Supporting Information

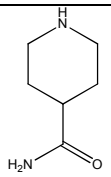
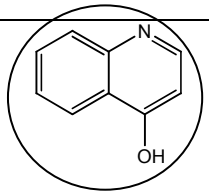
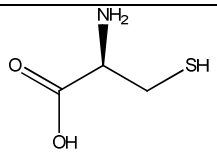
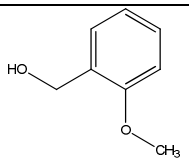
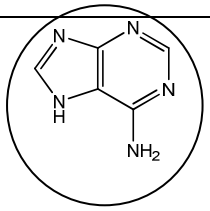
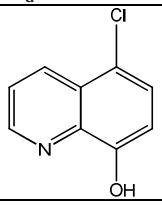
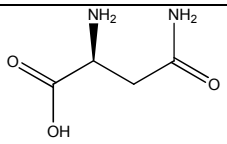
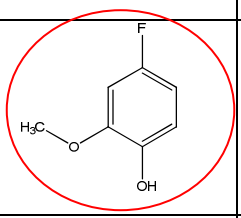
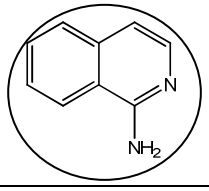
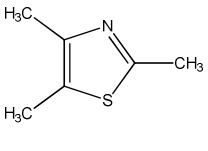
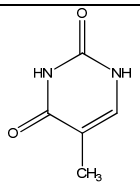
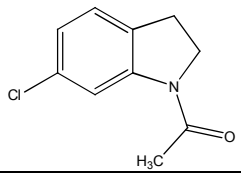
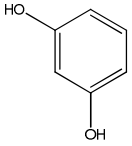
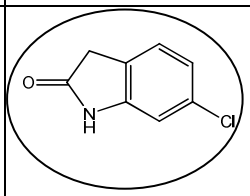
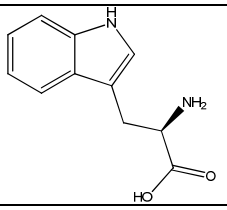
Missing Fragments: Detecting cooperative binding in fragment based drug design

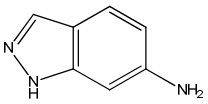
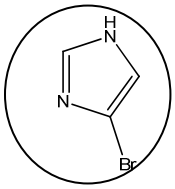
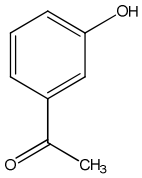
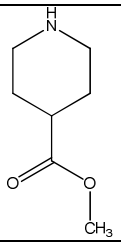
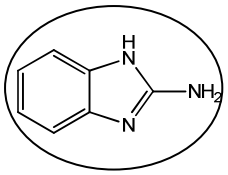
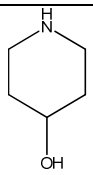
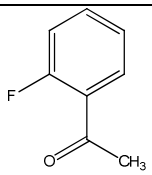
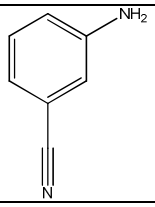
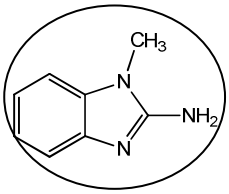
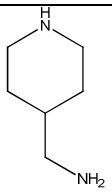
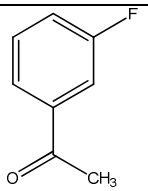
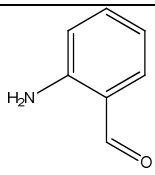
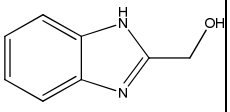
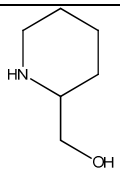
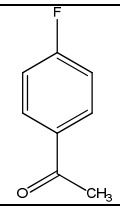
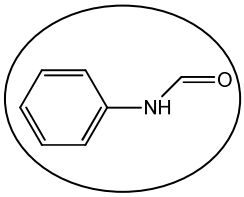
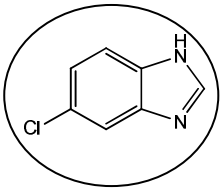
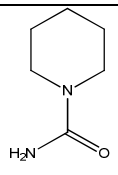
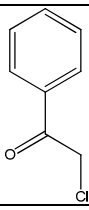
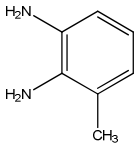
Pramod C. Nair[†], Alpeshkumar K. Malde[†], Nyssa Drinkwater[‡§] and Alan E. Mark*[†‡]

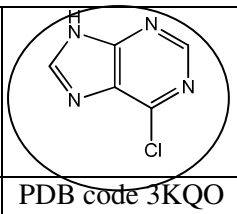
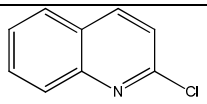
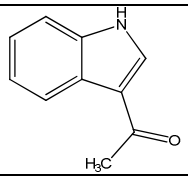
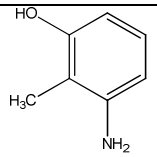
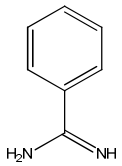
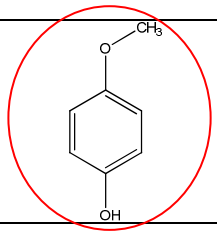
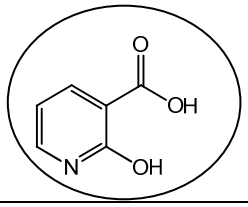
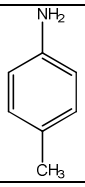
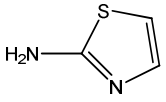
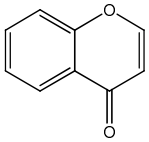
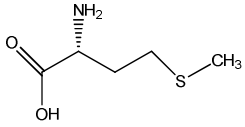
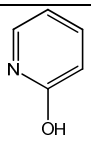
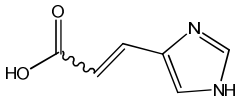
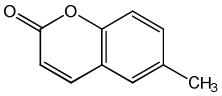
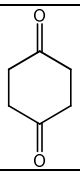
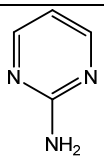
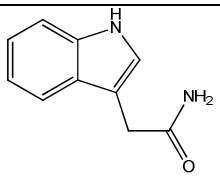
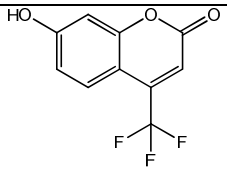
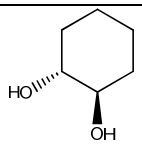
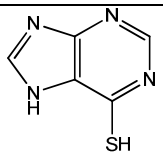
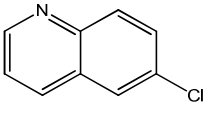
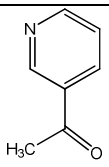
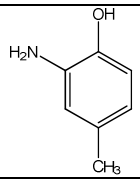
[†]School of Chemistry and Molecular Biosciences (SCMB), [‡]Institute for Molecular Bioscience (IMB), The University of Queensland (UQ), St Lucia Campus, Brisbane, QLD 4072 Australia.

[§]Randall Division of Cell and Molecular Biophysics, King's College London, London, UK SE1 1UL.

Table S1 List of fragment cocktails used for the X-ray screening by Drinkwater et al. The hits identified by the fragment based X-ray screen (FBS-X) and electron spray ionization Fourier transform mass spectrometry (ESI-FTMS) are encircled in black and red respectively. The PDB codes for FBS-X hits are given. The 4 fragments of each cocktail are referred as **F1:X**, **F2:X**, **F3:X** and **F4:X**, where X refers to the number of the cocktail.

	F1	**F2**	**F3**	**F4**	
					
Cocktail 1		PDB code 3KPU (Compound 1) K_d 690 ± 220			
					
Cocktail 2	PDB code 3KPV (Compound 2) K_d 180 ± 90				
					
Cocktail 3	PDB code 3KPW (Compound 3) K_d 14 ± 2.0				
					
Cocktail 4			PDB code 3KPY (Compound 4) Not detected by ITC		

					
		PDB code 3KQM Compound 5 K_d 170 ± 16			
Cocktail 5					
					
Cocktail 6	PDB code 3KQS Compound 6 K_d 6.3 ± 0.02				
					
Cocktail 7	PDB code 3KQT Compound 7 K_d 4.6 ± 0.69				
					
Cocktail 8				PDB code3KQV Compound 8 Not detected by ITC	
					
Cocktail 9	PDB code Compound 9 K_d 97 ± 5.9				

					
Cocktail 10	PDB code 3KQO Compound 10 K_d 140 ± 27				
					
Cocktail 11	PDB code 3KQP Compound 11 K_d 380 ± 44			PDB code 3KQQ Compound 12 Not detected by ITC	
					[a]False Positive
Cocktail 12					
					[a]False Positive
Cocktail 13					
					[a]False Positive
Cocktail 14					
					[a]False Positive
Cocktail 15					

[a] False positives refer to cases in which density was observed in the initial X-ray screen but no density was observed when the fragments were soaked into the crystals individually.

Table S2 Refinement statistics for the models with and without fragments.

Parameter	PDB 3KPY	Without fragment	With fragment F3	With fragments F1 & F2	With fragments F1 (2 binding modes) & F2
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit cell					
a,b (Å)	93.9	93.9	93.9	93.9	93.9
c (Å)	188.6	188.6	188.6	188.6	188.6
α,β,γ (◦)	90	90	90	90	90
No. unique reflections	30511	30487	30487	30487	30487
Resolution (Å)	45.56–2.40	45.56–2.40	45.56–2.40	45.56–2.40	45.56–2.40
Refinement					
R_{cryst}/R_{free} (%)	20.6/27.1	21.5/27.5	21.3/27.5	21.3/27.2	21.9/27.5

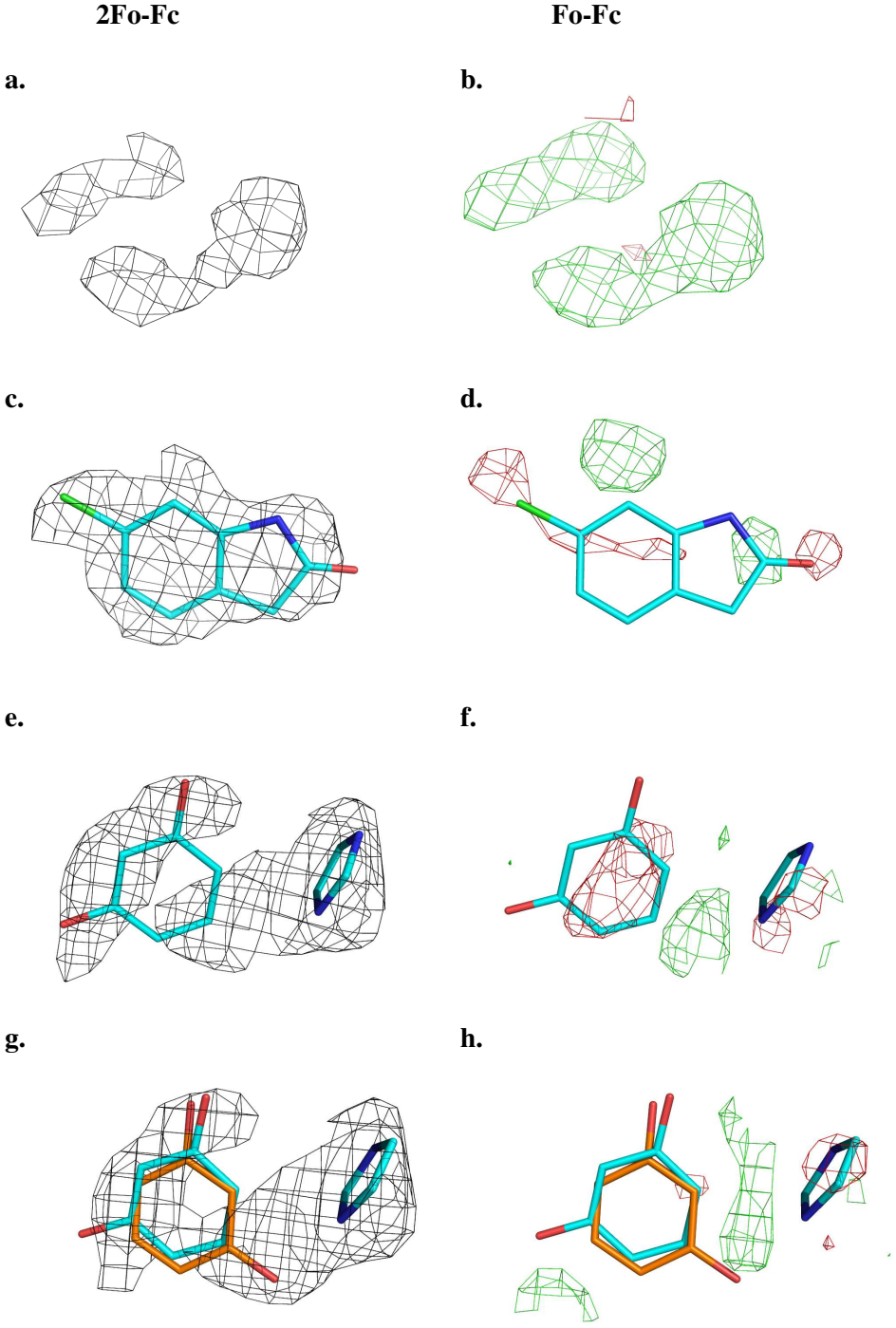


Figure S1. The electron density maps for PDB 3KPY (chain B) refined without any fragment (**a** and **b**), in presence of fragment F3 (**c** and **d**), in presence of fragments F1 and F2 (**e** and **f**), in presence of F2 and two binding modes of F1 (cyan and orange) each with partial occupancies (**g** and **h**). 2Fo-Fc maps (black) are contoured at 1.0σ level. Fo-Fc maps are contoured at +3.0σ level (positive 'green' density) and -3.0σ level (negative 'red' density).

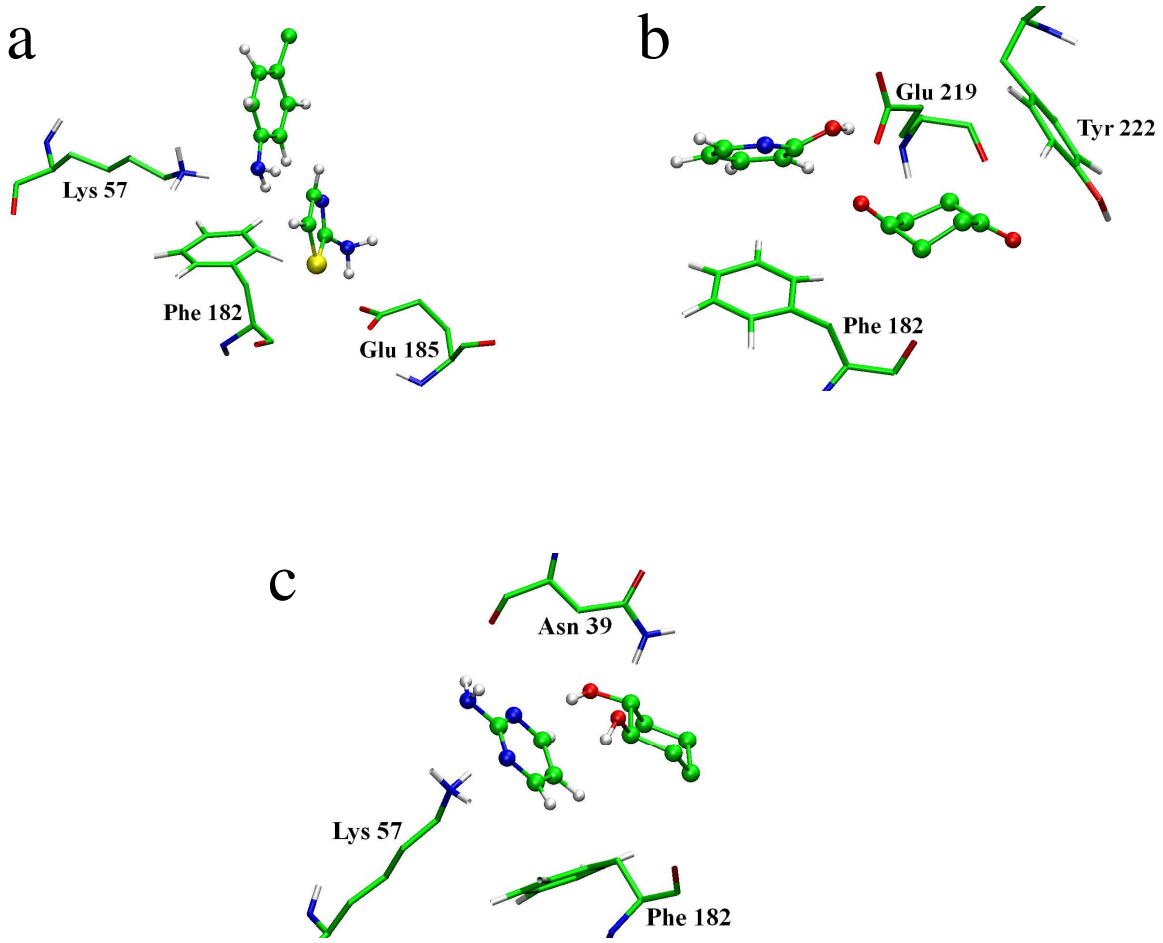


Figure S2. Combinations of fragments taken from cocktails that were considered to be false positives and that remained stable during 5ns simulations. The fragments and the residues are represented as ball and stick and stick models respectively. (a) **F1:12** and **F2:12** (b) **F1:13** and **F4:13** (c) **F1:14** and **F4:14**

Computational Details

All molecular dynamics (MD) simulations were performed using the GROMOS96[1] simulation package in conjunction with the GROMOS 53A6 force field.[2] The initial structure of human phenylethanolamine N-methyltransferase (PNMT) complexed with the cofactor *S-adenosyl-L-homocysteine* (SAH) and fragment **F1** was taken from the pdb entry 3KPV. The topologies of the fragment molecules (Figure 1) were generated using the 'Automated Topology Builder' (ATB) and are available from the ATB repository (http://compbio.biosci.uq.edu.au/atb)[3]: Fragment **F1:4** (RNME F030); Fragment **F2:4** (RNME F135); Fragment **F3:4** (RNME F249); and Fragment **F4:4** (RNME F345). Simulations of the ligands free in solution were performed by placing a given ligand in a periodic rectangular box containing 853 simple point charge (SPC) water molecules. The simulations in the protein were carried in a truncated octahedral box with ~7500 simple point charge (SPC) water molecules. For the systems in which the fragment was bound to the protein, the configuration of the solvent was relaxed by performing a steepest descent minimization in which the protein and ligand atoms were positionally restrained to their initial positions using a harmonic interaction potential with a force constant of 2.0 x 10^3 kJ/mol/nm^2. The system was then further equilibrated by performing a 200 ps simulation, with the heavy atoms of the protein positionally restrained, before a series of unrestrained molecular dynamics (MD) simulations were commenced. All the simulations were performed at constant temperature (298 K) and pressure (1 atm). This was achieved using a Berendsen thermostat with a coupling time of 0.1 ps and a Berendsen barostat with a coupling time of 0.5 ps.[4] The isothermal compressibility was set to 4.575×10^{-4} kJ/mol/nm^3. Nonbonded interactions were calculated using a twin-range cutoff. Interactions within the short-range cutoff of 0.8 nm were updated every time step. Interactions within the longer-range cutoff of 1.4 nm were updated every 5 time steps together with the pairlist. To correct for the truncation of electrostatic interactions beyond the 1.4 nm long-range

cutoff a reaction field correction was applied using an effective dielectric (ε) of 61.[5] The equations of motion were integrated using the leapfrog scheme with a 2 fs time step. Initial velocities at a given temperature were taken from a Maxwell-Boltzmann distribution. The lengths of all bonds were constrained to ideal values using the SHAKE[6] algorithm with a geometric tolerance of 0.0001.

Crystallographic refinement was performed using Phenix version 1.7.1-743.[7] The electron density maps were built using the structure factors and the co-ordinates deposited in the protein data bank (PDB), PDB code 3KPY. The topologies of the fragment molecules used for the X-ray crystallographic refinement were obtained using the program eLBOW[8] of Phenix.[7]

Free energy calculations

The change in Gibbs free energy (ΔG) associated with mutating the fragment from one form to another in water and in the protein was determined using the coupling parameter approach in conjunction with the thermodynamic integration formula (1),[9]

$$\Delta G_{0 \rightarrow 1} = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \qquad (1)$$

where λ = 0 corresponded to the initial state of the system and λ = 1 corresponded to the final state of the system. *H* is the Hamiltonian of the system and the brackets $\langle ... \rangle_{\lambda}$ correspond to an average over an equilibrium ensemble at λ. The relative free energy of binding $\Delta\Delta G$ was determined from the difference in the change in free energy of performing the same mutation free in solution and bound to the protein. Equation (1) was integrated by performing separate simulations at a series of 21 equally spaced λ points (0.0, 0.05, 0.1, 0.15, 0.2 … 0.8, 0.85 0.9, 0.95, 1.0). Ten additional λ points (0.025, 0.075, 1.25, 1.75, 2.25, 2.75, 0.825, 0.875, 0.925, 0.975) were included to smooth the integrand for both the bound and unbound states. For the mutations in water a 1 ns simulation was performed at each λ value. For the mutations in the protein, the system was first equilibrated for 0.5 ns, and 3.0 ns of sampling used to provide an estimate of $\langle \partial H/\partial \lambda \rangle_{\lambda}$. To determine the degree of convergence

thermodynamic cycles wherein the molecules were transformed from one to another in circular path in water, and when bound to the protein, were constructed. To transform the ligand from one state to another a dual topology approach was used. To remove a molecule the non-bonded and the columbic interactions of all atoms in the molecule were scaled to zero. The degree of convergence was also checked by performing the forward and backward mutations. To prevent numerical instabilities as atoms were created or destroyed the soft-core potential as described by Beutler et al.[10] was used with α_{LJ} = 0.5, α_C =0.5 nm^2. The area beneath the curve in (1) was estimated using a trapezoidal approximation. The statistical error at each λ-point was estimated by calculating the standard error for each value of λ and the total error obtained by integrating over the path. To calculate the free energy for different mutations in PNMT and water, a weak (10 kJ/mol/nm^2) harmonic distance restraint (0.5nm) was applied between the C atom of fragment **F1** and N atom of fragment **F2**. Note, this restraint was maintained in all legs of the thermodynamic cycles calculated in water and PNMT and thus makes no net contribution to the relative free energy.

References

1. Christen, M.; Hünenberger, P. H.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. F. The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry* **2005**, 26, 1719-1751.

2. Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, 25, 1656-1676.

3. Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated force field Topology Builder (ATB) and repository: version 1.0. *J. Chem. Theory Comput.* **2011**, 7, 4026-4037.

4. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, 81, 3684-3690.

5. Heinz, T. Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* **2001**, 115, 1125.

6. Ryckaert, J.; Ciccotti, G.; Berendsen, H. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, 23, 327-341.

7. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **2010**, 66, 213-221.

8. Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica Section D* **2009**, 65, 1074-1080.

9. Mark, A. E. Free energy perturbation calculations. In *Encyclopedia of Computational Chemistry*, Schleyer, P. v. R., Ed. John Wiley & Sons: 1998; pp 1070-1083.

10. Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **1994**, 222, 529-539.