Supporting Information for Publication: Challenges Encountered Applying Equilibrium and Non-equilibrium Binding Free Energy Calculations

Hannah M. Baumann,[†] Vytautas Gapsys,[‡] Bert L. de Groot,[‡] and David L. Mobley[*,†,¶]

†*Department of Pharmaceutical Sciences, University of California, Irvine*
‡*Computational Biomolecular Dynamics Group, Department of Theoretical and Computational Biophysics, Max Planck Institute for Biophysical Chemistry, D-37077 Göttingen, Germany*
¶*Department of Chemistry, University of California, Irvine*

E-mail: dmobley@mobleylab.org

S1. Hydration free energy calculations

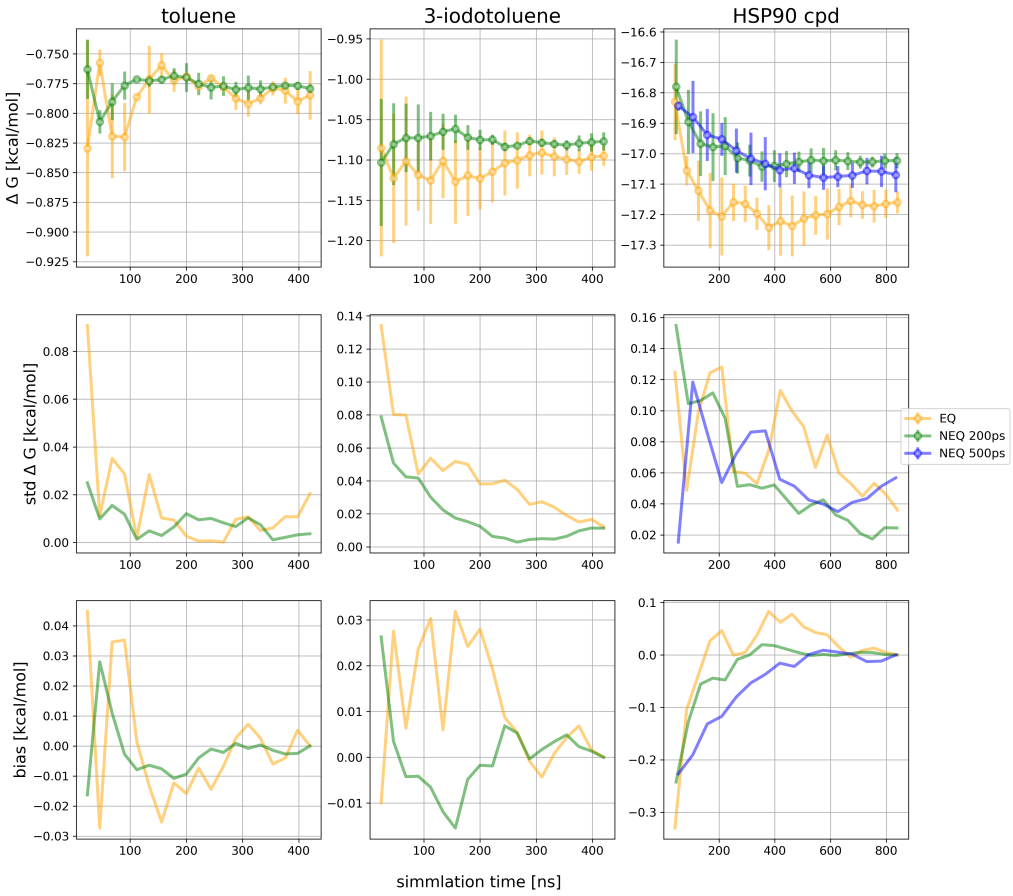


Figure S1. Hydration free energy, standard deviation (std) and bias as a function of total simulation time for toluene, 3-iodotoluene and the HSP90 compound. For the EQ approach (yellow) we show the mean free energy difference across three replicates in the top row while for the NEQ approach (green and blue) the free energy estimate was obtained by pooling work values from three replicates and estimating one free energy difference using BAR (Sec. 4.5). The uncertainty estimate in both approaches is the standard deviation across three replicates. Second and third row represents the standard deviation and bias. For the bias we assumed that the final value of the mean free energy is the theoretical hydration free energy of the compounds, meaning that it may be underestimated if simulations did not converge. Standard deviation and bias were reduced faster with the NEQ approach, suggesting that it is slightly more efficient in this test set. For the HSP90 compound the two methods did not converge to the same mean free energy within uncertainty. We tested a different NEQ protocol with longer, but fewer transitions (blue, 500 ps/transition). EQ and NEQ approaches converged to statistically the same result using this protocol.

S2. Restraints in binding free energy calculations

Table S1. Protein (P1, P2, P3) and ligand atoms (L1, L2, L3) involved in Boresch-style restraints. One distance (P3-L1), two angles (P2-P3-L1 and P3-L1-L2) and three dihedrals (P1-P2-P3-L1, P2-P3-L1-L2 and P3-L1-L2-L3) were restrained to their value in an equilibrated reference structure.

	P1	P2	P3	L1	L2	L3
toluene Pose I	Cys54 CA	Ala98 O	Ala99 O	C3	C5	C4
toluene Pose II	Cys54 CA	Ala98 O	Ala99 O	C3	C5	C4
3-iodotoluene	Cys54 CA	Ala98 O	Ala99 O	C1	C6	C3
HSP90 compound	Tyr201 CD2	Pro164 C	Gly82 O	N2	C1	C3

S3. Toluene binding to T4 lysozyme L99A

Table S2. Binding free energy ΔG° for the toluene/T4 lysozyme system. We show the free energy difference of each individual leg of the thermodynamic cycle as well as the binding free energy calculated by summation along the cycle. Reported uncertainties for decoupling the ligand in the binding site ($\Delta G_{decouple\ ligand}$) and in the solvent ($\Delta G_{solvate\ ligand}$) are the standard deviation of three independent replicates. $\Delta G_{restraints\ off}$ was calculated analytically, using the formula in Boresch *et al.*[1]

	$\Delta G_{restraints\ on}$ [kcal/mol]	$\Delta G_{decouple\ ligand}$ [kcal/mol]	$\Delta G_{restraints\ off}$ [kcal/mol]	$\Delta G_{solvate\ ligand}$ [kcal/mol]	ΔG^0 [kcal/mol]
Pose I EQ	1.96 ± 0.08	11.68 ± 0.08	-8.615	-0.79 ± 0.02	-4.2 ± 0.1
Pose I NEQ 200 ps		11.6 ± 0.3		-0.779 ± 0.004	-4.2 ± 0.3
Pose I NEQ 500 ps		11.7 ± 0.3			-4.3 ± 0.3
Pose I NEQ 1 ns		11.7 ± 0.2			-4.3 ± 0.2
Pose II EQ	1.3 ± 0.1	9.6 ± 0.1	-6.774	-0.79 ± 0.02	-3.3 ± 0.1
Pose II NEQ 200 ps		9.6 ± 0.1		-0.779 ± 0.004	-3.3 ± 0.1
Pose II NEQ 500 ps		9.51 ± 0.08			-3.3 ± 0.1
Pose II NEQ 1 ns		9.5 ± 0.2			-3.2 ± 0.2

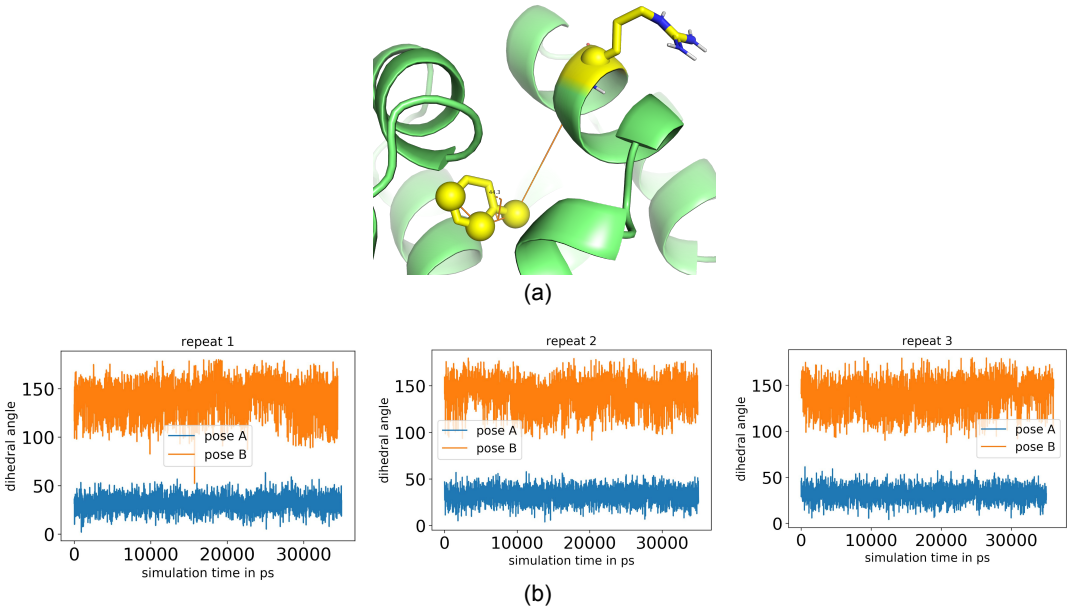


Figure S2. Two binding modes of toluene in T4 lysozyme L99A. (a) The dihedral angle used to differentiate the two bindings modes is defined by the C1, C3 and C5 atoms of toluene and the alpha carbon of arginine 119, depicted as spheres. (b) The crystallographic binding mode (pose I) is shown in blue and the alternative mode (pose II) in orange. Orientational restraints kept toluene from transitioning between binding modes. The two binding modes were restrained with different force constants. Pose I was restrained with a force constant of 20 $kcal/(mol * Å^2)$ for the bond and 20 $kcal/(mol * rad^2)$ for angles and dihedrals, while the alternative binding mode was restrained with 10 $kcal/(mol * Å^2)$ and 10 $kcal/(mol * rad^2)$.

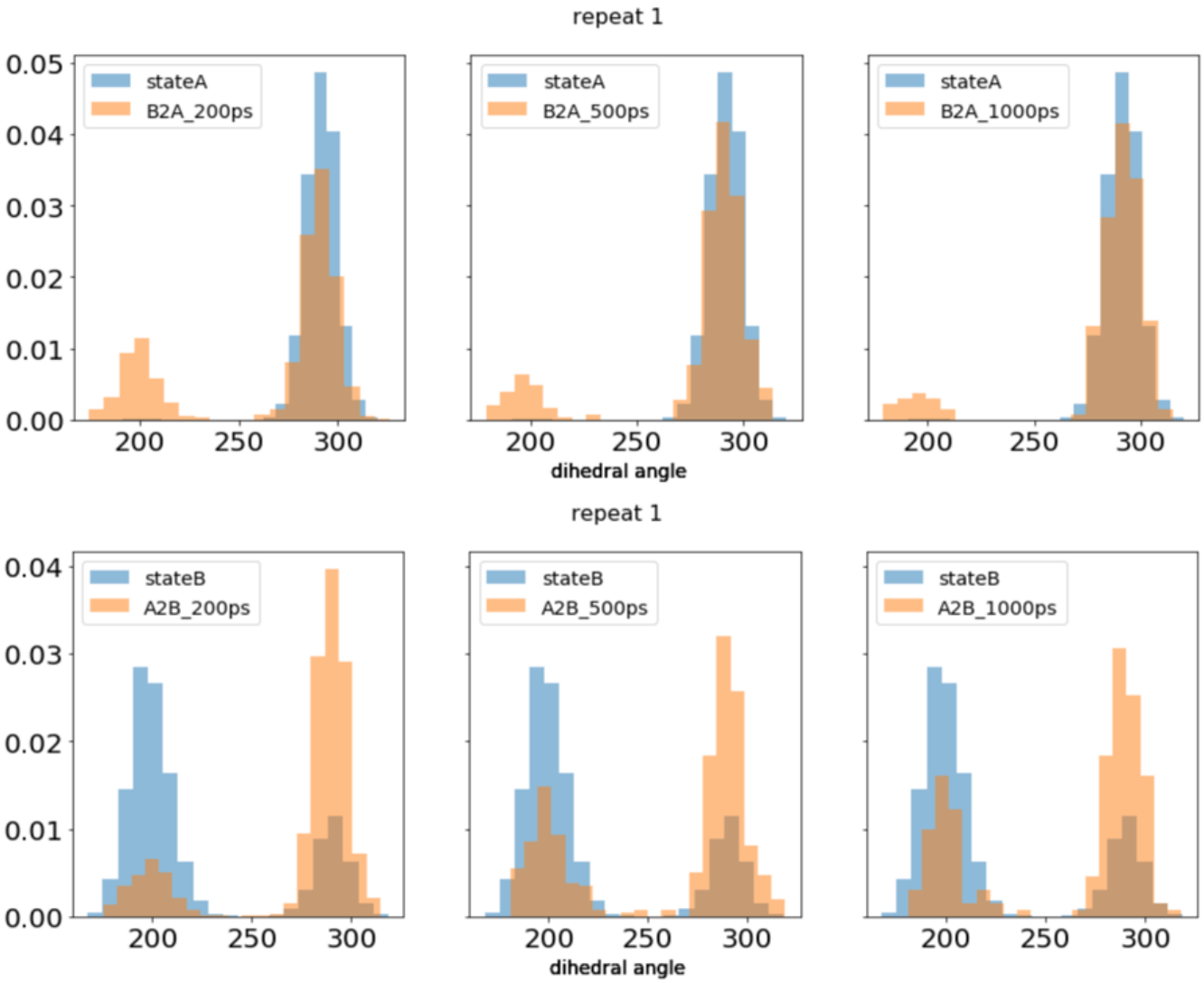


Figure S3. The distribution of the dihedral angle $\xi 1$ of Ile78 for toluene binding to T4 lysozyme L99A. Shown are values in the end states (blue) and the last frame of the switching transitions towards the end state (orange; A2B means going from state A to state B) for different transition lengths. Especially for the decoupling transitions (bottom, A2B: interacting to non-interacting), the transition length had an impact on how many transitions reoriented to the other Ile78 rotamer. This figure shows only one repeat in the crystallographic pose, all repeats and the two poses have very similar distributions and can be found in the SI.

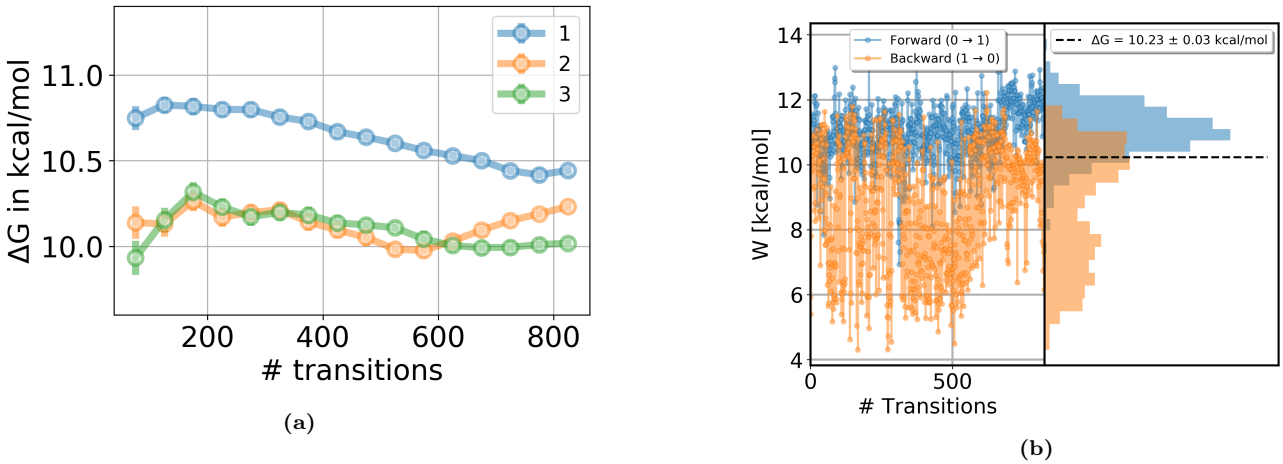


Figure S4. Protocol using weaker Boresch-style restraints (5 $kcal/(mol * Å^2)$ and 5 $kcal/(mol * rad^2)$). (a) Free energy difference for decoupling toluene in the binding site with the NEQ protocol using weaker Boresch-style restraints. The cumulative free energy difference is plotted against the number of transitions. The standard deviation across three independent replicates is 0.2 kcal/mol which is slightly lower than in the protocol with higher force constant on the restraints (0.3 kcal/mol). (b) Work values for toluene binding to T4 lysozyme L99A. Shown are values measured for each attempted transition, in forward (blue) and reverse (orange) direction, as well as the distribution of the work values. The reverse work distribution is similar to the one in the protocol using stronger restraints and the mean dissipation was independent of the restraint strength and in both protocols 1.7 kcal/mol in the reverse direction and 0.9 kcal/mol in the forward direction. Since it was previously shown [1] that the binding free energy is independent of the force constant used for the orientational restraints, we did not calculate the overall binding free energy for this protocol.

Table S3. Comparison of the free energy difference for decoupling toluene in T4 lysozyme L99A using Boresch-style restraints and center-of-mass (COM) - COM flat-bottom harmonic distance restraints. For the COM-COM restraint protocol we first turned the Boresch-style restraints off while turning on a distance restraint between the COM of the ligand and the COM of the side chain alanine 99. Then we decoupled the ligand and afterwards switched the COM-COM restraint back to the Boresch-style restraints in the noninteracting state. Reported uncertainties are the standard deviation of three independent replicates. In the NEQ approach, the dissipated work was higher in the COM-COM restraint protocol (3.8 kcal/mol in the reverse direction and 1.4 in the forward direction) than in the protocol using Boresch-style restraints. With a single COM-COM distance restraint, the ligand can sample the sphere around the radius of that distance restraint which presumably caused the larger dissipation.

		$\Delta G_{COM\ on,Boresch\ off}$ [kcal/mol]	$\Delta G_{decouple\ ligand}$ [kcal/mol]	$\Delta G_{COM\ off,Boresch\ on}$ [kcal/mol]	ΔG_{total} [kcal/mol]
Only Boresch restraints	NEQ		11.6 ± 0.3		11.6 ± 0.3
COM-COM distance restraints	EQ	-1.87 ± 0.03	5.9 ± 0.1	8.38 ± 0.04	12.4 ± 0.1
	NEQ		5.9 ± 0.3		12.4 ± 0.3

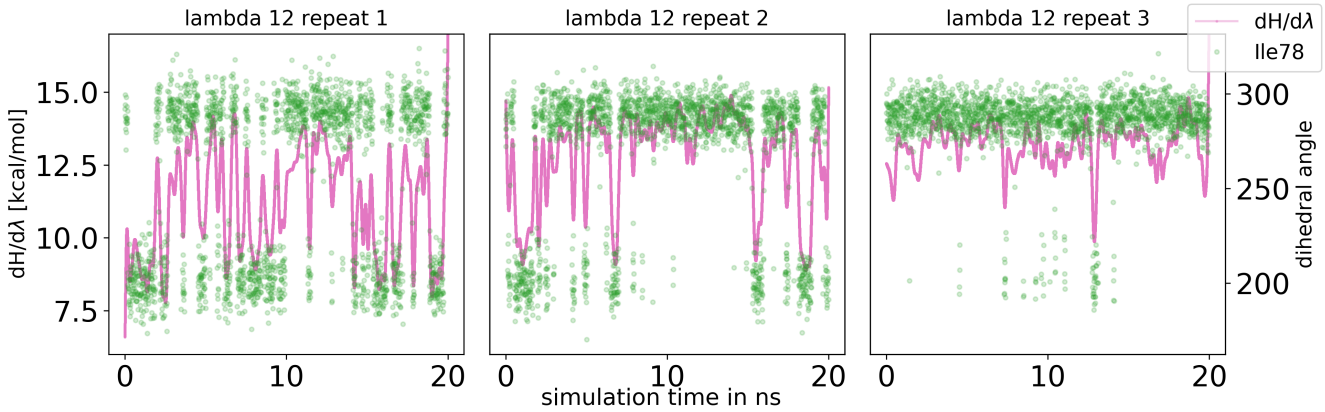


Figure S5. Correlation between dH/dλ values and the dihedral angle of Ile78 for toluene binding to T4 lysozyme L99A. A running average of the dH/dλ values (pink, averaged across 2000 data points) and the dihedral angle of Ile78 (green) is plotted as a function of simulation time. Here we only show λ 12 for the three independent replicates, additional λ windows can be found in the SI. Sudden changes in dH/dλ correlated with rotation of the Ile78 side chain.

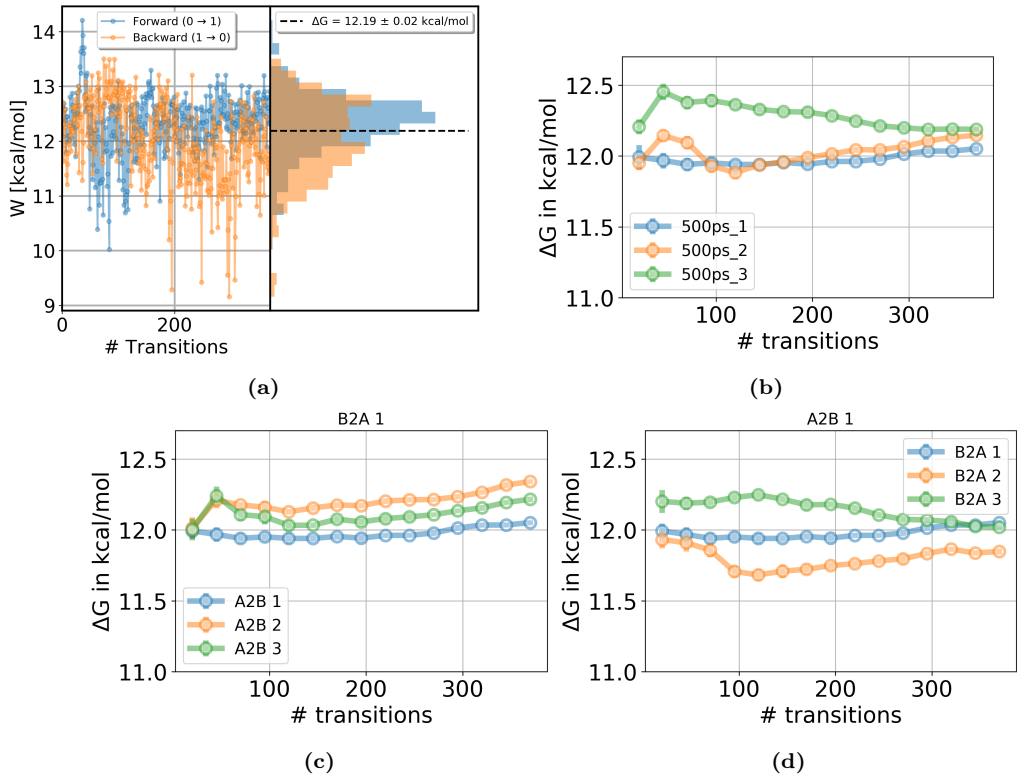


Figure S6. Restraining the Ile78 side chain in the NEQ protocol to further investigate its impact. (a) Work values for toluene binding to T4 lysozyme L99A while restraining the Ile78 side chain. Shown are values measured for each attempted transition, in forward (blue) and reverse (orange) direction, as well as the distribution of the work values. In contrast to the unrestrained protocol, the reverse work distribution (orange) is unimodal, which suggests that sampling of two Ile78 rotamers caused the bimodal work distributions in the unrestrained simulations. (b) Free energy difference for decoupling toluene in the binding site with the NEQ protocol using dihedral restraints. The cumulative free energy difference is plotted against the number of transitions. Three independent replicates seem to converge to the same free energy difference and the standard deviation was lower in this protocol compared to the unrestrained calculations (0.06 vs. 0.3 kcal/mol).(c) and (d) Analyzing the set of forward and reverse data of replicates separately. Three replicates give three sets of transitions in the forward direction and reverse direction which are independent from one another. We took a fixed set of transitions from one replicate for one direction (in (c) the set of transitions from the first replicate was used for the reverse direction (B2A 1)) and calculated the free energy difference using data from the other direction of all three replicates (in (c) A2B 1 in blue uses transitions in the forward direction from replicate one). The same was performed for the other direction in (d). The free energy difference of the second replicate (orange) is higher in the forward direction (c) and lower in the reverse direction (d). This shows that the apparent convergence (and low standard deviation) seen in (b) is likely a coincidence.

S4. 3-Iodotoluene binding to T4 lysozyme L99A

Table S4. Binding free energy ΔG° for the 3-iodotoluene/T4 lysozyme system. We show the free energy difference of each individual leg of the thermodynamic cycle as well as the binding free energy calculated by summation along the cycle. Reported uncertainties are the standard deviation of three independent replicates. $\Delta G_{restraints\ off}$ was calculated analytically, using the formula in Boresch *et al.*[1] Costs for restraining and releasing the restraints on Val111 were 0.1 and -0.1 kcal/mol.

	$\Delta G_{restraints\ on}$ [kcal/mol]	$\Delta G_{decouple\ ligand}$ [kcal/mol]	$\Delta G_{restraints\ off}$ [kcal/mol]	$\Delta G_{solvate\ ligand}$ [kcal/mol]	ΔG^0 [kcal/mol]
EQ	1.7 ± 0.1	12.8 ± 0.2	-8.638	-1.10 ± 0.01	-4.8 ± 0.2
NEQ 500 ps		13.6 ± 0.2		-1.08 ± 0.01	-5.6 ± 0.2
NEQ 1 ns		13.7 ± 0.1			-5.7 ± 0.1
NEQ 2 ns		13.5 ± 0.3			-5.5 ± 0.3
NEQ 4 ns		13.1 ± 0.1			-5.1 ± 0.1
restrain Val111 EQ		13.0 ± 0.4		-1.10 ± 0.01	-5.0 ± 0.4
restrain Val111 NEQ		12.77 ± 0.07		-1.08 ± 0.01	-4.8 ± 0.1

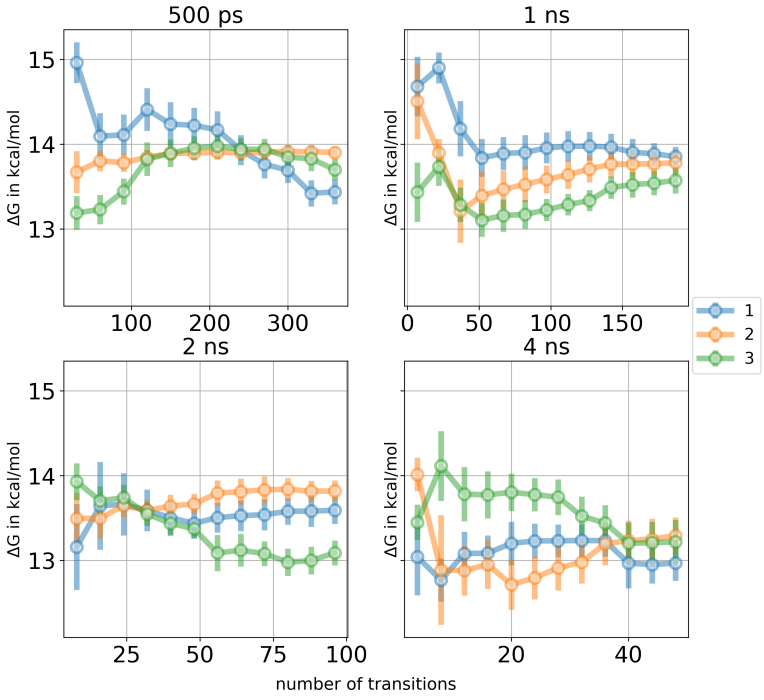


Figure S7. Free energy difference for decoupling 3-iodotoluene in the binding site using non-equilibrium FEC. The cumulative free energy difference is plotted as a function of the number of transitions. We tested four different protocols that differ in the length of the non-equilibrium transition and the number of transitions. Three independent replicates (blue, orange, green) were run for each protocol. Uncertainties were estimated via bootstrapping. The different protocols did not converge to the same free energy difference in the simulation time.

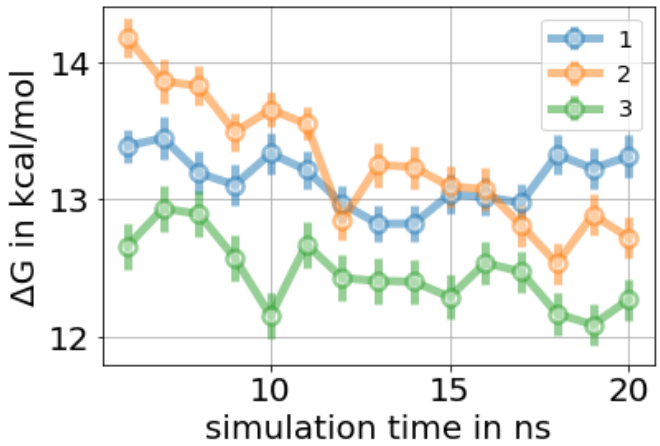


Figure S8. Free energy difference for decoupling 3-iodotoluene in the binding site using equilibrium FEC. The cumulative free energy difference of three independent replicates is plotted as a function of the simulation time per λ window. An analytical uncertainty estimate was used. Independent replicates did not converge to the same free energy difference within uncertainty in 20 ns per λ window simulation time.

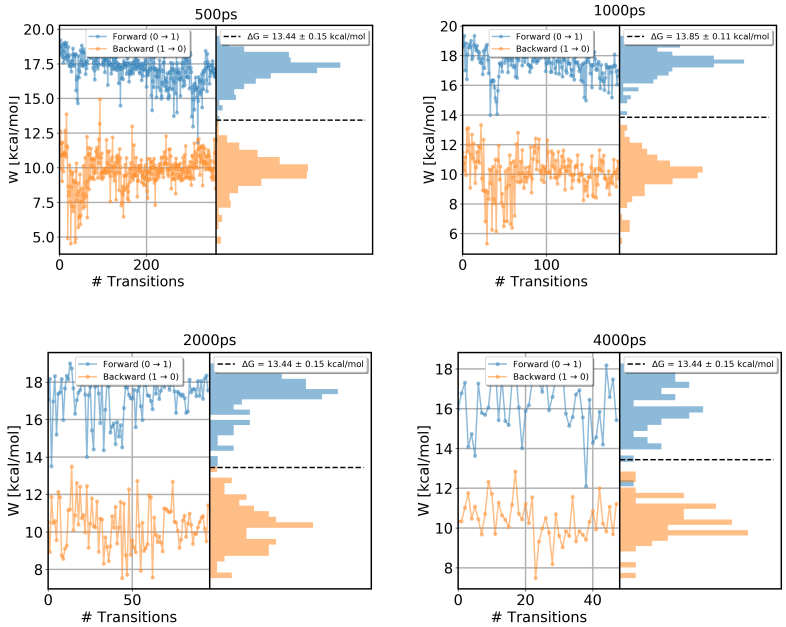


Figure S9. Work values for 3-iodotoluene binding to T4 lysozyme L99A, in forward (blue) and reverse (orange) direction. Shown are values measured for each attempted transition, as well as the distribution of the work values. The four plots show protocols that differ in the length of each NEQ switching transitions. The overlap of the work distributions in all protocols was poor, however improved slightly with transition length.

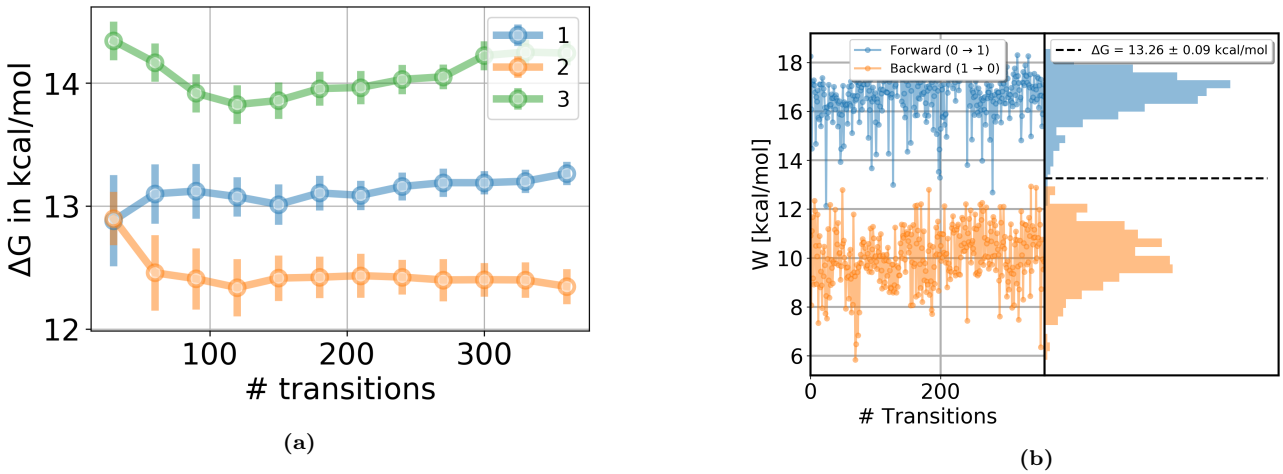


Figure S10. Weaker Boresch-style restraints (5 $kcal/(mol*\text{Å}^2)$ and 5 $kcal/(mol*rad^2)$). (a) Free energy difference for decoupling 3-iodotoluene in the binding site with the NEQ protocol using weaker Boresch-style restraints. The cumulative free energy difference is plotted against the number of transitions. The standard deviation across three independent replicates was 0.8 kcal/mol which is higher than in the protocol with stronger force constant on the restraints (0.2 kcal/mol). (b) Work values for 3-iodotoluene binding to T4 lysozyme L99A. Shown are values for replicate 1 measured for each attempted transition, in forward (blue) and reverse (orange) direction, as well as the distribution of the work values. The reverse work distribution is similar to the one in the protocol using stronger restraints and the mean dissipation was similar in both protocols, in the reverse direction 3.7 kcal/mol for the protocol with a weaker force constant and 3.9 kcal/mol for the other protocol and in the forward direction, 3.4 vs. 4.1 kcal/mol, respectively.

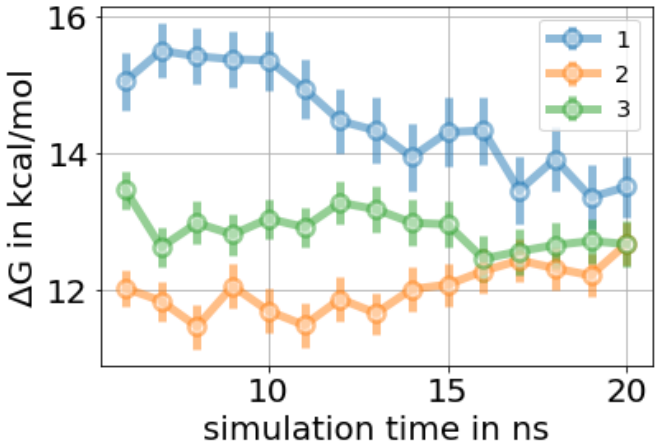


Figure S11. Decoupling 3-iodotoluene in the binding site while switching the Val111 side chain using the EQ approach. We show the free energy difference of three replicates as a function of simulation time.

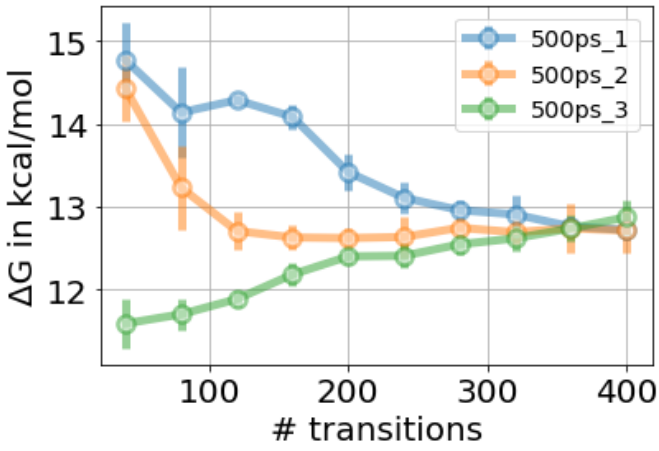


Figure S12. Decoupling 3-iodotoluene in the binding site while switching the Val111 side chain using the NEQ approach. We show the free energy difference of three replicates as a function of number of transitions. The estimate converged approximately after 320 transitions.

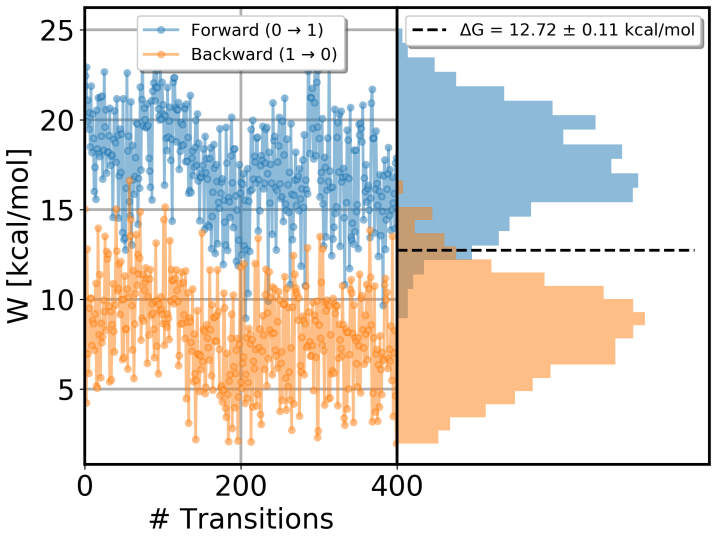


Figure S13. Work values and work distributions for decoupling 3-iodotoluene in the binding site of T4 lysozyme L99A while switching the Val111 side chain. Work values of the forward direction (decoupling, blue) and reverse direction (coupling, orange) are plotted as a function of transition number and the distribution of work values in the vertical plot. Transitions were run for 500 ps each. Forward and reverse work distributions overlapped well when the side chain was switched during the alchemical path, compared to the unrestrained protocol (Figure S9, 500 ps).

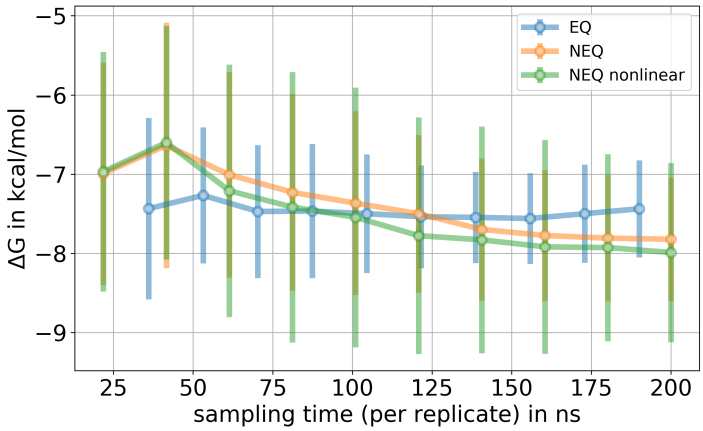


Figure S14. Free energy difference for releasing the dihedral restraints in the decoupled state of the 3-iodotoluene/ T4 lysozyme system in the protocol where multiple dihedrals were restrained. The alchemical path was nonlinear (blue and green curve) or linear (orange curve). For the EQ approach we show the mean free energy difference across six replicates while for the NEQ approach the free energy estimate was obtained by pooling work values from six replicates and estimating one free energy difference using BAR (Sec. 4.5). The uncertainty estimate in both approaches is the standard deviation across six replicates. All protocols converged to the same free energy difference within uncertainty. In the NEQ approach, a nonlinear alchemical path, where more sampling time can be spend in the more challenging part of the transformation (lower force constants) did not perform better than the protocol using a linear alchemical path. The standard deviation was high (0.6-1.1 kcal/mol) indicating sampling problems in this protocol where multiple additional side chains were restrained. Input topology files for the nonlinear NEQ protocol are provided in the SI files.

S5. The HSP90 system

Table S5. Binding free energy ΔG° for the HSP90 system. We show the free energy difference of each individual leg of the thermodynamic cycle as well as the binding free energy calculated by summation along the cycle. Reported uncertainties are the standard deviation of three independent replicates. $\Delta G_{restraints\ off}$ was calculated analytically.[1] The last two rows represent protocols, where all crystallographic water molecules were removed from the input structure, including three buried binding site waters.

	$\Delta G_{restraints\ on}$ [kcal/mol]	$\Delta G_{decouple\ ligand}$ [kcal/mol]	$\Delta G_{restraints\ off}$ [kcal/mol]	$\Delta G_{solvate\ ligand}$ [kcal/mol]	ΔG^0 [kcal/mol]
EQ	1.26 ± 0.08	36.7 ± 0.3	-9.003	-17.16 ± 0.04	-11.8 ± 0.3
HREX		35.9 ± 0.5			-11.0 ± 0.5
NEQ 1 ns		36.3 ± 0.4		-17.02 ± 0.02	-11.5 ± 0.4
NEQ 5 ns		36.2 ± 0.3			-11.4 ± 0.3
NEQ 10 ns		36.2 ± 0.5			-11.4 ± 0.5
EQ no water	1.4 ± 0.04	33.3 ± 1.3	-8.924	-17.16 ± 0.04	-8.6 ± 1.3
HREX no water		34.6 ± 0.8			-9.9 ± 0.8

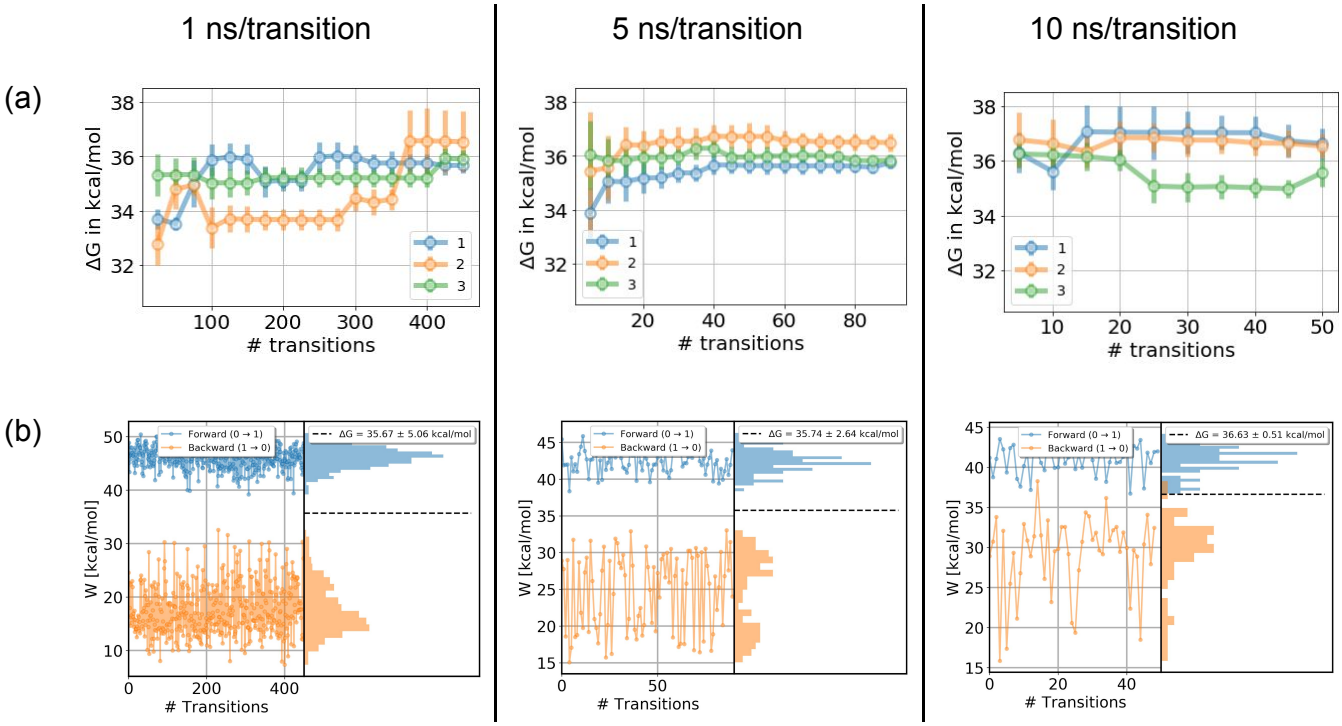


Figure S15. (a) Free energy difference of three independent replicates as a function of number of NEQ transitions in the HSP90 system. Three protocols using 1 ns, 5 ns or 10 ns per switching transition were used in the NEQ approach and uncertainty estimates obtained via bootstrapping. (b) Work values of forward (blue) and reverse (orange) transitions as a function of transition number and the distribution of work values. Shown is one example (replicate 1) per protocol, additional plots are provided in the SI. The uncertainty is the analytical error (see section 4.3). At poor overlap the uncertainty is high (5 kcal/mol for 1 ns/transition replicate 1), indicating that the estimate is likely incorrect.

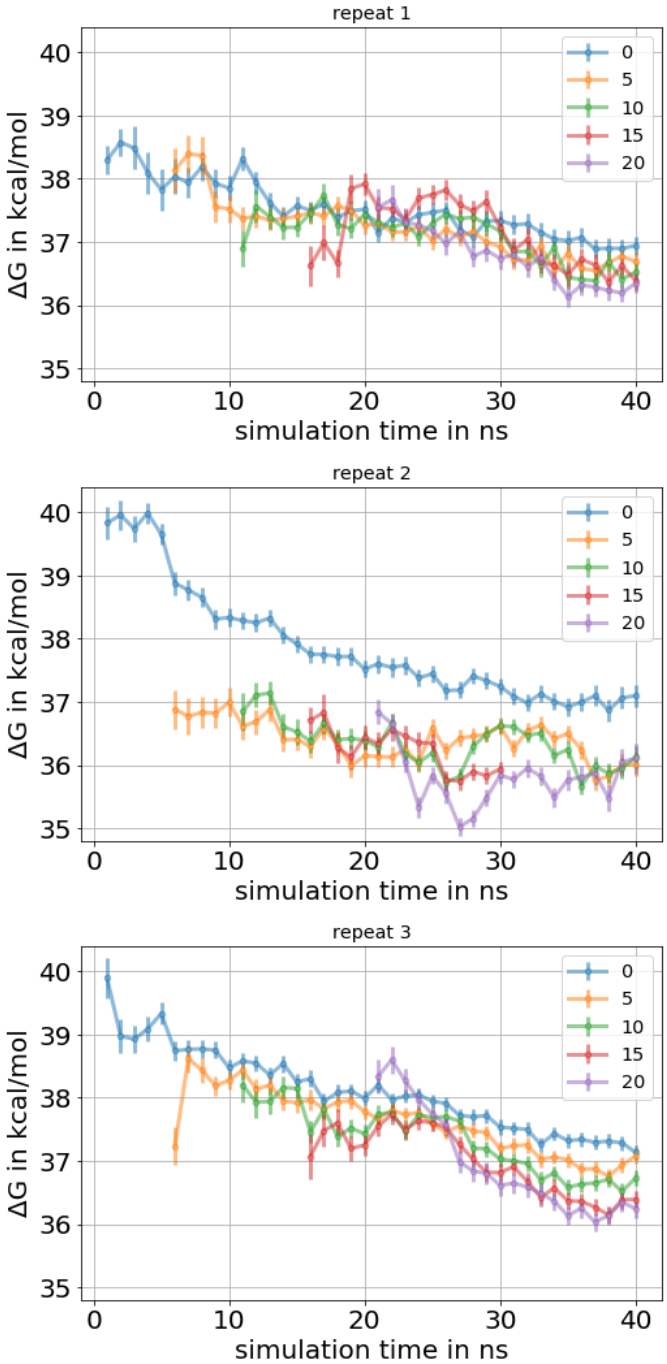


Figure S16. Impact of the equilibration time on the free energy difference in the HSP90 system. In a separate plot for every replicate, we show the cumulative free energy difference as a function of simulation time/lambda window. Different amount of data were discarded for equilibration, e.g. 0 means all data from the production run were used to estimate the free energy difference and 20 means 20 ns/λ window were discarded for equilibration. In all three repeats the final free energy difference depends on the amount of data discarded for equilibration indicating a slow DOF in the system.

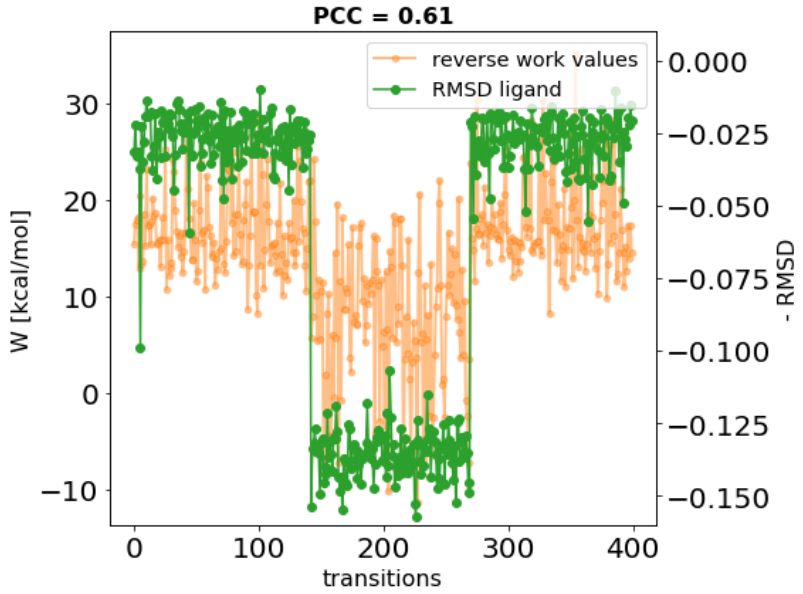


Figure S17. Correlation between reverse work values (Wr) and the ligand RMSD in the HSP90 system. The work values (orange) and the RMSD (green) are plotted as a function of the transition number. The RMSD was calculated from the set of first frames of the NEQ transitions. Work values correlated with the ligand RMSD, which is supported by a high Pearson correlation coefficient (PCC=0.61). The binding mode flip had an impact on the work values.

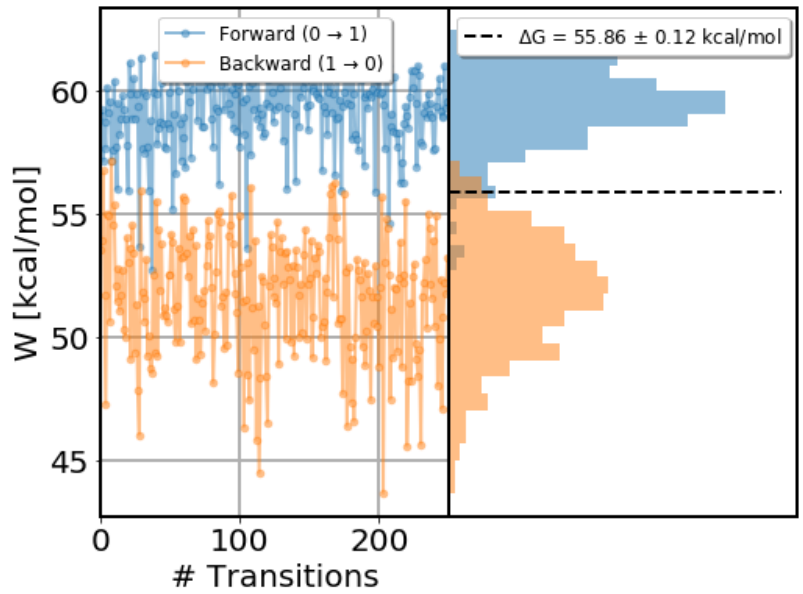


Figure S18. Work values for the ligand binding to HSP90, in forward (blue) and reverse (orange) direction in the presence of the bubble-ligand. Shown are values measured for each attempted transition, as well as the distribution of the work values. Transitions were run for 1 ns each. Forward and reverse work distributions overlapped well in the presence of the bubble-ligand, supporting that the poor overlap was caused by inadequate water sampling.

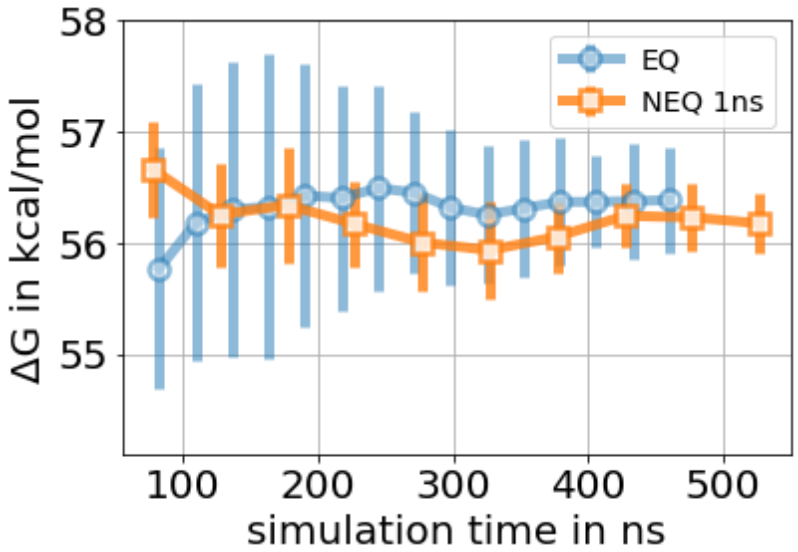


Figure S19. Free energy difference for decoupling the HSP90 ligand in the binding site in the presence of the bubble-ligand. For the EQ approach we show the mean free energy difference across three replicates while for the NEQ approach the free energy estimate was obtained by pooling work values from three replicates and estimating one free energy difference using BAR (Sec. 4.5). The uncertainty estimate in both approaches is the standard deviation across three replicates.

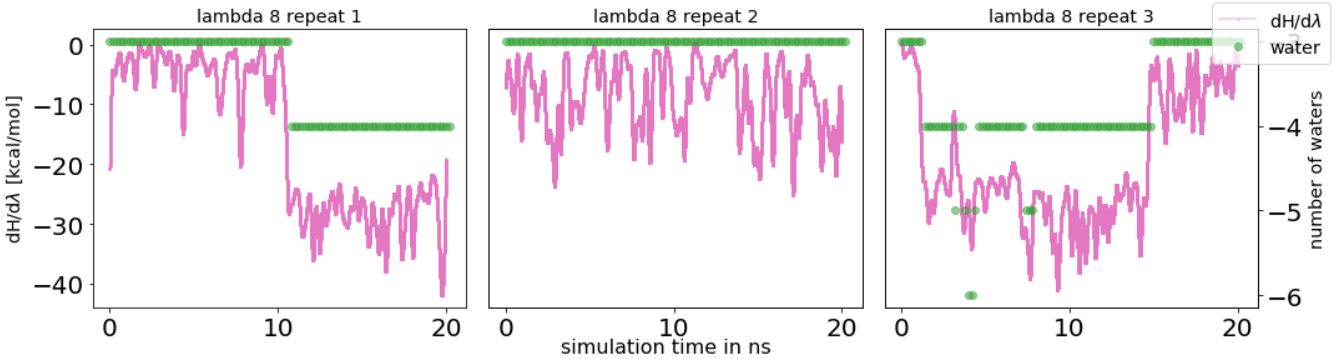
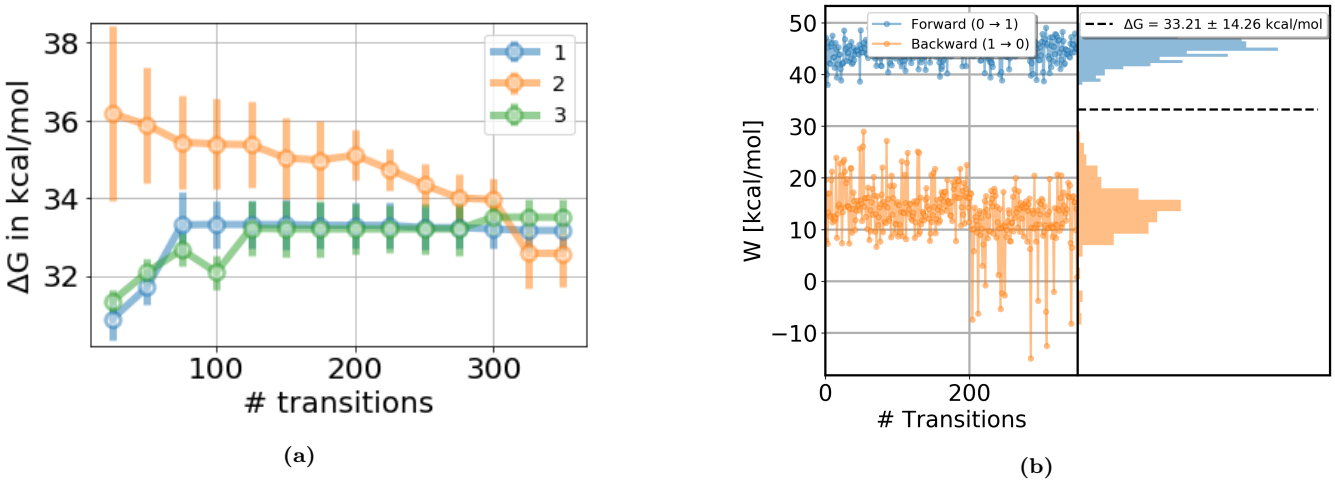


Figure S20. Water sampling problems when turning the interactions of the bubble-ligand off. Correlation between $dH/d\lambda$ values and number of water molecules in the binding site for the ligand binding to HSP90. A running average of the $dH/d\lambda$ values (pink, averaged across 2000 data points) and the number of water molecules (green) is plotted as a function of simulation time. Here we only show λ 8 for the three independent replicates, additional λ windows can be found in the SI. Sudden changes in $dH/d\lambda$ correlated with the entry/escape of water molecules into the space between protein and ligand. Water sampling was inadequate which shows that introducing the bubble-ligand only separates the water sampling from the decoupling of the ligand, but does not solve the problem.

Table S6. Binding free energy ΔG° for the HSP90 system using the bubble approach. We show the free energy difference of each individual leg of the thermodynamic cycle as well as the binding free energy calculated by summation along the cycle. Reported uncertainties are the standard deviation of three independent replicates. $\Delta G_{restraints\ off}$ was calculated analytically.[1]

	EQ	NEQ
$\Delta G_{restraints\ on}$ [kcal/mol]	1.26 ± 0.04	
$\Delta G_{restrain\ loop}$ [kcal/mol]	1.8 ± 0.2	
$\Delta G_{bubble\ on}$ [kcal/mol]	-6.979 ± 0.006	
$\Delta G_{decouple\ ligand}$ [kcal/mol]	56.4 ± 0.5	56.2 ± 0.3
$\Delta G_{bubble\ off}$ [kcal/mol]	-10.6 ± 0.2	
$\Delta G_{restraints\ loop\ off}$ [kcal/mol]	-3.1 ± 0.5	
$\Delta G_{restraints\ off}$ [kcal/mol]	-9.003	
$\Delta G_{solvate\ ligand}$ [kcal/mol]	-17.16 ± 0.04	-17.02 ± 0.02
ΔG^0 [kcal/mol]	-12.7 ± 0.8	-12.7 ± 0.6



(a) (b)

Figure S21. Weaker Boresch-style restraints (5 $kcal/(mol*\text{Å}^2)$ and 5 $kcal/(mol*rad^2)$). (a) Free energy difference for decoupling the HSP90 ligand in the binding site with the NEQ protocol using weaker Boresch-style restraints. The cumulative free energy difference is plotted against the number of transitions. Transitions were run for 1 ns. The standard deviation across three independent replicates was 0.4 kcal/mol which is higher than in the protocol with stronger force constant on the restraints (0.2 kcal/mol). (b) Work values for the ligand binding to HSP90. Shown are values for replicate 1 measured for each attempted transition, in forward (blue) and reverse (orange) direction, as well as the distribution of the work values. The reverse work distribution is similar to the one in the protocol using stronger restraints and the mean dissipation was similar in both protocols, in the reverse direction 18.7 kcal/mol for the protocol with a weaker force constant and 15.2 kcal/mol for the original protocol, and in the forward direction 11.5 vs. 13.9 kcal/mol.

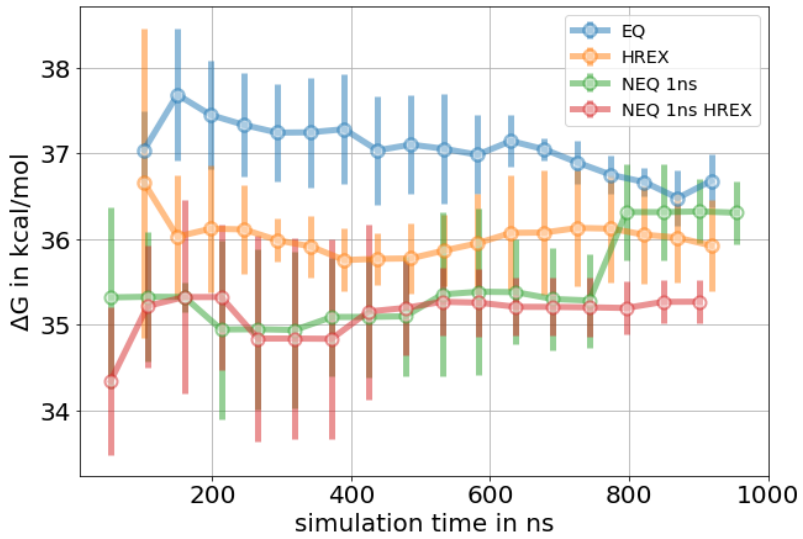
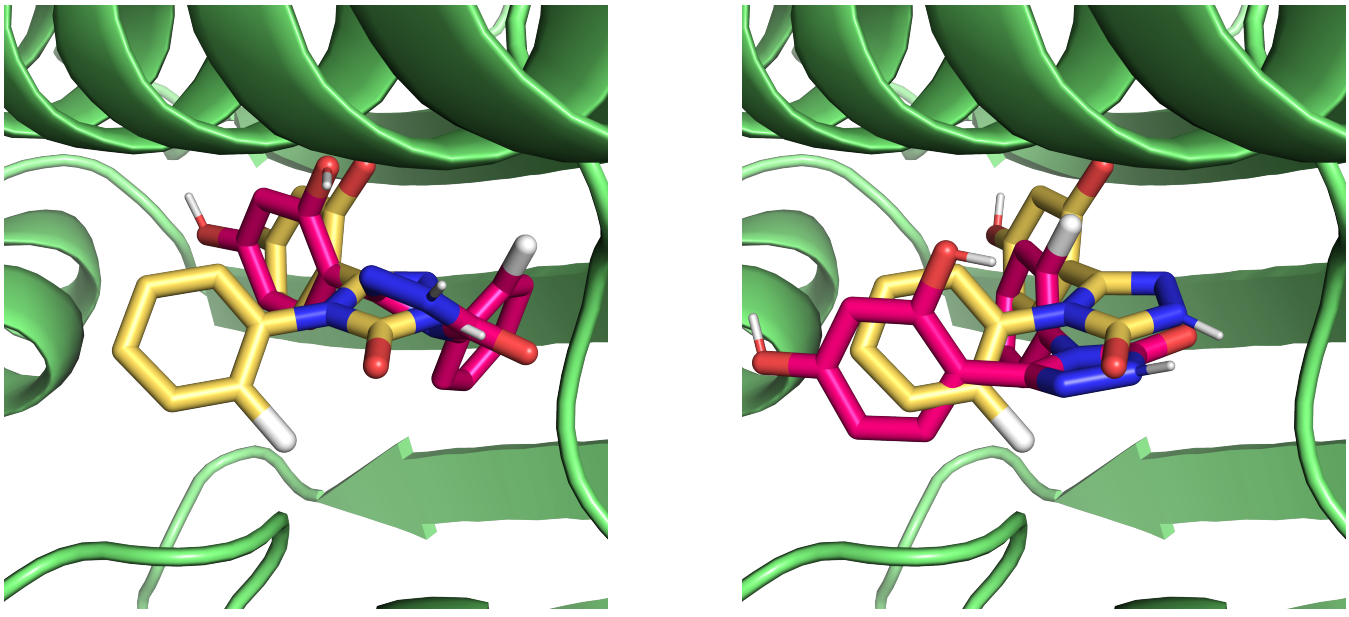


Figure S22. Free energy difference for decoupling the HSP90 ligand in the binding site. We show EQ and NEQ protocols with and without the use of HREX as an enhanced sampling technique. In the NEQ HREX approach, end state structures were used from the EQ HREX simulation, and no enhanced sampling was performed in the NEQ switches. For the EQ approach we show the mean free energy difference across three replicates while for the NEQ approach the free energy estimate was obtained by pooling work values from three replicates and estimating one free energy difference using BAR (Sec. 4.5). The uncertainty estimate in both approaches is the standard deviation across three replicates. Using structures from HREX simulations as input for the NEQ switches did not improve results in this case. On the contrary, additional sampling problems were introduced (see Figure S23).



(a) **(b)**

Figure S23. Flipped binding mode of the ligand binding to HSP90. The crystallographic binding pose is shown in yellow sticks, the flipped binding mode in magenta sticks. (a) In the non-interacting state the fluorophenyl ring was flipped in some frames. (b) Some NEQ switching transitions starting from the flipped binding mode in (a) crashed when the two phenyl rings flipped their position in the binding site.

References

(1) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *The Journal of Physical Chemistry B* **2003**, *107*, 9535–9551.