Supporting information for: Efficient determination of protein-protein standard binding free energies from first principles

James C. Gumbart,[†] Benoît Roux,[‡] and Christophe Chipot[*,¶,§]

School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, Department of Biochemistry and Molecular Biology and Gordon Center for Integrative Science, The University of Chicago, Chicago, IL 60637, and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana IL 61801

E-mail: chipot@ks.uiuc.edu

[*]To whom correspondence should be addressed

[†]School of Physics, Georgia Institute of Technology, Atlanta, GA 30332

[‡]Department of Biochemistry and Molecular Biology and Gordon Center for Integrative Science, The University of Chicago, Chicago, IL 60637

[¶]University of Illinois at Urbana-Champaign

[§]Équipe de dynamique des assemblages membranaires, UMR 7565, Université de Lorraine, BP 70239, 54506 Vandœuvre-lès-nancy cedex, France

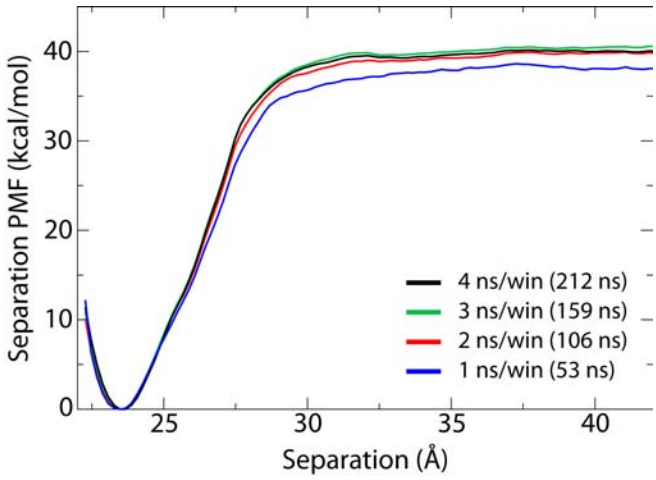

Figure S1: Separation PMF from REMD-US simulations as a function of time. The PMF is shown after 53 (blue), 106 (red), 159 (green), and 212 ns (black).

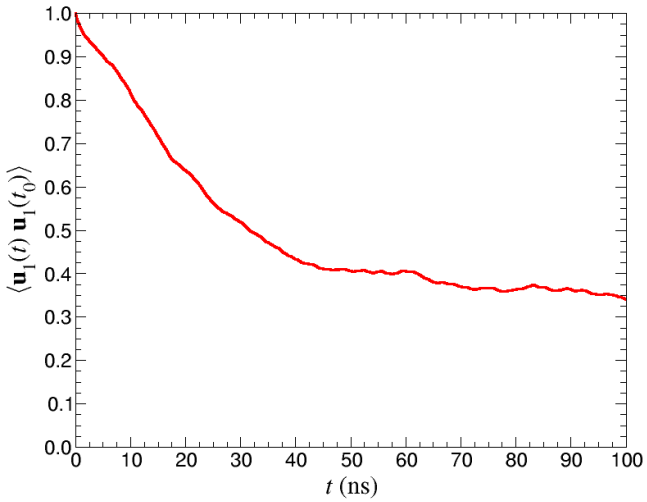

Figure S2: Orientational auto-correlation function for barstar alone in an aqueous environment based on a 150-ns simulation.

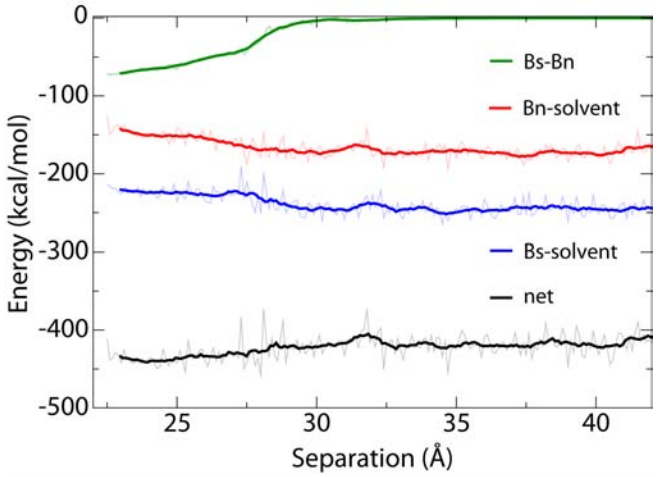

Figure S3: Instantaneous Lennard-Jones interaction energy as a function of separation during REMD-US simulations averaged over time. The interaction energy was calculated for barstar with barnase (green), barstar with solvent (blue), barnase with solvent (red), and then summed (black). The light lines represent original data and the heavy lines a running average.

Explicit derivation of the binding free energy from PMFs (372 ns)

Contributions from adding restraints in the bound state

All of the calculations in this section were run with the ABF approach using an extended coordinate. The temperature is $T = 300\,\mathrm{K}$. Specific atoms restrained are given, as is the time required for convergence. A uniform force constant of $k = 10\,\mathrm{kcal/mol{\cdot}\AA^2}$ was used for all conformational restraints, while $k = 0.1\,\mathrm{kcal/mol{\cdot}degree^2}$ for all orientational and positional restraints. The order in which the various restraints are listed also corresponds to the sequence in which they were applied.

RMSD restraint on barstar, applied to carbonyl carbons only (6 ns).

$$e^{\beta G^{\mathrm{site}}_{\mathrm{Bs,c}}} = \frac{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta U}}{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta [U + u_{\mathrm{Bs,c}}]}} = \frac{0.128221}{0.004628} = 27.707 \rightarrow G^{\mathrm{site}}_{\mathrm{Bs,c}} = 1.980\,\mathrm{kcal/mol} \quad (1)$$

RMSD restraint on barnase applied to carbonyl carbons only (12 ns).

$$e^{\beta G^{\mathrm{site}}_{\mathrm{Bn,c}}} = \frac{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta U + u_{\mathrm{Bs,c}}}}{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta [U + u_{\mathrm{Bs,c}} + u_{\mathrm{Bn,c}}]}} = \frac{0.197105}{0.001033} = 190.779 \rightarrow G^{\mathrm{site}}_{\mathrm{Bn,c}} = 3.130\,\mathrm{kcal/mol} \quad (2)$$

RMSD restraint on barstar residues 29 31 33 35 38 39 42 76, applied to all heavy atoms (12 ns).

$$e^{\beta G^{\mathrm{site}}_{\mathrm{Bs,res}}} = \frac{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta U + u_c}}{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta [U + u_c + u_{\mathrm{Bs,res}}]}} = \frac{0.140798}{0.006106} = 23.059 \rightarrow G^{\mathrm{site}}_{\mathrm{Bs,res}} = 1.871\,\mathrm{kcal/mol} \quad (3)$$

RMSD restraint on barnase residues 27 59 60 83 85 87 102, applied to all heavy atoms (24 ns).

$$e^{\beta G^{\mathrm{site}}_{\mathrm{Bn,res}}} = \frac{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta U + u_c + u_{\mathrm{Bs,res}}}}{\int_{\mathrm{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta [U + u_c + u_{\mathrm{Bs,res}} + u_{\mathrm{Bn,res}}]}} = \frac{0.722859}{0.002202} = 328.259 \rightarrow G^{\mathrm{site}}_{\mathrm{Bn,res}} = 3.454\,\mathrm{kcal/mol} \quad (4)$$

Orientational (20 ns) and positional (8 ns) restraints. The definition of each angle is given in parentheses (see Fig. 1A in the main text), with the reference points taken as the centers of the carbonyl carbons of the following groups of residues: all residues of barnase (P1), residues 70 to 73 and 89 to

92 of barnase (P2), 85 to 88 and 97 to 100 of barnase (P3), all residues of barstar (P1'), residues 13 to 24 of barstar (P2'), and finally residues 66 to 79 of barstar (P3'). Minima used for the restraining potentials are also given.

Orientational restraint on Θ (P1-P1'-P2'); $u_{\Theta,0} = 93^\circ$ (8 ns).

$$e^{\beta G_\Theta^{\text{site}}} = \frac{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}]}}{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_\Theta]}} = \frac{3.54799}{3.06692} = 1.1569 \rightarrow G_\Theta^{\text{site}} = 0.087\,\text{kcal/mol} \quad (5)$$

Orientational restraint on Φ (P1-P1'-P2'-P3'); $u_{\Phi,0} = 40^\circ$ (4 ns).

$$e^{\beta G_\Phi^{\text{site}}} = \frac{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_\Theta]}}{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_\Theta+u_\Phi]}} = \frac{8.10199}{4.51961} = 1.7926 \rightarrow G_\Phi^{\text{site}} = 0.348\,\text{kcal/mol} \quad (6)$$

Orientational restraint on Ψ (P2-P1-P1'-P2'); $u_{\Psi,0} = -124^\circ$ (8 ns).

$$e^{\beta G_\Psi^{\text{site}}} = \frac{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_\Theta+u_\Phi]}}{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_\Theta+u_\Phi+u_\Psi]}} = \frac{6.11885}{4.12578} = 1.4831 \rightarrow G_\Psi^{\text{site}} = 0.235\,\text{kcal/mol} \quad (7)$$

Positional restraint on θ (P1'-P1-P2); $u_{\theta,0} = 91^\circ$ (4 ns).

$$e^{\beta G_\theta^{\text{site}}} = \frac{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_o]}}{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_o+u_\theta]}} = \frac{2.35634}{1.91079} = 1.23318 \rightarrow G_\theta^{\text{site}} = 0.125\,\text{kcal/mol} \quad (8)$$

Positional restraint on ϕ (P1'-P1-P2-P3); $u_{\phi,0} = 73^\circ$ (4 ns).

$$e^{\beta G_\phi^{\text{site}}} = \frac{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_o+u_\theta]}}{\int_{\text{site}} d\mathbf{1} \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_o+u_\theta+u_\phi]}} = \frac{2.56531}{2.35208} = 1.090656 \rightarrow G_\phi^{\text{site}} = 0.052\,\text{kcal/mol} \quad (9)$$

Separation PMF from REMD-US (212 ns)

$$(1) I^* = \int_{\text{site}} dr e^{-\beta(W(r)-W(r^*))} = \int dr e^{-\beta(W(r)-W(41.5))} = \int dr e^{-\beta(W(r)-39.93)} = 7.865 \times 10^{28}\,\text{Å} \quad (10)$$

$$(2)I^* = \int dre^{-\beta(W(r)-W(39.5))} = \int dre^{-\beta(W(r)-40.07)} = 9.969 \times 10^{28}\,\text{Å} \tag{11}$$

$$(3)I^* = \int dre^{-\beta(W(r)-W(37.5))} = \int dre^{-\beta(W(r)-40.12)} = 1.080 \times 10^{29}\,\text{Å} \tag{12}$$

$$(1)S^* = (r^*)^2 \int_0^{\pi} \sin(\theta) d\theta \int_0^{2\pi} d\phi e^{-\beta u_a(\theta,\phi)} \tag{13}$$

$$= (r^*)^2 \int_0^{\pi} \sin(\theta) d\theta e^{-\beta*(0.5*0.1*(180/\pi)^2)*(\theta-91*\pi/180)^2} \tag{14}$$

$$\times \int_0^{2\pi} d\phi e^{-\beta*(0.5*0.1*(180/\pi)^2)*(\phi-73*\pi/180)^2} \tag{15}$$

$$= (41.5\,\text{Å})^2 \times (0.1067) \times (0.106814) = 19.629\,\text{Å}^2 \tag{16}$$

$$(2)S^* = (39.5\,\text{Å})^2 \times (0.1067) \times (0.106814) = 17.782\,\text{Å}^2 \tag{17}$$

$$(3)S^* = (37.5\,\text{Å})^2 \times (0.1067) \times (0.106814) = 16.027\,\text{Å}^2 \tag{18}$$

$$(1)S^*I^*(r^* = 41.5) = 1.544 \times 10^{30}\text{Å}^3 \tag{19}$$

$$(2)S^*I^*(r^* = 39.5) = 1.773 \times 10^{30}\text{Å}^3 \tag{20}$$

$$(3)S^*I^*(r^* = 37.5) = 1.730 \times 10^{30}\text{Å}^3 \tag{21}$$

Contributions from removing restraints in the unbound state

Orientational and positional restraints.

$$e^{-\beta G_{\Psi}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_{\Theta}+u_{\Phi}+u_{\Psi}]}}{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_{\Theta}+u_{\Phi}]}} \tag{22}$$

$$e^{-\beta G_{\Phi}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_{\Theta}+u_{\Phi}]}}{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_{\Theta}]}} \tag{23}$$

$$e^{-\beta G_{\Theta}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}+u_{\Theta}]}}{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{c,all}}]}} \tag{24}$$

$$e^{-\beta G_o^{\text{bulk}}} = \frac{1}{8\pi^2} \int_0^{\pi} \sin(\Theta) d\Theta \int_0^{2\pi} d\Phi \int_0^{2\pi} d\Psi e^{-\beta u_o(\Theta,\Phi,\Psi)} \tag{25}$$

$$= \frac{1}{8\pi^2} \int_0^{\pi} \sin(\Theta) d\Theta e^{-\beta*(0.5*0.1*(180/\pi)^2)*(\Theta-93*\pi/180)^2} \tag{26}$$

$$\times \int_0^{2\pi} d\Phi e^{-\beta*(0.5*0.1*(180/\pi)^2)*(\Phi-40*\pi/180)^2} \tag{27}$$

$$\times \int_0^{2\pi} d\Psi e^{-\beta*(0.5*0.1*(180/\pi)^2)*(\Psi-236*\pi/180)^2} \tag{28}$$

$$= \frac{1}{8\pi^2} \times (0.1067) \times (0.106814) \times (0.106814) = 1.5418 \times 10^{-5} \tag{29}$$

$$\rightarrow G_o^{\text{bulk}} = 6.605\,\text{kcal/mol} \tag{30}$$

RMSD restraint on barnase residues 27 59 60 83 85 87 102, applied to all heavy atoms (21 ns).

$$e^{-\beta G_{\text{Bn,res}}^{\text{bulk}}} = \times \frac{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_c+u_{\text{Bs,res}}+u_{\text{Bn,res}}]}}{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_c+u_{\text{Bs,res}}]}} = \frac{0.76632}{1.03936 \times 10^{-6}} = 737299.88 \tag{31}$$

$$\rightarrow G_{\text{Bn,res}}^{\text{bulk}} = 8.054\,\text{kcal/mol} \tag{32}$$

RMSD restraint on barstar residues 29 31 33 35 38 39 42 76, applied to all heavy atoms (15 ns).

$$e^{-\beta G^{\text{bulk}}_{\text{Bs,res}}} = \times \frac{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_c+u_{\text{Bs,res}}]}}{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_c]}} = \frac{0.725745}{0.000129096} = 5621.7466 \tag{33}$$

$$\rightarrow G^{\text{bulk}}_{\text{Bs,res}} = 5.147\,\text{kcal/mol} \tag{34}$$

RMSD restraint on barnase, applied to carbonyl carbons only (18 ns).

$$e^{-\beta G^{\text{bulk}}_{\text{Bn,c}}} = \times \frac{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{Bs,c}}+u_{\text{Bn,c}}]}}{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{Bs,c}}]}} = \frac{0.394126}{0.000375907} = 1048.4668 \tag{35}$$

$$\rightarrow G^{\text{bulk}}_{\text{Bn,c}} = 4.146\,\text{kcal/mol} \tag{36}$$

RMSD restraint on barstar, applied to carbonyl carbons only (24 ns).

$$e^{-\beta G^{\text{bulk}}_{\text{Bs,c}}} = \times \frac{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta[U+u_{\text{Bs,c}}]}}{\int_{\text{bulk}} d\mathbf{1}\delta(\mathbf{r}_1 - \mathbf{r}_1^*) \int d\mathbf{X} e^{-\beta U}} = \frac{0.359751}{0.00121979} = 294.9286 \tag{37}$$

$$\rightarrow G^{\text{bulk}}_{\text{Bs,c}} = 3.390\,\text{kcal/mol} \tag{38}$$

Final calculation of the binding constant and free energy

$$K_{\text{eq}} = S^* I^* e^{-\beta[(G^{\text{bulk}}_{\text{Bs,c}} - G^{\text{site}}_{\text{Bs,c}}) + (G^{\text{bulk}}_{\text{Bn,c}} - G^{\text{site}}_{\text{Bn,c}}) + (G^{\text{bulk}}_{\text{Bs,res}} - G^{\text{site}}_{\text{Bs,res}}) + (G^{\text{bulk}}_{\text{Bn,res}} - G^{\text{site}}_{\text{Bn,res}}) + (G^{\text{bulk}}_{o} - G^{\text{site}}_{o}) - G^{\text{site}}_{a}]} \tag{39}$$

$$= S^* I^* \times e^{-\beta[(3.390-1.980)+(4.146-3.130)+(5.147-1.871)+(8.054-3.454)} \tag{40}$$

$$+(6.605-(0.235+0.125+0.052))-(0.087+0.348)] \tag{41}$$

$$= S^* I^* \times e^{-16.060\beta} = S^* I^* \times 1.9922 \times 10^{-12} \tag{42}$$

$$(1) K_{\mathrm{eq}} = 1.544 \times 10^{30}\,\text{Å}^3 \times 1.9922 \times 10^{-12} = 3.0759 \times 10^{18}\,\text{Å}^3 (r^* = 41.5) \tag{43}$$

$$(2) 1.773 \times 10^{30}\,\text{Å}^3 \times 1.9922 \times 10^{-12} = 3.5321 \times 10^{18}\,\text{Å}^3 (r^* = 39.5) \tag{44}$$

$$(3) 1.730 \times 10^{30}\,\text{Å}^3 \times 1.9922 \times 10^{-12} = 3.4465 \times 10^{18}\,\text{Å}^3 (r^* = 37.5) \tag{45}$$

$$G_{\mathrm{bind}} = -kT \log(K_{\mathrm{eq}} C^{\circ}) = -kT \log(K_{\mathrm{eq}} \frac{1}{1661\,\text{Å}^3}) = \tag{46}$$

$$(1)\ 20.96\,\mathrm{kcal/mol} \qquad (r^* = 41.5) \tag{47}$$

$$(2)\ 21.04\,\mathrm{kcal/mol} \qquad (r^* = 39.5) \tag{48}$$

$$(3)\ 21.02\,\mathrm{kcal/mol} \qquad (r^* = 37.5) \tag{49}$$