can relate the thermodynamic variables  $T$ ,  $P$ , and  $\mu$  to  $\Omega(N, V, E)$  by combining Eq. 1.7 with Eq. 1.9:

$$\begin{aligned}\beta &= \left( \frac{\partial \ln \Omega}{\partial E} \right)_{N,V}, \\ \beta P &= \left( \frac{\partial \ln \Omega}{\partial V} \right)_{N,E}, \\ \beta \mu &= \left( \frac{\partial \ln \Omega}{\partial N} \right)_{V,E}.\end{aligned}\tag{1.10}$$

The normalized probability  $P(\mathbf{x})$  of finding a system in a given microstate  $\mathbf{x}$  can also be related to the inverse partition function

$$P(\mathbf{x}) = \frac{E_0}{N!h^{3N}} \frac{\delta(\mathcal{H}(\mathbf{x}) - E)}{\Omega(N, V, E)}.\tag{1.11}$$

The normalized probability defines the distribution of macrostates in equilibrium. Accordingly, any thermodynamic observable  $Y(N, V, E)$  can be calculated by multiplying the instantaneous quantity  $\mathcal{Y}(\mathbf{x})$  of a microstate with its probability and integrating over the whole phase space, *i.e.*

$$\begin{aligned}Y(N, V, E) &= \langle \mathcal{Y}(\mathbf{x}) \rangle \\ &= \int d\mathbf{x} P(\mathbf{x}) \mathcal{Y}(\mathbf{x})\end{aligned}\tag{1.12}$$

where  $\langle \dots \rangle$  denotes the ensemble average. Note that Eq. 1.12 is also valid for other ensembles, but the probability, given by Eq. 1.11 for a microcanonical ensemble, has to be adapted.

While being the simplest ensemble, the use of the microcanocial ensemble is limited, because most experiments are conducted under isothermal conditions.

### 1.3.2 THE CANONICAL ENSEMBLE

The canonical ensemble is an ensemble of systems with a constant number  $N$  particles in a container of constant volume  $V$  at temperature  $T$ , *i.e.* the systems are in contact with a heat bath allowing the exchange of energy (heat) in order to keep the temperature of the system constant.

The main difference to the microcanonical ensemble resides in the heat exchange with the environment. The total energy is not conserved and the microstates now obey a Boltzmann distribution, with probabilities proportional to  $\exp(-\beta\mathcal{H}(\mathbf{x}))$ , where  $\beta = (k_B T)^{-1}$  and  $k_B$  is the Boltzmann's constant,. The canonical partition function is an integral over the Boltzmann factor for all possible microstates

$$Q(N, V, T) = \frac{1}{h^{3N} N!} \int d\mathbf{x} \exp(-\beta\mathcal{H}(\mathbf{x})). \quad (1.13)$$

The energy of the ensemble  $E$  is the ensemble average, given by

Eq. 1.12, of the Hamiltonian.

$$\begin{aligned} E(N, V, T) &= \langle \mathcal{H}(\mathbf{x}) \rangle. \\ &= \int d\mathbf{x} P(\mathbf{x}) \mathcal{H}(\mathbf{x}) \\ &= -\frac{1}{Q(N, V, T)} \frac{\partial}{\partial \beta} Q(N, V, T) \\ &= -\frac{\partial}{\partial \beta} \ln Q(N, V, T). \end{aligned} \tag{1.14}$$

For a canonical ensemble, the entropy difference between two states does not tell anything about the spontaneity of a process, because the corresponding entropy change in the environment (resulting from the transfer of heat) has to be taken into account. Therefore it is convenient to define the Helmholtz free energy  $A$  as

$$A(N, V, T) = -\beta^{-1} \ln Q(N, V, T). \tag{1.15}$$

This quantity gives information about the spontaneity of a process. If the free-energy difference is negative, the process will occur. If it is zero, the two states are in equilibrium.

## 1.4 MOLECULAR DYNAMICS SIMULATION

Molecular dynamics (MD) simulations apply classical mechanics (Sect. 1.2) to molecular systems. Common applications include organic or biological molecules in solvent. In the following, the framework of MD is briefly introduced.

### 1.4.1 THE SYSTEM

A typical simulated system consists of one or more solute(s), *e.g.* an organic molecule, a protein, a lipid, a saccharide, a nucleic acid or a combination of those. The environment of this solute is either vacuum, implicit solvent (which models the average influence of solvent molecules *via* a continuous medium) or explicit solvent (solvent molecules are added to the system). The number of atoms can be on the order of several millions of atoms, which leads to system sizes on the order of  $10 \times 10 \times 10 \text{ nm}^3$ . The times simulated are on the order of nanoseconds to milliseconds per day computer time, depending on the available computing resources, on the system size, on the employed approximations and on the algorithms used.

The atoms to be simulated are described as  $N$  point particles having masses  $\{m_i\}$  and charges  $\{q_i\}$ . Often, a number of atoms are combined to one point particle, which is referred to as an united atom<sup>?</sup> (*e.g.* a methyl group) or a coarse grained bead<sup>??</sup> (*e.g.*, a whole ester functional group). The electrons and other elementary particles are not included explicitly. Only their average influence is modeled by means of the masses and charges. Therefore, MD simulations cannot describe electronically excited states, chemical reactions or light-matter interactions. In the case of explicit solvation, periodic boundary conditions are usually applied to avoid surface effects. The system is then modeled in a space-filling polyhedron, *e.g.*, a cube or a truncated octahedron, and an atom leaving the boundary on one side enters again on the opposite side. In other words, the system is allowed to interact with periodic copies of itself.

### 1.4.2 THE INTERACTION FUNCTION (FORCE FIELD)

The potential energy  $\mathcal{V}$  (Sect. 1.2) is a mathematical expression to describe the nature of the interactions within the system. In MD, this function is referred to as a force field. A force field consists of many terms, which can be classified as covalent and non-bonded terms.<sup>7</sup> Usually non-bonded terms are two-body terms, *i.e.* they include only the coordinates of two point particles. Some  $n$ -body terms with  $n > 2$  can be considered in exceptional cases.

The covalent terms describe the interactions within one molecule up to third-neighbor interactions, *i.e.* between atoms separated by up to three chemical bonds. They usually include: (i) bond stretching terms, which are harmonic potentials controlling the distance between two connected atoms; (ii) bond-angle bending terms, which are harmonic potentials controlling the angle spanned by three atoms; (iii) torsional-dihedral terms, which are usually cosine series controlling the configuration of four consecutively bonded atoms; (iv) improper-dihedral terms, which are harmonic potentials controlling the configuration of four atoms, of which three are bonded to a central atom. Other terms (*e.g.* covalent cross-terms) are possible, but seldomly used.

Interactions between atoms separated by more than two bonds (including pairs of atoms in different molecules) are described by non-bonded interaction terms. These terms depend on the distance between two atoms. There are typically two different terms considered which describe the electrostatic and van-der-Waals interactions between the atoms. The electrostatic interactions are defined by a Coulomb potential having a  $1/r$  dependence, where  $r$  is the distance of the atoms. These terms model the attraction of two atoms with partial charges of opposite signs, or the repulsion of atoms with partial charges of the same sign.

The van-der-Waals interactions are commonly modeled with a Lennard-Jones functional form, including the repulsion due to the Pauli exclusion principle for the electrons at short distances ( $r^{-12}$  dependence), and the attraction due to London dispersion forces ( $r^{-6}$  dependence) for longer distances. Because non-bonded interactions, especially the electrostatic interactions, decay very slowly, special care has to be taken for long-range non-bonded interactions. Otherwise one would need to simulate very large systems which is computationally expensive. The most prominent methods in this regard are the reaction field<sup>7</sup> (RF) and the particle-mesh Ewald<sup>7</sup> (PME) methods. In the RF method, a cutoff is applied which sets the interaction to zero at distances larger than the cutoff distance. To avoid the neglect of longer-range interactions, a reaction-field term is added, accounting for the effect of a dielectric environment beyond the cutoff sphere. Since the covalent terms take care of interactions between atoms separated by up to three bonds, the nonbonded interactions for first and second neighbors are usually excluded. Third-neighbor covalent interactions are usually scaled, or handled using special (reduced) non-bonded interaction parameters.

All force-field terms extensively make use of parameters like force constants, reference bond lengths, angles, torsions and improper dihedrals, partial charges and Lennard-Jones parameters. The number of parameters is on the order of number of degrees of freedom of the molecules considered. For example, a water ( $\text{H}_2\text{O}$ ) molecule is in principle described by 10 parameters, including 4 Lennard-Jones parameters (2 for the oxygen, 2 for the hydrogen, the latter sometimes set to zero), 2 partial charges, and two reference values and two force constants for the bond lengths and angle, respectively. These parameters have to be fitted in order to reproduce experimental properties in an MD simulation. Initial

guesses of parameters like force constants (energetic information) are taken from experimental spectroscopic measurements. For reference structural information (bond lengths, angles, torsions and improper dihedrals), data from X-ray crystallography can be used. All experimentally derivable parameters can be refined or even completely replaced by quantum-mechanical (QM) estimates. However, the initial parameters taken from experiments or calculations have to be refined to make them compatible to each other and to the employed algorithms. Therefore, force-field parameterization is a time consuming and difficult task, which is still the subject of ongoing research.

### 1.4.3 INTEGRATION OF THE EQUATIONS OF MOTION

Given a force field for the interactions of the particles and after defining the Hamiltonian  $\mathcal{H}$  (Eq. 1.1) of the system, the classical equations of motion in Eqs. 1.3 and 1.4 can be integrated. In view of the large size of the system ( $6N$  dimensions), the analytical integration of the equations of motion is impossible and has to be performed numerically instead. One employs a finite time step  $\Delta t$ , which has to be short enough to avoid quadrature errors and long enough to provide computational efficiency and to avoid round-off errors. It is typically chosen on the order of  $1/10$  of the fastest timescale of the system, determined by the bond vibrations, on a femtosecond timescale. Different algorithms aiming at reducing the integration error have been developed. One example is the leap-frog algorithm<sup>7</sup> which evolves the positions  $\mathbf{r}_i$  and velocities

$\mathbf{v}_i$  shifted by a half timestep

$$\mathbf{v}_i(t + \frac{\Delta t}{2}) = \mathbf{v}_i(t - \frac{\Delta t}{2}) + \frac{\mathbf{f}_i(\mathbf{r}(t))}{m_i} \Delta t \quad (1.16)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \frac{\Delta t}{2}) \Delta t. \quad (1.17)$$

This leads to a vanishing  $(\Delta t)^2$  term and, therefore, to a reduced error on the order of  $(\Delta t)^3$ .

#### 1.4.4 THERMOSTATTING AND BAROSTATTING

The integration of the equations of motion leads to a microcanonical ensemble (Sect. 1.2). To reproduce experimental conditions, it is necessary to simulate at constant temperature and, possibly, pressure. The temperature and pressure is regulated by employing thermostats and barostats.

The instantaneous temperature, *i.e.* the temperature of a specific state  $\mathbf{x}$  of a system, depends solely on the kinetic energy as defined by the momenta of the particles,

$$\begin{aligned} \mathcal{T} &= \frac{2\mathcal{K}}{\mathcal{N}_{\text{dof}}k_{\text{B}}}, \\ &= \sum_i \frac{\mathbf{p}_i^2}{m_i \mathcal{N}_{\text{dof}} k_{\text{B}}} \end{aligned} \quad (1.18)$$

where  $N_{\text{dof}} = 3N - 6$  is the number of degrees of freedom of the system with  $N$  particles. In principle one could just constrain the temperature to the desired target temperature by scaling the momenta of the particles at every integration step, but this leads to a constrained temperature, which is not correct for a canonical or Gibbs ensemble. Instead, the temperature distribution should be a Boltzmann distribution, where the instantaneous temperature is allowed to fluctuate. This temperature fluctuation is made

possible in the weak-coupling thermostat by Berendsen,<sup>7</sup> which rescales the velocities with the factor

$$\gamma = \left[ 1 + \frac{\Delta t}{\tau_T} \left( \frac{T}{\mathcal{T}} - 1 \right) \right]^{1/2}, \quad (1.19)$$

by employing a coupling time parameter  $\tau_T$ . The temperature distribution is still not a Boltzmann distribution, which requires a more advanced velocity-scaling thermostat like the Nosé-Hoover<sup>8 9 10</sup> or Nosé-Hoover chain<sup>11</sup> thermostats. Another means of temperature control relies on stochastic approaches like the Andersen thermostat,<sup>12</sup> where the velocities of the particles are reset to a Maxwell-Boltzmann distribution at constant time periods. Other equations of motion such as the Langevin equation of motion are also able to generate constant-temperature ensembles. Here, the forces are attenuated by a friction force while additional stochastic forces (“random kicks”) are applied to the particles. Pure Monte Carlo approaches with the Metropolis-Hastings acceptance criterion<sup>13</sup> also automatically generate an *NVT* ensemble.

The instantaneous pressure  $\mathcal{P}$ , on the other hand, depends on both the momenta and positions,

$$\mathcal{P} = \frac{1}{3V} \sum_i \frac{\mathbf{p}_i^2}{m_i} + \mathbf{r}_i \cdot \mathbf{f}_i, \quad (1.20)$$

where the product of position and force,  $\mathbf{r}_i \cdot \mathbf{f}_i$ , is the virial, which accounts for the deviation from ideal-gas behavior. To simulate at the correct average pressure, the positions of the particles and including the box size are usually rescaled, *e.g.*, by the Berendsen weak-coupling barostat.<sup>7</sup>

### 1.4.5 FREE-ENERGY CALCULATIONS

Many thermodynamic properties like the temperature, pressure, or the density can be extracted from a simulation by calculating the average over the corresponding instantaneous variable. For the calculation of free energies, however, it is not a viable alternative.<sup>??</sup> Although we can express Eq. 1.13 for the free energy as an ensemble average by noting that  $Q(N, V, T) = \langle \exp(\beta\mathcal{H}(\mathbf{x})) \rangle$ ,

$$A(N, V, T) = -\beta^{-1} \ln \langle \exp(\beta\mathcal{H}(\mathbf{x})) \rangle \quad (1.21)$$

it is not possible to calculate free energies in this way. The ensemble average will not converge, because configurations  $\mathbf{x}$  with high energies  $\mathcal{H}$  are seldom visited in a simulation (low Boltzmann weight), but have the strongest impact on the ensemble average. Thus, it would be impossible to obtain a converged result. But we can take advantage of the fact that one is usually interested in free-energy differences between two states A and B instead of the absolute free energy, *i.e.*

$$\Delta A(N, V, T) = A_B - A_A \quad (1.22)$$

$$= -\beta^{-1} \ln \frac{Q_B}{Q_A} \quad (1.23)$$

where the subscript A or B denote a property pertaining to state A or B, respectively.

One distinguishes between thermodynamic, conformational and alchemical free-energy differences, depending on the end states considered. Thermodynamic free-energy differences consider end states which differ in a thermodynamic property like the temperature, pressure or number of molecules. Conformational free-energy differences consider end states which are different parts of phase

space. In other words, the free-energy change upon changing conformations is observed along one (or more) degree(s) of freedom of the system. Alchemical free-energy differences consider end states which differ in the Hamiltonian. An alchemical coordinate is added to the system, along which the Hamiltonian is changed, while the degrees of freedom (number of atoms) remain unchanged. Whereas the former free-energy differences are physical, the alchemical one is unphysical, and has no experimental counterpart. But in MD simulations, it often offers a sampling advantage to conduct alchemical calculations and compare the results of such simulations with experimental data *via* a thermodynamic cycle. An example is the free-energy of binding of a drug molecule to a receptor, which is depicted in Fig. 1.2. One can alchemically mutate the drug molecule to a dummy skeleton (non-interacting particles), once bound to the receptor and once unbound. The comparison of these two differences gives access to the binding free-energy at considerably lower computational cost than directly calculating the conformational free-energy difference.

FIGURE 1.2: *Schematic illustration of a thermodynamic cycle considering the binding of a drug molecule A (magenta) to a receptor (brown).* The upper horizontal processe defines the conformational change from the unbound receptor and drug molecule to the bound complex, which means the drug molecule is displaced from the bulk water into the binding pocket. The corresponding free-energy difference is  $\Delta A_A^{\text{bound-unbound}}$ . The lower horizontal processe defines the same conformational change for the dummy state D (non-interacting skeleton). The corresponding free-energy difference is  $\Delta A_D^{\text{bound-unbound}}$  is zero, because the dummy D is not interacting. The vertical processes correspond to the alchemical transformation of molecule A to the dummy state D, once in the unbound state and once in the bound state. The total free-energy change around the cycle is zero. Therefore the free-energy difference of interest ( $\Delta A_A^{\text{bound-unbound}}$ ) can be inferred from the two alchemical calculations, which can be performed at considerably lower computational cost than the direct conformational calculation.

For conformational free-energy calculations, the Hamiltonian

remains unchanged. Therefore, the ratio between the partition functions can be expressed as ratio of probabilities  $P_A$  and  $P_B$  of being in state A or B,

$$\begin{aligned}\Delta A(N, V, T) &= -\beta^{-1} \ln \frac{\int_B d\mathbf{x} \exp(-\beta\mathcal{H}(\mathbf{x}))}{\int_A d\mathbf{x} \exp(-\beta\mathcal{H}(\mathbf{x}))} \\ &= -\beta^{-1} \ln \frac{P_B}{P_A}.\end{aligned}\quad (1.24)$$

Note that the integral is not running over the whole phase space anymore, but is restricted to the different regions A and B of the phase space. The probabilities of being in state A and state B are easily calculated from a simulation and can then be used to determine the free-energy difference. This approach is frequently used and termed direct counting. However, it requires sufficient sampling and interconversion between the two states A and B, which is only possible if the free-energy difference is small enough and the states are not separated by a too high barrier.

In alchemical calculations, when  $\mathcal{H}_A$  is not equal to  $\mathcal{H}_B$ , one applies the following transformation

$$\begin{aligned}\Delta A(N, V, T) &= -\beta^{-1} \ln \frac{\int d\mathbf{x} \exp(-\beta\mathcal{H}_B(\mathbf{x}))}{\int d\mathbf{x} \exp(-\beta\mathcal{H}_A(\mathbf{x}))} \\ &= -\beta^{-1} \ln \frac{\int d\mathbf{x} \exp(-\beta(\mathcal{H}_B(\mathbf{x}) - \mathcal{H}_A(\mathbf{x}))) \exp(-\beta\mathcal{H}_A(\mathbf{x}))}{\int d\mathbf{x} \exp(-\beta\mathcal{H}_A(\mathbf{x}))} \\ &= -\beta^{-1} \ln \int d\mathbf{x} P_A \exp(-\beta(\mathcal{H}_B(\mathbf{x}) - \mathcal{H}_A(\mathbf{x}))) \\ &= -\beta^{-1} \ln \langle \exp[-\beta(\mathcal{H}_B(\mathbf{x}) - \mathcal{H}_A(\mathbf{x}))] \rangle_A.\end{aligned}\quad (1.25)$$

This approach is called free-energy perturbation and Eq. 1.25 is

the Zwanzig formula.<sup>7</sup> In practice one simulates state A (with Hamiltonian  $\mathcal{H}_A$ ) and one infers the energy of state B from the configurations sampled at state A. This approach will only yield good results if there is enough phase-space overlap between states A and B, *i.e.* if the configurations sampled in A are also low-energy conformations for B.

If the phase-space overlap is not sufficient, a hybrid Hamiltonian can be constructed with a coupling parameter  $\lambda$  that defines a continuous transformation between the Hamiltonians of the physical end-states *A* and *B*. The hybrid Hamiltonian  $\mathcal{H}(\mathbf{x}; \lambda)$  must satisfy the boundary conditions  $\mathcal{H}(\mathbf{x}; 0) = \mathcal{H}_A(\mathbf{x})$  and  $\mathcal{H}(\mathbf{x}; 1) = \mathcal{H}_B(\mathbf{x})$ . This enables splitting up the alchemical transformation into several parts by simulating at intermediate states along the coupling parameter and calculating the free-energy differences for only a part of the whole transformation. The phase-space overlap between two neighboring states is increased, which leads to better convergence. Finally the total free-energy difference can be obtained as a sum over the parts.

A different and frequently used approach is the thermodynamic integration method.<sup>7, 8, 9</sup> Here one integrates the mean force along the alchemical coupling parameter,

$$\Delta A(N, V, T) = \int_0^1 \frac{\partial A(\lambda)}{\partial \lambda} d\lambda . \quad (1.26)$$

The derivative of the free energy with respect to the coupling

parameter is given by

$$\begin{aligned}
 \frac{\partial A(\lambda)}{\partial \lambda} &= \frac{\partial}{\partial \lambda} \left( -\beta^{-1} \ln \frac{1}{h^{3N} N!} \int \exp(\beta \mathcal{H}(\mathbf{x}; \lambda)) d\mathbf{x} \right) \\
 &= -\beta^{-1} \frac{\frac{\partial}{\partial \lambda} \int \exp(\beta \mathcal{H}(\mathbf{x}; \lambda)) d\mathbf{x}}{\int \exp(\beta \mathcal{H}(\mathbf{x}; \lambda)) d\mathbf{x}} \\
 &= -\beta^{-1} \frac{\int \frac{\partial \mathcal{H}(\mathbf{x}; \lambda)}{\partial \lambda} \exp(\beta \mathcal{H}(\mathbf{x}; \lambda)) d\mathbf{x}}{\int \exp(\beta \mathcal{H}(\mathbf{x}; \lambda)) d\mathbf{x}} \\
 &= \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_{\lambda}, \tag{1.27}
 \end{aligned}$$

*i.e.* it is exactly the ensemble average of the Hamiltonian derivative with respect to  $\lambda$ . The free-energy difference is thus calculated by simulating at different  $\lambda$ -values, calculating the ensemble average of the Hamiltonian derivative at the respective  $\lambda$  values and finally integrating numerically to obtain the final free-energy difference.

## 1.5 AIM OF THIS THESIS

This thesis deals with methodological developments of free-energy calculations and their application in MD simulations. In Chapter ??, a force-field for resorcin[4]arenes is presented, which is employed for the calculation of the free-energy difference between two distinct conformations, a closed VASE conformation with a cavity and an open KITE conformation with an expanded surface. To efficiently calculate the free-energy difference, a method called ball-and-stick local elevation umbrella sampling (B&S-LEUS) is used. Chapters 3-?? deal with the development and application of the so-called conveyor belt scheme. This scheme employs multiple coupled replicas which concertedly move on a forward-

turn-backward path, akin a conveyor belt, along a coordinate of interest. In Chapters 3 and ??, the coordinate of interest is the alchemical coupling parameter  $\lambda$ , and the scheme is therefore termed conveyor belt thermodynamic integration (CBTI). CBTI is introduced and applied to the aqueous annihilation of methanol. In Chapter ??, the performance is compared and tested on two other systems, namely to the alchemical mutations of parts of a tripeptide and a guanosine triphosphate. Finally, in Chapter ??, the conveyor belt scheme is extended to conformational changes, now termed conveyor belt umbrella sampling (CBUS), and applied to the calculation of binding free energies between ions and crown ethers in various solvents.



# Alchemical Free-Energy Calculations by Multiple-replica $\lambda$ -dynamics: The Conveyor Belt Thermodynamic Integration

# 2

*“Let us learn to dream, gentlemen, and then perhaps we shall learn the truth.”*

---

August Kekulé, 1865

A new method is proposed to calculate alchemical free-energy differences based on molecular dynamics (MD) simulations, called the conveyor belt thermodynamic integration (CBTI) scheme. As in thermodynamic integration (TI),  $K$  replicas of the system are simulated at different values of the alchemical coupling parameter  $\lambda$ . The number  $K$  is taken to be even and the replicas are equally spaced on a forward-turn-backward-turn path, akin to a conveyor belt (CB) between the two physical end-states. And as in  $\lambda$ -dynamics ( $\lambda$ D), the  $\lambda$ -values associated with the individual systems evolve

in time along the simulation. However, they do so in a concerted fashion, determined by the evolution of a single dynamical variable  $\Lambda$  of period  $2\pi$  controlling the advance of the entire CB. Thus, a change of  $\Lambda$  is always associated with  $K/2$  equispaced replicas moving forward and  $K/2$  equispaced replicas moving backward along  $\lambda$ . As a result, the effective free-energy profile of the replica system along  $\Lambda$  is periodic of period  $2\pi K^{-1}$  and the magnitude of its variations decreases rapidly upon increasing  $K$ , at least as  $K^{-1}$  in the limit of large  $K$ . When a sufficient number of replicas is used, these variations become small, which enables a complete and quasi-homogeneous coverage of the  $\lambda$ -range by the replica system, without application of any biasing potential. If desired, a memory-based biasing potential can still be added to further homogenize the sampling, the preoptimization of which is computationally inexpensive. The final free-energy profile along  $\lambda$  is calculated similarly to TI, by binning of the Hamiltonian  $\lambda$ -derivative as a function of  $\lambda$  considering all replicas jointly, followed by quadrature integration. The associated quadrature error can be kept very low owing to the continuous and quasi-homogeneous  $\lambda$ -sampling. The CBTI scheme can be viewed as a continuous/deterministic/dynamical analog of the Hamiltonian replica-exchange/permutation (HRE/HRP) schemes, or as a correlated multiple-replica analog of the  $\lambda$ D or  $\lambda$ -local elevation umbrella sampling ( $\lambda$ -LEUS) schemes. Compared to TI, it shares the advantage of the latter schemes in terms of enhanced orthogonal sampling, *i.e.* the avail-

ability of variable- $\lambda$  paths to circumvent conformational barriers present at specific  $\lambda$ -values. Compared to HRE/HRP, it permits a deterministic and continuous sampling of the  $\lambda$ -range, and bypasses the need to carefully preselect a  $\lambda$ -ladder and a swapping-attempt frequency. Compared to  $\lambda$ -LEUS, it eliminates (or drastically reduces) the dead time associated with the preoptimization of a biasing potential. The goal of this chapter is to provide the mathematical/physical formulation of the proposed CBTI scheme, along with an initial application of the method to the calculation of the hydration free energy of methanol.

## 2.1 INTRODUCTION

Classical molecular dynamics (MD) simulations provide insight into (bio-)molecular systems at atomic resolution, thereby explaining and complementing experimental observations. This often involves the calculation of free-energy differences,<sup>?</sup> <sup>?</sup> <sup>?</sup> which characterize the relative stabilities of two or more macroscopic states of the system. These states may differ thermodynamically (different pressures, temperatures or numbers of molecules), conformationally (different regions in a space spanned by a set of specific generalized coordinates) or alchemically (different Hamiltonian functions).

Alchemical free-energy calculations involve atom mutations or interaction alterations that have no experimental counterpart. However, by comparing the results of two such calculations in different environments (*e.g.* mutation of a molecule into another one in vacuum or in a solvent) *via* a thermodynamic cycle,<sup>?</sup> <sup>?</sup> the calculated values can be converted to experimentally accessible differences (*e.g.* relative solvation free energies of the two molecules in the given solvent). This indirect pathway *via* a cycle typically offers a strong sampling advantage relative to the calculation over a direct conformational path (*e.g.* reversibly displacing the two molecules from vacuum into the solvent across the liquid surface), while giving the same result at full convergence.<sup>?</sup> <sup>?</sup> <sup>?</sup> Over the last decades, numerous methods have been proposed to calculate alchemical free-energy differences between states  $A$ ,  $B$ ,  $C$ , ... of a molecular system involving the same numbers of atoms but different Hamiltonian functions. They can be roughly classified as reference-state methods and pathway-dependent methods.<sup>?</sup>

In reference-state methods, a single simulation is performed at

a reference state  $R$  and the relative free energy of a target state  $A$  is calculated using one-step free-energy perturbation<sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> (OSP) as a free-energy estimator. The state  $R$  may be unphysical, in which case the free-energy difference between two physical states  $A$  and  $B$  is obtained by comparing the results of two such calculations ( $R$  to  $A$  and  $R$  to  $B$ ). The accuracy of the method, *i.e.* its convergence at finite sampling times, depends crucially on the extent of Boltzmann-weighted phase-space overlap between the reference and target states,<sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> *i.e.* whether configurations relevant for  $A$  are well sampled in  $R$ . Methods to enhance this overlap include in particular: (i) the use of a reference state with soft-core<sup>?</sup> (SC) sites<sup>?</sup> <sup>?</sup> <sup>?</sup> or other softened force-field terms;<sup>?</sup> <sup>?</sup> <sup>?</sup> (ii) the construction of a reference state encompassing all the target states, as in enveloping distribution sampling<sup>?</sup> <sup>?</sup> <sup>?</sup> (EDS); (iii) the extension to the use of multiple reference states,<sup>?</sup> along with the application of the (multi-state<sup>?</sup>) Bennet acceptance ratio<sup>?</sup> <sup>?</sup> (BAR or MBAR) estimator instead of OSP. In principle, reference-state methods bear the promise of enabling the extrapolative calculation of the relative free energies of numerous arbitrary states ( $A$ ,  $B$ ,  $C$ , ...) based on a single reference-state simulation. In practice, however, they seldom hold this promise (or not in a sufficiently robust fashion based on finite simulations), because the design of a suitable reference state requires considerable (*e.g.* SC approach) or even complete (*e.g.* EDS approach) *a priori* knowledge of the target states it has to be appropriate for.

In pathway-dependent methods, a hybrid Hamiltonian is constructed by employing a coupling parameter  $\lambda$  that defines a continuous transformation between the Hamiltonians of the physical end-states  $A$  and  $B$ . Given such a path, the most established, robust and still frequently used method is multi-configuration<sup>?</sup>

thermodynamic integration<sup>7 8 9</sup> (MCTI or, simply, TI). In TI, a set of independent simulations are performed at different constant  $\lambda$ -values, and the ensemble average of the derivative of the hybrid Hamiltonian with respect to  $\lambda$  is subsequently integrated by numerical quadrature<sup>7 8 9</sup> or curve fitting.<sup>7 8 9 10</sup> A common alternative to TI is multi-configuration<sup>7</sup> free-energy perturbation<sup>7 11 12 13 14 15 16</sup> (MC-FEP or, simply, FEP), where the OSP estimator is used to evaluate the free-energy difference from one  $\lambda$ -point to the previous or/and next one. Refinements and improvements of the original TI protocol include in particular: (i) the design of Hamiltonian coupling schemes leading to high sampling efficiencies;<sup>7 8 9 17 18</sup> (ii) the partial automation of the TI protocol,<sup>7 19 20 21 22</sup> *i.e.* of the selection of  $\lambda$ -points along with associated initial configurations, equilibration times and sampling times; (iii) the use of free-energy estimators with improved statistical efficiencies over plain quadrature, *e.g.* extended TI (EXTI) estimator<sup>23</sup> or MBAR estimator;<sup>24</sup> (iv) the design of alternative single- or multiple-replica schemes where the  $\lambda$ -values are no longer fixed during the simulation.<sup>7 8 9 17 18</sup>

Although the possibility of changing the  $\lambda$ -values over the course of a simulation may appear at first sight to represent an unnecessary complication of the TI protocol, it often leads to a significant enhancement of the sampling efficiency. This is because it increases the likelihood of crossing barriers in the orthogonal space,<sup>7 25</sup> that is, the space spanned by all the degrees of freedom of the system excluding  $\lambda$  (*i.e.* all the conformational ones). These orthogonal barriers are typically higher at certain  $\lambda$ -values than at others. The highest ones may be seldom crossed in corresponding simulations at fixed  $\lambda$ , and allowing for variations of  $\lambda$  may open pathways to circumvent them. Note that, in many cases, this enhancement is not sufficient *per se* and other techniques must

be applied to further improve the orthogonal sampling along the  $\lambda$ -path.<sup>???</sup> There exist two main routes for performing variable- $\lambda$  simulations: (*i*) Hamiltonian replica exchange (HRE) or permutation (HRP), which involve multiple system replicas; (*ii*)  $\lambda$ -dynamics ( $\lambda$ D) or  $\lambda$ -Monte Carlo ( $\lambda$ MC), which consider a single system.

In the HRE scheme,<sup>???</sup> a series of system copies (walkers) are distributed over a set of fixed  $\lambda$ -values (replicas), as is the case in TI. However, at regular time intervals, swaps are attempted between pairs of systems corresponding to different  $\lambda$ -values (typically adjacent ones). These attempts are accepted or rejected according to a Metropolis-Hastings criterion<sup>??</sup> depending on the Boltzmann factor ratio of the replica system before and after the swap. Although the trajectories at each  $\lambda$ -point become discontinuous, the TI-like statistics is preserved and the data can be analyzed in the same way as for TI. Recent extensions of the method involve in particular the consideration of more advanced exchange schemes,<sup>???</sup> of replica reservoirs,<sup>???</sup> of frozen replicas,<sup>??</sup> of heating-quenching steps between the sampling periods,<sup>???</sup> of the infinite-swapping limit,<sup>???</sup> and of generalized-ensemble distributions.<sup>???</sup> They also include the implementation of  $\lambda$ -moves that go beyond pairwise swaps with a selection based on a Suwa-Todo criterion,<sup>???</sup> as implemented in the HRP method.<sup>???</sup> In the latter case, enabling arbitrary permutations and abandoning the detailed-balance constraint leads to a significant increase in the probability of exchange acceptance.

In the  $\lambda$ D scheme<sup>???</sup> (see also its ancestor  $\lambda$ MC), a single system is considered for which the  $\lambda$ -value evolves dynamically in the course of the simulation, *i.e.*  $\lambda$  is treated as an extra pseudo-conformational degree of freedom with an assigned

mass parameter  $m_\lambda$  and momentum  $p_\lambda$ . This momentum enters an extended Hamiltonian that includes an additional term for the corresponding kinetic energy. This results in a continuous sampling of the  $\lambda$ -range instead of the discrete sampling underlying TI or HRE/HRP. The free-energy difference is then typically estimated from the probability distribution along  $\lambda$ , with the drawback of requiring a threshold to define the end-states.<sup>?</sup> This issue can be alleviated by introducing a coordinate transformation with plateaus,<sup>??</sup> or by using a TI-like formula<sup>???</sup> or a Rao-Blackwell estimator,<sup>?</sup> the latter conceptually similar to MBAR.

The major advantages of  $\lambda$ D compared to HRE/HRP are that it involves a simpler single-system setup, is deterministic, samples the  $\lambda$ -range continuously, and does not require the specification (and optimization<sup>?????????????????</sup>) of a  $\lambda$ -ladder and of a swapping-attempt frequency. The main drawback<sup>?</sup> is that the sampling probability along  $\lambda$  is no longer imposed in the form of a fixed set of  $\lambda$ -points with equal sampling times, but entirely controlled by the free-energy profile along  $\lambda$ . As a result,  $\lambda$ -values with high relative free energies may be poorly represented (or not at all) and  $\lambda$ -barriers with high relative free energies may be seldom crossed (or not at all) in the course of a finite simulation. In addition, care must be taken to avoid the sampling of  $\lambda$ -values beyond the physical end-states of the alchemical coupling<sup>?</sup> (*i.e.* below 0 or above 1). Both the sampling inhomogeneity and the end-point issues can in principle be remedied<sup>?</sup> by employing an appropriate coordinate transformation<sup>?????????</sup> or/and by applying a biasing potential along  $\lambda$ .<sup>???????</sup> A combination of these two approaches underlies the  $\lambda$ -local elevation umbrella sampling method<sup>???</sup> ( $\lambda$ -LEUS), which relies in particular on an adaptive memory-based biasing potential. A similar principle is also at the heart of numerous other

methods such as the flat-histogram,<sup>??</sup>  $\lambda$ -metadynamics,<sup>???</sup> adaptive integration,<sup>?</sup> adaptive biasing force,<sup>?</sup> adaptively biased<sup>?</sup> and expanded-ensemble<sup>???????</sup> methods.

In  $\lambda$ -LEUS, a local elevation<sup>?</sup> (LE) build-up phase is used to construct a suitable biasing potential, and followed by an umbrella sampling<sup>??</sup> (US) phase where this potential is frozen<sup>?</sup> and biased equilibrium statistics gathered with quasi-homogeneous sampling of the  $\lambda$ -range. Clearly, the LE phase represents an efficiency loss in the method, *i.e.* a dead time. Note, however, that a similar (but generally shorter) dead time also exists in TI and HRE/HRP in the form of discarded initial equilibration times for all replicas. Another drawback of  $\lambda$ -LEUS is that it is not parallelizable in the same fashion as TI and HRE/HRP, where the simulations at different  $\lambda$ -points can be carried out in parallel (see, however, various swarm,<sup>???</sup> multiple-walker<sup>??</sup> and flying Gaussian<sup>??</sup> variants of memory-based biasing methods). Finally, some care must be taken to select an appropriate mass and thermostat-coupling scheme for the  $\lambda$ -variable, so as to ensure that this variable is adequately coupled to the configurational degrees of freedom.<sup>?</sup>

In the present chapter, we propose a new alchemical free-energy calculation scheme with the goal of combining the advantages and alleviating the shortcomings of both HRE/HRP and  $\lambda$ D. This scheme is termed conveyor belt thermodynamic integration (CBTI) and relies on the coupled  $\lambda$ D of a set of system replicas, in which the  $\lambda$ -distance between successive replicas is kept fixed along a forward-turn-backward-turn path, akin to a conveyor belt between the two physical end-states. The basic principle of CBTI is illustrated schematically in Fig. 3.1.

Considering a free-energy profile  $G(\lambda)$  presenting a constant uphill slope between  $A$  and  $B$ , a single system subjected to  $\lambda$ D

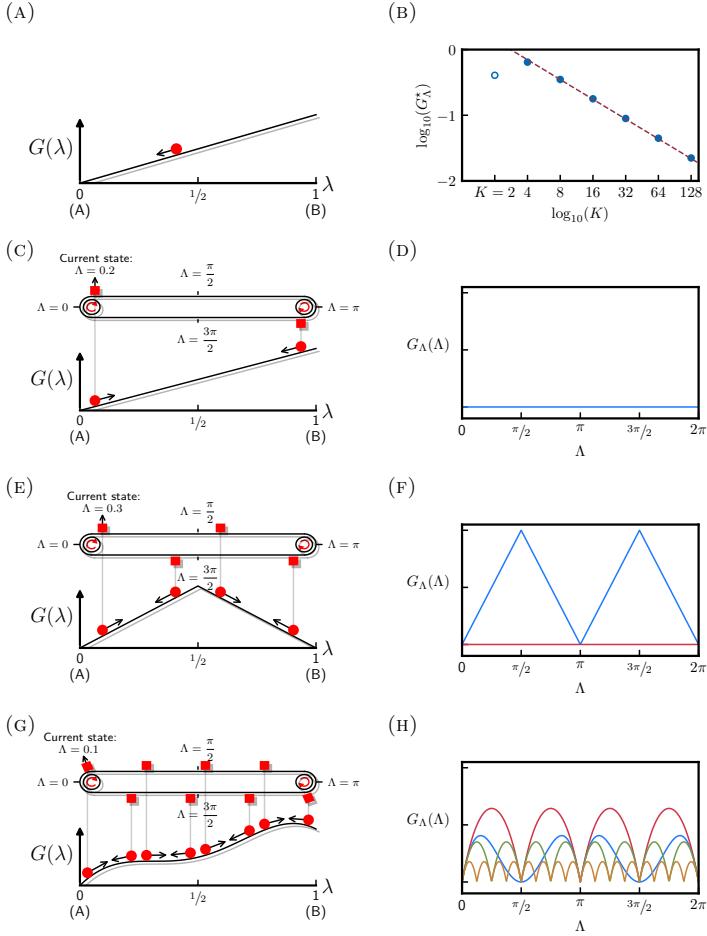


FIGURE 2.1: *Schematic illustration of the conveyor belt thermodynamic integration (CBTI) approach for different free-energy profiles  $G(\lambda)$  and numbers of replicas  $K$ .* The free-energy profile  $G(\lambda)$  is shown on the left along with the conveyor belt for an inclined plane (c), a piecewise-linear curve (e) and a more realistic curve (g), the latter curve corresponding to the function  $G(\lambda) = \sin(\pi\lambda)(\lambda - 0.5)^2 + 0.25\lambda$ . The corresponding free-energy profiles  $G_\Lambda(\Lambda)$  along the conveyor belt advance variable  $\Lambda$  (Eq. 3.16) are shown on the right for  $K = 2$  (blue),  $K = 4$  (red),  $K = 8$  (green) or  $K = 16$  (orange). The top-left panel (a) illustrates the situation of plain unbiased  $\lambda D$ , where the system would “roll down” the slope and keep sampling the neighborhood of the state with the lowest free energy. The top-right panel (b) shows the height  $G_\Lambda^*$  of the residual barriers in the free-energy profile  $G_\Lambda(\Lambda)$  considering this illustrative function and increasing values of  $K$ . The data is represented in logarithmic form and a linear line of slope  $-1$  is fitted to the filled circles.

in the absence of a biasing potential will “roll down” to  $A$  and keep sampling the neighborhood of this state (Fig. 3.1a). To circumvent this problem, one may decide to couple the  $\lambda D$  of two replicas 0 and 1 of the system in such a way that any downhill displacement of replica 0 implies an equivalent uphill displacement of replica 1 (Fig. 3.1c). This working principle is exploited in practice by vehicles like cable cars or funiculars, where the motion of the two vehicles are controlled by a stretched cable connected to two pulleys. It has also been exploited previously for MD in the context of twin-system<sup>??</sup> EDS, which couples forward and reverse alchemical changes performed in two different environments. In this situation, one may describe the  $\lambda$ -values of the two replicas by means of a single periodic angular variable  $\Lambda$  representing the advance of the cable from 0 (replica 0 in state  $A$ , replica 1 in state  $B$ ) to  $\pi$  (0 in  $B$ , 1 in  $A$ ) and then  $2\pi$  (back to the starting situation). Note that the  $\pi$  to  $2\pi$  return range of  $\Lambda$  deviates from the cable car or funicular analogy, where the vehicles never go “over the pulley”. If the free-energy profile  $G(\lambda)$  has a constant slope, it is easily seen that there will be no net driving force on the cable. In the  $\lambda D$  context, this means that the  $\Lambda$ -variable will undergo random diffusion with a homogeneous sampling of  $\Lambda$ , *i.e.* the corresponding free-energy profile  $G_\Lambda(\Lambda)$  will be flat (Fig. 3.1d). Accordingly, each of the two replicas will sample the entire  $\lambda$ -range homogeneously.

Consider now a free-energy profile  $G(\lambda)$  presenting a constant uphill slope from 0 to  $1/2$  and a constant downhill slope from  $1/2$  to 1 (Fig. 3.1e). In a setup with two replicas, the  $G_\Lambda(\Lambda)$  profile will no longer be flat (Fig. 3.1f, blue curve). From 0 to  $\pi/2$  and from  $\pi$  to  $3\pi/2$ , both replicas move uphill, whereas from  $\pi/2$  to  $\pi$  and from  $3\pi/2$  to  $2\pi$ , they both move downhill. This can be remedied by using four instead of only two replicas, placed at equal

distances along the cable. This working principle is now more reminiscent of the real-life situation of a conveyor belt. In this four-replica setup, the  $G_\Lambda(\Lambda)$  profile will again be flat (Fig. 3.1f, red curve), and each of the four replicas will sample the entire  $\lambda$ -range homogeneously.

Finally, consider a more realistic free-energy profile  $G(\lambda)$  (Fig. 3.1g). In the general case, the  $G_\Lambda(\Lambda)$  profile will never be exactly flat irrespective of the number  $K$  of replicas. However, by adding more and more replicas to the conveyor belt, the features of the  $G_\Lambda(\Lambda)$  profile can be progressively reduced in magnitude (as illustrated for  $K = 2, 4, 8$  or  $16$  replicas by the blue, red, green and orange curves of Fig. 3.1h). Note that we restrict the choice of  $K$  to even values, in order to have always the same number of replicas moving forward and backward. Although the choice of an odd number of replicas would be acceptable, it is likely to be less favorable (especially at small  $K$ ), because one extra replica would always move in either of the two directions.

Considering the illustrative  $G(\lambda)$  of Fig. 3.1g, the magnitude  $G_\Lambda^*$  of the variations in  $G_\Lambda(\Lambda)$  is shown in Fig. 3.1b as a function of  $K$ . As discussed in Sect. 3.A, these variations can be interpreted as the residual of a  $K$ -point trapezoid quadrature approximation to the vanishing integral of the derivative of an even periodic function over one period, namely that of the  $G(\lambda)$  profile after mirroring. Quantitatively, this interpretation shows that the magnitude  $G_\Lambda^*$  of these variations decreases at least as  $K^{-1}$  in the limit of large  $K$ . Note that the convergence to a flat profile (*i.e.* the decrease of the barrier heights towards zero upon increasing  $K$ ) does not need to be regular for small  $K$  (see *e.g.* the comparatively low variations for the blue curve in Fig. 3.1h) and that it can be stronger than  $K^{-1}$  for large  $K$  if  $G(\lambda)$  presents particular continuity/symmetry properties. By using a sufficient number of replicas, one may

thus ensure a quasi-homogeneous sampling of the  $\lambda$ -range by each replica, even in the absence of coordinate transformation or biasing potential. If desired, a memory-based biasing potential may still be applied to further homogenize the sampling. However, the LE build-up time can be considerably reduced relative to that needed in a corresponding single-system  $\lambda$ -LEUS simulation, because this biasing potential only needs to be applied to the  $\Lambda$ -variable and no longer to the  $\lambda$ -variable. Thus, it has to compensate for comparatively small  $G_\Lambda(\Lambda)$  variations. Furthermore, it only needs to be constructed over a limited  $\Lambda$ -range, considering that  $G_\Lambda(\Lambda)$  is periodic with period  $2\pi K^{-1}$  as well as even over this  $2\pi K^{-1}$  interval.

Irrespective whether a biasing potential is employed or not, the definition of a free-energy estimator for the CBTI scheme, *i.e.* a procedure to construct  $G(\lambda)$  based on the simulation results for the multiple-replica system, is not as trivial<sup>7</sup> as in the single-system  $\lambda$ D or  $\lambda$ -LEUS cases. Here, CBTI is analyzed using a TI-like estimator<sup>7</sup> relying on the quadrature integration of the average Hamiltonian  $\lambda$ -derivative binned along  $\lambda$  considering all replicas simultaneously. The associated quadrature error can be kept very low by using a large number of bins, which is rendered possible by the continuous and quasi-homogeneous  $\lambda$ -sampling. Although this estimator is very robust, we note that it may not be optimal in terms of statistical efficiency.<sup>7,8,9,10,11,12,13,14</sup>

In the present chapter, we provide the mathematical/physical formulation of the proposed CBTI scheme, and report an initial application of the method to an illustrative alchemical transformation: the mutation of a methanol molecule to a dummy (non-interacting) skeleton in water, giving access to the hydration free energy of the molecule.

## 2.2 THEORY

### 2.2.1 ALCHEMICAL FREE-ENERGY CALCULATIONS

The goal of path-dependent alchemical free-energy calculations is to evaluate the free-energy difference  $\Delta G$  between two states  $A$  and  $B$  of a molecular system, by introducing a coupling scheme relying on a parameter  $\lambda$ , and sampling along the so-defined  $\lambda$ -path. The two states have the same number  $3N$  of degrees of freedom, but distinct Hamiltonian functions  $\mathcal{H}_A(\mathbf{x})$  and  $\mathcal{H}_B(\mathbf{x})$ , respectively, where  $\mathbf{x} = (\mathbf{r}, \mathbf{p})$  is the  $6N$ -dimensional phase-space vector representative of a microscopic system configuration,  $\mathbf{r}$  and  $\mathbf{p}$  being the corresponding coordinate and momentum vectors. The coupling parameter is introduced into a hybrid Hamiltonian  $\mathcal{H}(\mathbf{x}; \lambda)$  satisfying the boundary conditions  $\mathcal{H}(\mathbf{x}; 0) = \mathcal{H}_A(\mathbf{x})$  and  $\mathcal{H}(\mathbf{x}; 1) = \mathcal{H}_B(\mathbf{x})$ , where the semi-colon indicates a parametric dependence.

Since the proposed CBTI scheme encompasses features of both  $\lambda$ D and TI, these two approaches are summarized briefly in the next two subsections. The following three subsections then describe in turn the basis of the CBTI scheme, its free-energy estimator and the application of a biasing potential.

### 2.2.2 $\lambda$ -DYNAMICS ( $\lambda$ D)

In the  $\lambda$ D scheme,<sup>7 8 9 10 11 12 13</sup> the coupling parameter  $\lambda$  is assigned a mass  $m_\lambda$  and a momentum  $p_\lambda$ , and considered to be an additional pseudo-conformational degree of freedom of the system.

The Hamiltonian of the extended system is defined as

$$\mathcal{H}^*(\mathbf{x}, \lambda) = \mathcal{H}(\mathbf{x}; \lambda) + \frac{p_\lambda^2}{2m_\lambda}, \quad (2.1)$$

where the star refers to an extended system in which  $\lambda$  is now a variable and no longer a parameter (thus the replacement of the semi-colon by a comma). This leads to the additional equation of motion

$$\ddot{\lambda} = \frac{\dot{p}_\lambda}{m_\lambda} = -\frac{1}{m_\lambda} \frac{\partial \mathcal{H}(\mathbf{x}; \lambda)}{\partial \lambda} \quad (2.2)$$

for propagating the  $\lambda$ -variable, where a dot over a variable indicates its time derivative.

The free-energy difference  $\Delta G$  between the two physical end-states can then in principle be calculated based on a single thermostated MD simulation of the extended system, as

$$\Delta G = -\frac{1}{\beta} \ln \frac{\langle \delta(\lambda - 1) \rangle^*}{\langle \delta(\lambda) \rangle^*}, \quad (2.3)$$

where  $\beta = (k_B T)^{-1}$ ,  $k_B$  being the Boltzmann constant and  $T$  the absolute temperature,  $\delta$  is the Dirac delta function, and  $\langle \dots \rangle^*$  denotes ensemble averaging for the extended system (*i.e.* over the joint trajectories of  $\mathbf{x}$  and  $\lambda$ ). In practice, the  $\delta$ -functions in Eq. 3.3 must be replaced by two finite end-state bins ( $\lambda$ -cutoff), sufficiently large for proper statistics but also sufficiently small for avoiding distortions due to averaging at the end-states.<sup>?</sup>

### 2.2.3 THERMODYNAMIC INTEGRATION (TI)

In the original TI scheme,<sup>???</sup> a set of  $K$  replicas of the system are simulated in parallel at fixed predefined  $\lambda$ -values in the range

$[0, 1]$ . Since the replicas are entirely decoupled from each other, the simulations can be performed serially as well. However, TI extensions including the HRE scheme<sup>???</sup> and the HRP scheme<sup>???</sup> introduce a coupling in the form of  $\lambda$ -value exchanges, in which case the simulations must really be carried out in parallel. The same will apply to the proposed CBTI scheme, where the coupling involves a synchronization of the dynamical  $\lambda$ -variations.

Considering all replicas  $k = 0 \dots K - 1$  as the members of a replica system, one may note the corresponding  $6K \times N$ -dimensional phase-space vector as  $\mathbf{X} = \{\mathbf{x}_k\}$  and the corresponding  $K$ -dimensional vector containing the fixed  $\lambda$ -values as  $\boldsymbol{\lambda} = \{\lambda_k\}$ . In plain TI, the Hamiltonian of the replica system is defined as

$$\mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda}) = \sum_{k=0}^{K-1} \mathcal{H}(\mathbf{x}_k; \lambda_k) , \quad (2.4)$$

where the dagger refers to a replica system, and  $\boldsymbol{\lambda}$  is here a parameter vector (thus the semi-colon). Because the Hamiltonian of Eq. 3.4 involves no coupling term between the replicas, the dynamics of a replica  $k$  is independent from that of the other replicas and solely depends on  $\lambda_k$ .

The free energy difference  $\Delta G$  between the two states can then be calculated based on a single thermostated MD simulation of the replica system, as

$$\begin{aligned} \Delta G &= \int_0^1 d\lambda' \left\langle \frac{\partial \mathcal{H}(\mathbf{x}; \lambda)}{\partial \lambda} \right\rangle_{\lambda'} \approx \sum_{k=0}^{K-1} w_k \left\langle \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \lambda_k} \right\rangle^\dagger \\ &= \sum_{k=0}^{K-1} w_k \left\langle \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \right\rangle^\dagger , \end{aligned} \quad (2.5)$$

where the  $w_k$  are quadrature weights for the numerical integra-

tion,<sup>???</sup>  $\langle \cdots \rangle_\lambda$  denotes ensemble averaging for a single system (*i.e.* over  $\mathbf{x}$ ) at the given  $\lambda$  value, and  $\langle \cdots \rangle^\dagger$  denotes ensemble averaging for the replica system (*i.e.* over  $\mathbf{X}$ ).

In the above form, TI has long been the workhorse of alchemical free-energy calculations. The method is extremely robust in the sense that the accuracy of the calculated  $\Delta G$  can always be systematically improved (more  $\lambda$ -points, longer equilibration or/and sampling times). However, it is not necessarily the most efficient method to determine  $\Delta G$  up to a certain accuracy, due to possible sub-optimalities in the coupling scheme,<sup>????? ?</sup> the protocol design<sup>????? ?</sup> the free-energy estimator,<sup>????? ? ? ? ?</sup> and the orthogonal sampling.<sup>????? ?</sup>

#### 2.2.4 CONVEYOR BELT THERMODYNAMIC INTEGRATION (CBTI)

The proposed CBTI scheme encompasses features of both  $\lambda$ D and TI. Similarly to TI, it is based on the simulation of a replica system involving  $K$  copies of the molecular system of interest, where  $K$  is taken to be even. And similarly to  $\lambda$ D, the individual replicas are extended systems, for which the associated  $\lambda_k$ -variable is allowed to evolve along the simulation. However, the evolutions of these  $\lambda_k$ -variables are not independent. They are coupled to each other by means of a sequence of hard constraints, so that they follow the course of a conveyor belt (CB). Thus, they are entirely determined by a single dynamical variable  $\Lambda$ , following the scenario depicted in Fig. 3.1 and discussed in the Introduction section.

The variable  $\Lambda$  is a continuous real variable representing the overall advance of the CB, successive multiples of  $2\pi$  corresponding to as many full rotations. Given  $\Lambda$  and  $K$ , the  $\lambda$ -value  $\lambda_k$

associated with a system  $k$  on the CB is obtained as

$$\lambda_k(\Lambda) = \zeta(\Lambda + k\Delta\Lambda) , \quad (2.6)$$

with

$$\Delta\Lambda = 2\pi K^{-1} . \quad (2.7)$$

Here, the function  $\zeta$  is a continuous and periodic zig-zag function of period  $2\pi$  and image range  $[0, 1]$ , defined over the reference period  $[0, 2\pi)$  as

$$\zeta(\theta) = \begin{cases} \pi^{-1}\theta & \text{if } \theta < \pi \\ 2 - \pi^{-1}\theta & \text{if } \theta \geq \pi \end{cases} \quad \text{for } \theta \in [0, 2\pi) , \quad (2.8)$$

where the  $[\cdot, \cdot)$  indicates an interval that is open to the right side, *e.g.*  $[0, 2\pi)$  includes 0 but excludes  $2\pi$ . An advance of the CB by  $\Delta\Lambda$  corresponds to a cyclic permutation of the  $K$  replicas, each system moving by one position forward along the CB, *i.e.*  $\lambda_k(\Lambda + \Delta\Lambda) = \lambda_{k+1}(\Lambda)$  for  $k < K - 1$ , along with  $\lambda_{K-1}(\Lambda + \Delta\Lambda) = \lambda_0(\Lambda)$ . For this reason, the increment  $\Delta\Lambda$  will be further referred to as one shift of the CB. The system  $k = 0$  can be viewed as a reference system, as  $\lambda_0 = \pi^{-1}\Lambda$  for  $0 \leq \Lambda < \pi$  and  $\lambda_0 = 2 - \pi^{-1}\Lambda$  for  $\pi \leq \Lambda < 2\pi$ . Since  $K$  is chosen to be even, an increase of  $\Lambda$  always corresponds to an increase of  $\lambda_k$  for half of the systems (forward-moving side of the CB) and a decrease of  $\lambda_k$  for the other half of the systems (backward-moving side of the CB). This choice also implies that  $\Lambda$ -values which are integer multiples of the CB shift  $\Delta\Lambda$  correspond to situations where there is one system in state  $A$  and one system in state  $B$ . Since an advance of the CB variable by  $2\pi$  leaves the replica system invariant, the variable  $\Lambda$  will commonly be refolded into the reference period  $[0, 2\pi)$  when

illustrating the results of the CBTI method.

The Hamiltonian of the extended replica system is defined as

$$\mathcal{H}^{\dagger\star}(\mathbf{X}, \Lambda) = \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda}) + \frac{p_\Lambda^2}{2m_\Lambda} \quad \text{with} \quad \boldsymbol{\lambda} = \boldsymbol{\lambda}(\Lambda), \quad (2.9)$$

where  $\mathcal{H}^\dagger$  is defined as in TI by Eq. 3.4 and  $\boldsymbol{\lambda}(\Lambda)$  by Eq. 3.6. In analogy with Eq. 3.2, the resulting equation of motion for  $\Lambda$  reads

$$\ddot{\Lambda} = \frac{\dot{p}_\Lambda}{m_\Lambda} = -\frac{1}{m_\Lambda} \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \Lambda} \quad (2.10)$$

where

$$\begin{aligned} \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \Lambda} &= \sum_{k=0}^{K-1} \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \frac{d\lambda_k}{d\Lambda} \\ &= \sum_{k=0}^{K-1} \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \zeta'(\Lambda + 2\pi K^{-1} k) . \end{aligned} \quad (2.11)$$

Here, the function  $\zeta'$  is the derivative of the zig-zag function of Eq. 3.8, given over the reference period  $[0, 2\pi]$  by

$$\zeta'(\theta) = \begin{cases} \pi^{-1} & \text{if } \theta < \pi \\ -\pi^{-1} & \text{if } \theta \geq \pi \end{cases} \quad \text{for } \theta \in [0, 2\pi) . \quad (2.12)$$

Formally, the derivative is not defined when  $\theta$  is an integer multiple of  $\pi$ , *i.e.* for a system that is exactly in one of the physical end-states  $A$  or  $B$ . In this case, the value of  $\zeta'$  has been arbitrarily set to  $\pi^{-1}$  for even multiples and  $-\pi^{-1}$  for odd multiples. This has a negligible impact in practice, as it only concerns a series of infinitesimal points over the entire  $\Lambda$ -range, *i.e.* infinitesimally few configurations along a CBTI simulation. For example, when using double-precision floating-point arithmetics (including denormalized numbers), their probability of occurrence is on the order

of  $10^{-324}K$  (*i.e.* a single expected occurrence over a simulation lasting about  $10^{292}K^{-1}$  times the age of the universe with a 2 fs timestep). Neither does going over the discontinuity within a timestep represent a source of non-conservativeness. The concerned replica will merely bounce back the corresponding physical end-state with a reversion of its velocity, akin to a particle reflected elastically by a hard wall (delta-function force). If desired, these exceptional points could be handled more formally by altering the definition of  $\zeta$ , *e.g.* by smoothing its tips in a narrow range around 0 and  $\pi$ .

The  $\lambda_k$ -dynamics of the individual systems is entirely specified by Eq. 3.10 to propagate  $\Lambda$  along with Eq. 3.6 to calculate the  $\lambda_k$ -values from the current  $\Lambda$ . Alternatively, one may write an equation of motion for the  $\lambda_k$ -variables of the individual replicas by combining the two equations (along with Eq. 3.11) as

$$\begin{aligned}\ddot{\lambda}_k &= \ddot{\Lambda} \frac{d\lambda_k}{d\Lambda} + \dot{\Lambda}^2 \frac{d^2\lambda_k}{d\Lambda^2} \\ &= -\frac{1}{m_\Lambda} \left( \sum_{l=0}^{K-1} \frac{\partial \mathcal{H}(\mathbf{x}_l; \lambda_l)}{\partial \lambda_l} \zeta' (\Lambda + 2\pi K^{-1} l) \right) \\ &\quad \zeta' (\Lambda + 2\pi K^{-1} k) .\end{aligned}\tag{2.13}$$

Note that the term in  $\dot{\Lambda}^2$  vanishes since the second derivative  $\zeta''$  of  $\zeta$  is zero (except at the exceptional singular points). However, it should not be overlooked if one decides to use a different function  $\zeta$ . Introducing the vector  $\mathbf{D}$  and the symmetric  $\Lambda$ -dependent matrix  $\underline{\mathbf{C}}$  defined by their components as

$$\begin{aligned}D_k &= \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \quad \text{and} \\ C_{kl}(\Lambda) &= \pi^2 \zeta' (\Lambda + 2\pi K^{-1} k) \zeta' (\Lambda + 2\pi K^{-1} l) ,\end{aligned}\tag{2.14}$$

Eq. 3.13 can be rewritten in an elegant matrix form as

$$\ddot{\boldsymbol{\lambda}} = -\frac{\mathbf{C}(\Lambda)}{\pi^2 m_\Lambda} \mathbf{D} . \quad (2.15)$$

The elements of the symmetric matrix  $\mathbf{C}$  are either -1 (pair of systems currently on opposite sides of the CB, and thus moving in opposite directions) or +1 (pair of systems currently on the same side of the CB, and thus moving in the same direction). The diagonal elements are all +1, and the other +1 values surround the diagonal (line- and column-wise), the rest being -1 values. Because  $K$  is even, the two types of values are always equally represented in the matrix, specific locations depending on  $\Lambda$ . Note that the variable  $\Lambda$  itself still needs to be explicitly propagated using Eq. 3.10.

For a given configuration  $\mathbf{X}$  of the replica system, the Hamiltonian  $\mathcal{H}^\dagger$  of Eq. 3.4 (together with Eq. 3.6) is periodic in  $\Lambda$  with a period  $2\pi$  corresponding to a full rotation of the CB. However, because the Hamiltonians of the individual replicas are identical, upon ensemble averaging over  $\mathbf{X}$ , one expects the calculated properties to be periodic over  $\Lambda$  with a smaller period  $\Delta\Lambda$ , corresponding to one shift of the CB. This is in particular the case for the probability distribution  $P(\Lambda)$  along  $\Lambda$  and the associated free-energy profile  $G_\Lambda(\Lambda)$ , given by

$$\begin{aligned} G_\Lambda(\Lambda) &= G_\Lambda(0) + \int_0^\Lambda d\Lambda' \left\langle \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \Lambda} \right\rangle_{\Lambda'}^\dagger \\ &= \tilde{G}_\Lambda(0) + \sum_{k=0}^{K-1} G(\lambda_k) \quad \text{with } \boldsymbol{\lambda} = \boldsymbol{\lambda}(\Lambda) , \end{aligned} \quad (2.16)$$

where  $\boldsymbol{\lambda}(\Lambda)$  is defined by Eq. 3.6,  $\langle \cdots \rangle_\Lambda^\dagger$  denotes ensemble averaging for the replica system (*i.e.* over  $\mathbf{X}$ ) at the given  $\Lambda$  value, and

the second equality follows from Eqs. 3.4 and 3.5 (the unknown constant  $\tilde{G}_\Lambda(0)$  is equal to  $G_\Lambda(0)$  increased by a sum of  $-G(\lambda_k(0))$  offsets).

Owing to this periodicity over a smaller interval, it is convenient to introduce a fractional advance variable  $\tilde{\Lambda}$  defined as

$$\tilde{\Lambda} = \gamma(\Lambda, \Delta\Lambda), \quad (2.17)$$

where

$$\gamma(\theta, \theta_o) = \theta_o (\theta_o^{-1}\theta - \lfloor \theta_o^{-1}\theta \rfloor) \quad (2.18)$$

returns the part of  $\theta$  in excess of the closest lower integer multiple of  $\theta_o$ . In contrast to  $\Lambda$ , which is an unbounded variable,  $\tilde{\Lambda}$  only spans a finite definition interval  $[0, \Delta\Lambda]$ . At full convergence, any average property binned as a function of  $\Lambda$  over the interval  $[0, 2\pi]$  will consist of  $K$  successive repeats of the same property binned as a function of  $\tilde{\Lambda}$  over its definition interval  $[0, \Delta\Lambda]$ , as observed in Fig. 3.1h for the free energy  $G_\Lambda(\Lambda)$ . Accordingly, in the absence of full convergence along  $\Lambda$ , binning as a function of  $\tilde{\Lambda}$  over the interval  $[0, \Delta\Lambda]$  followed by  $K$ -fold replication provides an efficient way to construct a more accurate representation of any  $\Lambda$ -resolved average quantity. In fact, the definition interval of  $\tilde{\Lambda}$  could be further halved by noting that, upon ensemble averaging over  $\mathbf{X}$  and for any  $\Lambda$  value that is an integer multiple of  $\Delta\Lambda$ , a forward move of the CB produces the same result as a backward move of the same magnitude. Consequently,  $\Lambda$ -resolved average properties are even over successive  $2\pi K^{-1}$  intervals, as also observed in Fig. 3.1h for the free energy  $G_\Lambda(\Lambda)$ . The corresponding information is thus entirely encompassed in an interval of size  $\Delta\Lambda/2$ .

The normalized probability distribution  $p(\lambda)$  along the cou-

pling variable  $\lambda$  considering all the replicas is defined by

$$p(\lambda) = K^{-1} \sum_{k=0}^{K-1} \langle \delta(\lambda_k - \lambda) \rangle^{\dagger\star} , \quad (2.19)$$

where  $\langle \dots \rangle^{\dagger\star}$  denotes ensemble averaging for the extended replica system (*i.e.* over the joint trajectories of  $\mathbf{X}$  and  $\boldsymbol{\lambda}$ ). At full convergence, this probability over the interval  $[0, 1]$  will consist of  $K/2$  successive repeats of the corresponding distribution over the interval  $[0, 2K^{-1}]$ . More precisely, the distribution  $p(\lambda)$  is related to the distribution  $\tilde{P}(\tilde{\Lambda})$  of  $\tilde{\Lambda}$  over interval  $[0, \Delta\Lambda]$  as

$$p(\lambda) = \Delta\Lambda \tilde{P}(\pi\gamma(\lambda, \pi^{-1}\Delta\Lambda)) . \quad (2.20)$$

In plain words, this means that  $\tilde{P}(\tilde{\Lambda})$  is the relevant quantity in terms of sampling along the coupling variable  $\lambda$ . If it is close to uniform over the range  $[0, \Delta\Lambda]$ , then  $p(\lambda)$  will also be close to uniform over the range  $[0, 1]$ . Here again, it is noted that  $p(\lambda)$  is also even over the interval  $[0, 2K^{-1}]$ , and could be mapped to a  $\tilde{\Lambda}$  value defined over an interval of size  $\Delta\Lambda/2$  instead of  $\Delta\Lambda$  if desired, as

$$p(\lambda) = \begin{cases} \Delta\Lambda \tilde{P}(\pi\gamma(\lambda, \pi^{-1}\Delta\Lambda)) & \text{if } \gamma(\lambda, \pi^{-1}\Delta\Lambda) < (2\pi)^{-1}\Delta\Lambda \\ \Delta\Lambda \tilde{P}(\pi\gamma(1 - \lambda, \pi^{-1}\Delta\Lambda)) & \text{otherwise} \end{cases} . \quad (2.21)$$

## 2.2.5 CBTI FREE-ENERGY ESTIMATOR

Due to the constraints coupling the  $\lambda_k$ -values of the  $K$  replicas, the function  $p(\lambda)$  of Eq. 3.19 is by no means a Boltzmann distribution in terms of the single-system Hamiltonian. In fact, as seen above,

it consists at full convergence of  $K/2$  successive repeats of the same even curve. In addition, compared to the Boltzmann distribution, it will be significantly flatter. On the one hand, the smaller amplitude of variations are desired, as they will lead to more homogeneous sampling and are expected to ease transitions along  $\Lambda$  (up to the limit imposed by the speed of random diffusion). On the other hand, it is no longer possible to evaluate the free-energy difference  $\Delta G$  directly from  $p(\lambda)$  in analogy with the  $\lambda D$  expression of Eq. 3.3. However, since the dynamics remains Hamiltonian and the coupling between replicas does not involve the configurational degrees of freedom, the change from TI to CBTI does not affect the conditional probabilities  $\mathcal{P}(\mathbf{x}|\lambda)$ . Thus, configurational ensemble averages sorted by  $\lambda$ -values will remain identical to those one would obtain from TI (or from HRE/HRP or  $\lambda D$ ). As a result,  $\Delta G$  can still be obtained by integrating over the average Hamiltonian derivative binned as a function of  $\lambda$  considering all replicas simultaneously, in analogy with the TI expression of Eq. 3.5. Note that the exceptional points of the function  $\zeta'$  (discussed previously in the context of Eq. 3.12) have no influence on the integration, as they represent finite discontinuities over infinitesimal ranges.

In practice,  $\Delta G$  is calculated here based on a single thermostated MD simulation of the extended replica system, as

$$\begin{aligned}\Delta G &= \int_0^1 d\lambda' K^{-1} \sum_{k=0}^{K-1} \left\langle \frac{\mathcal{H}(x_k; \lambda_k)}{\partial \lambda_k} \delta(\lambda_k - \lambda') \right\rangle^{\dagger*} \quad (2.22) \\ &\approx \sum_{j=0}^{J-1} \left\langle \frac{\sum_{k=0}^{K-1} \frac{\mathcal{H}(x_k; \lambda_k)}{\partial \lambda_k} \alpha(\lambda_k, j; J)}{\sum_{k=0}^{K-1} \alpha(\lambda_k, j; J)} \right\rangle^{\dagger*},\end{aligned}$$

where

$$\alpha(\theta, j; J) = \begin{cases} 1 & \text{if } j \leq J\theta < j + 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.23)$$

is a binning function corresponding to a discretization of the  $\lambda$ -interval  $[0, 1]$  using  $J$  bins. The approximation in Eq. 3.22 corresponds to a simple forward rectangular quadrature, where the Hamiltonian derivative is averaged over the  $J$  successive bins considering all replicas. Since  $\tilde{P}(\tilde{\Lambda})$ , and thus  $p(\lambda)$ , will typically be close to homogeneous,  $J$  can be taken very large, resulting in a negligible quadrature error. For example, if  $K$  replicas sample  $L$  configurations each, the number of data points per bin will be close to  $KL/J$ , with limited variations across bins. Defining the maximal allowed value  $J_{\max}$  as the highest value of  $J$  for which empty bins (vanishing denominator in the ensemble average of Eq. 3.22) never occur, a graph of  $\Delta G$  evaluated upon increasing  $J$  from 1 to  $J_{\max}$  will rapidly level off to a plateau when quadrature errors become negligible. Two variants which do not require the specification of a number of bins are also proposed in Sect. 3.C (Eqs. 3.C.1 and 3.C.2).

## 2.2.6 CBTI WITH MEMORY-BASED BIASING POTENTIAL

When using a large number of replicas, the sampling along  $\lambda$  afforded by the CBTI scheme will be close to homogeneous. However, for practical reasons (*e.g.* number of processors available on a computer node), one may wish to use a small number of replicas. In this case, the sampling homogeneity can be enhanced by addition of a biasing potential. It is sufficient to apply this potential

to the fractional advance variable  $\tilde{\Lambda}$  over the range  $[0, \Delta\Lambda]$ . With inclusion of a biasing potential  $\mathcal{B}$ , Eq. 3.10 becomes

$$\ddot{\Lambda} = -\frac{1}{m_{\Lambda}} \frac{\partial}{\partial \Lambda} \left( \mathcal{H}^{\dagger}(\mathbf{X}; \boldsymbol{\lambda}) + \mathcal{B}(\tilde{\Lambda}) \right)$$

with  $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\Lambda)$  and  $\tilde{\Lambda} = \tilde{\Lambda}(\Lambda)$  , (2.24)

where  $\mathcal{H}^{\dagger}$  is defined by Eq. 3.4,  $\boldsymbol{\lambda}(\Lambda)$  by Eq. 3.6 and  $\tilde{\Lambda}(\Lambda)$  by Eq. 3.17.

In analogy with the  $\lambda$ -LEUS scheme,<sup>7 8 9 10</sup> this biasing potential can be expressed as a sum of local grid-based spline functions, built in a LE preoptimization phase and frozen in a subsequent US sampling phase. However, the duration of the LE phase can be considerably reduced compared to a single-system  $\lambda$ -LEUS simulation, considering that  $\tilde{P}(\tilde{\Lambda})$  is already close to homogeneity in the absence of biasing and that the support interval is reduced to the  $\tilde{\Lambda}$ -range  $[0, \Delta\Lambda]$ . The latter interval can actually be further restricted to  $[0, \Delta\Lambda/2]$  considering the even symmetry of  $\tilde{P}(\tilde{\Lambda})$ , *i.e.* by enforcing an even symmetry of  $\mathcal{B}$  as well.

Since the application of a biasing potential that only involves the  $\lambda_k$ -variables does not alter the conditional probabilities  $\mathcal{P}(\mathbf{x}|\boldsymbol{\lambda})$ , Eq. 3.22 (or the variants of Eqs. 3.C.1 and 3.C.2) can still be employed without any modification to evaluate the free-energy change. In other words, in contrast to the  $\lambda$ -LEUS scheme, the CBTI scheme with the presented TI-like free-energy estimator does not require any reweighting.

## 2.3 COMPUTATIONAL DETAILS

### 2.3.1 TEST SYSTEM

As an initial application of the proposed CBTI scheme, we considered here a relatively simple perturbation, namely the conversion of methanol from a fully interacting molecule to a dummy skeleton (no intermolecular interactions) in an aqueous environment at  $P = 1$  bar and  $T = 298.15$  K. The calculations were performed using a modified version of the GROMOS11 program<sup>7, 18, 19</sup> along with the parameters of the GROMOS-compatible 2016H66 force field<sup>7</sup> for methanol<sup>7</sup> (united atom, rigid bonds, flexible bond angle) and the simple point charge (SPC) model<sup>7</sup> for water (fully rigid). Since the dummy skeleton retains the intramolecular interactions (here, only the bond angle), the calculated free-energy change  $\Delta G$  corresponds directly to minus the hydration free energy of methanol.

Possible issues related to the existence of a singularity<sup>7, 20, 21</sup> and the insufficient solute-solvent kinetic-energy exchange<sup>7, 22, 23</sup> close to  $\lambda = 1$  were alleviated in the usual way, by means of a soft-core scheme<sup>7</sup> for the alchemical coupling and of stochastic dynamics<sup>7, 24</sup> (SD) for thermostating the solute and solvent conformational degrees of freedom. In most CBTI simulations, the instantaneous temperature  $T_\Lambda$  of the CB advance variable  $\Lambda$  was also controlled separately by means of a Nosé-Hoover chain thermostat<sup>25</sup> at 298.15 K (eight successive thermostat variables), with a coupling time  $\tau_\Lambda$ .

### 2.3.2 SIMULATIONS SETS

The exploration of the CBTI scheme and the comparison of its performance with that of existing methods was carried out in five successive steps: (1) establishing reference TI results; (2) analyzing the influence of the CBTI parameters (number  $K$  of replicas along with the mass-parameter  $m_\Lambda$  and thermostat coupling time  $\tau_\Lambda$  of the CB advance variable); (3) investigating the use of a biasing potential; (4) examining the features of the TI-like free-energy estimator (effect of the number  $J$  of integration bins and use of equations without specification of  $J$ ); (5) comparing the results of CBTI with those of existing methods.

The reference TI calculations (Step 1) were performed using  $K_{\text{TI}} = 2^n + 1$  equidistant  $\lambda$ -points covering the range  $[0, 1]$  with  $n = 1, 2, \dots, 7$ . They involved initial configurations equilibrated for 0.2 ns starting from the equilibrated configuration at the previous  $\lambda$ -point, and a simulation time of  $100K_{\text{TI}}^{-1}$  ns per  $\lambda$ -point. Each of these calculations, involving a total single-system sampling time of 100 ns, was repeated ten times using different random initial velocities. The integration over the average Hamiltonian derivative was performed based on Eq. 3.5 using the Simpson quadrature rule.<sup>?</sup> <sup>?</sup> <sup>?</sup>

To explore the influence of the CBTI parameters (Step 2), various combinations of  $K$ ,  $m_\Lambda$  and  $\tau_\Lambda$  were considered in three series of calculations, namely: (i) the choices  $m_\Lambda = 16, 160, 800, 1600$  or  $3200 \text{ u nm}^2$  (where  $\text{u}$  stands for atomic mass unit, *i.e.* g mol<sup>-1</sup>), along with  $K = 16$  replicas in the absence of thermostat coupling for  $\Lambda$ , *i.e.* with  $\tau_\Lambda \rightarrow \infty$ ; (ii) the choices  $\tau_\Lambda = 0.05, 0.1, 0.5, 1$  or 2 ps along with  $K = 16$  replicas and  $m_\Lambda = 160 \text{ u nm}^2$ ; (iii) the choices  $K = 8, 16, 32, 64$  or 128 along with  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5$  ps; The parameters (and

results) of these three series of simulations, including their durations  $t_{\text{sim}}$ , are summarized in Tab. 3.1 (entries 1-15). All these simulations were preceded by 0.2 ns equilibration. For the the third series (entries 11-15),  $m_\Lambda$  was made proportional to  $K^{1/2}$ , an arbitrary parameter choice justified by arguments provided in Sect. 3.B, and the five simulations relied on the the same total single-system sampling time of 256 ns. This exploration showed that the CBTI method is rather robust with respect to the choice of its parameters. The values  $K = 16$ ,  $m_\Lambda = 160 \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  were retained as a good combination for the alchemical perturbation considered. For comparison with the TI results of Step 1, ten repeats of the calculation involving this specific choice were performed using different random initial velocities and a total single-system sampling time of 100 ns after 0.2 ns equilibration.

The application of CBTI with a biasing potential (Step 3) was investigated in the context of simulations with  $K = 8$  or  $16$ , both with  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$ . For  $K = 8$ , the biasing potential  $\mathcal{B}$  (Eq. 3.24) was constructed using  $N_{\text{gp}} = 34$  basis functions centered at equidistant grid-points  $i = 0, \dots, N_{\text{gp}} - 1$  over the  $\tilde{\Lambda}$ -range  $[0, \pi/4]$ . The coefficients of the basis-functions  $i$  and  $N_{\text{gp}} - 1 - i$  with  $i = 0, \dots, N_{\text{gp}}/2 - 1$  were constrained to be identical, considering the expected even symmetry of  $\tilde{P}(\tilde{\Lambda})$ . In terms of the CB advance variable  $\Lambda$ , this means that the biasing potential relied in effect on  $K(N_{\text{gp}} - 1) = 264$  local functions covering the  $\Lambda$ -range  $[0, 2\pi]$ , these functions being defined by only 17 independent coefficients. For  $K = 16$ ,  $\mathcal{B}$  relied on  $N_{\text{gp}} = 18$  basis functions over the  $\tilde{\Lambda}$ -range  $[0, \pi/8]$ , leading to 272 functions over the  $\Lambda$ -range  $[0, 2\pi]$  defined by 9 independent coefficients. Second-order splines<sup>??</sup> (of range  $\pm 2\delta$  with  $\delta = \pi/132$  or  $\pi/136$  for  $K = 8$  and  $16$ , respectively) were employed<sup>?</sup> as basis functions. An initial build-up force constant  $c_{\text{LE}} = 10^{-3} \text{ kJ mol}^{-1}$  was used,

which was multiplied by a reduction factor  $f_{\text{red}} = 0.1$  after each double-sweep of half the  $\tilde{\Lambda}$ -range  $[0, \pi K^{-1}]$ . The duration  $t_{\text{LE}}$  of the LE build-up phase for the replica system was = 0.15 ns for  $K = 8$  and 0.07 ns for  $K = 16$ , corresponding to only 1.1 – 1.2 ns total single-system simulation time. The parameters (and results) of these two simulations are summarized in Tab. 3.1 (entries 16 and 17). The duration  $t_{\text{sim}}$  of the US sampling phases for the replica system were 22 ns, corresponding to total single-system sampling times of 176 ns and 352 ns for  $K = 8$  and 16, respectively.

To examine the features of the TI-like free-energy estimator (Step 4), the number  $J$  of integration bins used in the rectangular quadrature to calculate  $\Delta G$  (Eq. 3.22) was varied considering the simulations of Steps 2 and 3 above (17 simulations of Tab. 3.1). The resulting  $\Delta G$  values were also compared to those of the variants  $\Delta G_{\text{alt}}$  (Eq. 3.C.1) and  $\Delta G_{\text{app}}$  (Eq. 3.C.2) proposed in Sect. 3.C.

Finally, the results of the CBTI simulations were compared to those of other methods (Step 5), namely TI or HRE using Simpson quadrature as estimator as well as TI using EXTI and MBAR as estimator. For Simpson quadrature, the TI simulations relied on  $K_{\text{TI}} = 3, 5, 9, 17, 65$  or 129 equispaced  $\lambda$ -points, and the HRE simulations relied on  $K_{\text{HRE}} = 17, 33$  or 65 equispaced replicas with exchange attempts every  $\tau_{\text{HRE}} = 0.2$  ps. The use of EXTI and MBAR was explored based on TI-like simulations relying on  $K_{\text{TI}} = 9$  and 17 equispaced  $\lambda$ -points. For EXTI, the average Hamiltonian derivative was extrapolated during the  $K_{\text{TI}}$  simulations to 129 equispaced virtual  $\lambda$ -points, and the latter 129 values used in the Simpson quadrature. For MBAR, the Hamiltonian was calculated at 129 equispaced virtual  $\lambda$ -points considering all the configurations sampled in the  $K_{\text{TI}}$ , and the data combined using the MBAR equation<sup>7</sup> as implemented in

TABLE 2.1: *Influence of the CBTI parameters in simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This table investigates the influence of the parameters selected for the CBTI scheme on the temperature and dynamics of the CB advance variable  $\Lambda$  and on the calculated free-energy change  $\Delta G$  (for the latter, considering a constant total single-system sampling time of 100 ns). For each simulation, the successive entries are: the index of the simulation (sim), the number  $K$  of replicas, the simulation time  $t_{\text{sim}}$  for the replica system, the mass-parameter  $m_\Lambda$ , the thermostat coupling time  $\tau_\Lambda$  ( $\infty$  indicates that no coupling is applied), the average temperature  $T_\Lambda$ , the root-mean-square fluctuation  $\sigma_\Lambda$  of  $\dot{\Lambda}$ , the autocorrelation time  $\tau_{\dot{\Lambda}}$  of  $\dot{\Lambda}$ , the diffusion coefficient  $D_\Lambda$  (Eq. 3.26), the free-energy difference  $\Delta G$  calculated using Eq. 3.22 with  $J = 500$  (except for entry 11,  $J = 200$ ), the alternative free-energy difference  $\Delta G_{\text{alt}}$  calculated using Eq. 3.C.1, and the approximate free-energy difference  $\Delta G_{\text{app}}$  calculated using Eq. 3.C.2. Error estimates obtained by bootstrapping (no Student  $t$ -factor included) are also reported between parentheses for  $\Delta G$ ,  $\Delta G_{\text{alt}}$  and  $\Delta G_{\text{app}}$ . Note that the simulations differ in terms of total single-system sampling time  $Kt_{\text{sim}}$ . To enable a fair comparison, the free energy-changes and associated errors have been calculated after truncating the all simulations to 100 ns single-system sampling time evenly distributed over all replicas. Associated graphs for the distributions  $P(\Lambda)$  of  $\Lambda$ ,  $P_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  and  $P_{\ddot{\Lambda}}$  of  $\ddot{\Lambda}$ , as well as the mean-square displacements  $d_\Lambda$  of  $\Lambda$  and autocorrelation functions  $c_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  can be found in Figs. 3.3, 3.5 and 3.7, or in Figs. Fig. 3.F.1 - Fig. 2.F.15 and Fig. 2.H.2-Fig. 2.H.4. Simulations 11, 13, 15, 16 and 17 are discussed in the main text. The other simulations are discussed in Sect. 3.B.

sim	$K$	$t_{\text{sim}}$ [ns]	$m_\Lambda$ [u nm <sup>2</sup> ]	$\tau_\Lambda$ [ps]	$T_\Lambda$ [K]	$\sigma_\Lambda$ [ps <sup>-1</sup> ]	$\tau_{\dot{\Lambda}}$ [ps]	$D_\Lambda$ [ns <sup>-1</sup> ]	$\Delta G$ [kJ mol <sup>-1</sup> ]	$\Delta G_{\text{alt}}$ [kJ mol <sup>-1</sup> ]	$\Delta G_{\text{app}}$ [kJ mol <sup>-1</sup> ]
1	16	10	16	$\infty$	308.0	0.40	0.02	16.7	21.14 (0.16)	21.23 (0.17)	20.21 (0.32)
2	16	10	160	$\infty$	297.9	0.12	0.61	13.2	21.27 (0.17)	21.32 (0.15)	20.29 (0.31)
3	16	10	800	$\infty$	296.3	0.05	1.90	10.2	21.25 (0.15)	21.34 (0.16)	20.27 (0.32)
4	16	10	1600	$\infty$	304.0	0.04	3.47	10.3	21.25 (0.15)	21.25 (0.17)	20.32 (0.32)
5	16	10	3200	$\infty$	297.7	0.03	5.63	7.4	21.23 (0.16)	21.24 (0.18)	20.11 (0.31)
6	16	10	160	0.05	284.0	0.12	0.37	6.9	21.23 (0.16)	21.04 (0.15)	20.22 (0.31)
7	16	10	160	0.10	291.7	0.12	0.36	7.4	21.48 (0.13)	21.41 (0.16)	20.47 (0.29)
8	16	10	160	0.50	296.4	0.12	0.41	9.1	21.42 (0.14)	21.48 (0.16)	20.45 (0.30)
9	16	10	160	1.00	299.4	0.12	0.49	12.9	21.57 (0.16)	21.48 (0.17)	20.59 (0.30)
10	16	10	160	2.00	296.8	0.12	0.56	13.1	21.48 (0.13)	21.73 (0.15)	20.45 (0.29)
11	8	32	113	0.50	298.4	0.15	0.23	4.4	21.69 (0.40)	21.69 (0.44)	15.65 (0.31)
12	16	16	160	0.50	298.3	0.12	0.41	10.2	21.42 (0.16)	21.48 (0.17)	20.45 (0.33)
13	32	8	226	0.50	300.6	0.10	0.38	6.2	21.44 (0.14)	21.33 (0.15)	21.34 (0.32)
14	64	3	320	0.50	297.2	0.09	0.31	2.5	21.32 (0.14)	21.27 (0.14)	21.29 (0.30)
15	128	2	452	0.50	304.1	0.08	0.24	2.3	21.43 (0.12)	21.61 (0.14)	21.41 (0.32)
16	8	22	113	0.50	296.4	0.15	0.43	15.3	21.48 (0.16)	21.61 (0.17)	19.65 (0.34)
17	16	22	160	0.50	299.3	0.12	0.43	9.4	21.30 (0.13)	21.43 (0.16)	20.61 (0.34)

pymbar.<sup>?</sup> All the above comparisons were performed at a total single-system sampling time of 100 ns distributed evenly over all replicas.

Error bars on the calculated  $\Delta G$  values were estimated in two different ways. For the calculations involving ten repeats (all TI calculations plus one CBTI simulation), the standard deviation  $\sigma$  of the mean was calculated by scaling that of the ten estimates by the square-root of nine, and the error  $\epsilon$  on the mean was calculated as  $\epsilon = 2.262\sigma$ , where 2.262 is the Student *t*-factor<sup>?</sup> for nine degrees of freedom and a two-sigma confidence interval of 95%. For the individual calculations that were not repeated, the statistical error was estimated by bootstrapping<sup>??</sup> (no Student *t*-factor included) using 100 bootstrap samples. If  $K$  replicas (CBTI) or  $\lambda$ -points (all other methods) have generated as many sets of  $L$  data points, a sample consists here of  $K$  sets of  $L$  data points selected randomly (possibly multiple times) from the  $K$  original data sets. Note that for TI/EXTI and TI/MBAR, it is essential to perform the bootstrapping based on the data from the  $K_{\text{TI}}$  real  $\lambda$ -points, and not from the  $129K_{\text{TI}}$  predicted values (the latter procedure would result in underestimated errors due to correlation in the derived data). The bootstrapping error will only be accurate provided that the data from the simulations, written to file every 2 ps, is uncorrelated in time. Normalized autocorrelation functions and characteristic times for the average Hamiltonian derivative in the different TI-simulations are provided in Fig. 3.D.1 and Tab. 3.D.1 to support this assumption.

### 2.3.3 SIMULATION PARAMETERS

The simulations involved a cubic computational box containing one methanol and 1000 water molecules under periodic boundary

conditions in the isothermal-isobaric ensemble at  $P = 1$  bar and  $T = 298.15$  K. They were performed using SD by integrating the Langevin equation of motion<sup>?</sup> using the leap-frog scheme<sup>?</sup> (SD variant<sup>?</sup>) with a timestep  $\Delta t = 2$  fs and a friction coefficient  $\gamma = 10\text{ ps}^{-1}$ . Since the kinetics of the system is irrelevant in this work, SD instead of thermostated MD was used to avoid problems related to insufficient solute-solvent kinetic-energy exchange<sup>???</sup> close to  $\lambda = 1$  (dummy-skeleton state). The value of  $\gamma$  corresponds to the coupling time of 0.1 ps commonly employed in GROMOS simulations<sup>??</sup> relying on a weak-coupling<sup>?</sup> thermostat. The average pressure was maintained close to its reference value by isotropic weak coupling<sup>?</sup> using a molecular virial, a coupling time  $\tau_P = 0.5$  ps and a compressibility  $\kappa = 4.575 \cdot 10^4 \text{ kJ mol}^{-1} \text{ nm}^{-3}$  as commonly used in GROMOS for aqueous biomolecular systems.<sup>??</sup> The bond rigidity of methanol and the full rigidity of water were enforced by application of the SHAKE algorithm<sup>?</sup> with a relative geometric tolerance of  $10^{-4}$ . The energies and Hamiltonian derivatives were written to file every 2 ps for analysis.

The non-bonded interactions were handled by means of a molecule-based twin-range cutoff scheme<sup>?</sup> with short- and long-range cutoff distances set to 0.8 and 1.4 nm, respectively, and an update frequency of 5 timesteps for the short-range pairlist and intermediate-range interactions. The molecule center was the center of geometry for methanol and the oxygen atom for water. A reaction-field correction<sup>??</sup> was applied to account for the mean effect of the electrostatic interactions beyond the long-range cutoff distance, using a relative dielectric permittivity of 61 as appropriate for the SPC model.<sup>?</sup> To alleviate issues related to the existence of a singularity<sup>???</sup> close to  $\lambda = 1$  (dummy-skeleton state), the alchemical transformation relied on

a soft-core scheme,<sup>7</sup> applied with the parameters  $\alpha_{\text{LJ}} = 0.5$  and  $\alpha_{\text{CRF}} = 0.5 \text{ nm}^2$ .

For the CBTI calculations, the propagation of the  $\Lambda$  variable (Eq. 3.10) preceded that of the conformational degrees of freedom, and was performed with the same timestep  $\Delta t$ . More precisely, the following leap-frog steps were carried out in sequence:  $\dot{\Lambda}(t - \Delta t/2) \rightarrow \dot{\Lambda}(t + \Delta t/2)$ ,  $\Lambda(t) \rightarrow \Lambda(t + \Delta t)$ , calculate  $\lambda$  from  $\Lambda$  using Eq. 3.6,  $\dot{\mathbf{r}}(t - \Delta t/2) \rightarrow \dot{\mathbf{r}}(t + \Delta t/2)$  and  $\mathbf{r}(t) \rightarrow \mathbf{r}(t + \Delta t)$ . Unless otherwise specified (explorative simulations), the CBTI simulations relied on a mass-parameter  $m_\Lambda$  set to  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$ , on a coupling time  $\tau_\Lambda$  set to 0.5 ps, and on the use of  $J = 500$  bins for evaluating  $\Delta G$  based on Eq. 3.22 ( $J = 200$  for the unbiased simulation with  $K = 8$ ).

### 2.3.4 TRAJECTORY ANALYSIS

The dynamics of the replica system in the CBTI simulations was characterized by monitoring the distribution  $P$  of the  $\Lambda$  variable, the average temperature  $T_\Lambda$ , the distribution  $P_{\dot{\Lambda}}$  of  $\dot{\Lambda}$ , the associated root-mean-square fluctuation  $\sigma_{\dot{\Lambda}}$ , the normalized autocorrelation function  $c_{\dot{\Lambda}}$  of  $\dot{\Lambda}$ , the associated autocorrelation time  $\tau_{\dot{\Lambda}}$ , the mean-square displacement  $d_\Lambda$  of  $\Lambda$  as a function of time, the associated diffusion coefficient  $D_\Lambda$ , and the distribution  $P_{\ddot{\Lambda}}$  of  $\ddot{\Lambda}$ .

The distribution  $P_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  can be compared to the analytical one-dimensional Maxwell-Boltzmann velocity distribution<sup>7</sup>

$$P_{\dot{\Lambda}}^{\text{MB}}(\dot{\Lambda}) = \left( \frac{\beta m_\Lambda}{2\pi} \right)^{\frac{1}{2}} e^{-\frac{\beta m_\Lambda}{2}\dot{\Lambda}^2}. \quad (2.25)$$

The diffusion coefficient  $D_\Lambda$  was calculated from the mean-square displacement  $d_\Lambda$  of  $\Lambda$  as a function of time  $t$  according to

the one-dimensional Einstein equation?

$$D_\Lambda = \lim_{t \rightarrow \infty} \frac{d_\Lambda(t)}{2t}$$

with  $d_\Lambda(t) = \langle [\Lambda(\tau + t) - \Lambda(\tau)]^2 \rangle_t$ , (2.26)

where  $\langle \dots \rangle_t$  denotes averaging over  $\tau$  (all possible time origins) at constant  $t$ . Note that this equation must be applied to the unbounded variable  $\Lambda$ , *i.e.* without refolding to the reference interval  $[0, 2\pi]$ . The infinite-time limit was replaced in practice by a linear least-squares fit over the time range 0 to 0.15 ns.

All the graphs presented in this chapter were generated with Python ([www.python.org](http://www.python.org)) and the Matplotlib library.<sup>?</sup>

## 2.4 RESULTS AND DISCUSSION

### 2.4.1 REFERENCE TI CALCULATIONS

The curve for the average Hamiltonian derivative as a function of  $\lambda$  corresponding to the aqueous methanol-to-dummy mutation is shown in Fig. 3.2a (blue curve, scale on the left), as averaged over the ten repeats of the 100 ns TI calculation with 129 equispaced  $\lambda$ -points. The corresponding running integral, *i.e.* the free-energy profile  $G(\lambda)$ , is also shown (orange curve, scale on the right). In view of the large amount of statistics (total 1  $\mu$ s sampling, *i.e.* clearly an overkill for such a calculation), the curve is perfectly smooth and extremely well converged. The average Hamiltonian derivative is large and positive below about 0.5 and becomes smaller and negative thereafter, leading to a maximum in the free-energy profile. The shape of these curves is determined both

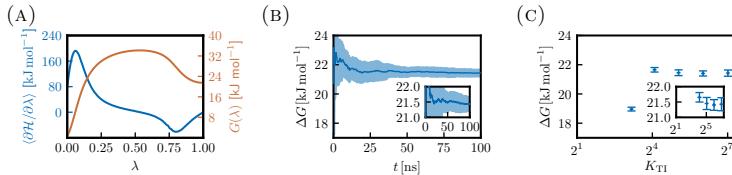


FIGURE 2.2: *Convergence properties of the TI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure provides the reference free-energy profile for the perturbation considered, and investigates the convergence properties of TI with different numbers  $K_{\text{TI}}$  of equidistant  $\lambda$ -points (based on ten repeats of 100 ns TI calculations). Panel (a) shows the average Hamiltonian derivative as a function of  $\lambda$  (blue curve, scale on the left) and the corresponding free-energy profile (orange, scale on the right). The curves are averaged over ten repeats of 100 ns TI calculations with 129 equidistant  $\lambda$ -points. Panel (b) shows the running  $\Delta G$  estimates for each of the ten repeats (dashed curves) along with the corresponding mean and error on the mean (95% confidence interval) over the repeats (thick blue curve and shaded area). Panel (c) shows  $\Delta G$  estimates considering ten repeats of 100 ns TI calculations using  $K_{\text{TI}} = 2^n + 1$  equidistant  $\lambda$ -points with  $n = 1, 2, \dots, 7$  (individual colored circles), along with the corresponding mean and error on the mean (95% confidence interval) over the repeats (blue crosses and error bars). Figs. equivalent to Panel (b) for the TI calculations with fewer  $\lambda$ -points are provided in Fig. 3.E.1. Numerical values at full sampling time are reported in Tab. 3.2 and Tab. 3.E.1.

by the hydration physics of methanol and by the choice of the employed (soft-core) coupling scheme.

The convergence properties of these TI calculations as a function of the total sampling time per repeat are illustrated in Fig. 3.2b, in the form of running  $\Delta G$  estimates for each of the ten repeats (dashed curves), along with the associated mean and error (95% confidence interval) on the mean over the repeats (thick blue curve and shaded area). The latter mean and error at full sampling time, reported in Tab. 3.2, evaluate to  $21.43 \pm 0.21 \text{ kJ mol}^{-1}$ . This result (with a minus sign) is in quantitative agreement with the experimental value for the hydration free energy of methanol at 298.15 K and 1 bar, namely?  $-21.4 \text{ kJ mol}^{-1}$ .

Corresponding graphs for TI calculations involving different numbers of equidistant  $\lambda$ -points ( $K_{\text{TI}} = 2^n + 1$  with  $n = 1, 2, \dots, 7$ ) at identical total sampling time (100 ns for each of the ten repeats) are shown in Fig. 3.E.1. The associated  $\Delta G$  means and errors at full sampling time are reported numerically in Tab. 3.2 (see Tab. 3.E.1 for the values of the individual repeats). The results of Tab. 3.2 also include a comparison of different error estimates (all excluding the Student  $t$ -factor), namely: (i) the standard deviation  $\sigma$  of the mean over the repeats; (ii) a bootstrap error  $\sigma_B$  calculated from the concatenated trajectories of the repeats at each  $\lambda$ -point; (iii) an alternative error  $\sigma_G$  obtained similarly by calculating the bootstrap errors on the average Hamiltonian derivatives at all  $\lambda$ -points, and propagating them assuming a Gaussian distribution of the estimates, as proposed in Refs. ???. The three values turn out to be very similar, so that further discussions will refer to the error  $\epsilon$  defined as  $\sigma$  amplified by the Student  $t$ -factor.

The effect of  $K_{\text{TI}}$  on the convergence properties is illustrated graphically in Fig. 3.2c, which shows the  $\Delta G$  estimates at full sampling time as a function of  $K_{\text{TI}}$  for each of the ten repeats (individual colored circles), along with the associated mean and error on the mean over the repeats (blue crosses and error bars). Using too few  $\lambda$ -points leads to quadrature errors, which are very significant for  $n \leq 4$  ( $K_{\text{TI}} \leq 17$ ), *i.e.* the shape of the average Hamiltonian derivative curve is not well captured by the Simpson quadrature. For  $n \geq 5$  ( $K_{\text{TI}} \geq 33$ ), the quadrature error becomes essentially negligible compared to the sampling error.

Interestingly, at constant total sampling time, the latter error appears to be essentially insensitive to the number of  $\lambda$ -points employed, with very similar  $\Delta G$  estimates and error bars of  $21.45 \pm 0.20$ ,  $21.40 \pm 0.17$  and  $21.43 \pm 0.21$  kJ mol<sup>-1</sup> for  $K_{\text{TI}} = 33, 65$  and

TABLE 2.2: *Repeat statistics over TI and CBTI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This table investigates the convergence properties of the calculated free-energy change  $\Delta G$  considering TI calculations with different numbers of  $\lambda$ -points along with a CBTI simulation with 16 replicas (average and various error estimates, based on ten calculation repeats involving each at a constant total single-system sampling time of 100 ns) The TI calculations involved  $K_{\text{TI}} = 2^n + 1$  equispaced  $\lambda$ -points with  $n = 1, 2, \dots, 7$ . The CBTI simulations involved  $K = 16$  replicas, along with  $m_{\Lambda} = 160 \text{ u nm}^2$  and  $\tau_{\Lambda} = 0.5 \text{ ps}$  (no biasing potential). In both cases, the total single-system sampling time was 100 ns, and the calculation was repeated ten times. Considering each set of repeats, the successive entries are: the free-energy change  $\Delta G_{\text{avg}}$  averaged over the repeats; the corresponding standard deviation  $\sigma$  of the mean; the associated error  $\epsilon$  on the mean; alternative error estimates  $\sigma_B$  and  $\sigma_G$  calculated by bootstrapping (no Student  $t$ -factor included). The quantity  $\sigma$  is calculated by scaling the standard deviation of the ten  $\Delta G$  values by the square-root of nine. The error estimate  $\epsilon$  is then calculated as  $\epsilon = 2.262\sigma$ , corresponding to a two-sigma confidence interval of 95%. The estimate  $\sigma_B$  corresponds to a bootstrap error calculated from the concatenated trajectories of the ten repeats (for TI, concatenated separately at each  $\lambda$ -point). The estimate  $\sigma_G$  (TI only) is obtained by calculating the bootstrap errors on the average Hamiltonian derivatives at all  $\lambda$ -point, and propagating them assuming a Gaussian distribution of the estimates, as proposed in Refs. [?]. The  $\Delta G$  values for the individual repeats can be found in Tabs. 3.E.1 and 2.G.1.

	TI							CBTI
$n$	1	2	3	4	5	6	7	-
$K_{\text{TI}}$ or $K$	3	5	9	17	33	65	129	16
$\Delta G_{\text{avg}}$ [kJ mol $^{-1}$ ]	15.61	4.81	18.97	21.65	21.45	21.40	21.43	21.39
$\epsilon$ [kJ mol $^{-1}$ ]	0.05	0.18	0.13	0.15	0.20	0.17	0.21	0.10
$\sigma$ [kJ mol $^{-1}$ ]	0.02	0.08	0.06	0.07	0.09	0.08	0.09	0.04
$\sigma_B$ [kJ mol $^{-1}$ ]	0.03	0.06	0.05	0.05	0.07	0.08	0.09	0.05
$\sigma_G$ [kJ mol $^{-1}$ ]	0.03	0.06	0.05	0.05	0.05	0.05	0.05	-

129, respectively (Tab. 3.2). Thus, for the system considered here and at constant total sampling time, the use of a larger number of  $\lambda$ -points with a reduced simulation time is advantageous as it reduces the quadrature error without incurring any penalty in terms of sampling error. However, it also induces an extra cost in the form of the preequilibration dead time, which is proportional to the number of  $\lambda$ -points employed. In addition, the insensitivity of the sampling error to the number of  $\lambda$ -points may still break down when the simulation times at the different  $\lambda$ -points become shorter than the corresponding orthogonal relaxation times.

The use of numerous  $\lambda$ -points with insufficient preequilibration is the plague affecting the slow-growth (SG) method,<sup>??</sup> which is rather inaccurate in practice<sup>???</sup> (large error and hysteresis), unless the results are exponentially averaged over repeats with different Boltzmann-distributed initial conditions, as in fast growth<sup>???</sup> (FG). In essence, the CBTI scheme aims at achieving this limit of a very large number of  $\lambda$ -points, but without suffering from issues related to the preequilibration dead time (TI), to an insufficient preequilibration (SG), or to the requirement of exponential averaging over initial conditions (FG).

#### 2.4.2 RESULTS USING THE CBTI SCHEME

The exploration of the influence of the CBTI mass parameter  $m_\Lambda$  and thermostat coupling time  $\tau_\Lambda$  considering different numbers  $K$  of replicas is presented in details in Sect. 3.B. The results are summarized in Tab. 3.1 (entries 1-10). They show that when selected within reasonable ranges, the parameters  $m_\Lambda$  and  $\tau_\Lambda$  have only a limited influence on the kinetic-energy exchange between the CB advance variable  $\Lambda$  and the conformational degrees of freedom, on the average temperature  $T_\Lambda$ , on the diffusion constant

$D_\Lambda$ , and on the calculated free-energy change  $\Delta G$ . Based on this exploration, a working choice  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  was selected for the following CBTI calculations.

Using this parameter setting, CBTI calculations were performed using  $K = 8, 16, 32, 64$  or  $128$  replicas along with  $256K^{-1} \text{ ns}$  simulation time for the replica system, *i.e.*  $256 \text{ ns}$  total single-system sampling time in all cases. The detailed results are shown graphically in Fig. 2.F.11 - Fig. 2.F.15 and summarized numerically in Tab. 3.1 (entries 11-15). Illustrative results for  $K = 8, 32$  or  $128$  (entries 11, 13 and 15) are also shown in Figs. 3.3 and 3.5, and discussed in details below.

The time series and probability distributions of  $\Lambda$  for these three simulations are displayed in Fig. 3.3a. For  $K = 8$  replicas (left panel),  $\Lambda$  covers the entire range  $[0, 2\pi]$ , but presents 8 regularly spaced probability peaks at integer multiples of  $2\pi/8$ . This corresponds to CB configurations where two of the replicas are at the physical end-states  $\lambda = 0$  and  $\lambda = 1$ , the six others being by pairs at  $\lambda = 0.25, 0.5$  and  $0.75$ . The occurrence of such probability peaks is not surprising considering the shape of the free-energy profile (Fig. 3.2a). This profile presents minima at  $\lambda = 0$  and  $\lambda = 1$ , the former one being the fully interacting methanol state, which is particularly favored. This significant inhomogeneity in the sampling of  $\Lambda$  correlates with a stepwise (hopping) dynamics in the  $\Lambda$  time series.

In contrast, for  $K = 128$  replicas (right panel),  $\Lambda$  only spans a limited subrange of the interval  $[0, 2\pi)$  and no longer presents regularly spaced probability peaks. For longer simulations, one would expect the probability curve to also cover the entire  $[0, 2\pi)$  range and present regularly spaced peaks at integer multiples of  $2\pi/128$ , but these would be of much smaller magnitudes compared to the  $K = 8$  case. For  $K = 128$ , the effect of these residual

variations becomes almost negligible, *i.e.* there is essentially no driving force on the  $\Lambda$  variable. As a result, instead of a hopping dynamics, the  $\Lambda$  variable now follows a purely diffusive (random-walk) trajectory.

The situation for  $K = 32$  replicas (middle panel) is intermediate. The  $[0, 2\pi)$  range is fully covered, but the distribution is irregular, presenting some small indentations at  $2\pi/32$  spacing.

The reduction of the coverage and homogeneity in terms of  $\Lambda$  over the interval  $[0, 2\pi)$  when going from  $K = 8$  to 32, and then to 128, is not due to a progressively slower (hopping *vs.* diffusive) dynamics, but simply to the different time spans of the simulations (32, 8 and 2 ns, respectively), imposed by the constant total single-system sampling time of 256 ns in the present simulations. This can be seen in Fig. 3.5.

Fig. 3.5a shows the time series and probability distribution of  $\dot{\Lambda}$ , the velocity associated with the  $\Lambda$ -variable. The Maxwell-Boltzmann distribution of Eq. 3.25 is also displayed for comparison. The observed width  $\sigma_{\dot{\Lambda}}$  of the distribution is reported numerically in Tab. 3.1. As expected from Eq. 3.25, it is proportional to  $m_{\Lambda}^{-1/2}$  which, in the present simulations, was itself made proportional to  $K^{1/2}$  (an arbitrary parameter choice justified in Sect. 3.B). Thus, the average magnitude of  $\dot{\Lambda}$  decreases somewhat when going from  $K = 8$  to 32, and then to 128 (scaling factor  $2^{-1/2} \approx 0.7$  each time). Fig. 3.4 shows the normalized autocorrelation functions  $c_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  for different  $m_{\Lambda}$  in the absence of thermostat coupling (entries 1-5 of Tab. 3.1). The corresponding correlation times  $\tau_{\dot{\Lambda}}$  are reported numerically in Tab. 3.1. Here, one observes that the correlation time increases with  $m_{\Lambda}$  as a result of the higher inertia of the CB. In addition, for the two lowest values of  $m_{\Lambda}$ , the initial decay of  $c_{\dot{\Lambda}}$  to negative values suggests that the velocities are frequently reversed at short time, when barriers cannot be

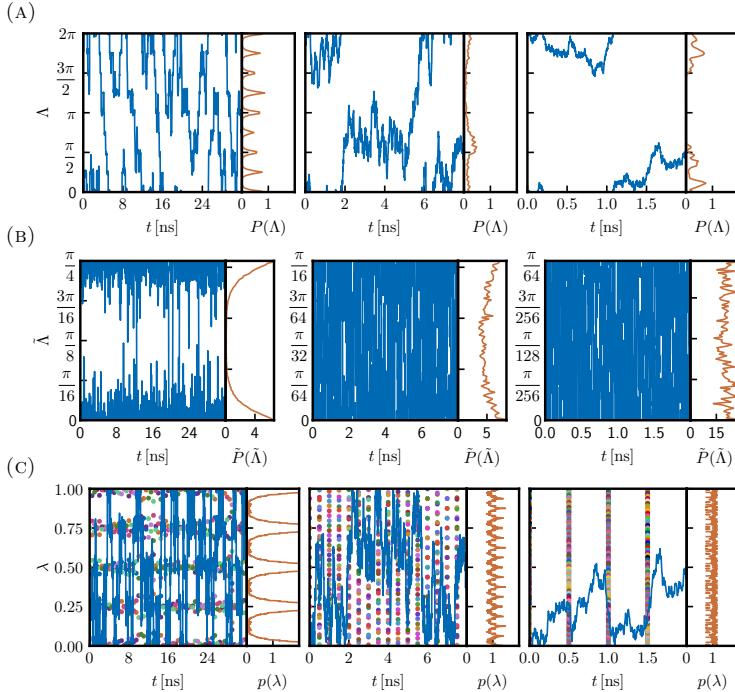


FIGURE 2.3: Time series and probability distributions of the relevant variables in unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. This figure compares the dynamics and distributions of the CB variables  $\Lambda$  (a),  $\tilde{\Lambda}$  (b) and  $\lambda$  (c) in CBTI simulations with  $K = 8$  (left), 32 (middle) or 128 (right) replicas (based on simulations with a total single-system sampling time of 256 ns). Panel (a) shows the time series  $\Lambda(t)$  and probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$ . Panel (b) shows the time series  $\tilde{\Lambda}(t)$  and probability distribution  $\tilde{P}(\tilde{\Lambda})$  of the CB fractional advance variable  $\tilde{\Lambda}$  (Eq. 3.17). Panel (c) shows the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas (Eq. 3.19). In Panel (c), the time series is shown as a blue curve for one reference replica  $k = 0$ , and as individual colored points at 0.5 ns interval for the  $K - 1$  other replicas, and the probability distribution  $p(\lambda)$  is calculated considering all replicas according to Eq. 3.19. All probability distributions are normalized to one. The CBTI simulations relied on  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  (no biasing potential). See also Tab. 3.1 entries 11, 13 and 15 for relevant numerical results. Corresponding graphs also including  $K = 16$  and 64 can be found in Figs. 2.F.11 - 2.F.15.

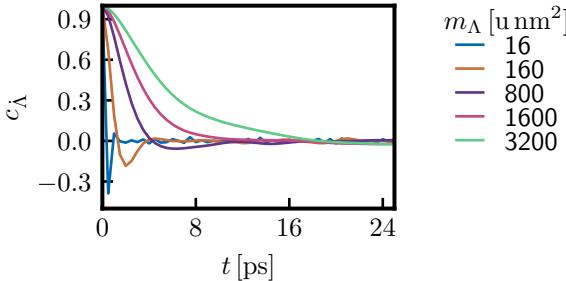


FIGURE 2.4: *Velocity autocorrelation functions of the CB advance variable in unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure shows the increase in the autocorrelation time of the CB velocity upon increasing the mass parameter in CBTI simulations with  $K = 16$  replicas and no thermostat coupling. The normalized autocorrelation function  $c_{\Lambda}$  of the velocity  $\dot{\Lambda}$  is shown for different values of the mass-parameter  $m_{\Lambda}$ , in the absence of thermostating of the  $\Lambda$ -variable ( $\tau_{\Lambda} \rightarrow \infty$ ). See also Tab. 3.1 entries 1-5 for relevant numerical results, including the corresponding autocorrelation times  $\tau_{\dot{\Lambda}}$ . Corresponding graphs for all simulations listed in Tab. 3.1 are shown in Figs. 3.F.1 - 2.F.15

crossed. Note, however, that this connection between  $m_{\Lambda}$  and  $\tau_{\dot{\Lambda}}$  is somewhat altered by the thermostat coupling in the other simulations (including those of entries 11-15 in Tab. 3.1).

Because the two above trends act in opposite directions, the actual rate of diffusion along  $\Lambda$  is not very different in the three simulations.

This is seen in Fig. 3.5c, which shows the mean-square displacement  $d_{\Lambda}$  of  $\Lambda$  as a function of time. The curves are essentially linear, as expected from the Einstein model (Eq. 3.26). The corresponding diffusion constants  $D_{\Lambda}$  are reported in Tab. 3.1, namely 4.4, 6.2 and  $2.3 \text{ ns}^{-1}$  for  $K = 8, 32$  and 128 replicas, respectively, *i.e.* not dramatically different. A similar observation holds for the probability distribution of  $\ddot{\Lambda}$ , the acceleration associated with the  $\Lambda$ -variable, shown in Fig. 3.5b. The choice of taking  $m_{\Lambda}$

proportional to  $K^{1/2}$  was made here precisely to achieve a CB acceleration that is essentially independent of  $K$  (see discussion in Sect. 3.A).

The evolution of  $D_\Lambda$  with  $K$  is clearly non-monotonic (see values for entries 11-15 in Tab. 3.1), which results in large part from the two opposite effects mentioned above concerning the effect of  $m_\Lambda$  on  $\sigma_{\tilde{\Lambda}}$  and  $\tau_{\tilde{\Lambda}}$ . On the one hand,  $\sigma_{\tilde{\Lambda}}$  (*i.e.* the average magnitude of the CB velocity) decreases upon increasing  $m_\Lambda$ , and the dynamics along the  $\Lambda$ -variable becomes intrinsically slower. On the other hand,  $\tau_{\tilde{\Lambda}}$  (*i.e.* the dynamical inertia of the CB) increases upon increasing  $m_\Lambda$ , and the ability to cross residual barriers in the  $G_\Lambda(\Lambda)$  free-energy profile is enhanced. The magnitude of these barriers itself decreases upon increasing  $K$ . In addition, inertia along  $\Lambda$  is also affected by the thermostat coupling. The combination of these effects is complex and the resulting trends in  $D_\Lambda$  upon varying  $K$  would be difficult to rationalize in details. Most importantly, however, they appear not to be extremely pronounced.

The time series and probability distributions of the CB fractional advance variable  $\tilde{\Lambda}$  are shown in Fig. 3.3b. In contrast to the distributions of  $\Lambda$  over the range  $[0, 2\pi]$  (Fig. 3.3a), the distributions of  $\tilde{\Lambda}$  over the range  $[0, \Delta\Lambda]$  are converged and cover the entire range, also for  $K = 128$ . For  $K = 8$ , the distribution of  $\tilde{\Lambda}$  is significantly biased towards 0 and  $2\pi/8$ , *i.e.* CB configurations with two of the replicas at the physical end-states. For  $K = 32$ , it is still somewhat biased towards 0 and  $2\pi/32$ , which corresponds to the same situation. And for  $K = 128$ , the probability distribution becomes very close to homogeneous.

Finally, the time series and probability distributions of the coupling variable  $\lambda$  are shown in Fig. 3.3c for all replicas. The time series is shown for one reference replica  $k = 0$  (blue curve)

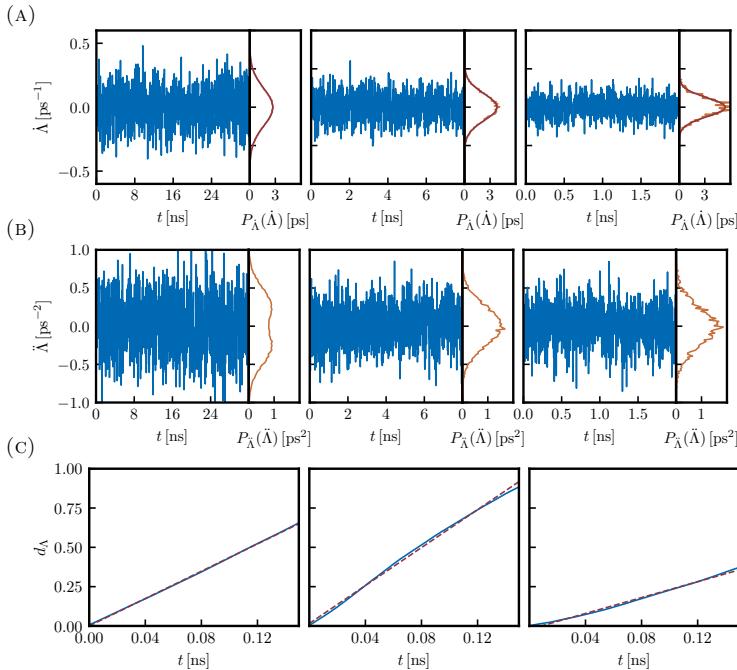


FIGURE 2.5: *Dynamical properties of the CB advance variable in unbiased CBTI calculations of the aqueous methanol-to-dummy mutation at  $298.15\text{ K}$  and  $1\text{ bar}$  with the 2016H66 force field.* This figure compares the dynamics and distributions of the CB variables  $\Delta$  (a) and  $\Delta\ddot{\Delta}$  (b), as well as the mean-square displacement  $d_\Delta$  along  $\Delta$  (c), in CBTI simulations with  $K = 8$  (left),  $32$  (middle) or  $128$  (right) replicas (based on simulations with a total single-system sampling time of  $256\text{ ns}$ ). Panel (a) shows the time series  $\Delta(t)$  and probability distribution  $P_{\Delta}(\Delta)$  of the velocity associated with the CB advance variable  $\Delta$ . Panel (b) shows the time series  $\Delta\ddot{\Delta}(t)$  and probability distribution  $P_{\Delta\ddot{\Delta}}(\Delta\ddot{\Delta})$  of the corresponding acceleration. Panel (c) shows the mean-square displacement  $d_\Delta(t)$  of  $\Delta$  (Eq. 3.26). In Panel (a), the probability distribution from the simulation (blue curve) is shown together with the analytical Maxwell-Boltzmann distribution  $P_{\Delta}^{\text{MB}}(\Delta)$  of Eq. 3.25 (orange curve). In Panel (c), the mean-square displacement  $d_\Delta$  from the simulation (blue curve) is shown along with a linear least-squares fit over the interval  $0 - 0.15\text{ ns}$  (brown dashed line). All probability distributions are normalized to one. The CBTI simulations relied on  $m_\Delta = 40K^{1/2}\text{ u nm}^2$  and  $\tau_\Delta = 0.5\text{ ps}$  (no biasing potential). See also Tab. 3.1, entries 11, 13 and 15 for relevant numerical results. Corresponding graphs also including  $K = 16$  and  $64$  can be found in Figs. 2.F.11 - 2.F.15.

as well as the other replicas  $k = 1 \dots K - 1$  (individual colored points at 0.5 ns interval). Compared to the reference replica, these representative points are merely shifted by  $2K^{-1}k$ , refolded by periodicity into the range  $[0, 2)$ , and the points in the range  $[1, 2)$  mirrored into the range  $[0, 1]$  according to Eqs. 3.6- 3.8. The probability distribution  $p(\lambda)$  is calculated according to Eq. 3.19, *i.e.* considering all replicas. Although the  $\Lambda$ -space in the range  $[0, 2\pi)$  is sampled only partially given the simulation lengths (Fig. 3.3a), the  $\lambda$ -space in the range  $[0, 1]$  (Fig. 3.3c), just like the  $\tilde{\Lambda}$ -space in the range  $[0, \Delta\Lambda)$  (Fig. 3.3b), is completely sampled in all cases. However, for  $K = 8$ , the barriers are still too high and the sampling is considerably biased towards  $\lambda$ -values that are integer multiples of  $1/4$ . The corresponding bias towards integer multiples of  $1/16$  is far less pronounced for  $K = 32$ , and the simulation with  $K = 128$  is very close to a homogeneous sampling of  $\lambda$ .

In summary, the sampling of  $\Lambda$  only needs to cover a  $\Delta\Lambda$  range to ensure the sampling of all possible  $\tilde{\Lambda}$ -values by the CB and thus, of all possible  $\lambda$ -values by at least one replica. However, the homogeneity of the sampling in  $\lambda$  depends crucially on  $K$ , quasi-homogeneous sampling requiring a large number of replicas.

This can be seen in Fig. 3.6, which shows in a logarithmic form the height  $G_{\Lambda}^*$  of the peaks in  $G_{\Lambda}(\Lambda)$  as a function of the number  $K$  of replicas (see also Fig. 2.I.1 for the individual curves). The dependency is approximately linear with a slope of minus one, suggesting that  $G_{\Lambda}^*$  scales as  $K^{-1}$ . Such a scaling is expected based on the considerations made in Sect. 3.A. Note that the proportionality constant depends on the derivatives of  $G(\lambda)$  at the two physical end-states. In a situation where these derivatives vanished, a scaling as  $K^{-2}$  would be expected instead.

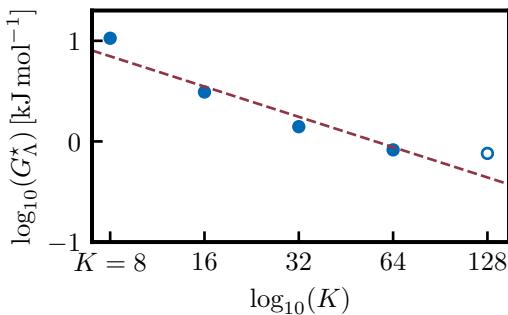


FIGURE 2.6: Height  $G_\Lambda^*$  of the residual barriers in the free-energy profile  $G_\Lambda(\Lambda)$  for unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. This figure shows the progressive decrease in the height of the barriers along  $\Lambda$  upon increasing the number  $K$  of replicas in a CBTI simulation. The data is represented in logarithmic form and a linear line of slope  $-1$  is fitted to the filled circles (intercept 1.75 at  $\log_{10}(K) = 0$ ). The open circle was omitted from the fit because  $G_\Lambda^*$  cannot be determined sufficiently accurately (within the noise). The simulations considered rely on  $K = 8, 16, 32, 64$  or  $128$  replicas, along with  $m_\Lambda = 40K^{1/2}$  u nm $^2$  and  $\tau_\Lambda = 0.5$  ps (no biasing potential). The corresponding free-energy profiles  $G_\Lambda(\Lambda)$  are shown graphically in Fig. 2.I.1.

### 2.4.3 CBTI WITH BIASING POTENTIAL

As seen above, simulations employing only few replicas may still be affected by a significant sampling inhomogeneity along  $\lambda$  due to high residual barriers along  $\Lambda$ . To prevent the undersampling of specific  $\lambda$ -ranges in this case, a memory-based biasing potential can be employed. This renders the CBTI method more robust since it becomes insensitive to the chosen number of replicas, and more compatible with possible constraints related to *e.g.* the number of cores in a specific computing node during the calculation. As discussed in the Theory section, it is sufficient to define such a biasing potential over a  $[0, \Delta\Lambda/2]$  range of  $\tilde{\Lambda}$ , and its magnitude is limited by the already largely homogeneous sampling of  $\tilde{\Lambda}$ . These features speed up the build-up procedure drastically, leading to a comparatively small simulation time for the LE build-up phase relative to single-system  $\lambda$ -LEUS.

Fig. 3.7 illustrates the sampling in a biased compared to that in the unbiased CBTI simulation with  $K = 8$  replicas. The biasing relied on 34 basis functions over half the  $\tilde{\Lambda}$ -range  $[0, \pi/4]$  (17 free coefficients) and a build-up phase of 0.15 ns (*i.e.* only 1.2 ns total single-system build-up time). During this time, the replica system performed three double sweeps over all grid-points, leading to a biasing potential with an energetic resolution of  $10^{-6} \text{ kJ mol}^{-1}$  (Figs. 3.7a and 3.7b).

The  $\Lambda$ -sampling in the biased CBTI simulation is much closer to uniform compared to that in the unbiased one (Fig. 3.7c *vs.* 3.7d), *i.e.* the relative probabilities between most and least sampled  $\Lambda$ -values has been reduced from 70 to 4.6. The time series of  $\lambda$  and the sampling distribution  $p(\lambda)$  show that the motion of the conveyor belt was hindered in the unbiased simulation, which it is no longer the case in the biased one (Fig. 3.7e *vs.* 3.7f).

Alternative build-up protocols were also explored for  $K = 8$  and  $K = 16$ , differing in the number  $N_{\text{gp}}$  of grid points, in the duration  $t_{\text{LE}}$  of the build-up phase, and in the build-up parameters  $c_{\text{LE}}$  and  $f_{\text{red}}$ . The corresponding results are shown graphically in Figs. 2.H.2 and 2.H.4 and summarized numerically in Tab. 2.H.1. All the tested protocols reduce the sampling inhomogeneity to an acceptable level, although residual inhomogeneities are difficult to remove entirely. Compared to the TI results with a sufficient number of  $\lambda$ -points (Tab. 3.2, entries with  $K_{\text{TI}} \geq 33$ ) and the unbiased CBTI results with a sufficient number of replicas (Tab. 3.1, entries 12-15 with  $K \geq 16$ ), the biased CBTI results with  $K = 8$  or 16 (entries 16 and 17 in Tab. 3.1) deliver similar  $\Delta G$  values and error bars, provided that an appropriate build-up protocol is employed.

#### 2.4.4 CBTI FREE-ENERGY ESTIMATOR

The calculation of free-energy differences from CBTI simulations according to Eq. 3.22 relies on the choice of a number  $J$  of bins. This number can be taken very large, thereby reducing the (rectangular) quadrature error to an essentially negligible amount, but should not exceed a threshold  $J_{\max}$  where empty bins occur. The influence of  $J$  is illustrated in Fig. 3.8, which shows the average Hamiltonian derivative curve obtained from the unbiased CBTI simulation with  $K = 32$  replicas (entry 13 in Tab. 3.1, restricted to total 100 ns single-system sampling time) for  $J = 10$  (Fig. 3.8a),  $J = 500$  (Fig. 3.8b) and  $J = 5000$  (Fig. 3.8c). A too low number of bins (*e.g.*  $J = 10$ ) will lead to an insufficient number of integration points, thereby increasing the quadrature error. A too high number of bins will lead to insufficient sampling of each bin, thereby increasing the statistical error. The optimal value of

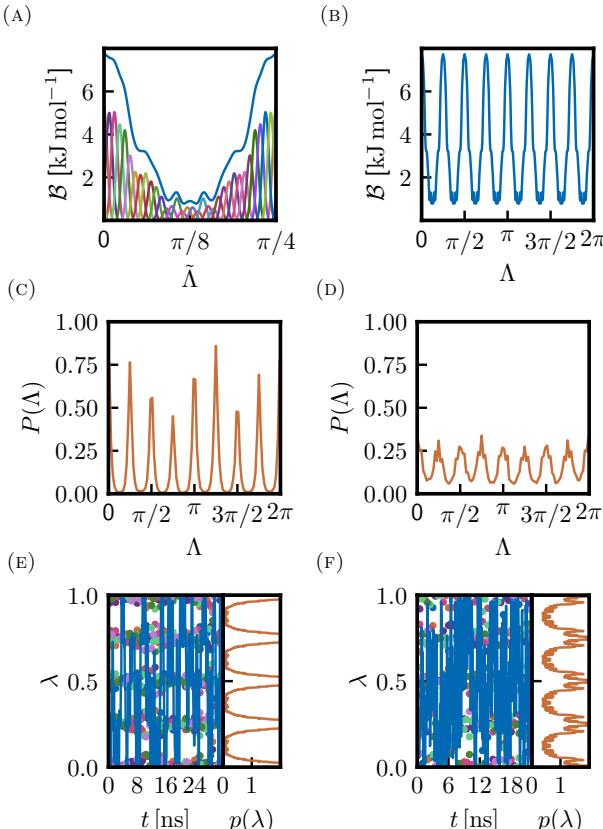


FIGURE 2.7: *Biassing potential and simulation results from unbiased vs. biased CBTI calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure illustrates how the use of a biassing potential permits to drastically reduce the sampling inhomogeneity of CBTI simulations performed with a small number of replicas (here,  $K = 8$  replicas). Panels (a) and (b) show the biassing potential along the  $\tilde{\Lambda}$ - and  $\Lambda$ -ranges, respectively. Panels (c) and (d) show the probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$  in the unbiased and biased simulation, respectively. Panels (e) and (f) show the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas in the unbiased and biased simulation, respectively. In Panel (a), the symmetric biassing potential (blue curve) is represented using 34 basis functions (curves of different colors) involving 17 free coefficients. In Panels (e) and (f), the time series is shown as a blue curve for one reference replica  $k = 0$ , and as single colored points at 0.5 ns spacing for the  $K - 1$  other replicas, and the probability distribution  $p(\lambda)$  is calculated considering all replicas according to Eq. 3.19. All probability distributions are normalized to one. The CBTI simulations were performed with or without biassing potential, using  $m_\Lambda = 113 \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$ . For the biased simulation, the biassing potential was constructed in a build-up phase of 0.15 ns duration. Results for biased simulations with  $K = 16$  and alternative build-up protocols are shown in Fig. 2.H.2 and Fig. 2.H.4, the numerical results being reported in Tab. 2.H.1.

$J$  will depend on the total sampling time and on the degree of homogeneity of the sampling along  $\lambda$ .

To investigate the influence of  $J$  on the calculated  $\Delta G$ , this number was systematically increased from 1 to  $J_{\max}$  in increments ranging from 10 to 1000. The results are shown in Fig. 3.8d, considering the unbiased CBTI simulations with  $K = 8, 16, 32, 64$  or 128 (entries 11-15 in Tab. 3.1, restricted to 100 ns total single-system sampling time) as well as the biased CBTI simulations with  $K = 8$  or 16 (entries 16 and 17 in Tab. 3.1, same restriction). For the unbiased simulation with  $K = 8$  (blue curve), the convergence of  $\Delta G$  upon increasing  $J$  is irregular and the value of  $J_{\max}$  is comparatively low (about 600), due to the inhomogeneous sampling along  $\lambda$ . All other curves present qualitatively similar convergence features. For a low  $J$  ( $J < 100$ ), the quadrature error leads to strong variations. For intermediate  $J$  ( $100 \leq J \leq 3000$ ), the curve presents a plateau, *i.e.* the  $\Delta G$  estimate becomes insensitive to changes in  $J$ . For larger  $J$  ( $3000 < J \leq J_{\max}$ , where  $J_{\max}$  ranges from about 5000 to 8000) the curves become somewhat noisy again due to undersampled bins. These results suggest that, as a rule of thumb, the value of  $J$  for the application of Eq. 3.22 should be selected between at least about 100 and at most about  $J_{\max}/2$ . For most of the  $\Delta G$  results presented in this chapter, a default value  $J = 500$  was selected as leading to a reasonably smooth Hamiltonian derivative curve (Fig. 3.8b) while keeping the quadrature error negligible (except for the CBTI simulation with  $K = 8$ , where  $J = 200$  was used). The results for the variants  $\Delta G_{\text{alt}}$  and  $\Delta G_{\text{app}}$  to  $\Delta G$  from Eq. 3.22 which do not require the specification of a number of bins (Eqs. 3.C.1 and 3.C.2) are also included in Tab. 3.1 for comparison. They are discussed in Sect. 3.C.

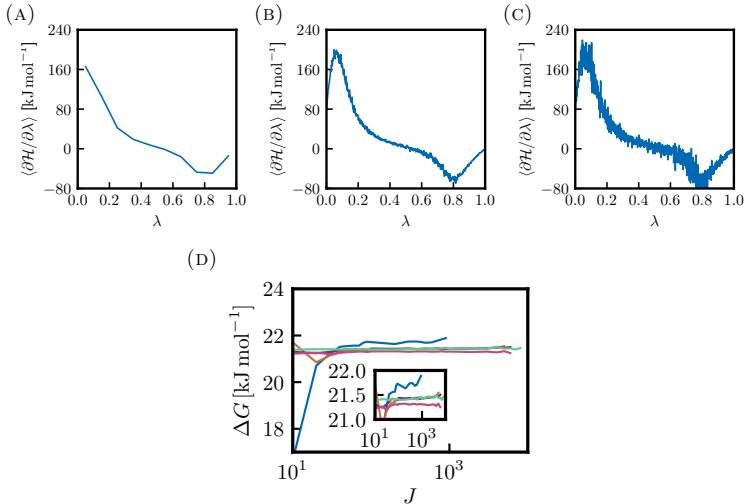


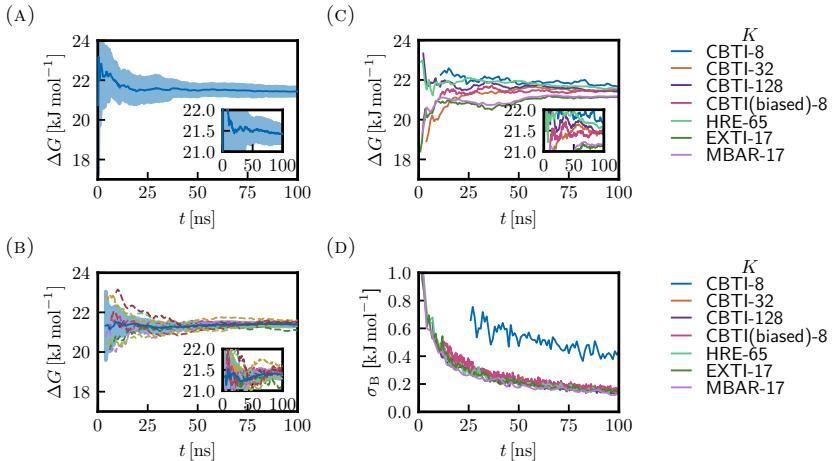
FIGURE 2.8: *Influence of the number of bins used in the free-energy estimator for CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure illustrates how the number  $J$  of bins employed in the estimator affects the quadrature and sampling quality (top) and the resulting estimated  $\Delta G$  (bottom). Panels (a), (b) and (c) show the average Hamiltonian derivative curve constructed with different numbers  $J$  of bins, namely  $J = 10$  (a), 100 (b) or 5000 (c). The curves are based on the CBTI simulation using  $K = 32$  replicas (entry 13 in Tab. 3.1). Panel (d) shows the  $\Delta G$  estimate from Eq. 3.22 for CBTI simulations with  $K = 8, 16, 32, 64$  or 128 replicas (the two former either unbiased or biased; entries 11-17 in Tab. 3.1) as a function of  $J$  from 1 to  $J_{\max}$ , where  $J_{\max}$  corresponds to the first occurrence of empty bins. All CBTI simulations were performed using  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$ , and their analysis is restricted here to 100 ns total single-system sampling time.

### 2.4.5 CBTI CONVERGENCE PROPERTIES

The convergence properties of TI and CBTI are compared in Figs. 3.9a and 3.9b, which shows the  $\Delta G$  values calculated from ten simulation repeats (individual dashed curves) along with the corresponding mean and error (95% confidence interval) on the mean (thick blue curve and shaded area) as a function of the total single-system sampling time per repeat. The simulation protocols compared are TI with  $K_{TI} = 129$   $\lambda$ -points (Fig. 3.9a, which is identical to Fig. 3.2b) and unbiased CBTI with  $K = 16$  (Fig. 3.9b). Note that in the latter case, the curves do not start at time zero because the application of Eq. 3.22 with  $J = 500$  to estimate  $\Delta G$  only becomes possible when  $J_{\max}$  exceeds 500 (*i.e.* each of the 500 bins encompasses at least one sampled  $\lambda$ -value). The mean  $\Delta G$  values over the repeats and the associated error bars at full single-system sampling time (100 ns per repeat) are reported in Tab. 3.2 (last two columns; see also Tab. 3.E.1 and Tab. 2.G.1 for the results of the individual repeats).

The comparison clearly evidences an enhanced convergence for CBTI at identical sampling time, both in terms of the running averages and of the final values. These final values are  $21.43 \pm 0.21$  kJ mol<sup>-1</sup> for TI, compared to  $21.39 \pm 0.10$  kJ mol<sup>-1</sup> for CBTI, *i.e.* identical within error bars, but with an error reduced by about a factor two for CBTI. This improved convergence is likely due to the improved orthogonal sampling, *i.e.* the diffusion of the systems along  $\lambda$  permits to circumvent orthogonal barriers present at specific  $\lambda$ -values.

The same observation holds for the CBTI simulations involving different numbers of replicas or/and a biasing potential. This is shown in Figs. 3.9c and 3.9d for  $K = 8, 32$  or  $128$  (unbiased) and for  $K = 8$  (biased). Here, a single simulation was performed



**FIGURE 2.9:** *Convergence properties of the CBTI calculations compared to other methods for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure compares the convergence properties of the calculated free-energy change  $\Delta G$  considering TI (a) vs. CBTI (b) based on ten repeats of calculations involving 100 ns total single-system sampling time, or considering CBTI (biased or unbiased) vs. TI, HRE, TI/EXTI and TI/MBAR in terms of estimated  $\Delta G$  (c) and error (d) based on one calculation involving 100 ns total single-system sampling time. Panel (a) is identical to Fig. 3.2b and shows the convergence of  $\Delta G$  for TI calculations with  $K_{\text{TI}} = 129$   $\lambda$ -points. Panel (b) displays the corresponding convergence of  $\Delta G$  for unbiased CBTI calculations with  $K = 16$  replicas. Panels (a) and (b) show the running  $\Delta G$  estimates for each of the ten repeats (dashed curves) as a function of the total sampling time  $t$  per repeat, along with the corresponding mean and error on the mean (95% confidence interval) over the repeats (thick blue curve and shaded area). The final  $\Delta G$  values for the individual repeats are reported in the Tab. 2.G.1, and the statistics over the repeats can be found in Tab. 3.2. Panel (c) compares convergence properties in terms of  $\Delta G$  based on single simulations. Panel (d) displays the associated errors  $\sigma_B$  calculated by bootstrapping (no Student  $t$ -factor included). Panels (c) and (d) consider unbiased CBTI ( $K = 8, 32$  or  $128$ ), biased CBTI ( $K = 8$ ) HRE ( $K_{\text{HRE}} = 65$ ), along with TI with EXTI or MBAR estimators ( $K_{\text{TI}} = 17$  and  $129$  virtual  $\lambda$ -points in both cases). The corresponding final  $\Delta G$  estimates are reported in Tab. 3.3.

TABLE 2.3: *Free-energy estimates from the CBTI calculations compared to other methods for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This table compares the convergence properties of the calculated free-energy change  $\Delta G$  considering CBTI (biased or unbiased), TI, HRE, EXTI and MBAR calculations with different numbers of replicas (all at a constant total single-system sampling time of 100 ns). The CBTI calculations involved  $K$  replicas, along with  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  (with or without biasing potential). The TI calculations involved  $K_{\text{TI}}$  equispaced  $\lambda$ -points. The HRE calculations involve  $K_{\text{HRE}}$  equispaced  $\lambda$ -points, along with  $\tau_{\text{HRE}} = 0.2 \text{ ps}$ . For EXTI and MBAR, the calculations relied on  $K_{\text{TI}}$  equispaced real  $\lambda$ -points, and consideration of 129 equispaced virtual  $\lambda$ -points. In all cases, the total single-system sampling time was 100 ns. The corresponding convergence properties are illustrated graphically in Fig. 3.9. Additional numerical results regarding the CBTI simulations are reported in Tab. 3.1, entries 11–15 (unbiased CBTI) and entries 16–17 (biased CBTI).

$K, K_{\text{TI}} - 1$ or $K_{\text{HRE}} - 1$	$\Delta G [\text{kJ mol}^{-1}]$					
	CBTI (unbiased)	CBTI (biased)	TI	HRE	EXTI	MBAR
8	21.69 (0.40)	21.48 (0.16)	19.22 (0.18)	-	21.24 (0.37)	21.37 (0.15)
16	21.42 (0.16)	21.30 (0.13)	21.87 (0.17)	21.52 (0.16)	21.12 (0.14)	21.17 (0.13)
32	21.44 (0.14)	-	21.48 (0.20)	21.39 (0.15)	-	-
64	21.32 (0.14)	-	21.18 (0.24)	21.47 (0.13)	-	-
128	21.43 (0.12)	-	21.45 (0.30)	-	-	-

(no repeats), and the running  $\Delta G$  value (Fig. 3.9c) is displayed along with the corresponding bootstrap error estimate (Fig. 3.9d; no Student  $t$ -factor included) as a function of the total single-system sampling time. The  $\Delta G$  estimates and bootstrap error bars at full single-system sampling time (100 ns) are reported in Tab. 3.3 (second and third columns). Except for the unbiased CBTI simulation with  $K = 8$ , where the sampling along  $\lambda$  is still heterogeneous, all the CBTI protocols follow very similar error curves upon increasing the sampling time (Fig. 3.9d), with final errors of  $0.12 - 0.16 \text{ kJ mol}^{-1}$ . In contrast, the convergence of TI is significantly slower, with corresponding final errors of  $0.17 - 0.30 \text{ kJ mol}^{-1}$  (fourth column in Tab. 3.3).

#### 2.4.6 COMPARISON OF CBTI WITH OTHER METHODS

The  $\Delta G$  values and associated bootstrap errors (no Student  $t$ -factor included) calculated using various CBTI protocols are compared to the results of other approaches in Tab. 3.3. The comparison involves calculations relying on 100 ns total single-system sampling time, *i.e.* essentially identical computational costs. The protocols compared are unbiased CBTI (with  $K = 8, 16, 32, 64$  or 128), biased CBTI (with  $K = 8$  or 16), TI (with  $K_{\text{TI}} = 9, 17, 33, 65$  or 129), HRE (with  $K_{\text{HRE}} = 17, 33$  or 65), along with TI/EXTI and TI/MBAR (both with  $K_{\text{TI}} = 9$  or 17, and using 129 virtual  $\lambda$ -points). More details on the HRE, TI/EXTI and TI/MBAR calculations can be found in Figs. 2.J.1 - 2.L.1.

As previously discussed, the results of the TI calculations with  $K_{\text{TI}} = 9$  or 17 are affected by large quadrature errors (see Fig. 3.2c and Tab. 3.2), and those of the unbiased CBTI simulation with  $K = 8$ , are affected by a significant heterogeneity of the sampling along  $\lambda$  (see Fig. 3.3). Excepting these three calculations, all free-energy estimates of Tab. 3.3 are essentially compatible, taking into account that the error bars exclude a Student  $t$ -factor.

The convergence of the HRE ( $K_{\text{HRE}} = 65$ ), TI/EXTI ( $K_{\text{TI}} = 17$ ) and TI/MBAR ( $K_{\text{TI}} = 17$ ) protocols are also compared to those of TI ( $K_{\text{TI}} = 129$ ), unbiased CBTI ( $K = 8, 32$  or 128) and biased CBTI ( $K = 8$ ) in Figs. 3.9c and 3.9d. Considering Fig. 3.9d, the uncertainty of the unbiased CBTI simulation with  $K = 8$  is found to be particularly large, owing to the sampling heterogeneity. Excepting this simulation, the uncertainty is the largest for the TI estimate. The different CBTI protocols reduce the error by about a factor two compared to TI, presumably due to enhanced orthogonal sampling. This sampling advantage is shared

by HRE, for which the error curve is very similar. Advanced free-energy estimators such as EXTI and MBAR also improve the convergence over plain TI in a different way, namely by increasing the statistical efficiency of the  $\Delta G$  estimation. The error reduction is also about a factor two compared to TI, which is probably not entirely coincidental. If two fixed- $\lambda$  simulations are trapped in different orthogonal configurational wells in TI, the variable- $\lambda$  approaches (CBTI, HRE) will promote well-transitions, whereas the advanced-estimator approaches (EXTI, MBAR) will import statistical information on the other well from the neighboring  $\lambda$ -point. One might refer to these two types of effects orthogonal-sampling *vs.* orthogonal-statistics advantages, respectively. For the simple system considered here, both results in comparable convergence improvements. This suggests that the two effects may not be cumulative if one generalizes the current CBTI estimator of Eq. 3.22 to an EXTI- or MBAR-type estimator.

## 2.5 CONCLUSION

In the present chapter, we proposed a new method called conveyor belt thermodynamic integration (CBTI) to calculate alchemical free-energy differences based on MD simulations. This approach borrows and combines ideas from thermodynamic integration<sup>???</sup> (TI), replica exchange<sup>???</sup> (HRE) or permutation<sup>???</sup> (HRP), and  $\lambda$ -dynamics<sup>???</sup> ( $\lambda$ D), along with the real-life working principle of the funicular.

In CBTI, one simulates in parallel a set of  $K$  equally spaced replicas (with  $K$  even) on a forward-turn-backward-turn path along the alchemical coupling variable  $\lambda$ , akin to a conveyor belt (CB) between the two physical end states. Because the  $\lambda$ -forces (Hamiltonian  $\lambda$ -derivative) exerted by the individual replicas on the CB largely compensate each other, the overall  $\Lambda$ -force on the CB advance variable  $\Lambda$  becomes increasingly small when  $K$  is made increasingly large (residual free-energy barriers decreasing at least as  $K^{-1}$  in the limit of large  $K$ , as shown in Sect. 3.A). As a result, for a sufficient number  $K$ , quasi-homogeneous sampling of the  $\lambda$ -range can be achieved without application of any biasing potential. If a smaller  $K$  is employed, a memory-based biasing potential can still be added to further homogenize the sampling, the preoptimization of which is computationally inexpensive. The results of a CBTI simulation (whether biased or not) can be analyzed similarly to TI, by binning of the average Hamiltonian  $\lambda$ -derivative as a function of  $\lambda$  considering all replicas jointly, followed by quadrature integration. In this case, the continuous and quasi-homogeneous sampling of the  $\lambda$ -range permits to use a large number of bins, thereby essentially eliminating quadrature errors.

As a first application, the CBTI scheme was employed here to calculate the hydration free energy of methanol. It was shown that the method is rather robust with respect to the choice of its parameters ( $K$  as well as the mass-parameter  $m_\Lambda$  and thermostat coupling time  $\tau_\Lambda$  of the CB), the most important sensitivity being relative to  $K$ . Upon increasing  $K$ , the distribution/dynamics of  $\Lambda$  evolves from regularly spaced preferential values with a hopping dynamics to quasi-homogeneous coverage with a diffusive dynamics. For the smallest number of replicas considered ( $K = 8$ ), application of a biasing potential is recommended. For larger numbers of replicas ( $K \geq 16$ ), it becomes unnecessary. The calculated  $\Delta G$  values compare well with those obtained using other methods.

The convergence is accelerated relative to TI with Simpson quadrature (smaller error bar at identical total single-system sampling time), owing to improved orthogonal sampling and reduced quadrature errors. It is comparable to HRE, which shares the same orthogonal-sampling advantage. It is also similar to TI with EXT1 or MBAR as free-energy estimator, which achieve a similar improvement *via* an orthogonal-statistics advantage, *i.e.* by effectively mixing information concerning distinct configurational wells across  $\lambda$ -points. It should be stressed, however, that the present mutation is rather non-challenging in terms of orthogonal sampling. Work is in progress to investigate other types of systems with more complicated orthogonal spaces: (*i*) the side-chain mutation in the central residue of a tripeptide considered in Refs. [11, 12]; (*ii*) the hydrogen-to-bromine mutation in the base of a nucleotide considered in Refs. [13, 14]. Here, it is expected that CBTI alone (just like HRE) will help overcoming barriers in the orthogonal space when these barriers are low at some  $\lambda$ -value (as in the first system mentioned), but may be insufficient on its

own when these barriers are high at all  $\lambda$ -values (as in the second system mentioned), in which case additional modifications must be applied to create artificially an orthogonal tunnel at least over a limited  $\lambda$ -range.

Compared to existing MD-based alchemical free-energy calculation methods, the CBTI scheme can be viewed in at least three different ways: (*i*) as a continuous/deterministic/dynamical (instead of discrete/stochastic) analog of the HRE scheme<sup>???</sup> or the HRP scheme;<sup>???</sup> (*ii*) as a correlated multiple-replica analog (reminiscent of other swarm,<sup>???</sup> multiple-walker<sup>??</sup> or flying-Gaussian<sup>??</sup> approaches) of the  $\lambda$ -local elevation umbrella sampling ( $\lambda$ -LEUS) scheme<sup>???</sup> (or the conceptually similar flat-histogram<sup>??</sup>  $\lambda$ -metadynamics,<sup>???</sup> adaptive integration,<sup>?</sup> adaptive biasing force,<sup>?</sup> adaptively biased<sup>?</sup> and expanded-ensemble<sup>???</sup> methods); (*iii*) as an equilibrium multiple-replica variant of the slow-growth<sup>??</sup> (SG) method (bypassing the associated hysteresis issues<sup>???</sup> or the requirement for exponential averaging over multiple repeats<sup>???</sup> ).

Compared to plain TI, it shares the advantage of HRE/HRP and  $\lambda$ -LEUS in terms of enhanced orthogonal sampling.<sup>???</sup> Compared to HRE/HRP, it permits a deterministic and continuous sampling of the  $\lambda$ -range, and bypasses the need for a careful preselection<sup>???</sup> of the  $\lambda$ -ladder and exchange-attempt interval. Compared to both TI and HRE/HRP, the quasi-homogeneous  $\lambda$ -sampling also essentially removes quadrature errors. Finally, compared to  $\lambda$ -LEUS, it eliminates (or drastically reduces) the dead time associated with the preoptimization of a biasing potential<sup>?</sup> or, alternatively, the use of this formally non-equilibrium statistics<sup>?</sup> in the production calculation.<sup>?</sup> For the above reasons, the CBTI scheme certainly represents a useful addition to the alchemical free-energy calculation toolkit.

Like TI and HRE/HRP, the CBTI method is also intrinsically parallel. However, assuming that the replicas are assigned to separate processors (including possible GPU implementations or/and cloud-computing applications), the requirement of an all-to-all information exchange between processors at every timestep might represent a drawback of the method relative to the no-exchange and infrequent exchange situations of TI and HRE/HRP, respectively. Although the communication is lightweight (Hamiltonian  $\lambda$ -derivative, *i.e.* a single real number), the synchronization requirement may cause a performance loss (reduced scalability and fault tolerance). Unless asynchronous variants<sup>??</sup> or multiple-timestep schemes<sup>??</sup> can be developed, this performance loss may represent a problem for parallel applications in situations involving many replicas of a small system, as these will involve more data exchange at a more frequent rate.

This scheme opens the way to at least two types of generalizations and extensions. First, a number of components of the scheme can be modified/generalized and in particular the following: (*i*) different functions  $\zeta$  may be used in Eq. 3.6 to modulate the replica density along  $\lambda$  (*e.g.* smoothing the tips or adding plateaus<sup>?</sup> for a denser sampling close to or at the physical end-states); (*ii*) different matrices  $\underline{\mathbf{C}}$  may be used in Eq. 3.14 corresponding to a non-uniform weighting of the  $\lambda$ -forces into the  $\Lambda$ -force, (*e.g.* canceling the effect of higher-order derivatives by alternating non-integer weights in analogy to standard quadrature methods); (*iii*) the coupling between replicas may be generalized from sequential pairwise constraints to possibly non-pairwise potentials (*e.g.* collective or harmonic); (*iv*) the TI-like free-energy estimator of Eq. 3.22 may be replaced by a statistically more powerful one of the MBAR type;<sup>???</sup> (*v*) CBTI would benefit from the use of an alchemical coupling path presenting a vanishing free-

energy derivative at the physical end-states (residual free-energy barriers along  $\Lambda$  decaying at least in  $K^{-2}$  instead of  $K^{-1}$  for large  $K$ , as shown in Sect. 3.A).

Second, the application range of CBTI, restricted here to alchemical processes, can be extended to encompass either thermodynamic or conformational free-energy changes. In the former case, extension to a CB variant of parallel tempering<sup>?</sup> (*i.e.* a CBPT scheme) appears relatively straightforward considering that scaling the temperature is equivalent to scaling the system potential energy. Such a CBPT scheme would represent a form of multicanonical sampling.<sup>???</sup> In the latter case, extension to a CB variant of umbrella sampling<sup>??</sup> (*i.e.* a CBUS scheme) could be designed *e.g.* by anchoring harmonic potentials to the CB and propagating the corresponding restraining forces onto the CB. Finally the extension of the CB approach to multistate problems<sup>???</sup> (*e.g.* path or network of CBs connecting the different states of a system) as well as multidimensional problems (sub-CBs anchored to a main CB) can also be envisioned.

## Appendix 2.A RELATIONSHIP BETWEEN CBTI AND QUADRATURE INTEGRATION

This appendix investigates the convergence properties of the CBTI scheme upon increasing the number  $K$  of replicas based on an analogy with quadrature integration. More precisely, it is shown that the variation amplitude  $G_\Lambda^*$  of  $G_\Lambda(\Lambda)$  decreases at least as  $K^{-1}$  in the limit of large  $K$  (with  $K$  even), as suggested by Fig. 3.1b and 3.6.

To this purpose, we first observe that the situation of a conveyor belt (CB) spanning the  $\lambda$ -range  $[0, 1]$  of the free-energy profile  $G(\lambda)$  with  $K$  equidistant replicas at spacings  $2K^{-1}$  along the cable (Fig. 3.A.1a) is mathematically equivalent to the situation of a train spanning a periodic function  $g(x)$  of period  $2\pi$  with  $K$  equidistant carriages at spacings  $2\pi K^{-1}$  spanning one period (Fig. 3.A.1b). The function  $g(x)$  is obtained by mirroring  $G(\lambda)$  at  $\lambda = 0$  or  $1$ , periodically translating the mirrored function by integer multiples of two, and stretching  $\lambda$  by a factor of  $\pi$  to define the new variable  $x$ .

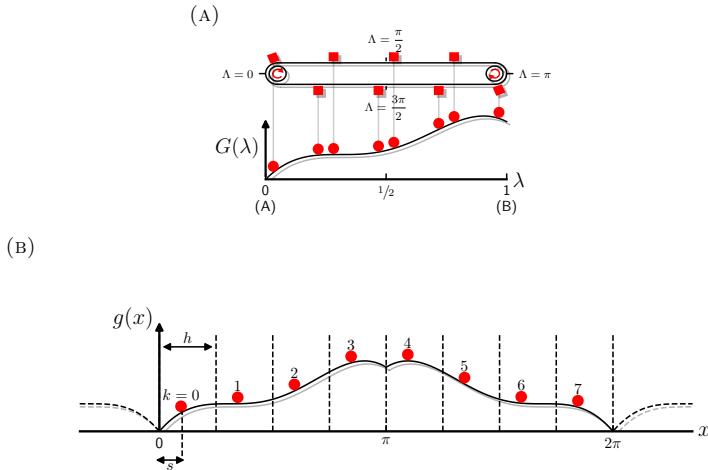


FIGURE 2.A.1: *Analogy between the CBTI scheme and the sampling of an even periodic function by a train of equispaced points.* The illustration considers a conveyor belt (CB) with  $K = 8$  replicas, analog to a train with 8 carriages.

In this analogy, we now have a periodic function  $g(x)$  of period  $2\pi$  that is even over the reference interval  $[0, 2\pi]$ . This function is sampled by  $K$  equidistant points at locations

$$x_k = s + kh \quad \text{with} \quad k = 0 \dots K-1 \quad \text{and} \quad h = 2\pi K^{-1}, \quad (2.A.1)$$

where  $s$  represents the offset position of the first point  $k = 0$ . Note that  $s$  can always be selected in the interval  $[0, h)$  by consideration of periodicity, along with an appropriate renumbering of the points carriages. We also introduce the negative first derivative of the function

$$f(x) = -g'(x). \quad (2.A.2)$$

In practical applications of CBTI, the Hamiltonian-coupling scheme will be defined by a continuously differentiable function of  $\lambda$ , in which case  $G(\lambda)$  is also continuously differentiable. The same applies to  $g(x)$  in the analogy of Fig. 3.A.1b, except for the possible occurrence of kinks at integer multiples of  $\pi$ . These occur because the derivative  $G'(\lambda)$  of  $G(\lambda)$  generally differs from zero at the physical end-states  $\lambda = 0$  or  $\lambda = 1$ . As a result,  $f(x)$  is continuous except at these points, where it will present jumps. An illustrative example for functions  $g(x)$  and  $f(x)$  with the above properties is shown in Figs. 3.A.2a and 3.A.2b, respectively.

In the analogy of Fig. 3.A.1b to the CB situation, the total force on the train of points (analog to the mean force on the  $\Lambda$ -variable) is given for a specific value of  $s$  (analog to the  $\tilde{\Lambda}$ -variable) by

$$F(s) = \sum_{k=0}^{K-1} f(kh + s). \quad (2.A.3)$$

As illustrated in Fig. 3.A.2c, this sum has a simple interpretation. It represents a trapezoidal quadrature estimate to the integral of the periodic function  $f(x)$  over the period  $[s, 2\pi + s]$ . Because  $f(x)$  is the negative derivative of the periodic function  $g(x)$ , its exact integral over one period must be zero. Thus, for any  $s$ , we expect  $F(s)$  to converge to zero in the limit  $K \rightarrow \infty$ . The following paragraphs investigate the corresponding convergence rate.

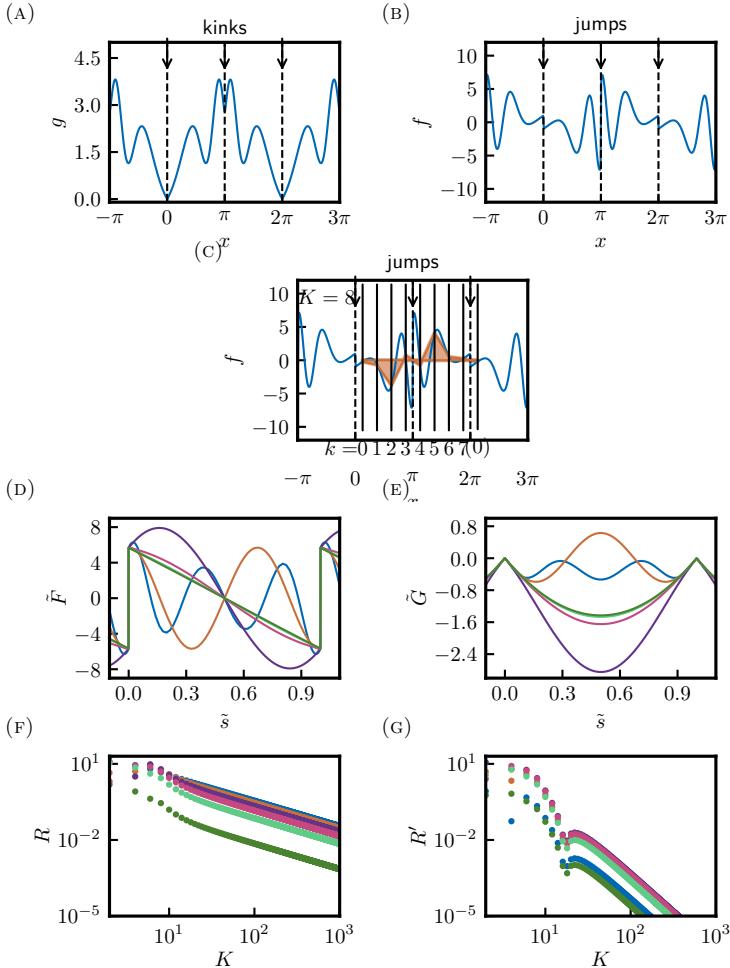


FIGURE 2.A.2: *Illustrative functions supporting the discussion in Sect. 3.A.* Panel (a) shows the illustrative function  $g(x)$  defined over the reference period  $[0; 2\pi[$  as  $x + \sin(x^2)$  if  $x < \pi$  or  $(2\pi - x) + \sin((2\pi - x)^2)$  if  $x \geq \pi$ . Panel (b) shows the negative derivative  $f(x)$  of  $g(x)$ . Panel (c) illustrates the trapezoidal integration of  $f(x)$  (blue) as a shaded area (orange) with  $K = 8$  and  $s = 0.4$ . The vertical lines show the positions of the carriages at values  $x_m$ . The trapezoids are seen to represent a poor approximation at the discontinuities. Panel (d) illustrates the total force  $\tilde{F}(\tilde{s})$  of Eq. 3.A.4 for different values of  $K$ . Panel (e) illustrates the potential energy  $\tilde{G}(\tilde{s})$  for different numbers of replicas  $K$ , which is the running integral of  $\tilde{F}$  for different numbers of replicas  $K$ . Panel (f) shows in a log-log form the residual integral  $R(\tilde{s})$  of Eq. 3.A.8 as a function of  $K$  for different values of  $\tilde{s}$ . Panel (g) shows in a log-log form the corrected residual integral  $R'(\tilde{s})$  of Eq. 3.A.10. Note that the logarithm of the absolute value of  $R$  or  $R'$  is displayed. The cusp in Panel (g) is explained by a change of sign.

The function  $F(s)$  is periodic of period  $h$ , odd over the reference interval  $[0, h)$  and generally discontinuous at  $s$  values which are integer multiples of  $h$ . Introducing

$$\tilde{s} = h^{-1}s \quad \text{and} \quad \tilde{F}(\tilde{s}) = F(h\tilde{s}) = \sum_{k=0}^{K-1} f((k + \tilde{s})h), \quad (2.A.4)$$

the function  $\tilde{F}(\tilde{s})$  is periodic of period 1, odd over the reference interval  $[0, 1)$  and generally discontinuous at integer values of  $\tilde{s}$ .

The corresponding potential energy of the train of points (analog to the free-energy profile  $G_\Lambda(\Lambda)$  of the CB over the interval  $[0, 2\pi K^{-1}]$ ) is given by

$$G(s) = - \int_0^s ds' F(s'). \quad (2.A.5)$$

The function  $G(s)$  is periodic of period  $h$ , even over the reference interval  $[0, h)$  with an extremum at  $h/2$ , and possibly presents kinks at  $s$  values which are integer multiples of  $h$ . Using the variable  $\tilde{s}$  and introducing

$$\tilde{G}(\tilde{s}) = - \int_0^{\tilde{s}} d\tilde{s}' \tilde{F}(\tilde{s}') , \quad (2.A.6)$$

one may write

$$G(s) = h\tilde{G}(h^{-1}s). \quad (2.A.7)$$

The function  $\tilde{G}(\tilde{s})$  is periodic of period 1, even over the reference interval  $[0, 1)$  with an extremum at  $1/2$ , and possibly presents kinks at integer values of  $\tilde{s}$ .

The functions  $\tilde{F}(\tilde{s})$  and  $\tilde{G}(\tilde{s})$  corresponding to the illustrative

functions  $g(x)$  and  $f(x)$  are shown in Figs. 3.A.2d and 3.A.2e, respectively, considering the choices  $K = 2, 4, 8, 16, 32$  or  $64$ . Convergence is observed upon increasing  $K$ , in which  $\tilde{F}(\tilde{s})$  becomes linear and  $\tilde{G}(\tilde{s})$  parabolic over the reference interval  $[0, 1]$ . Thus, in the limit of large  $K$ , one expects the maximal variation  $\tilde{G}^*$  in  $\tilde{G}(\tilde{s})$  to converge to a constant and, *via* Eq. 3.A.7, the maximal variation  $G^*$  in  $G(s)$  (analog to the corresponding maximal variation  $G_\Lambda^*$  of  $G_\Lambda(\Lambda)$  in the CB scheme) to scale as  $h$ , *i.e.* as  $K^{-1}$ .

This observation, made here in a special case, can be generalized as follows. Given

$$R(\tilde{s}) = h\tilde{F}(\tilde{s}) = h \sum_{k=0}^{K-1} f(h(k + \tilde{s})), \quad (2.A.8)$$

$G^*$  is the absolute extremum (largest absolute value) of the function

$$S(\tilde{s}) = h\tilde{G}(\tilde{s}) = \int_0^{\tilde{s}} d\tilde{s}' R(\tilde{s}') \quad (2.A.9)$$

over the  $\tilde{s}$ -interval  $[0, 1/2]$ . According to Fig. 3.A.2c, the quantity  $R(\tilde{s})$  is the residual of the trapezoidal quadrature approximation to the integral of  $f(x)$  over one period, where the first grid point is offset from the origin by a fraction  $\tilde{s}$  of the spacing  $h$ . If  $f(x)$  was continuous everywhere, the convergence would be quadratic in  $h$  (*i.e.* scale as  $K^{-2}$ ), as expected from a trapezoidal quadrature. However, the presence of discontinuities in  $f(x)$  at  $0$  and  $\pi$  introduces an error that is linear in  $h$ . To recover quadratic convergence, one would need to introduce a correcting term for the points  $0, K/2 - 1, K/2$  and  $K - 1$  surrounding the discontinuities, namely

$$\begin{aligned}
R'(\tilde{s}) = & R(\tilde{s}) + h \left\{ \frac{\tilde{s}-1}{2} \left[ f(h\tilde{s}) + f \left( h \left( \tilde{s} + \frac{K}{2} \right) \right) \right] \right. \\
& - \frac{\tilde{s}}{2} \left[ f(h(\tilde{s}+K-1)) + f \left( h \left( \tilde{s} + \frac{K}{2} - 1 \right) \right) \right] \\
& \left. + \left( \tilde{s} - \frac{1}{2} \right) (\Delta_0 + \Delta_\pi) \right\} \tag{2.A.10}
\end{aligned}$$

where  $\Delta_0$  and  $\Delta_\pi$  account for the magnitude of the discontinuities, *i.e.*

$$\Delta_0 = \lim_{x \xrightarrow{>} 0} f(x) \quad \text{and} \quad \Delta_\pi = \lim_{x \xrightarrow{>} \pi} f(x). \tag{2.A.11}$$

The convergence of  $R(\tilde{s})$  and  $R'(\tilde{s})$  for the illustrative functions  $g(x)$  and  $f(x)$  upon increasing  $K$  along with the choices  $\tilde{s} = 0.01, 0.2, 0.3, 0.4, 0.49$  is shown in logarithmic form in Figs. 3.A.2f and 3.A.2g, respectively. As expected, for large  $K$ ,  $R$  shows a linear convergence (slope -1) and  $R'$  a quadratic one (slope -2). The limiting cases  $\tilde{s} = 0$  and  $\tilde{s} = 1/2$  are special. For  $\tilde{s} = 0$ , Eq. 3.A.8 implies an evaluation of  $f(x)$  at the discontinuity. If one uses an average value of 0 for these points,  $R(0)$  evaluates to 0 for any  $K$  by symmetry. The same applies for  $\tilde{s} = 1/2$ ,  $R(1/2)$ . For all other values for  $\tilde{s}$ , the convergence is as expected, *i.e.* linear for  $R(\tilde{s})$  and quadratic for  $R'(\tilde{s})$ .

If  $R(\tilde{s})$  converges to zero as  $K^{-1}$  for (nearly) all  $\tilde{s}$  values, the function  $S(\tilde{s})$  of Eq. 3.A.9 will converge to zero with the same scaling, and so will its absolute extremum  $G^*$ . We conclude that the magnitude  $G_\Lambda^*$  of the residual variations in  $G_\Lambda(\Lambda)$  for the CBTI scheme scales at least as  $K^{-1}$  in the limit of large  $K$ . And if  $G(\lambda)$  has a vanishing derivative at the physical end-states, it will

scale at least as  $K^{-2}$ . It should be stressed, however, that these are worst-case scalings. A higher level of continuity or specific symmetry properties of  $G(\lambda)$  may tighten the scaling. To give an extreme case, for a free-energy profile  $G(\lambda) = \cos(2\pi\lambda)$ , one shows easily that the residual variations  $G_{\Lambda}^*$  entirely vanish irrespective of  $K$ . This follows from

$$\sum_{k=0}^{K-1} \cos(2\pi K^{-1}k + c) = 0 \quad \forall c \quad (\text{for } K \text{ even}). \quad (2.A.12)$$

The present analysis connects the convergence of the CBTI scheme with  $K$  to the properties of a quadrature integration. This connection opens interesting tracks for future work. In particular, it suggests that these convergence properties could be improved by altering the coupling scheme (*e.g.* to enforce vanishing  $G(\lambda)$  derivatives at the end-states), the function  $\zeta$  in Eq. 3.6 (*e.g.* higher system density close to discontinuities) and the CBTI weighting in Eq. 3.4 (*e.g.* from trapezoidal to higher-level quadrature, *e.g.* Simpson or even Romberg).

## Appendix 2.B CHOICE OF THE CBTI PARAMETERS

Here, we explore the influence of the CBTI mass parameter  $m_{\Lambda}$  and thermostat coupling time  $\tau_{\Lambda}$  considering different numbers  $K$  of replicas.

A first set of simulations involves the choices  $m_{\Lambda} = 16, 160, 800, 1600$  or  $3200 \text{ u nm}^2$ , along with  $K = 16$  replicas in the absence of separate thermostat coupling for  $\Lambda$ , *i.e.*  $\tau_{\Lambda} \rightarrow \infty$  (10 ns

simulations after 0.2 ns equilibration). The results are shown graphically in Figs. 3.F.1 - 3.F.5 and reported numerically in Tab. 3.1 (entries 1-5). In the absence of explicit thermostating for the  $\Lambda$ -variable, the average temperature  $T_\Lambda$  of this variable is controlled exclusively by its coupling to the conformational degrees of freedom, themselves thermostated by the application of SD with a reference temperature of 298.15 K and an effective coupling time of 0.1 ps. All choices of  $m_\Lambda$  considered lead to average temperatures  $T_\Lambda$  close to 298.15 K, suggesting an appropriate kinetic-energy exchange. The mass-parameter choices of  $m_\Lambda = 16 \text{ u nm}^2$  and  $1600 \text{ u nm}^2$  present the largest deviations of about 10 K and 6 K, respectively, while the deviation is at most 2 K for the other choices.

Expectedly, increasing  $m_\Lambda$  tendentially leads to a decrease in the diffusion coefficient  $D_\Lambda$ . This results from a smaller width  $\sigma_\Lambda$  of the  $\dot{\Lambda}$  distribution (Eq. 3.25; see also Fig. 3.5a where  $m_\Lambda$  increases from left to right), *i.e.* a lower average magnitude of the velocity along  $\Lambda$ . However,  $D_\Lambda$  is also affected by the presence of residual variations in the free-energy profile  $G_\Lambda(\Lambda)$ , which may be more easily overcome at higher  $m_\Lambda$  (more inertia). This is reflected in the increase of the  $\dot{\Lambda}$  autocorrelation time  $\tau_{\dot{\Lambda}}$  upon increasing  $m_\Lambda$  (see also Fig. 3.4 for the corresponding normalized autocorrelation functions). This second effect is likely to have more influence for small  $K$ , where the  $G_\Lambda(\Lambda)$  variations are more pronounced. In the present case with  $K = 16$ , these opposite effects probably explain why the trend in  $D_\Lambda$  upon increasing  $m_\Lambda$  is not strictly monotonic, with a comparatively high  $D_\Lambda$  for  $m_\Lambda = 1600 \text{ u nm}^2$  (see also Fig. 3.5c where the diffusion trend is non-monotonic upon increasing  $K$  when  $m_\Lambda$  is made proportional to  $K^{1/2}$ ).

In terms of the calculated free-energy change  $\Delta G$ , all simula-

tions provide the same values within the statistical error, with a slightly larger deviation for the simulation involving the lowest mass  $m_{\Lambda} = 16 \text{ u nm}^2$ , probably due to the comparatively large  $T_{\Lambda}$  deviation. Because it leads to accurate  $T_{\Lambda}$  and  $\Delta G$  values while presenting a high  $D_{\Lambda}$ , the second-to-lowest mass  $m_{\Lambda} = 160 \text{ u nm}^2$  was chosen for all following CBTI simulations employing  $K = 16$  replicas.

A second set of simulations involves a separate coupling of the  $\Lambda$  variable to a Nosé-Hoover chain thermostat with the choices  $\tau_{\Lambda} = 0.05, 0.1, 0.5, 1$  or  $2 \text{ ps}$  along with  $K = 16$  replicas and  $m_{\Lambda} = 160 \text{ u nm}^2$  ( $10 \text{ ns}$  simulations after  $0.2 \text{ ns}$  equilibration). The results are shown graphically in Figs. 3.F.6 - 2.F.10 and reported numerically in Tab. 3.1 (entries 6-10). Except for the two shortest coupling times, the average temperature  $T_{\Lambda}$  is very close to the reference temperature, with a maximal deviation of about  $2 \text{ K}$ . The larger deviations for  $\tau_{\Lambda} = 0.05$  or  $0.1 \text{ ps}$  suggest that the coupling of  $\Lambda$  to its thermostat should be made less tight than that of the conformational degrees of freedom to their thermostat (here SD with  $0.1 \text{ ps}$ ). The coupling parameter  $\tau_{\Lambda}$  has no influence on the width  $\sigma_{\dot{\Lambda}}$  of the  $\dot{\Lambda}$  distribution, which is the same for all five simulations and very close to the Maxwell-Boltzmann one (Eq. 3.25). On the other hand, a lower  $\tau_{\Lambda}$ , *i.e.* a tighter coupling, reduces the inertia of the conveyor belt, which results in a decrease of the autocorrelation time  $\tau_{\dot{\Lambda}}$ . This leads to a tendential decrease of  $D_{\Lambda}$  upon decreasing  $\tau_{\Lambda}$ . In terms of the calculated  $\Delta G$ , all simulations provide consistent values within the statistical error, with a slightly larger deviation for the simulation involving  $\tau_{\Lambda} = 0.05 \text{ ps}$ , probably again due to the comparatively large  $T_{\Lambda}$  deviation. In comparison to the first set of simulations with  $\tau_{\Lambda} \rightarrow \infty$ , the  $\Delta G$  values are slightly higher ( $0.1 - 0.3 \text{ kJ mol}^{-1}$ ), which could be due to the fact that the sampled  $\dot{\Lambda}$  distribution  $P_{\dot{\Lambda}}$  is even closer to

the Maxwell-Boltzmann distribution (Eq. 3.25), also for shorter time intervals. For  $\tau_\Lambda \geq 0.5$  ps, the choice of  $\tau_\Lambda$  has little influence on the calculation, and a coupling time  $\tau_\Lambda = 0.5$  ps was chosen for all following CBTI simulations. Note, however, that a looser coupling could be of advantage by leading to a higher  $D_\Lambda$ .

A third set of simulations considers the choices  $K = 8, 16, 32, 64$  or  $128$  along with  $m_\Lambda = 40K^{1/2}$  u nm<sup>2</sup> and  $\tau_\Lambda = 0.5$  ps ( $256K^{-1}$  ns simulations after 0.2 ns equilibration). The results are shown graphically in Figs. 2.F.11 - 2.F.15 and reported in Tab. 3.1 (entries 11-15). Simulations 11, 13 and 15 are discussed in details in Sect. 3.4 (see Figs. 3.3 and 3.5). For the five simulations, the average temperature  $T_\Lambda$  is close to the target temperature, with a maximal deviation of about 6 K. The free-energy differences  $\Delta G$  are consistent within the statistical error, except for the simulation employing  $K = 8$  replicas (entry 11), which is due to the non-uniform sampling of the  $\lambda$ -range as discussed in Sect. 3.4.

The rationale for making the mass  $m_\Lambda$  proportional to the square-root of  $K$ , an arbitrary parameter choice, is the following. If one wishes the CB variable  $\Lambda$  to evolve dynamically on comparable timescales irrespective of the number  $K$  of replicas attached to it, one should ensure that its acceleration  $\ddot{\Lambda}$  depends only weakly on  $K$ . Assuming that the forces exerted by the  $K$  replicas are essentially uncorrelated, their sum (*i.e.* the net force on the CB) will scale as  $K^{1/2}$ . Thus, an identical scaling of  $m_\Lambda$  is required to preserve an approximately constant  $\ddot{\Lambda}$ . The time series and distributions of  $\ddot{\Lambda}$  for different choices of  $K$  when using this scaling for  $m_\Lambda$  (see Fig. 3.5b) are indeed similar, supporting the above considerations.

In summary, this appendix shows that when selected within reasonable ranges, the parameters  $m_\Lambda$  and  $\tau_\Lambda$  have only a limited influence on the kinetic-energy exchange between  $\Lambda$  and the con-

formational degrees of freedom, on the temperature  $T_\Lambda$ , on the diffusion constant  $D_\Lambda$ , and on the calculated free energies  $\Delta G$ . In Sect. 3.4, working choice  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  is selected for all CBTI calculations.

## Appendix 2.C SIMPLIFIED FREE-ENERGY ESTIMATORS

The free-energy estimator employed here for CBTI is given by Eq. 3.22. The approximation involved corresponds to a simple forward rectangular quadrature, where the Hamiltonian derivative is averaged over  $J$  successive bins considering all replicas simultaneously. If  $K$  replicas sample  $L$  configurations each, and  $p(\lambda)$  is close to homogeneous, the number of data points per bin will be close to  $KL/J$ , with limited variations across bins. In this case, one may consider a simpler alternative to Eq. 3.22 that does not involve the specification of a number of bins.

Considering all the  $K$  replicas simultaneously, the  $KL$  pairs of  $\lambda$ -values and associated Hamiltonian derivatives are sorted in ascending order for  $\lambda$  (index  $i = 0, \dots, KL-1$ ). One then calculates

$$\Delta G_{\text{alt}} = (KL)^{-1} \sum_{i=0}^{KL-1} \frac{\lambda_{i+1} - \lambda_{i-1}}{2} \left( \frac{\partial \mathcal{H}}{\partial \lambda} \right)_i \quad (2.C.1)$$

using  $\lambda_{-1} = \lambda_{KL} = 0$ . This alternative estimate, noted  $\Delta G_{\text{alt}}$ , considers that each sample  $i$  defines its own single-point bin, of width determined by contact to the next lower and higher data points  $i-1$  and  $i+1$ , respectively, thereby accounting for a possible

heterogeneity in the  $\lambda$ -sampling.

In the limit where the sampling becomes sufficiently close to homogeneous, Eqs. 3.22 or 3.C.1 can be replaced by an even simpler approximate expression, namely

$$\Delta G_{\text{app}} \approx K^{-1} \sum_{k=0}^{K-1} \left\langle \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \right\rangle^{\dagger*}, \quad (2.C.2)$$

which corresponds to averaging the Hamiltonian derivative over all replicas and over the entire CBTI simulation. This approximate estimate, noted  $\Delta G_{\text{app}}$ , assumes that the  $KL$  data points are on average equispaced along  $\lambda$ , which only holds in the limit of homogeneous sampling.

The accuracy of the estimate  $\Delta G$  of Eq. 3.22 with  $J = 500$  (except for entry 11,  $J = 200$ ) is compared in Tab. 3.1 to those of the simpler expressions  $\Delta G_{\text{alt}}$  and  $\Delta G_{\text{app}}$ , which do not require the specification of a number of bins. The estimate  $\Delta G_{\text{alt}}$  is very close to  $\Delta G$  (within error bar), suggesting that Eq. 3.C.1, which assumes that each sampled  $\lambda$ -value is at the center of its own bin, is a viable parameter-free alternative to Eq. 3.22. On the other hand, the estimate  $\Delta G_{\text{app}}$ , which assumes that each bin encompasses the same number of sampled  $\lambda$ -points irrespective of  $J$ , is more approximate. As expected from the involved assumption,  $\Delta G_{\text{app}}$  is only a good approximation to  $\Delta G$  when the CBTI sampling is indeed close to uniform along  $\lambda$ . Considering entries 11-17, this is essentially the case for the unbiased simulations with  $K \geq 32$ . However, for the unbiased simulation with  $K = 8$  and the two biased simulations, the sampling is not sufficiently homogeneous and the discrepancy can be significant (about  $1 - 5 \text{ kJ mol}^{-1}$ ).

## Appendix 2.D AUTOCORRELATION OF THE HAMILTONIAN DERIVATIVE

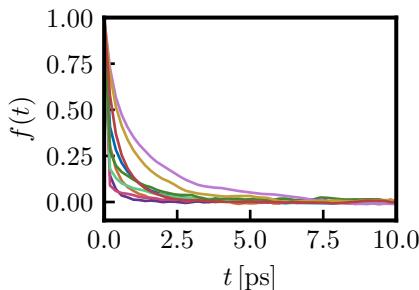


FIGURE 2.D.1: Normalized autocorrelation function  $f(t)$  of the Hamiltonian derivative for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. These functions are based on the TI calculation relying on  $K_{\text{TI}} = 17$   $\lambda$ -points. The functions displayed correspond to every second  $\lambda$ -point. The associated autocorrelation times  $\tau_f$  are reported in Tab. 3.D.1.

TABLE 2.D.1: Autocorrelation times  $\tau_f$  of the Hamiltonian derivative for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. These times are based on the TI calculation relying on  $K_{\text{TI}} = 17$   $\lambda$ -points. The associated autocorrelation functions  $f(t)$  are displayed in Fig. 3.D.1 for every second  $\lambda$ -point. The autocorrelation times were derived by performing an exponential fit to the  $f(t)$  curve.

$\lambda$	$\tau_f$ [ps]	$\lambda$	$\tau_f$ [ps]
0.000	0.339	0.562	0.161
0.062	0.297	0.625	0.224
0.125	0.207	0.688	0.377
0.188	0.150	0.750	1.080
0.250	0.109	0.812	1.119
0.312	0.085	0.875	0.799
0.375	0.087	0.938	0.555
0.438	0.112	1.000	0.442
0.500	0.139		

## Appendix 2.E REFERENCE TI CALCULATIONS

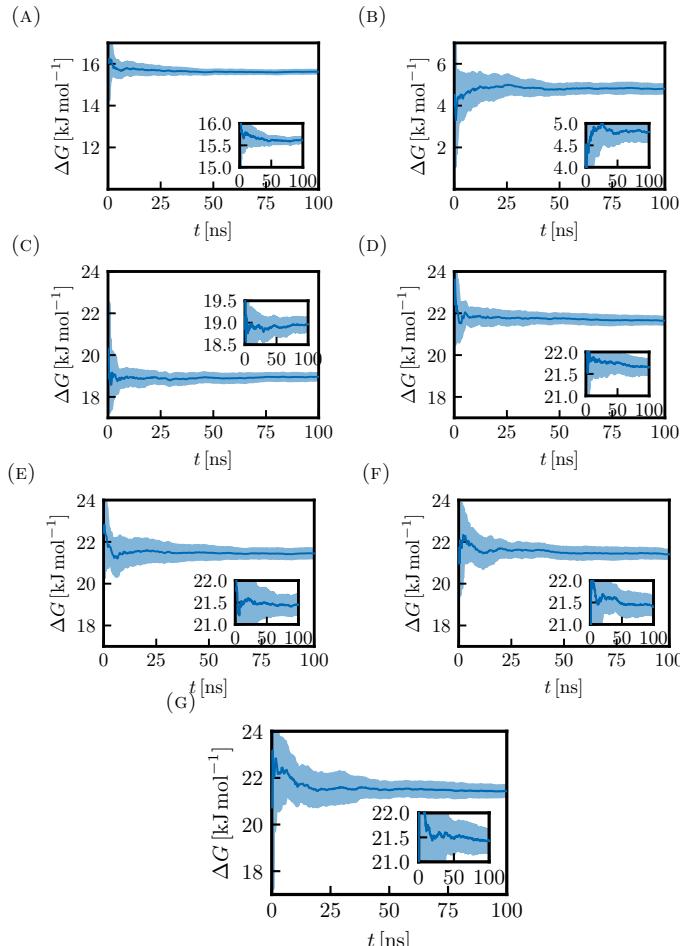


FIGURE 2.E.1: Convergence properties of the TI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. This figure complements Fig. 3.2b by also including the TI calculations involving fewer than 129  $\lambda$ -points. Here, the running  $\Delta G$  estimates are shown for calculations involving  $K_{\text{TI}} = 2^n + 1$  equidistant  $\lambda$ -points with  $n = 1, 2, \dots, 7$  resulting in  $K_{\text{TI}} = 3$  a, 5 b, 9 c, 17 d, 33 e, 65 f and 129 g. Note that the vertical ranges in Panels a and b differ from the range used in the other graphs. The corresponding numerical values at full sampling time are reported in Tabs. 3.2 and 3.E.1.

TABLE 2.E.1: *Repeat results over TI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* The free-energy change  $\Delta G$  is reported for TI calculations involving  $K_{\text{TI}} = 2^n + 1$   $\lambda$ -points with  $n = 1, 2, \dots, 7$ . For each choice of  $K_{\text{TI}}$  the results of ten repeats (successive rows) are listed, all involving a total sampling time of 100 ns equally distributed over the  $\lambda$ -points. For each repeat, the calculated  $\Delta G$  (Simpson quadrature) is reported along with a bootstrap error estimate (first parenthesis) and an error propagated from the bootstrap errors of the Hamiltonian derivative at each  $\lambda$ -point (second parenthesis). These two error estimates do not include a Student  $t$ -factor. Repeat statistics based on this data can be found in Tab. 3.2.

$K_{\text{TI}}$	$\Delta G [\text{kJ mol}^{-1}]$						
	1 3	2 5	3 9	4 17	5 33	6 65	7 129
1	15.68 (0.08) (0.10)	4.82 (0.18) (0.15)	19.22 (0.18) (0.16)	21.87 (0.17) (0.15)	21.48 (0.20) (0.15)	21.18 (0.24) (0.15)	21.45 (0.30) (0.14)
2	15.52 (0.11) (0.10)	4.32 (0.18) (0.14)	18.72 (0.17) (0.16)	21.83 (0.18) (0.14)	21.59 (0.19) (0.14)	21.72 (0.22) (0.14)	21.87 (0.29) (0.14)
3	15.64 (0.09) (0.09)	4.50 (0.22) (0.16)	18.92 (0.17) (0.14)	21.57 (0.20) (0.14)	21.72 (0.17) (0.14)	21.62 (0.20) (0.14)	21.60 (0.26) (0.14)
4	15.58 (0.10) (0.09)	4.98 (0.17) (0.14)	19.18 (0.17) (0.14)	21.88 (0.18) (0.14)	21.85 (0.19) (0.14)	21.66 (0.22) (0.14)	21.67 (0.31) (0.14)
5	15.71 (0.10) (0.09)	4.97 (0.18) (0.15)	18.77 (0.17) (0.15)	21.37 (0.18) (0.15)	21.37 (0.22) (0.15)	21.29 (0.25) (0.14)	21.30 (0.26) (0.14)
6	15.62 (0.10) (0.09)	5.11 (0.18) (0.16)	18.99 (0.17) (0.15)	21.54 (0.21) (0.15)	21.19 (0.17) (0.15)	21.66 (0.22) (0.14)	21.76 (0.27) (0.14)
7	15.66 (0.10) (0.08)	4.75 (0.19) (0.15)	18.92 (0.14) (0.15)	21.67 (0.14) (0.15)	21.71 (0.21) (0.14)	21.29 (0.22) (0.14)	21.39 (0.29) (0.14)
8	15.48 (0.09) (0.09)	4.79 (0.18) (0.15)	18.77 (0.16) (0.15)	21.37 (0.17) (0.15)	21.03 (0.19) (0.15)	21.34 (0.25) (0.14)	20.99 (0.33) (0.14)
9	15.58 (0.09) (0.08)	4.77 (0.19) (0.16)	18.98 (0.19) (0.15)	21.49 (0.18) (0.15)	21.12 (0.18) (0.14)	21.02 (0.23) (0.14)	21.16 (0.32) (0.14)
10	15.65 (0.10) (0.09)	5.06 (0.19) (0.14)	19.17 (0.17) (0.15)	21.95 (0.19) (0.15)	21.45 (0.18) (0.15)	21.24 (0.22) (0.14)	21.09 (0.26) (0.14)

## Appendix 2.F UNBIASED CBTI SIMULATIONS

### 2.F.1 EXPLORATION OF THE INFLUENCE OF $m_\Lambda$ IN CBTI

**Simulation #1,**  $m_\Lambda = 16 \text{ u nm}^2$

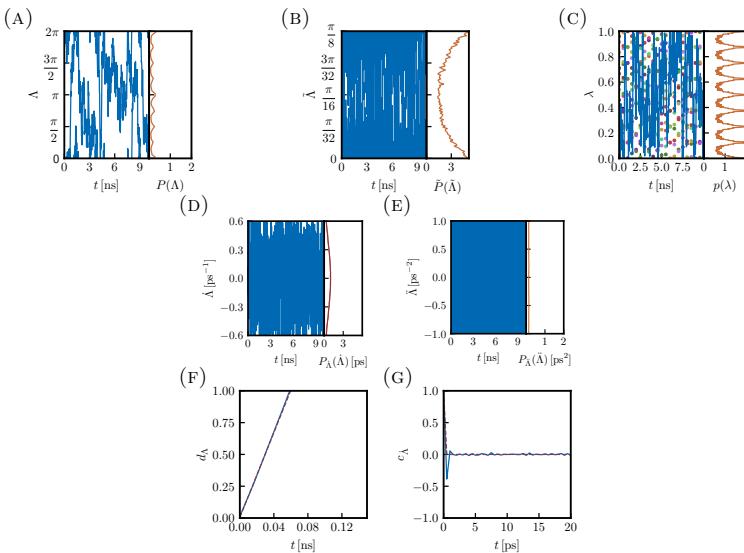


FIGURE 2.F.1: Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 16 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ). This simulation corresponds to entry 1 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\dot{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

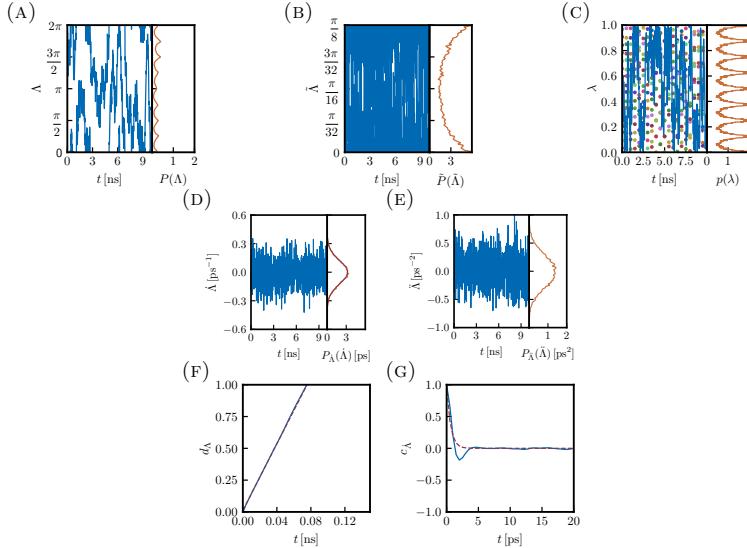
**Simulation #2,  $m_\Lambda = 160 \text{ u nm}^2$** 


FIGURE 2.F.2: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 2 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

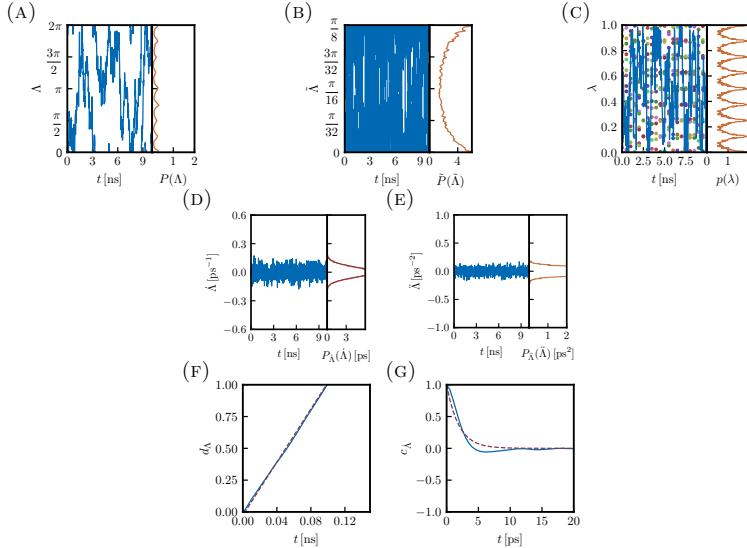
**Simulation #3,  $m_\Lambda = 800 \text{ u nm}^2$** 


FIGURE 2.F.3: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 800 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 3 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\bar{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\bar{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\bar{\lambda}^{\ddot{\cdot}}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\bar{\Lambda}$  (blue) along with an exponential fit (dashed brown).

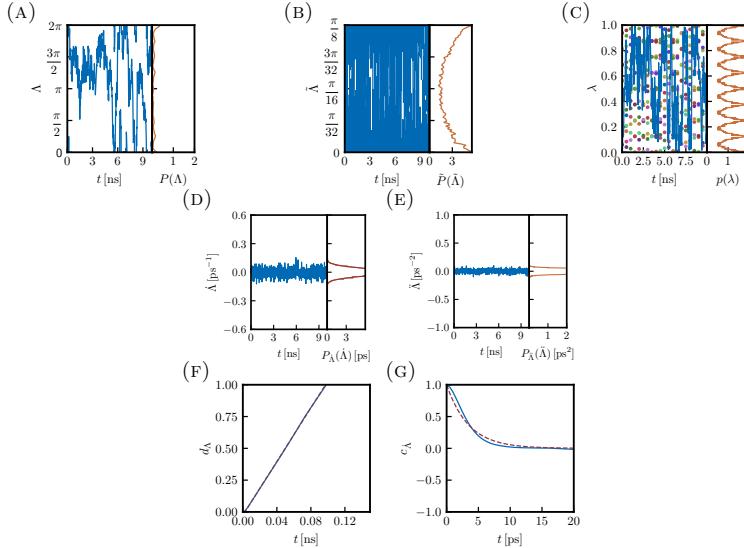
**Simulation #4,  $m_\Lambda = 1600 \text{ u nm}^2$** 

FIGURE 2.F.4: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 1600 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 4 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\bar{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\bar{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\bar{\lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\bar{\Lambda}$  (blue) along with an exponential fit (dashed brown).

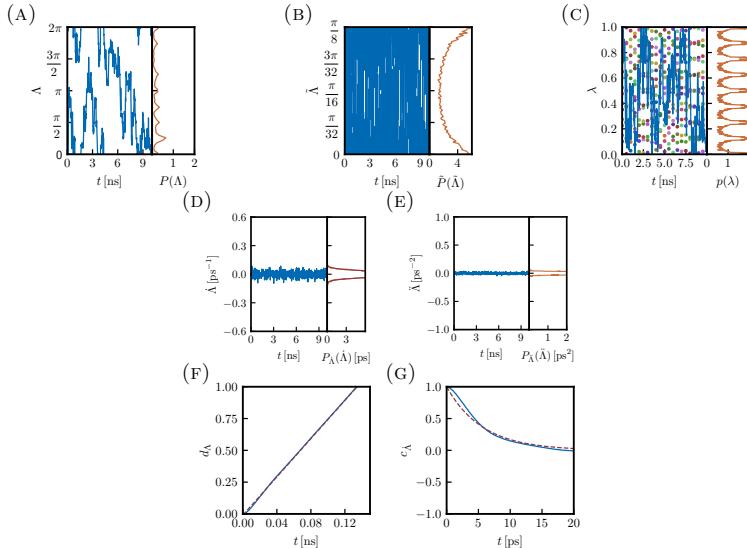
**Simulation #5,  $m_\Lambda = 3200 \text{ u nm}^2$** 

FIGURE 2.F.5: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 3200 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 5 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

## 2.F.2 EXPLORATION OF THE INFLUENCE OF $\tau_\Lambda$ IN CBTI

**Simulation #6,  $\tau_\Lambda = 0.05$  ps**

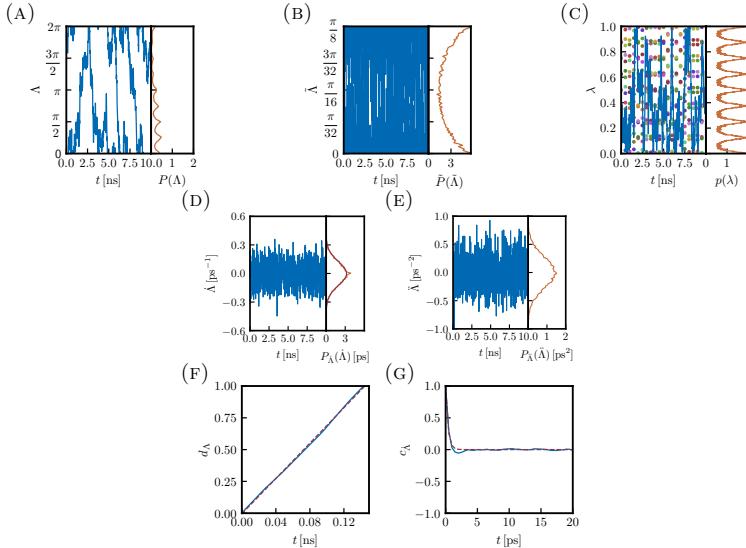


FIGURE 2.F.6: Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.05$  ps. This simulation corresponds to entry 6 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\dot{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

**Simulation #7,  $\tau_\Lambda = 0.1$  ps**

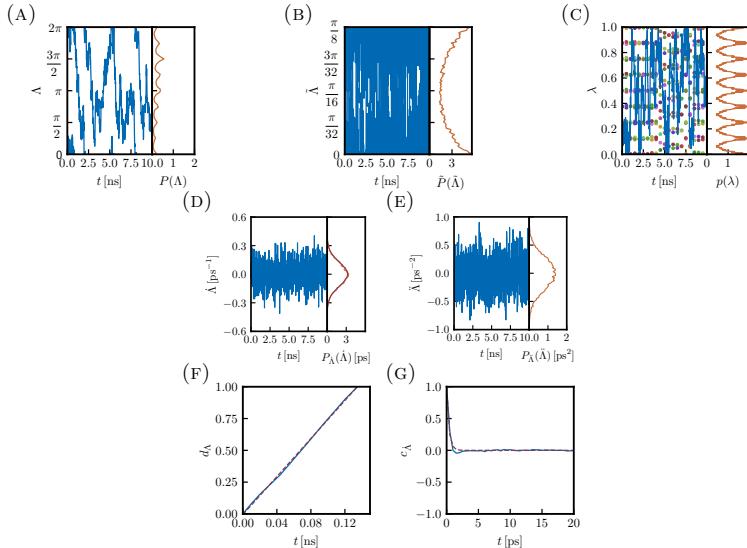


FIGURE 2.F.7: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.1$  ps.* This simulation corresponds to entry 7 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

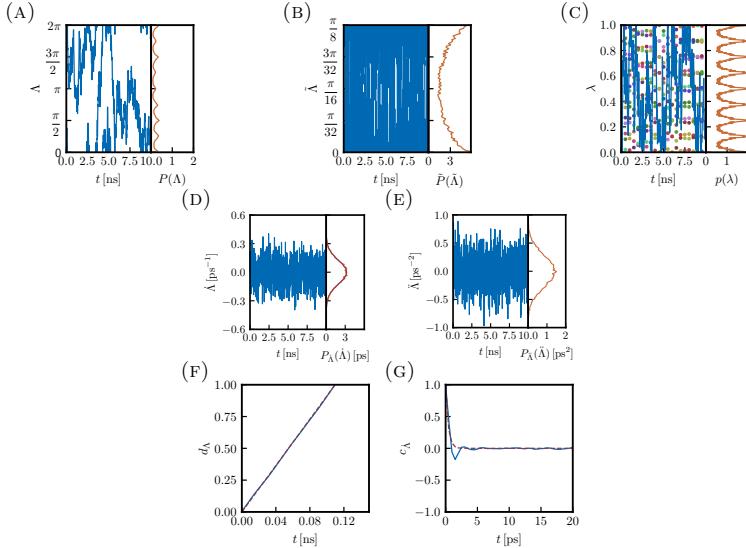
**Simulation #8,  $\tau_\Lambda = 0.5$  ps**


FIGURE 2.F.8: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5$  ps.* This simulation corresponds to entry 8 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

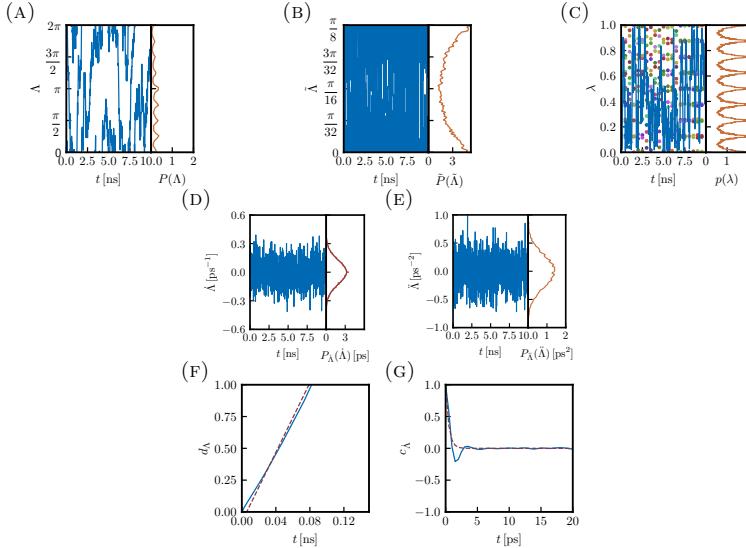
**Simulation #9,  $\tau_\Lambda = 1$  ps**

FIGURE 2.F.9: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 1$  ps. This simulation corresponds to entry 9 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).*

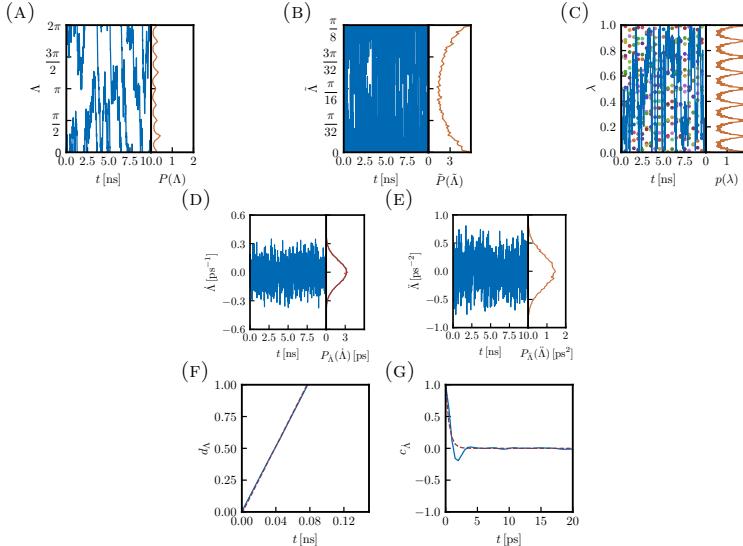
**Simulation #10,  $\tau_\Lambda = 2$  ps**


FIGURE 2.F.10: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 2$  ps.* This simulation corresponds to entry 10 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

### 2.F.3 EXPLORATION OF THE INFLUENCE OF $K$ IN CBTI

**Simulation #11,  $K = 8$**

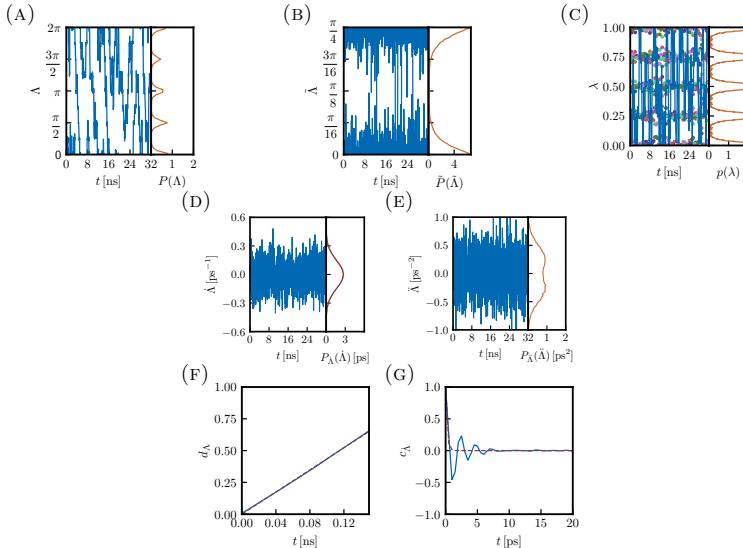


FIGURE 2.F.11: *Results from the CBTI simulation of 32 ns employing  $K = 8$  replicas with a mass-parameter  $m_\Lambda = 113 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 11 in Tab. 3.1.* Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\Lambda$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\Delta$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Delta}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

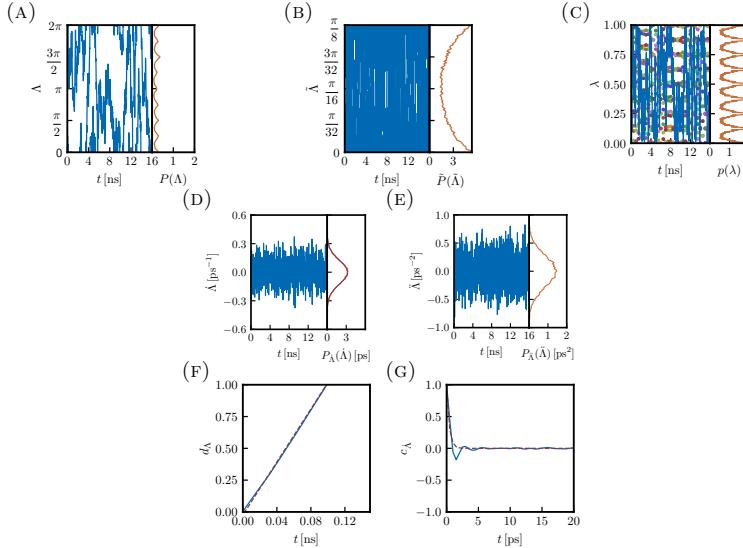
Simulation #12,  $K = 16$ 

FIGURE 2.F.12: *Results from the CBTI simulation of 16 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 12 in Tab. 3.1.* Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

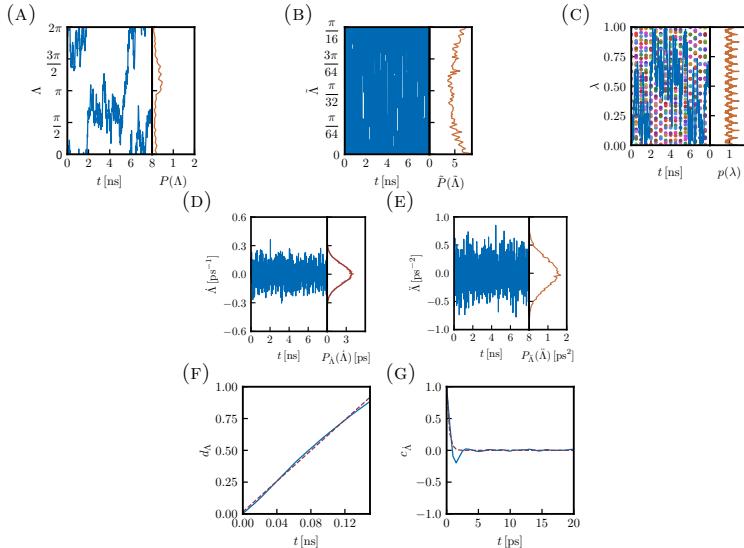
**Simulation #13,  $K = 32$** 

FIGURE 2.F.13: *Results from the CBTI simulation of 8 ns employing  $K = 32$  replicas with a mass-parameter  $m_\Lambda = 226 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 13 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).*

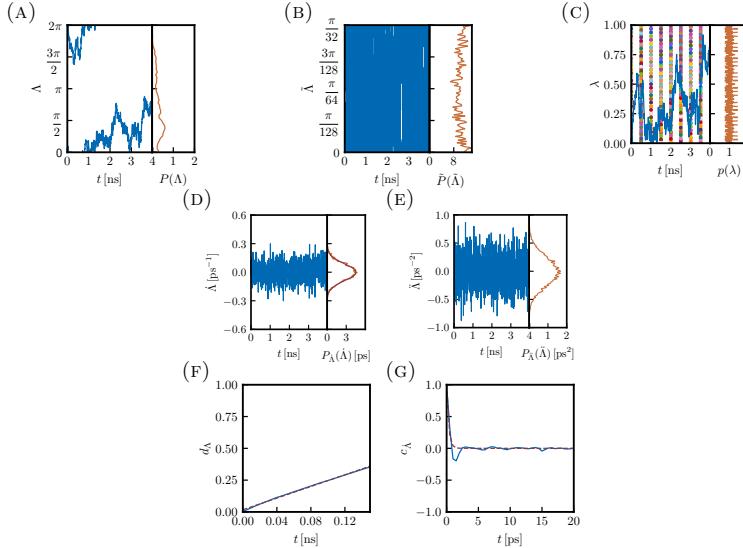
Simulation #14,  $K = 64$ 

FIGURE 2.F.14: *Results from the CBTI simulation of 4 ns employing  $K = 64$  replicas with a mass-parameter  $m_\Lambda = 320 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 14 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).*

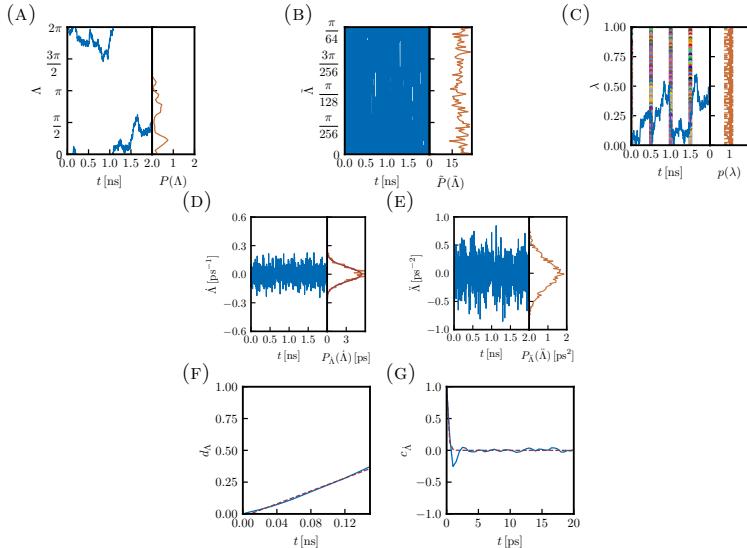
**Simulation #15,  $K = 128$** 

FIGURE 2.F.15: *Results from the CBTI simulation of 2 ns employing  $K = 128$  replicas with a mass-parameter  $m_\Lambda = 452 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 15 in Tab. 3.1.* Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

## Appendix 2.G REPEATS OF A CBTI SIMULATION

TABLE 2.G.1: *Repeat results of a CBTI calculation for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar using the 2061H66 force field.* The free-energy change  $\Delta G$  is reported for the CBTI calculation involving  $K = 16$  replicas along with  $m_A = 160 \text{ u nm}^2$  and  $\tau_A = 0.5 \text{ ps}$ . The results of ten repeats (successive rows) are listed, all of 6.25 ns duration (total single-system sampling time of 100 ns). For each repeat, the calculated  $\Delta G$  (Eq. 3.22 with  $J = 500$ ) is reported along with a bootstrap error estimate. This error does not include a Student  $t$ -factor. Repeat statistics based on this data can be found in Tab. 3.2.

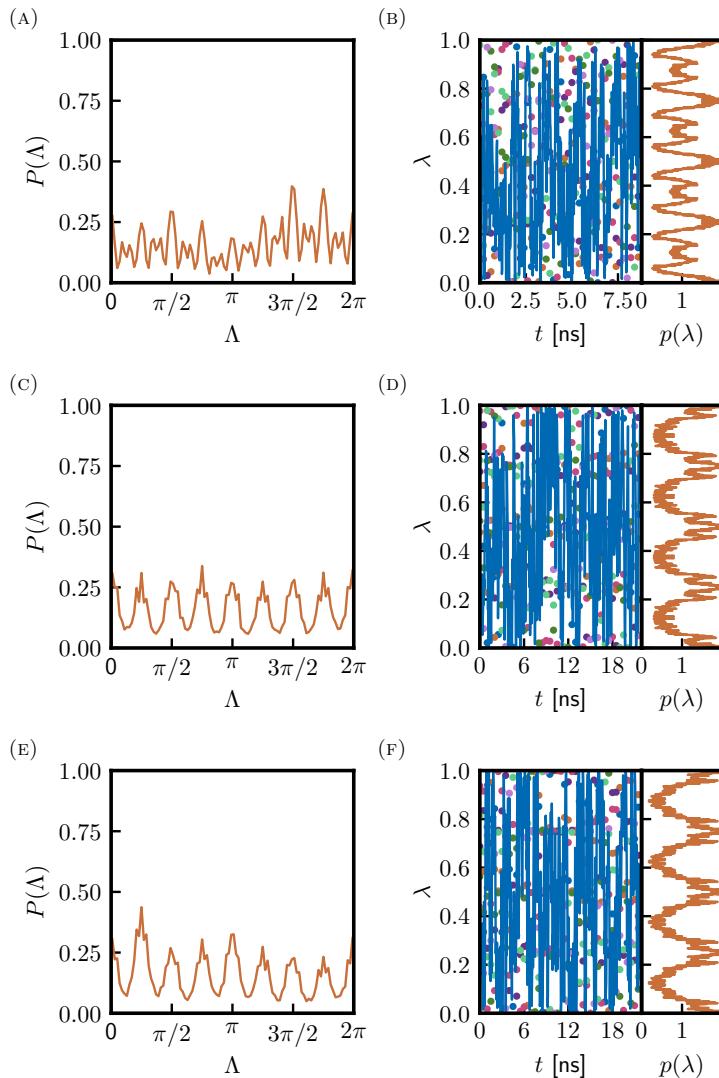
$\Delta G \text{ [kJ mol}^{-1}\text{]}$	
1	21.42 (0.16)
2	21.33 (0.16)
3	21.44 (0.17)
4	21.38 (0.15)
5	21.08 (0.14)
6	21.43 (0.15)
7	21.59 (0.15)
8	21.54 (0.15)
9	21.31 (0.15)
10	21.42 (0.16)

---

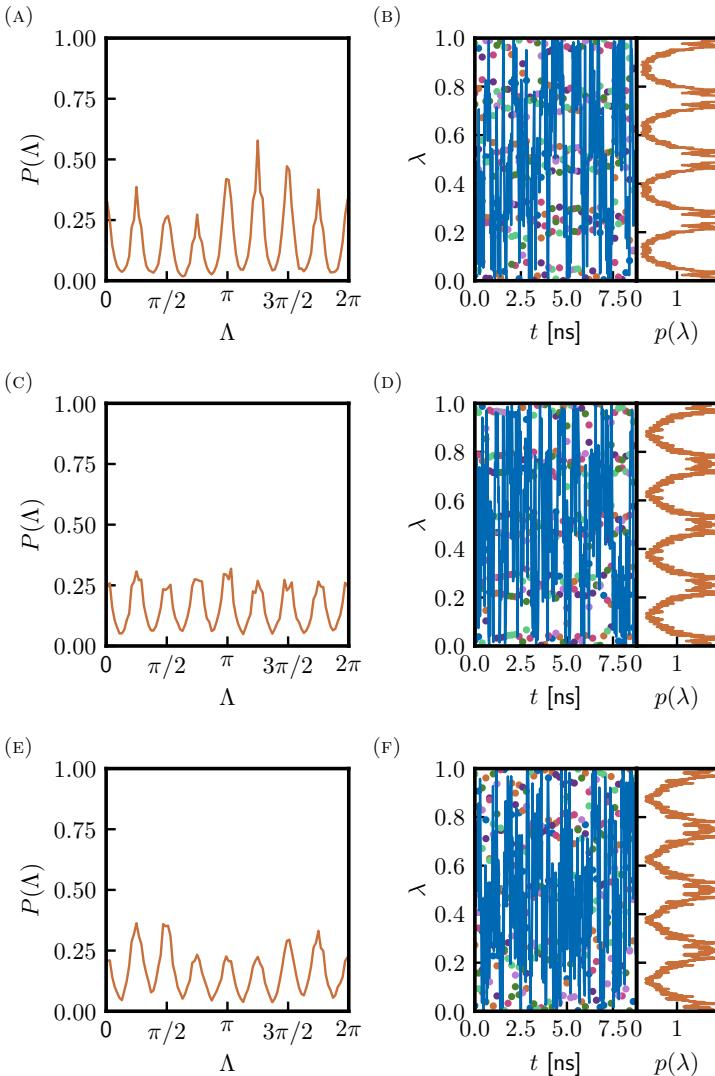
## Appendix 2.H CBTI SIMULATIONS WITH BIASING POTENTIAL

TABLE 2.H.1: *Parameters and results of biased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* The successive entries are the index of the simulation (sim), the number  $K$  of replicas, the number  $N_{\text{gp}}$  of gridpoints for  $\tilde{\Lambda}$  in the range  $[0; 2\pi/K]$ , the build-up force constant  $c_{\text{LE}}$ , the reduction factor  $f_{\text{red}}$ , the LE build-up time  $t_{\text{LE}}$  for the replica system, the number  $N_{\text{ds}}$  of double-sweeps over the  $\tilde{\Lambda}$  range during the build-up, the US umbrella sampling time  $t_{\text{US}}$  for the replica system, and the free-energy difference  $\Delta G$  calculated using Eq. 3.22 with  $J = 500$ . The corresponding results are illustrated graphically in Figs. 2.H.2 and 2.H.4. Only entries 2 and 8 are discussed in Sect. 3.4.

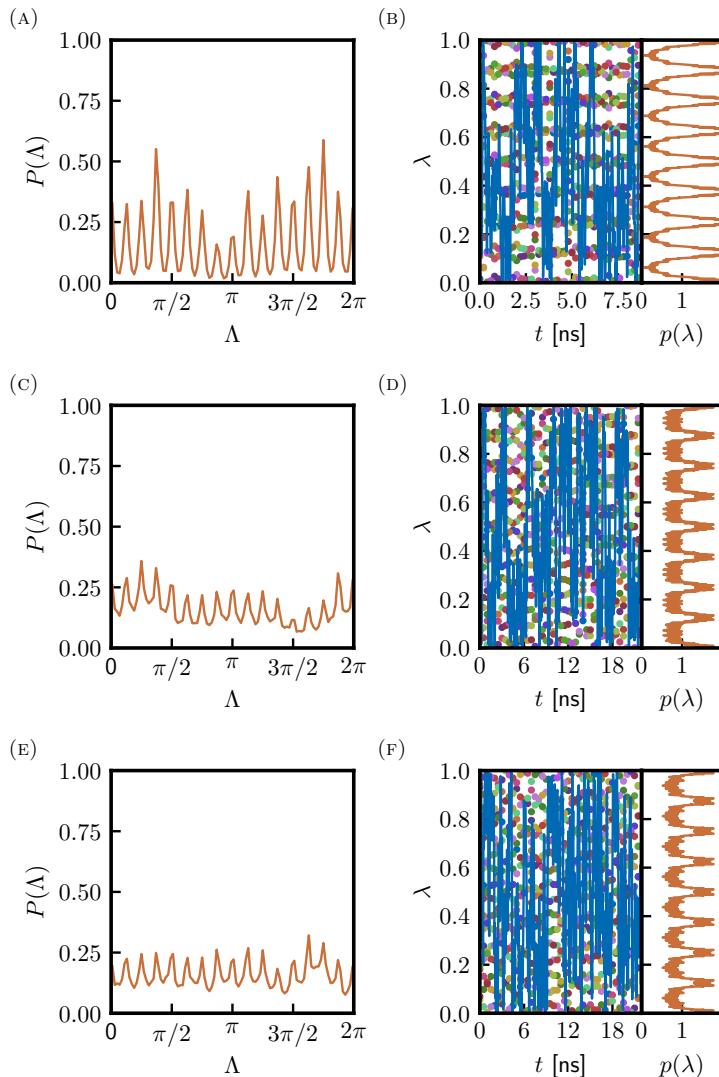
sim	$K$	$N_{\text{gp}}$	$c_{\text{LE}}$ [kJ mol $^{-1}$ ]	$f_{\text{red}}$	$t_{\text{LE}}$ [ns]	$N_{\text{ds}}$	$t_{\text{US}}$ [ns]	$\Delta G$ [kJ mol $^{-1}$ ]
1	8	10	0.01	0.1	0.050	3	22	$21.11 \pm 0.15$
2	8	34	0.001	0.1	0.150	3	22	$21.48 \pm 0.18$
3	8	34	0.001	0.1	0.200	6	22	$21.29 \pm 0.17$
4	8	34	0.001	0.8	0.054	3	22	$21.54 \pm 0.22$
5	8	34	0.001	0.8	0.128	6	22	$21.38 \pm 0.17$
6	8	34	0.001	0.8	0.156	9	22	$21.46 \pm 0.16$
7	16	6	0.01	0.1	0.050	3	22	$21.23 \pm 0.22$
8	16	18	0.001	0.1	0.070	3	22	$21.30 \pm 0.13$
9	16	18	0.001	0.1	0.083	6	22	$21.48 \pm 0.16$
10	16	18	0.001	0.8	0.030	3	10	$21.00 \pm 0.14$
11	16	18	0.001	0.8	0.050	6	20	$21.57 \pm 0.14$
12	16	18	0.001	0.8	0.084	9	20	$21.21 \pm 0.17$



*Continued next page.*



**FIGURE 2.H.2:** Time series and probability distributions of the relevant CB-variables in biased CBTI calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The calculations rely on  $K = 8$  replicas. Each row shows the probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$  (left) and the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas (right). The successive panels correspond to different protocol settings listed in Tab. 2.H.1. Panels (a,b) entry 1. Panels (c,d) entry 2. Panels (e,f) entry 3. Panels (g,h) entry 4. Panels (i,j) entry 5. and Panels (k,l) entry 6.



*Continued next page.*

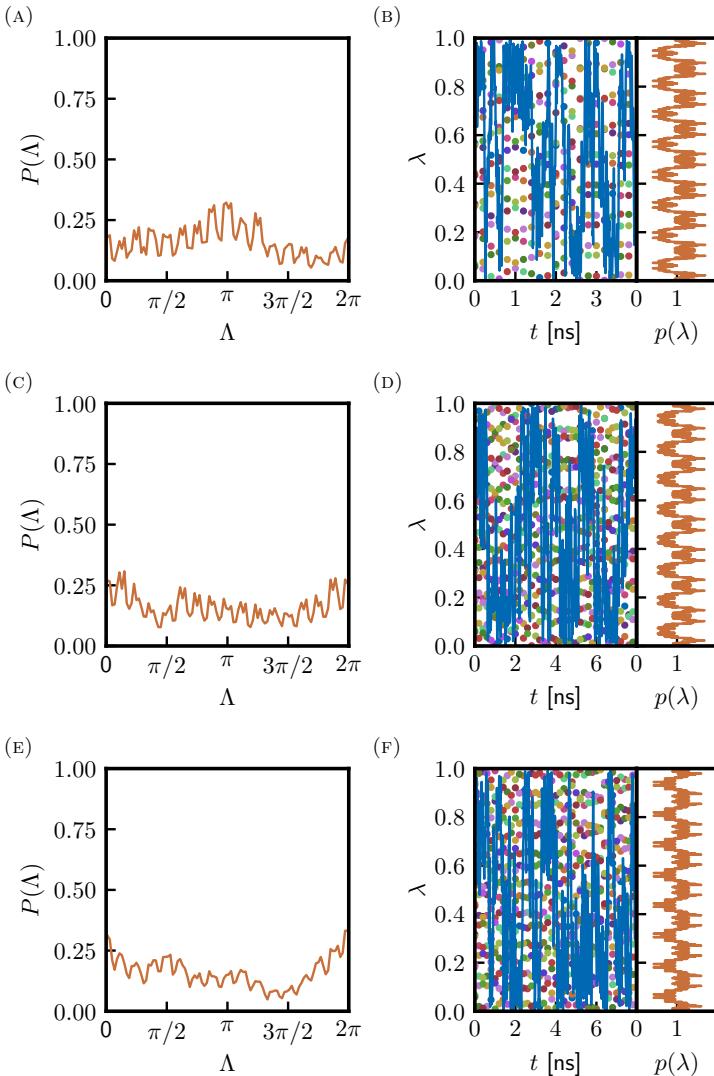


FIGURE 2.H.4: Time series and probability distributions of the relevant CB-variables in biased CBTI calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The calculations rely on  $K = 16$  replicas. Each row shows the probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$  (left) and the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas (right). The successive panels correspond to different protocol settings listed in Tab. 2.H.1. Panels (a,b) entry 7. Panels (c,d) entry 8. Panels (e,f) entry 9. Panels (g,h) entry 10. Panels (i,j) entry 11. and Panels (k,l) entry 12.

## Appendix 2.I FREE-ENERGY PROFILES $G_{\tilde{\Lambda}}(\tilde{\Lambda})$ ALONG $\tilde{\Lambda}$

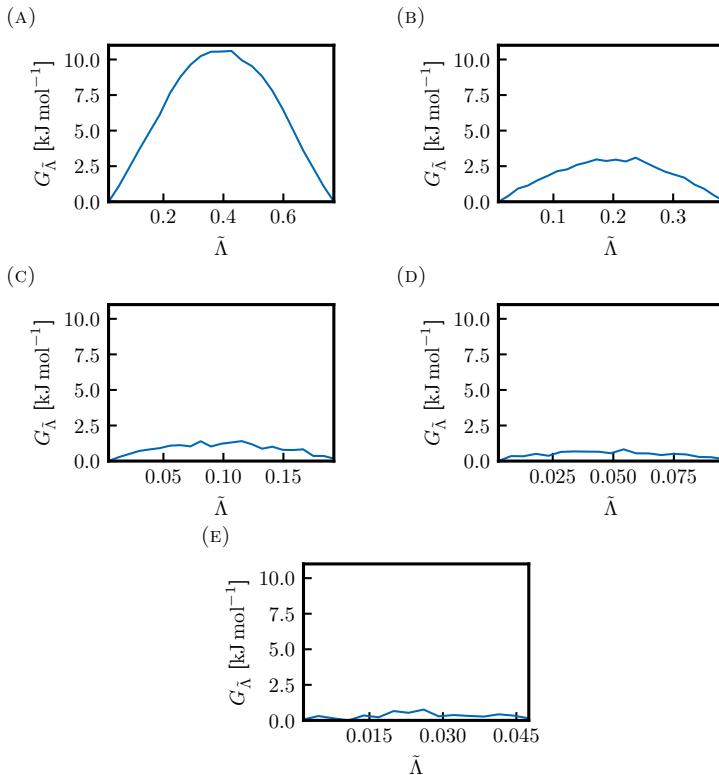


FIGURE 2.I.1: Free-energy profiles  $G_{\tilde{\Lambda}}(\tilde{\Lambda})$  in unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The simulations relied on  $K = 8$  a,  $K = 16$  b,  $K = 32$  c,  $K = 64$  d, and  $K = 128$  e replicas, a mass-parameter  $m_{\Lambda} = 40K^{1/2} \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_{\Lambda} = 0.5$  ps. The free-energy profiles were calculated as  $G_{\tilde{\Lambda}}(\Lambda) = -\beta^{-1} \ln P_{\tilde{\Lambda}}(\tilde{\Lambda})$ , where  $P_{\tilde{\Lambda}}(\tilde{\Lambda})$  is the normalized probability distribution of  $\tilde{\Lambda}$ , and anchored to zero at their minimum. The value at the maximum corresponds to  $G_{\tilde{\Lambda}}^*$ , shown graphically in Fig. 3.6.

## Appendix 2.J HRE SIMULATIONS

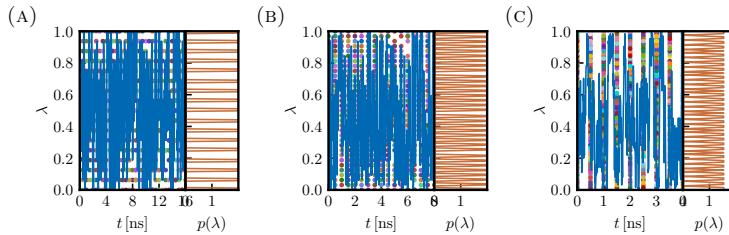


FIGURE 2.J.1: Time series  $\lambda(t)$  and probability distributions  $p(\lambda)$  of the coupling variable  $\lambda$  in HRE simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The time series is shown as a blue curve for replica  $k = 0$ , and as individual colored points at 0.5 ns interval for the  $K_{\text{HRE}} - 1$  other replicas. Panel a:  $K_{\text{HRE}} = 17$ . Panel b:  $K_{\text{HRE}} = 33$ . Panel c:  $K_{\text{HRE}} = 65$ .

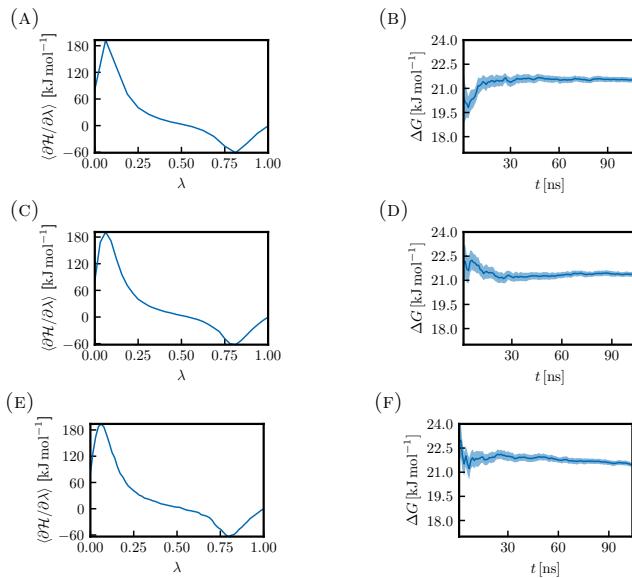


FIGURE 2.J.2: *Relevant results from HRE calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* Panels a, c and e show the Hamiltonian derivative curve considering a total single-system sampling time of 100 ns. Panel a:  $K_{\text{HRE}} = 17$ . Panel c:  $K_{\text{HRE}} = 33$ . Panel e:  $K_{\text{HRE}} = 65$ . Panels b, d and f show the Hamiltonian derivative curve considering a total single-system sampling time of 100 ns. Panel b:  $K_{\text{HRE}} = 17$ . Panel d:  $K_{\text{HRE}} = 33$ . Panel f:  $K_{\text{HRE}} = 65$ .

## Appendix 2.K TI/EXTI CALCULATIONS

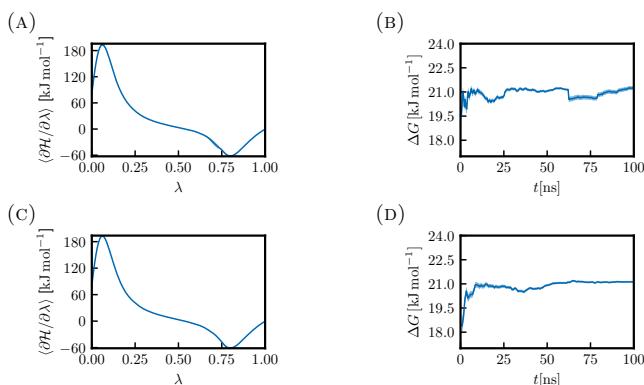


FIGURE 2.K.1: *Relevant results from TI/EXTI calulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* Panels a and c show the Hamiltonian derivative curve, which was predicted at 129  $\lambda$ -points. Panel a:  $K_{\text{TI}} = 9$ . Panel c:  $K_{\text{TI}} = 17$ . Panels b and d show the convergence of the  $\Delta G$  value dependent on the total single-system sampling time. Panel b:  $K_{\text{TI}} = 9$ . Panel d:  $K_{\text{TI}} = 17$ .

## Appendix 2.L TI/MBAR CALCULATIONS

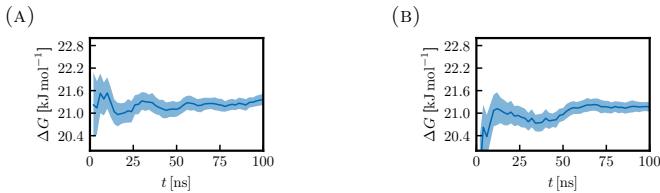


FIGURE 2.L.1: *Convergence properties of TI/MBAR calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* Panels b and d show the convergence of the  $\Delta G$  value dependent on the total single-system sampling time. Panel a:  $K_{\text{TI}} = 9$ . Panel b:  $K_{\text{TI}} = 17$ .



# Alchemical Free-Energy Calculations by Multiple-replica $\lambda$ -dynamics: The Conveyor Belt Thermodynamic Integration

# 3

*“Let us learn to dream, gentlemen, and then perhaps we shall learn the truth.”*

---

August Kekulé, 1865

A new method is proposed to calculate alchemical free-energy differences based on molecular dynamics (MD) simulations, called the conveyor belt thermodynamic integration (CBTI) scheme. As in thermodynamic integration (TI),  $K$  replicas of the system are simulated at different values of the alchemical coupling parameter  $\lambda$ . The number  $K$  is taken to be even and the replicas are equally spaced on a forward-turn-backward-turn path, akin to a conveyor belt (CB) between the two physical end-states. And as in  $\lambda$ -dynamics ( $\lambda$ D), the  $\lambda$ -values associated with the individual systems evolve

in time along the simulation. However, they do so in a concerted fashion, determined by the evolution of a single dynamical variable  $\Lambda$  of period  $2\pi$  controlling the advance of the entire CB. Thus, a change of  $\Lambda$  is always associated with  $K/2$  equispaced replicas moving forward and  $K/2$  equispaced replicas moving backward along  $\lambda$ . As a result, the effective free-energy profile of the replica system along  $\Lambda$  is periodic of period  $2\pi K^{-1}$  and the magnitude of its variations decreases rapidly upon increasing  $K$ , at least as  $K^{-1}$  in the limit of large  $K$ . When a sufficient number of replicas is used, these variations become small, which enables a complete and quasi-homogeneous coverage of the  $\lambda$ -range by the replica system, without application of any biasing potential. If desired, a memory-based biasing potential can still be added to further homogenize the sampling, the preoptimization of which is computationally inexpensive. The final free-energy profile along  $\lambda$  is calculated similarly to TI, by binning of the Hamiltonian  $\lambda$ -derivative as a function of  $\lambda$  considering all replicas jointly, followed by quadrature integration. The associated quadrature error can be kept very low owing to the continuous and quasi-homogeneous  $\lambda$ -sampling. The CBTI scheme can be viewed as a continuous/deterministic/dynamical analog of the Hamiltonian replica-exchange/permutation (HRE/HRP) schemes, or as a correlated multiple-replica analog of the  $\lambda$ D or  $\lambda$ -local elevation umbrella sampling ( $\lambda$ -LEUS) schemes. Compared to TI, it shares the advantage of the latter schemes in terms of enhanced orthogonal sampling, *i.e.* the avail-

ability of variable- $\lambda$  paths to circumvent conformational barriers present at specific  $\lambda$ -values. Compared to HRE/HRP, it permits a deterministic and continuous sampling of the  $\lambda$ -range, and bypasses the need to carefully preselect a  $\lambda$ -ladder and a swapping-attempt frequency. Compared to  $\lambda$ -LEUS, it eliminates (or drastically reduces) the dead time associated with the preoptimization of a biasing potential. The goal of this chapter is to provide the mathematical/physical formulation of the proposed CBTI scheme, along with an initial application of the method to the calculation of the hydration free energy of methanol.

### 3.1 INTRODUCTION

Classical molecular dynamics (MD) simulations provide insight into (bio-)molecular systems at atomic resolution, thereby explaining and complementing experimental observations. This often involves the calculation of free-energy differences,<sup>?</sup> <sup>?</sup> <sup>?</sup> which characterize the relative stabilities of two or more macroscopic states of the system. These states may differ thermodynamically (different pressures, temperatures or numbers of molecules), conformationally (different regions in a space spanned by a set of specific generalized coordinates) or alchemically (different Hamiltonian functions).

Alchemical free-energy calculations involve atom mutations or interaction alterations that have no experimental counterpart. However, by comparing the results of two such calculations in different environments (*e.g.* mutation of a molecule into another one in vacuum or in a solvent) *via* a thermodynamic cycle,<sup>?</sup> <sup>?</sup> the calculated values can be converted to experimentally accessible differences (*e.g.* relative solvation free energies of the two molecules in the given solvent). This indirect pathway *via* a cycle typically offers a strong sampling advantage relative to the calculation over a direct conformational path (*e.g.* reversibly displacing the two molecules from vacuum into the solvent across the liquid surface), while giving the same result at full convergence.<sup>?</sup> <sup>?</sup> <sup>?</sup> Over the last decades, numerous methods have been proposed to calculate alchemical free-energy differences between states  $A$ ,  $B$ ,  $C$ , ... of a molecular system involving the same numbers of atoms but different Hamiltonian functions. They can be roughly classified as reference-state methods and pathway-dependent methods.<sup>?</sup>

In reference-state methods, a single simulation is performed at

a reference state  $R$  and the relative free energy of a target state  $A$  is calculated using one-step free-energy perturbation<sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> (OSP) as a free-energy estimator. The state  $R$  may be unphysical, in which case the free-energy difference between two physical states  $A$  and  $B$  is obtained by comparing the results of two such calculations ( $R$  to  $A$  and  $R$  to  $B$ ). The accuracy of the method, *i.e.* its convergence at finite sampling times, depends crucially on the extent of Boltzmann-weighted phase-space overlap between the reference and target states,<sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> <sup>?</sup> *i.e.* whether configurations relevant for  $A$  are well sampled in  $R$ . Methods to enhance this overlap include in particular: (i) the use of a reference state with soft-core<sup>?</sup> (SC) sites<sup>?</sup> <sup>?</sup> <sup>?</sup> or other softened force-field terms;<sup>?</sup> <sup>?</sup> <sup>?</sup> (ii) the construction of a reference state encompassing all the target states, as in enveloping distribution sampling<sup>?</sup> <sup>?</sup> <sup>?</sup> (EDS); (iii) the extension to the use of multiple reference states,<sup>?</sup> along with the application of the (multi-state<sup>?</sup>) Bennet acceptance ratio<sup>?</sup> <sup>?</sup> (BAR or MBAR) estimator instead of OSP. In principle, reference-state methods bear the promise of enabling the extrapolative calculation of the relative free energies of numerous arbitrary states ( $A$ ,  $B$ ,  $C$ , ...) based on a single reference-state simulation. In practice, however, they seldom hold this promise (or not in a sufficiently robust fashion based on finite simulations), because the design of a suitable reference state requires considerable (*e.g.* SC approach) or even complete (*e.g.* EDS approach) *a priori* knowledge of the target states it has to be appropriate for.

In pathway-dependent methods, a hybrid Hamiltonian is constructed by employing a coupling parameter  $\lambda$  that defines a continuous transformation between the Hamiltonians of the physical end-states  $A$  and  $B$ . Given such a path, the most established, robust and still frequently used method is multi-configuration<sup>?</sup>

thermodynamic integration<sup>???</sup> (MCTI or, simply, TI). In TI, a set of independent simulations are performed at different constant  $\lambda$ -values, and the ensemble average of the derivative of the hybrid Hamiltonian with respect to  $\lambda$  is subsequently integrated by numerical quadrature<sup>???</sup> or curve fitting.<sup>???</sup> A common alternative to TI is multi-configuration<sup>?</sup> free-energy perturbation<sup>?????????</sup> (MC-FEP or, simply, FEP), where the OSP estimator is used to evaluate the free-energy difference from one  $\lambda$ -point to the previous or/and next one. Refinements and improvements of the original TI protocol include in particular: (i) the design of Hamiltonian coupling schemes leading to high sampling efficiencies;<sup>???</sup> (ii) the partial automation of the TI protocol,<sup>???</sup> *i.e.* of the selection of  $\lambda$ -points along with associated initial configurations, equilibration times and sampling times; (iii) the use of free-energy estimators with improved statistical efficiencies over plain quadrature, *e.g.* extended TI (EXTI) estimator<sup>?</sup> or MBAR estimator;<sup>?</sup> (iv) the design of alternative single- or multiple-replica schemes where the  $\lambda$ -values are no longer fixed during the simulation.<sup>???</sup>

Although the possibility of changing the  $\lambda$ -values over the course of a simulation may appear at first sight to represent an unnecessary complication of the TI protocol, it often leads to a significant enhancement of the sampling efficiency. This is because it increases the likelihood of crossing barriers in the orthogonal space,<sup>??</sup> that is, the space spanned by all the degrees of freedom of the system excluding  $\lambda$  (*i.e.* all the conformational ones). These orthogonal barriers are typically higher at certain  $\lambda$ -values than at others. The highest ones may be seldom crossed in corresponding simulations at fixed  $\lambda$ , and allowing for variations of  $\lambda$  may open pathways to circumvent them. Note that, in many cases, this enhancement is not sufficient *per se* and other techniques must

be applied to further improve the orthogonal sampling along the  $\lambda$ -path.<sup>???</sup> There exist two main routes for performing variable- $\lambda$  simulations: (*i*) Hamiltonian replica exchange (HRE) or permutation (HRP), which involve multiple system replicas; (*ii*)  $\lambda$ -dynamics ( $\lambda$ D) or  $\lambda$ -Monte Carlo ( $\lambda$ MC), which consider a single system.

In the HRE scheme,<sup>???</sup> a series of system copies (walkers) are distributed over a set of fixed  $\lambda$ -values (replicas), as is the case in TI. However, at regular time intervals, swaps are attempted between pairs of systems corresponding to different  $\lambda$ -values (typically adjacent ones). These attempts are accepted or rejected according to a Metropolis-Hastings criterion<sup>??</sup> depending on the Boltzmann factor ratio of the replica system before and after the swap. Although the trajectories at each  $\lambda$ -point become discontinuous, the TI-like statistics is preserved and the data can be analyzed in the same way as for TI. Recent extensions of the method involve in particular the consideration of more advanced exchange schemes,<sup>???</sup> of replica reservoirs,<sup>???</sup> of frozen replicas,<sup>??</sup> of heating-quenching steps between the sampling periods,<sup>???</sup> of the infinite-swapping limit,<sup>???</sup> and of generalized-ensemble distributions.<sup>???</sup> They also include the implementation of  $\lambda$ -moves that go beyond pairwise swaps with a selection based on a Suwa-Todo criterion,<sup>???</sup> as implemented in the HRP method.<sup>???</sup> In the latter case, enabling arbitrary permutations and abandoning the detailed-balance constraint leads to a significant increase in the probability of exchange acceptance.

In the  $\lambda$ D scheme<sup>???</sup> (see also its ancestor  $\lambda$ MC), a single system is considered for which the  $\lambda$ -value evolves dynamically in the course of the simulation, *i.e.*  $\lambda$  is treated as an extra pseudo-conformational degree of freedom with an assigned

mass parameter  $m_\lambda$  and momentum  $p_\lambda$ . This momentum enters an extended Hamiltonian that includes an additional term for the corresponding kinetic energy. This results in a continuous sampling of the  $\lambda$ -range instead of the discrete sampling underlying TI or HRE/HRP. The free-energy difference is then typically estimated from the probability distribution along  $\lambda$ , with the drawback of requiring a threshold to define the end-states.<sup>?</sup> This issue can be alleviated by introducing a coordinate transformation with plateaus,<sup>??</sup> or by using a TI-like formula<sup>???</sup> or a Rao-Blackwell estimator,<sup>?</sup> the latter conceptually similar to MBAR.

The major advantages of  $\lambda$ D compared to HRE/HRP are that it involves a simpler single-system setup, is deterministic, samples the  $\lambda$ -range continuously, and does not require the specification (and optimization<sup>?????????????????</sup>) of a  $\lambda$ -ladder and of a swapping-attempt frequency. The main drawback<sup>?</sup> is that the sampling probability along  $\lambda$  is no longer imposed in the form of a fixed set of  $\lambda$ -points with equal sampling times, but entirely controlled by the free-energy profile along  $\lambda$ . As a result,  $\lambda$ -values with high relative free energies may be poorly represented (or not at all) and  $\lambda$ -barriers with high relative free energies may be seldom crossed (or not at all) in the course of a finite simulation. In addition, care must be taken to avoid the sampling of  $\lambda$ -values beyond the physical end-states of the alchemical coupling<sup>?</sup> (*i.e.* below 0 or above 1). Both the sampling inhomogeneity and the end-point issues can in principle be remedied<sup>?</sup> by employing an appropriate coordinate transformation<sup>?????????</sup> or/and by applying a biasing potential along  $\lambda$ .<sup>???????</sup> A combination of these two approaches underlies the  $\lambda$ -local elevation umbrella sampling method<sup>???</sup> ( $\lambda$ -LEUS), which relies in particular on an adaptive memory-based biasing potential. A similar principle is also at the heart of numerous other

methods such as the flat-histogram,<sup>??</sup>  $\lambda$ -metadynamics,<sup>???</sup> adaptive integration,<sup>?</sup> adaptive biasing force,<sup>?</sup> adaptively biased<sup>?</sup> and expanded-ensemble<sup>???????</sup> methods.

In  $\lambda$ -LEUS, a local elevation<sup>?</sup> (LE) build-up phase is used to construct a suitable biasing potential, and followed by an umbrella sampling<sup>??</sup> (US) phase where this potential is frozen<sup>?</sup> and biased equilibrium statistics gathered with quasi-homogeneous sampling of the  $\lambda$ -range. Clearly, the LE phase represents an efficiency loss in the method, *i.e.* a dead time. Note, however, that a similar (but generally shorter) dead time also exists in TI and HRE/HRP in the form of discarded initial equilibration times for all replicas. Another drawback of  $\lambda$ -LEUS is that it is not parallelizable in the same fashion as TI and HRE/HRP, where the simulations at different  $\lambda$ -points can be carried out in parallel (see, however, various swarm,<sup>???</sup> multiple-walker<sup>??</sup> and flying Gaussian<sup>??</sup> variants of memory-based biasing methods). Finally, some care must be taken to select an appropriate mass and thermostat-coupling scheme for the  $\lambda$ -variable, so as to ensure that this variable is adequately coupled to the configurational degrees of freedom.<sup>?</sup>

In the present chapter, we propose a new alchemical free-energy calculation scheme with the goal of combining the advantages and alleviating the shortcomings of both HRE/HRP and  $\lambda$ D. This scheme is termed conveyor belt thermodynamic integration (CBTI) and relies on the coupled  $\lambda$ D of a set of system replicas, in which the  $\lambda$ -distance between successive replicas is kept fixed along a forward-turn-backward-turn path, akin to a conveyor belt between the two physical end-states. The basic principle of CBTI is illustrated schematically in Fig. 3.1.

Considering a free-energy profile  $G(\lambda)$  presenting a constant uphill slope between  $A$  and  $B$ , a single system subjected to  $\lambda$ D

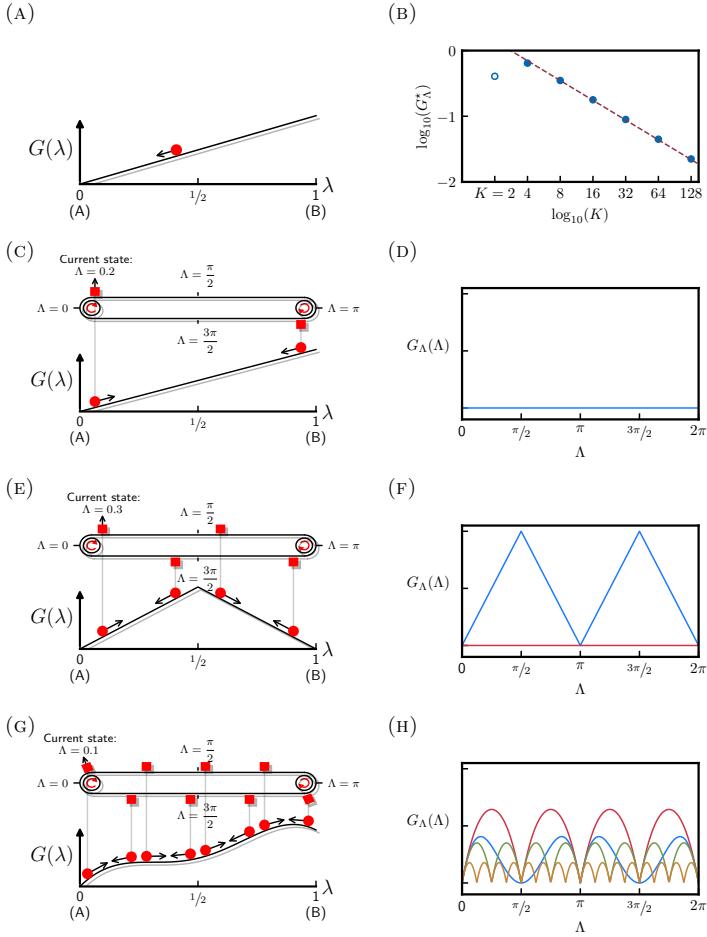


FIGURE 3.1: *Schematic illustration of the conveyor belt thermodynamic integration (CBTI) approach for different free-energy profiles  $G(\lambda)$  and numbers of replicas  $K$ .* The free-energy profile  $G(\lambda)$  is shown on the left along with the conveyor belt for an inclined plane (c), a piecewise-linear curve (e) and a more realistic curve (g), the latter curve corresponding to the function  $G(\lambda) = \sin(\pi\lambda)(\lambda - 0.5)^2 + 0.25\lambda$ . The corresponding free-energy profiles  $G_\Lambda(\Lambda)$  along the conveyor belt advance variable  $\Lambda$  (Eq. 3.16) are shown on the right for  $K = 2$  (blue),  $K = 4$  (red),  $K = 8$  (green) or  $K = 16$  (orange). The top-left panel (a) illustrates the situation of plain unbiased  $\lambda D$ , where the system would “roll down” the slope and keep sampling the neighborhood of the state with the lowest free energy. The top-right panel (b) shows the height  $G_\Lambda^*$  of the residual barriers in the free-energy profile  $G_\Lambda(\Lambda)$  considering this illustrative function and increasing values of  $K$ . The data is represented in logarithmic form and a linear line of slope  $-1$  is fitted to the filled circles.

in the absence of a biasing potential will “roll down” to  $A$  and keep sampling the neighborhood of this state (Fig. 3.1a). To circumvent this problem, one may decide to couple the  $\lambda D$  of two replicas 0 and 1 of the system in such a way that any downhill displacement of replica 0 implies an equivalent uphill displacement of replica 1 (Fig. 3.1c). This working principle is exploited in practice by vehicles like cable cars or funiculars, where the motion of the two vehicles are controlled by a stretched cable connected to two pulleys. It has also been exploited previously for MD in the context of twin-system<sup>??</sup> EDS, which couples forward and reverse alchemical changes performed in two different environments. In this situation, one may describe the  $\lambda$ -values of the two replicas by means of a single periodic angular variable  $\Lambda$  representing the advance of the cable from 0 (replica 0 in state  $A$ , replica 1 in state  $B$ ) to  $\pi$  (0 in  $B$ , 1 in  $A$ ) and then  $2\pi$  (back to the starting situation). Note that the  $\pi$  to  $2\pi$  return range of  $\Lambda$  deviates from the cable car or funicular analogy, where the vehicles never go “over the pulley”. If the free-energy profile  $G(\lambda)$  has a constant slope, it is easily seen that there will be no net driving force on the cable. In the  $\lambda D$  context, this means that the  $\Lambda$ -variable will undergo random diffusion with a homogeneous sampling of  $\Lambda$ , *i.e.* the corresponding free-energy profile  $G_\Lambda(\Lambda)$  will be flat (Fig. 3.1d). Accordingly, each of the two replicas will sample the entire  $\lambda$ -range homogeneously.

Consider now a free-energy profile  $G(\lambda)$  presenting a constant uphill slope from 0 to  $1/2$  and a constant downhill slope from  $1/2$  to 1 (Fig. 3.1e). In a setup with two replicas, the  $G_\Lambda(\Lambda)$  profile will no longer be flat (Fig. 3.1f, blue curve). From 0 to  $\pi/2$  and from  $\pi$  to  $3\pi/2$ , both replicas move uphill, whereas from  $\pi/2$  to  $\pi$  and from  $3\pi/2$  to  $2\pi$ , they both move downhill. This can be remedied by using four instead of only two replicas, placed at equal

distances along the cable. This working principle is now more reminiscent of the real-life situation of a conveyor belt. In this four-replica setup, the  $G_\Lambda(\Lambda)$  profile will again be flat (Fig. 3.1f, red curve), and each of the four replicas will sample the entire  $\lambda$ -range homogeneously.

Finally, consider a more realistic free-energy profile  $G(\lambda)$  (Fig. 3.1g). In the general case, the  $G_\Lambda(\Lambda)$  profile will never be exactly flat irrespective of the number  $K$  of replicas. However, by adding more and more replicas to the conveyor belt, the features of the  $G_\Lambda(\Lambda)$  profile can be progressively reduced in magnitude (as illustrated for  $K = 2, 4, 8$  or  $16$  replicas by the blue, red, green and orange curves of Fig. 3.1h). Note that we restrict the choice of  $K$  to even values, in order to have always the same number of replicas moving forward and backward. Although the choice of an odd number of replicas would be acceptable, it is likely to be less favorable (especially at small  $K$ ), because one extra replica would always move in either of the two directions.

Considering the illustrative  $G(\lambda)$  of Fig. 3.1g, the magnitude  $G_\Lambda^*$  of the variations in  $G_\Lambda(\Lambda)$  is shown in Fig. 3.1b as a function of  $K$ . As discussed in Sect. 3.A, these variations can be interpreted as the residual of a  $K$ -point trapezoid quadrature approximation to the vanishing integral of the derivative of an even periodic function over one period, namely that of the  $G(\lambda)$  profile after mirroring. Quantitatively, this interpretation shows that the magnitude  $G_\Lambda^*$  of these variations decreases at least as  $K^{-1}$  in the limit of large  $K$ . Note that the convergence to a flat profile (*i.e.* the decrease of the barrier heights towards zero upon increasing  $K$ ) does not need to be regular for small  $K$  (see *e.g.* the comparatively low variations for the blue curve in Fig. 3.1h) and that it can be stronger than  $K^{-1}$  for large  $K$  if  $G(\lambda)$  presents particular continuity/symmetry properties. By using a sufficient number of replicas, one may

thus ensure a quasi-homogeneous sampling of the  $\lambda$ -range by each replica, even in the absence of coordinate transformation or biasing potential. If desired, a memory-based biasing potential may still be applied to further homogenize the sampling. However, the LE build-up time can be considerably reduced relative to that needed in a corresponding single-system  $\lambda$ -LEUS simulation, because this biasing potential only needs to be applied to the  $\Lambda$ -variable and no longer to the  $\lambda$ -variable. Thus, it has to compensate for comparatively small  $G_\Lambda(\Lambda)$  variations. Furthermore, it only needs to be constructed over a limited  $\Lambda$ -range, considering that  $G_\Lambda(\Lambda)$  is periodic with period  $2\pi K^{-1}$  as well as even over this  $2\pi K^{-1}$  interval.

Irrespective whether a biasing potential is employed or not, the definition of a free-energy estimator for the CBTI scheme, *i.e.* a procedure to construct  $G(\lambda)$  based on the simulation results for the multiple-replica system, is not as trivial<sup>?</sup> as in the single-system  $\lambda$ D or  $\lambda$ -LEUS cases. Here, CBTI is analyzed using a TI-like estimator<sup>?</sup> relying on the quadrature integration of the average Hamiltonian  $\lambda$ -derivative binned along  $\lambda$  considering all replicas simultaneously. The associated quadrature error can be kept very low by using a large number of bins, which is rendered possible by the continuous and quasi-homogeneous  $\lambda$ -sampling. Although this estimator is very robust, we note that it may not be optimal in terms of statistical efficiency.<sup>???</sup>

In the present chapter, we provide the mathematical/physical formulation of the proposed CBTI scheme, and report an initial application of the method to an illustrative alchemical transformation: the mutation of a methanol molecule to a dummy (non-interacting) skeleton in water, giving access to the hydration free energy of the molecule.

## 3.2 THEORY

### 3.2.1 ALCHEMICAL FREE-ENERGY CALCULATIONS

The goal of path-dependent alchemical free-energy calculations is to evaluate the free-energy difference  $\Delta G$  between two states  $A$  and  $B$  of a molecular system, by introducing a coupling scheme relying on a parameter  $\lambda$ , and sampling along the so-defined  $\lambda$ -path. The two states have the same number  $3N$  of degrees of freedom, but distinct Hamiltonian functions  $\mathcal{H}_A(\mathbf{x})$  and  $\mathcal{H}_B(\mathbf{x})$ , respectively, where  $\mathbf{x} = (\mathbf{r}, \mathbf{p})$  is the  $6N$ -dimensional phase-space vector representative of a microscopic system configuration,  $\mathbf{r}$  and  $\mathbf{p}$  being the corresponding coordinate and momentum vectors. The coupling parameter is introduced into a hybrid Hamiltonian  $\mathcal{H}(\mathbf{x}; \lambda)$  satisfying the boundary conditions  $\mathcal{H}(\mathbf{x}; 0) = \mathcal{H}_A(\mathbf{x})$  and  $\mathcal{H}(\mathbf{x}; 1) = \mathcal{H}_B(\mathbf{x})$ , where the semi-colon indicates a parametric dependence.

Since the proposed CBTI scheme encompasses features of both  $\lambda$ D and TI, these two approaches are summarized briefly in the next two subsections. The following three subsections then describe in turn the basis of the CBTI scheme, its free-energy estimator and the application of a biasing potential.

### 3.2.2 $\lambda$ -DYNAMICS ( $\lambda$ D)

In the  $\lambda$ D scheme,  $\lambda$  is assigned a mass  $m_\lambda$  and a momentum  $p_\lambda$ , and considered to be an additional pseudo-conformational degree of freedom of the system.

The Hamiltonian of the extended system is defined as

$$\mathcal{H}^*(\mathbf{x}, \lambda) = \mathcal{H}(\mathbf{x}; \lambda) + \frac{p_\lambda^2}{2m_\lambda}, \quad (3.1)$$

where the star refers to an extended system in which  $\lambda$  is now a variable and no longer a parameter (thus the replacement of the semi-colon by a comma). This leads to the additional equation of motion

$$\ddot{\lambda} = \frac{\dot{p}_\lambda}{m_\lambda} = -\frac{1}{m_\lambda} \frac{\partial \mathcal{H}(\mathbf{x}; \lambda)}{\partial \lambda} \quad (3.2)$$

for propagating the  $\lambda$ -variable, where a dot over a variable indicates its time derivative.

The free-energy difference  $\Delta G$  between the two physical end-states can then in principle be calculated based on a single thermostated MD simulation of the extended system, as

$$\Delta G = -\frac{1}{\beta} \ln \frac{\langle \delta(\lambda - 1) \rangle^*}{\langle \delta(\lambda) \rangle^*}, \quad (3.3)$$

where  $\beta = (k_B T)^{-1}$ ,  $k_B$  being the Boltzmann constant and  $T$  the absolute temperature,  $\delta$  is the Dirac delta function, and  $\langle \dots \rangle^*$  denotes ensemble averaging for the extended system (*i.e.* over the joint trajectories of  $\mathbf{x}$  and  $\lambda$ ). In practice, the  $\delta$ -functions in Eq. 3.3 must be replaced by two finite end-state bins ( $\lambda$ -cutoff), sufficiently large for proper statistics but also sufficiently small for avoiding distortions due to averaging at the end-states.<sup>?</sup>

### 3.2.3 THERMODYNAMIC INTEGRATION (TI)

In the original TI scheme,<sup>???</sup> a set of  $K$  replicas of the system are simulated in parallel at fixed predefined  $\lambda$ -values in the range

$[0, 1]$ . Since the replicas are entirely decoupled from each other, the simulations can be performed serially as well. However, TI extensions including the HRE scheme<sup>???</sup> and the HRP scheme<sup>???</sup> introduce a coupling in the form of  $\lambda$ -value exchanges, in which case the simulations must really be carried out in parallel. The same will apply to the proposed CBTI scheme, where the coupling involves a synchronization of the dynamical  $\lambda$ -variations.

Considering all replicas  $k = 0 \dots K - 1$  as the members of a replica system, one may note the corresponding  $6K \times N$ -dimensional phase-space vector as  $\mathbf{X} = \{\mathbf{x}_k\}$  and the corresponding  $K$ -dimensional vector containing the fixed  $\lambda$ -values as  $\boldsymbol{\lambda} = \{\lambda_k\}$ . In plain TI, the Hamiltonian of the replica system is defined as

$$\mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda}) = \sum_{k=0}^{K-1} \mathcal{H}(\mathbf{x}_k; \lambda_k) , \quad (3.4)$$

where the dagger refers to a replica system, and  $\boldsymbol{\lambda}$  is here a parameter vector (thus the semi-colon). Because the Hamiltonian of Eq. 3.4 involves no coupling term between the replicas, the dynamics of a replica  $k$  is independent from that of the other replicas and solely depends on  $\lambda_k$ .

The free energy difference  $\Delta G$  between the two states can then be calculated based on a single thermostated MD simulation of the replica system, as

$$\begin{aligned} \Delta G &= \int_0^1 d\lambda' \left\langle \frac{\partial \mathcal{H}(\mathbf{x}; \lambda)}{\partial \lambda} \right\rangle_{\lambda'} \approx \sum_{k=0}^{K-1} w_k \left\langle \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \lambda_k} \right\rangle^\dagger \\ &= \sum_{k=0}^{K-1} w_k \left\langle \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \right\rangle^\dagger , \end{aligned} \quad (3.5)$$

where the  $w_k$  are quadrature weights for the numerical integra-

tion,<sup>???</sup>  $\langle \dots \rangle_\lambda$  denotes ensemble averaging for a single system (*i.e.* over  $\mathbf{x}$ ) at the given  $\lambda$  value, and  $\langle \dots \rangle^\dagger$  denotes ensemble averaging for the replica system (*i.e.* over  $\mathbf{X}$ ).

In the above form, TI has long been the workhorse of alchemical free-energy calculations. The method is extremely robust in the sense that the accuracy of the calculated  $\Delta G$  can always be systematically improved (more  $\lambda$ -points, longer equilibration or/and sampling times). However, it is not necessarily the most efficient method to determine  $\Delta G$  up to a certain accuracy, due to possible sub-optimalities in the coupling scheme,<sup>????? ?</sup> the protocol design<sup>????? ?</sup> the free-energy estimator,<sup>????? ? ? ? ?</sup> and the orthogonal sampling.<sup>????? ?</sup>

### 3.2.4 CONVEYOR BELT THERMODYNAMIC INTEGRATION (CBTI)

The proposed CBTI scheme encompasses features of both  $\lambda$ D and TI. Similarly to TI, it is based on the simulation of a replica system involving  $K$  copies of the molecular system of interest, where  $K$  is taken to be even. And similarly to  $\lambda$ D, the individual replicas are extended systems, for which the associated  $\lambda_k$ -variable is allowed to evolve along the simulation. However, the evolutions of these  $\lambda_k$ -variables are not independent. They are coupled to each other by means of a sequence of hard constraints, so that they follow the course of a conveyor belt (CB). Thus, they are entirely determined by a single dynamical variable  $\Lambda$ , following the scenario depicted in Fig. 3.1 and discussed in the Introduction section.

The variable  $\Lambda$  is a continuous real variable representing the overall advance of the CB, successive multiples of  $2\pi$  corresponding to as many full rotations. Given  $\Lambda$  and  $K$ , the  $\lambda$ -value  $\lambda_k$

associated with a system  $k$  on the CB is obtained as

$$\lambda_k(\Lambda) = \zeta(\Lambda + k\Delta\Lambda) , \quad (3.6)$$

with

$$\Delta\Lambda = 2\pi K^{-1} . \quad (3.7)$$

Here, the function  $\zeta$  is a continuous and periodic zig-zag function of period  $2\pi$  and image range  $[0, 1]$ , defined over the reference period  $[0, 2\pi)$  as

$$\zeta(\theta) = \begin{cases} \pi^{-1}\theta & \text{if } \theta < \pi \\ 2 - \pi^{-1}\theta & \text{if } \theta \geq \pi \end{cases} \quad \text{for } \theta \in [0, 2\pi) , \quad (3.8)$$

where the  $[\cdot, \cdot)$  indicates an interval that is open to the right side, *e.g.*  $[0, 2\pi)$  includes 0 but excludes  $2\pi$ . An advance of the CB by  $\Delta\Lambda$  corresponds to a cyclic permutation of the  $K$  replicas, each system moving by one position forward along the CB, *i.e.*  $\lambda_k(\Lambda + \Delta\Lambda) = \lambda_{k+1}(\Lambda)$  for  $k < K - 1$ , along with  $\lambda_{K-1}(\Lambda + \Delta\Lambda) = \lambda_0(\Lambda)$ . For this reason, the increment  $\Delta\Lambda$  will be further referred to as one shift of the CB. The system  $k = 0$  can be viewed as a reference system, as  $\lambda_0 = \pi^{-1}\Lambda$  for  $0 \leq \Lambda < \pi$  and  $\lambda_0 = 2 - \pi^{-1}\Lambda$  for  $\pi \leq \Lambda < 2\pi$ . Since  $K$  is chosen to be even, an increase of  $\Lambda$  always corresponds to an increase of  $\lambda_k$  for half of the systems (forward-moving side of the CB) and a decrease of  $\lambda_k$  for the other half of the systems (backward-moving side of the CB). This choice also implies that  $\Lambda$ -values which are integer multiples of the CB shift  $\Delta\Lambda$  correspond to situations where there is one system in state  $A$  and one system in state  $B$ . Since an advance of the CB variable by  $2\pi$  leaves the replica system invariant, the variable  $\Lambda$  will commonly be refolded into the reference period  $[0, 2\pi)$  when

illustrating the results of the CBTI method.

The Hamiltonian of the extended replica system is defined as

$$\mathcal{H}^{\dagger\star}(\mathbf{X}, \Lambda) = \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda}) + \frac{p_\Lambda^2}{2m_\Lambda} \quad \text{with } \boldsymbol{\lambda} = \boldsymbol{\lambda}(\Lambda), \quad (3.9)$$

where  $\mathcal{H}^\dagger$  is defined as in TI by Eq. 3.4 and  $\boldsymbol{\lambda}(\Lambda)$  by Eq. 3.6. In analogy with Eq. 3.2, the resulting equation of motion for  $\Lambda$  reads

$$\ddot{\Lambda} = \frac{\dot{p}_\Lambda}{m_\Lambda} = -\frac{1}{m_\Lambda} \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \Lambda} \quad (3.10)$$

where

$$\begin{aligned} \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \Lambda} &= \sum_{k=0}^{K-1} \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \frac{d\lambda_k}{d\Lambda} \\ &= \sum_{k=0}^{K-1} \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \zeta'(\Lambda + 2\pi K^{-1}k) . \end{aligned} \quad (3.11)$$

Here, the function  $\zeta'$  is the derivative of the zig-zag function of Eq. 3.8, given over the reference period  $[0, 2\pi]$  by

$$\zeta'(\theta) = \begin{cases} \pi^{-1} & \text{if } \theta < \pi \\ -\pi^{-1} & \text{if } \theta \geq \pi \end{cases} \quad \text{for } \theta \in [0, 2\pi) . \quad (3.12)$$

Formally, the derivative is not defined when  $\theta$  is an integer multiple of  $\pi$ , *i.e.* for a system that is exactly in one of the physical end-states  $A$  or  $B$ . In this case, the value of  $\zeta'$  has been arbitrarily set to  $\pi^{-1}$  for even multiples and  $-\pi^{-1}$  for odd multiples. This has a negligible impact in practice, as it only concerns a series of infinitesimal points over the entire  $\Lambda$ -range, *i.e.* infinitesimally few configurations along a CBTI simulation. For example, when using double-precision floating-point arithmetics (including denormalized numbers), their probability of occurrence is on the order

of  $10^{-324}K$  (*i.e.* a single expected occurrence over a simulation lasting about  $10^{292}K^{-1}$  times the age of the universe with a 2 fs timestep). Neither does going over the discontinuity within a timestep represent a source of non-conservativeness. The concerned replica will merely bounce back the corresponding physical end-state with a reversion of its velocity, akin to a particle reflected elastically by a hard wall (delta-function force). If desired, these exceptional points could be handled more formally by altering the definition of  $\zeta$ , *e.g.* by smoothing its tips in a narrow range around 0 and  $\pi$ .

The  $\lambda_k$ -dynamics of the individual systems is entirely specified by Eq. 3.10 to propagate  $\Lambda$  along with Eq. 3.6 to calculate the  $\lambda_k$ -values from the current  $\Lambda$ . Alternatively, one may write an equation of motion for the  $\lambda_k$ -variables of the individual replicas by combining the two equations (along with Eq. 3.11) as

$$\begin{aligned}\ddot{\lambda}_k &= \ddot{\Lambda} \frac{d\lambda_k}{d\Lambda} + \dot{\Lambda}^2 \frac{d^2\lambda_k}{d\Lambda^2} \\ &= -\frac{1}{m_\Lambda} \left( \sum_{l=0}^{K-1} \frac{\partial \mathcal{H}(\mathbf{x}_l; \lambda_l)}{\partial \lambda_l} \zeta' (\Lambda + 2\pi K^{-1} l) \right) \\ &\quad \zeta' (\Lambda + 2\pi K^{-1} k) .\end{aligned}\tag{3.13}$$

Note that the term in  $\dot{\Lambda}^2$  vanishes since the second derivative  $\zeta''$  of  $\zeta$  is zero (except at the exceptional singular points). However, it should not be overlooked if one decides to use a different function  $\zeta$ . Introducing the vector  $\mathbf{D}$  and the symmetric  $\Lambda$ -dependent matrix  $\underline{\mathbf{C}}$  defined by their components as

$$\begin{aligned}D_k &= \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \quad \text{and} \\ C_{kl}(\Lambda) &= \pi^2 \zeta' (\Lambda + 2\pi K^{-1} k) \zeta' (\Lambda + 2\pi K^{-1} l) ,\end{aligned}\tag{3.14}$$

Eq. 3.13 can be rewritten in an elegant matrix form as

$$\ddot{\boldsymbol{\lambda}} = -\frac{\mathbf{C}(\Lambda)}{\pi^2 m_\Lambda} \mathbf{D} . \quad (3.15)$$

The elements of the symmetric matrix  $\mathbf{C}$  are either -1 (pair of systems currently on opposite sides of the CB, and thus moving in opposite directions) or +1 (pair of systems currently on the same side of the CB, and thus moving in the same direction). The diagonal elements are all +1, and the other +1 values surround the diagonal (line- and column-wise), the rest being -1 values. Because  $K$  is even, the two types of values are always equally represented in the matrix, specific locations depending on  $\Lambda$ . Note that the variable  $\Lambda$  itself still needs to be explicitly propagated using Eq. 3.10.

For a given configuration  $\mathbf{X}$  of the replica system, the Hamiltonian  $\mathcal{H}^\dagger$  of Eq. 3.4 (together with Eq. 3.6) is periodic in  $\Lambda$  with a period  $2\pi$  corresponding to a full rotation of the CB. However, because the Hamiltonians of the individual replicas are identical, upon ensemble averaging over  $\mathbf{X}$ , one expects the calculated properties to be periodic over  $\Lambda$  with a smaller period  $\Delta\Lambda$ , corresponding to one shift of the CB. This is in particular the case for the probability distribution  $P(\Lambda)$  along  $\Lambda$  and the associated free-energy profile  $G_\Lambda(\Lambda)$ , given by

$$\begin{aligned} G_\Lambda(\Lambda) &= G_\Lambda(0) + \int_0^\Lambda d\Lambda' \left\langle \frac{\partial \mathcal{H}^\dagger(\mathbf{X}; \boldsymbol{\lambda})}{\partial \Lambda} \right\rangle_{\Lambda'}^\dagger \\ &= \tilde{G}_\Lambda(0) + \sum_{k=0}^{K-1} G(\lambda_k) \quad \text{with } \boldsymbol{\lambda} = \boldsymbol{\lambda}(\Lambda) , \end{aligned} \quad (3.16)$$

where  $\boldsymbol{\lambda}(\Lambda)$  is defined by Eq. 3.6,  $\langle \cdots \rangle_\Lambda^\dagger$  denotes ensemble averaging for the replica system (*i.e.* over  $\mathbf{X}$ ) at the given  $\Lambda$  value, and

the second equality follows from Eqs. 3.4 and 3.5 (the unknown constant  $\tilde{G}_\Lambda(0)$  is equal to  $G_\Lambda(0)$  increased by a sum of  $-G(\lambda_k(0))$  offsets).

Owing to this periodicity over a smaller interval, it is convenient to introduce a fractional advance variable  $\tilde{\Lambda}$  defined as

$$\tilde{\Lambda} = \gamma(\Lambda, \Delta\Lambda), \quad (3.17)$$

where

$$\gamma(\theta, \theta_o) = \theta_o (\theta_o^{-1}\theta - \lfloor \theta_o^{-1}\theta \rfloor) \quad (3.18)$$

returns the part of  $\theta$  in excess of the closest lower integer multiple of  $\theta_o$ . In contrast to  $\Lambda$ , which is an unbounded variable,  $\tilde{\Lambda}$  only spans a finite definition interval  $[0, \Delta\Lambda]$ . At full convergence, any average property binned as a function of  $\Lambda$  over the interval  $[0, 2\pi]$  will consist of  $K$  successive repeats of the same property binned as a function of  $\tilde{\Lambda}$  over its definition interval  $[0, \Delta\Lambda]$ , as observed in Fig. 3.1h for the free energy  $G_\Lambda(\Lambda)$ . Accordingly, in the absence of full convergence along  $\Lambda$ , binning as a function of  $\tilde{\Lambda}$  over the interval  $[0, \Delta\Lambda]$  followed by  $K$ -fold replication provides an efficient way to construct a more accurate representation of any  $\Lambda$ -resolved average quantity. In fact, the definition interval of  $\tilde{\Lambda}$  could be further halved by noting that, upon ensemble averaging over  $\mathbf{X}$  and for any  $\Lambda$  value that is an integer multiple of  $\Delta\Lambda$ , a forward move of the CB produces the same result as a backward move of the same magnitude. Consequently,  $\Lambda$ -resolved average properties are even over successive  $2\pi K^{-1}$  intervals, as also observed in Fig. 3.1h for the free energy  $G_\Lambda(\Lambda)$ . The corresponding information is thus entirely encompassed in an interval of size  $\Delta\Lambda/2$ .

The normalized probability distribution  $p(\lambda)$  along the cou-

pling variable  $\lambda$  considering all the replicas is defined by

$$p(\lambda) = K^{-1} \sum_{k=0}^{K-1} \langle \delta(\lambda_k - \lambda) \rangle^{\dagger\star} , \quad (3.19)$$

where  $\langle \dots \rangle^{\dagger\star}$  denotes ensemble averaging for the extended replica system (*i.e.* over the joint trajectories of  $\mathbf{X}$  and  $\boldsymbol{\lambda}$ ). At full convergence, this probability over the interval  $[0, 1]$  will consist of  $K/2$  successive repeats of the corresponding distribution over the interval  $[0, 2K^{-1}]$ . More precisely, the distribution  $p(\lambda)$  is related to the distribution  $\tilde{P}(\tilde{\Lambda})$  of  $\tilde{\Lambda}$  over interval  $[0, \Delta\Lambda]$  as

$$p(\lambda) = \Delta\Lambda \tilde{P}(\pi\gamma(\lambda, \pi^{-1}\Delta\Lambda)) . \quad (3.20)$$

In plain words, this means that  $\tilde{P}(\tilde{\Lambda})$  is the relevant quantity in terms of sampling along the coupling variable  $\lambda$ . If it is close to uniform over the range  $[0, \Delta\Lambda]$ , then  $p(\lambda)$  will also be close to uniform over the range  $[0, 1]$ . Here again, it is noted that  $p(\lambda)$  is also even over the interval  $[0, 2K^{-1}]$ , and could be mapped to a  $\tilde{\Lambda}$  value defined over an interval of size  $\Delta\Lambda/2$  instead of  $\Delta\Lambda$  if desired, as

$$p(\lambda) = \begin{cases} \Delta\Lambda \tilde{P}(\pi\gamma(\lambda, \pi^{-1}\Delta\Lambda)) & \text{if } \gamma(\lambda, \pi^{-1}\Delta\Lambda) < (2\pi)^{-1}\Delta\Lambda \\ \Delta\Lambda \tilde{P}(\pi\gamma(1 - \lambda, \pi^{-1}\Delta\Lambda)) & \text{otherwise} \end{cases} . \quad (3.21)$$

### 3.2.5 CBTI FREE-ENERGY ESTIMATOR

Due to the constraints coupling the  $\lambda_k$ -values of the  $K$  replicas, the function  $p(\lambda)$  of Eq. 3.19 is by no means a Boltzmann distribution in terms of the single-system Hamiltonian. In fact, as seen above,

it consists at full convergence of  $K/2$  successive repeats of the same even curve. In addition, compared to the Boltzmann distribution, it will be significantly flatter. On the one hand, the smaller amplitude of variations are desired, as they will lead to more homogeneous sampling and are expected to ease transitions along  $\Lambda$  (up to the limit imposed by the speed of random diffusion). On the other hand, it is no longer possible to evaluate the free-energy difference  $\Delta G$  directly from  $p(\lambda)$  in analogy with the  $\lambda D$  expression of Eq. 3.3. However, since the dynamics remains Hamiltonian and the coupling between replicas does not involve the configurational degrees of freedom, the change from TI to CBTI does not affect the conditional probabilities  $\mathcal{P}(\mathbf{x}|\lambda)$ . Thus, configurational ensemble averages sorted by  $\lambda$ -values will remain identical to those one would obtain from TI (or from HRE/HRP or  $\lambda D$ ). As a result,  $\Delta G$  can still be obtained by integrating over the average Hamiltonian derivative binned as a function of  $\lambda$  considering all replicas simultaneously, in analogy with the TI expression of Eq. 3.5. Note that the exceptional points of the function  $\zeta'$  (discussed previously in the context of Eq. 3.12) have no influence on the integration, as they represent finite discontinuities over infinitesimal ranges.

In practice,  $\Delta G$  is calculated here based on a single thermostated MD simulation of the extended replica system, as

$$\begin{aligned}\Delta G &= \int_0^1 d\lambda' K^{-1} \sum_{k=0}^{K-1} \left\langle \frac{\mathcal{H}(x_k; \lambda_k)}{\partial \lambda_k} \delta(\lambda_k - \lambda') \right\rangle^{\dagger*} \quad (3.22) \\ &\approx \sum_{j=0}^{J-1} \left\langle \frac{\sum_{k=0}^{K-1} \frac{\mathcal{H}(x_k; \lambda_k)}{\partial \lambda_k} \alpha(\lambda_k, j; J)}{\sum_{k=0}^{K-1} \alpha(\lambda_k, j; J)} \right\rangle^{\dagger*},\end{aligned}$$

where

$$\alpha(\theta, j; J) = \begin{cases} 1 & \text{if } j \leq J\theta < j + 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.23)$$

is a binning function corresponding to a discretization of the  $\lambda$ -interval  $[0, 1]$  using  $J$  bins. The approximation in Eq. 3.22 corresponds to a simple forward rectangular quadrature, where the Hamiltonian derivative is averaged over the  $J$  successive bins considering all replicas. Since  $\tilde{P}(\tilde{\Lambda})$ , and thus  $p(\lambda)$ , will typically be close to homogeneous,  $J$  can be taken very large, resulting in a negligible quadrature error. For example, if  $K$  replicas sample  $L$  configurations each, the number of data points per bin will be close to  $KL/J$ , with limited variations across bins. Defining the maximal allowed value  $J_{\max}$  as the highest value of  $J$  for which empty bins (vanishing denominator in the ensemble average of Eq. 3.22) never occur, a graph of  $\Delta G$  evaluated upon increasing  $J$  from 1 to  $J_{\max}$  will rapidly level off to a plateau when quadrature errors become negligible. Two variants which do not require the specification of a number of bins are also proposed in Sect. 3.C (Eqs. 3.C.1 and 3.C.2).

### 3.2.6 CBTI WITH MEMORY-BASED BIASING POTENTIAL

When using a large number of replicas, the sampling along  $\lambda$  afforded by the CBTI scheme will be close to homogeneous. However, for practical reasons (*e.g.* number of processors available on a computer node), one may wish to use a small number of replicas. In this case, the sampling homogeneity can be enhanced by addition of a biasing potential. It is sufficient to apply this potential

to the fractional advance variable  $\tilde{\Lambda}$  over the range  $[0, \Delta\Lambda]$ . With inclusion of a biasing potential  $\mathcal{B}$ , Eq. 3.10 becomes

$$\ddot{\Lambda} = -\frac{1}{m_{\Lambda}} \frac{\partial}{\partial \Lambda} \left( \mathcal{H}^{\dagger}(\mathbf{X}; \boldsymbol{\lambda}) + \mathcal{B}(\tilde{\Lambda}) \right)$$

with  $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\Lambda)$  and  $\tilde{\Lambda} = \tilde{\Lambda}(\Lambda)$  , (3.24)

where  $\mathcal{H}^{\dagger}$  is defined by Eq. 3.4,  $\boldsymbol{\lambda}(\Lambda)$  by Eq. 3.6 and  $\tilde{\Lambda}(\Lambda)$  by Eq. 3.17.

In analogy with the  $\lambda$ -LEUS scheme,<sup>7 8 9 10</sup> this biasing potential can be expressed as a sum of local grid-based spline functions, built in a LE preoptimization phase and frozen in a subsequent US sampling phase. However, the duration of the LE phase can be considerably reduced compared to a single-system  $\lambda$ -LEUS simulation, considering that  $\tilde{P}(\tilde{\Lambda})$  is already close to homogeneity in the absence of biasing and that the support interval is reduced to the  $\tilde{\Lambda}$ -range  $[0, \Delta\Lambda]$ . The latter interval can actually be further restricted to  $[0, \Delta\Lambda/2]$  considering the even symmetry of  $\tilde{P}(\tilde{\Lambda})$ , *i.e.* by enforcing an even symmetry of  $\mathcal{B}$  as well.

Since the application of a biasing potential that only involves the  $\lambda_k$ -variables does not alter the conditional probabilities  $\mathcal{P}(\mathbf{x}|\boldsymbol{\lambda})$ , Eq. 3.22 (or the variants of Eqs. 3.C.1 and 3.C.2) can still be employed without any modification to evaluate the free-energy change. In other words, in contrast to the  $\lambda$ -LEUS scheme, the CBTI scheme with the presented TI-like free-energy estimator does not require any reweighting.

### 3.3 COMPUTATIONAL DETAILS

#### 3.3.1 TEST SYSTEM

As an initial application of the proposed CBTI scheme, we considered here a relatively simple perturbation, namely the conversion of methanol from a fully interacting molecule to a dummy skeleton (no intermolecular interactions) in an aqueous environment at  $P = 1$  bar and  $T = 298.15$  K. The calculations were performed using a modified version of the GROMOS11 program<sup>7 8 9</sup> along with the parameters of the GROMOS-compatible 2016H66 force field<sup>7</sup> for methanol<sup>7</sup> (united atom, rigid bonds, flexible bond angle) and the simple point charge (SPC) model<sup>7</sup> for water (fully rigid). Since the dummy skeleton retains the intramolecular interactions (here, only the bond angle), the calculated free-energy change  $\Delta G$  corresponds directly to minus the hydration free energy of methanol.

Possible issues related to the existence of a singularity<sup>7 8 9</sup> and the insufficient solute-solvent kinetic-energy exchange<sup>7 8 9</sup> close to  $\lambda = 1$  were alleviated in the usual way, by means of a soft-core scheme<sup>7</sup> for the alchemical coupling and of stochastic dynamics<sup>7 8</sup> (SD) for thermostating the solute and solvent conformational degrees of freedom. In most CBTI simulations, the instantaneous temperature  $T_\Lambda$  of the CB advance variable  $\Lambda$  was also controlled separately by means of a Nosé-Hoover chain thermostat<sup>7</sup> at 298.15 K (eight successive thermostat variables), with a coupling time  $\tau_\Lambda$ .

### 3.3.2 SIMULATIONS SETS

The exploration of the CBTI scheme and the comparison of its performance with that of existing methods was carried out in five successive steps: (1) establishing reference TI results; (2) analyzing the influence of the CBTI parameters (number  $K$  of replicas along with the mass-parameter  $m_\Lambda$  and thermostat coupling time  $\tau_\Lambda$  of the CB advance variable); (3) investigating the use of a biasing potential; (4) examining the features of the TI-like free-energy estimator (effect of the number  $J$  of integration bins and use of equations without specification of  $J$ ); (5) comparing the results of CBTI with those of existing methods.

The reference TI calculations (Step 1) were performed using  $K_{\text{TI}} = 2^n + 1$  equidistant  $\lambda$ -points covering the range  $[0, 1]$  with  $n = 1, 2, \dots, 7$ . They involved initial configurations equilibrated for 0.2 ns starting from the equilibrated configuration at the previous  $\lambda$ -point, and a simulation time of  $100K_{\text{TI}}^{-1}$  ns per  $\lambda$ -point. Each of these calculations, involving a total single-system sampling time of 100 ns, was repeated ten times using different random initial velocities. The integration over the average Hamiltonian derivative was performed based on Eq. 3.5 using the Simpson quadrature rule.<sup>?</sup> <sup>?</sup> <sup>?</sup>

To explore the influence of the CBTI parameters (Step 2), various combinations of  $K$ ,  $m_\Lambda$  and  $\tau_\Lambda$  were considered in three series of calculations, namely: (i) the choices  $m_\Lambda = 16, 160, 800, 1600$  or  $3200 \text{ u nm}^2$  (where  $\text{u}$  stands for atomic mass unit, *i.e.* g mol<sup>-1</sup>), along with  $K = 16$  replicas in the absence of thermostat coupling for  $\Lambda$ , *i.e.* with  $\tau_\Lambda \rightarrow \infty$ ; (ii) the choices  $\tau_\Lambda = 0.05, 0.1, 0.5, 1$  or 2 ps along with  $K = 16$  replicas and  $m_\Lambda = 160 \text{ u nm}^2$ ; (iii) the choices  $K = 8, 16, 32, 64$  or 128 along with  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5$  ps; The parameters (and

results) of these three series of simulations, including their durations  $t_{\text{sim}}$ , are summarized in Tab. 3.1 (entries 1-15). All these simulations were preceded by 0.2 ns equilibration. For the the third series (entries 11-15),  $m_\Lambda$  was made proportional to  $K^{1/2}$ , an arbitrary parameter choice justified by arguments provided in Sect. 3.B, and the five simulations relied on the the same total single-system sampling time of 256 ns. This exploration showed that the CBTI method is rather robust with respect to the choice of its parameters. The values  $K = 16$ ,  $m_\Lambda = 160 \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  were retained as a good combination for the alchemical perturbation considered. For comparison with the TI results of Step 1, ten repeats of the calculation involving this specific choice were performed using different random initial velocities and a total single-system sampling time of 100 ns after 0.2 ns equilibration.

The application of CBTI with a biasing potential (Step 3) was investigated in the context of simulations with  $K = 8$  or  $16$ , both with  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$ . For  $K = 8$ , the biasing potential  $\mathcal{B}$  (Eq. 3.24) was constructed using  $N_{\text{gp}} = 34$  basis functions centered at equidistant grid-points  $i = 0, \dots, N_{\text{gp}} - 1$  over the  $\tilde{\Lambda}$ -range  $[0, \pi/4]$ . The coefficients of the basis-functions  $i$  and  $N_{\text{gp}} - 1 - i$  with  $i = 0, \dots, N_{\text{gp}}/2 - 1$  were constrained to be identical, considering the expected even symmetry of  $\tilde{P}(\tilde{\Lambda})$ . In terms of the CB advance variable  $\Lambda$ , this means that the biasing potential relied in effect on  $K(N_{\text{gp}} - 1) = 264$  local functions covering the  $\Lambda$ -range  $[0, 2\pi]$ , these functions being defined by only 17 independent coefficients. For  $K = 16$ ,  $\mathcal{B}$  relied on  $N_{\text{gp}} = 18$  basis functions over the  $\tilde{\Lambda}$ -range  $[0, \pi/8]$ , leading to 272 functions over the  $\Lambda$ -range  $[0, 2\pi]$  defined by 9 independent coefficients. Second-order splines<sup>??</sup> (of range  $\pm 2\delta$  with  $\delta = \pi/132$  or  $\pi/136$  for  $K = 8$  and  $16$ , respectively) were employed<sup>?</sup> as basis functions. An initial build-up force constant  $c_{\text{LE}} = 10^{-3} \text{ kJ mol}^{-1}$  was used,

which was multiplied by a reduction factor  $f_{\text{red}} = 0.1$  after each double-sweep of half the  $\tilde{\Lambda}$ -range  $[0, \pi K^{-1}]$ . The duration  $t_{\text{LE}}$  of the LE build-up phase for the replica system was = 0.15 ns for  $K = 8$  and 0.07 ns for  $K = 16$ , corresponding to only 1.1 – 1.2 ns total single-system simulation time. The parameters (and results) of these two simulations are summarized in Tab. 3.1 (entries 16 and 17). The duration  $t_{\text{sim}}$  of the US sampling phases for the replica system were 22 ns, corresponding to total single-system sampling times of 176 ns and 352 ns for  $K = 8$  and 16, respectively.

To examine the features of the TI-like free-energy estimator (Step 4), the number  $J$  of integration bins used in the rectangular quadrature to calculate  $\Delta G$  (Eq. 3.22) was varied considering the simulations of Steps 2 and 3 above (17 simulations of Tab. 3.1). The resulting  $\Delta G$  values were also compared to those of the variants  $\Delta G_{\text{alt}}$  (Eq. 3.C.1) and  $\Delta G_{\text{app}}$  (Eq. 3.C.2) proposed in Sect. 3.C.

Finally, the results of the CBTI simulations were compared to those of other methods (Step 5), namely TI or HRE using Simpson quadrature as estimator as well as TI using EXTI and MBAR as estimator. For Simpson quadrature, the TI simulations relied on  $K_{\text{TI}} = 3, 5, 9, 17, 65$  or 129 equispaced  $\lambda$ -points, and the HRE simulations relied on  $K_{\text{HRE}} = 17, 33$  or 65 equispaced replicas with exchange attempts every  $\tau_{\text{HRE}} = 0.2$  ps. The use of EXTI and MBAR was explored based on TI-like simulations relying on  $K_{\text{TI}} = 9$  and 17 equispaced  $\lambda$ -points. For EXTI, the average Hamiltonian derivative was extrapolated during the  $K_{\text{TI}}$  simulations to 129 equispaced virtual  $\lambda$ -points, and the latter 129 values used in the Simpson quadrature. For MBAR, the Hamiltonian was calculated at 129 equispaced virtual  $\lambda$ -points considering all the configurations sampled in the  $K_{\text{TI}}$ , and the data combined using the MBAR equation<sup>7</sup> as implemented in

TABLE 3.1: *Influence of the CBTI parameters in simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This table investigates the influence of the parameters selected for the CBTI scheme on the temperature and dynamics of the CB advance variable  $\Lambda$  and on the calculated free-energy change  $\Delta G$  (for the latter, considering a constant total single-system sampling time of 100 ns). For each simulation, the successive entries are: the index of the simulation (sim), the number  $K$  of replicas, the simulation time  $t_{\text{sim}}$  for the replica system, the mass-parameter  $m_\Lambda$ , the thermostat coupling time  $\tau_\Lambda$  ( $\infty$  indicates that no coupling is applied), the average temperature  $T_\Lambda$ , the root-mean-square fluctuation  $\sigma_\Lambda$  of  $\dot{\Lambda}$ , the autocorrelation time  $\tau_{\dot{\Lambda}}$  of  $\dot{\Lambda}$ , the diffusion coefficient  $D_\Lambda$  (Eq. 3.26), the free-energy difference  $\Delta G$  calculated using Eq. 3.22 with  $J = 500$  (except for entry 11,  $J = 200$ ), the alternative free-energy difference  $\Delta G_{\text{alt}}$  calculated using Eq. 3.C.1, and the approximate free-energy difference  $\Delta G_{\text{app}}$  calculated using Eq. 3.C.2. Error estimates obtained by bootstrapping (no Student  $t$ -factor included) are also reported between parentheses for  $\Delta G$ ,  $\Delta G_{\text{alt}}$  and  $\Delta G_{\text{app}}$ . Note that the simulations differ in terms of total single-system sampling time  $Kt_{\text{sim}}$ . To enable a fair comparison, the free energy-changes and associated errors have been calculated after truncating the all simulations to 100 ns single-system sampling time evenly distributed over all replicas. Associated graphs for the distributions  $P(\Lambda)$  of  $\Lambda$ ,  $P_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  and  $P_{\ddot{\Lambda}}$  of  $\ddot{\Lambda}$ , as well as the mean-square displacements  $d_\Lambda$  of  $\Lambda$  and autocorrelation functions  $c_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  can be found in Figs. 3.3, 3.5 and 3.7, or in Figs. Fig. 3.F.1 - Fig. 2.F.15 and Fig. 2.H.2-Fig. 2.H.4. Simulations 11, 13, 15, 16 and 17 are discussed in the main text. The other simulations are discussed in Sect. 3.B.

sim	$K$	$t_{\text{sim}}$ [ns]	$m_\Lambda$ [u nm <sup>2</sup> ]	$\tau_\Lambda$ [ps]	$T_\Lambda$ [K]	$\sigma_\Lambda$ [ps <sup>-1</sup> ]	$\tau_{\dot{\Lambda}}$ [ps]	$D_\Lambda$ [ns <sup>-1</sup> ]	$\Delta G$ [kJ mol <sup>-1</sup> ]	$\Delta G_{\text{alt}}$ [kJ mol <sup>-1</sup> ]	$\Delta G_{\text{app}}$ [kJ mol <sup>-1</sup> ]
1	16	10	16	$\infty$	308.0	0.40	0.02	16.7	21.14 (0.16)	21.23 (0.17)	20.21 (0.32)
2	16	10	160	$\infty$	297.9	0.12	0.61	13.2	21.27 (0.17)	21.32 (0.15)	20.29 (0.31)
3	16	10	800	$\infty$	296.3	0.05	1.90	10.2	21.25 (0.15)	21.34 (0.16)	20.27 (0.32)
4	16	10	1600	$\infty$	304.0	0.04	3.47	10.3	21.25 (0.15)	21.25 (0.17)	20.32 (0.32)
5	16	10	3200	$\infty$	297.7	0.03	5.63	7.4	21.23 (0.16)	21.24 (0.18)	20.11 (0.31)
6	16	10	160	0.05	284.0	0.12	0.37	6.9	21.23 (0.16)	21.04 (0.15)	20.22 (0.31)
7	16	10	160	0.10	291.7	0.12	0.36	7.4	21.48 (0.13)	21.41 (0.16)	20.47 (0.29)
8	16	10	160	0.50	296.4	0.12	0.41	9.1	21.42 (0.14)	21.48 (0.16)	20.45 (0.30)
9	16	10	160	1.00	299.4	0.12	0.49	12.9	21.57 (0.16)	21.48 (0.17)	20.59 (0.30)
10	16	10	160	2.00	296.8	0.12	0.56	13.1	21.48 (0.13)	21.73 (0.15)	20.45 (0.29)
11	8	32	113	0.50	298.4	0.15	0.23	4.4	21.69 (0.40)	21.69 (0.44)	15.65 (0.31)
12	16	16	160	0.50	298.3	0.12	0.41	10.2	21.42 (0.16)	21.48 (0.17)	20.45 (0.33)
13	32	8	226	0.50	300.6	0.10	0.38	6.2	21.44 (0.14)	21.33 (0.15)	21.34 (0.32)
14	64	3	320	0.50	297.2	0.09	0.31	2.5	21.32 (0.14)	21.27 (0.14)	21.29 (0.30)
15	128	2	452	0.50	304.1	0.08	0.24	2.3	21.43 (0.12)	21.61 (0.14)	21.41 (0.32)
16	8	22	113	0.50	296.4	0.15	0.43	15.3	21.48 (0.16)	21.61 (0.17)	19.65 (0.34)
17	16	22	160	0.50	299.3	0.12	0.43	9.4	21.30 (0.13)	21.43 (0.16)	20.61 (0.34)

pymbar.<sup>?</sup> All the above comparisons were performed at a total single-system sampling time of 100 ns distributed evenly over all replicas.

Error bars on the calculated  $\Delta G$  values were estimated in two different ways. For the calculations involving ten repeats (all TI calculations plus one CBTI simulation), the standard deviation  $\sigma$  of the mean was calculated by scaling that of the ten estimates by the square-root of nine, and the error  $\epsilon$  on the mean was calculated as  $\epsilon = 2.262\sigma$ , where 2.262 is the Student  $t$ -factor<sup>?</sup> for nine degrees of freedom and a two-sigma confidence interval of 95%. For the individual calculations that were not repeated, the statistical error was estimated by bootstrapping<sup>??</sup> (no Student  $t$ -factor included) using 100 bootstrap samples. If  $K$  replicas (CBTI) or  $\lambda$ -points (all other methods) have generated as many sets of  $L$  data points, a sample consists here of  $K$  sets of  $L$  data points selected randomly (possibly multiple times) from the  $K$  original data sets. Note that for TI/EXTI and TI/MBAR, it is essential to perform the bootstrapping based on the data from the  $K_{\text{TI}}$  real  $\lambda$ -points, and not from the  $129K_{\text{TI}}$  predicted values (the latter procedure would result in underestimated errors due to correlation in the derived data). The bootstrapping error will only be accurate provided that the data from the simulations, written to file every 2 ps, is uncorrelated in time. Normalized autocorrelation functions and characteristic times for the average Hamiltonian derivative in the different TI-simulations are provided in Fig. 3.D.1 and Tab. 3.D.1 to support this assumption.

### 3.3.3 SIMULATION PARAMETERS

The simulations involved a cubic computational box containing one methanol and 1000 water molecules under periodic boundary

conditions in the isothermal-isobaric ensemble at  $P = 1$  bar and  $T = 298.15$  K. They were performed using SD by integrating the Langevin equation of motion<sup>?</sup> using the leap-frog scheme<sup>?</sup> (SD variant<sup>?</sup>) with a timestep  $\Delta t = 2$  fs and a friction coefficient  $\gamma = 10\text{ ps}^{-1}$ . Since the kinetics of the system is irrelevant in this work, SD instead of thermostated MD was used to avoid problems related to insufficient solute-solvent kinetic-energy exchange<sup>???</sup> close to  $\lambda = 1$  (dummy-skeleton state). The value of  $\gamma$  corresponds to the coupling time of 0.1 ps commonly employed in GROMOS simulations<sup>??</sup> relying on a weak-coupling<sup>?</sup> thermostat. The average pressure was maintained close to its reference value by isotropic weak coupling<sup>?</sup> using a molecular virial, a coupling time  $\tau_P = 0.5$  ps and a compressibility  $\kappa = 4.575 \cdot 10^4 \text{ kJ mol}^{-1} \text{ nm}^{-3}$  as commonly used in GROMOS for aqueous biomolecular systems.<sup>??</sup> The bond rigidity of methanol and the full rigidity of water were enforced by application of the SHAKE algorithm<sup>?</sup> with a relative geometric tolerance of  $10^{-4}$ . The energies and Hamiltonian derivatives were written to file every 2 ps for analysis.

The non-bonded interactions were handled by means of a molecule-based twin-range cutoff scheme<sup>?</sup> with short- and long-range cutoff distances set to 0.8 and 1.4 nm, respectively, and an update frequency of 5 timesteps for the short-range pairlist and intermediate-range interactions. The molecule center was the center of geometry for methanol and the oxygen atom for water. A reaction-field correction<sup>??</sup> was applied to account for the mean effect of the electrostatic interactions beyond the long-range cutoff distance, using a relative dielectric permittivity of 61 as appropriate for the SPC model.<sup>?</sup> To alleviate issues related to the existence of a singularity<sup>???</sup> close to  $\lambda = 1$  (dummy-skeleton state), the alchemical transformation relied on

a soft-core scheme,<sup>7</sup> applied with the parameters  $\alpha_{\text{LJ}} = 0.5$  and  $\alpha_{\text{CRF}} = 0.5 \text{ nm}^2$ .

For the CBTI calculations, the propagation of the  $\Lambda$  variable (Eq. 3.10) preceded that of the conformational degrees of freedom, and was performed with the same timestep  $\Delta t$ . More precisely, the following leap-frog steps were carried out in sequence:  $\dot{\Lambda}(t - \Delta t/2) \rightarrow \dot{\Lambda}(t + \Delta t/2)$ ,  $\Lambda(t) \rightarrow \Lambda(t + \Delta t)$ , calculate  $\lambda$  from  $\Lambda$  using Eq. 3.6,  $\dot{\mathbf{r}}(t - \Delta t/2) \rightarrow \dot{\mathbf{r}}(t + \Delta t/2)$  and  $\mathbf{r}(t) \rightarrow \mathbf{r}(t + \Delta t)$ . Unless otherwise specified (explorative simulations), the CBTI simulations relied on a mass-parameter  $m_\Lambda$  set to  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$ , on a coupling time  $\tau_\Lambda$  set to 0.5 ps, and on the use of  $J = 500$  bins for evaluating  $\Delta G$  based on Eq. 3.22 ( $J = 200$  for the unbiased simulation with  $K = 8$ ).

### 3.3.4 TRAJECTORY ANALYSIS

The dynamics of the replica system in the CBTI simulations was characterized by monitoring the distribution  $P$  of the  $\Lambda$  variable, the average temperature  $T_\Lambda$ , the distribution  $P_{\dot{\Lambda}}$  of  $\dot{\Lambda}$ , the associated root-mean-square fluctuation  $\sigma_{\dot{\Lambda}}$ , the normalized autocorrelation function  $c_{\dot{\Lambda}}$  of  $\dot{\Lambda}$ , the associated autocorrelation time  $\tau_{\dot{\Lambda}}$ , the mean-square displacement  $d_\Lambda$  of  $\Lambda$  as a function of time, the associated diffusion coefficient  $D_\Lambda$ , and the distribution  $P_{\ddot{\Lambda}}$  of  $\ddot{\Lambda}$ .

The distribution  $P_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  can be compared to the analytical one-dimensional Maxwell-Boltzmann velocity distribution<sup>7</sup>

$$P_{\dot{\Lambda}}^{\text{MB}}(\dot{\Lambda}) = \left( \frac{\beta m_\Lambda}{2\pi} \right)^{\frac{1}{2}} e^{-\frac{\beta m_\Lambda}{2}\dot{\Lambda}^2}. \quad (3.25)$$

The diffusion coefficient  $D_\Lambda$  was calculated from the mean-square displacement  $d_\Lambda$  of  $\Lambda$  as a function of time  $t$  according to

the one-dimensional Einstein equation?

$$D_\Lambda = \lim_{t \rightarrow \infty} \frac{d_\Lambda(t)}{2t}$$

with  $d_\Lambda(t) = \langle [\Lambda(\tau + t) - \Lambda(\tau)]^2 \rangle_t$ , (3.26)

where  $\langle \dots \rangle_t$  denotes averaging over  $\tau$  (all possible time origins) at constant  $t$ . Note that this equation must be applied to the unbounded variable  $\Lambda$ , *i.e.* without refolding to the reference interval  $[0, 2\pi]$ . The infinite-time limit was replaced in practice by a linear least-squares fit over the time range 0 to 0.15 ns.

All the graphs presented in this chapter were generated with Python ([www.python.org](http://www.python.org)) and the Matplotlib library.<sup>?</sup>

## 3.4 RESULTS AND DISCUSSION

### 3.4.1 REFERENCE TI CALCULATIONS

The curve for the average Hamiltonian derivative as a function of  $\lambda$  corresponding to the aqueous methanol-to-dummy mutation is shown in Fig. 3.2a (blue curve, scale on the left), as averaged over the ten repeats of the 100 ns TI calculation with 129 equispaced  $\lambda$ -points. The corresponding running integral, *i.e.* the free-energy profile  $G(\lambda)$ , is also shown (orange curve, scale on the right). In view of the large amount of statistics (total 1  $\mu$ s sampling, *i.e.* clearly an overkill for such a calculation), the curve is perfectly smooth and extremely well converged. The average Hamiltonian derivative is large and positive below about 0.5 and becomes smaller and negative thereafter, leading to a maximum in the free-energy profile. The shape of these curves is determined both

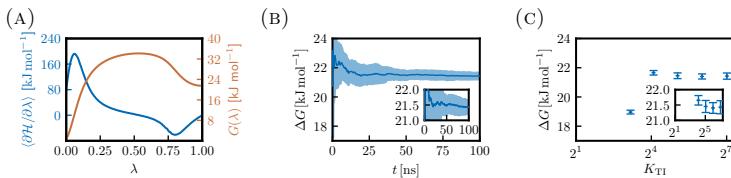


FIGURE 3.2: *Convergence properties of the TI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure provides the reference free-energy profile for the perturbation considered, and investigates the convergence properties of TI with different numbers  $K_{\text{TI}}$  of equidistant  $\lambda$ -points (based on ten repeats of 100 ns TI calculations). Panel (a) shows the average Hamiltonian derivative as a function of  $\lambda$  (blue curve, scale on the left) and the corresponding free-energy profile (orange, scale on the right). The curves are averaged over ten repeats of 100 ns TI calculations with 129 equidistant  $\lambda$ -points. Panel (b) shows the running  $\Delta G$  estimates for each of the ten repeats (dashed curves) as a function of the total sampling time  $t$  per repeat, along with the corresponding mean and error on the mean (95% confidence interval) over the repeats (thick blue curve and shaded area). Panel (c) shows  $\Delta G$  estimates considering ten repeats of 100 ns TI calculations using  $K_{\text{TI}} = 2^n + 1$  equidistant  $\lambda$ -points with  $n = 1, 2, \dots, 7$  (individual colored circles), along with the corresponding mean and error on the mean (95% confidence interval) over the repeats (blue crosses and error bars). Figs. equivalent to Panel (b) for the TI calculations with fewer  $\lambda$ -points are provided in Fig. 3.E.1. Numerical values at full sampling time are reported in Tab. 3.2 and Tab. 3.E.1.

by the hydration physics of methanol and by the choice of the employed (soft-core) coupling scheme.

The convergence properties of these TI calculations as a function of the total sampling time per repeat are illustrated in Fig. 3.2b, in the form of running  $\Delta G$  estimates for each of the ten repeats (dashed curves), along with the associated mean and error (95% confidence interval) on the mean over the repeats (thick blue curve and shaded area). The latter mean and error at full sampling time, reported in Tab. 3.2, evaluate to  $21.43 \pm 0.21 \text{ kJ mol}^{-1}$ . This result (with a minus sign) is in quantitative agreement with the experimental value for the hydration free energy of methanol at 298.15 K and 1 bar, namely?  $-21.4 \text{ kJ mol}^{-1}$ .

Corresponding graphs for TI calculations involving different numbers of equidistant  $\lambda$ -points ( $K_{\text{TI}} = 2^n + 1$  with  $n = 1, 2, \dots, 7$ ) at identical total sampling time (100 ns for each of the ten repeats) are shown in Fig. 3.E.1. The associated  $\Delta G$  means and errors at full sampling time are reported numerically in Tab. 3.2 (see Tab. 3.E.1 for the values of the individual repeats). The results of Tab. 3.2 also include a comparison of different error estimates (all excluding the Student  $t$ -factor), namely: (i) the standard deviation  $\sigma$  of the mean over the repeats; (ii) a bootstrap error  $\sigma_B$  calculated from the concatenated trajectories of the repeats at each  $\lambda$ -point; (iii) an alternative error  $\sigma_G$  obtained similarly by calculating the bootstrap errors on the average Hamiltonian derivatives at all  $\lambda$ -points, and propagating them assuming a Gaussian distribution of the estimates, as proposed in Refs. ???. The three values turn out to be very similar, so that further discussions will refer to the error  $\epsilon$  defined as  $\sigma$  amplified by the Student  $t$ -factor.

The effect of  $K_{\text{TI}}$  on the convergence properties is illustrated graphically in Fig. 3.2c, which shows the  $\Delta G$  estimates at full sampling time as a function of  $K_{\text{TI}}$  for each of the ten repeats (individual colored circles), along with the associated mean and error on the mean over the repeats (blue crosses and error bars). Using too few  $\lambda$ -points leads to quadrature errors, which are very significant for  $n \leq 4$  ( $K_{\text{TI}} \leq 17$ ), *i.e.* the shape of the average Hamiltonian derivative curve is not well captured by the Simpson quadrature. For  $n \geq 5$  ( $K_{\text{TI}} \geq 33$ ), the quadrature error becomes essentially negligible compared to the sampling error.

Interestingly, at constant total sampling time, the latter error appears to be essentially insensitive to the number of  $\lambda$ -points employed, with very similar  $\Delta G$  estimates and error bars of  $21.45 \pm 0.20$ ,  $21.40 \pm 0.17$  and  $21.43 \pm 0.21$  kJ mol<sup>-1</sup> for  $K_{\text{TI}} = 33, 65$  and

TABLE 3.2: *Repeat statistics over TI and CBTI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This table investigates the convergence properties of the calculated free-energy change  $\Delta G$  considering TI calculations with different numbers of  $\lambda$ -points along with a CBTI simulation with 16 replicas (average and various error estimates, based on ten calculation repeats involving each at a constant total single-system sampling time of 100 ns) The TI calculations involved  $K_{\text{TI}} = 2^n + 1$  equispaced  $\lambda$ -points with  $n = 1, 2, \dots, 7$ . The CBTI simulations involved  $K = 16$  replicas, along with  $m_{\Lambda} = 160 \text{ u nm}^2$  and  $\tau_{\Lambda} = 0.5 \text{ ps}$  (no biasing potential). In both cases, the total single-system sampling time was 100 ns, and the calculation was repeated ten times. Considering each set of repeats, the successive entries are: the free-energy change  $\Delta G_{\text{avg}}$  averaged over the repeats; the corresponding standard deviation  $\sigma$  of the mean; the associated error  $\epsilon$  on the mean; alternative error estimates  $\sigma_B$  and  $\sigma_G$  calculated by bootstrapping (no Student  $t$ -factor included). The quantity  $\sigma$  is calculated by scaling the standard deviation of the ten  $\Delta G$  values by the square-root of nine. The error estimate  $\epsilon$  is then calculated as  $\epsilon = 2.262\sigma$ , corresponding to a two-sigma confidence interval of 95%. The estimate  $\sigma_B$  corresponds to a bootstrap error calculated from the concatenated trajectories of the ten repeats (for TI, concatenated separately at each  $\lambda$ -point). The estimate  $\sigma_G$  (TI only) is obtained by calculating the bootstrap errors on the average Hamiltonian derivatives at all  $\lambda$ -point, and propagating them assuming a Gaussian distribution of the estimates, as proposed in Refs. ? ? . The  $\Delta G$  values for the individual repeats can be found in Tabs. 3.E.1 and 2.G.1.

	TI							CBTI
$n$	1	2	3	4	5	6	7	-
$K_{\text{TI}}$ or $K$	3	5	9	17	33	65	129	16
$\Delta G_{\text{avg}}$ [kJ mol $^{-1}$ ]	15.61	4.81	18.97	21.65	21.45	21.40	21.43	21.39
$\epsilon$ [kJ mol $^{-1}$ ]	0.05	0.18	0.13	0.15	0.20	0.17	0.21	0.10
$\sigma$ [kJ mol $^{-1}$ ]	0.02	0.08	0.06	0.07	0.09	0.08	0.09	0.04
$\sigma_B$ [kJ mol $^{-1}$ ]	0.03	0.06	0.05	0.05	0.07	0.08	0.09	0.05
$\sigma_G$ [kJ mol $^{-1}$ ]	0.03	0.06	0.05	0.05	0.05	0.05	0.05	-

129, respectively (Tab. 3.2). Thus, for the system considered here and at constant total sampling time, the use of a larger number of  $\lambda$ -points with a reduced simulation time is advantageous as it reduces the quadrature error without incurring any penalty in terms of sampling error. However, it also induces an extra cost in the form of the preequilibration dead time, which is proportional to the number of  $\lambda$ -points employed. In addition, the insensitivity of the sampling error to the number of  $\lambda$ -points may still break down when the simulation times at the different  $\lambda$ -points become shorter than the corresponding orthogonal relaxation times.

The use of numerous  $\lambda$ -points with insufficient preequilibration is the plague affecting the slow-growth (SG) method,<sup>??</sup> which is rather inaccurate in practice<sup>???</sup> (large error and hysteresis), unless the results are exponentially averaged over repeats with different Boltzmann-distributed initial conditions, as in fast growth<sup>???</sup> (FG). In essence, the CBTI scheme aims at achieving this limit of a very large number of  $\lambda$ -points, but without suffering from issues related to the preequilibration dead time (TI), to an insufficient preequilibration (SG), or to the requirement of exponential averaging over initial conditions (FG).

### 3.4.2 RESULTS USING THE CBTI SCHEME

The exploration of the influence of the CBTI mass parameter  $m_\Lambda$  and thermostat coupling time  $\tau_\Lambda$  considering different numbers  $K$  of replicas is presented in details in Sect. 3.B. The results are summarized in Tab. 3.1 (entries 1-10). They show that when selected within reasonable ranges, the parameters  $m_\Lambda$  and  $\tau_\Lambda$  have only a limited influence on the kinetic-energy exchange between the CB advance variable  $\Lambda$  and the conformational degrees of freedom, on the average temperature  $T_\Lambda$ , on the diffusion constant

$D_\Lambda$ , and on the calculated free-energy change  $\Delta G$ . Based on this exploration, a working choice  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  was selected for the following CBTI calculations.

Using this parameter setting, CBTI calculations were performed using  $K = 8, 16, 32, 64$  or  $128$  replicas along with  $256K^{-1} \text{ ns}$  simulation time for the replica system, *i.e.*  $256 \text{ ns}$  total single-system sampling time in all cases. The detailed results are shown graphically in Fig. 2.F.11 - Fig. 2.F.15 and summarized numerically in Tab. 3.1 (entries 11-15). Illustrative results for  $K = 8, 32$  or  $128$  (entries 11, 13 and 15) are also shown in Figs. 3.3 and 3.5, and discussed in details below.

The time series and probability distributions of  $\Lambda$  for these three simulations are displayed in Fig. 3.3a. For  $K = 8$  replicas (left panel),  $\Lambda$  covers the entire range  $[0, 2\pi]$ , but presents 8 regularly spaced probability peaks at integer multiples of  $2\pi/8$ . This corresponds to CB configurations where two of the replicas are at the physical end-states  $\lambda = 0$  and  $\lambda = 1$ , the six others being by pairs at  $\lambda = 0.25, 0.5$  and  $0.75$ . The occurrence of such probability peaks is not surprising considering the shape of the free-energy profile (Fig. 3.2a). This profile presents minima at  $\lambda = 0$  and  $\lambda = 1$ , the former one being the fully interacting methanol state, which is particularly favored. This significant inhomogeneity in the sampling of  $\Lambda$  correlates with a stepwise (hopping) dynamics in the  $\Lambda$  time series.

In contrast, for  $K = 128$  replicas (right panel),  $\Lambda$  only spans a limited subrange of the interval  $[0, 2\pi)$  and no longer presents regularly spaced probability peaks. For longer simulations, one would expect the probability curve to also cover the entire  $[0, 2\pi)$  range and present regularly spaced peaks at integer multiples of  $2\pi/128$ , but these would be of much smaller magnitudes compared to the  $K = 8$  case. For  $K = 128$ , the effect of these residual

variations becomes almost negligible, *i.e.* there is essentially no driving force on the  $\Lambda$  variable. As a result, instead of a hopping dynamics, the  $\Lambda$  variable now follows a purely diffusive (random-walk) trajectory.

The situation for  $K = 32$  replicas (middle panel) is intermediate. The  $[0, 2\pi)$  range is fully covered, but the distribution is irregular, presenting some small indentations at  $2\pi/32$  spacing.

The reduction of the coverage and homogeneity in terms of  $\Lambda$  over the interval  $[0, 2\pi)$  when going from  $K = 8$  to 32, and then to 128, is not due to a progressively slower (hopping *vs.* diffusive) dynamics, but simply to the different time spans of the simulations (32, 8 and 2 ns, respectively), imposed by the constant total single-system sampling time of 256 ns in the present simulations. This can be seen in Fig. 3.5.

Fig. 3.5a shows the time series and probability distribution of  $\dot{\Lambda}$ , the velocity associated with the  $\Lambda$ -variable. The Maxwell-Boltzmann distribution of Eq. 3.25 is also displayed for comparison. The observed width  $\sigma_{\dot{\Lambda}}$  of the distribution is reported numerically in Tab. 3.1. As expected from Eq. 3.25, it is proportional to  $m_{\Lambda}^{-1/2}$  which, in the present simulations, was itself made proportional to  $K^{1/2}$  (an arbitrary parameter choice justified in Sect. 3.B). Thus, the average magnitude of  $\dot{\Lambda}$  decreases somewhat when going from  $K = 8$  to 32, and then to 128 (scaling factor  $2^{-1/2} \approx 0.7$  each time). Fig. 3.4 shows the normalized autocorrelation functions  $c_{\dot{\Lambda}}$  of  $\dot{\Lambda}$  for different  $m_{\Lambda}$  in the absence of thermostat coupling (entries 1-5 of Tab. 3.1). The corresponding correlation times  $\tau_{\dot{\Lambda}}$  are reported numerically in Tab. 3.1. Here, one observes that the correlation time increases with  $m_{\Lambda}$  as a result of the higher inertia of the CB. In addition, for the two lowest values of  $m_{\Lambda}$ , the initial decay of  $c_{\dot{\Lambda}}$  to negative values suggests that the velocities are frequently reversed at short time, when barriers cannot be

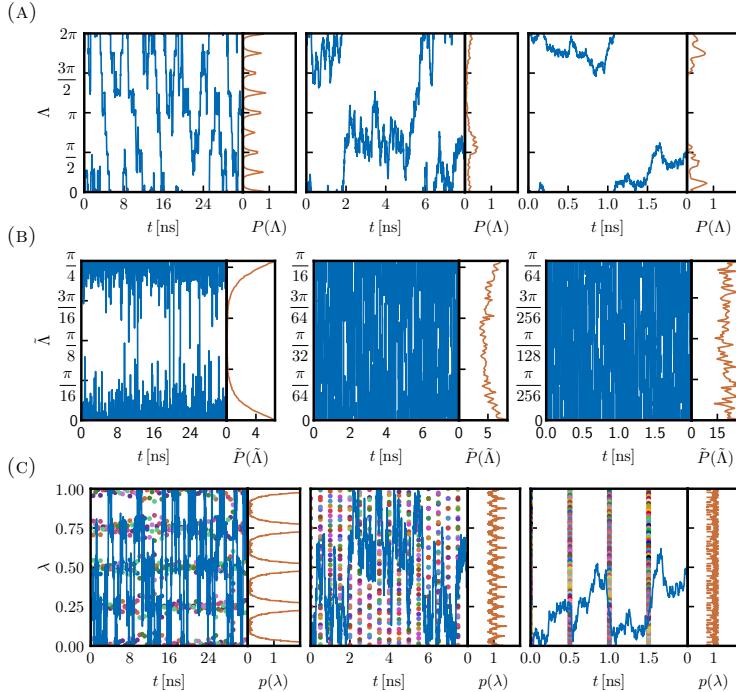


FIGURE 3.3: Time series and probability distributions of the relevant variables in unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. This figure compares the dynamics and distributions of the CB variables  $\Lambda$  (a),  $\tilde{\Lambda}$  (b) and  $\lambda$  (c) in CBTI simulations with  $K = 8$  (left), 32 (middle) or 128 (right) replicas (based on simulations with a total single-system sampling time of 256 ns). Panel (a) shows the time series  $\Lambda(t)$  and probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$ . Panel (b) shows the time series  $\tilde{\Lambda}(t)$  and probability distribution  $\tilde{P}(\tilde{\Lambda})$  of the CB fractional advance variable  $\tilde{\Lambda}$  (Eq. 3.17). Panel (c) shows the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas (Eq. 3.19). In Panel (c), the time series is shown as a blue curve for one reference replica  $k = 0$ , and as individual colored points at 0.5 ns interval for the  $K - 1$  other replicas, and the probability distribution  $p(\lambda)$  is calculated considering all replicas according to Eq. 3.19. All probability distributions are normalized to one. The CBTI simulations relied on  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  (no biasing potential). See also Tab. 3.1 entries 11, 13 and 15 for relevant numerical results. Corresponding graphs also including  $K = 16$  and 64 can be found in Figs. 2.F.11 - 2.F.15.

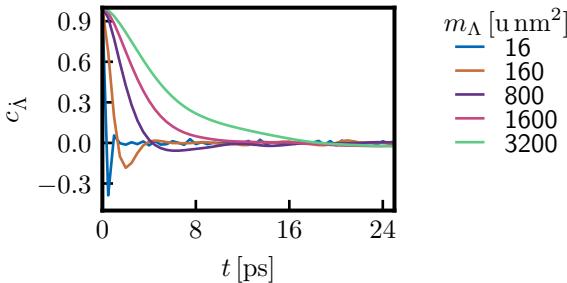


FIGURE 3.4: *Velocity autocorrelation functions of the CB advance variable in unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure shows the increase in the autocorrelation time of the CB velocity upon increasing the mass parameter in CBTI simulations with  $K = 16$  replicas and no thermostat coupling. The normalized autocorrelation function  $c_{\Lambda}$  of the velocity  $\dot{\Lambda}$  is shown for different values of the mass-parameter  $m_{\Lambda}$ , in the absence of thermostating of the  $\Lambda$ -variable ( $\tau_{\Lambda} \rightarrow \infty$ ). See also Tab. 3.1 entries 1–5 for relevant numerical results, including the corresponding autocorrelation times  $\tau_{\dot{\Lambda}}$ . Corresponding graphs for all simulations listed in Tab. 3.1 are shown in Figs. 3.F.1 – 2.F.15

crossed. Note, however, that this connection between  $m_{\Lambda}$  and  $\tau_{\dot{\Lambda}}$  is somewhat altered by the thermostat coupling in the other simulations (including those of entries 11–15 in Tab. 3.1).

Because the two above trends act in opposite directions, the actual rate of diffusion along  $\Lambda$  is not very different in the three simulations.

This is seen in Fig. 3.5c, which shows the mean-square displacement  $d_{\Lambda}$  of  $\Lambda$  as a function of time. The curves are essentially linear, as expected from the Einstein model (Eq. 3.26). The corresponding diffusion constants  $D_{\Lambda}$  are reported in Tab. 3.1, namely 4.4, 6.2 and  $2.3 \text{ ns}^{-1}$  for  $K = 8, 32$  and 128 replicas, respectively, *i.e.* not dramatically different. A similar observation holds for the probability distribution of  $\ddot{\Lambda}$ , the acceleration associated with the  $\Lambda$ -variable, shown in Fig. 3.5b. The choice of taking  $m_{\Lambda}$

proportional to  $K^{1/2}$  was made here precisely to achieve a CB acceleration that is essentially independent of  $K$  (see discussion in Sect. 3.A).

The evolution of  $D_\Lambda$  with  $K$  is clearly non-monotonic (see values for entries 11-15 in Tab. 3.1), which results in large part from the two opposite effects mentioned above concerning the effect of  $m_\Lambda$  on  $\sigma_{\tilde{\Lambda}}$  and  $\tau_{\tilde{\Lambda}}$ . On the one hand,  $\sigma_{\tilde{\Lambda}}$  (*i.e.* the average magnitude of the CB velocity) decreases upon increasing  $m_\Lambda$ , and the dynamics along the  $\Lambda$ -variable becomes intrinsically slower. On the other hand,  $\tau_{\tilde{\Lambda}}$  (*i.e.* the dynamical inertia of the CB) increases upon increasing  $m_\Lambda$ , and the ability to cross residual barriers in the  $G_\Lambda(\Lambda)$  free-energy profile is enhanced. The magnitude of these barriers itself decreases upon increasing  $K$ . In addition, inertia along  $\Lambda$  is also affected by the thermostat coupling. The combination of these effects is complex and the resulting trends in  $D_\Lambda$  upon varying  $K$  would be difficult to rationalize in details. Most importantly, however, they appear not to be extremely pronounced.

The time series and probability distributions of the CB fractional advance variable  $\tilde{\Lambda}$  are shown in Fig. 3.3b. In contrast to the distributions of  $\Lambda$  over the range  $[0, 2\pi]$  (Fig. 3.3a), the distributions of  $\tilde{\Lambda}$  over the range  $[0, \Delta\Lambda]$  are converged and cover the entire range, also for  $K = 128$ . For  $K = 8$ , the distribution of  $\tilde{\Lambda}$  is significantly biased towards 0 and  $2\pi/8$ , *i.e.* CB configurations with two of the replicas at the physical end-states. For  $K = 32$ , it is still somewhat biased towards 0 and  $2\pi/32$ , which corresponds to the same situation. And for  $K = 128$ , the probability distribution becomes very close to homogeneous.

Finally, the time series and probability distributions of the coupling variable  $\lambda$  are shown in Fig. 3.3c for all replicas. The time series is shown for one reference replica  $k = 0$  (blue curve)

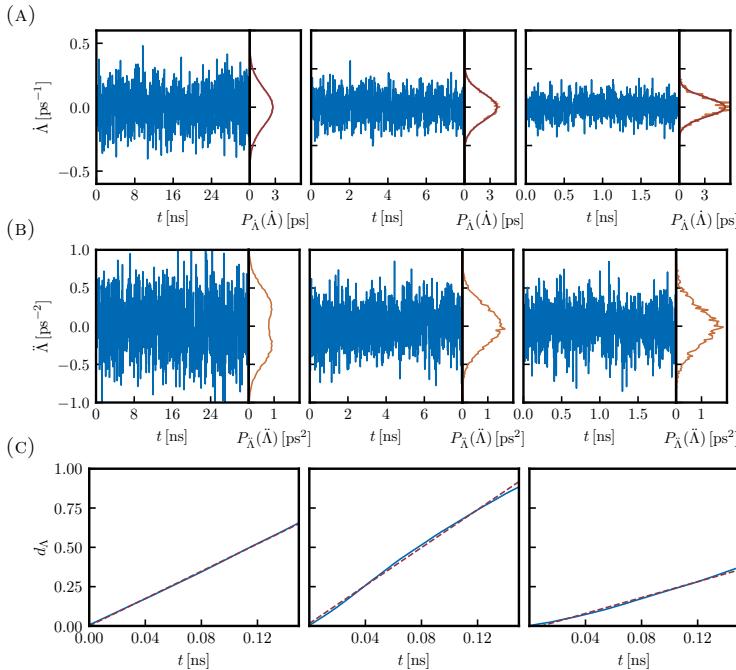


FIGURE 3.5: *Dynamical properties of the CB advance variable in unbiased CBTI calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure compares the dynamics and distributions of the CB variables  $\dot{\Lambda}$  (a) and  $\ddot{\Lambda}$  (b), as well as the mean-square displacement  $d_{\Lambda}$  along  $\Lambda$  (c), in CBTI simulations with  $K = 8$  (left), 32 (middle) or 128 (right) replicas (based on simulations with a total single-system sampling time of 256 ns). Panel (a) shows the time series  $\dot{\Lambda}(t)$  and probability distribution  $P_{\dot{\Lambda}}(\dot{\Lambda})$  of the velocity associated with the CB advance variable  $\Lambda$ . Panel (b) shows the time series  $\ddot{\Lambda}(t)$  and probability distribution  $P_{\ddot{\Lambda}}(\ddot{\Lambda})$  of the corresponding acceleration. Panel (c) shows the mean-square displacement  $d_{\Lambda}(t)$  of  $\Lambda$  (Eq. 3.26). In Panel (a), the probability distribution from the simulation (blue curve) is shown together with the analytical Maxwell-Boltzmann distribution  $P_{\dot{\Lambda}}^{\text{MB}}(\dot{\Lambda})$  of Eq. 3.25 (orange curve). In Panel (c), the mean-square displacement  $d_{\Lambda}$  from the simulation (blue curve) is shown along with a linear least-squares fit over the interval 0 - 0.15 ns (brown dashed line). All probability distributions are normalized to one. The CBTI simulations relied on  $m_{\Lambda} = 40K^{1/2} \text{ u nm}^2$  and  $\tau_{\Lambda} = 0.5 \text{ ps}$  (no biasing potential). See also Tab. 3.1, entries 11, 13 and 15 for relevant numerical results. Corresponding graphs also including  $K = 16$  and 64 can be found in Figs. 2.F.11 - 2.F.15.

as well as the other replicas  $k = 1 \dots K - 1$  (individual colored points at 0.5 ns interval). Compared to the reference replica, these representative points are merely shifted by  $2K^{-1}k$ , refolded by periodicity into the range  $[0, 2)$ , and the points in the range  $[1, 2)$  mirrored into the range  $[0, 1]$  according to Eqs. 3.6- 3.8. The probability distribution  $p(\lambda)$  is calculated according to Eq. 3.19, *i.e.* considering all replicas. Although the  $\Lambda$ -space in the range  $[0, 2\pi)$  is sampled only partially given the simulation lengths (Fig. 3.3a), the  $\lambda$ -space in the range  $[0, 1]$  (Fig. 3.3c), just like the  $\tilde{\Lambda}$ -space in the range  $[0, \Delta\Lambda)$  (Fig. 3.3b), is completely sampled in all cases. However, for  $K = 8$ , the barriers are still too high and the sampling is considerably biased towards  $\lambda$ -values that are integer multiples of  $1/4$ . The corresponding bias towards integer multiples of  $1/16$  is far less pronounced for  $K = 32$ , and the simulation with  $K = 128$  is very close to a homogeneous sampling of  $\lambda$ .

In summary, the sampling of  $\Lambda$  only needs to cover a  $\Delta\Lambda$  range to ensure the sampling of all possible  $\tilde{\Lambda}$ -values by the CB and thus, of all possible  $\lambda$ -values by at least one replica. However, the homogeneity of the sampling in  $\lambda$  depends crucially on  $K$ , quasi-homogeneous sampling requiring a large number of replicas.

This can be seen in Fig. 3.6, which shows in a logarithmic form the height  $G_{\Lambda}^*$  of the peaks in  $G_{\Lambda}(\Lambda)$  as a function of the number  $K$  of replicas (see also Fig. 2.I.1 for the individual curves). The dependency is approximately linear with a slope of minus one, suggesting that  $G_{\Lambda}^*$  scales as  $K^{-1}$ . Such a scaling is expected based on the considerations made in Sect. 3.A. Note that the proportionality constant depends on the derivatives of  $G(\lambda)$  at the two physical end-states. In a situation where these derivatives vanished, a scaling as  $K^{-2}$  would be expected instead.

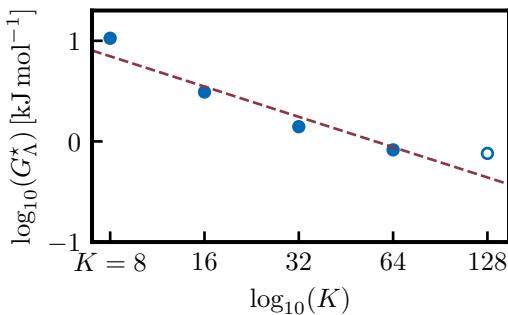


FIGURE 3.6: Height  $G_\Lambda^*$  of the residual barriers in the free-energy profile  $G_\Lambda(\Lambda)$  for unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. This figure shows the progressive decrease in the height of the barriers along  $\Lambda$  upon increasing the number  $K$  of replicas in a CBTI simulation. The data is represented in logarithmic form and a linear line of slope  $-1$  is fitted to the filled circles (intercept 1.75 at  $\log_{10}(K) = 0$ ). The open circle was omitted from the fit because  $G_\Lambda^*$  cannot be determined sufficiently accurately (within the noise). The simulations considered rely on  $K = 8, 16, 32, 64$  or  $128$  replicas, along with  $m_\Lambda = 40K^{1/2}$  u nm<sup>2</sup> and  $\tau_\Lambda = 0.5$  ps (no biasing potential). The corresponding free-energy profiles  $G_\Lambda(\Lambda)$  are shown graphically in Fig. 2.I.1.

### 3.4.3 CBTI WITH BIASING POTENTIAL

As seen above, simulations employing only few replicas may still be affected by a significant sampling inhomogeneity along  $\lambda$  due to high residual barriers along  $\Lambda$ . To prevent the undersampling of specific  $\lambda$ -ranges in this case, a memory-based biasing potential can be employed. This renders the CBTI method more robust since it becomes insensitive to the chosen number of replicas, and more compatible with possible constraints related to *e.g.* the number of cores in a specific computing node during the calculation. As discussed in the Theory section, it is sufficient to define such a biasing potential over a  $[0, \Delta\Lambda/2]$  range of  $\tilde{\Lambda}$ , and its magnitude is limited by the already largely homogeneous sampling of  $\tilde{\Lambda}$ . These features speed up the build-up procedure drastically, leading to a comparatively small simulation time for the LE build-up phase relative to single-system  $\lambda$ -LEUS.

Fig. 3.7 illustrates the sampling in a biased compared to that in the unbiased CBTI simulation with  $K = 8$  replicas. The biasing relied on 34 basis functions over half the  $\tilde{\Lambda}$ -range  $[0, \pi/4]$  (17 free coefficients) and a build-up phase of 0.15 ns (*i.e.* only 1.2 ns total single-system build-up time). During this time, the replica system performed three double sweeps over all grid-points, leading to a biasing potential with an energetic resolution of  $10^{-6} \text{ kJ mol}^{-1}$  (Figs. 3.7a and 3.7b).

The  $\Lambda$ -sampling in the biased CBTI simulation is much closer to uniform compared to that in the unbiased one (Fig. 3.7c *vs.* 3.7d), *i.e.* the relative probabilities between most and least sampled  $\Lambda$ -values has been reduced from 70 to 4.6. The time series of  $\lambda$  and the sampling distribution  $p(\lambda)$  show that the motion of the conveyor belt was hindered in the unbiased simulation, which it is no longer the case in the biased one (Fig. 3.7e *vs.* 3.7f).

Alternative build-up protocols were also explored for  $K = 8$  and  $K = 16$ , differing in the number  $N_{\text{gp}}$  of grid points, in the duration  $t_{\text{LE}}$  of the build-up phase, and in the build-up parameters  $c_{\text{LE}}$  and  $f_{\text{red}}$ . The corresponding results are shown graphically in Figs. 2.H.2 and 2.H.4 and summarized numerically in Tab. 2.H.1. All the tested protocols reduce the sampling inhomogeneity to an acceptable level, although residual inhomogeneities are difficult to remove entirely. Compared to the TI results with a sufficient number of  $\lambda$ -points (Tab. 3.2, entries with  $K_{\text{TI}} \geq 33$ ) and the unbiased CBTI results with a sufficient number of replicas (Tab. 3.1, entries 12–15 with  $K \geq 16$ ), the biased CBTI results with  $K = 8$  or 16 (entries 16 and 17 in Tab. 3.1) deliver similar  $\Delta G$  values and error bars, provided that an appropriate build-up protocol is employed.

### 3.4.4 CBTI FREE-ENERGY ESTIMATOR

The calculation of free-energy differences from CBTI simulations according to Eq. 3.22 relies on the choice of a number  $J$  of bins. This number can be taken very large, thereby reducing the (rectangular) quadrature error to an essentially negligible amount, but should not exceed a threshold  $J_{\max}$  where empty bins occur. The influence of  $J$  is illustrated in Fig. 3.8, which shows the average Hamiltonian derivative curve obtained from the unbiased CBTI simulation with  $K = 32$  replicas (entry 13 in Tab. 3.1, restricted to total 100 ns single-system sampling time) for  $J = 10$  (Fig. 3.8a),  $J = 500$  (Fig. 3.8b) and  $J = 5000$  (Fig. 3.8c). A too low number of bins (*e.g.*  $J = 10$ ) will lead to an insufficient number of integration points, thereby increasing the quadrature error. A too high number of bins will lead to insufficient sampling of each bin, thereby increasing the statistical error. The optimal value of

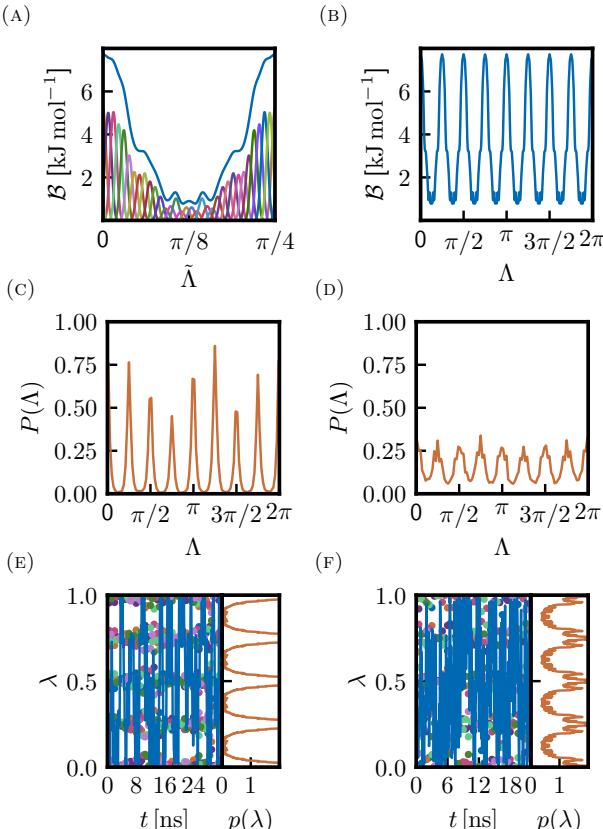


FIGURE 3.7: *Biassing potential and simulation results from unbiased vs. biased CBTI calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure illustrates how the use of a biassing potential permits to drastically reduce the sampling inhomogeneity of CBTI simulations performed with a small number of replicas (here,  $K = 8$  replicas). Panels (a) and (b) show the biassing potential along the  $\tilde{\Lambda}$ - and  $\Lambda$ -ranges, respectively. Panels (c) and (d) show the probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$  in the unbiased and biased simulation, respectively. Panels (e) and (f) show the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas in the unbiased and biased simulation, respectively. In Panel (a), the symmetric biassing potential (blue curve) is represented using 34 basis functions (curves of different colors) involving 17 free coefficients. In Panels (e) and (f), the time series is shown as a blue curve for one reference replica  $k = 0$ , and as single colored points at 0.5 ns spacing for the  $K - 1$  other replicas, and the probability distribution  $p(\lambda)$  is calculated considering all replicas according to Eq. 3.19. All probability distributions are normalized to one. The CBTI simulations were performed with or without biassing potential, using  $m_\Lambda = 113 \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$ . For the biased simulation, the biassing potential was constructed in a build-up phase of 0.15 ns duration. Results for biased simulations with  $K = 16$  and alternative build-up protocols are shown in Fig. 2.H.2 and Fig. 2.H.4, the numerical results being reported in Tab. 2.H.1.

$J$  will depend on the total sampling time and on the degree of homogeneity of the sampling along  $\lambda$ .

To investigate the influence of  $J$  on the calculated  $\Delta G$ , this number was systematically increased from 1 to  $J_{\max}$  in increments ranging from 10 to 1000. The results are shown in Fig. 3.8d, considering the unbiased CBTI simulations with  $K = 8, 16, 32, 64$  or 128 (entries 11-15 in Tab. 3.1, restricted to 100 ns total single-system sampling time) as well as the biased CBTI simulations with  $K = 8$  or 16 (entries 16 and 17 in Tab. 3.1, same restriction). For the unbiased simulation with  $K = 8$  (blue curve), the convergence of  $\Delta G$  upon increasing  $J$  is irregular and the value of  $J_{\max}$  is comparatively low (about 600), due to the inhomogeneous sampling along  $\lambda$ . All other curves present qualitatively similar convergence features. For a low  $J$  ( $J < 100$ ), the quadrature error leads to strong variations. For intermediate  $J$  ( $100 \leq J \leq 3000$ ), the curve presents a plateau, *i.e.* the  $\Delta G$  estimate becomes insensitive to changes in  $J$ . For larger  $J$  ( $3000 < J \leq J_{\max}$ , where  $J_{\max}$  ranges from about 5000 to 8000) the curves become somewhat noisy again due to undersampled bins. These results suggest that, as a rule of thumb, the value of  $J$  for the application of Eq. 3.22 should be selected between at least about 100 and at most about  $J_{\max}/2$ . For most of the  $\Delta G$  results presented in this chapter, a default value  $J = 500$  was selected as leading to a reasonably smooth Hamiltonian derivative curve (Fig. 3.8b) while keeping the quadrature error negligible (except for the CBTI simulation with  $K = 8$ , where  $J = 200$  was used). The results for the variants  $\Delta G_{\text{alt}}$  and  $\Delta G_{\text{app}}$  to  $\Delta G$  from Eq. 3.22 which do not require the specification of a number of bins (Eqs. 3.C.1 and 3.C.2) are also included in Tab. 3.1 for comparison. They are discussed in Sect. 3.C.

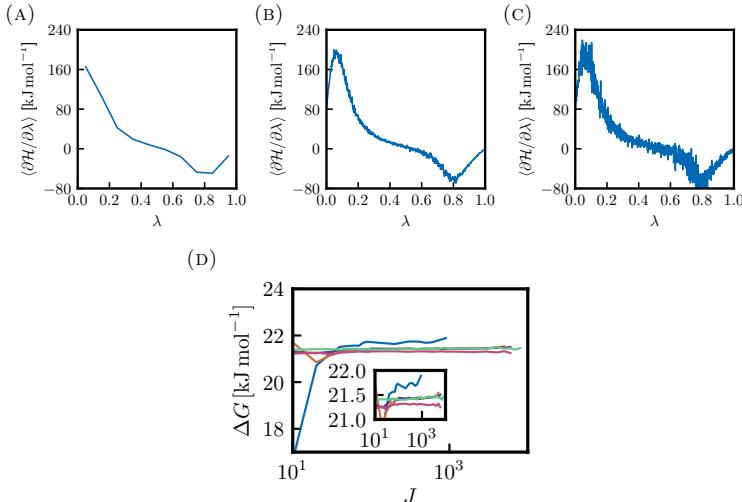


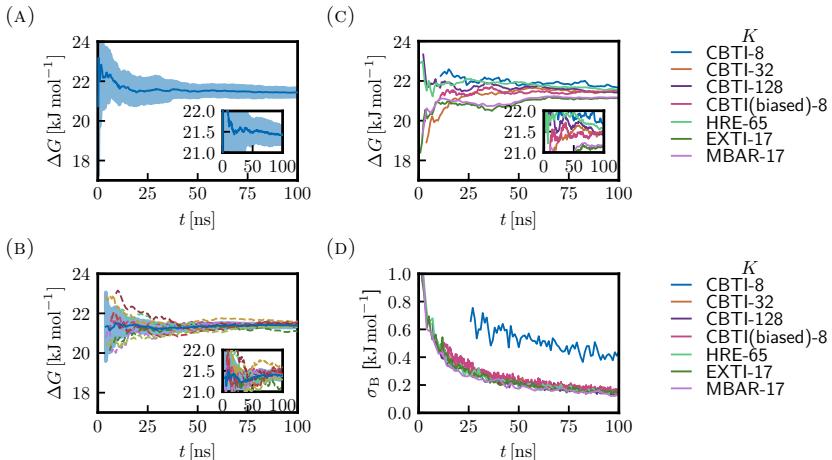
FIGURE 3.8: *Influence of the number of bins used in the free-energy estimator for CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure illustrates how the number  $J$  of bins employed in the estimator affects the quadrature and sampling quality (top) and the resulting estimated  $\Delta G$  (bottom). Panels (a), (b) and (c) show the average Hamiltonian derivative curve constructed with different numbers  $J$  of bins, namely  $J = 10$  (a), 100 (b) or 5000 (c). The curves are based on the CBTI simulation using  $K = 32$  replicas (entry 13 in Tab. 3.1). Panel (d) shows the  $\Delta G$  estimate from Eq. 3.22 for CBTI simulations with  $K = 8, 16, 32, 64$  or 128 replicas (the two former either unbiased or biased; entries 11-17 in Tab. 3.1) as a function of  $J$  from 1 to  $J_{\max}$ , where  $J_{\max}$  corresponds to the first occurrence of empty bins. All CBTI simulations were performed using  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$ , and their analysis is restricted here to 100 ns total single-system sampling time.

### 3.4.5 CBTI CONVERGENCE PROPERTIES

The convergence properties of TI and CBTI are compared in Figs. 3.9a and 3.9b, which shows the  $\Delta G$  values calculated from ten simulation repeats (individual dashed curves) along with the corresponding mean and error (95% confidence interval) on the mean (thick blue curve and shaded area) as a function of the total single-system sampling time per repeat. The simulation protocols compared are TI with  $K_{TI} = 129$   $\lambda$ -points (Fig. 3.9a, which is identical to Fig. 3.2b) and unbiased CBTI with  $K = 16$  (Fig. 3.9b). Note that in the latter case, the curves do not start at time zero because the application of Eq. 3.22 with  $J = 500$  to estimate  $\Delta G$  only becomes possible when  $J_{\max}$  exceeds 500 (*i.e.* each of the 500 bins encompasses at least one sampled  $\lambda$ -value). The mean  $\Delta G$  values over the repeats and the associated error bars at full single-system sampling time (100 ns per repeat) are reported in Tab. 3.2 (last two columns; see also Tab. 3.E.1 and Tab. 2.G.1 for the results of the individual repeats).

The comparison clearly evidences an enhanced convergence for CBTI at identical sampling time, both in terms of the running averages and of the final values. These final values are  $21.43 \pm 0.21$  kJ mol<sup>-1</sup> for TI, compared to  $21.39 \pm 0.10$  kJ mol<sup>-1</sup> for CBTI, *i.e.* identical within error bars, but with an error reduced by about a factor two for CBTI. This improved convergence is likely due to the improved orthogonal sampling, *i.e.* the diffusion of the systems along  $\lambda$  permits to circumvent orthogonal barriers present at specific  $\lambda$ -values.

The same observation holds for the CBTI simulations involving different numbers of replicas or/and a biasing potential. This is shown in Figs. 3.9c and 3.9d for  $K = 8, 32$  or  $128$  (unbiased) and for  $K = 8$  (biased). Here, a single simulation was performed



**FIGURE 3.9:** *Convergence properties of the CBTI calculations compared to other methods for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This figure compares the convergence properties of the calculated free-energy change  $\Delta G$  considering TI (a) vs. CBTI (b) based on ten repeats of calculations involving 100 ns total single-system sampling time, or considering CBTI (biased or unbiased) vs. TI, HRE, TI/EXTI and TI/MBAR in terms of estimated  $\Delta G$  (c) and error (d) based on one calculation involving 100 ns total single-system sampling time. Panel (a) is identical to Fig. 3.2b and shows the convergence of  $\Delta G$  for TI calculations with  $K_{\text{TI}} = 129$   $\lambda$ -points. Panel (b) displays the corresponding convergence of  $\Delta G$  for unbiased CBTI calculations with  $K = 16$  replicas. Panels (a) and (b) show the running  $\Delta G$  estimates for each of the ten repeats (dashed curves) as a function of the total sampling time  $t$  per repeat, along with the corresponding mean and error on the mean (95% confidence interval) over the repeats (thick blue curve and shaded area). The final  $\Delta G$  values for the individual repeats are reported in the Tab. 2.G.1, and the statistics over the repeats can be found in Tab. 3.2. Panel (c) compares convergence properties in terms of  $\Delta G$  based on single simulations. Panel (d) displays the associated errors  $\sigma_B$  calculated by bootstrapping (no Student  $t$ -factor included). Panels (c) and (d) consider unbiased CBTI ( $K = 8, 32$  or  $128$ ), biased CBTI ( $K = 8$ ) HRE ( $K_{\text{HRE}} = 65$ ), along with TI with EXTI or MBAR estimators ( $K_{\text{TI}} = 17$  and  $129$  virtual  $\lambda$ -points in both cases). The corresponding final  $\Delta G$  estimates are reported in Tab. 3.3.

TABLE 3.3: *Free-energy estimates from the CBTI calculations compared to other methods for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* This table compares the convergence properties of the calculated free-energy change  $\Delta G$  considering CBTI (biased or unbiased), TI, HRE, EXTI and MBAR calculations with different numbers of replicas (all at a constant total single-system sampling time of 100 ns). The CBTI calculations involved  $K$  replicas, along with  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  (with or without biasing potential). The TI calculations involved  $K_{\text{TI}}$  equispaced  $\lambda$ -points. The HRE calculations involve  $K_{\text{HRE}}$  equispaced  $\lambda$ -points, along with  $\tau_{\text{HRE}} = 0.2 \text{ ps}$ . For EXTI and MBAR, the calculations relied on  $K_{\text{TI}}$  equispaced real  $\lambda$ -points, and consideration of 129 equispaced virtual  $\lambda$ -points. In all cases, the total single-system sampling time was 100 ns. The corresponding convergence properties are illustrated graphically in Fig. 3.9. Additional numerical results regarding the CBTI simulations are reported in Tab. 3.1, entries 11–15 (unbiased CBTI) and entries 16–17 (biased CBTI).

$K, K_{\text{TI}} - 1$ or $K_{\text{HRE}} - 1$	$\Delta G [\text{kJ mol}^{-1}]$					
	CBTI (unbiased)	CBTI (biased)	TI	HRE	EXTI	MBAR
8	21.69 (0.40)	21.48 (0.16)	19.22 (0.18)	-	21.24 (0.37)	21.37 (0.15)
16	21.42 (0.16)	21.30 (0.13)	21.87 (0.17)	21.52 (0.16)	21.12 (0.14)	21.17 (0.13)
32	21.44 (0.14)	-	21.48 (0.20)	21.39 (0.15)	-	-
64	21.32 (0.14)	-	21.18 (0.24)	21.47 (0.13)	-	-
128	21.43 (0.12)	-	21.45 (0.30)	-	-	-

(no repeats), and the running  $\Delta G$  value (Fig. 3.9c) is displayed along with the corresponding bootstrap error estimate (Fig. 3.9d; no Student  $t$ -factor included) as a function of the total single-system sampling time. The  $\Delta G$  estimates and bootstrap error bars at full single-system sampling time (100 ns) are reported in Tab. 3.3 (second and third columns). Except for the unbiased CBTI simulation with  $K = 8$ , where the sampling along  $\lambda$  is still heterogeneous, all the CBTI protocols follow very similar error curves upon increasing the sampling time (Fig. 3.9d), with final errors of  $0.12 - 0.16 \text{ kJ mol}^{-1}$ . In contrast, the convergence of TI is significantly slower, with corresponding final errors of  $0.17 - 0.30 \text{ kJ mol}^{-1}$  (fourth column in Tab. 3.3).

### 3.4.6 COMPARISON OF CBTI WITH OTHER METHODS

The  $\Delta G$  values and associated bootstrap errors (no Student  $t$ -factor included) calculated using various CBTI protocols are compared to the results of other approaches in Tab. 3.3. The comparison involves calculations relying on 100 ns total single-system sampling time, *i.e.* essentially identical computational costs. The protocols compared are unbiased CBTI (with  $K = 8, 16, 32, 64$  or 128), biased CBTI (with  $K = 8$  or 16), TI (with  $K_{\text{TI}} = 9, 17, 33, 65$  or 129), HRE (with  $K_{\text{HRE}} = 17, 33$  or 65), along with TI/EXTI and TI/MBAR (both with  $K_{\text{TI}} = 9$  or 17, and using 129 virtual  $\lambda$ -points). More details on the HRE, TI/EXTI and TI/MBAR calculations can be found in Figs. 2.J.1 - 2.L.1.

As previously discussed, the results of the TI calculations with  $K_{\text{TI}} = 9$  or 17 are affected by large quadrature errors (see Fig. 3.2c and Tab. 3.2), and those of the unbiased CBTI simulation with  $K = 8$ , are affected by a significant heterogeneity of the sampling along  $\lambda$  (see Fig. 3.3). Excepting these three calculations, all free-energy estimates of Tab. 3.3 are essentially compatible, taking into account that the error bars exclude a Student  $t$ -factor.

The convergence of the HRE ( $K_{\text{HRE}} = 65$ ), TI/EXTI ( $K_{\text{TI}} = 17$ ) and TI/MBAR ( $K_{\text{TI}} = 17$ ) protocols are also compared to those of TI ( $K_{\text{TI}} = 129$ ), unbiased CBTI ( $K = 8, 32$  or 128) and biased CBTI ( $K = 8$ ) in Figs. 3.9c and 3.9d. Considering Fig. 3.9d, the uncertainty of the unbiased CBTI simulation with  $K = 8$  is found to be particularly large, owing to the sampling heterogeneity. Excepting this simulation, the uncertainty is the largest for the TI estimate. The different CBTI protocols reduce the error by about a factor two compared to TI, presumably due to enhanced orthogonal sampling. This sampling advantage is shared

by HRE, for which the error curve is very similar. Advanced free-energy estimators such as EXTI and MBAR also improve the convergence over plain TI in a different way, namely by increasing the statistical efficiency of the  $\Delta G$  estimation. The error reduction is also about a factor two compared to TI, which is probably not entirely coincidental. If two fixed- $\lambda$  simulations are trapped in different orthogonal configurational wells in TI, the variable- $\lambda$  approaches (CBTI, HRE) will promote well-transitions, whereas the advanced-estimator approaches (EXTI, MBAR) will import statistical information on the other well from the neighboring  $\lambda$ -point. One might refer to these two types of effects orthogonal-sampling *vs.* orthogonal-statistics advantages, respectively. For the simple system considered here, both results in comparable convergence improvements. This suggests that the two effects may not be cumulative if one generalizes the current CBTI estimator of Eq. 3.22 to an EXTI- or MBAR-type estimator.

### 3.5 CONCLUSION

In the present chapter, we proposed a new method called conveyor belt thermodynamic integration (CBTI) to calculate alchemical free-energy differences based on MD simulations. This approach borrows and combines ideas from thermodynamic integration<sup>???</sup> (TI), replica exchange<sup>???</sup> (HRE) or permutation<sup>???</sup> (HRP), and  $\lambda$ -dynamics<sup>???</sup> ( $\lambda$ D), along with the real-life working principle of the funicular.

In CBTI, one simulates in parallel a set of  $K$  equally spaced replicas (with  $K$  even) on a forward-turn-backward-turn path along the alchemical coupling variable  $\lambda$ , akin to a conveyor belt (CB) between the two physical end states. Because the  $\lambda$ -forces (Hamiltonian  $\lambda$ -derivative) exerted by the individual replicas on the CB largely compensate each other, the overall  $\Lambda$ -force on the CB advance variable  $\Lambda$  becomes increasingly small when  $K$  is made increasingly large (residual free-energy barriers decreasing at least as  $K^{-1}$  in the limit of large  $K$ , as shown in Sect. 3.A). As a result, for a sufficient number  $K$ , quasi-homogeneous sampling of the  $\lambda$ -range can be achieved without application of any biasing potential. If a smaller  $K$  is employed, a memory-based biasing potential can still be added to further homogenize the sampling, the preoptimization of which is computationally inexpensive. The results of a CBTI simulation (whether biased or not) can be analyzed similarly to TI, by binning of the average Hamiltonian  $\lambda$ -derivative as a function of  $\lambda$  considering all replicas jointly, followed by quadrature integration. In this case, the continuous and quasi-homogeneous sampling of the  $\lambda$ -range permits to use a large number of bins, thereby essentially eliminating quadrature errors.

As a first application, the CBTI scheme was employed here to calculate the hydration free energy of methanol. It was shown that the method is rather robust with respect to the choice of its parameters ( $K$  as well as the mass-parameter  $m_\Lambda$  and thermostat coupling time  $\tau_\Lambda$  of the CB), the most important sensitivity being relative to  $K$ . Upon increasing  $K$ , the distribution/dynamics of  $\Lambda$  evolves from regularly spaced preferential values with a hopping dynamics to quasi-homogeneous coverage with a diffusive dynamics. For the smallest number of replicas considered ( $K = 8$ ), application of a biasing potential is recommended. For larger numbers of replicas ( $K \geq 16$ ), it becomes unnecessary. The calculated  $\Delta G$  values compare well with those obtained using other methods.

The convergence is accelerated relative to TI with Simpson quadrature (smaller error bar at identical total single-system sampling time), owing to improved orthogonal sampling and reduced quadrature errors. It is comparable to HRE, which shares the same orthogonal-sampling advantage. It is also similar to TI with EXT1 or MBAR as free-energy estimator, which achieve a similar improvement *via* an orthogonal-statistics advantage, *i.e.* by effectively mixing information concerning distinct configurational wells across  $\lambda$ -points. It should be stressed, however, that the present mutation is rather non-challenging in terms of orthogonal sampling. Work is in progress to investigate other types of systems with more complicated orthogonal spaces: (*i*) the side-chain mutation in the central residue of a tripeptide considered in Refs. ? ? ? ; (*ii*) the hydrogen-to-bromine mutation in the base of a nucleotide considered in Refs. ? ? ?. Here, it is expected that CBTI alone (just like HRE) will help overcoming barriers in the orthogonal space when these barriers are low at some  $\lambda$ -value (as in the first system mentioned), but may be insufficient on its

own when these barriers are high at all  $\lambda$ -values (as in the second system mentioned), in which case additional modifications must be applied to create artificially an orthogonal tunnel at least over a limited  $\lambda$ -range.

Compared to existing MD-based alchemical free-energy calculation methods, the CBTI scheme can be viewed in at least three different ways: (*i*) as a continuous/deterministic/dynamical (instead of discrete/stochastic) analog of the HRE scheme<sup>???</sup> or the HRP scheme;<sup>???</sup> (*ii*) as a correlated multiple-replica analog (reminiscent of other swarm,<sup>???</sup> multiple-walker<sup>??</sup> or flying-Gaussian<sup>??</sup> approaches) of the  $\lambda$ -local elevation umbrella sampling ( $\lambda$ -LEUS) scheme<sup>???</sup> (or the conceptually similar flat-histogram<sup>??</sup>  $\lambda$ -metadynamics,<sup>???</sup> adaptive integration,<sup>?</sup> adaptive biasing force,<sup>?</sup> adaptively biased<sup>?</sup> and expanded-ensemble<sup>???</sup> methods); (*iii*) as an equilibrium multiple-replica variant of the slow-growth<sup>??</sup> (SG) method (bypassing the associated hysteresis issues<sup>???</sup> or the requirement for exponential averaging over multiple repeats<sup>???</sup> ).

Compared to plain TI, it shares the advantage of HRE/HRP and  $\lambda$ -LEUS in terms of enhanced orthogonal sampling.<sup>???</sup> Compared to HRE/HRP, it permits a deterministic and continuous sampling of the  $\lambda$ -range, and bypasses the need for a careful preselection<sup>???</sup> of the  $\lambda$ -ladder and exchange-attempt interval. Compared to both TI and HRE/HRP, the quasi-homogeneous  $\lambda$ -sampling also essentially removes quadrature errors. Finally, compared to  $\lambda$ -LEUS, it eliminates (or drastically reduces) the dead time associated with the preoptimization of a biasing potential<sup>?</sup> or, alternatively, the use of this formally non-equilibrium statistics<sup>?</sup> in the production calculation.<sup>?</sup> For the above reasons, the CBTI scheme certainly represents a useful addition to the alchemical free-energy calculation toolkit.

Like TI and HRE/HRP, the CBTI method is also intrinsically parallel. However, assuming that the replicas are assigned to separate processors (including possible GPU implementations or/and cloud-computing applications), the requirement of an all-to-all information exchange between processors at every timestep might represent a drawback of the method relative to the no-exchange and infrequent exchange situations of TI and HRE/HRP, respectively. Although the communication is lightweight (Hamiltonian  $\lambda$ -derivative, *i.e.* a single real number), the synchronization requirement may cause a performance loss (reduced scalability and fault tolerance). Unless asynchronous variants<sup>??</sup> or multiple-timestep schemes<sup>??</sup> can be developed, this performance loss may represent a problem for parallel applications in situations involving many replicas of a small system, as these will involve more data exchange at a more frequent rate.

This scheme opens the way to at least two types of generalizations and extensions. First, a number of components of the scheme can be modified/generalized and in particular the following: (*i*) different functions  $\zeta$  may be used in Eq. 3.6 to modulate the replica density along  $\lambda$  (*e.g.* smoothing the tips or adding plateaus<sup>?</sup> for a denser sampling close to or at the physical end-states); (*ii*) different matrices  $\underline{\mathbf{C}}$  may be used in Eq. 3.14 corresponding to a non-uniform weighting of the  $\lambda$ -forces into the  $\Lambda$ -force, (*e.g.* canceling the effect of higher-order derivatives by alternating non-integer weights in analogy to standard quadrature methods); (*iii*) the coupling between replicas may be generalized from sequential pairwise constraints to possibly non-pairwise potentials (*e.g.* collective or harmonic); (*iv*) the TI-like free-energy estimator of Eq. 3.22 may be replaced by a statistically more powerful one of the MBAR type;<sup>???</sup> (*v*) CBTI would benefit from the use of an alchemical coupling path presenting a vanishing free-

energy derivative at the physical end-states (residual free-energy barriers along  $\Lambda$  decaying at least in  $K^{-2}$  instead of  $K^{-1}$  for large  $K$ , as shown in Sect. 3.A).

Second, the application range of CBTI, restricted here to alchemical processes, can be extended to encompass either thermodynamic or conformational free-energy changes. In the former case, extension to a CB variant of parallel tempering<sup>?</sup> (*i.e.* a CBPT scheme) appears relatively straightforward considering that scaling the temperature is equivalent to scaling the system potential energy. Such a CBPT scheme would represent a form of multicanonical sampling.<sup>???</sup> In the latter case, extension to a CB variant of umbrella sampling<sup>??</sup> (*i.e.* a CBUS scheme) could be designed *e.g.* by anchoring harmonic potentials to the CB and propagating the corresponding restraining forces onto the CB. Finally the extension of the CB approach to multistate problems<sup>???</sup> (*e.g.* path or network of CBs connecting the different states of a system) as well as multidimensional problems (sub-CBs anchored to a main CB) can also be envisioned.

## Appendix 3.A RELATIONSHIP BETWEEN CBTI AND QUADRATURE INTEGRATION

This appendix investigates the convergence properties of the CBTI scheme upon increasing the number  $K$  of replicas based on an analogy with quadrature integration. More precisely, it is shown that the variation amplitude  $G_\Lambda^*$  of  $G_\Lambda(\Lambda)$  decreases at least as  $K^{-1}$  in the limit of large  $K$  (with  $K$  even), as suggested by Fig. 3.1b and 3.6.

To this purpose, we first observe that the situation of a conveyor belt (CB) spanning the  $\lambda$ -range  $[0, 1]$  of the free-energy profile  $G(\lambda)$  with  $K$  equidistant replicas at spacings  $2K^{-1}$  along the cable (Fig. 3.A.1a) is mathematically equivalent to the situation of a train spanning a periodic function  $g(x)$  of period  $2\pi$  with  $K$  equidistant carriages at spacings  $2\pi K^{-1}$  spanning one period (Fig. 3.A.1b). The function  $g(x)$  is obtained by mirroring  $G(\lambda)$  at  $\lambda = 0$  or  $1$ , periodically translating the mirrored function by integer multiples of two, and stretching  $\lambda$  by a factor of  $\pi$  to define the new variable  $x$ .

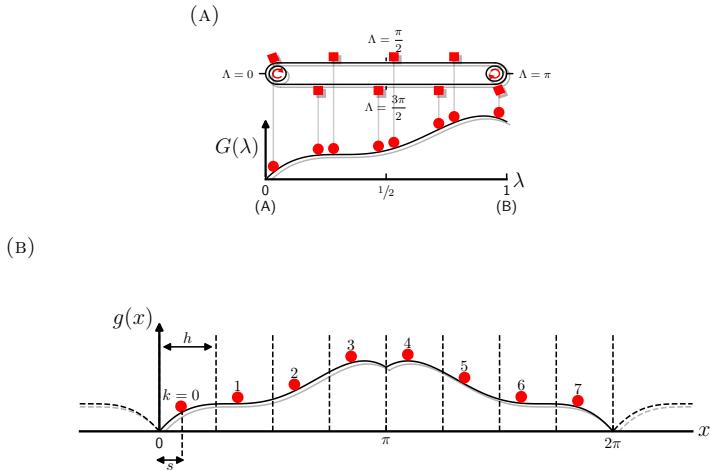


FIGURE 3.A.1: *Analogy between the CBTI scheme and the sampling of an even periodic function by a train of equispaced points.* The illustration considers a conveyor belt (CB) with  $K = 8$  replicas, analog to a train with 8 carriages.

In this analogy, we now have a periodic function  $g(x)$  of period  $2\pi$  that is even over the reference interval  $[0, 2\pi]$ . This function is sampled by  $K$  equidistant points at locations

$$x_k = s + kh \quad \text{with} \quad k = 0 \dots K-1 \quad \text{and} \quad h = 2\pi K^{-1}, \quad (3.A.1)$$

where  $s$  represents the offset position of the first point  $k = 0$ . Note that  $s$  can always be selected in the interval  $[0, h)$  by consideration of periodicity, along with an appropriate renumbering of the points carriages. We also introduce the negative first derivative of the function

$$f(x) = -g'(x). \quad (3.A.2)$$

In practical applications of CBTI, the Hamiltonian-coupling scheme will be defined by a continuously differentiable function of  $\lambda$ , in which case  $G(\lambda)$  is also continuously differentiable. The same applies to  $g(x)$  in the analogy of Fig. 3.A.1b, except for the possible occurrence of kinks at integer multiples of  $\pi$ . These occur because the derivative  $G'(\lambda)$  of  $G(\lambda)$  generally differs from zero at the physical end-states  $\lambda = 0$  or  $\lambda = 1$ . As a result,  $f(x)$  is continuous except at these points, where it will present jumps. An illustrative example for functions  $g(x)$  and  $f(x)$  with the above properties is shown in Figs. 3.A.2a and 3.A.2b, respectively.

In the analogy of Fig. 3.A.1b to the CB situation, the total force on the train of points (analog to the mean force on the  $\Lambda$ -variable) is given for a specific value of  $s$  (analog to the  $\tilde{\Lambda}$ -variable) by

$$F(s) = \sum_{k=0}^{K-1} f(kh + s). \quad (3.A.3)$$

As illustrated in Fig. 3.A.2c, this sum has a simple interpretation. It represents a trapezoidal quadrature estimate to the integral of the periodic function  $f(x)$  over the period  $[s, 2\pi + s]$ . Because  $f(x)$  is the negative derivative of the periodic function  $g(x)$ , its exact integral over one period must be zero. Thus, for any  $s$ , we expect  $F(s)$  to converge to zero in the limit  $K \rightarrow \infty$ . The following paragraphs investigate the corresponding convergence rate.

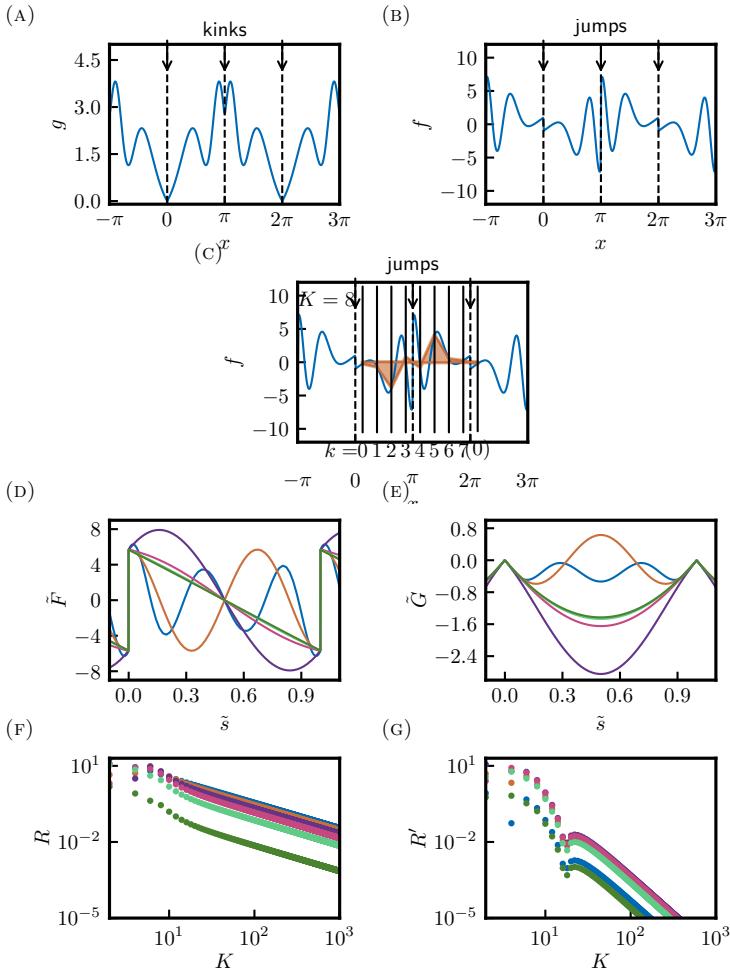


FIGURE 3.A.2: *Illustrative functions supporting the discussion in Sect. 3.A.* Panel (a) shows the illustrative function  $g(x)$  defined over the reference period  $[0; 2\pi]$  as  $x + \sin(x^2)$  if  $x < \pi$  or  $(2\pi - x) + \sin((2\pi - x)^2)$  if  $x \geq \pi$ . Panel (b) shows the negative derivative  $f(x)$  of  $g(x)$ . Panel (c) illustrates the trapezoidal integration of  $f(x)$  (blue) as a shaded area (orange) with  $K = 8$  and  $s = 0.4$ . The vertical lines show the positions of the carriages at values  $x_m$ . The trapezoids are seen to represent a poor approximation at the discontinuities. Panel (d) illustrates the total force  $\tilde{F}(\tilde{s})$  of Eq. 3.A.4 for different values of  $K$ . Panel (e) illustrates the potential energy  $\tilde{G}(\tilde{s})$  for different numbers of replicas  $K$ , which is the running integral of  $\tilde{F}$  for different numbers of replicas  $K$ . Panel (f) shows in a log-log form the residual integral  $R(\tilde{s})$  of Eq. 3.A.8 as a function of  $K$  for different values of  $\tilde{s}$ . Panel (g) shows in a log-log form the corrected residual integral  $R'(\tilde{s})$  of Eq. 3.A.10. Note that the logarithm of the absolute value of  $R$  or  $R'$  is displayed. The cusp in Panel (g) is explained by a change of sign.

The function  $F(s)$  is periodic of period  $h$ , odd over the reference interval  $[0, h)$  and generally discontinuous at  $s$  values which are integer multiples of  $h$ . Introducing

$$\tilde{s} = h^{-1}s \quad \text{and} \quad \tilde{F}(\tilde{s}) = F(h\tilde{s}) = \sum_{k=0}^{K-1} f((k + \tilde{s})h), \quad (3.A.4)$$

the function  $\tilde{F}(\tilde{s})$  is periodic of period 1, odd over the reference interval  $[0, 1)$  and generally discontinuous at integer values of  $\tilde{s}$ .

The corresponding potential energy of the train of points (analog to the free-energy profile  $G_\Lambda(\Lambda)$  of the CB over the interval  $[0, 2\pi K^{-1}]$ ) is given by

$$G(s) = - \int_0^s ds' F(s'). \quad (3.A.5)$$

The function  $G(s)$  is periodic of period  $h$ , even over the reference interval  $[0, h)$  with an extremum at  $h/2$ , and possibly presents kinks at  $s$  values which are integer multiples of  $h$ . Using the variable  $\tilde{s}$  and introducing

$$\tilde{G}(\tilde{s}) = - \int_0^{\tilde{s}} d\tilde{s}' \tilde{F}(\tilde{s}') , \quad (3.A.6)$$

one may write

$$G(s) = h\tilde{G}(h^{-1}s). \quad (3.A.7)$$

The function  $\tilde{G}(\tilde{s})$  is periodic of period 1, even over the reference interval  $[0, 1)$  with an extremum at  $1/2$ , and possibly presents kinks at integer values of  $\tilde{s}$ .

The functions  $\tilde{F}(\tilde{s})$  and  $\tilde{G}(\tilde{s})$  corresponding to the illustrative

functions  $g(x)$  and  $f(x)$  are shown in Figs. 3.A.2d and 3.A.2e, respectively, considering the choices  $K = 2, 4, 8, 16, 32$  or  $64$ . Convergence is observed upon increasing  $K$ , in which  $\tilde{F}(\tilde{s})$  becomes linear and  $\tilde{G}(\tilde{s})$  parabolic over the reference interval  $[0, 1]$ . Thus, in the limit of large  $K$ , one expects the maximal variation  $\tilde{G}^*$  in  $\tilde{G}(\tilde{s})$  to converge to a constant and, *via* Eq. 3.A.7, the maximal variation  $G^*$  in  $G(s)$  (analog to the corresponding maximal variation  $G_\Lambda^*$  of  $G_\Lambda(\Lambda)$  in the CB scheme) to scale as  $h$ , *i.e.* as  $K^{-1}$ .

This observation, made here in a special case, can be generalized as follows. Given

$$R(\tilde{s}) = h\tilde{F}(\tilde{s}) = h \sum_{k=0}^{K-1} f(h(k + \tilde{s})), \quad (3.A.8)$$

$G^*$  is the absolute extremum (largest absolute value) of the function

$$S(\tilde{s}) = h\tilde{G}(\tilde{s}) = \int_0^{\tilde{s}} d\tilde{s}' R(\tilde{s}') \quad (3.A.9)$$

over the  $\tilde{s}$ -interval  $[0, 1/2]$ . According to Fig. 3.A.2c, the quantity  $R(\tilde{s})$  is the residual of the trapezoidal quadrature approximation to the integral of  $f(x)$  over one period, where the first grid point is offset from the origin by a fraction  $\tilde{s}$  of the spacing  $h$ . If  $f(x)$  was continuous everywhere, the convergence would be quadratic in  $h$  (*i.e.* scale as  $K^{-2}$ ), as expected from a trapezoidal quadrature. However, the presence of discontinuities in  $f(x)$  at  $0$  and  $\pi$  introduces an error that is linear in  $h$ . To recover quadratic convergence, one would need to introduce a correcting term for the points  $0, K/2 - 1, K/2$  and  $K - 1$  surrounding the discontinuities, namely

$$\begin{aligned}
R'(\tilde{s}) = & R(\tilde{s}) + h \left\{ \frac{\tilde{s}-1}{2} \left[ f(h\tilde{s}) + f \left( h \left( \tilde{s} + \frac{K}{2} \right) \right) \right] \right. \\
& - \frac{\tilde{s}}{2} \left[ f(h(\tilde{s} + K - 1)) + f \left( h \left( \tilde{s} + \frac{K}{2} - 1 \right) \right) \right] \\
& \left. + \left( \tilde{s} - \frac{1}{2} \right) (\Delta_0 + \Delta_\pi) \right\} \tag{3.A.10}
\end{aligned}$$

where  $\Delta_0$  and  $\Delta_\pi$  account for the magnitude of the discontinuities, *i.e.*

$$\Delta_0 = \lim_{x \xrightarrow{>} 0} f(x) \quad \text{and} \quad \Delta_\pi = \lim_{x \xrightarrow{>} \pi} f(x). \tag{3.A.11}$$

The convergence of  $R(\tilde{s})$  and  $R'(\tilde{s})$  for the illustrative functions  $g(x)$  and  $f(x)$  upon increasing  $K$  along with the choices  $\tilde{s} = 0.01, 0.2, 0.3, 0.4, 0.49$  is shown in logarithmic form in Figs. 3.A.2f and 3.A.2g, respectively. As expected, for large  $K$ ,  $R$  shows a linear convergence (slope -1) and  $R'$  a quadratic one (slope -2). The limiting cases  $\tilde{s} = 0$  and  $\tilde{s} = 1/2$  are special. For  $\tilde{s} = 0$ , Eq. 3.A.8 implies an evaluation of  $f(x)$  at the discontinuity. If one uses an average value of 0 for these points,  $R(0)$  evaluates to 0 for any  $K$  by symmetry. The same applies for  $\tilde{s} = 1/2$ ,  $R(1/2)$ . For all other values for  $\tilde{s}$ , the convergence is as expected, *i.e.* linear for  $R(\tilde{s})$  and quadratic for  $R'(\tilde{s})$ .

If  $R(\tilde{s})$  converges to zero as  $K^{-1}$  for (nearly) all  $\tilde{s}$  values, the function  $S(\tilde{s})$  of Eq. 3.A.9 will converge to zero with the same scaling, and so will its absolute extremum  $G^*$ . We conclude that the magnitude  $G_\Lambda^*$  of the residual variations in  $G_\Lambda(\Lambda)$  for the CBTI scheme scales at least as  $K^{-1}$  in the limit of large  $K$ . And if  $G(\lambda)$  has a vanishing derivative at the physical end-states, it will

scale at least as  $K^{-2}$ . It should be stressed, however, that these are worst-case scalings. A higher level of continuity or specific symmetry properties of  $G(\lambda)$  may tighten the scaling. To give an extreme case, for a free-energy profile  $G(\lambda) = \cos(2\pi\lambda)$ , one shows easily that the residual variations  $G_{\Lambda}^*$  entirely vanish irrespective of  $K$ . This follows from

$$\sum_{k=0}^{K-1} \cos(2\pi K^{-1}k + c) = 0 \quad \forall c \quad (\text{for } K \text{ even}). \quad (3.A.12)$$

The present analysis connects the convergence of the CBTI scheme with  $K$  to the properties of a quadrature integration. This connection opens interesting tracks for future work. In particular, it suggests that these convergence properties could be improved by altering the coupling scheme (*e.g.* to enforce vanishing  $G(\lambda)$  derivatives at the end-states), the function  $\zeta$  in Eq. 3.6 (*e.g.* higher system density close to discontinuities) and the CBTI weighting in Eq. 3.4 (*e.g.* from trapezoidal to higher-level quadrature, *e.g.* Simpson or even Romberg).

## Appendix 3.B CHOICE OF THE CBTI PARAMETERS

Here, we explore the influence of the CBTI mass parameter  $m_{\Lambda}$  and thermostat coupling time  $\tau_{\Lambda}$  considering different numbers  $K$  of replicas.

A first set of simulations involves the choices  $m_{\Lambda} = 16, 160, 800, 1600$  or  $3200 \text{ u nm}^2$ , along with  $K = 16$  replicas in the absence of separate thermostat coupling for  $\Lambda$ , *i.e.*  $\tau_{\Lambda} \rightarrow \infty$  (10 ns

simulations after 0.2 ns equilibration). The results are shown graphically in Figs. 3.F.1 - 3.F.5 and reported numerically in Tab. 3.1 (entries 1-5). In the absence of explicit thermostating for the  $\Lambda$ -variable, the average temperature  $T_\Lambda$  of this variable is controlled exclusively by its coupling to the conformational degrees of freedom, themselves thermostated by the application of SD with a reference temperature of 298.15 K and an effective coupling time of 0.1 ps. All choices of  $m_\Lambda$  considered lead to average temperatures  $T_\Lambda$  close to 298.15 K, suggesting an appropriate kinetic-energy exchange. The mass-parameter choices of  $m_\Lambda = 16 \text{ u nm}^2$  and  $1600 \text{ u nm}^2$  present the largest deviations of about 10 K and 6 K, respectively, while the deviation is at most 2 K for the other choices.

Expectedly, increasing  $m_\Lambda$  tendentially leads to a decrease in the diffusion coefficient  $D_\Lambda$ . This results from a smaller width  $\sigma_\Lambda$  of the  $\dot{\Lambda}$  distribution (Eq. 3.25; see also Fig. 3.5a where  $m_\Lambda$  increases from left to right), *i.e.* a lower average magnitude of the velocity along  $\Lambda$ . However,  $D_\Lambda$  is also affected by the presence of residual variations in the free-energy profile  $G_\Lambda(\Lambda)$ , which may be more easily overcome at higher  $m_\Lambda$  (more inertia). This is reflected in the increase of the  $\dot{\Lambda}$  autocorrelation time  $\tau_{\dot{\Lambda}}$  upon increasing  $m_\Lambda$  (see also Fig. 3.4 for the corresponding normalized autocorrelation functions). This second effect is likely to have more influence for small  $K$ , where the  $G_\Lambda(\Lambda)$  variations are more pronounced. In the present case with  $K = 16$ , these opposite effects probably explain why the trend in  $D_\Lambda$  upon increasing  $m_\Lambda$  is not strictly monotonic, with a comparatively high  $D_\Lambda$  for  $m_\Lambda = 1600 \text{ u nm}^2$  (see also Fig. 3.5c where the diffusion trend is non-monotonic upon increasing  $K$  when  $m_\Lambda$  is made proportional to  $K^{1/2}$ ).

In terms of the calculated free-energy change  $\Delta G$ , all simula-

tions provide the same values within the statistical error, with a slightly larger deviation for the simulation involving the lowest mass  $m_\Lambda = 16 \text{ u nm}^2$ , probably due to the comparatively large  $T_\Lambda$  deviation. Because it leads to accurate  $T_\Lambda$  and  $\Delta G$  values while presenting a high  $D_\Lambda$ , the second-to-lowest mass  $m_\Lambda = 160 \text{ u nm}^2$  was chosen for all following CBTI simulations employing  $K = 16$  replicas.

A second set of simulations involves a separate coupling of the  $\Lambda$  variable to a Nosé-Hoover chain thermostat with the choices  $\tau_\Lambda = 0.05, 0.1, 0.5, 1$  or  $2 \text{ ps}$  along with  $K = 16$  replicas and  $m_\Lambda = 160 \text{ u nm}^2$  ( $10 \text{ ns}$  simulations after  $0.2 \text{ ns}$  equilibration). The results are shown graphically in Figs. 3.F.6 - 2.F.10 and reported numerically in Tab. 3.1 (entries 6-10). Except for the two shortest coupling times, the average temperature  $T_\Lambda$  is very close to the reference temperature, with a maximal deviation of about  $2 \text{ K}$ . The larger deviations for  $\tau_\Lambda = 0.05$  or  $0.1 \text{ ps}$  suggest that the coupling of  $\Lambda$  to its thermostat should be made less tight than that of the conformational degrees of freedom to their thermostat (here SD with  $0.1 \text{ ps}$ ). The coupling parameter  $\tau_\Lambda$  has no influence on the width  $\sigma_{\dot{\Lambda}}$  of the  $\dot{\Lambda}$  distribution, which is the same for all five simulations and very close to the Maxwell-Boltzmann one (Eq. 3.25). On the other hand, a lower  $\tau_\Lambda$ , *i.e.* a tighter coupling, reduces the inertia of the conveyor belt, which results in a decrease of the autocorrelation time  $\tau_{\dot{\Lambda}}$ . This leads to a tendential decrease of  $D_\Lambda$  upon decreasing  $\tau_\Lambda$ . In terms of the calculated  $\Delta G$ , all simulations provide consistent values within the statistical error, with a slightly larger deviation for the simulation involving  $\tau_\Lambda = 0.05 \text{ ps}$ , probably again due to the comparatively large  $T_\Lambda$  deviation. In comparison to the first set of simulations with  $\tau_\Lambda \rightarrow \infty$ , the  $\Delta G$  values are slightly higher ( $0.1 - 0.3 \text{ kJ mol}^{-1}$ ), which could be due to the fact that the sampled  $\dot{\Lambda}$  distribution  $P_{\dot{\Lambda}}$  is even closer to

the Maxwell-Boltzmann distribution (Eq. 3.25), also for shorter time intervals. For  $\tau_\Lambda \geq 0.5$  ps, the choice of  $\tau_\Lambda$  has little influence on the calculation, and a coupling time  $\tau_\Lambda = 0.5$  ps was chosen for all following CBTI simulations. Note, however, that a looser coupling could be of advantage by leading to a higher  $D_\Lambda$ .

A third set of simulations considers the choices  $K = 8, 16, 32, 64$  or  $128$  along with  $m_\Lambda = 40K^{1/2}$  u nm<sup>2</sup> and  $\tau_\Lambda = 0.5$  ps ( $256K^{-1}$  ns simulations after 0.2 ns equilibration). The results are shown graphically in Figs. 2.F.11 - 2.F.15 and reported in Tab. 3.1 (entries 11-15). Simulations 11, 13 and 15 are discussed in details in Sect. 3.4 (see Figs. 3.3 and 3.5). For the five simulations, the average temperature  $T_\Lambda$  is close to the target temperature, with a maximal deviation of about 6 K. The free-energy differences  $\Delta G$  are consistent within the statistical error, except for the simulation employing  $K = 8$  replicas (entry 11), which is due to the non-uniform sampling of the  $\lambda$ -range as discussed in Sect. 3.4.

The rationale for making the mass  $m_\Lambda$  proportional to the square-root of  $K$ , an arbitrary parameter choice, is the following. If one wishes the CB variable  $\Lambda$  to evolve dynamically on comparable timescales irrespective of the number  $K$  of replicas attached to it, one should ensure that its acceleration  $\ddot{\Lambda}$  depends only weakly on  $K$ . Assuming that the forces exerted by the  $K$  replicas are essentially uncorrelated, their sum (*i.e.* the net force on the CB) will scale as  $K^{1/2}$ . Thus, an identical scaling of  $m_\Lambda$  is required to preserve an approximately constant  $\ddot{\Lambda}$ . The time series and distributions of  $\ddot{\Lambda}$  for different choices of  $K$  when using this scaling for  $m_\Lambda$  (see Fig. 3.5b) are indeed similar, supporting the above considerations.

In summary, this appendix shows that when selected within reasonable ranges, the parameters  $m_\Lambda$  and  $\tau_\Lambda$  have only a limited influence on the kinetic-energy exchange between  $\Lambda$  and the con-

formational degrees of freedom, on the temperature  $T_\Lambda$ , on the diffusion constant  $D_\Lambda$ , and on the calculated free energies  $\Delta G$ . In Sect. 3.4, working choice  $m_\Lambda = 40K^{1/2} \text{ u nm}^2$  and  $\tau_\Lambda = 0.5 \text{ ps}$  is selected for all CBTI calculations.

### Appendix 3.C SIMPLIFIED FREE-ENERGY ESTIMATORS

The free-energy estimator employed here for CBTI is given by Eq. 3.22. The approximation involved corresponds to a simple forward rectangular quadrature, where the Hamiltonian derivative is averaged over  $J$  successive bins considering all replicas simultaneously. If  $K$  replicas sample  $L$  configurations each, and  $p(\lambda)$  is close to homogeneous, the number of data points per bin will be close to  $KL/J$ , with limited variations across bins. In this case, one may consider a simpler alternative to Eq. 3.22 that does not involve the specification of a number of bins.

Considering all the  $K$  replicas simultaneously, the  $KL$  pairs of  $\lambda$ -values and associated Hamiltonian derivatives are sorted in ascending order for  $\lambda$  (index  $i = 0, \dots, KL-1$ ). One then calculates

$$\Delta G_{\text{alt}} = (KL)^{-1} \sum_{i=0}^{KL-1} \frac{\lambda_{i+1} - \lambda_{i-1}}{2} \left( \frac{\partial \mathcal{H}}{\partial \lambda} \right)_i \quad (3.C.1)$$

using  $\lambda_{-1} = \lambda_{KL} = 0$ . This alternative estimate, noted  $\Delta G_{\text{alt}}$ , considers that each sample  $i$  defines its own single-point bin, of width determined by contact to the next lower and higher data points  $i-1$  and  $i+1$ , respectively, thereby accounting for a possible

heterogeneity in the  $\lambda$ -sampling.

In the limit where the sampling becomes sufficiently close to homogeneous, Eqs. 3.22 or 3.C.1 can be replaced by an even simpler approximate expression, namely

$$\Delta G_{\text{app}} \approx K^{-1} \sum_{k=0}^{K-1} \left\langle \frac{\partial \mathcal{H}(\mathbf{x}_k; \lambda_k)}{\partial \lambda_k} \right\rangle^{\dagger*}, \quad (3.C.2)$$

which corresponds to averaging the Hamiltonian derivative over all replicas and over the entire CBTI simulation. This approximate estimate, noted  $\Delta G_{\text{app}}$ , assumes that the  $KL$  data points are on average equispaced along  $\lambda$ , which only holds in the limit of homogeneous sampling.

The accuracy of the estimate  $\Delta G$  of Eq. 3.22 with  $J = 500$  (except for entry 11,  $J = 200$ ) is compared in Tab. 3.1 to those of the simpler expressions  $\Delta G_{\text{alt}}$  and  $\Delta G_{\text{app}}$ , which do not require the specification of a number of bins. The estimate  $\Delta G_{\text{alt}}$  is very close to  $\Delta G$  (within error bar), suggesting that Eq. 3.C.1, which assumes that each sampled  $\lambda$ -value is at the center of its own bin, is a viable parameter-free alternative to Eq. 3.22. On the other hand, the estimate  $\Delta G_{\text{app}}$ , which assumes that each bin encompasses the same number of sampled  $\lambda$ -points irrespective of  $J$ , is more approximate. As expected from the involved assumption,  $\Delta G_{\text{app}}$  is only a good approximation to  $\Delta G$  when the CBTI sampling is indeed close to uniform along  $\lambda$ . Considering entries 11-17, this is essentially the case for the unbiased simulations with  $K \geq 32$ . However, for the unbiased simulation with  $K = 8$  and the two biased simulations, the sampling is not sufficiently homogeneous and the discrepancy can be significant (about  $1 - 5 \text{ kJ mol}^{-1}$ ).

### Appendix 3.D AUTOCORRELATION OF THE HAMILTONIAN DERIVATIVE

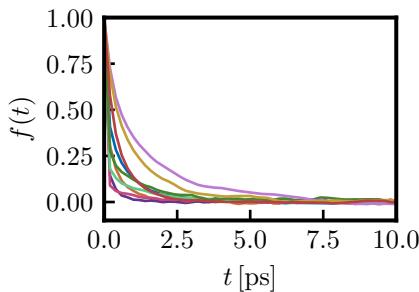


FIGURE 3.D.1: Normalized autocorrelation function  $f(t)$  of the Hamiltonian derivative for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. These functions are based on the TI calculation relying on  $K_{\text{TI}} = 17$   $\lambda$ -points. The functions displayed correspond to every second  $\lambda$ -point. The associated autocorrelation times  $\tau_f$  are reported in Tab. 3.D.1.

TABLE 3.D.1: Autocorrelation times  $\tau_f$  of the Hamiltonian derivative for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. These times are based on the TI calculation relying on  $K_{\text{TI}} = 17$   $\lambda$ -points. The associated autocorrelation functions  $f(t)$  are displayed in Fig. 3.D.1 for every second  $\lambda$ -point. The autocorrelation times were derived by performing an exponential fit to the  $f(t)$  curve.

$\lambda$	$\tau_f$ [ps]	$\lambda$	$\tau_f$ [ps]
0.000	0.339	0.562	0.161
0.062	0.297	0.625	0.224
0.125	0.207	0.688	0.377
0.188	0.150	0.750	1.080
0.250	0.109	0.812	1.119
0.312	0.085	0.875	0.799
0.375	0.087	0.938	0.555
0.438	0.112	1.000	0.442
0.500	0.139		

### Appendix 3.E REFERENCE TI CALCULATIONS

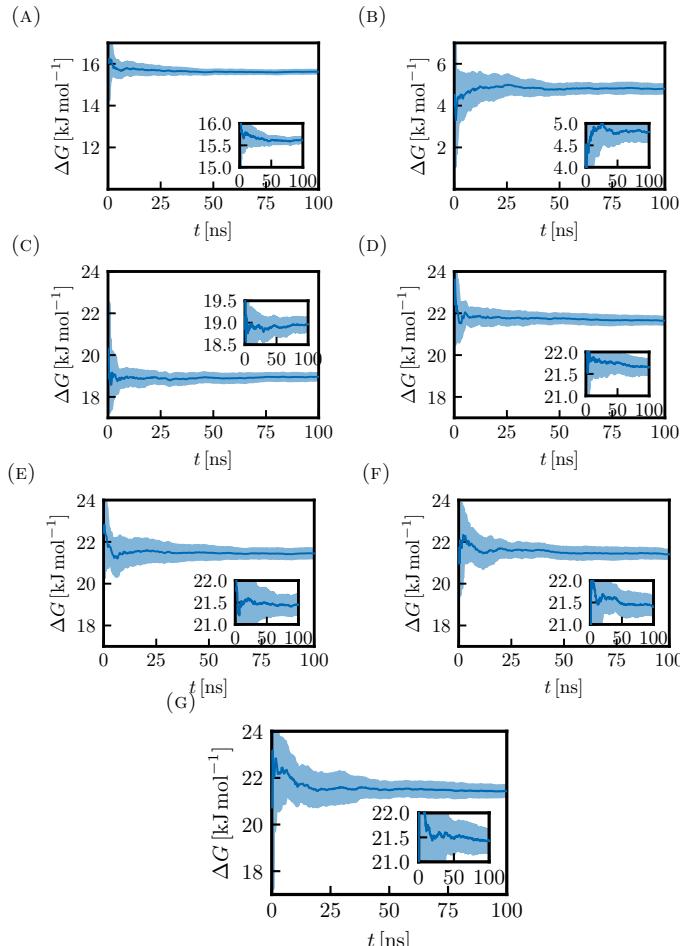


FIGURE 3.E.1: Convergence properties of the TI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. This figure complements Fig. 3.2b by also including the TI calculations involving fewer than 129  $\lambda$ -points. Here, the running  $\Delta G$  estimates are shown for calculations involving  $K_{\text{TI}} = 2^n + 1$  equidistant  $\lambda$ -points with  $n = 1, 2, \dots, 7$  resulting in  $K_{\text{TI}} = 3$  a, 5 b, 9 c, 17 d, 33 e, 65 f and 129 g. Note that the vertical ranges in Panels a and b differ from the range used in the other graphs. The corresponding numerical values at full sampling time are reported in Tabs. 3.2 and 3.E.1.

TABLE 3.E.1: *Repeat results over TI calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* The free-energy change  $\Delta G$  is reported for TI calculations involving  $K_{\text{TI}} = 2^n + 1$   $\lambda$ -points with  $n = 1, 2, \dots, 7$ . For each choice of  $K_{\text{TI}}$  the results of ten repeats (successive rows) are listed, all involving a total sampling time of 100 ns equally distributed over the  $\lambda$ -points. For each repeat, the calculated  $\Delta G$  (Simpson quadrature) is reported along with a bootstrap error estimate (first parenthesis) and an error propagated from the bootstrap errors of the Hamiltonian derivative at each  $\lambda$ -point (second parenthesis). These two error estimates do not include a Student  $t$ -factor. Repeat statistics based on this data can be found in Tab. 3.2.

$K_{\text{TI}}$	$\Delta G [\text{kJ mol}^{-1}]$						
	1 3	2 5	3 9	4 17	5 33	6 65	7 129
1	15.68 (0.08) (0.10)	4.82 (0.18) (0.15)	19.22 (0.18) (0.16)	21.87 (0.17) (0.15)	21.48 (0.20) (0.15)	21.18 (0.24) (0.15)	21.45 (0.30) (0.14)
2	15.52 (0.11) (0.10)	4.32 (0.18) (0.14)	18.72 (0.17) (0.16)	21.83 (0.18) (0.14)	21.59 (0.19) (0.14)	21.72 (0.22) (0.14)	21.87 (0.29) (0.14)
3	15.64 (0.09) (0.09)	4.50 (0.22) (0.16)	18.92 (0.17) (0.14)	21.57 (0.20) (0.14)	21.72 (0.17) (0.14)	21.62 (0.20) (0.14)	21.60 (0.26) (0.14)
4	15.58 (0.10) (0.09)	4.98 (0.17) (0.14)	19.18 (0.17) (0.14)	21.88 (0.18) (0.14)	21.85 (0.19) (0.14)	21.66 (0.22) (0.14)	21.67 (0.31) (0.14)
5	15.71 (0.10) (0.09)	4.97 (0.18) (0.15)	18.77 (0.17) (0.15)	21.37 (0.18) (0.15)	21.37 (0.22) (0.15)	21.29 (0.25) (0.14)	21.30 (0.26) (0.14)
6	15.62 (0.10) (0.09)	5.11 (0.18) (0.16)	18.99 (0.17) (0.15)	21.54 (0.21) (0.15)	21.19 (0.17) (0.15)	21.66 (0.22) (0.14)	21.76 (0.27) (0.14)
7	15.66 (0.10) (0.08)	4.75 (0.19) (0.15)	18.92 (0.14) (0.15)	21.67 (0.14) (0.15)	21.71 (0.21) (0.14)	21.29 (0.22) (0.14)	21.39 (0.29) (0.14)
8	15.48 (0.09) (0.09)	4.79 (0.18) (0.15)	18.77 (0.16) (0.15)	21.37 (0.17) (0.15)	21.03 (0.19) (0.15)	21.34 (0.25) (0.14)	20.99 (0.33) (0.14)
9	15.58 (0.09) (0.08)	4.77 (0.19) (0.16)	18.98 (0.19) (0.15)	21.49 (0.18) (0.15)	21.12 (0.18) (0.14)	21.02 (0.23) (0.14)	21.16 (0.32) (0.14)
10	15.65 (0.10) (0.09)	5.06 (0.19) (0.14)	19.17 (0.17) (0.15)	21.95 (0.19) (0.15)	21.45 (0.18) (0.15)	21.24 (0.22) (0.14)	21.09 (0.26) (0.14)

## Appendix 3.F UNBIASED CBTI SIMULATIONS

### 3.F.1 EXPLORATION OF THE INFLUENCE OF $m_\Lambda$ IN CBTI

**Simulation #1,**  $m_\Lambda = 16 \text{ u nm}^2$

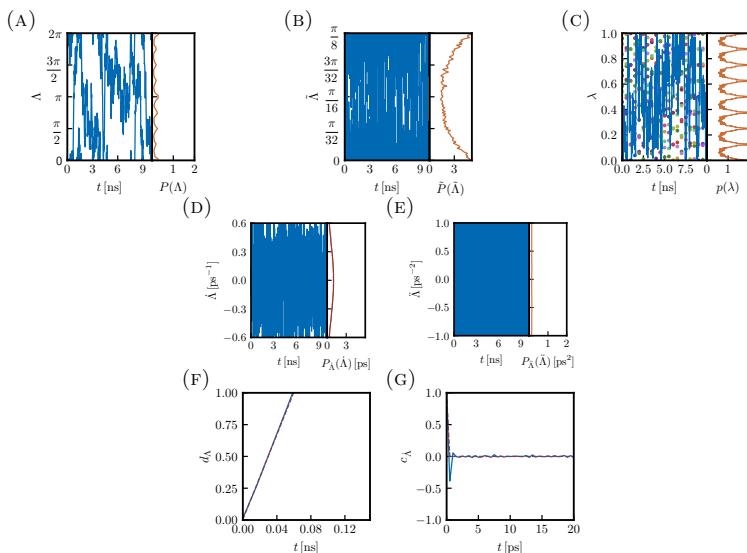


FIGURE 3.F.1: Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 16 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ). This simulation corresponds to entry 1 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\dot{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

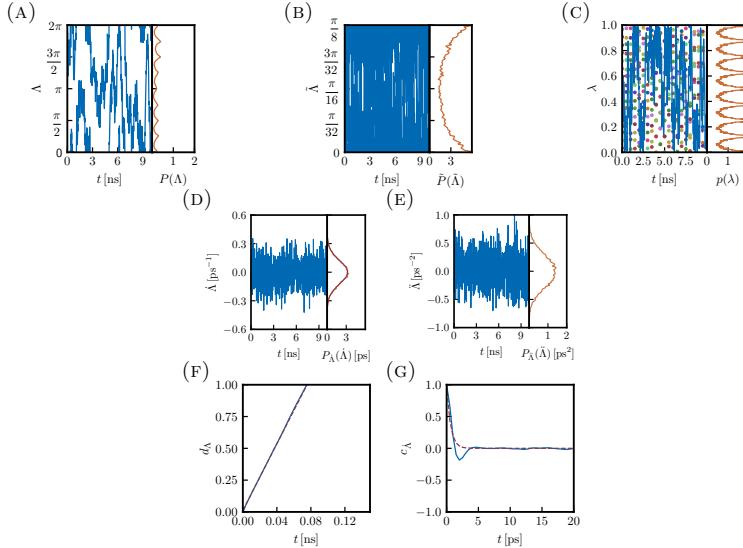
**Simulation #2,  $m_\Lambda = 160 \text{ u nm}^2$** 

FIGURE 3.F.2: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 2 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

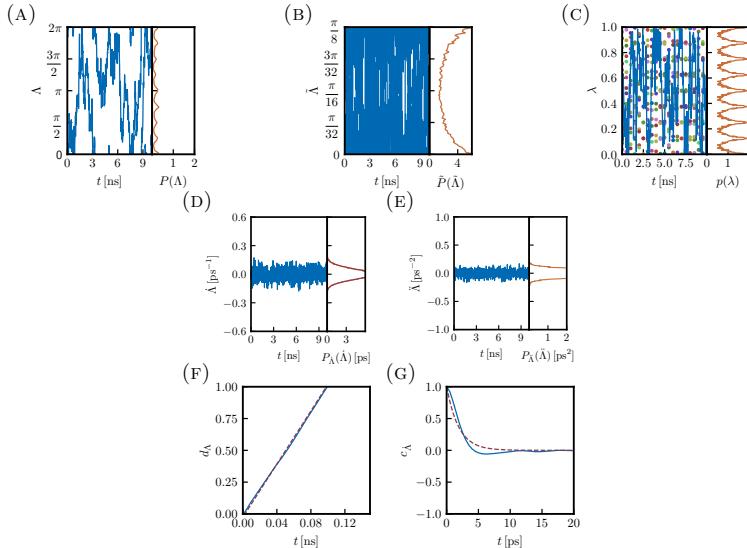
**Simulation #3,  $m_\Lambda = 800 \text{ u nm}^2$** 

FIGURE 3.F.3: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 800 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 3 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\bar{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\bar{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\bar{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\bar{\Lambda}$  (blue) along with an exponential fit (dashed brown).

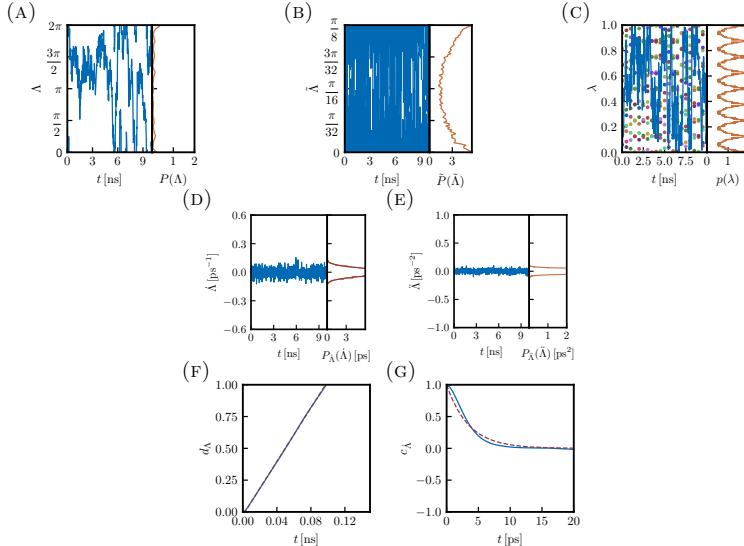
**Simulation #4,  $m_\Lambda = 1600 \text{ u nm}^2$** 

FIGURE 3.F.4: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 1600 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 4 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

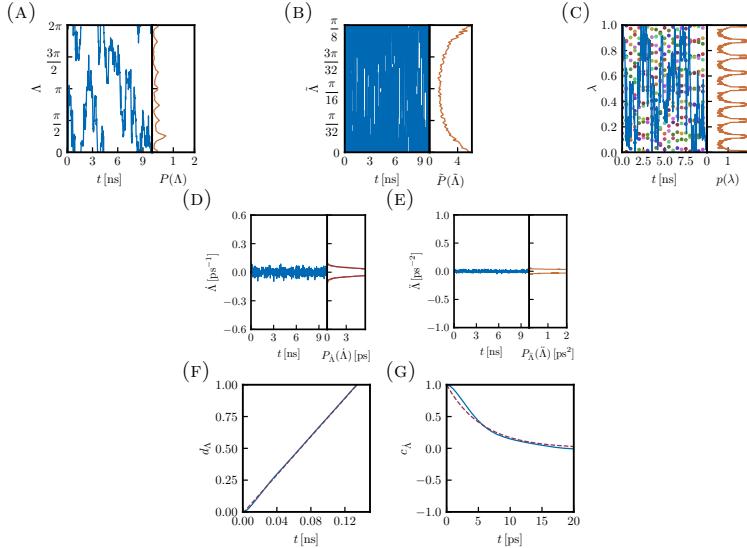
**Simulation #5,  $m_\Lambda = 3200 \text{ u nm}^2$** 

FIGURE 3.F.5: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 3200 \text{ u nm}^2$  and no thermostat coupling of the  $\Lambda$ -variable ( $\tau_\Lambda \rightarrow \infty$ ).* This simulation corresponds to entry 5 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

### 3.F.2 EXPLORATION OF THE INFLUENCE OF $\tau_\Lambda$ IN CBTI

**Simulation #6,  $\tau_\Lambda = 0.05$  ps**

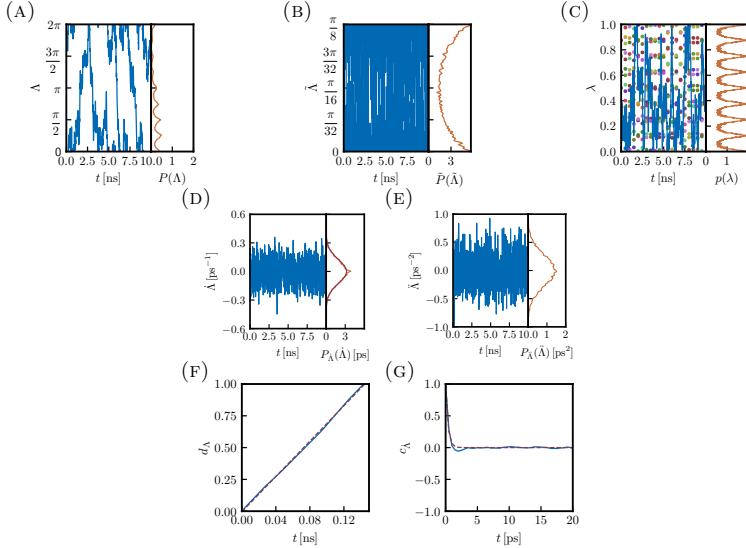


FIGURE 3.F.6: Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.05$  ps. This simulation corresponds to entry 6 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\bar{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\bar{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\bar{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\bar{\Lambda}$  (blue) along with an exponential fit (dashed brown).

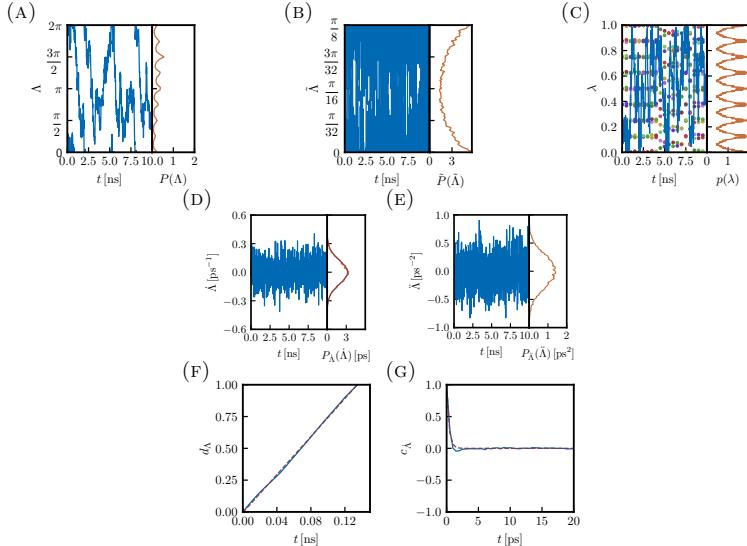
**Simulation #7,  $\tau_\Lambda = 0.1$  ps**

FIGURE 3.F.7: Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.1$  ps. This simulation corresponds to entry 7 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

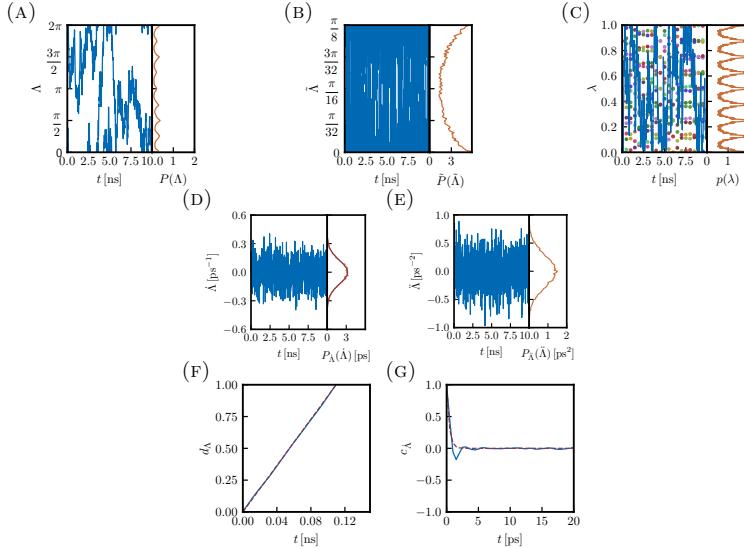
**Simulation #8,  $\tau_\Lambda = 0.5$  ps**


FIGURE 3.F.8: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5$  ps.* This simulation corresponds to entry 8 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

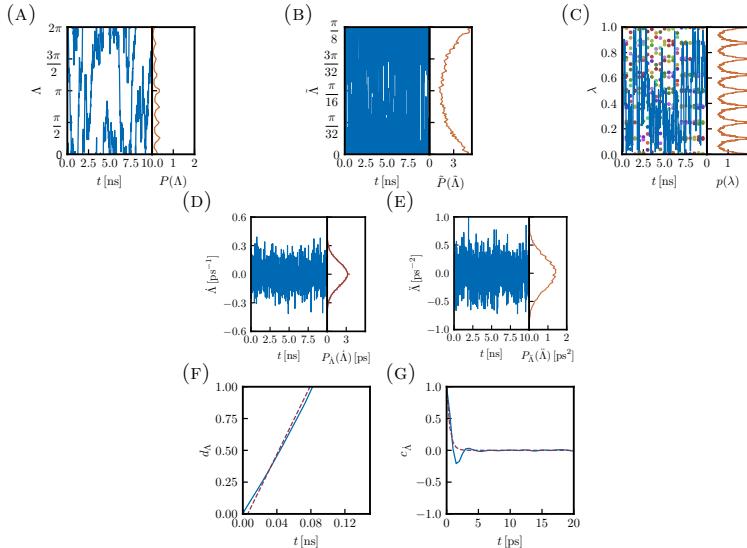
**Simulation #9,  $\tau_\Lambda = 1$  ps**

FIGURE 3.F.9: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 1$  ps.* This simulation corresponds to entry 9 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

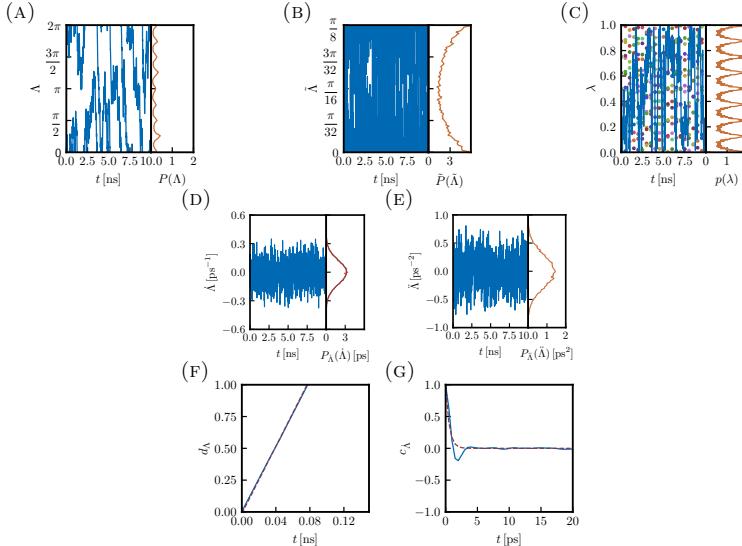
**Simulation #10,  $\tau_\Lambda = 2$  ps**


FIGURE 3.F.10: *Results from the CBTI simulation of 10 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 2$  ps.* This simulation corresponds to entry 10 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

### 3.F.3 EXPLORATION OF THE INFLUENCE OF $K$ IN CBTI

**Simulation #11,  $K = 8$**

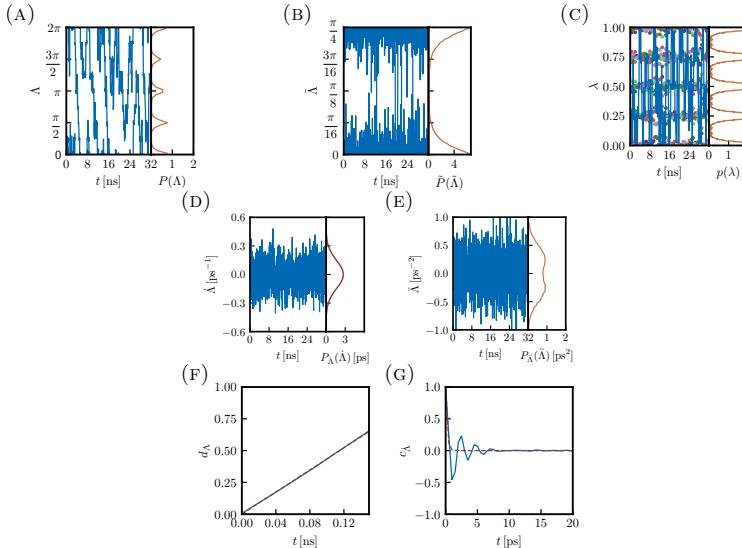


FIGURE 3.F.11: *Results from the CBTI simulation of 32 ns employing  $K = 8$  replicas with a mass-parameter  $m_\Lambda = 113 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 11 in Tab. 3.1.* Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

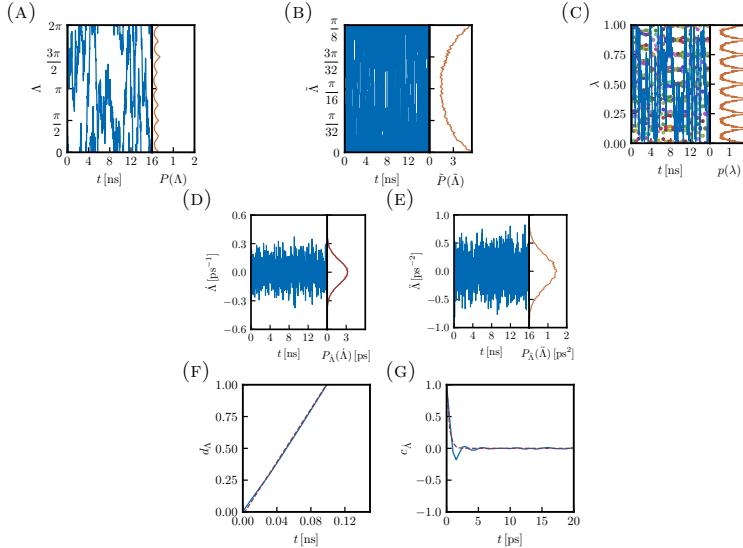
Simulation #12,  $K = 16$ 

FIGURE 3.F.12: *Results from the CBTI simulation of 16 ns employing  $K = 16$  replicas with a mass-parameter  $m_\Lambda = 160 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 12 in Tab. 3.1.* Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

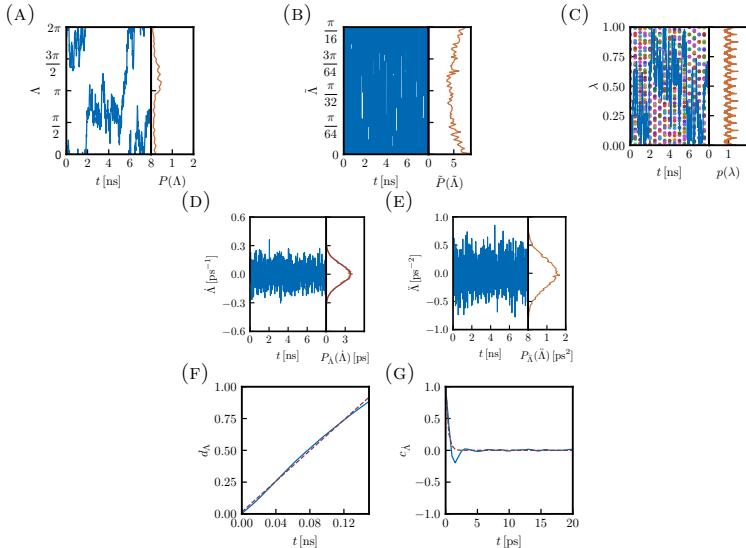
**Simulation #13,  $K = 32$** 

FIGURE 3.F.13: *Results from the CBTI simulation of 8 ns employing  $K = 32$  replicas with a mass-parameter  $m_{\Lambda} = 226 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_{\Lambda} = 0.5 \text{ ps}$ . This simulation corresponds to entry 13 in Tab. 3.1. Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).*

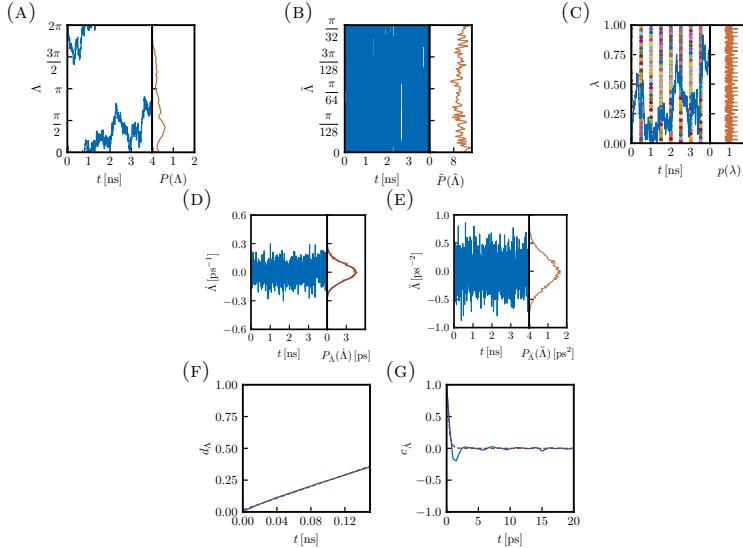
Simulation #14,  $K = 64$ 

FIGURE 3.F.14: *Results from the CBTI simulation of 4 ns employing  $K = 64$  replicas with a mass-parameter  $m_\Lambda = 320 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 14 in Tab. 3.1.* Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

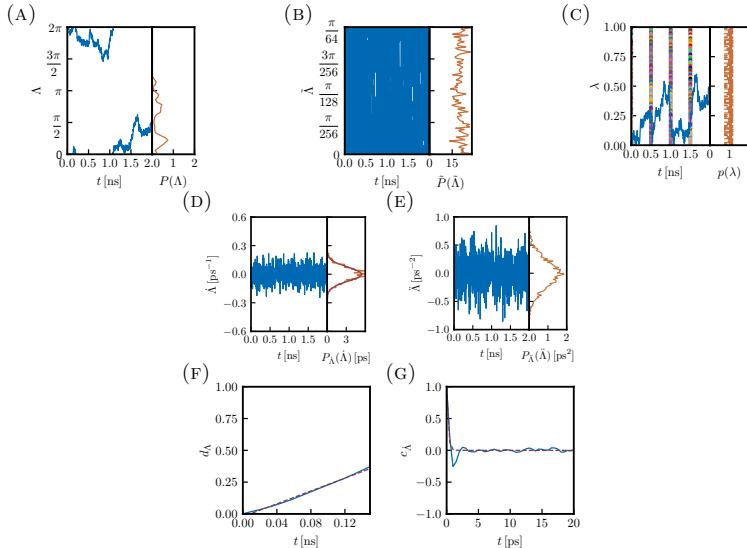
**Simulation #15,  $K = 128$** 

FIGURE 3.F.15: *Results from the CBTI simulation of 2 ns employing  $K = 128$  replicas with a mass-parameter  $m_\Lambda = 452 \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_\Lambda = 0.5 \text{ ps}$ . This simulation corresponds to entry 15 in Tab. 3.1.* Panel a time series (blue) and distribution (orange) of  $\Lambda$ . Panel b time series (blue) and distribution (orange) of  $\tilde{\Lambda}$ . Panel c time series of  $\lambda$  for replica  $k = 0$  (blue) along with the  $\lambda$ -values for the  $K - 1$  other replicas (colored dots) as well as distribution of  $\lambda$  considering all replicas (orange). Panel d time series (blue) and distribution (blue) of the velocity  $\dot{\Lambda}$  along with the analytical Maxwell-Boltzmann distribution (orange). Panel e time series (blue) and distribution (blue) of the acceleration  $\ddot{\Lambda}$ . Panel f time series of the mean-square displacement of  $\Lambda$  (blue) along with a linear least-square fit (dashed brown). Panel g autocorrelation function of  $\dot{\Lambda}$  (blue) along with an exponential fit (dashed brown).

## Appendix 3.G REPEATS OF A CBTI SIMULATION

TABLE 3.G.1: *Repeat results of a CBTI calculation for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar using the 2061H66 force field.* The free-energy change  $\Delta G$  is reported for the CBTI calculation involving  $K = 16$  replicas along with  $m_A = 160 \text{ u nm}^2$  and  $\tau_A = 0.5 \text{ ps}$ . The results of ten repeats (successive rows) are listed, all of 6.25 ns duration (total single-system sampling time of 100 ns). For each repeat, the calculated  $\Delta G$  (Eq. 3.22 with  $J = 500$ ) is reported along with a bootstrap error estimate. This error does not include a Student  $t$ -factor. Repeat statistics based on this data can be found in Tab. 3.2.

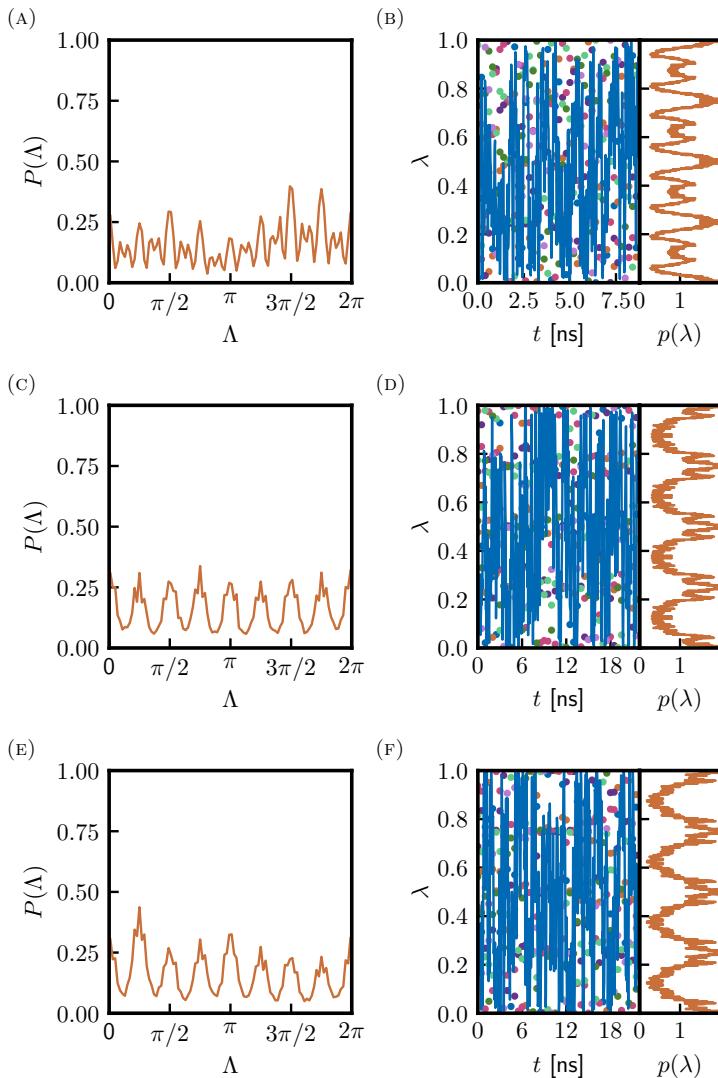
	$\Delta G \text{ [kJ mol}^{-1}\text{]}$
1	21.42 (0.16)
2	21.33 (0.16)
3	21.44 (0.17)
4	21.38 (0.15)
5	21.08 (0.14)
6	21.43 (0.15)
7	21.59 (0.15)
8	21.54 (0.15)
9	21.31 (0.15)
10	21.42 (0.16)

---

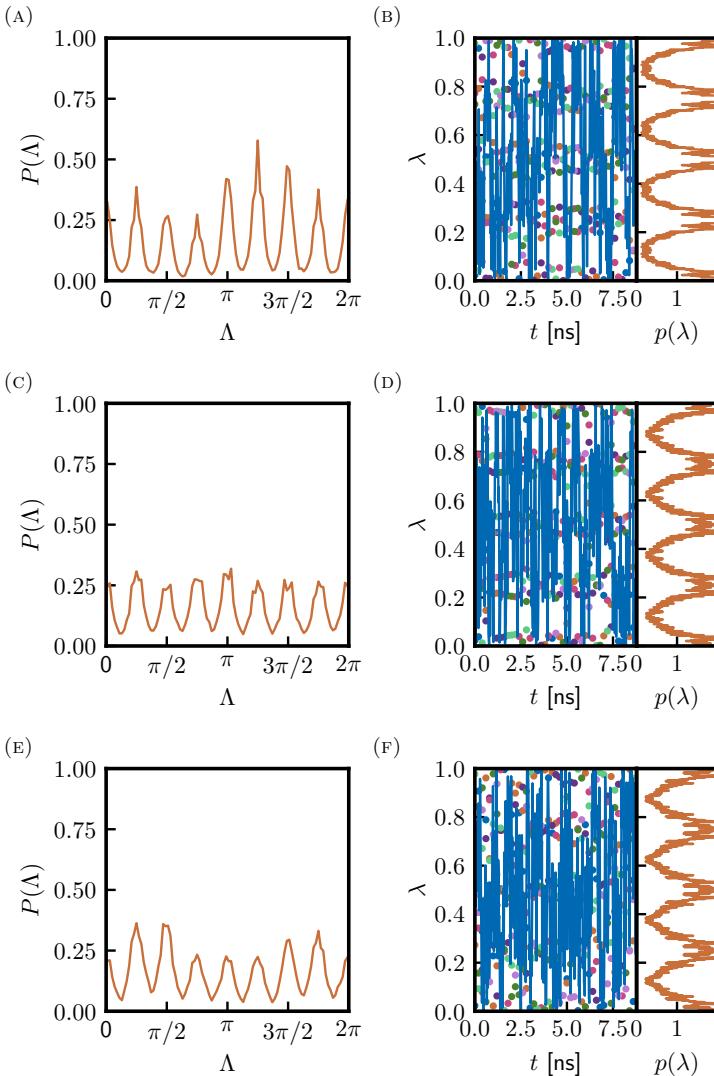
## Appendix 3.H CBTI SIMULATIONS WITH BIASING POTENTIAL

TABLE 3.H.1: *Parameters and results of biased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* The successive entries are the index of the simulation (sim), the number  $K$  of replicas, the number  $N_{\text{gp}}$  of gridpoints for  $\tilde{\Lambda}$  in the range  $[0; 2\pi/K]$ , the build-up force constant  $c_{\text{LE}}$ , the reduction factor  $f_{\text{red}}$ , the LE build-up time  $t_{\text{LE}}$  for the replica system, the number  $N_{\text{ds}}$  of double-sweeps over the  $\tilde{\Lambda}$  range during the build-up, the US umbrella sampling time  $t_{\text{US}}$  for the replica system, and the free-energy difference  $\Delta G$  calculated using Eq. 3.22 with  $J = 500$ . The corresponding results are illustrated graphically in Figs. 2.H.2 and 2.H.4. Only entries 2 and 8 are discussed in Sect. 3.4.

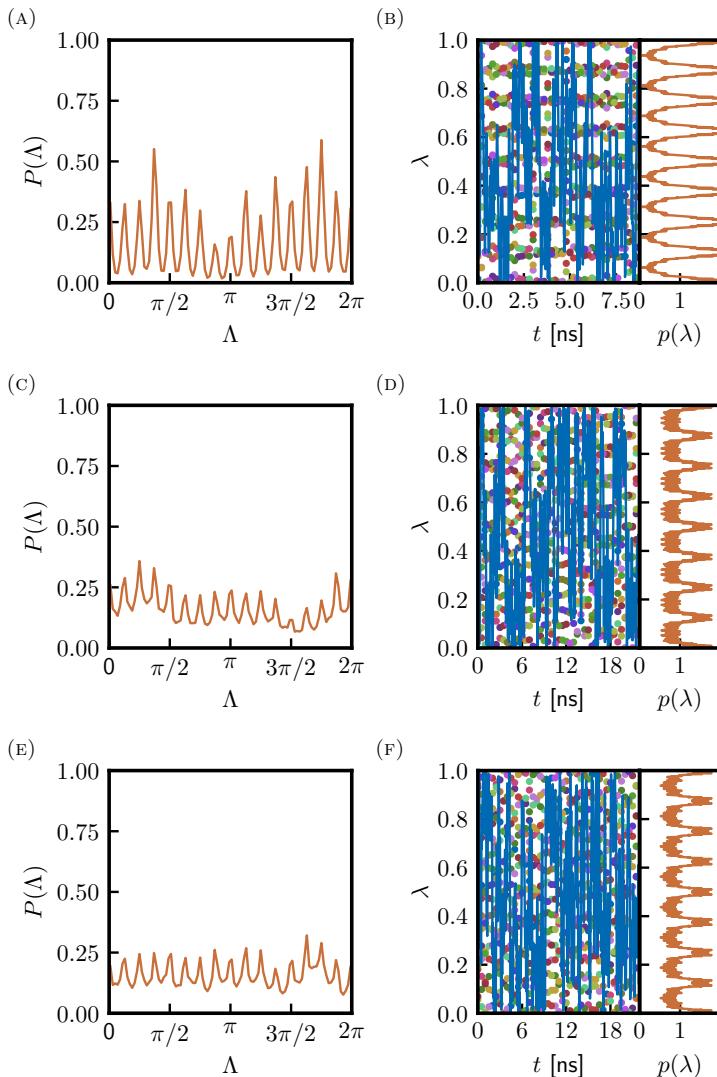
sim	$K$	$N_{\text{gp}}$	$c_{\text{LE}}$ [kJ mol $^{-1}$ ]	$f_{\text{red}}$	$t_{\text{LE}}$ [ns]	$N_{\text{ds}}$	$t_{\text{US}}$ [ns]	$\Delta G$ [kJ mol $^{-1}$ ]
1	8	10	0.01	0.1	0.050	3	22	$21.11 \pm 0.15$
2	8	34	0.001	0.1	0.150	3	22	$21.48 \pm 0.18$
3	8	34	0.001	0.1	0.200	6	22	$21.29 \pm 0.17$
4	8	34	0.001	0.8	0.054	3	22	$21.54 \pm 0.22$
5	8	34	0.001	0.8	0.128	6	22	$21.38 \pm 0.17$
6	8	34	0.001	0.8	0.156	9	22	$21.46 \pm 0.16$
7	16	6	0.01	0.1	0.050	3	22	$21.23 \pm 0.22$
8	16	18	0.001	0.1	0.070	3	22	$21.30 \pm 0.13$
9	16	18	0.001	0.1	0.083	6	22	$21.48 \pm 0.16$
10	16	18	0.001	0.8	0.030	3	10	$21.00 \pm 0.14$
11	16	18	0.001	0.8	0.050	6	20	$21.57 \pm 0.14$
12	16	18	0.001	0.8	0.084	9	20	$21.21 \pm 0.17$



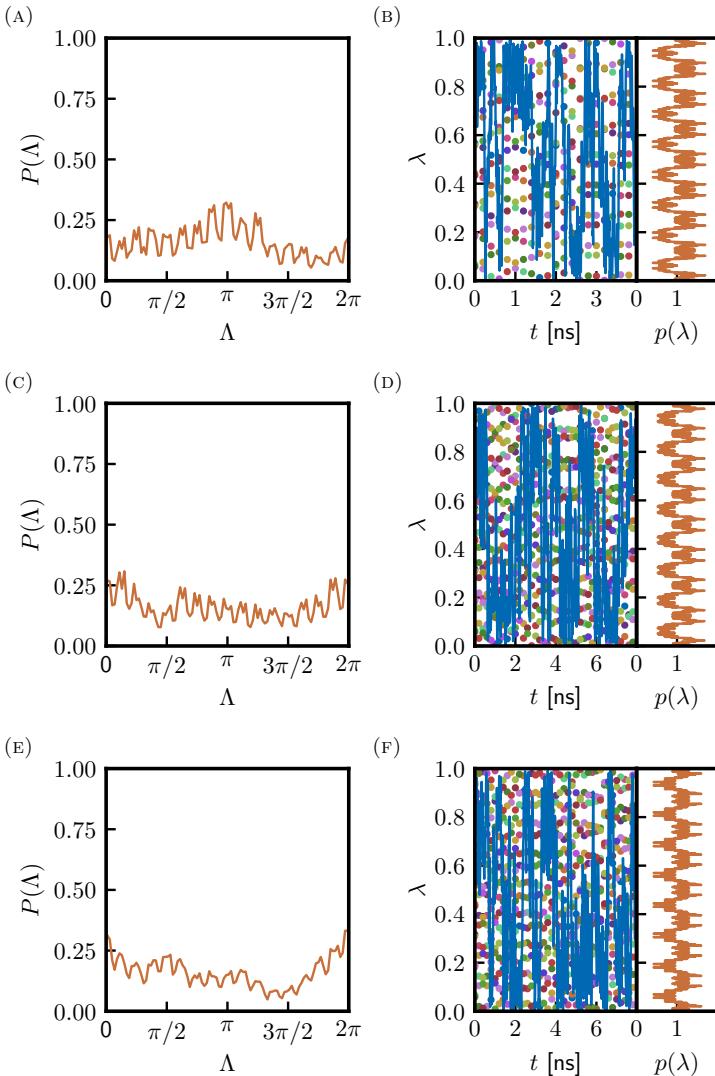
*Continued next page.*



**FIGURE 3.H.2:** Time series and probability distributions of the relevant CB-variables in biased CBTI calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The calculations rely on  $K = 8$  replicas. Each row shows the probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$  (left) and the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas (right). The successive panels correspond to different protocol settings listed in Tab. 2.H.1. Panels (a,b) entry 1. Panels (c,d) entry 2. Panels (e,f) entry 3. Panels (g,h) entry 4. Panels (i,j) entry 5. and Panels (k,l) entry 6.



*Continued next page.*



**FIGURE 3.H.4:** Time series and probability distributions of the relevant CB-variables in biased CBTI calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The calculations rely on  $K = 16$  replicas. Each row shows the probability distribution  $P(\Lambda)$  of the CB advance variable  $\Lambda$  (left) and the time series  $\lambda(t)$  and probability distribution  $p(\lambda)$  of the coupling variable  $\lambda$  for all replicas (right). The successive panels correspond to different protocol settings listed in Tab. 2.H.1. Panels (a,b) entry 7. Panels (c,d) entry 8. Panels (e,f) entry 9. Panels (g,h) entry 10. Panels (i,j) entry 11. and Panels (k,l) entry 12.

Appendix 3.I FREE-ENERGY  
PROFILES  $G_{\tilde{\Lambda}}(\tilde{\Lambda})$   
ALONG  $\tilde{\Lambda}$

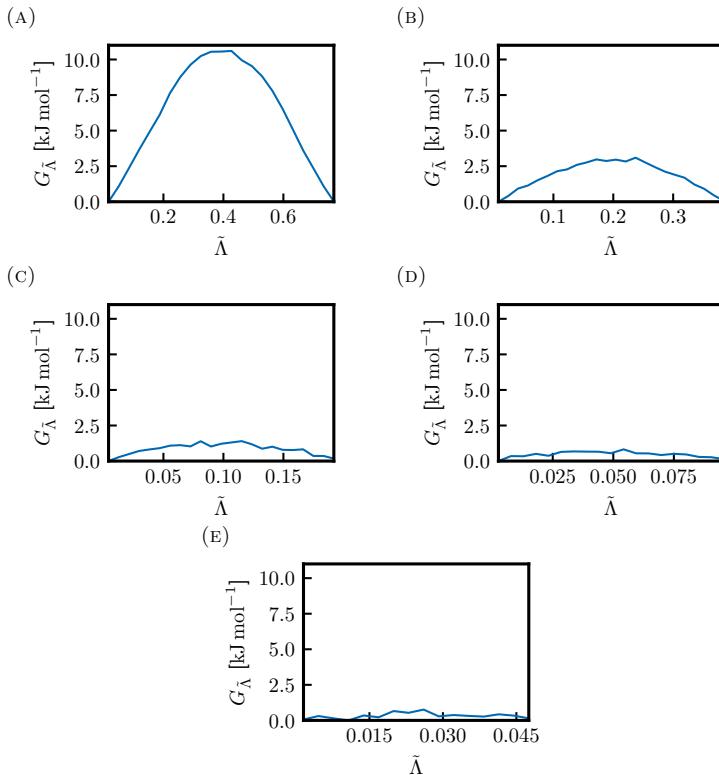


FIGURE 3.I.1: Free-energy profiles  $G_{\tilde{\Lambda}}(\tilde{\Lambda})$  in unbiased CBTI simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The simulations relied on  $K = 8$  a,  $K = 16$  b,  $K = 32$  c,  $K = 64$  d, and  $K = 128$  e replicas, a mass-parameter  $m_{\Lambda} = 40K^{1/2} \text{ u nm}^2$  and thermostat coupling of the  $\Lambda$ -variable with  $\tau_{\Lambda} = 0.5$  ps. The free-energy profiles were calculated as  $G_{\tilde{\Lambda}}(\Lambda) = -\beta^{-1} \ln P_{\tilde{\Lambda}}(\tilde{\Lambda})$ , where  $P_{\tilde{\Lambda}}(\tilde{\Lambda})$  is the normalized probability distribution of  $\tilde{\Lambda}$ , and anchored to zero at their minimum. The value at the maximum corresponds to  $G_{\tilde{\Lambda}}^*$ , shown graphically in Fig. 3.6.

### Appendix 3.J HRE SIMULATIONS

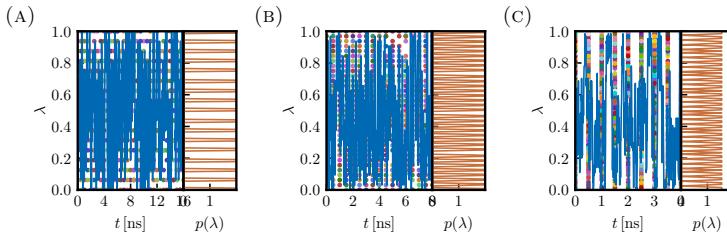


FIGURE 3.J.1: Time series  $\lambda(t)$  and probability distributions  $p(\lambda)$  of the coupling variable  $\lambda$  in HRE simulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field. The time series is shown as a blue curve for replica  $k = 0$ , and as individual colored points at 0.5 ns interval for the  $K_{\text{HRE}} - 1$  other replicas. Panel a:  $K_{\text{HRE}} = 17$ . Panel b:  $K_{\text{HRE}} = 33$ . Panel c:  $K_{\text{HRE}} = 65$ .

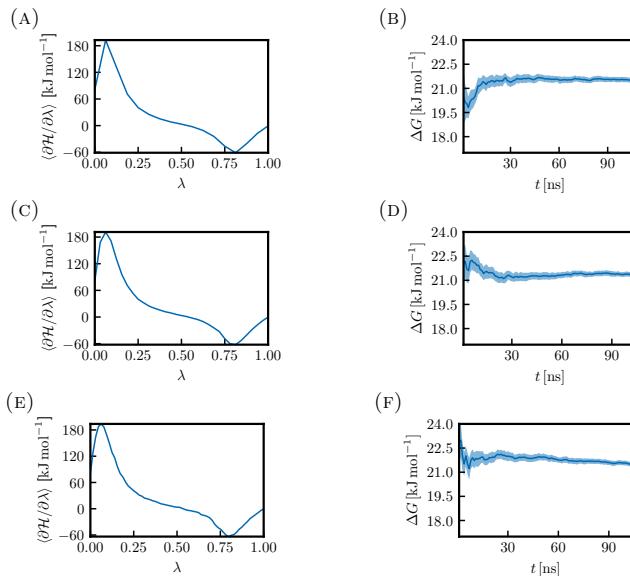


FIGURE 3.J.2: *Relevant results from HRE calculations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* Panels a, c and e show the Hamiltonian derivative curve considering a total single-system sampling time of 100 ns. Panel a:  $K_{\text{HRE}} = 17$ . Panel c:  $K_{\text{HRE}} = 33$ . Panel e:  $K_{\text{HRE}} = 65$ . Panels b, d and f show the Hamiltonian derivative curve considering a total single-system sampling time of 100 ns. Panel b:  $K_{\text{HRE}} = 17$ . Panel d:  $K_{\text{HRE}} = 33$ . Panel f:  $K_{\text{HRE}} = 65$ .

## Appendix 3.K TI/EXTI CALCULATIONS

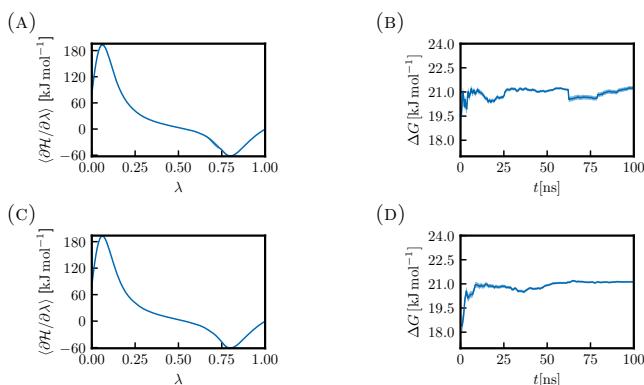


FIGURE 3.K.1: *Relevant results from TI/EXTI calulations of the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* Panels a and c show the Hamiltonian derivative curve, which was predicted at 129  $\lambda$ -points. Panel a:  $K_{\text{TI}} = 9$ . Panel c:  $K_{\text{TI}} = 17$ . Panels b and d show the convergence of the  $\Delta G$  value dependent on the total single-system sampling time. Panel b:  $K_{\text{TI}} = 9$ . Panel d:  $K_{\text{TI}} = 17$ .

## Appendix 3.L TI/MBAR CALCULATIONS

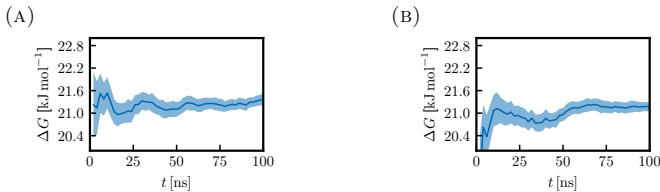


FIGURE 3.L.1: *Convergence properties of TI/MBAR calculations for the aqueous methanol-to-dummy mutation at 298.15 K and 1 bar with the 2016H66 force field.* Panels b and d show the convergence of the  $\Delta G$  value dependent on the total single-system sampling time. Panel a:  $K_{\text{TI}} = 9$ . Panel b:  $K_{\text{TI}} = 17$ .



# Outlook

# 4

*“ tata ”*

---

tutu

## 4.1 IMPROVEMENTS FOR RE-0EDS



## References

- [1] J. Mansfeld, *Die Vorsokratiker. Griechisch/Deutsch.*; Stuttgart: Reclam., **1983**.
- [2] H. Diels, W. Kranz, *Die Fragmente der Vorsokratiker. Edition 6.*; Weidmann Berlin, **1951**.
- [3] C.C.W. Taylor, *The Atomists: Leucippus and Democritus. Fragments, A Text and Translation with Commentary.*; Toronto: University of Toronto Press, **1999**.
- [4] J.L. Stocks, *The works of Aristotle translated into English - De Caelo Volume 2. Edition Ross, W.D..*; Clarendon Press, Oxford, **1930**.
- [5] A. Chalmers, *The Scientist's Atom and the Philosopher's Stone. How Science Succeeded and Philosophy Failed to Gain Knowledge of Atoms.*; Springer Netherlands, **2009**.
- [6] I. Newton, *The Principia: Mathematical principles of natural philosophy. (A new translation by I. B. Cohen and A. Whitman).*; University of California Press, Berkeley, USA, **1999**.
- [7] F. Cajori, *Sir Isaac Newton's mathematical principles of natural philosophy and his system of the world.*; Univ. California Press, Berkeley, USA, **1962**.
- [8] A. Chalmers, Atomism from the 17th to the 20th Century In: *The Stanford Encyclopedia of Philosophy*; E.N. Zalta, Ed.; Metaphysics Research Lab, Stanford Univ., USA, **2014**; <https://plato.stanford.edu/archives/win2014/entries/atomism-modern/>.
- [9] M.F. Schlecht, *Molecular modeling on the PC.*; Wiley-VCH, New York, **1998**.
- [10] R.B. Kelly, *Can. J. Chem.* **1957**, *35*, 149: A relationship between the conformations of cyclohexane derivatives and their physical properties.
- [11] A. Kekulé, *Justus Liebigs Ann. Chem.* **1866**, *137*, 129: Untersuchungen über aromatische Verbindungen.
- [12] A. Kekulé, *Justus Liebigs Ann. Chem.* **1872**, *162*, 77: Über einige Condensationsprodukte des Aldehyds.
- [13] A. Rocke, *Ann. Sci.* **1985**, *42*, 355: Hypothesis and experiment in the early development of Kekulé's benzene theory.
- [14] N. Hopwood, S. de Chadarevian, *Models: The third dimension of science.*; Stanford University Press, Stanford, **2004**.

- [15] Bunn Crowfoot D., Rogers-Low C.W., B.W., A. Turner-Jones, X-ray crystallographic investigation of the structure of penicillin. In: *Chemistry of penicillin.*; H.T. Clarke, J.R. Johnson, R. Robinson, Eds.; Princeton University Press, Princeton, New Jersey, **1949**; pp 310-367.
- [16] L. Pauling, R.B. Corey, H.R. Branson, *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 205: The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain.
- [17] F.H. Crick, J.D. Watson, *Proc. R. Soc. London Ser. A* **1954**, *223*, 80: The complementary structure of desoxyribonucleic acid.
- [18] J. Loschmidt, *Konstitutionsformeln der organischen Chemie in graphischer Darstellung*; Engelmann, Leipzig, **1861**.
- [19] J. C. Kendrew, G. Bodo, H. M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips, *Nature* **1958**, *181*, 662: A three-dimensional model of the myoglobin molecule obtained by X-ray analysis.
- [20] M.F. Perutz, M.G. Rossmann, A.F. Cullis, H. Muirhead, G. Will, A.C.T. North, *Nature* **1960**, *185*, 416: Structure of haemoglobin: A three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by x-ray analysis.
- [21] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *J. Chem. Phys.* **1953**, *21*, 1087: Equation of state calculations by fast computing machines.
- [22] B.J. Alder, T.E. Wainwright, *J. Chem. Phys.* **1957**, *27*, 1208: Phase transition for a hard sphere system.
- [23] A. Rahman, *Phys. Rev.* **1964**, *136*, A405: Correlations in the motion of atoms in liquid argon.
- [24] R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman lectures on physics. Volume 1.*; Addison-Wesley, Boston, USA, **1963**.
- [25] F.H. Stillinger, A. Rahman, *J. Chem. Phys.* **1974**, *60*, 1545: Improved simulation of liquid water by molecular dynamics.
- [26] J.A. McCammon, B.R. Gelin, M. Karplus, *Nature* **1977**, *267*, 585: Dynamics of folded proteins.
- [27] H.L. Friedman, *A course in statistical mechanics*; Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA, **1985**.
- [28] R. Kubo, *Statistical Mechanics. An Advanced Course with Problems and Solutions. Edition 8.*; Elsevier Science Publisher B.V., Amsterdam, The Netherlands, **1990**.
- [29] D.A. McQuarrie, *Statistical mechanics.*; University Science Books, Sausalito, CA, USA, **2000**.
- [30] L.K. Nash, *Elements of statistical Thermodynamics Edition 2.*; Dover Publications Inc., Mineola, New York, USA, **2006**.
- [31] K. Huang, *Introduction to statistical physics. Edition 2nd ed..*; Chapman & Hall/CRC Taylor & Francis Group, Boca Raton, USA, **2010**.

- [32] G. Gallavotti, *Eur. Phys. J. H* **2016**, *41*, 181: Ergodicity: a historical perspective. Equilibrium and nonequilibrium.
- [33] X. Daura, A.E. Mark, W.F. van Gunsteren, *J. Comput. Chem.* **1998**, *19*, 535: Parametrization of aliphatic  $\text{CH}_n$  united atoms of GROMOS96 force field.
- [34] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, *J. Phys. Chem. B* **2007**, *111*, 7812: The MARTINI force field: Coarse grained model for biomolecular simulation.
- [35] G.A. Voth, *Coarse-graining of condensed phase and biomolecular systems.*; Taylor & Francis, Inc., New York, USA, **2008**.
- [36] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656: A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6.
- [37] H.J.C. Berendsen, W.F. van Gunsteren, H.R.J. Zwinderman, R.G. Geurtsen, *Ann. New York Acad. Sci.* **1986**, *482*, 269: Simulations of proteins in water.
- [38] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089: Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems.
- [39] R.W. Hockney, *Methods Comput. Phys.* **1970**, *9*, 135: The potential calculation and some applications.
- [40] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. di Nola, J.R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684: Molecular dynamics with coupling to an external bath.
- [41] S. Nosé, *Mol. Phys.* **1984**, *52*, 255: A molecular dynamics method for simulations in the canonical ensemble.
- [42] S. Nosé, *J. Chem. Phys.* **1984**, *81*, 511: A unified formulation of the constant temperature molecular dynamics methods.
- [43] W.G. Hoover, *Phys. Rev. A* **1985**, *31*, 1695: Canonical dynamics: Equilibrium phase-space distributions.
- [44] G.J. Martyna, M.L. Klein, M. Tuckerman, *J. Chem. Phys.* **1992**, *97*, 2635: Nosé-Hoover chains: The canonical ensemble via continuous dynamics.
- [45] H.C. Andersen, *J. Chem. Phys.* **1980**, *72*, 2384: Molecular dynamics simulations at constant pressure and/or temperature.
- [46] W.K. Hastings, *Biometrika* **1970**, *57*, 97: Monte Carlo sampling methods using Markov chains and their applications.
- [47] P.M. King, Free energy via molecular simulation : A primer. In: *Computer simulation of biomolecular systems, theoretical and experimental applications. Volume 2.*; W.F. van Gunsteren, P.K. Weiner, A.J. Wilkinson, Eds.; ESCOM Science Publishers, B.V., Leiden, The Netherlands, **1993**; pp 267-314.
- [48] M.E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation.*; Oxford University Press, New York, USA, **2010**.
- [49] R.W. Zwanzig, *J. Chem. Phys.* **1954**, *22*, 1420: High-temperature equation of state by a perturbation method. I. Nonpolar gases.

- [50] J.G. Kirkwood, *Phys. Rev.* **1933**, *44*, 31: Quantum statistics of almost classical assemblies.
- [51] J.G. Kirkwood, *Phys. Rev.* **1934**, *45*, 116: Quantum statistics of almost classical assemblies.
- [52] J.G. Kirkwood, *J. Chem. Phys.* **1935**, *3*, 300: Statistical mechanics of fluid mixtures.
- [53] J.R. Moran, S. Karbach, D.J. Cram, *J. Am. Chem. Soc.* **1982**, *104*, 5826: Cavitands: synthetic molecular vessel.
- [54] J.R. Moran, J.L. Ericson, E. Dalcanale, J.A. Bryant, C.B. Knobler, D.J. Cram, *J. Am. Chem. Soc.* **1991**, *113*, 5707: Vases and kites as cavitands.
- [55] D.J. Cram, H.-J. Choi, J.A. Bryant, C.B. Knobler, *J. Am. Chem. Soc.* **1992**, *114*, 7748: Solvophobic and entropic driving forces for forming vel-crapplexes, which are four-fold, lock-key dimers in organic media.
- [56] F. Hof, S.L. Craig, C. Nuckolls, J. Rebek Jr., *Angew. Chem. Int. Ed.* **2002**, *41*, 1488: Molecular encapsulation.
- [57] S.K. Körner, F.C. Tucci, D.M. Rudkevich, T. Heinz, J. Rebek, *Chem. Eur. J.* **2000**, *6*, 187: A self-assembled cylindrical capsule: new supramolecular phenomena through encapsulation..
- [58] R.J. Hooley, H.J. Van Anda, J. Rebek Jr., *J. Am. Chem. Soc.* **2006**, *128*, 3894: Cavitands with revolving doors regulate binding selectivities and rates in water.
- [59] R.J. Hooley, J. Rebek, *Org. Lett.* **2007**, *9*, 1179: Self-complex deep cavitands: Alkyl chains coil into a nearby cavity.
- [60] R.J. Hooley, S.R. Shenoy, J. Rebek, *Org. Lett.* **2008**, *10*, 5397: Electronic and steric effects in binding of deep cavitands.
- [61] D. Fankhauser, D. Kolarski, W.R. Grüning, F. Diederich, *Eur. J. Org.* **2014**, *Chem.*, 3575: Resorcin[4]arene-based molecular baskets and water-soluble container molecules: Synthesis and <sup>1</sup>H NMR Host-guest complexation studies.
- [62] V.A. Azov, A. Schlegel, F. Diederich, *Bull. Chem. Soc. Jpn.* **2006**, *79*, 1926: Functionalized calix[4]resorcinarene cavitands. Versatile platforms for the modular construction of extended molecular switches.
- [63] E. Dalcanale, P. Soncini, G. Bacchilega, F. Ugozzoli, *J. Chem. Soc. Chem. Commun.* **1989**, *8*, 500: Selective complexation of neutral molecules in organic solvents. Host-Guest complexes and cavitates between cavitands and aromatic compounds.
- [64] P. Soncini, S. Bonsignore, E. Dalcanale, F. Ugozzoli, *J. Org. Chem.* **1992**, *57*, 4608: Cavitands as versatile molecular receptors.
- [65] J. Chen, J. Rebek, *Org. Lett.* **2002**, *4*, 327: Selectivity in an encapsulated cycloaddition reaction.
- [66] F.H. Zelder, R. Salvio, J. Rebek, *Chem. Commun. (Cambridge U.)* **2006**, *K.*, 1280: A synthetic receptor for phosphocholine esters.
- [67] O.B. Berryman, A.C. Sather, A. Lledó, J. Rebek, *Angew. Chem. Int. Ed.* **2011**, *50*, 9400: Switchable catalysis with a light-responsive cavitand.

- [68] I. Pochorovski, M.O. Ebert, J.P. Gisselbrecht, C. Boudon, W.B. Schweizer, F. Diederich, *J. Am. Chem. Soc.* **2012**, *134*, 14702: Redox-switchable resorcin[4]arene cavitands: Molecular grippers.
- [69] V. Balzani, M. Gómez-López, J.F. Stoddart, *Acc. Chem. Res.* **1998**, *31*, 405: Molecular machines.
- [70] V. Balzani, A. Credi, F. Raymo, J. Stoddart, *Angew. Chem. Int. Ed.* **2000**, *39*, 3348: Artificial molecular machines.
- [71] E.R. Kay, D.A. Leigh, *Angew. Chem. Int. Ed.* **2015**, *54*, 10080: Rise of the molecular machines.
- [72] K.G. Untch, R.J. Kurland, *J. Am. Chem. Soc.* **1963**, *85*, 346: The conformational equilibration of cis, cis, cis-1,4,7-cyclononatriene; a determination of activation energy and entropy by N.M.R. II.
- [73] K.G. Untch, *J. Am. Chem. Soc.* **1963**, *85*, 345: *sym*-cis,cis,cis-1,4,7-Cyclononatriene, an unusual cyclic six  $\pi$  electron system.
- [74] P. Radlick, S. Winstein, *J. Am. Chem. Soc.* **1963**, *85*, 344: cis-cis-cis-d 1,4,7-Cyclononatriene, a homoconjugated six  $\pi$ -electron system.
- [75] M.R. Detty, L.A. Paquette, *Tetrahedron Lett.* **1977**, *18*, 347: The question of hexahomobenzene automerization. New synthetic approaches to cis<sup>3</sup>-1,4,7-cyclononatriene.
- [76] F.A.L. Anet, M. Ghiasi, *J. Am. Chem. Soc.* **1980**, *102*, 2528: Ring inversion in 1,4,7-cyclononatriene (trishomobenzene). Force-field calculations for abnormally large bond angles.
- [77] M.A. Zamora, F.D. Suvire, R.D. Enriz, *J. Comput. Chem.* **2008**, *29*, 280: Ring inversion in 1,4,7 cyclononatriene and analogues: Ab initio and DFT calculations and topological analysis.
- [78] G.M. Robinson, *J. Chem. Soc. Trans.* **1915**, *107*, 267: XXX. - A reaction of homopiperonyl and of homoveratryl alcohols.
- [79] A.S. Lindsey, *J. Chem. Soc.*, **1965**, 1685: 316. The structure of cyclotrimeratrylene (10,15-dihydro-2,3,7,8,12,13-hexamethoxy-5H-tribenzo[a,d,g]cyclononene) and related compounds.
- [80] B. Miller, B.D. Gesner, *Tetrahedron Lett.* **1965**, *6*, 3351: Conformation of cyclotrimeratrylene by nuclear magnetic resonance measurements.
- [81] H. Zimmermann, P. Tolstoy, H. Limbach, R. Poupko, Z. Luz, *J. Phys. Chem. B* **2004**, *108*, 18772: The saddle form of cyclotrimeratrylene.
- [82] M.J. Hardie, *Chem. Soc. Rev.* **2010**, *39*, 516: Recent advances in the chemistry of cyclotrimeratrylene.
- [83] G.P. Moss, *Pure Appl. Chem.* **2009**, *68*, 2193: Basic terminology of stereochemistry (IUPAC Recommendations 1996).
- [84] D. Cremer, J.A. Pople, *J. Am. Chem. Soc.* **1975**, *97*, 1354: A general definition of ring puckering coordinates.
- [85] V.A. Azov, B. Jaun, F. Diederich, *Helv. Chim. Acta* **2004**, *87*, 449: NMR investigations into the vase-kite conformational switching of resorcin[4]arene cavitands.

- [86] V.A. Azov, F. Diederich, Switching Processes in Cavitands, Containers and Capsules. In: *Molecular Switches, Second Edition.*; B.L. Feringa, W.R. Browne, Eds.; Wiley-VCH Verlag GmbH, Co. KGaA, Weinheim, Germany., **2011**; pp 257-300.
- [87] I. Pochorovski, F. Diederich, *Acc. Chem. Res.* **2014**, *47*, 2096: Development of redox-switchable resorcin[4]arene cavitands.
- [88] C. Groppe, B.L. Quigley, F. Diederich, *J. Am. Chem. Soc.* **2018**, *140*, 2705: Molecular recognition with resorcin[4]arene cavitands: Switching, halogen-bonded capsules, and enantioselective complexation.
- [89] T. Haino, D.M. Rudkevich, J. Rebek, *J. Am. Chem. Soc.* **1999**, *121*, 11253: Kinetically stable caviplexes in water.
- [90] S. Javor, J. Rebek, *J. Am. Chem. Soc.* **2011**, *133*, 17473: Activation of a water-soluble resorcinarene cavitand at the water-phosphocholine micelle interface.
- [91] A. Galán, E.C. Escudero-Adán, A. Frontera, P. Ballester, *J. Org. Chem.* **2014**, *79*, 5545: Synthesis, structure, and binding properties of lipophilic cavitands based on a calix[4]pyrrole-resorcinarene hybrid scaffold.
- [92] D.J. Cram, S. Karbach, H.Y. Kim, L. Baczyński, G.W. Kallemeyn, *J. Am. Chem. Soc.* **1985**, *107*, 2575: Shell closure of two cavitands forms carcerand complexes with components of the medium as permanent guest.
- [93] D.J. Cram, S. Karbach, H.E. Kim, C.B. Knobler, E.F. Maverick, J.L. Ericson, R.C. Helgeson, *J. Am. Chem. Soc.* **1988**, *110*, 2229: Host-guest complexation. 46. Cavitands as open molecular vessels form solvates.
- [94] T. Amaya, J. Rebek, *J. Am. Chem. Soc.* **2004**, *126*, 14149: Hydrogen-bonded encapsulation complexes in protic solvents.
- [95] T. Gottschalk, P.D. Jarowski, F. Diederich, *Tetrahedron* **2008**, *64*, 8307: Reversibly controllable guest binding in precisely defined cavities: selectivity, induced fit, and switching in novel resorcin[4]arene-based container molecules.
- [96] J. Hornung, D. Fankhauser, L.D. Shirtcliff, A. Praetorius, W.B. Schweizer, F. Diederich, *Chem. Eur. J.* **2011**, *17*, 12362: Cycloalkane and alicyclic heterocycle complexation by new switchable resorcin[4]arene-based container molecules: NMR and ITC binding studies.
- [97] D. Ajami, J. Rebek, *Acc. Chem. Res.* **2013**, *46*, 990: More chemistry in small spaces.
- [98] D.M. Rudkevich, G. Hilmersson, J. Rebek, *J. Am. Chem. Soc.* **1997**, *119*, 9911: Intramolecular hydrogen bonding controls the exchange rates of guests in a cavitand.
- [99] P. Amrhein, P.L. Wash, A. Shivanuk, J. Rebek, *Org. Lett.* **2002**, *4*, 319: Metal ligation regulates conformational equilibria and binding properties of cavitands.
- [100] V.A. Azov, A. Beeby, M. Cacciarini, A.G. Cheetham, F. Diederich, M. Frei, J.K. Gimzewski, V. Gramlich, B. Hecht, B. Jaun, T. Latychevskaia, A. Lieb, Y. Lill, F. Marotti, A. Schlegel, R.R. Schlittler, P.J. Skinner, P. Seiler, Y. Yamakoshi, *Adv. Funct. Mater.* **2006**, *16*, 147: Resorcin[4]arene cavitand-based molecular switches.

- [101] V.A. Azov, A. Schlegel, F. Diederich, *Angew. Chem. Int. Ed.* **2005**, *44*, 4635: Geometrically precisely defined multinanometer expansion/contraction motions in a resorcin[4]arene cavitand based molecular switch.
- [102] V.A. Azov, R. Gómez, J. Stelten, *Tetrahedron* **2008**, *64*, 1909: Synthesis of electrochemically responsive TTF-based molecular tweezers: Evidence of tight intramolecular TTF pairing in solution.
- [103] I. Pochorovski, B. Breiten, W.B. Schweizer, F. Diederich, *Chem. Eur. J.* **2010**, *16*, 12590: FRET studies on a series of BODIPY-dye-labeled switchable resorcin[4]arene cavitands.
- [104] O.B. Berryman, A.C. Sather, J. Rebek, *Chem. Commun. (Cambridge U.K.)* **2011**, *47*, 656: A light controlled cavitand wall regulates guest binding.
- [105] O.B. Berryman, A.C. Sather, J. Rebek, *Org. Lett.* **2011**, *13*, 5232: A deep cavitand with a fluorescent wall functions as an ion sensor.
- [106] I. Pochorovski, J. Milić, D. Kolarski, C. Gropp, W.B. Schweizer, F. Diederich, *J. Am. Chem. Soc.* **2014**, *136*, 3852: Evaluation of hydrogen-bond acceptors for redox-switchable resorcin[4]arene cavitands.
- [107] J. Milić, M. Zalibera, I. Pochorovski, N. Trapp, J. Nomrowski, D. Neshchadin, L. Ruhlmann, C. Boudon, O.S. Wenger, A. Savitsky, W. Lubitz, G. Gescheidt, F. Diederich, *J. Phys. Chem. Lett.* **2016**, *7*, 2470: Paramagnetic molecular grippers: The elements of six-state redox switches.
- [108] J. Milić, M. Zalibera, D. Talaat, J. Nomrowski, N. Trapp, L. Ruhlmann, C. Boudon, O.S. Wenger, A. Savitsky, W. Lubitz, F. Diederich, *Chem. Eur. J.* **2017**, *24*, 1431: Photoredox-switchable resorcin[4]arene cavitands: radical control of molecular gripping machinery via hydrogen bonding.
- [109] V. García-López, J. Milić, M. Zalibera, D. Neshchadin, M. Kuss-Petermann, L. Ruhlmann, J. Nomrowski, N. Trapp, C. Boudon, G. Gescheidt, O.S. Wenger, F. Diederich, *Tetrahedron* **2018**, *74*, 5615: Light-actuated resorcin[4]arene cavitands.
- [110] J.V. Milić, T. Schneeberger, M. Zalibera, K.Z. Milowska, Q.K. Ong, N. Trapp, L. Ruhlmann, C. Boudon, C. Thilgen, F. Diederich, *Helv. Chim. Acta* **2019**, *102*, e1800225/1: Thioether-functionalized quinone-based resorcin[4]arene cavitands: Electroswitchable molecular actuators.
- [111] P. Roncucci, L. Pirondini, G. Paderni, C. Massera, E. Dalcanale, V.A. Azov, F. Diederich, *Chem. Eur. J.* **2006**, *12*, 4775: Conformational behavior of pyrazine-bridged and mixed-bridged cavitands: A general model for solvent effects on thermal "vase-kite" switching.
- [112] P.J. Skinner, A.G. Cheetham, A. Beeby, V. Gramlich, F. Diederich, *Helv. Chim. Acta* **2001**, *84*, 2146: Conformational switching of resorcin[4]arene cavitands by protonation.
- [113] M. Frei, F. Marotti, F. Diederich, *Chem. Commun.* **2004**, *2004*, 1362: ZnII-induced conformational control of amphiphilic cavitands in Langmuir monolayers.
- [114] I. Pochorovski, C. Boudon, J.P. Gisselbrecht, M.O. Ebert, W.B. Schweizer, F. Diederich, *Angew. Chem. Int. Ed.* **2012**, *51*, 262: Quinone-based, redox-active resorcin[4]arene cavitands.

- [115] V.A. Azov, P.J. Skinner, Y. Yamakoshi, P. Seiler, V. Gramlich, F. Diederich, *Helv. Chim. Acta* **2003**, *86*, 3648: Functionalized and partially or differentially bridged resorcin[4]arene cavitands: synthesis and solid-state structures.
- [116] I. Mohammed-Ziegler, F. Billes, *J. Inclusion Phenom. Macrocyclic Chem.* **2007**, *58*, 19: Optical spectroscopy and theoretical studies in calixarene chemistry.
- [117] B.H. Boo, H.S. Kim, S.G. Koh, M. Lee, K. No, *J. Mol. Struct.* **2013**, *1031*, 119: Time-resolved fluorescence and fluorescence anisotropy of calix[4]arene: Elucidation of the excitation energies of various conformers.
- [118] F. Liu, R.C. Helgeson, K.N. Houk, *Acc. Chem. Res.* **2014**, *47*, 2168: Building on Cram's legacy: Stimulated gating in hemicarcerands.
- [119] C. Gropp, T. Husch, N. Trapp, M. Reiher, F. Diederich, *J. Am. Chem. Soc.* **2017**, *139*, 12190: Dispersion and halogen-bonding interactions: Binding of the axial conformers of monohalo- and ( $\pm$ )-*trans*-1,2-dihalocyclohexanes in enantiopure alleno-acetylenic cage.
- [120] J.P. Otto, L. Wang, I. Pochorovski, S.M. Blau, A. Aspuru-Guzik, Z. Bao, G.S. Engel, M. Chiu, *Chem. Sci.* **2018**, *9*, 3694: Disentanglement of excited-state dynamics with implications for FRET measurements: two-dimensional electronic spectroscopy of a BODIPY-functionalized cavitand.
- [121] F.M. Ytreberg, *J. Phys. Chem. B* **2010**, *114*, 5431: Computational study of small molecule binding for both tethered and free condition.
- [122] C. Bonaccorso, G. Brancatelli, G. Forte, G. Arena, S. Geremia, D. Sciotto, C. Sgarlata, *RSC Adv.* **2014**, *4*, 53575: Factors driving the self-assembly of water-soluble calix[4]arene and gemini guests: a combined solution, computational and solid-state study.
- [123] I. Pochorovski, T. Knehans, D. Nettels, A.M. Müller, W.B. Schweizer, A. Cafisch, B. Schuler, F. Diederich, *J. Am. Chem. Soc.* **2014**, *136*, 2441: Experimental and computational study of BODIPY dye-labeled cavitand dynamics.
- [124] P.H. Hünenberger, J.K. Granwehr, J.-N. Aebscher, N. Ghoneim, E. Haselbach, W.F. van Gunsteren, *J. Am. Chem. Soc.* **1997**, *119*, 7533: Experimental and theoretical approaches to hydrogen-bonded diastereomeric interactions in a model complex.
- [125] G.M. Torrie, J.P. Valleau, *J. Comput. Phys.* **1977**, *23*, 187: Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.
- [126] J.P. Valleau, G.M. Torrie, A guide to Monte Carlo for statistical mechanics: 1. Highways. In: *Modern theoretical chemistry. Volume 5.*; B.J. Berne, Ed.; Plenum Press, New York, USA, **1977**; pp 137-168.
- [127] T.C. Beutler, W.F. van Gunsteren, *J. Chem. Phys.* **1994**, *100*, 1492: The computation of a potential of mean force: Choice of the biasing potential in the umbrella sampling technique.
- [128] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, *J. Comput. Chem.* **1995**, *16*, 1339: Multidimensional free-energy calculations using the weighted histogram analysis method.

- [129] T.N. Heinz, W.F. van Gunsteren, P.H. Hünenberger, *J. Chem. Phys.* **2001**, *115*, 1125: Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations.
- [130] E. Piccinini, M. Ceccarelli, F. Affinito, R. Brunetti, C. Jacoboni, *J. Chem. Theory Comput.* **2008**, *4*, 173: Biased molecular simulations for free-energy mapping: A comparison on the KcsA channel as a test case.
- [131] G.H. Paine, H.A. Scheraga, *Biopolymers* **1985**, *24*, 1391: Prediction of the native conformation of a polypeptide by a statistical-mechanical procedure. I. Backbone structure of enkephalin.
- [132] M. Mezei, *J. Comput. Phys.* **1987**, *68*, 237: Adaptive umbrella sampling. Self-consistent determination of the non-Boltzmann bias.
- [133] R.W.W. Hooft, B.P. van Eijck, J. Kroon, *J. Chem. Phys.* **1992**, *97*, 6690: An adaptive umbrella sampling procedure in conformational analysis using molecular dynamics and its application to glycol.
- [134] R.A. Friedman, M. Mezei, *J. Chem. Phys.* **1995**, *102*, 419: The potentials of mean force of sodium chloride and sodium dimethylphosphate in water: An application of adaptive umbrella sampling.
- [135] C. Bartels, M. Karplus, *J. Comput. Chem.* **1997**, *18*, 1450: Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations.
- [136] J. Wang, Y. Gu, H. Liu, *J. Chem. Phys.* **2006**, *125*, 094907/1: Determination of conformational free energies of peptides by multidimensional adaptive umbrella sampling.
- [137] V. Babin, C. Roland, T.A. Darden, C. Sagui, *J. Chem. Phys.* **2006**, *125*, 204909/1: The free energy landscape of small peptides as obtained from metadynamics with umbrella sampling corrections.
- [138] S. Marsili, A. Barducci, R. Chelli, P. Procacci, V. Schettino, *J. Phys. Chem. B* **2006**, *110*, 14011: Self-healing umbrella sampling: A non-equilibrium approach for quantitative free energy calculations.
- [139] T. Lelièvre, M. Roussel, G. Stoltz, *J. Chem. Phys.* **2007**, *126*, 134111/1: Computation of free energy profiles with parallel adaptive dynamics.
- [140] A. van der Vaart, M. Karplus, *J. Chem. Phys.* **2007**, *126*, 164106/1: Minimum free energy pathways and free energy profiles for conformational transitions based on atomistic molecular dynamics simulations.
- [141] V. Babin, C. Roland, C. Sagui, *J. Chem. Phys.* **2008**, *128*, 134101/1: Adaptively biased molecular dynamics for free energy calculations.
- [142] C.B. Barnett, K.J. Naidoo, *Mol. Phys.* **2009**, *107*, 1243: Free energies from adaptive reaction coordinate forces (FEARCF): An application to ring puckering.
- [143] A. Laio, M. Parrinello, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562: Escaping free-energy minima.
- [144] A. Laio, A. Rodriguez-Fortea, F.L. Gervasio, M. Ceccarelli, M. Parrinello, *J. Phys. Chem. B* **2005**, *109*, 6714: Assessing the accuracy of metadynamics.

- [145] E. Darve, D. Rodriguez-Gomez, A. Pohorille, *J. Chem. Phys.* **2008**, *128*, 144120/1: Adaptive biasing force method for scalar and vector free energy calculation.
- [146] H.S. Hansen, P.H. Hünenberger, *J. Comput. Chem.* **2010**, *31*, 1: Using the local elevation method to construct optimized umbrella sampling potentials: Calculation of the relative free energies and interconversion barriers of glucopyranose ring conformers in water.
- [147] H.S. Hansen, X. Daura, P.H. Hünenberger, *J. Chem. Theory Comput.* **2010**, *6*, 2598: Enhanced conformational sampling in molecular dynamics simulations of solvated peptides: fragment-based local elevation umbrella sampling.
- [148] H.S. Hansen, P.H. Hünenberger, *J. Chem. Theory Comput.* **2010**, *6*, 2622: Ball-and-stick local elevation umbrella sampling: molecular simulations involving enhanced sampling within conformational or alchemical subspaces of low internal dimensionalities, minimal irrelevant volume and problem-adapted geometries.
- [149] N.S. Bieler, P.H. Hünenberger, *J. Chem. Phys.* **2015**, *142*, 165102/1: On the ambiguity of conformational states: a B&S-LEUS simulation study of the helical conformations of decaalanine in water.
- [150] C.D. Christ, W.F. van Gunsteren, *J. Chem. Phys.* **2007**, *126*, 184110/1: Enveloping distribution sampling: A method to calculate free energy differences from a single simulation.
- [151] C.D. Christ, W.F. van Gunsteren, *J. Chem. Phys.* **2008**, *128*, 174112/1: Multiple free energies from a single simulation: Extending enveloping distribution sampling to nonoverlapping phase-space distributions.
- [152] T. Huber, A.E. Torda, W.F. van Gunsteren, *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695: Local elevation: A method for improving the searching properties of molecular dynamics simulation.
- [153] L. Perić-Hassler, H.S. Hansen, R. Baron, P.H. Hünenberger, *Carbohydr. Res.* **2010**, *345*, 1781: Conformational properties of glucose-based disaccharides using molecular dynamics simulations with local elevation umbrella sampling.
- [154] N.S. Bieler, R. Häuselmann, P.H. Hünenberger, *J. Chem. Theory Comput.* **2014**, *10*, 3006: Local elevation umbrella sampling applied to the calculation of alchemical free-energy changes via  $\lambda$ -dynamics: The  $\lambda$ -LEUS scheme.
- [155] N.S. Bieler, P.H. Hünenberger, *J. Chem. Phys.* **2014**, *141*, 201101/1: Communication: Estimating the initial biasing potential for  $\lambda$ -local-elevation umbrella-sampling ( $\lambda$ -LEUS) simulations via slow growth.
- [156] N.S. Bieler, P.H. Hünenberger, *J. Comput. Chem.* **2015**, *36*, 1686: Orthogonal sampling in free-energy calculations of residue mutations in a tripeptide: T1 vs.  $\lambda$ -LEUS.
- [157] N.S. Bieler, J.P. Tschopp, P.H. Hünenberger, *J. Chem. Theory Comput.* **2015**, *11*, 2575: Multistate  $\lambda$ -local-elevation umbrella-sampling (MS- $\lambda$ -LEUS): Method and application to the complexation of cations by crown ethers.
- [158] D.F. Hahn, P.H. Hünenberger, *J. Chem. Theory Comput.* **2019**, *15*, 2392: Alchemical free-energy calculations by multiple-replica  $\lambda$ -dynamics: The conveyor belt thermodynamic integration (CBTI) scheme.

- [159] H. Ihm, J.-S. Ahn, M.S. Lah, Y.H. Ko, K. Paek, *Org. Lett.* **2004**, *6*, 3893: Oligobisvelcraplex: Self-assembled linear oligomer by solvophobic  $\pi - \pi$  stacking interaction of bisvelcrands based on resorcin[4]arene.
- [160] C.R. Groom, I.J. Bruno, M.P. Lightfoot, S.C. Ward, *Acta Cryst.* **2016**, *B72*, 171: The Cambridge Structural Database.
- [161] E.F. Krause, *Taxicab Geometry.*; Dover Publications, N.Y., USA, **1986**.
- [162] C. Habermann, F. Kindermann, *Comput. Econ.* **2007**, *30*, 153: Multidimensional spline interpolation: Theory and applications.
- [163] P. Wu, X. Hu, W. Yang, *J. Phys. Chem. Lett.* **2011**, *2*, 2099:  $\lambda$ -metadynamics approach to compute absolute solvation free energy.
- [164] W.F. van Gunsteren, *The GROMOS software for biomolecular simulation. Available at: <http://www.gromos.net>* **05/05/2011**.
- [165] N. Schmid, C.D. Christ, M. Christen, A.P. Eichenberger, W.F. van Gunsteren, *Comput. Phys. Commun.* **2012**, *183*, 890: Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation.
- [166] A.-P. E. Kunz, J.R. Allison, D.P. Geerke, B.A.C. Horta, P.H. Hünenberger, S. Riniker, N. Schmid, W.F. van Gunsteren, *J. Comput. Chem.* **2012**, *33*, 340: New functionalities in the GROMOS biomolecular simulation software.
- [167] B.A.C. Horta, P.T. Merz, P. Fuchs, J. Dolenc, S. Riniker, P.H. Hünenberger, *J. Chem. Theory Comput.* **2016**, *12*, 3825: A GROMOS-compatible force field for small organic molecules in the condensed phase: The 2016H66 parameter set.
- [168] I.G. Tironi, W.F. van Gunsteren, *Mol. Phys.* **1994**, *83*, 381: A molecular dynamics study of chloroform.
- [169] W.F. van Gunsteren, H.J.C. Berendsen, *Mol. Simul.* **1988**, *1*, 173: A leap-frog algorithm for stochastic dynamics.
- [170] P. Langevin, *C. R. Acad. Sci. Paris* **1908**, *146*, 530: Sur la théorie du mouvement brownien.
- [171] S. Yun-yu, W. Lu, W.F. van Gunsteren, *Mol. Simul.* **1988**, *1*, 369: On the approximation of solvent effects on the conformation and dynamics of cyclosporin A by stochastic dynamics simulation techniques.
- [172] J.A. Barker, R.O. Watts, *Mol. Phys.* **1973**, *26*, 789: Monte Carlo studies of the dielectric properties of water-like models.
- [173] I.G. Tironi, R. Sperb, P.E. Smith, W.F. van Gunsteren, *J. Chem. Phys.* **1995**, *102*, 5451: A generalized reaction field method for molecular dynamics simulations.
- [174] J.A. Riddick, W.B. Bunger, T.K. Sakano, *Techniques of chemistry II: Organic solvents, physical properties and methods of purification. Volume 2. Edition 4.*; John Wiley & Sons, New York, USA, **1986**.
- [175] C. Wohlfahrt, Pure liquids: Data. In: *Landolt-Börnstein. Numerical data and functional relationships in science and technology. Volume 6.*; O. Madelung, Ed.; Springer, Berlin, Germany, **1991**; pp 5-228.

- [176] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327: Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes.
- [177] D.F. Hahn, *Animations of Resorcin[4]arene Simulations*. Available at: <https://doi.org/10.3929/ethz-b-000376729> **2019**.
- [178] C.A. Schalley, *Analytical Methods in Supramolecular Chemistry Edition 2nd, Completely Revised and Enlarged Edition.*; Wiley-VCH, Weinheim, Germany, **2012**.
- [179] C.M. Baker, P.E.M. Lopes, X. Zhu, B. Roux, A.D. McKerell, *J. Chem. Theory Comput.* **2010**, *6*, 1181: Accurate calculation of hydration free energies using pair-specific Lennard-Jones parameters in the CHARMM Drude polarizable force field.
- [180] P.S. Nernberg, B. Jo, C. So, A. Tripathy, T. Head-Gordon, *J. Phys. Chem. B* **2012**, *116*, 4524: Optimizing solute-water van der Waals interactions to reproduce solvation free energies.
- [181] D.E. Chapman, J.K. Steck, P.S. Nernberg, *J. Chem. Theory Comput.* **2014**, *10*, 273: Optimizing protein-protein van der Waals interactions for the AMBER ff9x/ff12 force field.
- [182] J.P.M. Jämbbeck, A.P. Lyubartsev, *J. Phys. Chem. B* **2014**, *118*, 3793: Update of the general Amber force field for small solutes with an emphasis on free energies of hydration.
- [183] S. Piana, A.G. Donchev, P. Robustelli, D.E. Shaw, *J. Phys. Chem. B* **2015**, *119*, 5113: Water dispersion interactions strongly influence simulated structural properties of disordered protein states.
- [184] M. Yang, A.H. Aytenfisu, A.D. MacKerell Jr., *Carbohydr. Res.* **2018**, *457*, 41: Proper balance of solvent-solute and solute-solute interactions in the treatment of the diffusion of glucose using the Drude polarizable force field.
- [185] Z. Gong, H. Sun, B.E. Eichinger, *J. Chem. Theory. Comput.* **2018**, *14*, 3595: On the temperature transferability of force field parameters for dispersion interactions.
- [186] S. Piana, K. Lindorff-Larsen, D.E. Shaw, *Biophys. J.* **2011**, *100*, L47: How robust are protein folding simulations with respect to force field parameterization.
- [187] F. Martín-García, E. Papaleo, P. Gomez-Puertas, W. Boomsma, K. Lindorff-Larsen, *PLoS ONE* **2015**, *10*, e0121114/1: Comparing molecular dynamics force fields in the essential subspace.
- [188] F. Vitalini, A.S.J.S. Mey, F. Noé, B.G. Keller, *J. Chem. Phys.* **2015**, *142*, 084101/1: Dynamic properties of force fields.
- [189] S. Riniker, *J. Chem. Inf. Model.* **2018**, *58*, 565: Fixed-charge atomistic force fields for molecular dynamics simulations in the condensed phase: An overview.
- [190] A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227: Theoretical studies of enzymic reactions - dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme.

- [191] M.J. Field, P.A. Bash, M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700: A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations.
- [192] H.M. Senn, W. Thiel, *Topics Curr. Chem.* **2007**, *268*, 173: QM/MM methods for biological systems.
- [193] K. Meier, N. Schmid, W.F. van Gunsteren, *J. Comput. Chem.* **2012**, *33*, 2108: Interfacing the GROMOS (bio)molecular simulation software to quantum-chemical program packages.
- [194] K. Meier, A. Choutko, J. Dolenc, A.P. Eichenberger, S. Riniker, W.F. van Gunsteren, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 2820: Multi-resolution simulation of biomolecular systems: A review of methodological issues.
- [195] C.J. Cramer, *Essentials of computational chemistry: Theories and models. Edition 2.*; John Wiley & Sons, Chichester, UK, **2004**.
- [196] M.J.S. Dewar, W. Thiel, *J. Am. Chem. Soc.* **1977**, *99*, 4899: Ground-states of molecules. 38. MNDO method - approximations and parameters.
- [197] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902: The development and use of quantum-mechanical molecular-models. 76. AM1 - a new general-purpose quantum-mechanical molecular-model.
- [198] J.J.P. Stewart, *J. Comput. Chem.* **1989**, *10*, 209: Optimization of parameters for semiempirical methods. 1. Method.
- [199] K. Govender, J. Gao, K.J. Naidoo, *J. Chem. Theory Comput.* **2014**, *10*, 4694: AM1/d-CB1: A semiempirical model for QM/MM simulations of chemical biology systems.
- [200] T.A. Manz, D.S. Sholl, *J. Chem. Theory Comput.* **2010**, *6*, 2455: Chemically meaningful atomic charges that reproduce the electrostatic potential in periodic and nonperiodic materials.
- [201] T.A. Manz, D.S. Sholl, *J. Chem. Theory Comput.* **2012**, *8*, 2844: Improved atoms-in-molecule charge partitioning functional for simultaneously reproducing the electrostatic potential and chemical states in periodic and nonperiodic materials.
- [202] T.A. Manz, N.G. Limas, *RSC Adv.* **2016**, *6*, 47771: Introducing DDEC6 atomic population analysis: Part 1. Charge partitioning theory and methodology.
- [203] P. Bleiziffer, K. Schaller, S. Riniker, *J. Chem. Inf. Model.* **2018**, *58*, 579: Machine learning of partial charges derived from high-quality quantum-mechanical calculations.
- [204] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*, 165: Electronic structure calculations on workstation computers: The program system TURBOMOLE.
- [205] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864: Inhomogeneous electron gas.
- [206] W. Kohn, L.J. Sham, *Phys. Rev.* **1965**, *140*, A1133: Self-consistent equations including exchange and correlation effects.

- [207] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297: Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy.
- [208] F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057: Accurate Coulomb-fitting basis sets for H to Rn.
- [209] V.N. Staroverov, G.E. Scuseria, J. Tao, J.P. Perdew, *J. Chem. Phys.* **2003**, *119*, 12129: Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes.
- [210] A. Klamt, V. Jonas, T. Burger, J.C.W. Lohrenz, *J. Phys. Chem. A* **1998**, *102*, 5074: Refinement and parametrization of COSMO-RS.
- [211] F. Eckert, A. Klamt, *AIChE J.* **2002**, *48*, 369: Fast solvent screening via quantum chemistry: COSMO-RS approach.
- [212] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, *Gaussian 09, Revision D.01*; Gaussian, Inc., Wallingford CT, **2009**.
- [213] N. Godbout, D.R. Salahub, J. Andzelm, E. Wimmer, *Can. J. Chem.* **1992**, *70*, 560: Optimization of Gaussian-type basis sets for local spin density functional calculations. Part I. Boron through neon, optimization technique and validation.
- [214] C. Sosa, J. Andzelm, B.C. Elkin, E. Wimmer, K.D. Dobbs, D.A. Dixon, *J. Phys. Chem.* **1992**, *96*, 6630: A local density functional study of the structure and vibrational frequencies of molecular transition-metal compounds.
- [215] S.H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200: Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis.
- [216] C. Lee, W. Yang, R.G. Parr, *Phys. Rev. B* **1988**, *37*, 785: Development of the Colic-Salvetti correlation-energy formula into a functional of the electron density.
- [217] A.D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648: Density-functional thermochemistry. III. The role of exact exchange.
- [218] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623: Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Field.
- [219] M. Head-Gordon, J.A. Popla, M.J. Frisch, *Chem. Phys. Lett.* **1988**, *153*, 503: MP2 energy evaluation by direct methods.

- [220] M.J. Frisch, M. Head-Gordon, J.A. Pople, *Chem. Phys. Lett.* **1990**, *166*, 275: A direct MP2 gradient method.
- [221] T.H. Dunning Jr., *J. Chem. Phys.* **1989**, *90*, 1007: Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen.
- [222] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33: VMD - Visual molecular dynamics.
- [223] C.D. Christ, A.E. Mark, W.F. van Gunsteren, *J. Comput. Chem.* **2010**, *31*, 1569: Basic ingredients of free energy calculations: A review.
- [224] J.D. Chodera, D.L. Mobley, M.R. Shirts, R.W. Dixon, K. Branson, V.S. Pande, *Curr. Opin. Struct. Biol.* **2011**, *21*, 150: Alchemical free energy methods for drug discovery: Progresses and challenges.
- [225] N. Hansen, W.F. van Gunsteren, *J. Chem. Theory Comput.* **2014**, *10*, 2632: Practical aspects of free-energy calculations: A review.
- [226] W.F. van Gunsteren, H.J.C. Berendsen, *J. Comput.-Aided Mol. Design* **1987**, *1*, 171: Thermodynamic cycle integration by computer simulation as a tool for obtaining free energy differences in molecular chemistry.
- [227] W.L. Jorgensen, J.K. Buckner, S. Boudon, J. Tirado-Rives, *J. Chem. Phys.* **1988**, *89*, 3742: Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water.
- [228] J. Comer, B. Roux, C. Chipot, *Mol. Sim.* **2014**, *40*, 218: Achieving ergodic sampling using replica-exchange free-energy calculations.
- [229] D. Markthaler, J. Gebhardt, S. Jakobtorweihen, N. Hansen, *Chem. Ing. Tech.* **2017**, *89*, 1306: Molecular simulations of thermodynamic properties for the system  $\alpha$ -cyclodextrin/alcohol in aqueous solution.
- [230] J. Baz, J. Gebhardt, H. Kraus, D. Markthaler, N. Hansen, *Chem. Ing. Tech.* **2018**, *90*, 1864: Insights into non-covalent binding obtained from molecular dynamics simulations.
- [231] M.R. Shirts, D.L. Mobley, J.D. Chodera, *Annu. Rep. Comput. Chem.* **2007**, *3*, 41: Alchemical free energy calculations: Ready for prime time.
- [232] A.E. Mark, Y. Xu, H. Liu, W.F. van Gunsteren, *Acta Biochim. Pol.* **1995**, *42*, 525: Rapid non-empirical approaches for estimating relative binding free energies.
- [233] H. Liu, A.E. Mark, W.F. van Gunsteren, *J. Phys. Chem.* **1996**, *100*, 9485: Estimating the relative free energy of different molecular states with respect to a single reference state.
- [234] A.E. Mark, H. Schäfer, H. Liu, W.F. van Gunsteren, Estimating relative free energies from a single simulation of the initial state. In: *Computational molecular dynamics: challenges, methods, ideas, Proceedings of the 2nd Intl. Symp. on Algorithms for Macromol. Mod.*; P. Deufhard, J. Hermans, B. Leimkuhler, A.E. Mark, S. Reich, R.D. Skeel, Eds.; Springer-Verlag, Berlin, Germany, **1999**; pp 149-162.
- [235] H. Schäfer, W.F. van Gunsteren, A.E. Mark, *J. Comput. Chem.* **1999**, *20*, 1604: Estimating relative free energies from a single ensemble: Hydration free energies.

- [236] E.P. Raman, K. Vanommeslaeghe, A.D. MacKerell Jr., *J. Chem. Theory Comput.* **2012**, *8*, 3513: Site-specific fragment identification guided by single-step free energy perturbation calculation.
- [237] E.P. Raman, S.K. Lakkaraju, R.A. Denny, A.D. MacKerell Jr., *J. Comput. Chem.* **2017**, *38*, 1238: Estimation of relative free energies of binding using precomputed ensembles based on the single-step free energy perturbation and the site-identification by ligand competitive saturation approaches.
- [238] S. Boresch, H.L. Woodcock, *Mol. Phys.* **2017**, *115*, 1200: Convergence of single-step free energy perturbation.
- [239] N. Lu, D.A. Kofke, *J. Chem. Phys.* **2001**, *114*, 7303: Accuracy of free-energy perturbation calculations in molecular simulation. I. Modeling.
- [240] N. Lu, D.A. Kofke, *J. Chem. Phys.* **2001**, *115*, 6866: Accuracy of free-energy perturbation calculations in molecular simulation. II. Heuristics.
- [241] T.C. Beutler, A.E. Mark, R. van Schaik, P.R. Gerber, W.F. van Gunsteren, *Chem. Phys. Lett.* **1994**, *222*, 529: Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations.
- [242] C. Oostenbrink, W.F. van Gunsteren, *J. Comput. Chem.* **2003**, *24*, 1730: Single-step perturbations to calculate free energy differences from unphysical reference states: limits on size, flexibility and character.
- [243] C. Oostenbrink, W.F. van Gunsteren, *Proteins: Struct. Func. Genet.* **2004**, *54*, 237: Free energies of binding of polychlorinated biphenyls to the estrogen receptor from a single simulation.
- [244] I.V. Khavrutskii, A. Wallqvist, *J. Chem. Theory Comput.* **2011**, *7*, 3001: Improved binding free energy predictions from single-reference thermodynamic integration augmented with Hamiltonian replica exchange.
- [245] A.E. Mark, W.F. van Gunsteren, H.J.C. Berendsen, *J. Chem. Phys.* **1991**, *94*, 3808: Calculation of relative free energy via indirect pathways.
- [246] A. Mentes, N.-J. Deng, R.S.K. Vijayan, J. Xia, E. Gallicchio, R.M. Levy, *J. Chem. Theory Comput.* **2016**, *12*, 2459: Binding energy distribution analysis method: Hamiltonian replica exchange with torsional flattening for binding mode prediction and binding free energy estimation.
- [247] L. Wang, B.J. Berne, *J. Chem. Phys.* **2018**, *149*, 072306/1: Efficient sampling of puckering states of monosaccharides through replica exchange with solute tempering and bond softening.
- [248] J. Xia, W. Flynn, R.M. Levy, *J. Chem. Inf. Model.* **2018**, *58*, 1356: Improving prediction accuracy of binding free energies and poses of HIV integrase complexes using the binding energy distribution analysis method with flattening potentials.
- [249] C.D. Christ, W.F. van Gunsteren, *J. Chem. Theory Comput.* **2009**, *5*, 276: Simple, efficient, and reliable computation of multiple free energy differences from a single simulation: A reference Hamiltonian parameter update scheme for enveloping distribution sampling (EDS).
- [250] C.D. Christ, W.F. van Gunsteren, *J. Comput. Chem.* **2009**, *30*, 1664: Comparison of three enveloping distribution sampling Hamiltonians for the estimation of multiple free energy differences from a single simulation.

- [251] M.R. Shirts, J.D. Chodera, *J. Chem. Phys.* **2008**, *129*, 124105/1: Statistically optimal analysis of samples from multiple equilibrium states.
- [252] C.H. Bennett, *J. Comput. Phys.* **1976**, *22*, 245: Efficient estimation of free energy differences from Monte Carlo data.
- [253] M.R. Shirts, E. Bair, G. Hooker, V.S. Pande, *Phys. Rev. Lett.* **2003**, *91*, 140601/1: Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods.
- [254] T.P. Straatsma, J.A. McCammon, *J. Chem. Phys.* **1991**, *95*, 1175: Multiconfiguration thermodynamic integration.
- [255] M. Jorge, N.M. Garrido, A.J. Queimada, I.G. Economou, E.A. Macedo, *J. Chem. Theory Comput.* **2010**, *6*, 1018: Effect of the integration method on the accuracy and computational efficiency of free energy calculations using thermodynamic integration.
- [256] S. Bruckner, S. Boresch, *J. Comput. Chem.* **2011**, *32*, 1303: Efficiency of alchemical free energy simulations. I. A practical comparison of the exponential formula, thermodynamic integration, and Bennett's acceptance ratio method.
- [257] S. Bruckner, S. Boresch, *J. Comput. Chem.* **2011**, *32*, 1320: Efficiency of alchemical free energy simulations. II. Improvements for thermodynamic integration.
- [258] C. Shyu, F.M. Ytreberg, *J. Comput. Chem.* **2009**, *30*, 2297: Reducing the bias and uncertainty of free energy estimates by using regression to fit thermodynamic integration data.
- [259] J.Z. Vilseck, O. Acevedo, *Annu. Rep. Comput. Chem.* **2010**, *6*, 37: Computing free-energy profiles using multidimensional potentials of mean force and polynomial quadrature methods.
- [260] C. Shyu, F.M. Ytreberg, *J. Comput. Chem.* **2011**, *32*, 134: Accurate estimation of solvation free energy using polynomial fitting techniques.
- [261] W.L. Jorgensen, L.L. Thomas, *J. Chem. Theory Comput.* **2008**, *4*, 869: Perspective on free-energy perturbation calculations for chemical equilibria.
- [262] A.J. Cross, *Chem. Phys. Lett.* **1986**, *128*, 198: Influence of Hamiltonian parameterization on convergence of Kirkwood free energy calculations.
- [263] A. Blondel, *J. Comput. Chem.* **2004**, *25*, 985: Ensemble variance in free energy calculations by thermodynamic integration: theory, optimal "alchemical" path, and practical solutions.
- [264] T.T. Pham, M.R. Shirts, *J. Chem. Phys.* **2011**, *135*, 034114/1: Identifying low variance pathways for free energy calculations of molecular transformations in solution phase.
- [265] T.T. Pham, M.R. Shirts, *J. Chem. Phys.* **2012**, *136*, 124120/1: Optimal pairwise and non-pairwise alchemical pathways for free energy calculations of molecular transformations in solution phase.
- [266] L.N. Naden, T.T. Pham, M.R. Shirts, *J. Chem. Theory Comput.* **2014**, *10*, 1128: Linear basis function approach to efficient alchemical free energy calculations. 1. Removal of uncharged atomic sites.

- [267] L.N. Naden, M.R. Shirts, *J. Chem. Theory Comput.* **2015**, *11*, 2536: Linear basis function approach to efficient alchemical free energy calculations. 2. Inserting and deleting particles with Coulombic interactions.
- [268] M. Mezei, *Mol. Simul.* **1993**, *10*, 225: Calculation of solvation free-energy differences for large solute change from computer simulation with quadrature-based nearly linear thermodynamic integration.
- [269] S. Hug, M. Schwarzfischer, J. Hasenauer, C. Marr, F.J. Theis, *Stat. Comput.* **2016**, *26*, 663: An adaptive scheduling scheme for calculating Bayes factors with thermodynamic integration using Simpson's rule.
- [270] A. Mecklenfeld, G. Raabe, *Mol. Phys.* **2017**, *115*, 1322: Efficient solvation free energy simulations: impact of soft-core potential and a new adaptive  $\lambda$ -spacing method.
- [271] Z.X. Sun, X.H. Wang, J.Z.H. Zhang, *Phys. Chem. Chem. Phys.* **2017**, *19*, 15005: BAR-based optimum adaptive sampling regime for variance minimization in alchemical transformation.
- [272] J. Dakka, K. Farkas-Pall, V. Balasubramanian, M. Turilli, S. Wan, D.W. Wright, S. Zasada, P.V. Coveney, S. Jha, *IEEE Conf. 2018, eScience*, : Concurrent and adaptive extreme scale binding free energy calculations.
- [273] A. de Ruiter, C. Oostenbrink, *J. Chem. Theory Comput.* **2016**, *12*, 4476: Extended thermodynamic integration: Efficient prediction of lambda derivatives at nonsimulated points.
- [274] Z. Liu, B.J. Berne, *J. Chem. Phys.* **1993**, *99*, 6071: Method for accelerating chain folding and mixing.
- [275] B. Tidor, *J. Phys. Chem.* **1993**, *97*, 1069: Simulated annealing on free energy surfaces by a combined molecular dynamics and Monte Carlo approach.
- [276] X. Kong, C.L. Brooks III, *J. Chem. Phys.* **1996**, *105*, 2414:  $\lambda$ -dynamics: A new approach to free energy calculations.
- [277] Y. Sugita, Y. Okamoto, *Chem. Phys. lett.* **1999**, *134*, 141: Replica-exchange molecular dynamics method for protein folding.
- [278] A. Mitsutake, Y. Sugita, Y. Okamoto, *Biopolymers* **2001**, *60*, 96: Generalized-ensemble algorithms for molecular simulations of biopolymers.
- [279] Y. Iba, *Int. J. Mod. Phys. C* **2001**, *12*, 623: Extended ensemble Monte Carlo.
- [280] M.M.H. Graf, M. Maurer, C. Oostenbrink, *J. Comput. Chem.* **2016**, *37*, 2597: Free-energy calculations of residue mutations in a tripeptide using various methods to overcome inefficient sampling.
- [281] C.J. Woods, J.W. Essex, *J. Phys. Chem. B* **2003**, *107*, 13703: The development of replica-exchange-based free-energy methods.
- [282] C.J. Woods, J.W. Essex, M.A. King, *J. Phys. Chem. B* **2003**, *107*, 13711: Enhanced configurational sampling in binding free-energy calculations.
- [283] J. Hritz, C. Oostenbrink, *J. Chem. Phys.* **2007**, *127*, 204104/1: Optimization of replica exchange molecular dynamics by fast mimicking interactions.

- [284] Y. Deng, B. Roux, *J. Chem. Phys.* **2008**, *128*, 115103/1: Computation of binding free energy with molecular dynamics and grand canonical MC simulations.
- [285] C.A.F. de Oliveira, D. Hamelberg, J.A. McCammon, *J. Chem. Theory Comput.* **2008**, *4*, 1516: Coupling accelerated molecular dynamics methods with thermodynamic integration simulations.
- [286] M. Fajer, D. Hamelberg, J.A. McCammon, *J. Chem. Theory Comput.* **2008**, *4*, 1565: Replica-exchange accelerated molecular dynamics (REX-AMD) applied to thermodynamic integration.
- [287] J. Hritz, C. Oostenbrink, *J. Chem. Phys.* **2008**, *128*, 144121/1: Hamiltonian replica exchange molecular dynamics using soft-core interactions.
- [288] J. Hritz, C. Oostenbrink, *J. Phys. Chem. B* **2009**, *113*, 12711: Efficient free energy calculations for compounds with multiple stable conformations separated by high energy barriers.
- [289] I.V. Khavrutskii, A. Wallqvist, *J. Chem. Theory Comput.* **2010**, *6*, 3427: Computing relative free energies of solvation using single reference thermodynamic integration augmented with Hamiltonian replica exchange.
- [290] U. Doshi, D. Hamelberg, *J. Chem. Theory Comput.* **2012**, *8*, 4004: Improved statistical sampling and accuracy with accelerated molecular dynamics on rotatable torsions.
- [291] L. Wang, B.J. Berne, R.A. Friesner, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1937: On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities.
- [292] J.A. Garate, C. Oostenbrink, *J. Comput. Chem.* **2013**, *34*, 1398: Free energy differences between states with different conformational ensembles.
- [293] J.W. Kaus, M. Arrar, J.A. McCammon, *J. Phys. Chem. B* **2014**, *118*, 5109: Accelerated adaptive integration method.
- [294] H. Fukunishi, O. Watanabe, S. Takada, *J. Chem. Phys.* **2002**, *116*, 9058: On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction.
- [295] B.W. Zhang, W. Dai, E. Gallicchio, P. He, J. Xia, Z. Tan, R.M. Levy, *J. Phys. Chem. B* **2016**, *120*, 8289: Simulating replica exchange: Markov state models, proposal schemes, and the infinite swapping limit.
- [296] F. Calvo, *J. Chem. Phys.* **2005**, *123*, 124106/1: All-exchanges parallel tempering.
- [297] P. Brenner, C.R. Sweet, D. VonHandorf, J.A. Izquierdo, *J. Chem. Phys.* **2007**, *126*, 074103/1: Accelerating the replica exchange method through an efficient all-pairs exchange.
- [298] J.D. Chodera, M.R. Shirts, *J. Chem. Phys.* **2011**, *135*, 194110/1: Replica exchange and expanded ensemble simulations as Gibbs sampling: Simple improvements for enhanced mixing.
- [299] H. Li, G. Li, B.A. Berg, W. Yang, *J. Chem. Phys.* **2006**, *125*, 144902/1: Finite reservoir replica exchange to enhance canonical sampling in rugged energy surfaces.

- [300] A. Okur, D.R. Roe, G. Cui, V. Hornak, C. Simmerling, *J. Chem. Theory Comput.* **2007**, *3*, 557: Improving convergence of replica-exchange simulations through coupling to a high-temperature structure reservoir.
- [301] A.E. Roitberg, A. Okur, C. Simmerling, *J. Phys. Chem. B* **2007**, *111*, 2415: Coupling of replica exchange simulations to a non-Boltzmann structure reservoir.
- [302] J.Z. Ruscio, N.L. Fawzi, T. Head-Gordon, *J. Comput. Chem.* **2010**, *31*, 620: How hot? Systematic convergence of the replica exchange method using multiple reservoirs.
- [303] A. Okur, B.T. Miller, K. Joo, J. Lee, B.R. Brooks, *J. Chem. Theory Comput.* **2013**, *9*, 1115: Generating reservoir conformations for replica exchange through the use of the conformational space annealing method.
- [304] N.M. Henriksen, D.R. Roe, T.E. Cheatham III, *J. Phys. Chem. B* **2013**, *117*, 4014: Reliable oligonucleotide conformational ensemble generation in explicit solvent for force field assessment using reservoir replica exchange molecular dynamics simulations.
- [305] A. Damjanovic, B.T. Miller, A. Okur, B.R. Brooks, *J. Chem. Phys.* **2018**, *149*, 072321/1: Reservoir pH replica exchange.
- [306] C. Chen, Y. Xiao, Y. Huang, *Phys. Rev. E* **2015**, *91*, 052708/1: Improving the replica-exchange molecular-dynamics method for efficient sampling in the temperature space.
- [307] C. Chen, Y. Huang, *J. Comput. Chem.* **2016**, *37*, 1565: Walking freely in the energy and temperature space by the modified replica exchange molecular dynamics method.
- [308] X. Li, C.P. O'Brien, G. Collier, N.A. Vellore, F. Wang, R.A. Latour, D.A. Bruce, S.J. Stuart, *J. Chem. Phys.* **2007**, *127*, 164116/1: An improved replica-exchange sampling method: Temperature intervals with global energy reassignment.
- [309] X. Li, R.A. Latour, S.J. Stuart, *J. Chem. Phys.* **2009**, *130*, 174106/1: TIGER2: An improved algorithm for temperature intervals with global exchange of replicas.
- [310] X. Li, J.A. Snyder, S.J. Stuart, R.A. Latour, *J. Chem. Phys.* **2015**, *143*, 144105/1: TIGER2 with solvent energy averaging (TIGER2A): An accelerated sampling method for large molecular systems with explicit representation of solvent.
- [311] M. Kulke, N. Geist, D. Möller, W. Langel, *J. Phys. Chem. B* **2018**, *122*, 7295: Replica-based protein structure sampling methods: Compromising between explicit and implicit solvents.
- [312] D. Sindhwikara, Y. Meng, A.E. Roitberg, *J. Chem. Phys.* **2008**, *128*, 024103/1: Exchange frequency in replica exchange molecular dynamics.
- [313] N. Plattner, J.D. Doll, J.D. Dupuis, H. Wang, Y. Liu, J.E. Gubernatis, *J. Chem. Phys.* **2011**, *135*, 134111/1: An infinite swapping approach to the rare-event sampling problem.
- [314] P. Dupuis, Y. Liu, N. Plattner, J.D. Doll, *Multiscale Model. Simul.* **2012**, *10*, 986: On the infinite swapping limit for parallel tempering.

- [315] N. Plattner, J.D. Doll, M. Meuwly, *J. Chem. Theory Comput.* **2013**, *9*, 4215: Overcoming the rare event sampling problem in biological systems with infinite swapping.
- [316] J. Lu, E. Vanden-Eijnden, *J. Chem. Phys.* **2013**, *138*, 084105/1: Infinite swapping replica exchange molecular dynamics leads to a simple simulation patch using mixture potentials.
- [317] J. Kim, T. Keyes, J.E. Straub, *J. Chem. Phys.* **2010**, *132*, 224107/1: Generalized replica exchange method.
- [318] Q. Lu, J. Kim, J.E. Straub, *J. Phys. Chem. B* **2012**, *116*, 8654: Exploring the solid-liquid phase change of an adapted Dzugutov model using generalized replica exchange method.
- [319] Q. Lu, J. Kim, J.E. Straub, *J. Chem. Phys.* **2013**, *138*, 104119/1: Order parameter free enhanced sampling of the vapor-liquid transition using the generalized replica exchange method.
- [320] Q. Lu, J. Kim, J.D. Farrell, D.J. Wales, J.E. Straub, *J. Chem. Phys.* **2014**, *141*, 18C525/1: Investigating the solid-liquid phase transition of water nanofilms using the generalized replica exchange method.
- [321] E. Małłolepsza, M. Secor, T. Keyes, *J. Phys. Chem. B* **2015**, *119*, 13379: Isobaric molecular dynamics version of the generalized replica exchange method (gREM): Liquid-vapor equilibrium.
- [322] E. Małłolepsza, T. Keyes, *J. Chem. Theory Comput.* **2015**, *11*, 5613: Water freezing and ice melting.
- [323] H. Suwa, S. Todo, *Phys. Rev. Lett.* **2010**, *105*, 120603/1: Markov chain Monte Carlo method without detailed balance.
- [324] S. Todo, H. Suwa, *J. Phys. Conf. Ser.* **2013**, *473*, 012013/1: Geometric allocation approaches in Markov chain Monte Carlo.
- [325] Y. Mori, H. Okumura, *J. Comput. Chem.* **2015**, *36*, 2344: Simulated tempering based on global balance or detailed balance conditions: Suwa-Todo, heat bath, and Metropolis algorithms.
- [326] S.G. Itoh, H. Okumura, *J. Chem. Theory Comput.* **2013**, *9*, 570: Replica-permutation method with the Suwa-Todo algorithm beyond the replica-exchange method.
- [327] S.G. Itoh, H. Okumura, *J. Comput. Chem.* **2013**, *34*, 2493: Hamiltonian replica-permutation method and its applications to an alanine dipeptide and amyloid- $\beta$ (29-42) peptides.
- [328] M. Yamauchi, H. Okumura, *J. Chem. Phys.* **2017**, *147*, 184107/1: Development of isothermal-isobaric replica-permutation method for molecular dynamics and Monte Carlo simulations and its application to reveal temperature and pressure dependence of folded, misfolded, and unfolded states of chignolin.
- [329] K.V. Damodaran, S. Banba, C.L. Brooks III, *J. Phys. Chem. B* **2001**, *105*, 9316: Application of multiple topology  $\lambda$ -dynamics to a host-guest system:  $\beta$ -cyclodextrin with substituted benzenes.
- [330] Z. Guo, J. Durkin, T. Fischermann, R. Ingram, A. Prongay, R. Zhang, V. Madison, *J. Med. Chem.* **2003**, *46*, 5360: Application of the  $\lambda$ -dynamics method to evaluate the relative binding free energies of inhibitors to HCV protease.

- [331] J.L. Knight, C.L. Brooks III, *J. Comput. Chem.* **2009**, *30*, 1692:  $\lambda$ -dynamics free energy simulation methods.
- [332] J.L. Knight, C.L. Brooks III, *J. Chem. Theory Comput.* **2011**, *7*, 2728: Multisite  $\lambda$ -dynamics for simulated structure-activity relationship studies.
- [333] S. Donnini, F. Tegeler, G. Groenhof, H. Grubmüller, *J. Chem. Theory Comput.* **2011**, *7*, 1962: Constant pH molecular dynamics in explicit solvent with  $\lambda$ -dynamics.
- [334] K.A. Armacost, G.B. Goh, C.L. Brooks III, *J. Chem. Theory Comput.* **2015**, *11*, 1267: Biasing potential replica exchange multisite dynamics for efficient free energy calculations.
- [335] R.L. Hayes, K.A. Armacost, J.Z. Vilseck, C.L. Brooks III, *J. Phys. Chem. B* **2017**, *121*, 3626: Adaptive landscape flattening accelerates sampling of alchemical space in multisite  $\lambda$ -dynamics.
- [336] J.L. Knight, C.L. Brooks III, *J. Comput. Chem.* **2011**, *32*, 3423: Applying efficient implicit nongeometric constraints in alchemical free energy simulations.
- [337] J. Kästner, W. Thiel, *J. Chem. Phys.* **2005**, *123*, 144104/1: Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: Umbrella integration.
- [338] J. Kästner, W. Thiel, *J. Chem. Phys.* **2006**, *124*, 234106/1: Analysis of the statistical error in umbrella sampling simulations by umbrella integration.
- [339] J. Kästner, *J. Chem. Phys.* **2012**, *136*, 234102/1: Umbrella integration with higher-order correction terms.
- [340] X. Ding, J.Z. Vilseck, R.L. Hayes, C.L. Brooks III, *J. Chem. Theory Comput.* **2017**, *13*, 2501: Gibbs sampler-based  $\lambda$ -dynamics and Rao-Blackwell estimator for alchemical free energy calculation.
- [341] A. Kone, D.A. Kofke, *J. Chem. Phys.* **2005**, *122*, 206101/1: Selection of temperature intervals for parallel-tempering simulations.
- [342] N. Rathore, M. Chopra, J.J. de Pablo, *J. Chem. Phys.* **2005**, *122*, 024111/1: Optimal allocation of replicas in parallel tempering simulations.
- [343] S. Trebst, M. Troyer, U.H.E. Hansmann, *J. Chem. Phys.* **2006**, *124*, 174903/1: Optimized parallel tempering simulations of proteins.
- [344] W. Nadler, U.H.E. Hansmann, *J. Phys. Chem. B Lett.* **2008**, *112*, 10386: Optimized explicit-solvent replica exchange molecular dynamics from scratch.
- [345] W. Nadler, J.H. Meinke, U.H.E. Hansmann, *Phys. Rev. E* **2008**, *78*, 061905/1: Folding proteins by first-passage-times-optimized replica exchange.
- [346] T. Vogel, D. Perez, *Physics Procedia* **2015**, *68*, 125: Accelerating the convergence of replica exchange simulations using Gibbs sampling and adaptive temperature sets.
- [347] T. Vogel, D. Perez, *Phys. Rev. Lett.* **2015**, *115*, 190602/1: Towards an optimal flow: Density-of-states-informed replica-exchange simulations.

- [348] D. Sidler, A. Schwaninger, S. Riniker, *J. Chem. Phys.* **2016**, *145*, 154114/1: Replica exchange enveloping distribution sampling (RE-EDS): A robust method to estimate multiple free-energy differences from a single simulation.
- [349] D. Sidler, M. Cristòfol-Clough, S. Riniker, *J. Chem. Theory Comput.* **2017**, *13*, 3020: Efficient round-trip time optimization for replica-exchange enveloping distribution sampling (RE-EDS).
- [350] J.L. MacCallum, M.I. Muniyat, K. Gaalswyk, *J. Phys. Chem. B* **2018**, *122*, 5448: Online optimization of total acceptance in Hamiltonian replica exchange simulations.
- [351] S. Donnini, F. Tegeler, G. Groenhof, H. Grubmüller, *J. Chem. Theory Comput.* **2013**, *9*, 3261: Correction to constant pH molecular dynamics in explicit solvent with  $\lambda$ -dynamics.
- [352] L. Zheng, W. Yang, *J. Chem. Theory Comput.* **2012**, *8*, 810: Practically efficient and robust free energy calculations: Double-integration orthogonal space tempering.
- [353] Z. Guo, C.L. Brooks III, X. Kong, *J. Phys. Chem. B* **1998**, *102*, 2032: Efficient and flexible algorithm for free energy calculations using the  $\lambda$ -dynamics approach.
- [354] Z. Guo, C.L. Brooks III, *J. Am. Chem. Soc.* **1998**, *120*, 1920: Rapid screening of binding affinities: Applications of the  $\lambda$ -dynamics method to a trypsin-inhibitor system.
- [355] M. Souaille, B. Roux, *Comput. Phys. Commun.* **2001**, *135*, 40: Extension to the weighted histogram analysis method: Combining umbrella sampling with free energy calculations.
- [356] F. Wang, D.P. Landau, *Phys. Rev. Lett.* **2001**, *86*, 2050: Efficient, multiple-range random walk algorithm to calculate the density of states.
- [357] D.P. Landau, S.-H. Tsai, M. Exler, *J. Chem. Phys.* **2004**, *72*, 1294: A new approach to Monte Carlo simulations in statistical physics: Wang-Landau sampling.
- [358] A. Barducci, G. Bussi, M. Parrinello, *Phys. Rev. Lett.* **2008**, *100*, 020603/1: Well-tempered metadynamics: A smoothly converging and tunable free-energy method.
- [359] M. Fasnacht, R.H. Swendsen, J.M. Rosenberg, *Phys. Rev. E* **2004**, *69*, 056704/1: Adaptive integration method for Monte Carlo simulations.
- [360] A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, *J. Chem. Phys.* **1992**, *96*, 1776: New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles.
- [361] A.P. Lyubartsev, A. Laaksonen, P.N. Vorontsov-Velyaminov, *Mol. Phys.* **1994**, *82*, 455: Free energy calculations for Lennard-Jones systems and water using the expanded ensemble method. A Monte Carlo and molecular dynamics simulation study.
- [362] A.P. Lyubartsev, A. Laaksonen, P.N. Vorontsov-Velyaminov, *Mol. Simul.* **1996**, *18*, 43: Determination of free energy from chemical potentials: Application of the expanded ensemble method.

- [363] F.A. Escobedo, F.J. Martínez-Veracoecha, *J. Chem. Phys.* **2007**, *127*, 174103/1: Optimized expanded ensembles for simulations involving molecular insertions and deletions. I. Closed systems.
- [364] F.A. Escobedo, F.J. Martínez-Veracoecha, *J. Chem. Phys.* **2007**, *127*, 174104/1: Optimized expanded ensembles for simulations involving molecular insertions and deletions. II. Open systems.
- [365] A.S. Paluch, D.L. Mobley, E.J. Maginn, *J. Chem. Theory Comput.* **2011**, *7*, 2910: Small molecule solvation free energy: Enhanced conformational sampling using expanded ensemble molecular dynamics simulation.
- [366] B. K. Radak, S. Donghyuk, B. Roux, *J. Chem. Phys.* **2018**, *149*, 072315/1: A generalized linear response framework for expanded ensemble and replica exchange simulations.
- [367] G.M. Torrie, J.P. Valleau, *Chem. Phys. Lett.* **1974**, *28*, 578: Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid.
- [368] T. Huber, W.F. van Gunsteren, *J. Phys. Chem. A* **1998**, *102*, 5937: SWARM-MD: Searching conformational space by cooperative molecular dynamics.
- [369] K.K. Burusco, N.J. Bruce, I. Alibay, R.A. Bryce, *Chem. Phys. Chem.* **2015**, *16*, 3233: Free energy calculations using a swarm-enhanced sampling molecular dynamics approach.
- [370] H. Kamberaj, *J. Mol. Graph. Model.* **2018**, *81*, 32: Faster protein folding using enhanced conformational sampling of molecular dynamics simulation.
- [371] I. Alibay, K.K. Burusco, N.J. Bruce, R.A. Bryce, *J. Phys. Chem. B* **2018**, *122*, 2462: Identification of rare Lewis oligosaccharide conformers in aqueous solution using enhanced sampling molecular dynamics.
- [372] P. Raiteri, A. Laio, F.L. Gervasio, C. Micheletti, M. Parrinello, *J. Phys. Chem. B* **2006**, *110*, 3533: Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics.
- [373] J. Comer, J.C. Phillips, K. Schulten, C. Chipot, *J. Chem. Theory Comput.* **2014**, *10*, 5276: Multiple-replica strategies for free-energy calculations in NAMD: Multiple-walker adaptive biasing force and walker selection rules.
- [374] Z. Šućur, V. Spiwok, *J. Chem. Theory Comput.* **2016**, *12*, 4644: Sampling enhancement and free energy prediction by the flying Gaussian method.
- [375] P. Kříž, Z. Šućur, V. Spiwok, *J. Phys. Chem. B* **2017**, *121*, 10479: Free-energy surface prediction by flying Gaussian method: Multisystem representation.
- [376] N. Hansen, P.H. Hünenberger, W.F. van Gunsteren, *J. Chem. Theory Comput.* **2013**, *9*, 1334: Efficient combination of environment change and alchemical perturbation within the enveloping distribution sampling (EDS) scheme: Twin-system EDS and application to the determination of octanol-water partition coefficients.
- [377] J. Gebhardt, N. Hansen, *Fluid Phase Equilib.* **2016**, *422*, 1: Calculation of binding affinities for linear alcohols to  $\alpha$ -cyclodextrin by twin-system enveloping distribution sampling simulations.

- [378] N. Lu, D.A. Kofke, T.B. Woolf, *J. Comput. Chem.* **2004**, *25*, 28: Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods.
- [379] M.R. Shirts, V.S. Pande, *J. Chem. Phys.* **2005**, *122*, 144107/1: Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration.
- [380] M. Fajer, R.V. Swift, J.A. McCammon, *J. Comput. Chem.* **2009**, *30*, 1719: Using multistate free energy techniques to improve the efficiency of replica exchange accelerated molecular dynamics.
- [381] Z. Tan, E. Gallicchio, M. Lapejosa, R.M. Levy, *J. Chem. Phys.* **2012**, *136*, 144102/1: Theory of binless multi-state free energy estimation with applications to protein-ligand binding.
- [382] B.W. Zhang, N. Deng, Z. Tan, R.M. Levy, *J. Chem. Theory Comput.* **2017**, *13*, 4660: Stratified UWHAM and its stochastic approximation for multicanonical simulations which are far from equilibrium.
- [383] B.A.C. Horta, P.F.J. Fuchs, W.F. van Gunsteren, P.H. Hünenberger, *J. Chem. Theory Comput.* **2011**, *7*, 1016: New interaction parameters for oxygen compounds in the GROMOS force field: Improved pure-liquid and solvation properties for alcohols, ethers, aldehydes, ketones, carboxylic acids and esters.
- [384] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration. In: *Intermolecular Forces*; B. Pullman, Ed.; Reidel, Dordrecht, The Netherlands, **1981**; pp 331-342.
- [385] T. Simonson, *Mol. Phys.* **1993**, *80*, 441: Free energy of particle insertion. An exact analysis of the origin singularity for simple liquids.
- [386] M. Zacharias, T.P. Straatsma, J.A. McCammon, *J. Chem. Phys.* **1994**, *100*, 9025: Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration.
- [387] M.R. Shirts, J.W. Pitera, W.C. Swope, V.S. Pande, *J. Chem. Phys.* **2003**, *119*, 5740: Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins.
- [388] M.R. Shirts, V.S. Pande, *J. Chem. Phys.* **2005**, *122*, 134508/1: Solvation free energies of amino acid side chain analogs for common molecular mechanics water models.
- [389] D.L. Mobley, E. Dumont, J.D. Chodera, K.A. Dill, *J. Phys. Chem. B* **2007**, *111*, 2242: Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent.
- [390] M. Lingenheil, R. Denschlag, R. Reichold, P. Tavan, *J. Chem. Theory Comput.* **2008**, *4*, 1293: The "hot-solvent/cold-solute" problem revisited.
- [391] C. de Boor, *A practical guide to splines. Applied mathematical sciences. Volume 27.*; Springer New York, **2001**.
- [392] J. Chodera, M. Shirts, *A Python implementation of the multistate Bennett acceptance ratio (MBAR)* Available at: <https://github.com/choderalab/pymbar/> **2017**.

- [393] W.S. Gosset (as "Student"), *Biometrika* **1908**, *6*, 1: The probable error of a mean.
- [394] B. Efron, *Ann. Statist.* **1979**, *7*, 1: Bootstrap methods: Another look at the jackknife.
- [395] B. Efron, R.J. Tibshirani, *An introduction to the bootstrap.*; Chapman, Hall / CRC Taylor, Francis Group, **1998**.
- [396] 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.*; Verlag der Fachvereine, Zürich, Switzerland, **1996**.
- [397] J.C. Maxwell, *Philos. Mag.* **1860**, *19*, 19: V. Illustrations of the dynamical theory of gases. — Part I. On the motions and collisions of perfectly elastic spheres.
- [398] A. Einstein, *Ann. Phys.* **1905**, *322*, 549: Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen.
- [399] J.D. Hunter, *Comput. Sci. Eng.* **2007**, *9*, 90: Matplotlib: A 2D graphics environment.
- [400] S. Cabani, P. Gianni, P. Mollica, L. Lepori, *J. Solut. Chem.* **1981**, *10*, 563: Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution.
- [401] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, Statistical mechanics and molecular dynamics: The calculation of free energy. In: *Molecular dynamics and protein structure (Proceedings Workshop 13-18 May 1984, at University of North Carolina).*; J. Hermans, Ed.; Polycrystal Book Service, P.O. Box 27, Western Springs, Illinois 60558, USA, **1985**; pp 43-46.
- [402] T.P. Straatsma, H.J.C. Berendsen, J.P.M. Postma, *J. Chem. Phys.* **1986**, *85*, 6720: Free energy of hydrophobic hydration: A molecular dynamics study of noble gases in water.
- [403] D.A. Pearlman, P.A. Kollman, *J. Chem. Phys.* **1989**, *91*, 7831: The lag between the Hamiltonian and the system configuration in free energy perturbation calculations.
- [404] J. Hermans, *J. Phys. Chem.* **1991**, *95*, 9029: Simple analysis of noise and hysteresis in (slow-growth) free energy simulations.
- [405] A.E. van Helden Mark, Smith S.P., Janssen P.E., van Gunsteren L.H.M., *J. Am. Chem. Soc.* **1994**, *116*, 6293: Convergence properties of free energy calculations:  $\alpha$ -cyclodextrin complexes as a case study.
- [406] C. Jarzynski, *Phys. Rev. Lett.* **1997**, *78*, 2690: Nonequilibrium equality for free energy differences.
- [407] F.E. Crooks, *Phys. Rev. E* **2000**, *61*, 2361: Path-ensemble averages in systems driven far from equilibrium.
- [408] D.A. Hendrix, C. Jarzynski, *J. Chem. Phys.* **2001**, *114*, 5974: A "fast growth" method of computing free energy differences.

- [409] G. Hummer, *Mol. Simul.* **2002**, *28*, 81: Fast-growth thermodynamic integration: Results for sodium ion hydration.
- [410] E. Gallicchio, J. Xia, W.F. Flynn, B. Zhang, S. Samlalsingh, A. Mentes, R.M. Levy, *Comput. Phys. Commun.* **2015**, *196*, 236: Asynchronous replica exchange software for grid and heterogeneous computing.
- [411] J. Xia, W.F. Flynn, E. Gallicchio, B.W. Zhang, P. He, Z. Tan, R.M. Levy, *J. Comput. Chem.* **2015**, *36*, 1772: Large-scale asynchronous and distributed multidimensional replica exchange molecular simulations and efficiency analysis.
- [412] J.A. Morrone, T.E. Markland, M. Ceriotti, B.J. Berne, *J. Chem. Phys.* **2011**, *134*, 014103/1: Efficient multiple time scale molecular dynamics: Using colored noise thermostats to stabilize resonances.
- [413] D.T. Margul, M.E. Tuckerman, *J. Chem. Theory Comput.* **2016**, *12*, 2170: A stochastic, resonance-free multiple time-step algorithm for polarizable models that permits very large time steps.
- [414] B. Baumann, *Nucl. Phys. B* **1987**, *285*, 391: Noncanonical path and surface simulation.
- [415] B.A. Berg, T. Neuhaus, *Phys. Lett.* **1991**, *267*, 249: Multicanonical algorithms for first order phase transitions.
- [416] B.A. Berg, T. Neuhaus, *Phys. Rev. Lett.* **1992**, *68*, 9: Multicanonical ensemble: A new approach to simulate first-order phase transitions.
- [417] J.Z. Vilseck, K.A. Armacost, R.L. Hayes, G.B. Goh, C.L. Brooks III, *J. Phys. Chem. Lett.* **2018**, *9*, 3328: Predicting binding free energies in a large combinatorial chemical space using multisite lambda dynamics.
- [418] B.L. Tembe, J.A. McCammon, *Comput. Chem.* **1984**, *8*, 281: Ligand-receptor interactions.
- [419] M. Maurer, N. Hansen, C. Oostenbrink, *J. Comput. Chem.* **2018**, *39*, 2226: Comparison of free-energy methods using a tripeptide-water model system.
- [420] F. Haber, *Ber. Dtsch. Physik. Ges.* **1919**, *21*, 750: Betrachtungen zur Theorie der Wärmetonung.
- [421] M. Christen, A.-P.E. Kunz, W.F. van Gunsteren, *J. Phys. Chem. B* **2006**, *110*, 8488: Sampling of rare events using hidden restraints.
- [422] M. Christen, A.-P.E. Kunz, W.F. van Gunsteren, *J. Phys. Chem. B* **2008**, *112*, 11446: Erratum to: "Sampling of rare events using hidden restraints" [J. Phys. Chem. B 110, 8488-8498 (2006)].
- [423] S. Piana, A. Laio, *J. Phys. Chem. B* **2007**, *111*, 4553: A bias-exchange approach to protein folding.
- [424] M. Yang, A.D. MacKerell Jr., *J. Chem. Theory Comput.* **2015**, *11*, 788: Conformational sampling of oligosaccharides using Hamiltonian replica exchange with two-dimensional dihedral biasing potentials and the weighted histogram analysis method (WHAM).
- [425] M. Yang, J. Huang, A.D. MacKerell Jr., *J. Chem. Theory Comput.* **2015**, *11*, 2855: Enhanced conformational sampling using replica exchange with concurrent solute scaling and Hamiltonian biasing realized in one dimension.

- [426] P. Liu, B. Kim, R.A. Friesner, B.J. Berne, *Proc. Natl. Acad. Sci.* **2005**, *102*, 13749: Replica exchange with solute tempering: A method for sampling biological systems in explicit water.
- [427] X. Huang, M. Hagen, B. Kim, R.A. Friesner, R. Zhou, B.J. Berne, *J. Phys. Chem. B* **2007**, *111*, 5405: Replica exchange with solute tempering: Efficiency in large scale systems.
- [428] L. Wang, R.A. Friesner, B.J. Berne, *J. Phys. Chem. B* **2011**, *115*, 9431: Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2).
- [429] L. Wang, R.A. Friesner, B.J. Berne, *J. Phys. Chem. B* **2011**, *115*, 11305: Erratum to “Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2)” [J. Phys. Chem. B 115, 9431–9438 (2011)].
- [430] J.R. Tame, G.N. Murshudov, E.J. Dodson, T.K. Neil, G.G. Dodson, C.F. Higgins, A.J. Wilkinson, *Science* **1994**, *264*, 1578: The structural basis of sequence-independent peptide binding by OppA protein.
- [431] J.R. Tame, E.J. Dodson, G.N. Murshudov, C.F. Higgins, A.J. Wilkinson, *Structure* **1995**, *3*, 1395: The crystal structures of the oligopeptide-binding protein OppA complexed with tripeptide and tetrapeptide ligand.
- [432] T.G. Davies, R.E. Hubbard, J.R. Tame, *Protein Sci.* **1999**, *8*, 1432: Relating structure to thermodynamics: the crystal structures and binding affinity of eight OppA-peptide complexes.
- [433] S.H. Sleigh, P.R. Seavers, A.J. Wilkinson, J.E. Ladbury, J.R.H Tame, *J. Mol. Biol.* **1999**, *291*, 393: Crystallographic and calorimetric analysis of peptide binding to OppA protein.
- [434] A. Warshel, *Acc. Chem. Res.* **2002**, *35*, 385: Molecular dynamics simulations of biological reactions.
- [435] M. Maurer, S.B.A. de Beer, C. Oostenbrink, *Molecules* **2016**, *21*, 499/1: Calculation of relative binding free energy in the water-filled active site of oligopeptide-binding protein A.
- [436] M. Serpi, V. Ferrari, F. Pertusati, *J. Med. Chem.* **2016**, *59*, 10343: Nucleoside derived antibiotics to fight microbial drug resistance: New utilities for an established class of drugs.
- [437] T. Läppchen, A.F. Hartog, V.A. Pinas, G.-J. Koomen, T. den Blaauwen, *Biochemistry* **2005**, *44*, 7879: GTP analogue inhibits polymerization and GTPase activity of the bacterial protein FtsZ without affecting its eukaryotic homologue tubuli.
- [438] T. Läppchen, V.A. Pinas, A.F. Hartog, G.-J. Koomen, C. Andreu Schaffner-Barbero, Trambaiolo J.M., Löwe D., Juhem J., Popov A., T. A.V. den Blaauwen, *Chem. & Biol.* **2008**, *15*, 189: Probing FtsZ and tubulin with C8-substituted GTP analogs reveals differences in their nucleotide binding sites.
- [439] F. Marcelo, S. Huecas, L.B. Ruiz-Ávila, F.J. Cañada, A. Perona, A. Poveda, S. Martín-Santamaría, A. Morreale, J. Jimínez-Barbero, J.M. Andreu, *J. Am. Chem. Soc.* **2013**, *135*, 16418: Interactions of bacterial cell division protein FtsZ with C8-substituted guanine nucleotide inhibitors. A combined NMR, biochemical and molecular modeling perspective.

- [440] R. Stolarski, C.-E. Hagberg, D. Shugar, *Eur. J. Biochem.* **1984**, *138*, 187: Studies on the dynamic syn-anti equilibrium in purine nucleosides and nucleotides with the aid of  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectroscop.
- [441] I.N. Bronshtein, L.A. Semendyayev, G. Musiol, H. Mühlig, *Handbook of Mathematics Edition 5.*; Springer Berlin Heidelberg, **2007**.
- [442] R.D. Lins, P.H. Hünenberger, *J. Comput. Chem.* **2005**, *26*, 1400: A new GROMOS force field for hexopyranose-based carbohydrates.
- [443] J.W. Steed, J.L. Atwood, *Supramolecular Chemistry. Edition 2.*; John Wiley & Sons, Chichester, UK, **2009**.
- [444] D. Kern, E.R.P. Zuidervaeg, *Curr. Opin. Struct. Biol.* **2003**, *13*, 748: The role of dynamics in allosteric regulation.
- [445] J.D. McGeagh, K.E. Ranaghan, A.J. Mulholland, *Biochim. Biophys. Acta* **2011**, *1814*, 1077: Protein dynamics and enzyme catalysis: Insights from simulation.
- [446] G.A. Somorjai, Y. Li, *Introduction to surface chemistry and catalysis.*; John Wiley & Sons, Hoboken, New Jersey, **2010**.
- [447] R.S. Bohacek, C. McMurtin, W.C. Guida, *Med. Res. Rev.* **1996**, *16*, 3: The art and practice of structure-based drug design: A molecular modeling perspective.
- [448] N. Bodor, P. Buchwald, *Med. Res. Rev.* **2000**, *20*, 58: Soft drug design: General principles and recent application.
- [449] W.F. van Gunsteren, H.J.C. Berendsen, *Angew. Chem. Int. Ed.* **1990**, *29*, 992: Computer simulation of molecular dynamics: Methodology, applications and perspectives in chemistry.
- [450] 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*, 4064: Biomolecular modelling: goals, problems, perspectives.
- [451] M.K. Gilson, J.A. Given, B.L. Bush, J.A. McCammon, *Biophys. J.* **1997**, *72*, 1047: The statistical thermodynamic basis for computation of binding-affinities. A critical review.
- [452] S. Shobana, B. Roux, O.S. Andersen, *J. Phys. Chem. B* **2000**, *104*, 5179: Free energy simulations: Thermodynamic reversibility and variability.
- [453] D.H. de Jong, L. Schäfer, A.H. de Vries, S.J. Marrink, H.J.C. Berendsen, H. Grubmüller, *J. Comput. Chem.* **2011**, *32*, 1919: Determining equilibrium constants for dimerization reactions from molecular dynamics simulations.
- [454] J. van Eerden, W.J. Briels, S. Harkema, D. Feil, *Chem. Phys. Lett.* **1989**, *164*, 370: Potential of mean force by thermodynamic integration: Molecular-dynamics simulation of decomplexation.
- [455] J.P. Valleau, S.G. Whittington, A guide to Monte Carlo for statistical mechanics: 2. Byways. In: *Modern theoretical chemistry. Volume 5.*; B.J. Berne, Ed.; Plenum Press, New York, USA, **1977**; pp 169-194.

- [456] E. Gallicchio, M. Andrec, A.K. Felts, R.M. Levy, *J. Phys. Chem. B* **2005**, *109*, 6722: Temperature weighted histogram analysis method, replica exchange, and transition paths.
- [457] A.M. Ferrenberg, R.H. Swendsen, *Phys. Rev. Lett.* **1988**, *61*, 2635: New Monte Carlo technique for studying phase transitions.
- [458] M. Feig, B.M. Pettitt, *Biopolymers* **1998**, *48*, 199: A molecular simulation picture of DNA hydration around A- and B-DNA.
- [459] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, J.M. Rosenberg, *J. Comput. Chem.* **1992**, *13*, 1011: The weighted histogram analysis method for free-energy calculations on biomolecules. I. The Method.
- [460] B. Roux, *Comput. Phys. Commun.* **1995**, *91*, 275: The calculation of the potential of mean force using computer simulation.
- [461] S. Sakuraba, A. Kitao, *J. Comput. Chem.* **2009**, *30*, 1850–1858: Multiple Markov transition matrix method: Obtaining the stationary probability distribution from multiple simulations.
- [462] H. Wu, F. No'e, *Multiscale Model. Simul.* **2014**, *12*, 25–54: Optimal estimation of free energies and stationary densities from multiple biased simulations.
- [463] H. Wu, A.S.J.S. Mey, E. Rosta, F. Noé, *J. Chem. Phys.* **2014**, *141*, 214106/1: Statistically optimal analysis of state-discretized trajectory data from multiple thermodynamic states.
- [464] A.S.J.S. Mey, H. Wu, F. Noé, *Phys. Rev. X* **2014**, *4*, 041018/1: xTRAM: Estimating equilibrium expectations from time-correlated simulation data at multiple thermodynamic states.
- [465] H. Wu, F. Paul, C. Wehmeyer, F. Noé, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E3221: Multiensemble Markov models of molecular thermodynamics and kinetics.
- [466] E. Rosta, G. Hummer, *J. Chem. Theory Comput.* **2015**, *11*, 276: Free energies from dynamic weighted histogram analysis using unbiased Markov state model.
- [467] L.S. Stelzl, A. Kells, E. Rosta, G. Hummer, *J. Chem. Theory Comput.* **2017**, *13*, 6328: Dynamic histogram analysis to determine free energies and rates from biased simulations.
- [468] S. Boresch, M. Karplus, *J. Chem. Phys.* **1996**, *105*, 5145: The Jacobian factor in free energy simulations.
- [469] L. Wang, J. Hermans, *J. Chem. Phys.* **1994**, *100*, 9129: Change of bond lengths in free-energy simulations: Algorithmic improvements, but when is it necessary.
- [470] J. Wang, Y. Deng, B. Roux, *Biophys. J.* **2006**, *91*, 2798: Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials.
- [471] S. Boresch, F. Tettinger, M. Leitgeb, M. Karplus, *J. Phys. Chem. B* **2003**, *107*, 9535: Absolute binding free energies: A quantitative approach for their calculation.

- [472] Y. Deng, B. Roux, *J. Chem. Theory Comput.* **2006**, *2*, 1255: Calculation of standard binding free energies: Aromatic molecules in the T4 lysozyme L99A mutant.
- [473] G.J. Rocklin, S.E. Boyce, M. Fischer, I. Fish, D.L. Mobley, B.K. Stoichet, K.A. Dill, *J. Mol. Biol.* **2013**, *425*, 4569: Blind prediction of charged ligand binding affinities in a model binding site.
- [474] G.J. Rocklin, D.L. Mobley, K.A. Dill, P.H. Hünenberger, *J. Chem. Phys.* **2013**, *139*, 184103/1: Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects.
- [475] F. Zhu, G. Hummer, *J. Comput. Chem.* **2012**, *33*, 453: Convergence and error estimation in free energy calculations using the weighted histogram analysis method.
- [476] D.F. Hahn, G. König, P.H. Hünenberger, *J. Chem. Theory Comput.* **2019**, *XXX*, XXX: Overcoming orthogonal barriers in alchemical free-energy calculations: On the relative merits of  $\lambda$ -variations,  $\lambda$ -extrapolations, and biasing.
- [477] H. Kokubo, T. Tanaka, Y. Okamoto, *J. Comput. Chem.* **2011**, *32*, 2810: *Ab initio* prediction of protein-ligand binding structures by replica-exchange umbrella sampling simulations.
- [478] H. Jiang, L. Maragliano, B. Roux, *J. Chem. Theory Comput.* **2012**, *8*, 4672: Calculation of free energy landscape in multi-dimensions with Hamiltonian exchange umbrella sampling on petascale supercomputer.
- [479] D.S. Dashti, A.E. Roitberg, *J. Chem. Theory Comput.* **2013**, *9*, 4692: Optimization of umbrella sampling replica exchange molecular dynamics by replica positioning.
- [480] S. Park, W. Im, *J. Chem. Theory Comput.* **2013**, *9*, 13: Two dimensional window exchange umbrella sampling for transmembrane helix assembly.
- [481] C.J. Pedersen, *J. Am. Chem. Soc.* **1967**, *89*, 7017: Cyclic polyether and their complexes with metal salts.
- [482] G.W. Gokel, D.M. Goli, C. Minganti, L. Contribution from Echegoyen, *J. Am. Chem. Soc.* **1983**, *Soc.*, 105: Clarification of the hole-size cation-diameter relationship in crown ethers and a new method for determining calcium cation homogeneous equilibrium binding constants.
- [483] G.D. Beresford, J.F. Stoddart, *Tetrahedron Lett.* **1980**, *21*, 867: The synthesis and complexing properties of oxo-12-crown-3 and 18-crown-5.
- [484] J.D. Lamb, R.M. Izatt, C.S. Swain, J.J. Christensen, *J. Am. Chem. Soc.* **1980**, *Soc.*, 102: A systematic study of the effect macrocycle ring size and donor atom type on the  $\log K$ ,  $\Delta H$ , and  $T\Delta S$  of reactions at  $25^\circ\text{C}$  in methanol of mono- and divalent cations with crown ethers.
- [485] A.J. Smetana, A.I. Popov, *J. Sol. Chem.* **1980**, *8*, 183: Lithium-7 nuclear magnetic resonance and calorimetric study of lithium crown complexes in various solvents.
- [486] F. Arnaud-Neu, R. Delgado, S. Chaves, *Pure Appl. Chem.* **2003**, *75*, 71: Critical evaluation of stability constants and thermodynamic functions of metal complexes of crown ethers.

- [487] R.M. Izatt, R.E. Terry, B.L. Haymore, L.D. Hansen, N.K. Dalley, A.G. Ayondet, J.J. Christensen, *Chem. Rev.* **1976**, *85*, 271: Calorimetric titration study of the interaction of several uni- and bivalent cations with 15-crown-5, 18-crown-6, and two isomers of dicyclohexo-18-crown-6 in aqueous solution at 25°C and  $\mu = 0.1$ .
- [488] H. Hoiland, J.A. Ringseth, T.S. Brun, *J. Sol.* **1979**, *Chem.*, **8**: Cation-crown ether complex formation in water. II. Alkali and alkaline earth cations and 12-crown-4, 15-crown-5, and 18-crown-6.
- [489] M.A. Thompson, E.D. Glendening, D. Feller, *J. Phys. Chem.* **1994**, *98*, 10465: The nature of  $K^+$ /crown ether interactions: A hybrid quantum mechanical-molecular mechanical study.
- [490] L.X. Dang, P.A. Kollman, *J. Phys. Chem.* **1995**, *99*, 55: Free energy of association of the  $K^+$ :18-crown-6 complex in water: A new molecular dynamics study.
- [491] L.X. Dang, *J. Am. Chem. Soc.* **1995**, *117*, 6954: Mechanism and thermodynamics of ion selectivity in aqueous solutions of 18-crown-6 ether: A molecular-dynamics study.
- [492] B. Martínez-Haya, P. Hurtado, A.R. Hortal, S. Hamad, J.D. Steill, J. Oomens, *J. Phys. Chem. A* **2010**, *114*, 7048: Emergence of symmetry and chirality in crown ether complexes with alkali metal cations.
- [493] B. Cooke, S.C. Schmidler, *J. Chem. Phys.* **2008**, *129*, 164112/1: Preserving the Boltzmann ensemble in replica-exchange molecular dynamics.
- [494] I. Fukuda, *J. Chem. Phys.* **2010**, *132*, 127101/1: Comment on “Preserving the Boltzmann ensemble in replica-exchange molecular dynamics” [*J. Chem. Phys.* **129** (2008) 164112].
- [495] E. Rosta, N.-V. Buchete, G. Hummer, *J. Chem. Theory Comput.* **2009**, *5*, 1393: Thermostat artifacts in replica exchange molecular dynamics simulations.
- [496] Z. Lin, W.F. van Gunsteren, *J. Chem. Phys.* **2015**, *143*, 034110/1: On the use of a weak-coupling thermostat in replica-exchange molecular dynamics simulations.
- [497] J.E. Basner, C. Jarzynski, *J. Phys. Chem. B* **2008**, *112*, 12722: Binless estimation of the potential of mean force.
- [498] B.W. Zhang, J. Xia, Z. Tan, R.M. Levy, *J. Phys. Chem. Lett.* **2015**, *6*, 3834: A stochastic solution to the unbinned WHAM equations.
- [499] Z. Tan, J. Xia, B.W. Zhang, R.M. Levy, *J. Chem. Phys.* **2016**, *144*, 034107/1: Locally weighted histogram analysis and stochastic solution for large-scale multi-state free energy estimation.
- [500] M.A. Kastenholz, T.U. Schwartz, P.H. Hünenberger, *Biophys. J.* **2006**, *91*, 2976: The transition between the B and Z conformations of DNA investigated by targeted molecular dynamics simulations with explicit solvation.
- [501] T.-S. Lee, B.K. Radak, A. Pabis, D.M. York, *J. Chem. Theory Comput.* **2013**, *9*, 153: A new maximum likelihood approach for free energy profile construction from molecular simulations.

- [502] T.-S. Lee, B.K. Radak, M. Huang, K.-Y. Wong, D.M. York, *J. Chem. Theory Comput.* **2014**, *10*, 24: Roadmaps through free energy landscapes calculated using the multidimensional vFEP approach.
- [503] J. Kästner, *WIREs Comput. Mol. Sci.* **2011**, *1*, 932: Umbrella sampling.
- [504] I. General, *J. Chem. Theory Comput.* **2010**, *6*, 2520: A note on the standard state's binding free energy.
- [505] P. Procacci, R. Chelli, *J. Chem. Theory Comput.* **2017**, *13*, 1924: Statistical mechanics of ligand-receptor noncovalent association, revisited: Binding site and standard state volumes in modern alchemical theories.
- [506] P.F.J. Fuchs, H.S. Hansen, P.H. Hünenberger, B.A.C. Horta, *J. Chem. Theory Comput.* **2012**, *8*, 3943: A GROMOS parameter set for vicinal diether functions: properties polyethylenoxide and polyethyleneglycol.
- [507] M.M. Reif, P.H. Hünenberger, *J. Chem. Phys.* **2011**, *134*, 144104/1: Computation of methodology-independent single-ion solvation properties from molecular simulations. IV. Optimized Lennard-Jones parameter sets for the alkali and halide ions in water.
- [508] P.H. Hünenberger, M.M. Reif, *Single-ion solvation: Experimental and theoretical approaches to elusive thermodynamic quantities. Edition 1, ISBN: 978-1-84755-187-0.*; Royal Society of Chemistry (Theoretical and Computational Chemistry Series), London, UK, **2011**.
- [509] T.S. Hofer, P.H. Hünenberger, *J. Chem. Phys.* **2018**, *148*, 222814/1: Absolute proton hydration free energy, surface potential of water, and redox potential of the hydrogen electrode from first principles: QM/MM MD free-energy simulations of sodium and potassium hydration.
- [510] N. Prasetyo, P.H. Hünenberger, T. Hofer, *J. Chem. Theory Comput.* **2018**, *14*, 6443: Single-ion thermodynamics from first principles: Calculation of the absolute hydration free energy and single-electrode potential of aqueous Li<sup>+</sup> using *ab initio* quantum mechanical/molecular mechanical molecular dynamics simulations.
- [511] D.P. Geerke, C. Oostenbrink, N.F.A. van der Vegt, W.F. van Gunsteren, *J. Phys. Chem. B* **2004**, *108*, 1436: An effective force field for molecular dynamics simulations of dimethyl sulfoxide and dimethyl sulfoxide-water mixtures.
- [512] R. Walser, A.E. Mark, W.F. van Gunsteren, M. Lauterbach, G. Wipff, *J. Chem. Phys.* **2000**, *112*, 10450: The effect of force-field parameters on properties of liquids: Parametrization of a simple three-site model for methanol.
- [513] T. Todorova, P.H. Hünenberger, J. Hutter, *J. Chem. Theory Comput.* **2008**, *4*, 779: Car-Parrinello molecular dynamics simulations of CaCl<sub>2</sub> aqueous solutions.
- [514] M.M. Reif, C. Oostenbrink, *J. Comput. Chem.* **2014**, *35*, 227: Net charge changes in the calculation of relative ligand-binding free energies *via* classical atomistic molecular dynamics simulations.
- [515] D.R. Lide, *CRC Handbook of Chemistry and Physics. Edition 80.*; CRC Press, Boca Raton, Florida, **1999**.
- [516] M. Rigby, E.B. Smith, W.A. Wakeham, G.C. Maitland, *The forces between molecules.*; Clarendon Press, Oxford, UK, **1986**.

- [517] M.S. Green, *J. Chem. Phys.* **1954**, *22*, 398: Markoff random processes and the statistical mechanics of time-dependent phenomena II: Irreversible processes in fluids.
- [518] R. Kubo, *J. Phys. Soc. Jpn.* **1957**, *12*, 570: Statistical-mechanical theory of irreversible processes I: General theory and simple applications to magnetic and conduction problems.
- [519] O. Nagano, *Acta Cryst. B* **1979**, *35*, 465: Structure of the potassium molybdate complex of 1,4,7,10,13,16-hexaoxacyclooctadecane (18-crown-6).
- [520] D. Petrov, C. Margreitter, M. Grandits, C. Oostenbrink, B. Zagrovic, *PLOS Comput. Biol.* **2013**, *9*, e1003154/1: A systematic framework for molecular dynamics simulations of protein post-translational modifications.
- [521] A. Renevey, S. Riniker, *Eur. Biophys. J.* **2017**, *25*, 363: The importance of N-methylations for the stability of the  $\beta^{6.3}$ -helical conformation of polytheonamide B.
- [522] A.K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P.C. Nair, C. Oostenbrink, A.E. Mark, *J. Chem. Theory Comput.* **2011**, *7*, 4026: An automated force field topology builder (ATB) and repository: version 1.0.
- [523] J. Kästner, H.M. Senn, S. Thiel, N. Otte, W. Thiel, *J. Chem. Theory Comput.* **2006**, *2*, 452: QM/MM free-energy perturbation compared to thermodynamic integration and umbrella sampling: Application to an enzymatic reaction.

# Curriculum Vitæ

BENJAMIN JOACHIM RIES

15.03.1991  
Ettlingen, Germany  
German citizen

## EDUCATION

- 2017 – 2021 PhD, ETH Zürich
- 2012 – 2014 Bioinformatics MSc., Universität Tübingen
- 2009 – 2013 Biochemistry BSc., Universität Tübingen
- 2002 – 2011 secondary school examinations (Abitur), Albertus Magnus Gymnasium, Ettlingen, Germany

## EXPERIENCE

- 2016 – 2017 Summerschool 2016, University of Jyväskylä, Finland
- 2016 – 2017 Erasmus+ internship, Uppsala Universitet, Sweden
- ?? Lab Assistant, MPI, Developmental Biology, Tübingen