Biomolecular modeling

Marcus Elstner and Tomáš Kubař

Theoretical Chemistry, TU Braunschweig

(Dated: December 10, 2010)

IX. FREE ENERGY SIMULATIONS

When searching for a physical quantity that is of most interest in chemistry, we could hardly find anything more appropriate than free energies – Helmholtz F or Gibbs G. Truly, these represent the holy grail of computational chemistry, both for their importance and because they are difficult to calculate.

These difficulties were hinted at in one of previous chapters. Recall that we can write

$$F = k_\mathrm{B}T \ln \iint \exp[\beta E(\vec{r},\vec{p})] \cdot \rho(\vec{r},\vec{p})\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p} + c \tag{IX.1}$$

The problem is that the large energy values (far from the minimum of energy) enter an exponential term, so that these high-energy regions may contribute significantly to the free energy F. So, in a simulation, if we have too few points in these high-energy regions of the phase space (*undersampling*), we may introduce sizeable errors in the calculated averages.

There are two fundamental approaches to overcome this difficulty: *free energy perturbation* and *thermodynamic integration*. Also, several computational tricks may be used for particular types of reactions, like *alchemical simulations* or *umbrella sampling*. An important observation is that it is not necessary to find the absolute value of the free energy. When considering a chemical reaction,[1] it is important to know merely the *free energy difference* (ΔF, ΔG) between the involved states (reactant A and product B).

A. Free energy perturbation (FEP)

For these states with energies $E_A(\vec{r},\vec{p})$ and $E_B(\vec{r},\vec{p})$, and partition functions Q_A and Q_B, free energy difference may be derived as

$$\begin{aligned}
\Delta F &= F_B - F_A = -k_\mathrm{B}T \ln \frac{Q_B}{Q_A} = -k_\mathrm{B}T \ln \frac{\iint \exp[-\beta E_B]\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p}}{\iint \exp[-\beta E_A]\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p}} \\
&= -k_\mathrm{B}T \ln \frac{\iint \exp[-\beta E_B]\ \exp[\beta E_A]\exp[-\beta E_A]\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p}}{\iint \exp[-\beta E_A]\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p}} \\
&= -k_\mathrm{B}T \ln \iint \exp[-\beta E_B]\exp[\beta E_A] \cdot \rho_A(\vec{r},\vec{p})\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p} \\
&= -k_\mathrm{B}T \ln \iint \exp[-\beta (E_B - E_A)] \cdot \rho_A(\vec{r},\vec{p})\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p}
\end{aligned} \tag{IX.2}$$

[1] in a very general sense of a reaction that need not involve chemical bonds being created or broken – ligand binding a protein, passage of a molecule through membrane, or protein folding are reactions as well

The integral has the form of an average of a property S taken with the phase space density of state A

$$\langle S \rangle_A = \iint S(\vec{r}, \vec{p}) \cdot \rho_A(\vec{r}, \vec{p}) \, \mathrm{d}\vec{r} \, \mathrm{d}\vec{p} \tag{IX.3}$$

and so we can write equivalently

$$\begin{aligned} \Delta F(A \to B) &= -k_\mathrm{B} T \ln \langle \exp[-\beta(E_B - E_A)] \rangle_A \\ \Delta F(B \to A) &= -k_\mathrm{B} T \ln \langle \exp[-\beta(E_A - E_B)] \rangle_B \end{aligned} \tag{IX.4}$$

which is the free energy formula by Zwanzig (1954) and the essence of the FEP method. Thus, in principle, it is possible to perform a simulation of state A and obtain the free energy by averaging the exponential of the difference of energies of states B and A, or vice versa. Practically, we start an MD in state A to get the phase space density ρ_A, and then calculate the difference between the energies of states B and A along the trajectory.

- Free energy of deprotonation of an amino acid side chain in a protein. We would run the dynamics for the protonated species, and then evaluate the energy difference between protonated and unprotonated species to get the average of $\exp[-\beta(E_B - E_A)]$. This would only work if the conformations of the protein, and above all the configuration of water molecules, sampled along the MD were very similar with both forms. Usually, this is not the case.

- The ionization of a molecule. Again, we would perform a simulation of the neutral species and evaluate the energy differences. Alas, the configuration of water would be quite different here, too, leading to a very small overlap of phase space densities.

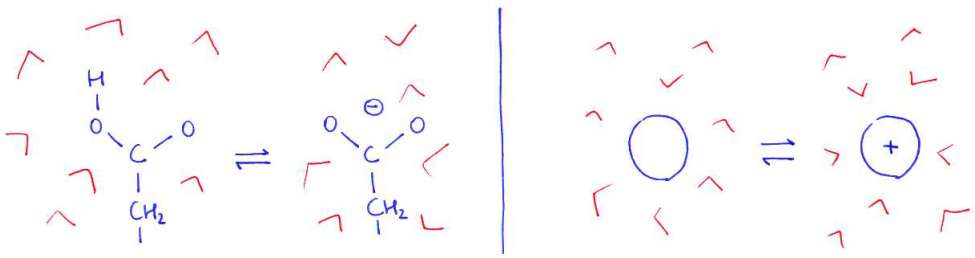

FIG. 1: Deprotonation of an amino acid (left) and ionization of a molecule (right), both in water.

Once again, let us emphasize the advantage of FEP over the direct evaluation of free energies. In the latter case, two simulations would be performed, one for each state A and

B, and the free energy difference would follow (using Eq. IX.1) as

$$\Delta F(A \to B) = k_\mathrm{B}T \ln \langle \exp[\beta E_B] \rangle_B - k_\mathrm{B}T \ln \langle \exp[\beta E_A] \rangle_A \tag{IX.5}$$

Here, note that the free energy difference is very small, of a few kcal/mol, while the total energies are very large, of hundreds or thousands kcal/mol, if the solvent or the like is included. So, we have to subtract two large numbers in order to get a small one. However, a small relative uncertainty (error) of the large values would be huge in comparison with the possibly small resulting free energy difference. Therefore, it is necessary to obtain these large values extremely accurate, which would mean the necessity to perform exceedingly long MD simulations – so long that we will never be able to afford it!

That is why we avoid performing individual simulations for the end states and rather evaluate the free energy difference directly in one simulation. Then, it is no longer necessary to sample the regions of the molecular system which do not change and are not in contact with the regions that are changing, as these do not contribute to the energy difference $E_B - E_A$. The region of phase space that has to be sampled thoroughly is much smaller, and the necessary simulation length may become feasible.

For the following, the concept of *overlap in phase space* or *overlap of phase space densities* is crucial. In a straightforward way, the more similar the states A and B are, the more similar are also the corresponding phase space densities, and they may exhibit an overlap, see Fig. 2. If the phase space densities for states A and B are similar (overlapping, Fig. 2 right), then

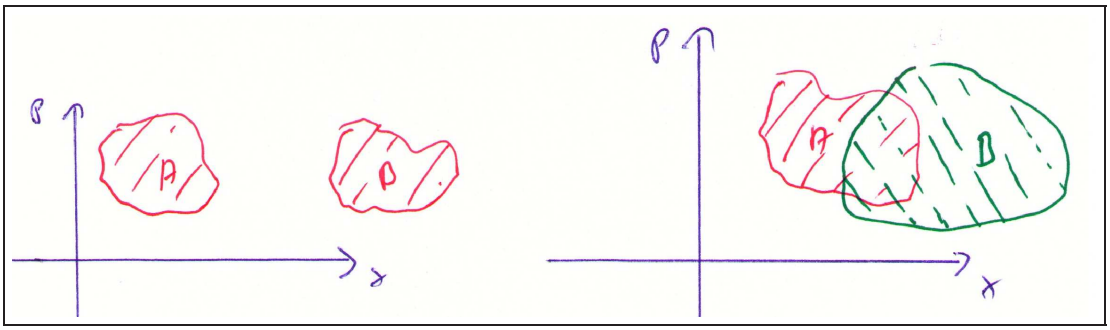

FIG. 2: Large (right) and no (left) overlap of phase space densities corresponding to two states.

the low-energy regions of state B may be sampled well even in the simulation of state A, and the free energy difference $\Delta F(A \to B)$ in Eq. IX.4 may converge. If this is not the case (like in Fig. 2 left), then the simulation of state A hardly comes to the region of phase space

where the state B has low energy; this region is undersampled, the averaging of the energy E_B is wrong, and the calculation will not converge. As a rule of thumb, this is the case if

$$|E_B - E_A| > k_\mathrm{B}T \tag{IX.6}$$

A way to overcome this problem is to insert an intermediate state (designated '1') which overlaps with both A and B, as in Fig. 3. The underlying idea is to make use of the fact

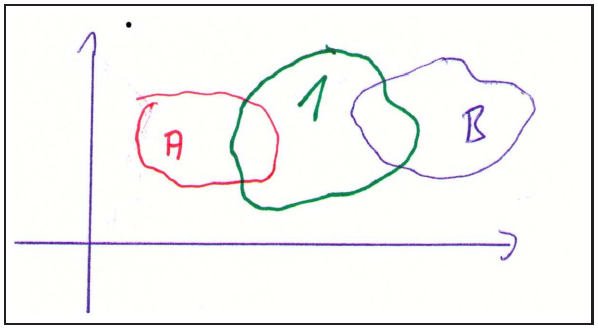

FIG. 3: Intermediate state '1' overlapping with state A and B

that free energy is a *state function*, and so

$$\Delta F(A \to B) = \Delta F(A \to 1) + \Delta F(1 \to B) \tag{IX.7}$$

Therefore, we can perform *two* MD simulations, one for each of the states A and 1, and evaluate free energies for the two subprocesses. These may be expected to converge better, and their sum gives the free energy of $A \to B$:

$$\begin{aligned} \Delta F &= -k_\mathrm{B}T \ln \left[\frac{Q_1}{Q_A} \cdot \frac{Q_B}{Q_1} \right] = \\ &= -k_\mathrm{B}T \ln \left\langle \exp[-\beta(E_1 - E_A)] \right\rangle_A - k_\mathrm{B}T \ln \left\langle \exp[-\beta(E_B - E_1)] \right\rangle_1 \end{aligned} \tag{IX.8}$$

Obviously, it is possible to insert more than one intermediate state between A and B, if these differ exceedingly. For N intermediate states $1, 2, \ldots, N$, we obtain

$$\begin{aligned} \Delta F = &- k_\mathrm{B}T \ln \left[\frac{Q_1}{Q_A} \cdot \frac{Q_2}{Q_1} \cdot \ldots \cdot \frac{Q_B}{Q_N} \right] = \\ = &- k_\mathrm{B}T \ln \left\langle \exp[-\beta(E_1 - E_A)] \right\rangle_A - k_\mathrm{B}T \ln \left\langle \exp[-\beta(E_2 - E_1)] \right\rangle_1 - \\ &- \ldots - k_\mathrm{B}T \ln \left\langle \exp[-\beta(E_B - E_N)] \right\rangle_N \end{aligned} \tag{IX.9}$$

and we have to perform $N + 1$ simulations, e.g. of states $A, 1, 2, \ldots, N$.

The description of this procedure may sound complicated, but it is implemented in the common simulation packages in a convenient way. Since we can change the chemical identities of the atoms or functional groups, this practice is often called *computational alchemy*. Typically, one introduces a parameter λ which 'converts' the force-field parameters (i.e. the Hamiltonian) from these of state A to those of state B:

$$E_\lambda = (1-\lambda)\cdot E_A + \lambda \cdot E_B \tag{IX.10}$$

- The (solvation) free energy difference of argon and xenon in aqueous solution. The two atoms differ only in the vdW parameters – the well depth ε and the radius σ. To transmutate the energy function from that of one species to the other, we interpolate:

$$\varepsilon_\lambda = (1-\lambda)\cdot \varepsilon_A + \lambda \cdot \varepsilon_B \tag{IX.11}$$

$$\sigma_\lambda = (1-\lambda)\cdot \sigma_A + \lambda \cdot \sigma_B \tag{IX.12}$$

  In the simulation, we start from $\lambda = 0$, i.e. an argon atom, and change it in subsequent steps to 1. For each step (called *window*), we perform an MD with the corresponding values of the vdW parameters, and calculate the relative free energies.

- A true chemical reaction like HCN $\rightarrow$ CNH. The situation is more complicated as we need the topologies of both molecules. Thus, a *dual-topology* simulation is performed: we have both molecules simultaneously in the simulation. These two molecules do *not* interact with each other, and we gradually switch off the interaction of one species with the solvent during the simulation while we switch on the other at the same time.

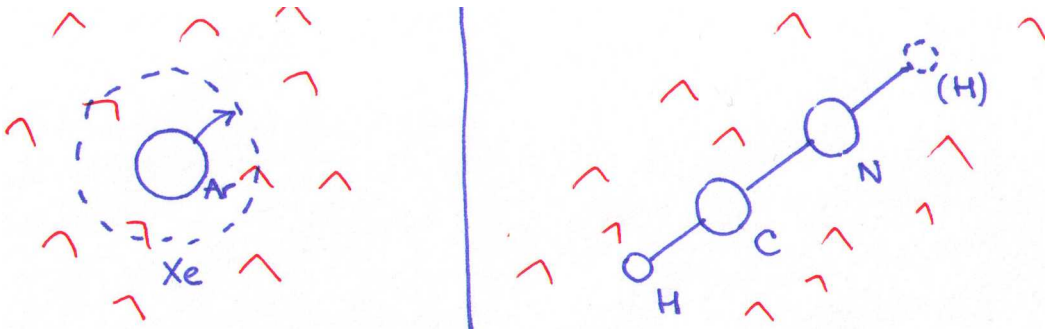

FIG. 4: Examples of 'alchemical' simulations.

B. Thermodynamic integration (TI)

In the last chapter, we have written the energy E as a function of the parameter λ. This means, that the free energy also becomes dependent on λ:

$$F = F(\lambda) \tag{IX.13}$$

with $F(0) = F(A)$ and $F(1) = F(B)$. Thus, we can write

$$\Delta F = F(B) - F(A) = \int_0^1 \frac{\partial F(\lambda)}{\partial \lambda} \mathrm{d}\lambda \tag{IX.14}$$

with

$$F(\lambda) = -k_\mathrm{B} T \ln Q(\lambda) \tag{IX.15}$$

The derivative of F rearranges to

$$\begin{aligned}
\frac{\partial F}{\partial \lambda}(\lambda) &= -k_\mathrm{B}T \frac{\partial \ln Q}{\partial \lambda}(\lambda) = -k_\mathrm{B}T \frac{1}{Q(\lambda)} \cdot \frac{\partial Q}{\partial \lambda}(\lambda) = -k_\mathrm{B}T \frac{1}{Q(\lambda)} \cdot \frac{\partial}{\partial \lambda} \iint \exp[-\beta E_\lambda] \, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p} = \\
&= -k_\mathrm{B}T \frac{1}{Q(\lambda)} \cdot \iint (-\beta) \frac{\partial E_\lambda}{\partial \lambda} \exp[-\beta E_\lambda] \, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p} = \\
&= -k_\mathrm{B}T \cdot (-\beta) \cdot \iint \frac{\partial E_\lambda}{\partial \lambda} \frac{\exp[-\beta E_\lambda]}{Q(\lambda)} \, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p} \\
&= 1 \cdot \iint \frac{\partial E_\lambda}{\partial \lambda} \rho_\lambda(\vec{r},\vec{p}) \, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p} = \left\langle \frac{\partial E_\lambda}{\partial \lambda} \right\rangle_\lambda
\end{aligned} \tag{IX.16}$$

This is the essence of TI – the derivative of free energy F with respect to the coupling parameter λ is calculated as the average of derivative of total MM energy E, which can be directly evaluated in the simulation. Then, the free energy difference follows simply as

$$\Delta F = \int_0^1 \left\langle \frac{\partial E_\lambda}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \tag{IX.17}$$

Practically, we perform a MD simulation for each chosen value of λ; it is usual to take equidistant values in the interval (0,1) like 0, 0.05,..., 0.95 and 1. Each of these simulations produces a value of $\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_\lambda$, so that we obtain the derivative of free energy in discrete points for $\lambda \in (0, 1)$. This function is then integrated numerically, and the result is the desired free energy difference ΔF.

An example of the TI simulation is shown in Fig. 5. An atom of rare gas (neon) is dissolved in water; in course of the NPT simulation, the van der Walls parameters of the

neon atom are being gradually switched off by means of the λ parameter, so that the atom is effectively disappearing. The derivative of total energy with respect to λ is evaluated for several (21) values of λ ranging from 0 to 1. Eq. IX.17 is then used to obtain the (Gibbs) free energy difference of the two states: (i) a neon atom in water, and (ii) no neon atom in water, i.e. outside of the solution in vacuo. Thus, the calculated free energy difference corresponds directly to the solvation free energy, a quantity which is of considerable value in chemistry.

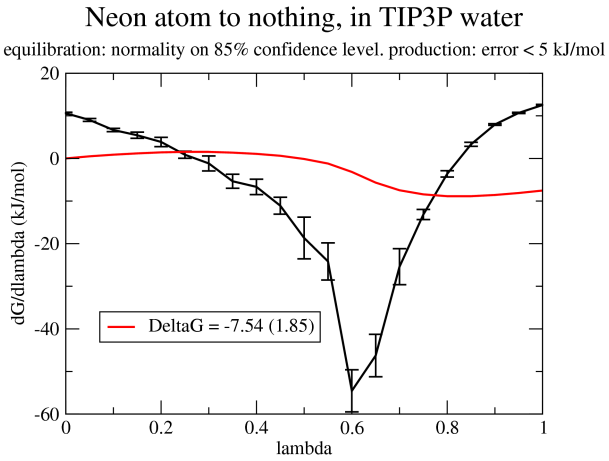

FIG. 5: TI simulation of a neon atom in water being disappeared. See text for explanation.

Finally, let us summarize the features of FEP and TI. Irrespective of the nature of the studied reaction, both FEP and TI require the introduction of a coupling parameter λ, which plays the role of the reaction coordinate with $\lambda = 0$ corresponding to the reactant and $\lambda = 1$ to the product. The fact that free energy is a state function guarantees the independence of the result on the chosen path between the reactant and the product, and so it does not matter if the reaction coordinate corresponds to an unphysical process like a change of chemical identity of one or more atoms (as is the case in the alchemical simulations).

The remaining open question regards the necessary number of windows in the simulation. We would like to have as few windows as possible, without compromising numerical precision of the calculation. In FEP, the assumption is that while simulating the state A, the low-energy regions of state B are sampled well. The closer the windows are, the better is this condition fulfilled. On the other hand, the free energy derivative is always evaluated for *one* λ-value with TI, and the problem present in FEP does not occur here. It is the numerical integration of the free energy derivative that brings on the numerical inaccuracy of TI.

C. Free energy from non-equilibrium simulations

A major disadvantage of the described methodology – TI using equilibrium simulations for discrete values of λ – is the very slow convergence of $\partial G/\partial\lambda$ once the alchemical change becomes large. So, it is often possible to describe the mutation of a single amino acid side chain in a protein provided the structure of the protein remains the same, but this should be considered a practical limit of the method.

To avoid this problem, the current development of free-energy methods makes use of *non-equilibrium simulations*. Here, the usual process of "equilibration" of the system for every of the selected values of λ followed by a "production phase" is not used; a non-equilibrium simulation consists of n MD steps, where the parameter λ starts at 0 and increases by $1/n$ in every MD step. This way, the simulation does not describe the system in equilibrium in any moment, as the external parameter λ is changing all the time. Whereas a single simulation of this kind is probably worthless, the remarkable equality by Jarzynski provides a link between an *ensemble* of such simulations and the desired free energy:

$$\exp[-\beta\Delta F] = \langle\exp[-\beta W]\rangle \tag{IX.18}$$

The true value of free energy ΔF is obtained as a special kind of ensemble average, for the ensemble of non-equilibrium TI simulations yielding "free energies" W. These values $W = \int_0^1 \partial E/\partial\lambda \,\mathrm{d}\lambda$ are no free energies whatsoever; instead, they may be called *(irreversible) work*. Since no convergence of any quantity is required within a single non-equilibrium simulation, these simulations may be very short – and this is the actual practice. However, the sampling problem persists because the largest statistical weight is carried by rarely occuring simulations (due to the unfavorable averaging in Eq. IX.18).

This sampling issue may be circumvented by *exponential work averaging* with gaussian approximation. An ensemble of simulations is performed for the 'forward' process $0 \to 1$ as well as for the 'reverse' process $1 \to 0$, and the obtained distributions of forward and backward irreversible work are approximated by gaussians with mean and standard deviation W_f, σ_f and W_r, σ_r, respectively. The free energy is calculated as an average of values

$$\begin{aligned} \Delta F_f &= W_f - \frac{1}{2}\beta\sigma_f^2 \\ \Delta F_r &= -W_r + \frac{1}{2}\beta\sigma_r^2 \end{aligned} \tag{IX.19}$$

A more general expression (than the Jarzynski equality) is the *Crooks fluctuation theorem* (CFS), according to which the distributions of forward and reverse work are related like

$$\frac{P_f(W)}{P_r(-W)} = \exp[\beta(W - \Delta F)] \tag{IX.20}$$

Then, once we have obtained well-converged distributions P_f and P_r, it is possible to apply *Bennett's acceptance ratio* for an equal number of forward and reverse simulation; the free energy follows from

$$\left\langle \frac{1}{1 + \exp[\beta(W - \Delta F)]} \right\rangle_f = \left\langle \frac{1}{1 + \exp[-\beta(W - \Delta F)]} \right\rangle_r \tag{IX.21}$$

It is possible to apply CFS more directly. A closer look at Eq. IX.20 reveals that the free energy corresponds to the value of work W for which the probabilities P_f and P_r are equal – to the intersection point of the distributions. To determine this point readily from the distributions may be difficult and a source of large errors if the overlap of the distributions is very small. Again, this issue can be solved by the assumption of normal distribution of the forward and reverse work, which was proven for a system with a large number of degrees of freedom. The procedure thus requires to perform a number of forward and reverse simulations sufficient to perform a good-quality gaussian fit to the resulting distributions of irreversible work. The free energy is calculated directly as the intersection points of these gaussian curves.

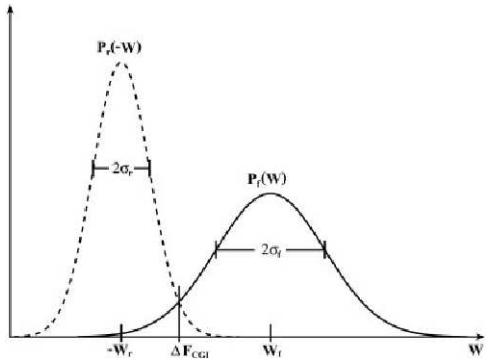

FIG. 6: The Crooks gaussian intersection (from Goette and Grubmüller 2009).

D. Thermodynamic cycles

Quite often, we are interested not in the absolute free energies and not even in the reaction free energies, but rather in the difference (Δ) of reaction free energies (ΔF) corresponding to two similar reactions. These may the be denoted as $\Delta\Delta F$ or $\Delta\Delta G$.

Consider as an example the binding of an inhibitor molecule I to an enzyme E, as shown in Fig. 7 left. Usually, we are interested in differences of binding free energies, for instance of an inhibitor I to two very similar enzymes E and E′:

$$\begin{aligned} \mathrm{E} + \mathrm{I} &\rightleftharpoons \mathrm{EI} \qquad \Delta G_1 \\ \mathrm{E}' + \mathrm{I} &\rightleftharpoons \mathrm{E}'\mathrm{I} \qquad \Delta G_2 \end{aligned} \tag{IX.22}$$

The binding of the inhibitor can induce large structural changes in the enzyme, and it would be very difficult (if not impossible) to describe this reaction in a simulation both correctly and efficiently at the same time. So, significant errors would seem to be inevitable. A way to solve this would be to simulate not the reaction of binding but rather the alchemical transmutation of enzyme E to E′. As we consider the enzymes to be very similar,[2] it is plausible to assume the structure of complexes EI and E′I to be similar as well. Then, the alchemical simulation may well be successful. As free energy is a state function, the sum of

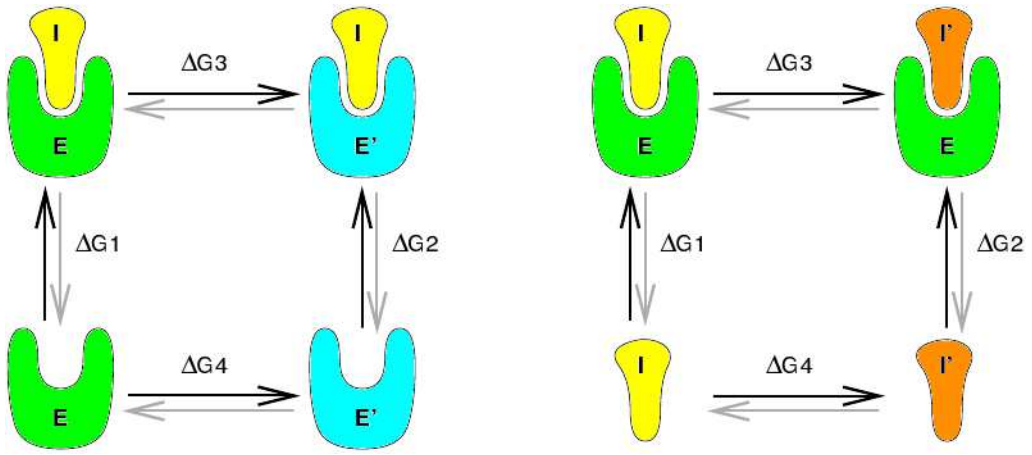

FIG. 7: Examples of the thermodynamic cycle.

free energies around a *thermodynamic cycle* vanishes (e.g. clockwise in Fig. 7 left):

$$\Delta F_1 + \Delta F_3 - \Delta F_2 - \Delta F_4 = 0 \tag{IX.23}$$

[2] Imagine E′ to be derived from E by a mutation of a single amino acid, e.g. leucine to valine.

The difference of binding free energies then follows to be equal the difference of free energies calculated in alchemical simulations:

$$\Delta\Delta F = \Delta F_1 - \Delta F_2 = \Delta F_3 - \Delta F_4 \tag{IX.24}$$

Similarly, it is possible to calculate the free energy difference of binding of two similar ligands to the same enzyme (Fig. 7 right), or the difference of solvation energy of two similar molecules. In the latter case, two alchemical simulations would be performed: one in vacuo and the other in solvent.

E. Potentials of mean force (PMF) and umbrella sampling

Sometimes, we wish to know not only the free energy difference of two states (the reactant and the product), but rather the free energy along the *reaction coordinate* q within a certain interval; the free energy is then a function of q while it is integrated over all other degrees of freedom. Such a free energy function $F(q)$ is called the *potential of mean force*. Examples of such a reaction coordinate q may be the distance between two particles if the dissociation of a complex is studied, the position of a proton for a reaction of proton transfer, or the dihedral angle when dealing with some conformational changes.

To separate the degree of freedom spanned by the reaction coordinate, we perform a coordinate transformation from $\vec{r} = (r_1, r_2, \ldots, r_{3N})$ to a set $(u_1, u_2, \ldots, u_{3N-1}, q)$, where the $(3N-1)$-dimensional vector $\vec{u}$ represents all remaining degrees of freedom, and we can write

$$\mathrm{d}\vec{r} = \mathrm{d}\vec{u} \cdot \mathrm{d}q \tag{IX.25}$$

Looking for the free energy at a certain value of q, all remaining degrees of freedom are averaged over (or 'integrated out'). One could think of performing an MD simulation and sampling *all* degrees of freedom except for q.

An example would be the free energy of formation of an ion pair in solution, as shown in Fig. 8. An MD simulation would be performed to calculate the free energy for every value of the reaction coordinate q.

The free energy is given by:

$$F = -k_\mathrm{B}T \ln \iint \exp[-\beta E(\vec{r}, \vec{p})]\, \mathrm{d}\vec{r}\, \mathrm{d}\vec{p} \tag{IX.26}$$

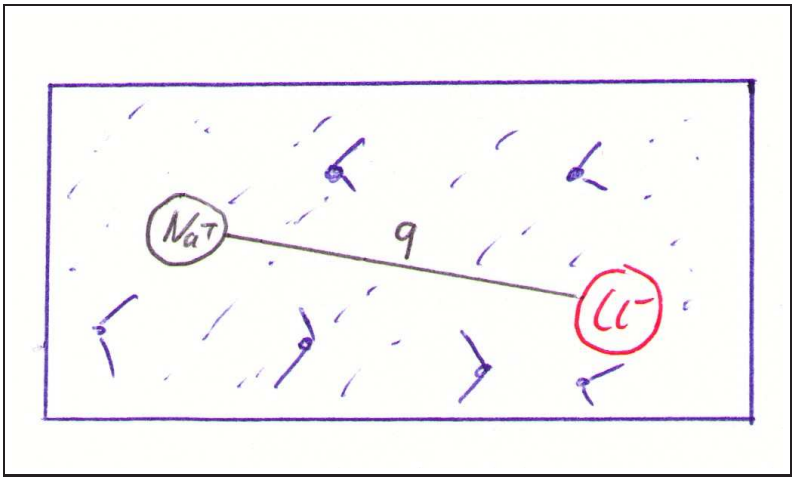

FIG. 8: Na^+ and Cl^- in water solution. The distance between the ions is the reaction coordinate q, and all other degrees of freedom (water) are represented by $\vec{u}$ and are free to vary.

If we wish to evaluate an expression for a coordinate q taking a certain value q_0, it is convenient to use the *Dirac delta function*,[3] $\delta(q - q_0)$. With that, we can write the free energy for the fixed reaction coordinate q_0 as

$$\begin{aligned} F(q_0) &= -k_\mathrm{B}T \ln \iint \delta(q - q_0) \exp[-\beta E(\vec{r}, \vec{p})] \, \mathrm{d}\vec{p} \, \mathrm{d}\vec{u} \, \mathrm{d}q \\ &= -k_\mathrm{B}T \ln \left[Q \cdot \iint \delta(q - q_0) \frac{\exp[-\beta E(\vec{r}, \vec{p})]}{Q} \, \mathrm{d}\vec{p} \, \mathrm{d}\vec{u} \, \mathrm{d}q \right] \\ &= -k_\mathrm{B}T \ln \left[Q \cdot \iint \delta(q - q_0) \cdot \rho(\vec{r}, \vec{p}) \, \mathrm{d}\vec{p} \, \mathrm{d}\vec{u} \, \mathrm{d}q \right] \\ &= -k_\mathrm{B}T \ln \left[Q \cdot \langle \delta(q - q_0) \rangle \right] \\ &= -k_\mathrm{B}T \ln Q - k_\mathrm{B}T \ln \langle \delta(q - q_0) \rangle \end{aligned} \tag{IX.27}$$

How to interpret this? $\rho(\vec{r}, \vec{p})$ is the probability, that the system is at the point $(\vec{r}, \vec{p})$. Then,

$$P(q_0) = \iint \delta(q - q_0) \cdot \rho(\vec{r}, \vec{p}) \, \mathrm{d}\vec{r} \, \mathrm{d}\vec{p} = \langle \delta(q - q_0) \rangle \tag{IX.28}$$

is the *probability* that the reaction coordinate q in the system takes the value of q_0 , because the integral proceeds over the whole phase space and the delta function 'cancels out' all points, where the reaction coordinate is *not equal* q_0! So, the integration collects all points in phase space, where the reaction coordinate has this specific value.

[3] This is a generalized function representing an infinitely sharp peak bounding unit area; $\delta(x)$ has the value of zero everywhere, except at $x = 0$ where its value is infinitely large in such a way that its integral is 1.

What would it work like in the example of the ion pair? We perform an MD simulation for the system, and then *count* how many times the reaction coordinate takes the specified value, in other words we calculate the probability $P(q_0)$ of finding the system at q_0.

Then, the free energy difference of two states A and B is:

$$\begin{aligned} F_B - F_A &= -k_\mathrm{B}T \ln Q - k_\mathrm{B}T \ln \langle \delta(q - q_B) \rangle - (-k_\mathrm{B}T \ln Q + k_\mathrm{B}T \ln \langle \delta(q - q_A) \rangle) \\ &= -k_\mathrm{B}T \ln \frac{\langle \delta(q - q_B) \rangle}{\langle \delta(q - q_A) \rangle} \\ &= -k_\mathrm{B}T \ln \frac{P(q_B)}{P(q_A)} \end{aligned} \tag{IX.29}$$

which is actually the known definition of the *equilibrium constant* $P(B)/P(A)$.

So, the task is clear: perform a MD, specify a coordinate, and then just count, how often the system is at special values of the reaction coordinate. The ratio of these numbers gives the free energy difference!

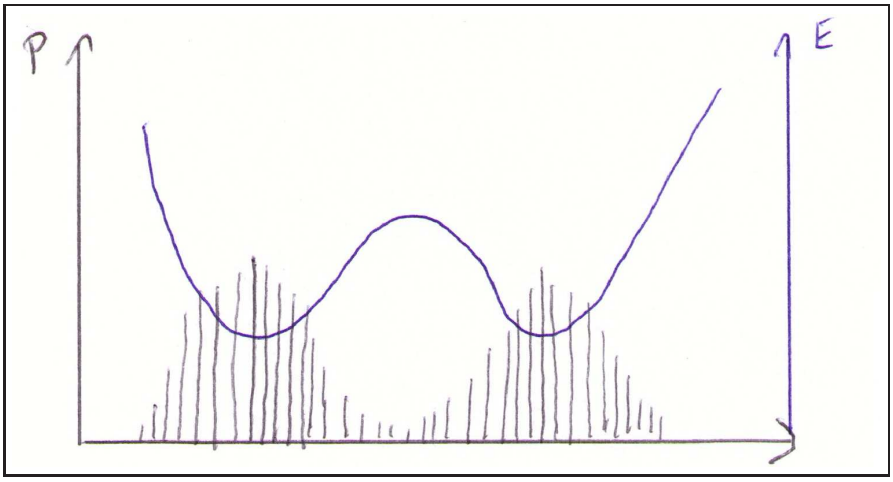

FIG. 9: Energy profile and probability distribution along the reaction coordinate. Note the undersampled region of the barrier.

This is very good, in principle. But, we also know the problem: If we there is a high barrier to be crossed along the reaction coordinate to come from A to B, a pure (*unbiased*) MD simulation will hardly make it,[4] and even if it does, the high-energy region (barrier) will be sampled quite poorly.

Then, a straightforward idea is to apply an *additional potential*, also called *biasing potential* in order to make the system spend a larger amount of time in that (those) region(s)

[4] In other words, the ergodicity of the simulation is hindered.

of phase space that would otherwise remain undersampled. This is the underlying principle of the *umbrella sampling*.[5] The additional potential shall depend only on the reaction coordinate: $V = V(q)$.[6] Then, what will the free energy look like in such a biased case? Let us start with the previously obtained expression:

$$\begin{aligned}
F(q_0) &= -k_\mathrm{B}T \ln \left[\frac{\iint \delta(q-q_0)\, \exp[-\beta E]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}}{\iint \exp[-\beta E]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}} \right] \\
&= -k_\mathrm{B}T \ln \left[\frac{\iint \delta(q-q_0) \exp[\beta V]\, \exp[-\beta(E+V)]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}}{\iint \exp[-\beta(E+V)]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}} \cdot \frac{\iint \exp[-\beta(E+V)]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}}{\iint \exp[-\beta E]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}} \right] \\
&= -k_\mathrm{B}T \ln \left[\langle \delta(q-q_0) \exp[\beta V] \rangle_{E+V} \frac{\iint \exp[-\beta(E+V)]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}}{\iint \exp[\beta V] \exp[-\beta(E+V)]\, \mathrm{d}\vec{r}\,\mathrm{d}\vec{p}} \right] \\
&= -k_\mathrm{B}T \ln \left[\langle \delta(q-q_0) \exp[\beta V] \rangle_{E+V} \frac{1}{\langle \exp[\beta V] \rangle_{E+V}} \right] \\
&= -k_\mathrm{B}T \ln \left[\exp[\beta V(q_0)]\, \langle \delta(q-q_0) \rangle_{E+V} \frac{1}{\langle \exp[\beta V] \rangle_{E+V}} \right] \\
&= -k_\mathrm{B}T \ln \langle \delta(q-q_0) \rangle_{E+V} - V(q_0) + k_\mathrm{B}T \ln \langle \exp[\beta V \rangle_{E+V} \\
&= -k_\mathrm{B}T \ln P^*(q_0) - V(q_0) + k_\mathrm{B}T \ln \langle \exp[\beta V] \rangle_{E+V}
\end{aligned} \tag{IX.30}$$

giving the free energy as function of reaction coordinate, or PMF in the form

$$F(q) = -k_\mathrm{B}T \ln P^*(q) - V(q) + K \tag{IX.31}$$

This result is very interesting: We have added an arbitrary potential $V(q)$ to our system. Now, we have to calculate the ensemble averages with the biased potential $E+V$ as indicated by $\langle\rangle_{E+V}$. We obtain the *biased probability* $P^*(q)$ of finding the system at the value of the reaction coordinate for the ensemble $E+V$, which can obviously be very different from that of the unbiased ensemble $P(q)$. Yet, we still get the right (unbiased) free energy $F(q)$, once we take the biased probability $P^*(q)$, subtract the biasing potential $V(q)$ at the value of the reaction coordinate and add the term K.

We can use this scheme efficiently, by way of moving the biasing (harmonic) potential along the reaction coordinate as shown in Fig. 10. In this case, we perform k simulations with the potentials V_k and get:

$$F(q) = -k_\mathrm{B}T \ln P^*(q) - V_k(q) + K_k \tag{IX.32}$$

[5] This should evoke the image of an interval of the reaction coordinate being covered by an umbrella.
[6] In such a case, $\langle \delta(q-q_0) \cdot \exp[\beta V] \rangle = \langle \delta(q-q_0) \cdot \exp[\beta V(q_0)] \rangle = \exp[\beta V(q_0)] \cdot \langle \delta(q-q_0) \rangle$ in the following.

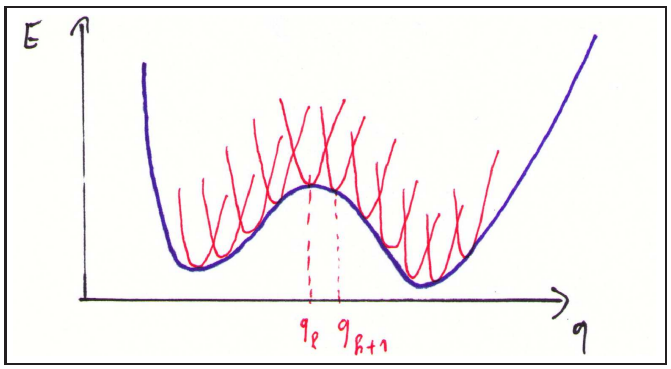

FIG. 10: Harmonic biasing potentials keep the system in the desired regions of reaction coordinate.

For each of these k simulations, we extract the probability $P^*(q)$ for every value of q and easily calculate $V^k(q)$. The curves of $-k_{\mathrm{B}}T \ln P^*(q) - V^k(q)$ for the simulations k and $k+1$ differ by a constant shift, which corresponds to the difference of K values, as shown in Fig. 11. The main task is to match the pieces together. One way is to fit the K_k in order

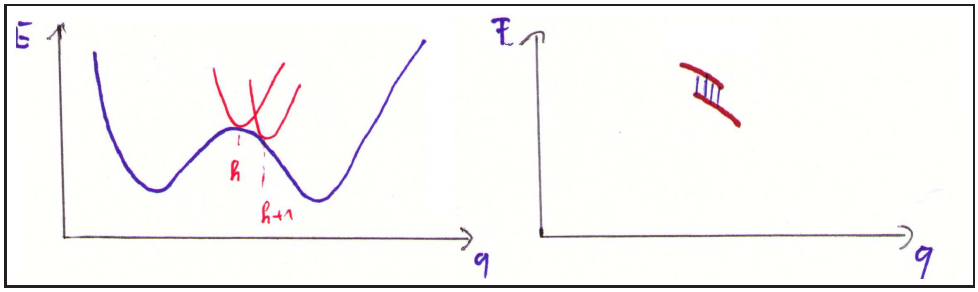

FIG. 11: The offset of free energy curves between two simulations k and $k+1$ is given by $K_k - K_{k+1}$

to get a smooth total $F(q)$ curve. This is possible if the pieces k and $k+1$ have sufficient 'overlap'.

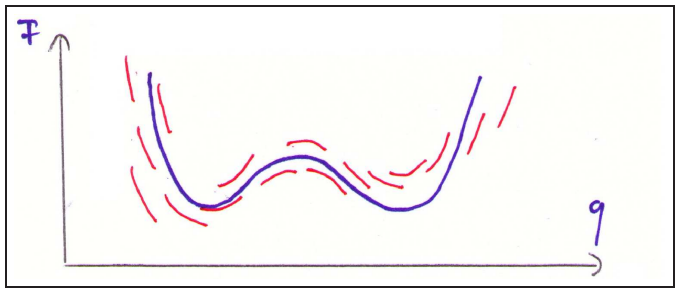

FIG. 12: Matching of histograms from different simulations

Another, quite involved method is the weighted histogram analysis method (WHAM). The starting point is the requirement of a perfect match, minimizing the total error. The

unbiased probabilities $P(x_j)$ of coordinate x falling into the bin j of the histogram and the shifts K_i are obtained by a self-consistent solution of a set of equations

$$\begin{aligned} P(x_j) &= \frac{\sum_{i=1}^{N} n_i(x_j)\ \exp[-\beta V_i(x_j)]}{\sum_{i=1}^{N} N_i\ \exp[-\beta(V_i(x_j) - K_i)]} \\ K_i &= -k\mathrm{T}\ \log \sum_{j}^{\mathrm{bins}} P(x_j)\ \exp[-\beta V_i(x_j)] \end{aligned} \tag{IX.33}$$

(for a total of N simulations, i-th simulation contains N_i frames, $n_i(x_j)$ is the number of hits in bin j in simulation i). The WHAM procedure is included in a range of modern packages for MD simulations.