Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information (ESI) for: Solvent Effects on Ligand Binding to a Serine Protease

Srinivasa M. Gopal,[†,¶] Fabian Klumpers,[‡,¶] Christian Herrmann,[‡] and Lars V. Schäfer[*,†]

†*Center for Theoretical Chemistry, Ruhr-University Bochum, D-44780 Bochum, Germany*

‡*Physical Chemistry I, Ruhr-University Bochum, D-44780 Bochum, Germany*

¶*Contributed equally to this work*

E-mail: lars.schaefer@ruhr-uni-bochum.de

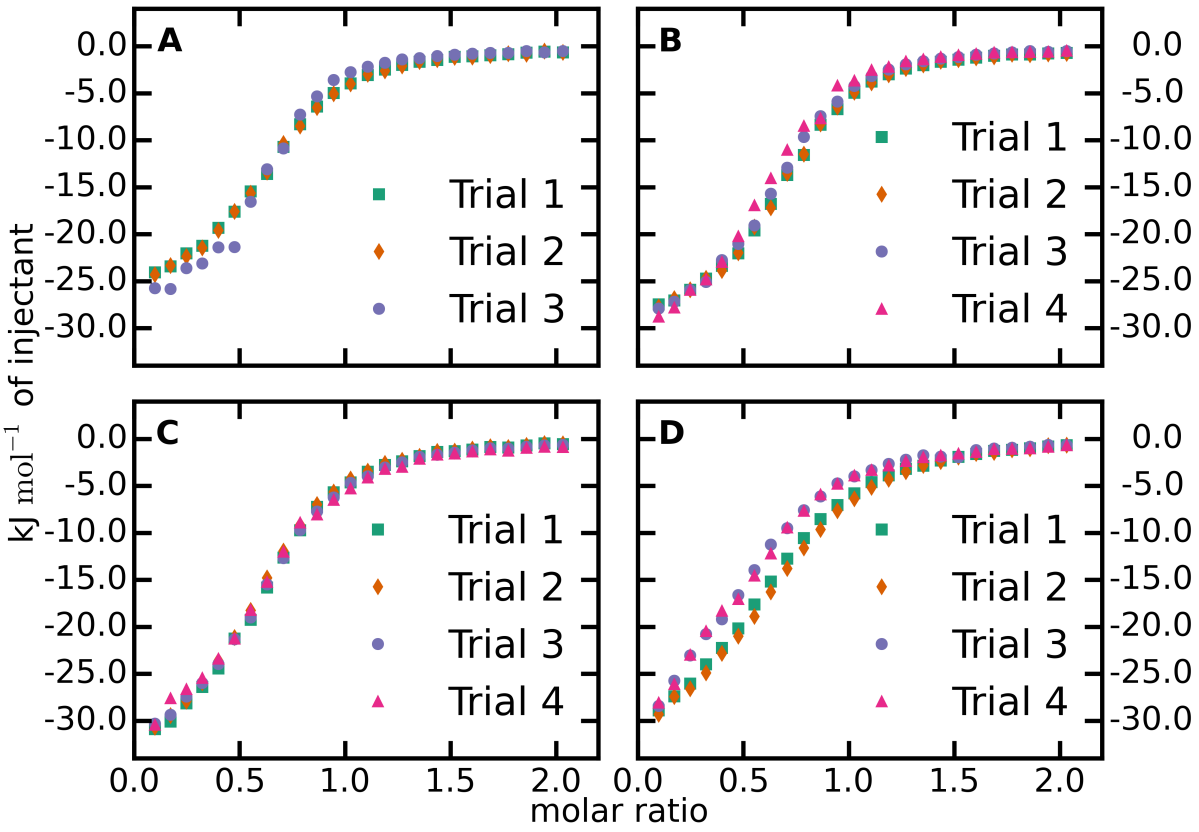


Fig. S1 Independent ITC titrations in 0% (A), 10% (B), 20% (C), and 30% (D) methanol/water (v/v) mixtures.

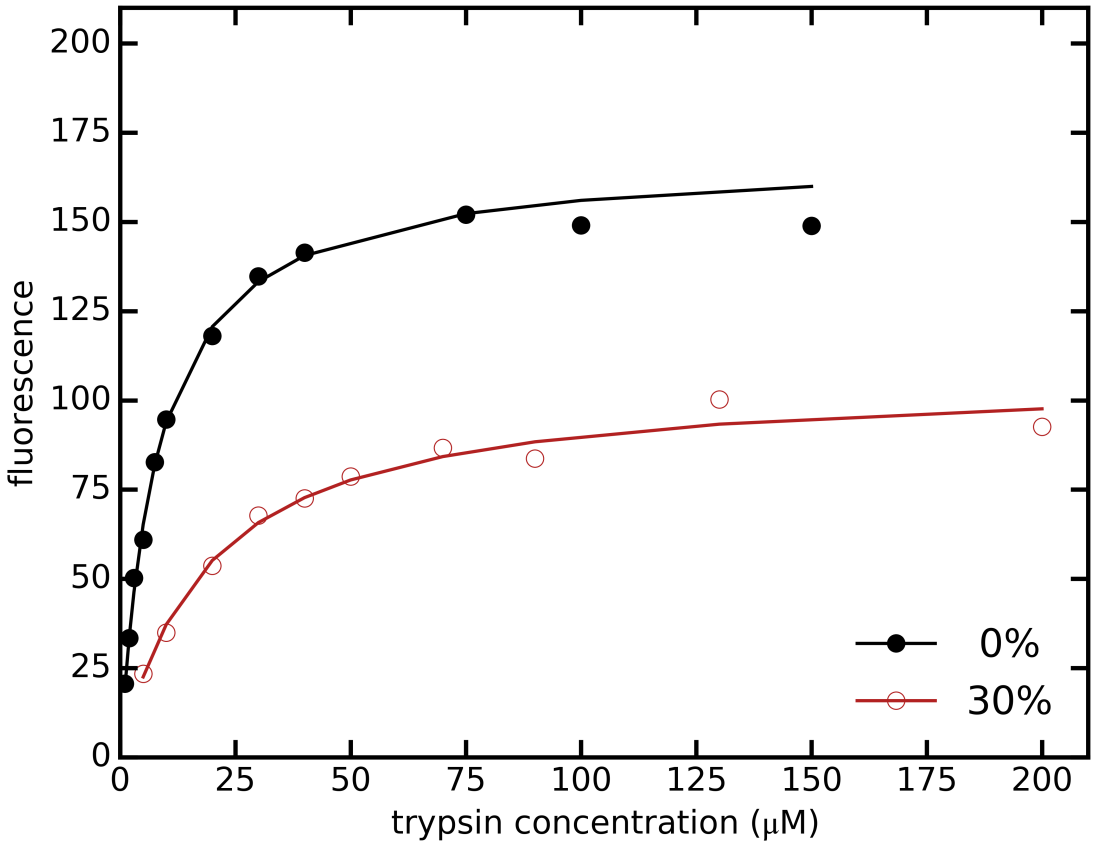


Fig. S2 Fluorescence titrations in 0% and 30% methanol/water (v/v) mixtures.

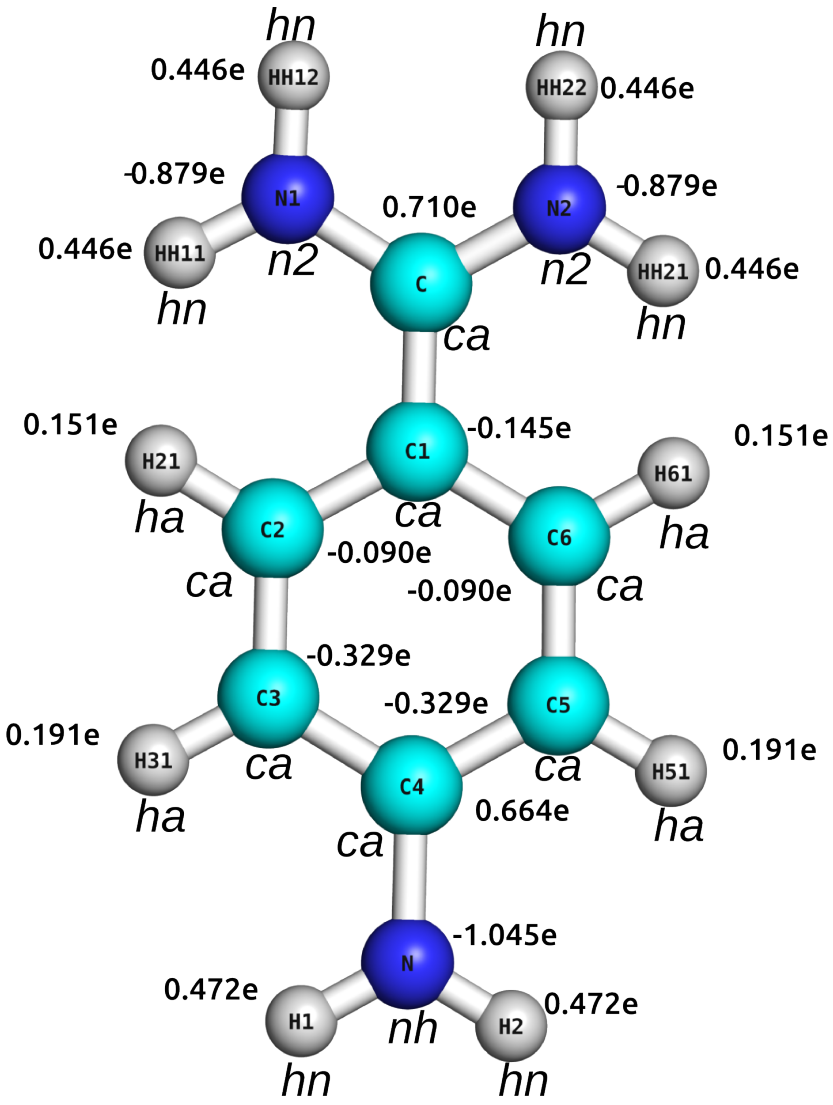


Fig. S3 Partial charges and GAFF atom types of PAB.

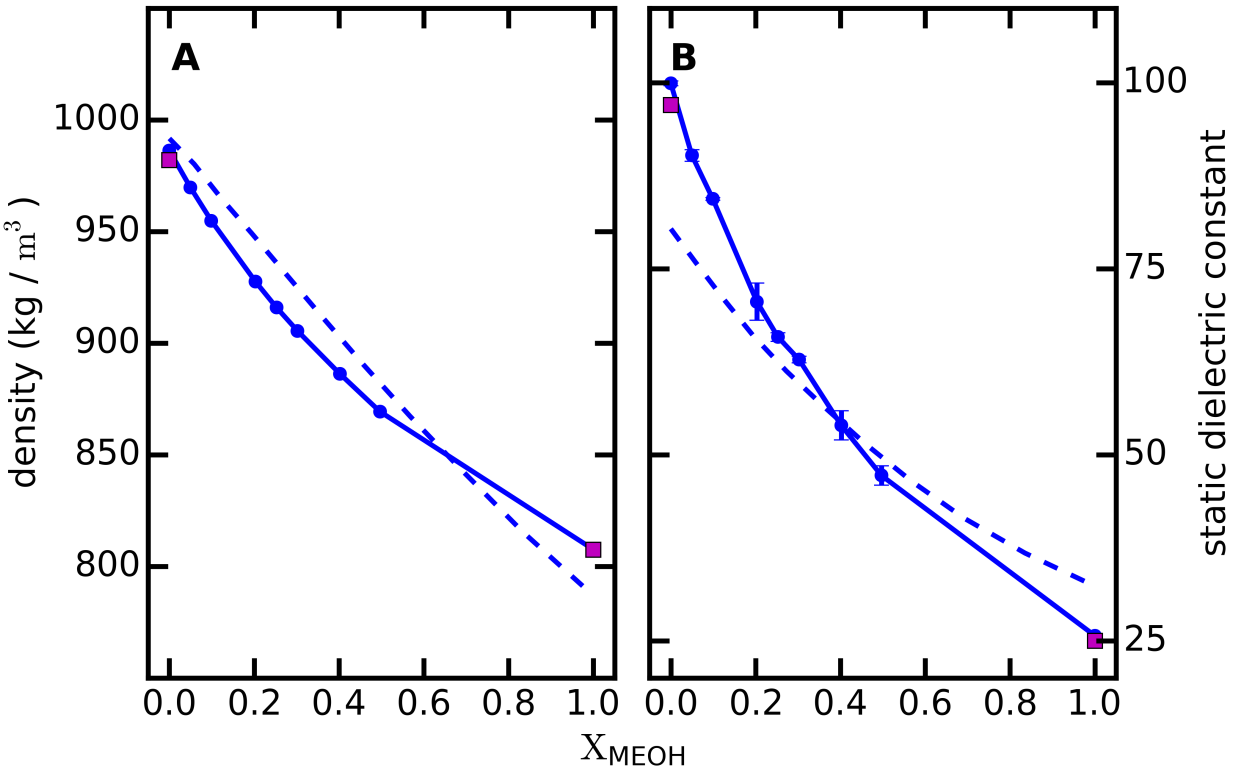


Fig. S4 Comparison of simulated (solid lines with symbols) and experimental[S1,S2] (dashed lines) properties of water/methanol mixtures. (A) Density. (B) Static dielectric constant. The reference values from the literature for TIP3P water[S3] and GAFF methanol[S4] are also shown (magenta squares). The static dielectric constant was obtained from the Neumann dipole fluctuation formula.[S5] MD simulations of the binary mixtures were 100 ns in length.

Tab. S1 Overview of equilibrium MD simulations performed in this work. Independent simulations were initiated with different random seeds for generating the initial atomic velocities according to a Maxwell-Boltzmann distribution at 298 K.

system	# independent simulations	duration (ns)	# water	# methanol	# Na^+	# Cl^-	box (nm)
			Trypsin–PAB complex				
0%	3	500	11891	-	36	44	7.31
10%	1	200	9881	1011	33	41	7.33
20%	1	200	7666	1801	29	37	7.23
30%	3	500	6397	2338	26	34	7.21
			Apo trypsin				
0%	3	250	11891	-	36	43	7.30
10%	1	200	9881	1011	33	40	7.33
20%	1	200	7666	1801	29	36	7.23
30%	3	250	6397	2338	26	33	7.20
			Free PAB				
0%	1	200	11891	-	36	37	7.12
10%	1	200	9881	1011	33	34	7.15
20%	1	200	7666	1801	29	30	7.05
30%	1	200	6397	2338	26	27	7.03

Free Energy of Binding

According to the thermodynamic cycle shown in Figure S5, the absolute free energy of binding, ΔG_{bind}, is given as

$$\Delta G_{PL}^{all} = \Delta G_{PL} + \Delta G_{PL}^{res} \tag{S1}$$

$$\Delta G_{bind}^{raw} = -(\Delta G_{PL}^{all} + \Delta G_{solv} + \Delta G^{res}) + \Delta G^{symm} \tag{S2}$$

$$\Delta G_{bind} = \Delta G_{bind}^{raw} + \Delta G_{bind}^{corr} \tag{S3}$$

where

- ΔG_{PL} is the free energy of binding a restrained ligand.
- $\Delta G_{\mathrm{PL}}^{\mathrm{res}}$ is the free energy of restraining the ligand.
- ΔG_{solv} is the free energy of solvation of the ligand.
- ΔG^{res} is the free energy of releasing the ligand to the standard volume at 1 M concentration ($V_0 = 1.66$ nm^3).
- $\Delta G^{\mathrm{symm}} = -k_B T \ln 2$ is the correction due to the two-fold rotational symmetry of the ligand. This is necessary, because rotations around this axis are not sampled in the complex (bound state), whereas they are sampled in the free (unbound) state.
- $\Delta G_{\mathrm{bind}}^{\mathrm{corr}}$ is the correction due to the net-charged periodic system (see below).

ΔG^{res} is calculated analytically following the approach of Boresch and co-workers[S7],

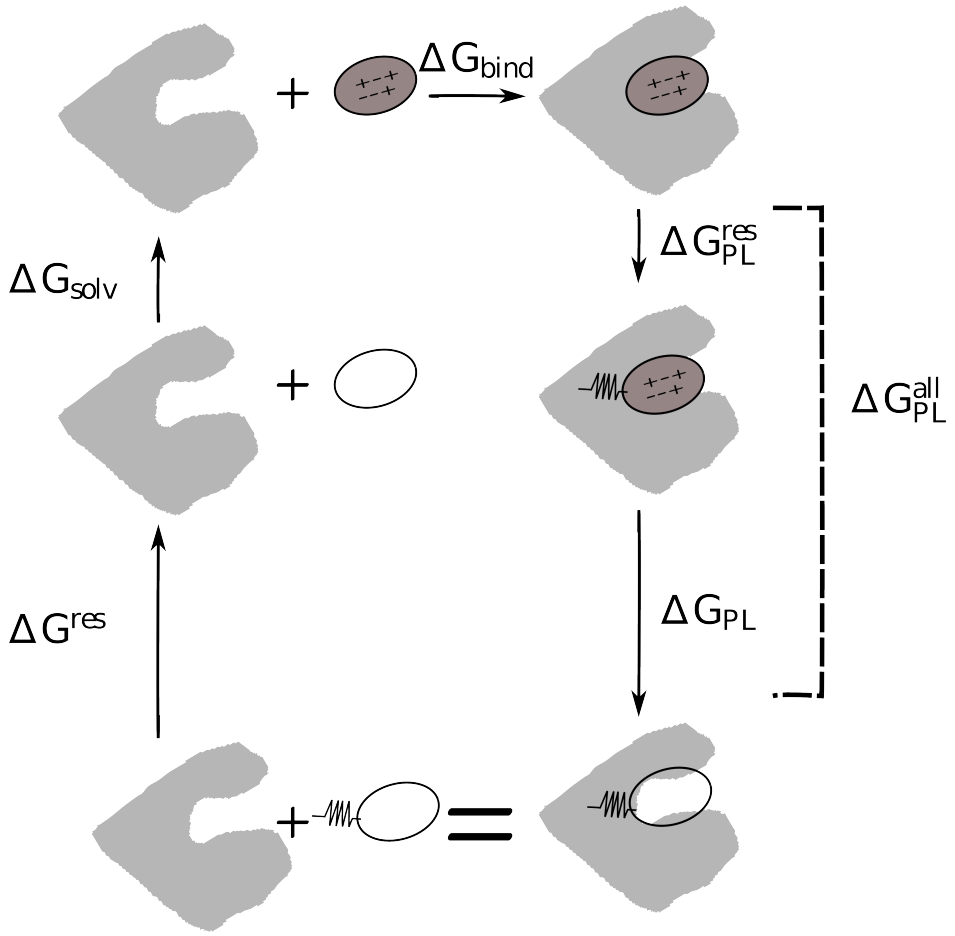


Fig. S5 Thermodynamic cycle for ligand binding[S6]. The steps involved are (1) restraining the ligand in the binding site ($\Delta \mathrm{G}_{\mathrm{PL}}^{\mathrm{res}}$), (2) decoupling of the non-bonded interactions between the restrained ligand and the environment ($\Delta \mathrm{G}_{\mathrm{PL}}$), (3) analytical correction for releasing the restraints ($\Delta \mathrm{G}^{\mathrm{res}}$), (4) solvating the ligand in the bulk ($\Delta \mathrm{G}_{\mathrm{solv}}$).

$$\Delta G^{res} = -k_B T \left(\frac{8\pi^2 V_0}{r^2 \sin\theta_A \sin\theta_B} \frac{K_{r_A} K_{\theta_A} K_{\theta_B} K_{\phi_A} K_{\phi_B} K_{\phi_C}}{(2\pi k_B T)^3} \right) \tag{S4}$$

where one distance (r_A), two angles (θ_A and θ_B) and three torsions (ϕ_A, ϕ_B, ϕ_C) were used to restrain PAB relative to the protein. Three atoms of Asp189 (C_γ, C_β, C_α) and PAB (C, C_4, C_6) are involved (Figure S6). The equilibrium values of the above coordinates, as obtained from our simulations, are given in Table S2.

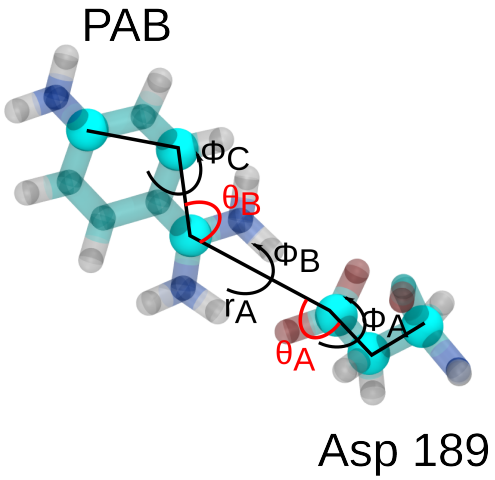


Fig. S6 Restraints used in the ligand binding free energy calculations.

Tab. S2 Average values of bond, angle, and dihedral coordinates (Figure S6) at various methanol concentrations.

System	r_A (nm)	$\theta_A(°)$	$\theta_B(°)$	$\phi_A(°)$	$\phi_B(°)$	$\phi_C(°)$
X-ray	0.39	133.3	136.7	138.9	12.5	-143.1
X-ray (minimized)	0.39	132.2	136.9	141.7	15.4	-143.0
0%	0.39	133.8	136.3	149.3	3.0	-130.0
10%	0.39	133.3	135.0	148.9	6.3	-131.7
20%	0.39	133.1	134.3	150.2	5.7	-130.5
30%	0.39	134.2	133.1	148.2	4.8	-132.7

$\Delta G^{res} = -28.9\,kJ\,mol^{-1}$ results from the force constants of $K_{r_A} = 4184\ kJ\,mol^{-1}\,nm^{-2}$, $K_{\theta_A} = K_{\theta_B} = 41.8\ kJ\,mol^{-1}$, and $K_{\phi_A} = K_{\phi_B} = K_{\phi_C} = 41.8\ kJ\,mol^{-1}$, and the equilibrium distances and angles listed above.

Comparison of ΔG_{PL}^{all} obtained from BAR, TI, and MBAR

Tab. S3 ΔG_{PL}^{all} from BAR. The standard deviation (from the three independent sets of simulations) is indicated in brackets. Units are kJ mol^{-1}.

System	run1	run2	run3	ΔG_{PL}^{all}
0%	272.9	268.8	269.4	270.4 (1.8)
10%	279.3	274.9	280.0	278.1 (2.3)
20%	288.9	285.9	283.8	286.2 (2.1)
30%	287.7	295.9	295.6	293.1 (3.7)

Tab. S4 ΔG_{PL}^{all} from TI. The standard deviation (from the three independent sets of simulations) is indicated in brackets. Units are kJ mol^{-1}.

System	run1	run2	run3	ΔG_{PL}^{all}
0%	272.0	267.9	268.1	269.3 (1.9)
10%	278.4	275.5	280.9	278.3 (2.2)
20%	288.6	286.0	283.1	285.9 (2.2)
30%	288.3	295.4	295.8	293.2 (3.4)

Tab. S5 ΔG_{PL}^{all} from MBAR. The standard deviation (from the three independent sets of simulations) is indicated in brackets. Units are kJ mol^{-1}.

System	run1	run2	run3	ΔG_{PL}^{all}
0%	272.0	267.9	268.1	269.3 (1.9)
10%	275.3	272.6	278.3	275.4 (2.3)
20%	286.0	286.2	284.1	285.4 (0.9)
30%	292.0	296.5	299.0	295.8 (2.9)

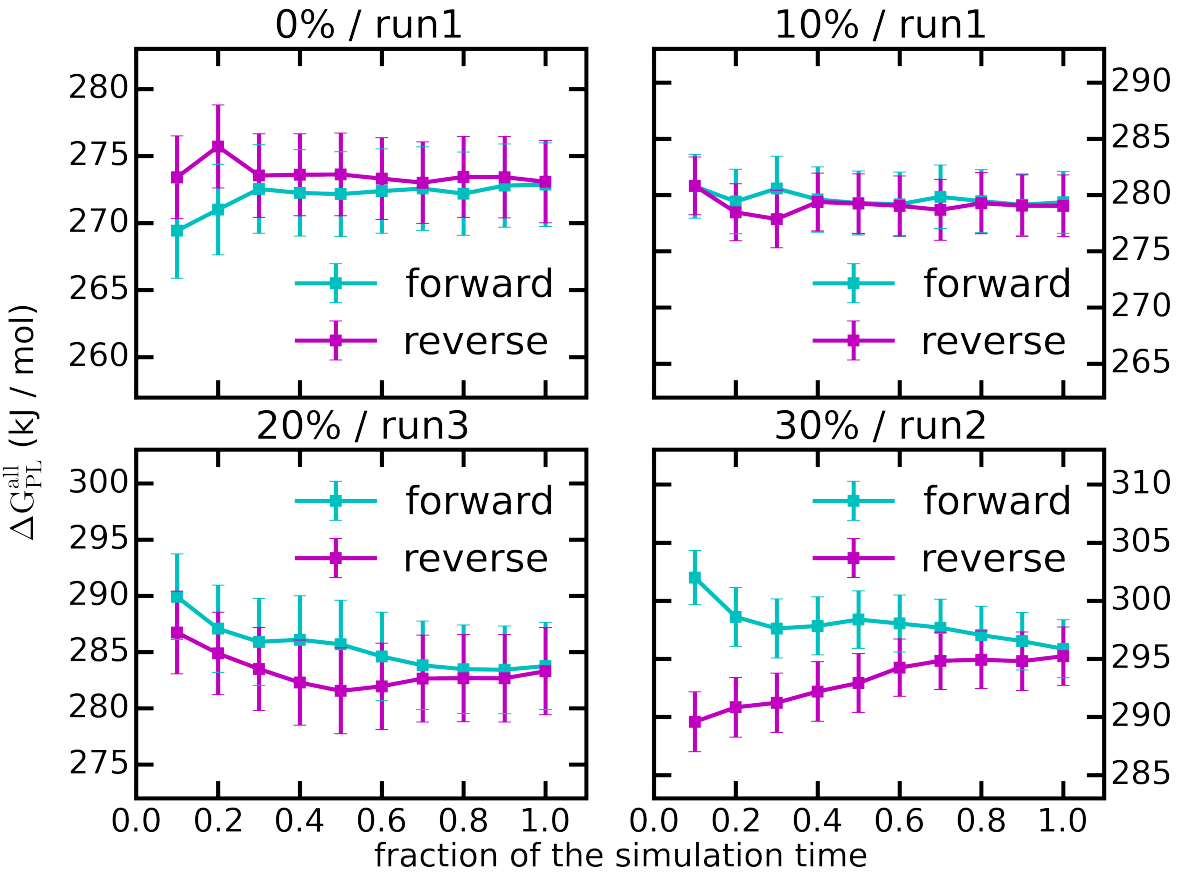


Fig. S7 Convergence of ΔG^{all}_{PL} estimated from forward and backward time series as suggested by Klimovich et al.[S8]

Free Energy Corrections for Charged Periodic Systems

Corrections to Free Energy of Binding

The following, we briefly summarize the corrections for charged periodic systems, as outlined by Rocklin and co-workers[S9]. We adopt the nomenclature from the original article.

The free energy directly obtained from the MD simulations under periodic boundary conditions (PBC), ΔG^{raw}_{bind}, suffers from finite-size effects, i.e., includes a box-size dependent contribution that needs to be corrected. According to Rocklin et al., this correction involves two parts, an analytical correction (ANA) and a discrete solvent correction (DSC) specific to the explicit-solvent MD approach.

$$\Delta G_{bind} = \Delta G^{raw}_{bind}(L) + \Delta\Delta G_{ANA}(L) + \Delta\Delta G_{DSC}(L) \tag{S5a}$$

$$\Delta\Delta G_{ANA}(L) = \Delta\Delta G_{NET}(L) + \Delta\Delta G_{USV}(L) + \Delta\Delta G_{RIP}(L) + \Delta\Delta G_{EMP}(L) \tag{S5b}$$

Here,

- L is the length of the box vector.
- $\Delta\Delta G_{NET}(L)$ corrects for net charge interactions due to PBC,

$$\Delta\Delta G_{NET} = -\frac{\xi_{LS}}{8\pi\epsilon_0}\left[\frac{(Q_P+Q_L)^2 - Q_P^2}{L}\right], \tag{S6}$$

  where ξ_{LS} = -2.837 is the cubic lattice-sum integration constant, and Q_P and Q_L are the net charges of the protein and ligand, respectively.

- $\Delta\Delta G_{USV}(L)$ corrects the under solvation due to net charges,

$$\Delta\Delta G_{USV} = -\left(1 - \frac{1}{\epsilon_S}\right)\Delta\Delta G_{NET}, \tag{S7}$$

where ϵ_S is the static dielectric constant of the solvent. For the water/methanol mixtures, we used the ϵ_S of the mixture (Figure S4).

- $\Delta\Delta G_{RIP}(L)$ accounts for the fact that protein and ligand are not point charges (obtained from a non-periodic Poisson-Boltzmann calculation),

$$\Delta\Delta G_{RIP} = \frac{[(I_P + I_L)(Q_P + Q_L) - I_P Q_P]}{L^3}, \tag{S8}$$

where I_P and I_L are the residual integrated potential (RIP) of protein and ligand, respectively. Since counterions were used in the MD simulations, $Q_P = 0$ was used for calculating $\Delta\Delta G_{RIP}$.

The RIPs for the protein (I_P) and the ligand (I_L) were calculated using the protocol and scripts provided by Rocklin and co-workers[S9] using the Adaptive-Poisson-Boltzmann-Software (APBS).[S10] The solute dielectric constant was set to 1 and solvent dielectric constant varied from 97 (pure water) to 63 (30% methanol), see Fig. S4. The grid spacing and length were 0.05 nm and 12.8 nm, respectively. The solvent probe radius was 0.14 nm. No ions were used in the PB calculations. Multiple Debye-Hückel boundary conditions were used. Other APBS options were set to the recommended values (quartic B-spline discretization (spl4), harmonic averaging (spl4), surface density 4000 points/nm^2). The calculations were carried out for 500 snapshots taken from the

MD simulations for each solvent mixture.

- $\Delta\Delta G_{DSC}$ corrects for the discrete nature of the solvent,

$$\Delta\Delta G_{DSC} = \Delta\Delta G_{DSI} + \Delta\Delta G_{DSF}(L) \tag{S9}$$

$$\Delta\Delta G_{DSF} = -\Delta\Delta G_{DSI}\frac{V_c}{L^3} \tag{S10}$$

$$\Delta\Delta G_{DSI} = -\frac{\gamma_S \rho_S Q_L}{6\epsilon_0 M_S} \tag{S11}$$

where $\Delta\Delta G_{DSI}$ is an infinite-system discrete solvent correction term, γ_S is the quadrupole-moment trace of the solvent model, V_c the excluded volume of the solute, and M_S and ρ_S are the molecular mass and density of the solvent, respectively. For TIP3P water, $\Delta\Delta G_{DSI}$ is $-74.1\,\mathrm{kJ\,mol^{-1}}$.

For the water/methanol mixtures, $\Delta\Delta G_{DSC}$ can be estimated using the following relation,

$$\Delta\Delta G_{DSC} = f_M \Delta\Delta G_{DSI,M} + f_W \Delta\Delta G_{DSI,W} + \Delta\Delta G_{DSF}(L) \tag{S12}$$

$$\Delta\Delta G_{DSF}(L) = -f_M \Delta G_{DSI,M}\frac{V_c}{L^3} - f_W \Delta G_{DSI,W}\frac{V_c}{L^3} \tag{S13}$$

where f_W and f_M are molar fractions of water and methanol respectively. For methanol, in the orientational disorder limit (ODL), γ_S is approximately 0.0062 e nm^2 (with respect to the center of mass of methanol). Using this value and $\rho_S = 807.6$ kg m^{-3} for methanol, $\Delta\Delta G_{DSI,M}$ is $-27.5\,\mathrm{kJ\,mol^{-1}}$. The $\Delta\Delta G_{DSF}(L)$ values of the protein in

the solvent mixtures are listed below.

$\Delta\Delta G_{DSF}(L)$ for methanol/water mixtures. The excluded volume of the protein is about 42.0 nm^3.

System	f_W	f_M	$\Delta\Delta G_{DSF}(L)$ (kJ mol^{-1})
0%	1.00	-	8.0
10%	0.91	0.09	7.5
20%	0.81	0.19	7.3
30%	0.63	0.37	6.9

- Adding $\Delta\Delta G_{EMP}(L)$ to the analytical correction makes it exact for a special case of a single point charge in a spherical cavity,

$$\Delta\Delta G_{EMP} = -\frac{1}{8\pi\epsilon_0}\frac{16\pi^2}{45}\left(1-\frac{1}{\epsilon_S}\right)\left[(Q_P+Q_L)^2-Q_P^2\right]\frac{R_L^5}{L^6}, \tag{S14}$$

where R_L is the effective radius of the ligand within the protein-ligand complex, evaluated from the residual integrated potential of the ligand (I_L). $\Delta\Delta G_{EMP}$ is ignored in the current context, as it depends on inverse sixth power of box-size and is thus smaller than 0.4 kJ/mol for the boxes used in ($L \geq 7$ nm).

Tab. S6 Corrections to ΔG_{PL}^{all}. Uncertainties are indicated in brackets. Units are kJ mol^{-1}.

System	$\Delta\Delta G_{NET}$	$\Delta\Delta G_{USV}$	$\Delta\Delta G_{RIP}$	$\Delta\Delta G_{DSF}$	$\Delta\Delta G_{DSI}$	$\Delta G_{PL}^{all\ corr}$
0%	27.6	-27.4	9.0 (0.2)	8.0	-74.1	-56.9 (0.2)
10%	27.5	-27.2	10.6 (0.3)	7.5	-69.9	-51.5 (0.3)
20%	28.0	-27.6	12.6 (0.3)	7.3	-65.2	-44.9 (0.3)
30%	28.0	-27.6	14.6 (0.3)	6.9	-61.5	-39.6 (0.3)

Corrections to Free Energy of Solvation

Tab. S7 Corrections to ΔG_{solv}. Units are kJ mol^{-1}.

System	$\Delta\Delta G_{NET}$	$\Delta\Delta G_{USV}$	$\Delta\Delta G_{RIP}$	$\Delta\Delta G_{DSF}$	$\Delta\Delta G_{DSI}$	ΔG_{solv}^{corr}
0%	27.6	-27.4	-0.01	0.1	-74.1	-73.8
10%	27.5	-27.2	-0.01	0.1	-69.9	-69.5
20%	28.0	-27.6	-0.01	0.1	-65.2	-64.7
30%	28.0	-27.6	-0.01	0.1	-61.5	-61.0

Comparison of ΔG_{solv} obtained from TI, BAR, and MBAR

Tab. S8 Comparison of ΔG_{solv} obtained from BAR, TI, and MBAR. The ΔG_{solv}^{raw} is the average of two independent sets of simulations. Uncertainties are indicated in brackets. Units are kJ mol^{-1}.

System	ΔG_{solv}^{raw}			ΔG_{solv}^{corr}	ΔG_{solv}		
	BAR	TI	MBAR		BAR	TI	MBAR
0%	-195.5 (0.1)	-195.7 (0.1)	-195.5 (0.1)	-73.8	-269.3 (0.1)	-269.5 (0.1)	-269.3 (0.1)
10%	-203.9 (0.2)	-204.9 (0.3)	-203.9 (0.4)	-69.5	-273.4 (0.2)	-274.4 (0.3)	-273.4 (0.4)
20%	-210.8 (0.3)	-210.9 (0.3)	-210.8 (0.3)	-64.7	-275.5 (0.3)	-275.6 (0.3)	-275.5 (0.3)
30%	-215.6 (0.3)	-215.9 (0.1)	-215.8 (0.1)	-61.0	-276.6 (0.3)	-276.9 (0.1)	-276.8 (0.1)

Comparison of ΔG_{bind} obtained from BAR, TI, and MBAR

The correction scheme described above corresponds to a process of binding a ligand from vacuum to the binding site. However, in the thermodynamic cycle (Figure S5), the reaction proceeds in the opposite direction. Thus, the signs of the corrections described in Table S6 have to be reversed. The final correction, $\Delta G_{bind}^{corr} = -(-\Delta G_{PL}^{all\ corr} + \Delta G_{solv}^{corr})$, is given in Table S9.

Tab. S9 Final correction for the entire binding cycle, $\Delta G_{bind}^{corr} = -(-\Delta G_{PL}^{all\ corr} + \Delta G_{solv}^{corr})$, calculated from Tabs. S6 and S7. Uncertainties are indicated in brackets. Units are $kJ\ mol^{-1}$.

System	$\Delta G_{PL}^{all\ corr}$	ΔG_{solv}^{corr}	ΔG_{bind}^{corr}
0%	-56.9 (0.2)	-73.8	16.9 (0.2)
10%	-51.5 (0.3)	-69.5	18.0 (0.3)
20%	-44.9 (0.3)	-64.7	19.8 (0.3)
30%	-39.6 (0.3)	-61.0	21.4 (0.3)

Tab. S10 Results from BAR. ΔG^{res} is $-28.9\,kJ\,mol^{-1}$. ΔG^{symm} is $-1.7\,kJ\,mol^{-1}$. Uncertainties are indicated in brackets. Units are $kJ\ mol^{-1}$.

System	ΔG_{solv}^{raw}	ΔG_{PL}^{all}	ΔG_{bind}^{raw}	ΔG_{bind}^{corr}	ΔG_{bind}
0%	-195.5 (0.1)	270.4 (1.8)	-47.7 (1.8)	16.9 (0.2)	-30.8 (1.8)
10%	-203.7 (0.2)	278.1 (2.3)	-47.2 (2.3)	18.0 (0.3)	-29.2 (2.3)
20%	-210.8 (0.3)	286.2 (2.1)	-48.2 (2.1)	19.8 (0.3)	-28.4 (2.1)
30%	-215.6 (0.3)	293.1 (3.7)	-50.3 (3.7)	21.4 (0.3)	-28.9 (3.7)

Tab. S11 Results from TI. ΔG^{res} is $-28.9\,kJ\,mol^{-1}$. ΔG^{symm} is $-1.7\,kJ\,mol^{-1}$. Uncertainties are indicated in brackets. Units are $kJ\ mol^{-1}$.

System	ΔG_{solv}^{raw}	ΔG_{PL}^{all}	ΔG_{bind}^{raw}	ΔG_{bind}^{corr}	ΔG_{bind}
0%	-195.7 (0.1)	269.3 (1.9)	-46.4 (1.9)	16.9 (0.2)	-29.5 (1.9)
10%	-204.9 (0.2)	278.3 (2.2)	-46.2 (2.2)	18.0 (0.3)	-28.2 (2.2)
20%	-210.9 (0.3)	285.9 (2.2)	-47.8 (2.2)	19.8 (0.3)	-27.9 (2.2)
30%	-215.9 (0.3)	293.2 (3.4)	-50.1 (3.4)	21.4 (0.3)	-28.7 (3.4)

Tab. S12 Results from MBAR. ΔG^{res} is $-28.9\,kJ\,mol^{-1}$. ΔG^{symm} is $-1.7\,kJ\,mol^{-1}$. Uncertainties are indicated in brackets. Units are $kJ\ mol^{-1}$.

System	ΔG_{solv}^{raw}	ΔG_{PL}^{all}	ΔG_{bind}^{raw}	ΔG_{bind}^{corr}	ΔG_{bind}
0%	-195.5 (0.1)	269.3 (1.9)	-46.6 (1.9)	16.9 (0.2)	-29.7 (1.9)
10%	-203.9 (0.4)	275.4 (2.3)	-44.3 (2.3)	18.0 (0.3)	-26.2 (2.3)
20%	-210.8 (0.3)	285.4 (0.9)	-47.4 (0.9)	19.8 (0.3)	-27.6 (1.0)
30%	-215.8 (0.1)	295.8 (2.9)	-52.8 (2.9)	21.4 (0.3)	-31.4 (2.9)

Tab. S13 Thermodynamics of PAB binding to trypsin in Tris Buffer. Uncertainties are indicated in brackets. Units are kJ mol^{-1}.

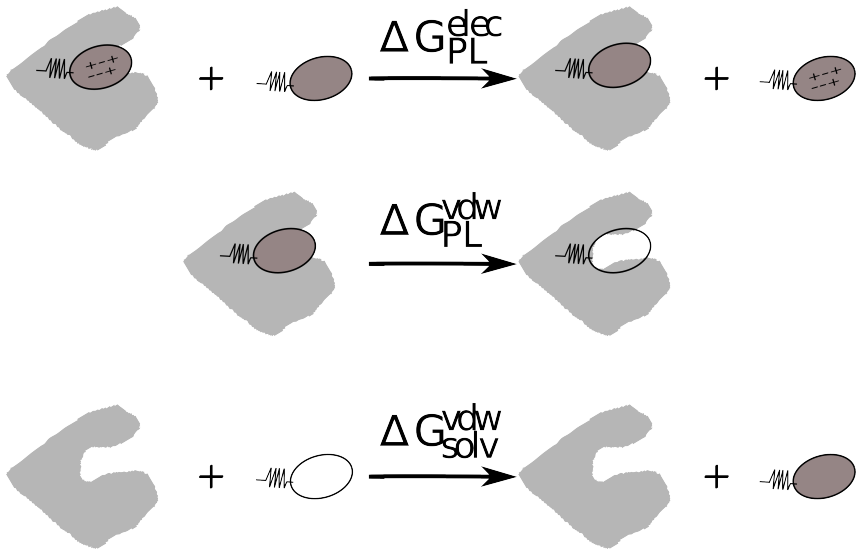
System	ΔG_{bind}	ΔH_{bind}	$-T\Delta S_{bind}$	$\Delta\Delta G_{bind}$	$\Delta\Delta H_{bind}$	$-T\Delta\Delta S_{bind}$
0%	-29.5 (0.2)	-28.5 (0.2)	-1.0 (0.3)	-	-	-
10%	-27.9 (0.3)	-29.8 (0.4)	+1.9 (0.1)	+1.6	-1.3	+2.9
20%	-27.4 (0.3)	-34.1 (0.5)	+6.7 (0.5)	+2.1	-5.6	+7.7
30%	-25.7 (0.5)	-37.3 (1.0)	+11.6 (0.8)	+3.6	-8.8	+12.6

Alternative approach to calculate free energy of binding

To independently check and validate our above results, we carried out additional sets of free energy simulations that do not require net charge corrections to the free energy. Figure S8 shows the alternative thermodynamic cycle. It involves three sets of simulations (instead of two in the previous approach): trypsin–ligand complex plus an extra uncharged ligand in the same simulation box (placed in the bulk solvent), only trypsin–ligand complex, and only ligand. The electrostatic contribution to the free energy is determined in a charge-neutral box by simultaneously introducing the electrostatic interactions of the ligand in the bulk solvent (PAB is kept at a distance of ca. 3.5 nm from the center of the protein by applying harmonic restraining potentials with force constant 1000 kJ/mol/nm^2 to the ligand heavy atoms), while its counterpart in the binding pocket is being electrostatically decoupled. The ΔG_{PL}^{elec} now includes the electrostatic contribution of both binding and solvation. The van der Waals contribution to the binding free energy, ΔG_{PL}^{vdW}, is calculated separately using a single uncharged ligand in the binding pocket, whereas the solvation contribution, ΔG_{solv}^{vdW}, is done as before with a single ligand in a box (Figure S8).

$$\Delta G_{bind} = -(\Delta G_{PL}^{res} + \Delta G_{PL}^{elec} + \Delta G_{PL}^{vdW} + \Delta G_{solv}^{vdW} + \Delta G^{res}) + \Delta G^{symm} \qquad \text{(S15)}$$

Fig. S8 Alternative thermodynamic cycle for binding. ΔG_{PL}^{elec} includes the electrostatic contribution to ΔG_{solv}.



Two independent sets of free energy simulations were performed for every solvent. Each λ-point (29 for electrostatics, 40 for van der Waals) was sampled for either 10 ns (PAB in binding pocket) or 5 ns (PAB in bulk solvent). The cumulative sampling time of these free energy simulations is $\approx 6\mu$s. The results are summarized in Table S14.

Tab. S14 ΔG_{bind} from the alternative approach (Figure S8). ΔG^{res} is $-28.9\,\mathrm{kJ\,mol^{-1}}$. ΔG_{PL}^{res} is ≈ 3 kJ mol^{-1}. Uncertainties are indicated in brackets. Units are kJ mol^{-1}.

Sys.	ΔG_{PL}^{elec}			ΔG_{PL}^{vdW}			ΔG_{solv}^{vdW}			ΔG_{bind}
	run1	run2	avg.	run1	run2	avg.	run1	run2	avg.	
0%	80	83.1	81.5 (1.5)	-28.0	-26.1	-27.1 (0.9)	-0.1	-0.3	-0.2 (0.1)	-29.9 (1.8)
10%	83.3	85.9	84.6 (1.3)	-25.6	-27.2	-26.4 (0.8)	-5.8	-5.7	-5.7 (0.1)	-28.2 (1.5)
20%	87.1	87.7	87.4 (0.3)	-23.2	-23.6	-23.4 (0.2)	-10.1	-10.0	-10.0 (0.1)	-29.8 (0.4)
30%	83.0	87.0	85.0 (2.0)	-22.5	-23.8	-23.1 (0.7)	-12.5	-12.7	-12.6 (0.1)	-25.2 (2.1)

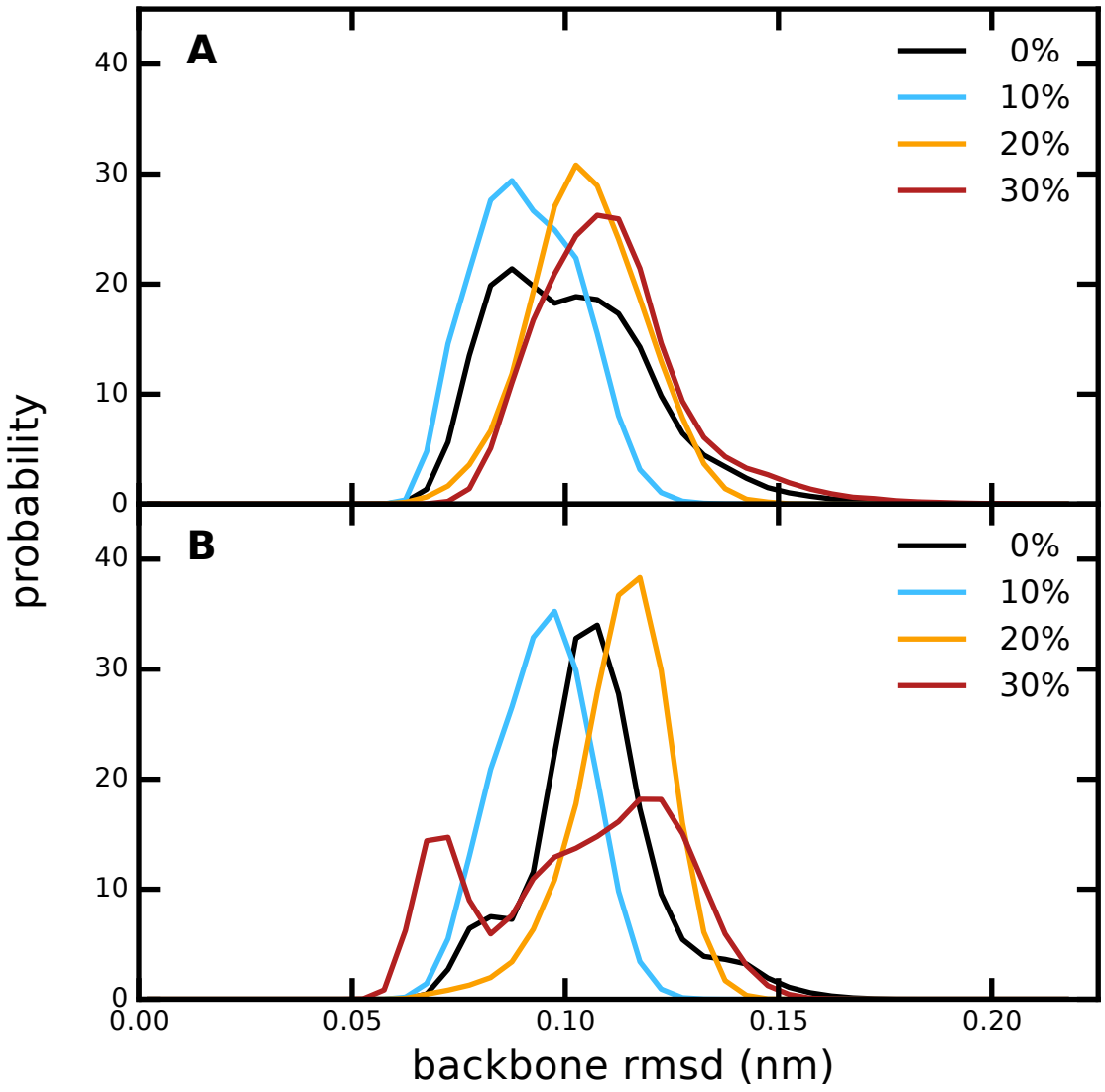


Fig. S9 Backbone RMSD distributions of trypsin at various methanol concentrations. (A) Trypsin-PAB complex. (B) Apo trypsin.

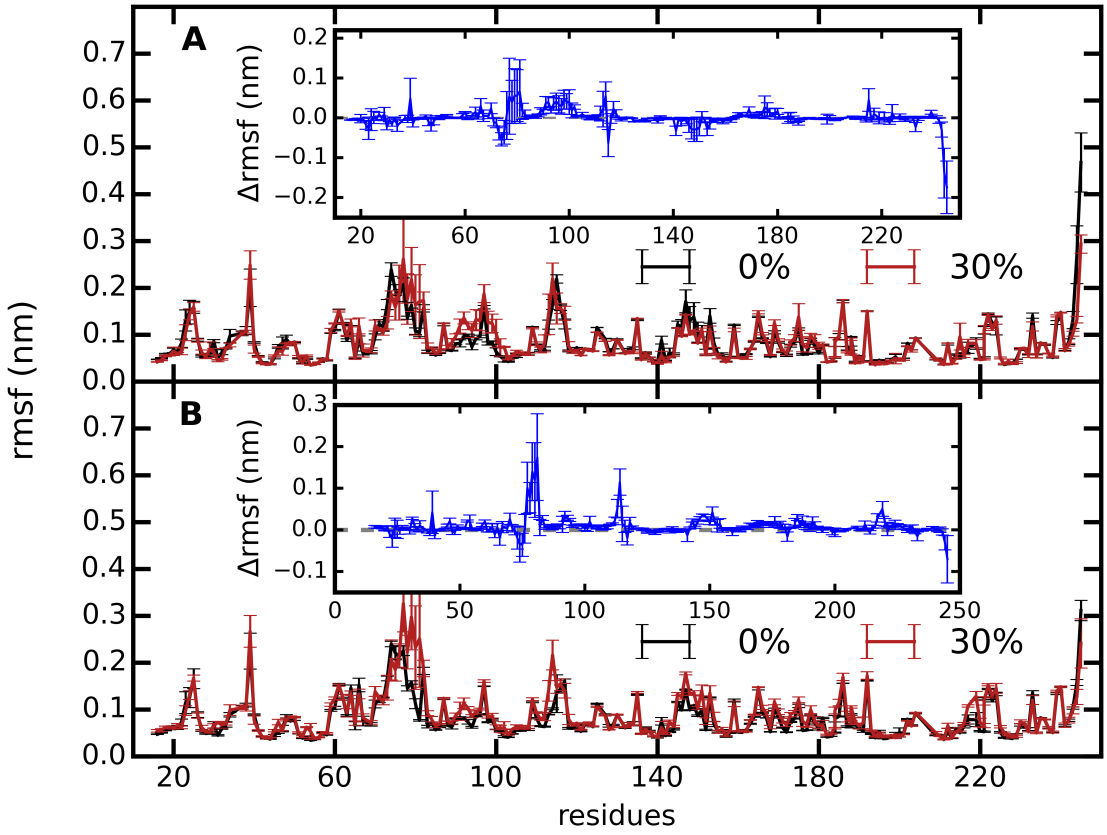


Fig. S10 RMS fluctuations of A) trypsin-PAB complex, and B) apo trypsin. The insets show the RMSF differences between 0% and 30% methanol.

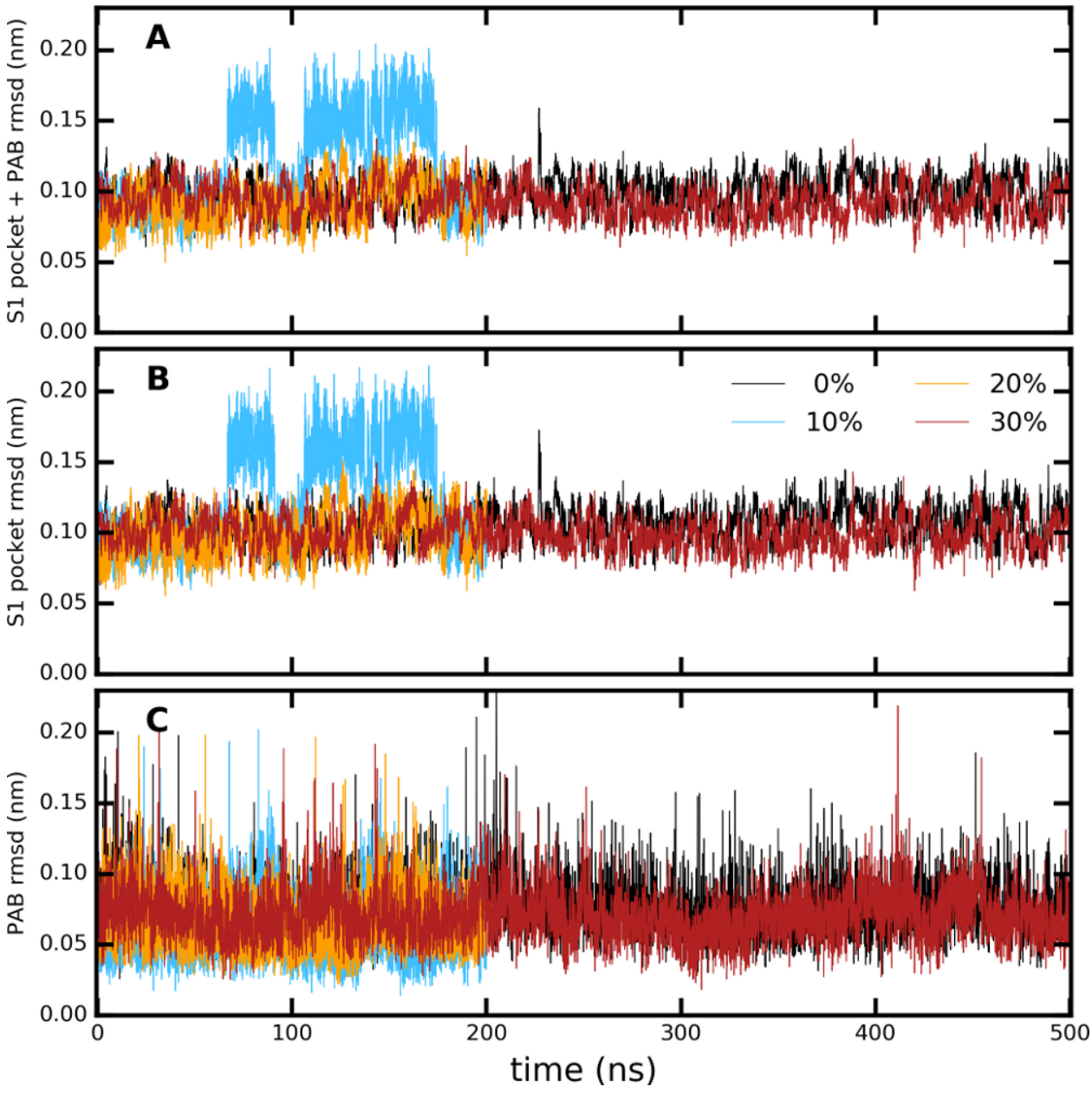


Fig. S11 RMSD time traces. (A) S1 pocket and PAB. (B) S1 pocket. (C) PAB after fitting to the protein backbone. The two distinct states observed for 10% methanol correspond to the two conformational states of Trp215.

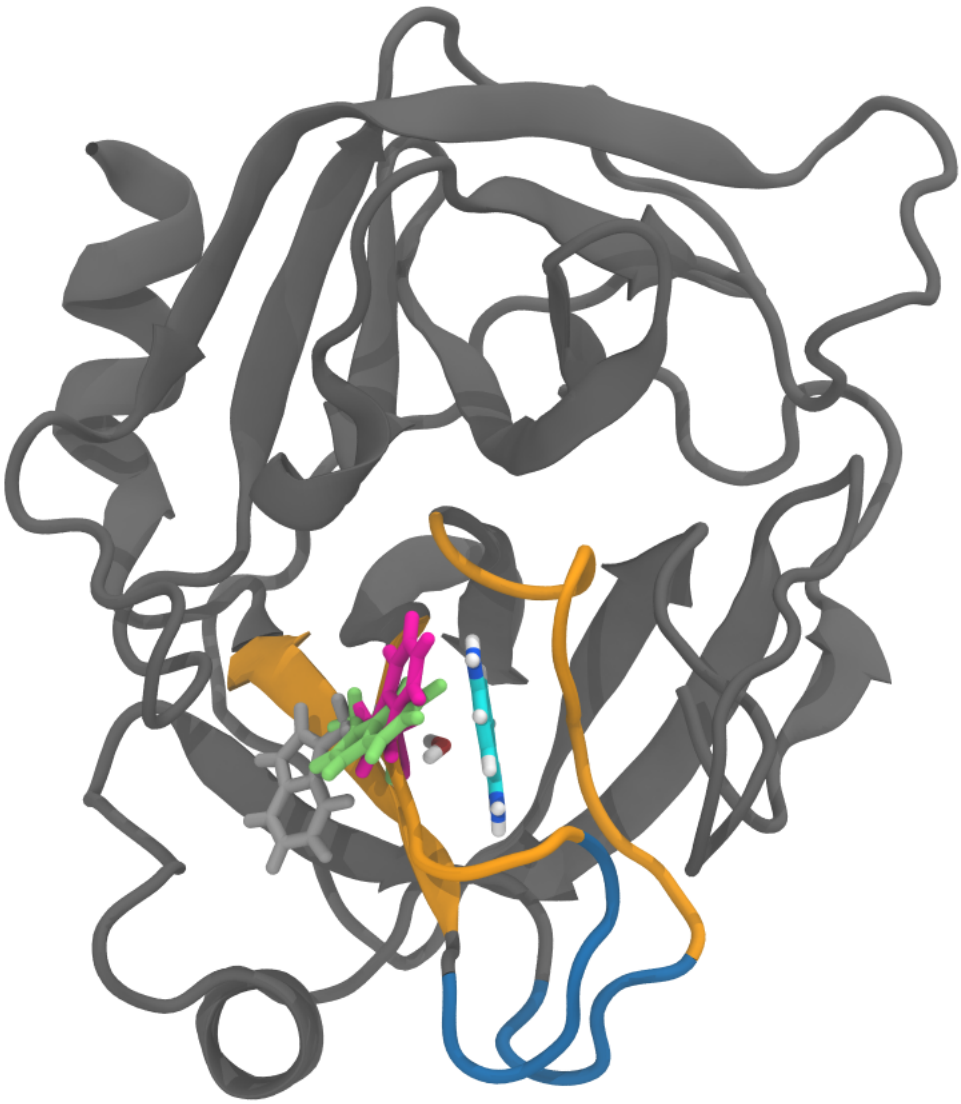


Fig. S12 Representative structures of closed (magenta and green) and open (light grey) states of Trp215. The open state is observed in the x-ray crystal structure of the complex (PDB: 3PTB).

Tab. S15 Average number of hydrogen bonds formed by PAB in the binding pocket with protein and solvent. Also shown are hydrogen bonds of free PAB in bulk. Uncertainties are indicated in brackets.

System	Binding Pocket			Bulk	
	Protein	Water	Methanol	Water	Methanol
0%	3.12 (0.07)	2.71 (0.06)	–	6.10 (0.01)	–
10%	3.21 (0.09)	2.46 (0.07)	0.01 (0.01)	5.51 (0.01)	0.58 (0.01)
20%	2.89 (0.11)	2.70 (0.08)	0.19 (0.01)	4.90 (0.02)	1.17 (0.02)
30%	3.29 (0.04)	2.33 (0.04)	0.24 (0.01)	4.43 (0.02)	1.60 (0.02)

Tab. S16 Average number of hydrogen bonds between PAB and residues in the binding site. WAT is the bound crystal water. Uncertainties are indicated in brackets.

System	Asp189	Gly219	Ser190	Ser195	WAT
0%	2.04 (0.02)	0.78 (0.20)	0.16 (0.05)	0.10 (0.01)	0.77 (0.06)
10%	1.99 (0.03)	0.76 (0.08)	0.16 (0.06)	0.19 (0.05)	0.75 (0.06)
20%	2.06 (0.01)	0.59 (0.23)	0.13 (0.11)	0.10 (0.03)	0.81 (0.14)
30%	2.04 (0.02)	0.87 (0.07)	0.19 (0.05)	0.11 (0.02)	0.74 (0.07)

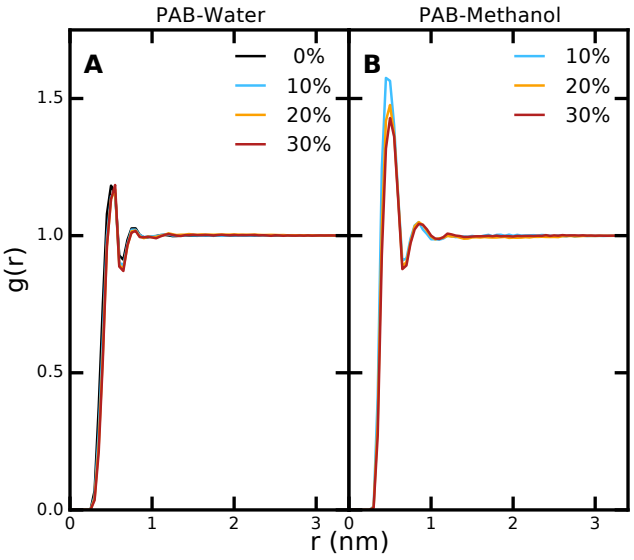


Fig. S13 Radial distribution functions from simulations of PAB free in solution. (A) PAB-water, (B) PAB-methanol.

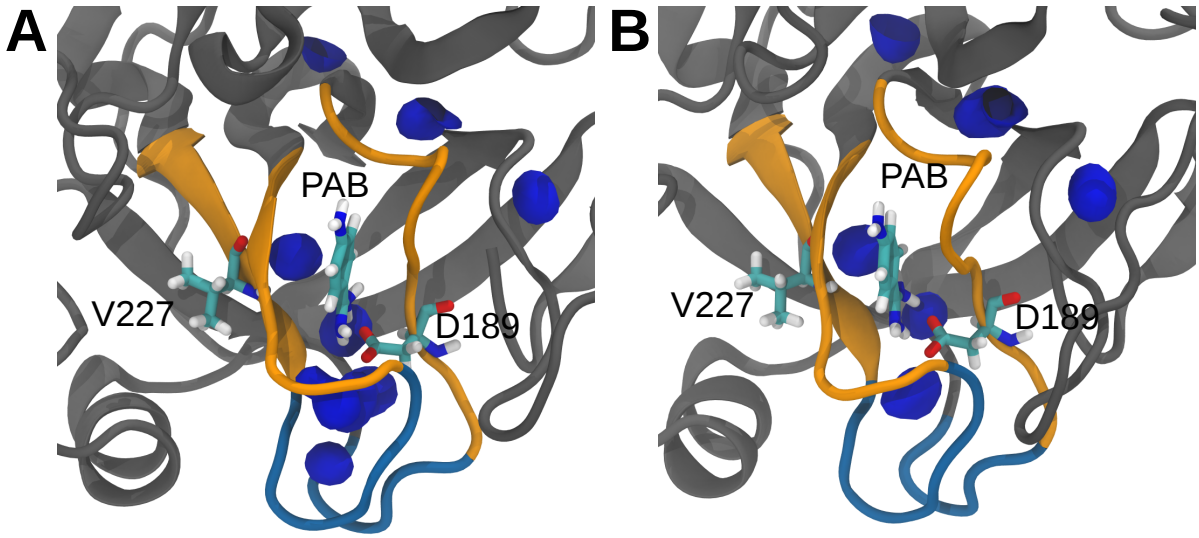


Fig. S14 Spatial density map of water in the S1 pocket region for the trypsin-PAB complex in (A) 10% and (B) 20% water/methanol. Only those regions where the number density of water molecules is equal to or greater than the bulk density are shown.

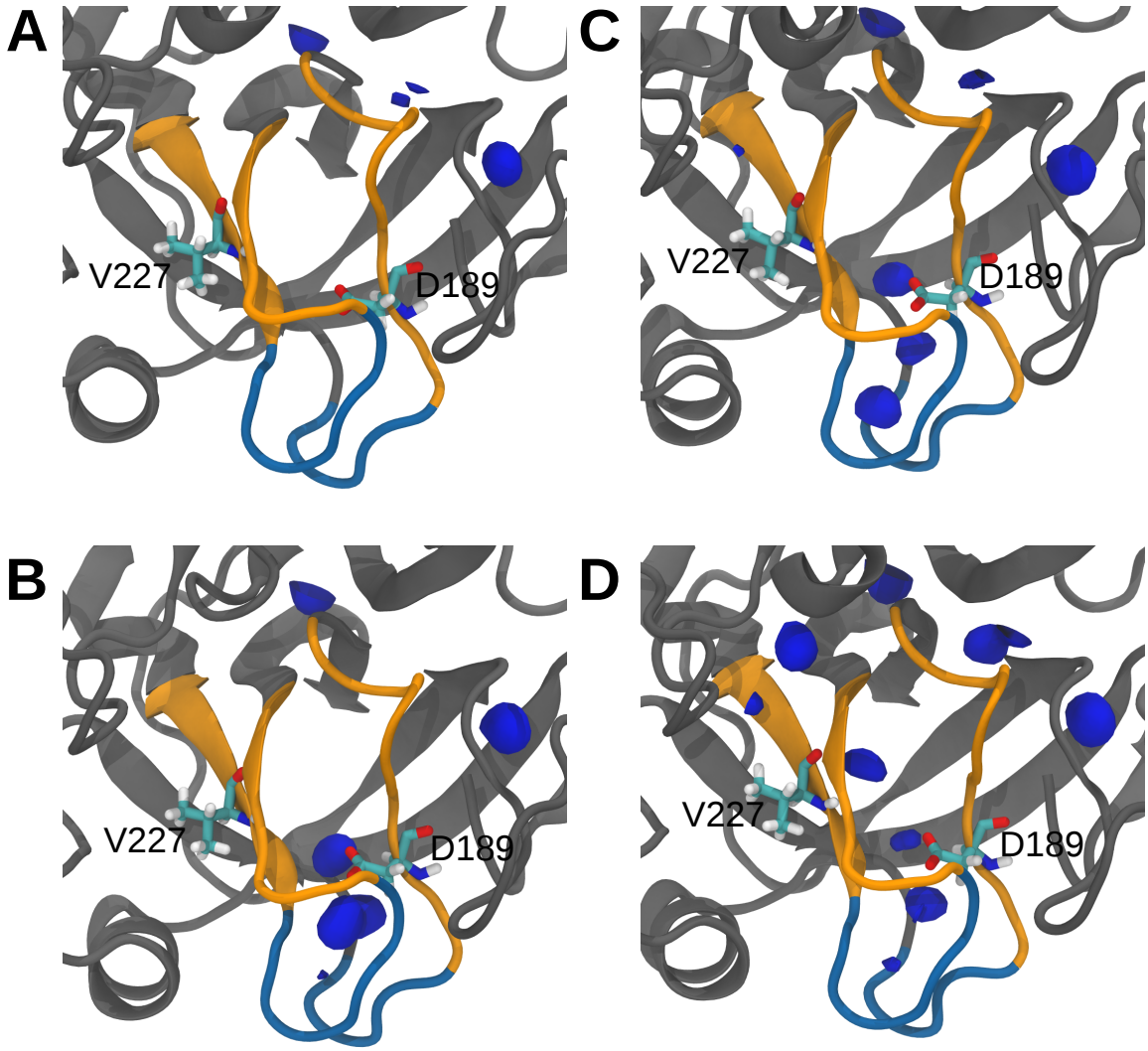


Fig. S15 Spatial density map of water in the S1 pocket region of apo trypsin in (A) pure water, (B) 10% methanol, (C) 20% methanol, (D) and 30% water/methanol. Only those regions where the number density of water molecules is equal to or greater than the bulk density are shown.

Tab. S17 Interaction energies of PAB with binding site residues from 3 independent 500 ns simulations in pure water. SOL indicates all water molecules including the crystallographic water molecule WAT. Uncertainties are shown in brackets.

Residue	Average (kJ/mol)	run1 (kJ/mol)	run2 (kJ/mol)	run3 (kJ/mol)
Ala221	3.1 (0.0)	3.0	3.1	3.1
Asn223	0.0 (0.0)	0.0	0.0	0.0
Asp102	-0.1 (0.0)	-0.1	-0.1	-0.1
Asp189	-119.1 (0.1)	-118.9	-119.1	-119.4
Asp194	-1.4 (0.0)	-1.3	-1.4	-1.4
Cys191	-17.9 (0.1)	-18.2	-17.6	-17.9
Cys220	-4.2 (0.2)	-3.8	-4.7	-4.1
Gln192	-9.5 (0.2)	-9.4	-10.0	-9.2
Gln221	0.2 (0.0)	0.1	0.2	0.2
Glu186	0.0 (0.0)	0.0	0.0	0.0
Gly187	0.0 (0.0)	0.0	0.0	0.0
Gly188	0.0 (0.0)	-0.0	-0.0	-0.0
Gly193	-0.3 (0.0)	-0.3	-0.3	-0.3
Gly196	0.0 (0.0)	0.0	0.0	0.0
Gly216	-7.9 (0.7)	-7.1	-9.6	-6.9
Gly219	-13.1 (5.0)	-4.9	-25.1	-9.2
Gly226	1.7 (0.1)	1.7	2.0	1.5
His57	-2.1 (0.1)	-2.0	-2.3	-2.0
Leu185	0.0 (0.0)	0.0	0.0	0.0
Lys188	-0.1 (0.0)	-0.1	-0.1	-0.1
Lys222	0.0 (0.0)	-0.0	-0.0	-0.0
Lys224	-3.2 (0.3)	-3.6	-2.5	-3.5
Lys230	0.0 (0.0)	0.0	0.0	0.0
Pro225	-2.8 (0.6)	-3.8	-1.4	-3.1
Ser190	-27.7 (1.4)	-29.9	-24.3	-28.9
Ser195	-2.4 (0.6)	-1.2	-3.9	-2.2
Ser214	-6.3 (0.2)	-6.5	-6.4	-5.9
Ser217	-1.8 (0.2)	-1.6	-2.2	-1.6
Thr229	0.0 (0.0)	-0.0	-0.0	-0.0
Trp215	-14.4 (0.3)	-14.5	-15.1	-13.7
Tyr228	-1.7 (0.1)	-1.9	-1.7	-1.5
Val227	-2.7 (0.1)	-2.7	-2.8	-2.5
SOL	-88.7 (3.5)	-94.7	-80.4	-91.1

Tab. S18 Interaction energies of PAB with binding site residues from 3 independent 500 ns simulations in 30% methanol. SOL indicates all water molecules including the crystallographic water molecule WAT; MTL all methanol molecules. Uncertainties are shown in brackets.

Residue	Average (kJ/mol)	run1 (kJ/mol)	run2 (kJ/mol)	run3 (kJ/mol)
Ala221	3.1 (0.0)	3.0	3.1	3.2
Asn223	0.0 (0.0)	0.0	0.0	0.0
Asp102	-0.1 (0.0)	-0.1	-0.1	-0.1
Asp189	-117.3 (0.3)	-117.9	-116.7	-117.2
Asp194	-1.4 (0.0)	-1.4	-1.4	-1.3
Cys191	-17.5 (0.1)	-17.4	-17.4	-17.6
Cys220	-5.0 (0.1)	-5.0	-4.9	-5.2
Gln192	-9.9 (0.3)	-10.4	-10.2	-9.1
Gln221	0.3 (0.0)	0.3	0.3	0.3
Glu186	0.0 (0.0)	0.0	0.0	0.0
Gly187	0.0 (0.0)	0.0	0.0	0.0
Gly188	0.0 (0.0)	0.0	0.0	0.0
Gly193	-0.4 (0.0)	-0.4	-0.4	-0.3
Gly196	0.0 (0.0)	0.0	0.0	0.0
Gly216	-10.9 (0.7)	-9.3	-11.5	-11.9
Gly219	-32.4 (1.4)	-29.8	-31.9	-35.5
Gly226	1.8 (0.2)	2.2	2.0	1.2
His57	-2.2 (0.0)	-2.1	-2.2	-2.2
Leu185	0.0 (0.0)	0.0	0.0	0.0
Lys188	-0.1 (0.0)	-0.1	-0.1	-0.1
Lys222	0.0 (0.0)	-0.0	-0.0	-0.0
Lys224	-1.9 (0.0)	-1.9	-2.0	-1.8
Lys230	0.0 (0.0)	0.0	0.0	0.0
Pro225	-0.5 (0.2)	-0.9	-0.3	-0.3
Ser190	-25.0 (1.9)	-20.5	-26.6	-27.9
Ser195	0.1 (0.3)	0.1	-0.6	0.7
Ser214	-9.1 (0.3)	-9.1	-8.4	-9.7
Ser217	-2.3 (0.2)	-1.9	-2.5	-2.5
Thr229	0.0 (0.0)	0.0	0.0	0.0
Trp215	-16.8 (0.2)	-17.1	-16.9	-16.3
Tyr228	-1.9 (0.2)	-2.1	-1.5	-2.1
Val227	-3.1 (0.0)	-3.1	-3.0	-3.1
SOL	-69.3 (0.9)	-71.2	-69.3	-67.5
MTL	-13.6 (0.1)	-13.6	-13.7	-13.4

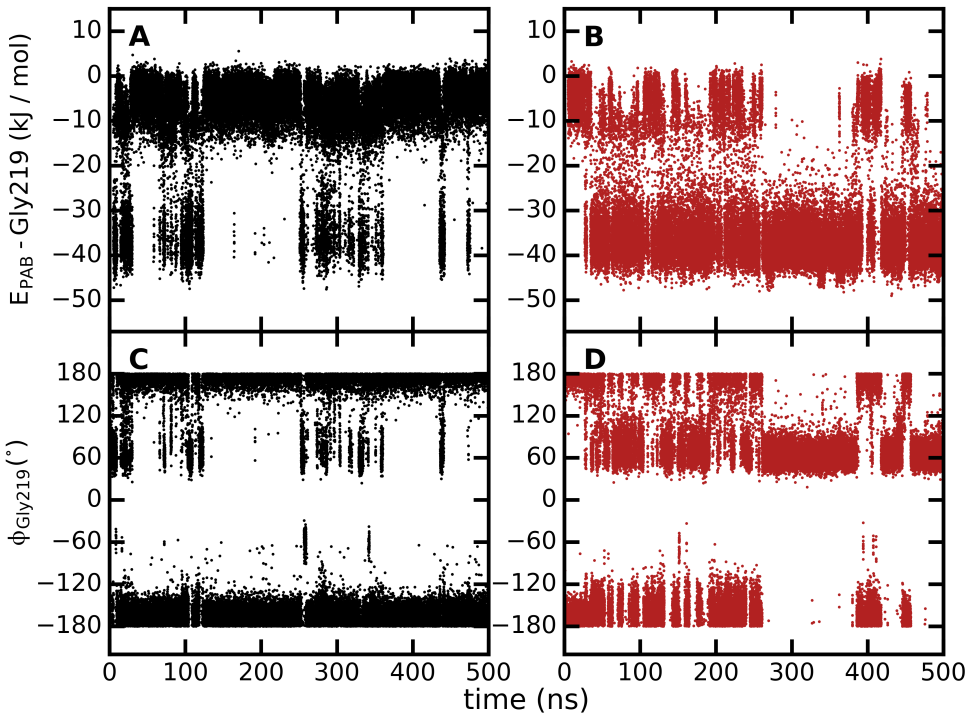


Fig. S16 Representative MD trajectories for 0% (A, C) and 30% (B, D) methanol. The Gly219–PAB interaction energy (top panel) correlates with the ϕ backbone dihedral angle of Gly219 (lower panel).

Tab. S19 Protein–protein interaction energies from independent MD simulations. The interaction energies were calculated between all amino acids listed in Table S17/S18, and between these residues and the rest of the protein. Uncertainties are indicated in brackets.

System	Average (kJ/mol)	run1 (kJ/mol)	run2 (kJ/mol)	run3 (kJ/mol)
		Complex		
0%	-14369.0 (11.3)	-14364.2	-14390.5	-14352.2
30%	-14406.7 (19.8)	-14368.9	-14415.1	-14436.0
		Apo		
0%	-14347.3 (8.0)	-14352.4	-14358.0	-14331.6
30%	-14342.7 (20.8)	-14384.1	-14319.0	-14325.1

Tab. S20 Protein – solvent interaction energies from independent MD simulations. As before, only the interactions of binding site residues with the solvent were considered. Uncertainties are indicated in brackets.

System	Average (kJ/mol)	run1 (kJ/mol)	run2 (kJ/mol)	run3 (kJ/mol)
		Complex		
0%	-2228.3 (3.5)	-2223.8	-2225.9	-2235.1
30%	-2297.6 (13.8)	-2322.2	-2274.3	-2296.3
		Apo		
0%	-2387.9 (14.4)	-2416.6	-2373.9	-2373.2
30%	-2492.8 (0.1)	-2492.9	-2492.6	-2492.8

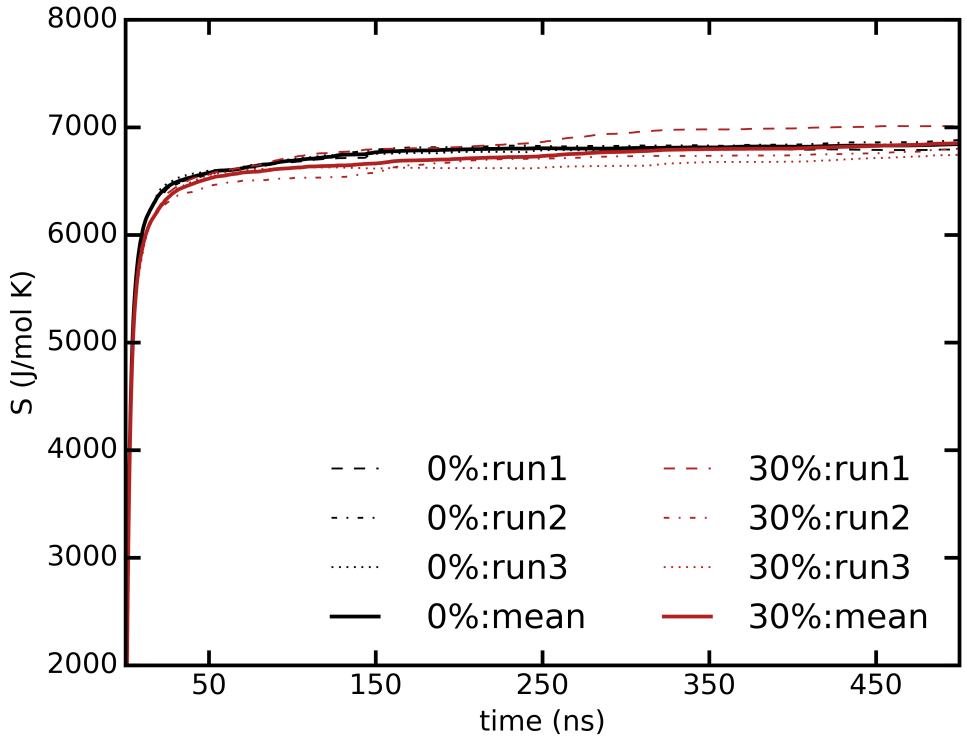


Fig. S17 Configurational entropy ($\mathrm{S}^{\mathrm{conf}}$) of complex calculated from three independent simulations using the quasi-harmonic approximation as formulated by Schlitter[S11]. All Cα atoms used to construct the covariance matrix of particle fluctuations. In pure water, $\mathrm{S}^{\mathrm{conf}}$ converges to a value of 6843.3 (26.7) J $\mathrm{K}^{-1}\mathrm{mol}^{-1}$. In 30% methanol, $\mathrm{S}^{\mathrm{conf}}$ is 6852.3 (81.2) J $\mathrm{K}^{-1}\mathrm{mol}^{-1}$.

Tab. S21 Examples of tightly bound water molecules in x-ray crystal structures of trypsin-benzamidine complexes. Distances and angles between the ligand, WAT, and Val227 are listed. PDB entry of the structure used in this paper is shown in bold.

PDB id	Resid of WAT	$N1/N2_{Lig} - O_{WAT}$ (nm)	$N_{Val227} - O_{WAT}$ (nm)	$N1/N2_{Lig} - O_{WAT} - N_{Val227}$ (°)
1BTY	268	0.29	0.35	93.7
1C5P	325	0.31	0.34	90.8
1CE5	732	0.32	0.36	84.4
1DPO	325	0.30	0.36	88.1
1H4W	2115	0.31	0.34	90.9
1HJ8	2216	0.29	0.35	88.5
1J16	502	0.29	0.35	89.1
1L2E	309	0.31	0.36	87.2
1OSS	292	0.29	0.34	88.5
1S0R	254	0.30	0.35	88.7
1S6H	317	0.30	0.36	87.8
1TIO	289	0.31	0.36	92.7
1V2L	613	0.29	0.35	88.6
1XUF	268	0.29	0.36	96.2
2BLW	2286	0.30	0.35	88.3
2EEK	535	0.31	0.35	90.0
3PTB	**416**	**0.31**	**0.34**	**89.8**

Free energy of tying up a water molecule in the binding site

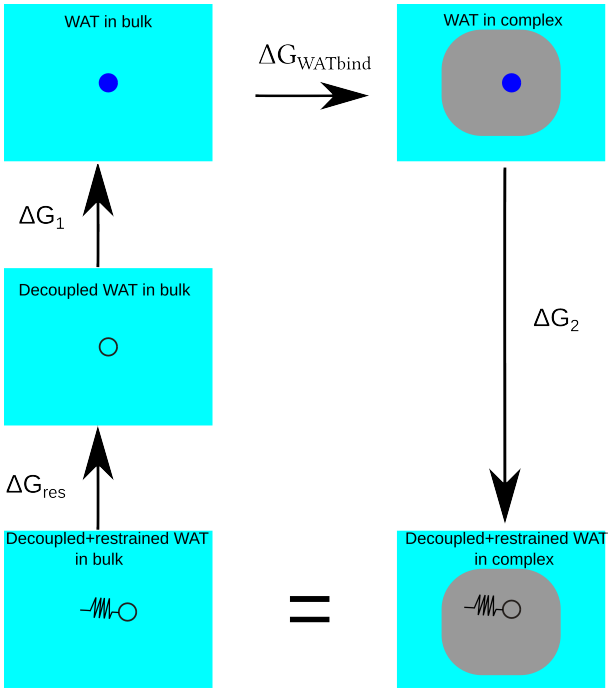


Fig. S18 Thermodynamic cycle of tying up the crystallographic WAT molecule from bulk to the binding site cavity in the trypsin-PAB complex. $\Delta G_{\mathrm{WATbind}} = -(\Delta G_1 + \Delta G_2 + \Delta G^{res})$.

Tab. S22 Average WAT–PAB and PAB–Val227 distances as well as PAB-WAT-Val227 angle.

System	$\mathrm{N1_{PAB}} - \mathrm{O_{WAT}}$ (nm)	$\mathrm{N_{Val227}} - \mathrm{O_{WAT}}$ (nm)	$\mathrm{N1_{PAB}} - \mathrm{O_{WAT}} - \mathrm{N_{Val227}}$ (°)
0%	0.32 (0.02)	0.34 (0.01)	90.2 (2.0)
10%	0.31 (0.01)	0.34 (0.01)	90.4 (0.8)
20%	0.33 (0.02)	0.34 (0.01)	89.3 (2.7)
30%	0.32 (0.01)	0.34 (0.01)	88.2 (0.6)

The crystal water molecule (WAT) that bridges between PAB and Val227 stays bound in the binding site over the entire simulation at all methanol concentrations (Table S22). To preserve this bound state during the free energy simulations, a harmonic restraint potential energy function was applied to the oxygen atom of WAT. The force constant of the restraint

was obtained from the equilibrium simulations, following the approach suggested in the literature[S12].

$$\Delta G^{res} = k_B T \ln \left(V_{Lig} / V_0 \right) \tag{S16}$$

$$V_{Lig} = \left(\frac{2\pi k_B T}{k^{res}} \right)^{3/2} \tag{S17}$$

$$k^{res} = 3 k_B T / \langle \delta r \rangle^2 \tag{S18}$$

where ΔG^{res} is the correction due to the restraint (Table S23), V_0 is the standard volume of water, V_{Lig} is the volume available to the ligand, and k^{res} is the force constant obtained from the atomic fluctuations δr of WAT in the bound state. Unlike for the ligand, no symmetry correction ΔG^{symm} needs to be applied in this case, since the restraining potential on the water oxygen atom does not prevent rotations around its C_2-axis, and these rotations indeed occur on the 5 ns time scale sampled at each λ-point.

Tab. S23 Free energy contributions due to restraints and standard volume. The standard volume is calculated from our simulation boxes. The corresponding standard concentration is ca. 55 M for pure water and accordingly lower for the water/methanol mixtures.

System	δr (nm)	k^{res} (kJ/mol/nm^2)	V_0 (nm^3)	ΔG^{res} (kJ/mol)
0%	0.059	2106.6	0.0294	-9.5
10%	0.058	2164.5	0.0357	-10.1
20%	0.061	1958.8	0.0440	-10.2
30%	0.057	2263.8	0.0522	-11.2

Tab. S24 The enthalpy ($\Delta H_{WATbind}$) of tying up a water molecule in the binding site is calculated using the relation: $\Delta H_{WATbind} = (E_{Prot\text{-}WAT} + E_{PAB\text{-}WAT} + E_{Solvent\text{-}WAT}) - E_{bulk}$. Uncertainties are indicated in brackets. Units are kJ mol^{-1}.

System	$E_{complex-WAT}$				E_{bulk}	$\Delta H_{WATbind}$
	$E_{Prot-WAT}$	$E_{PAB-WAT}$	$E_{Solvent-WAT}$	Total		
0%	-38.5 (0.2)	-34.2 (0.8)	-0.6 (0.1)	-73.3 (0.9)	-40.1 (0.1)	-33.2 (0.9)
10%	-39.7 (0.2)	-32.5 (0.3)	-0.5 (0.1)	-72.7 (0.5)	-43.0 (0.1)	-29.7 (0.5)
20%	-36.2 (0.1)	-28.4 (0.2)	-0.2 (0.1)	-64.8 (0.2)	-46.0 (0.1)	-18.8 (0.2)
30%	-37.6 (0.1)	-26.6 (0.5)	-0.3 (0.1)	-64.5 (0.5)	-48.4 (0.1)	-16.1 (0.5)

References

(S1) Akerlof, G. *J. Am. Chem. Soc.* **1932**, *54*, 4125–4139.

(S2) Mikhail, S.; Kimel, W. *J. Chem. Eng. Data* **1961**, *6*, 533–537.

(S3) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(S4) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. *J. Chem. Theory Comput.* **2012**, *8*, 61–74.

(S5) Neumann, M. *Mol. Phys.* **1983**, *50*, 841–858.

(S6) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, 084902.

(S7) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.

(S8) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. *J. Comput. Aided Mol. Des.* **2015**, *29*, 397–411.

(S9) Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. *J. Chem. Phys.* **2013**, *139*, 184103.

(S10) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc Natl Acad Sci U S A* **2001**, *98*, 10037–10041.

(S11) Schlitter, J. *Chem. Phys. Lett.* **1993**, *215*, 617–621.

(S12) Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 7683–7689.