

# Supporting Information for “I. Dissociation free energies in drug-receptor systems via non equilibrium alchemical simulations: theoretical framework”

Piero Procacci\*

*Department of Chemistry, University of Florence, Italy*

E-mail: procacci@unifi.it

According to the Crooks theorem, the simulation volume dependent annihilation free energy of the bound ligand is given by

$$\begin{aligned}\Delta G_{\text{sim}}^{(b)} &= -k_B T \ln \left[ (1-c)e^{-\beta \Delta G_x} + ce^{-\beta \Delta G_{sx}} \right] \\ &= -k_B T \ln \left\{ (1-c)e^{-\beta \Delta G_x} \left[ 1 + \frac{c}{1-c} e^{\beta(\Delta G_x - \Delta G_{sx})} \right] \right\}\end{aligned}\quad (1)$$

with  $\Delta G_x = \langle W_{1 \rightarrow 0} \rangle - \frac{\beta \sigma_1^2}{2}$  and  $\Delta G_{sx} = \langle W_{2 \rightarrow 0} \rangle - \frac{\beta \sigma_2^2}{2}$ , the latter term being due the shadow normal component ( $c \ll 1$ ) of the unbound ligand in MD box of standard size. For normal work distributions,  $\Delta G_{sx}$ , i.e. the annihilation free energy of the free ligand in a box of volume  $V_{\text{box}}$  containing the protein, is strictly related to the hydration free energy of the pure ligand and, clearly,  $\Delta G_{sx} \rightarrow \Delta G_s$  as  $V_{\text{box}} \rightarrow \infty$ . We make the ansatz that

$$\Delta G_{sx} = \Delta G_s (1 - e^{-V_{\text{box}}/V_{\text{RL}}}) = \Delta G_s - \delta \Delta G_s \quad (2)$$

---

\*To whom correspondence should be addressed

where  $V_{\text{RL}}$  is defined by the above equation and is related to the sum of the volumes of the ligand and of the protein in standard conditions and  $\delta\Delta G_s$  is a correction due to the presence of the protein in the box. As discussed in the paper, the quantity  $\Delta G_x$ , the annihilation free energy starting from the bound ligand states, is instead independent of the box size and must be hence related to some bound state physical observable. We have seen in the paper that  $\Delta G_0^b = \Delta G_x + k_B T \ln(V_{\text{site}}/V_0)$ . In an ideal equilibrium simulation in box of volume  $V_{\text{box}}$ , the volume dependent dissociation free energy is related to the standard dissociation free energy as

$$\begin{aligned}\Delta G_{\text{sim}} &= \Delta G_0^b - \Delta G_s - k_B T \ln \frac{V_{\text{box}}}{V_0} \\ \Delta G_0^b - \Delta G_s &= \Delta G_{\text{sim}} + k_B T \ln \frac{V_{\text{box}}}{V_0} \\ \Delta G_x - \Delta G_s &= \Delta G_{\text{sim}} + k_B T \ln \frac{V_{\text{box}}}{V_{\text{site}}}\end{aligned}\quad (3)$$

so that, conventionally,  $\Delta G_{\text{sim}} = \Delta G_0^b - \Delta G_s = \Delta G_0$  when the box volume equals that of the standard state. Upon substitution of Eq. 3 into Eq. 1 and exploiting Eq. 2, we find that

$$\Delta G_{\text{sim}}^{(b)} = -k_B T \ln \left\{ (1-c)e^{-\beta\Delta G_x} \left[ 1 + \frac{c}{1-c} e^{\beta\left(\Delta G_{\text{sim}} + \delta\Delta G_s + k_B T \ln \frac{V_{\text{box}}}{V_{\text{site}}}\right)} \right] \right\} \quad (4)$$

We also have that the ratio of the weights of the two normal components in Eq. 1 is related to the *equilibrium* probability ratio of the bound and unbound state in the simulation volume  $V_{\text{box}}$ , that is

$$\frac{c}{1-c} = e^{-\beta\Delta G_{\text{sim}}} \quad (5)$$

Using Eq. 5 in Eq. 4

$$\begin{aligned}\Delta G_{\text{sim}}^{(b)} &= -k_B T \ln \left[ (1-c)e^{-\beta\Delta G_x} \left( 1 + \frac{V_{\text{box}}}{V_{\text{site}}} e^{\beta\delta\Delta G_s} \right) \right] \\ \Delta G_{\text{sim}}^{(b)} &= \Delta G_x - k_B T \ln \left[ (1-c) \left( 1 + \frac{V_{\text{box}}}{V_{\text{site}}} e^{\beta\delta\Delta G_s} \right) \right]\end{aligned}\quad (6)$$

We now set in Eq. 1  $c = 0.5$ , such that the weights of the starting states are equal. We obtain

$$\begin{aligned}\Delta G_{\text{sim}}^{(b)} &= k_B T \ln 2 + \ln \left[ e^{-\beta \Delta G_x} + e^{-\beta \Delta G_{sx}} \right] \\ &\simeq k_B T \ln 2 + \Delta G_{sx}\end{aligned}\quad (7)$$

where the latter equation is essentially exact for tight binding ligands. The  $c=1/2$   $\Delta G_{\text{sim}}$  differs by a term  $k_B T \ln 2$  with respect to the hydration free energy of the ligand  $\Delta G_{sx} \simeq \Delta G_s$ . This is due to the fact that the two equal weight normal components referring to the starting bound and unbound states in FS-DAM are normalized to 1/2 rather than to 1, so that the crossing point with the reverse distribution  $P(-W)$ , exhibiting a single normal component with weight (essentially) 1 peaked at  $\Delta G_s + 0.5\beta\sigma_s^2$ , get up-shifted. Consequently we do expect that also the volume  $V'_*$  for which  $c = 1/2$  is multiplied by a factor of two with respect to the volume,  $V_*$ , at which  $\Delta G_{\text{sim}}(c = 0.5) = \Delta G_s$ , i.e.

$$V'_* = 2V_* \quad (8)$$

Setting  $c = 1/2$  in Eq. 6, and using Eq. 7 we find that the equal weight volume in FS-DAM satisfies the relation:

$$\Delta G_x - \Delta G_{sx} = k_B T \ln \left[ \left( 1 + \frac{V'_*}{V_{\text{site}}} e^{\beta \delta \Delta G_s} \right) \right] \quad (9)$$

so that

$$\begin{aligned}\frac{V'_*}{V_{\text{site}}} &= e^{\beta(\Delta G_x - \Delta G_s)} - e^{-\beta \delta \Delta G_s} \\ \frac{V_*}{V_{\text{site}}} &\simeq \frac{1}{2} e^{\beta(\Delta G_x - \Delta G_s)}\end{aligned}\quad (10)$$

where in the last equation we have exploited the fact that  $\Delta G_{sx}$  is essentially equal to  $\Delta G_s$  for large box volumes and used Eq. 8. On the other hand, the equilibrium constant of the reaction  $RL \rightleftharpoons R + L$  does not depend on the concentrations of the species but the ratio of the bound and

unbound states probability does. The concentration  $c_*$  for which the two probability are equal is trivially given by

$$\begin{aligned}\frac{[\mathbf{R}][\mathbf{L}]}{\mathbf{RL}} &= \frac{\frac{c_*}{2} \frac{c_*}{2}}{\frac{c_*}{2}} = \frac{c_*}{2} = K \\ c_* &= 2K\end{aligned}\tag{11}$$

Eq. 11 can be written in the equivalent form

$$\begin{aligned}\frac{V_*}{V_0} &= \frac{1}{2} e^{\beta(\Delta G_0^b - \Delta G_s)} \\ \frac{V_*}{V_{\text{site}}} &= \frac{1}{2} e^{\beta(\Delta G_x - \Delta G_s)}\end{aligned}\tag{12}$$

where in Eq. 12 we have used the relation  $\Delta G_0^b = \Delta G_x + k_B T \ln(V_{\text{site}}/V_0)$  as a definition for the exclusion zone  $V_{\text{site}}$ . The dissociation constant can be hence computed via FS-DAM calculations as

$$K = \frac{e^{-\beta(\Delta G_x - \Delta G_s)}}{V_{\text{site}}}\tag{13}$$

that is equivalent to Eq. 23 of the main paper. It should be noted that Eq. 13 (or Eq. 23 of the main paper) makes no reference whatsoever to the speed of the NE process and is hence a *general* equation that is valid also for alchemical transformations done at infinitely low speed. In this latter case one should assume that a (equilibrium) DAM simulation yields an infinitely narrow distribution peaked at  $\Delta G_x$  and  $\Delta G_s$  for the annihilation of the bound and unbound ligand, respectively.