

Supporting Information

Accurate and Rigorous Prediction of the Changes in Protein Free Energies in a Large-Scale Mutation Scan

Vytautas Gapsys, Servaas Michielssens, Daniel Seeliger, and Bert L. de Groot**

anie_201510054_sm_miscellaneous_information.pdf

Methods

Free energy calculation. To estimate double free energy differences reporting on the protein thermal stability a thermodynamic cycle was constructed according to [1]. The unfolded protein was represented by a capped tripeptide consisting of two glycine residues at the termini and the amino acid of interest in the middle. For the charge conserving mutations MD simulations were performed for the folded and unfolded states separately. In the case of charge changing mutations, the overall charge of the system was kept constant by performing mutations in the folded and unfolded states simultaneously in opposite directions by putting both states in the same simulation box [2]. The tripeptide was positioned ~ 3 nm away from barnase. To prevent the peptide and barnase getting closer to each other than the van der Waals and short range electrostatic cutoffs, position restraints were applied to the $C\alpha$ atom of the middle residue of a tripeptide. Also, linear center of mass removal was performed for the barnase and the rest of the simulated system separately. For the protein-protein binding free energy calculations the thermodynamic cycle was constructed by estimating ΔG upon mutation in a protein complexed with its binding partner and free in solution. Equilibrium 20 ns MD simulations were performed for the two states. Subsequently, snapshots from the equilibrium runs were extracted to spawn 100 simulations of 50 ps each to alchemically morph between the two states of the system. The work values over every non-equilibrium transition were extracted and further used to estimate the free energy differences relying on the Crooks Fluctuation Theorem [3] and utilizing Bennett acceptance ratio as a maximum likelihood estimator [4].

Force fields. The following force fields were used in the study: Amber family (Amber99sb [5] and Amber99sb*ILDN [6, 7]), OPLS [8, 9], and Charmm family (Charmm22* [10], Charmm36 [11]). In addition, for the Charmm36 force field the effects of two water models were tested: TIP3P model [12] without Lennard-Jones parameters on hydrogen atoms and the native Charmm TIP3P water model, denoted Charmm36H. Calculations for barnase were performed with all six force field variants. For the staphylococcal nuclease and protein binding simulations Amber99sb*ILDN and Charmm36H were used. Neurotensin receptor 1 was simulated with the Amber99sb*ILDN force field.

Molecular dynamics simulations. MD simulations were performed using the Gromacs [13] simulation package versions 4.6 and 5.0. Hybrid structures and topologies were prepared with the pmx [14] software. For the mutations considered in the protein-protein binding free energy calculations we constructed a specialized mutation library. In the library, for any pair of amino acids a hybrid structure was constructed by morphing backbone atoms and $C\beta$ atom directly from one state to another, while all the rest of the side chain atoms were turned to dummies. Particle Mesh Ewald [15] was used to treat electrostatic interactions in the simulation. For all the transitions and equilibrium simulations of the tripeptides, charge conserving mutations in barnase, staphylococcal nuclease and NTR1 triple mutant a short range cutoff for the electrostatic interactions of 1.2 nm was used. For the charge changing mutations in barnase, single and double mutants in NTR1 a shorter electrostatic interaction cutoff of 1.1 nm was used. During the alchemical transitions the non-bonded interactions were soft-cored using the default parameter set [16]. All

bonds were constrained using LINCS [17]. Temperature was controlled by the velocity rescaling thermostat [18]. The barnase and NTR1 simulations were performed at 300 K temperature using a time constant of 0.1 ps. Staphylococcal nuclease was simulated at 293 K temperature with a 0.1 ps time constant, α -chymotrypsin complexed with turkey ovomucoid third domain was simulated at 294 K temperature (0.1 ps time constant), antibody HyHEL-10 Fv complexed with hen egg lysozyme was simulated at 303 K temperature (0.1 ps time constant). The Parrinello-Rahman barostat [19] was used for pressure coupling. In all the simulations the pressure was kept at 1 bar using a time constant of 5 ps. The van der Waals interactions were switched off between 1 and 1.1 nm.

The structures used for the molecular dynamics simulations: barnase, 1BNI [20], staphylococcal nuclease 1STN [21], neurotensin receptor 1 4BUO [22], α -chymotrypsin complexed with turkey ovomucoid third domain 1CHO [23], antibody HyHEL-10 Fv complexed with hen egg lysozyme [24]. The proteins and peptides were solvated in TIP3P [12] water and 150 mM of Na^+ and Cl^- were added. Prior to starting the simulations, a steepest descent energy minimization was performed. 20 ns equilibrium MD runs were performed for all the simulation cases. From the generated trajectories the first 2 ns were discarded for all protein simulations. From the tripeptide trajectories the first 4 ns were discarded. 100 snapshots were equidistantly extracted from the resulting trajectories. For the cases of barnase, staphylococcal nuclease, tripeptides and protein-protein binding simulations hybrid structures for the mutated residues were introduced into the snapshots. Subsequently, every system was energy minimized and subjected to a 20 ps equilibrium simulation. Afterwards, an alchemical transition between the states of a hybrid structure was performed in 50 ps. In the case of NTR1, hybrid structures for the mutants were generated prior to starting the equilibrium simulations. Therefore, the alchemical transitions were started directly after extracting the snapshots. The transitions in the neurotensin receptor 1 were performed in 200 ps for the single and double mutants, whereas, to ensure convergence in the triple mutant, 1600 ps transitions were used.

Error estimation. The error bars throughout the manuscript report on the standard error of the mean of a measurement. For the average unsigned error estimates, AUE was calculated for every $\Delta\Delta G$ value and subsequently the standard error was assessed over multiple AUEs. For the correlations and percentage of the predictions within 1 kcal/mol accuracy, a bootstrapping procedure was employed (1000 resampling runs for every error estimate).

Rosetta ddg_monomer protocol. Mini-Rosetta version Release 3.3 from SVN 42942 was used for calculations. Prior to running the ddg_monomer calculations pre-minimization was performed using distance constraints on the $\text{C}\alpha$ atoms (minimize_with_cst protocol). Afterwards, the protocols were executed following the description by Kellogg et al. [25]. We used three protocols termed "row 3", "row 6" and "row 16" according to the naming in Table 1 of [25]. The exact commands with the parameters for the protocols were taken from the supporting information of [25]. The barnase calculations were performed starting from two crystal structures: 1BNI [20], resolution of 2.1 Å and 1A2P [26], resolution of 1.5 Å. The structures used for protein-protein binding calculations: 1CHO [23], resolution of 1.8 Å, 2DQJ [24], resolution of 1.8 Å.

Partial Least Squares (PLS) regression model. PLS regression was built using the specialized package [27] implemented in the statistical software package R [28]. For all the constructed models 1 PLS component was used. To build the regression three approaches were considered, which yielded similar results (Figure S11 in the Supporting Information).

In the first approach, the set of charge conserving mutations was randomly divided into training and testing subsets by a 1:1 ratio. The training subset was used for model building, while the testing subset was left for crossvalidation. The procedure was repeated 10000 times and an average model over all the runs was calculated. In the second approach, all the charge conserving mutations were used for model building and the charge changing mutations served as a subset for crossvalidation. For the third approach, all the charge changing mutations were used for model building, whereas the charge conserving mutations served as a subset for crossvalidation. The performance of the model from the second approach against the training and crossvalidation subsets can be seen in Fig. 1c,d and Figures S9a,S10a in the Supporting Information. The model from the first approach yielded results highly similar to the second approach. The PLS model based on the charged mutations also showed similar performance to the other two approaches: crossvalidation $\text{cor}=0.77$, $\text{AUE}=2.83$ kJ/mol, 73% within 1 kcal/mol accuracy; for all the mutations $\text{cor}=0.74$, $\text{AUE}=2.94$ kJ/mol, 72% within 1 kcal/mol accuracy.

Database analysis. ProTherm was parsed to extract the $\Delta\Delta G$ values. Whenever available, free energy differences of unfolding in water were used. To compare experimental values for the same mutation that has been reported multiple times, firstly, all the entries for identical mutations coming from the same source organism and the same protein have been extracted. Afterwards, the set of entries for a given mutation were randomly shuffled and paired by creating a chain of values for comparison. For the analysis in Fig. 2b a selected pair of $\Delta\Delta G$ values was accepted if for both experiments the reported temperatures, pH and ion concentrations (if reported in ProTherm) were equivalent. For the comparisons in Fig. 2c and Figure S13 in the Supporting Information values of those mutations were considered that were used for the alchemical free energy calculations; experimental conditions were required to have identical temperatures and pH.

ProTherm version containing 26045 entries was downloaded on 22.01.2014.

Literature sources for the ΔT_m vs $\Delta\Delta G$ analysis. Ribose-binding protein^a [29], ribose-binding protein^b [30], lysozyme^c [31], lysozyme^d [32], rubredoxin [33], SNase^e [29], SNase^f [34], fibroblast growth factor [35], neurotensin receptor 1^g [22], neurotensin receptor 1^h [36].

Literature sources for the $\Delta\Delta G$ values for protein-protein binding analysis. α -chymotrypsin complexed with turkey ovomucoid third domain 1CHO [37], antibody HyHEL-10 Fv complexed with hen egg lysozyme [38–40]. CC/PBSA results were taken from [41].

Average unsigned errors by amino acid type. To render the impact of the force field parameters clearer the average unsigned errors (AUE) were decomposed

into contributions for every amino acid separately. The results are presented in the supporting figures S15 and S16. The AUE values were calculated for every amino acid in each force field and compared to the results obtained with the consensus approach. For those cases where consensus outperformed the other force fields, it could be deduced that the parameterization of that particular amino acid in both force fields might benefit from a closer investigation. Glutamine and lysine serve as an example for such a situation (Fig. S16): the consensus approach gives a smaller AUE than any individual force field, hence, Amber99sb*ILDN and Charmm36H must be pointing in opposite directions (with respect to the experimental values) when predicting free energy changes for these residues. Such analysis also provides useful information when the consensus approach does not outperform the other force fields. For example, in figure S16 isoleucine involving free energy changes are predicted best by Charmm36H, indicating that both force fields make an error in the same direction (with respect to the experimental values), but one of the force fields has a better representation of this amino acid.

Protocol ^a	row 3	row 6	row 16	row 3	row 6	row 16
PDB ^b	1BNI	1BNI	1BNI	1A2P	1A2P	1A2P
All mutations	0.49	0.46	0.48	0.23	0.48	0.47
Charge conserving	0.61	0.67	0.64	0.36	0.72	0.71
Charge changing	-0.04	-0.19	-0.17	-0.19	-0.19	-0.26

Table S1: Correlations of the Rosetta based $\Delta\Delta G$ estimates and experimental measurements

^a Corresponds to the protocol in Table 1 in [25].

^b Resolution: 1BNI 2.1 Å, 1A2P 1.5 Å.

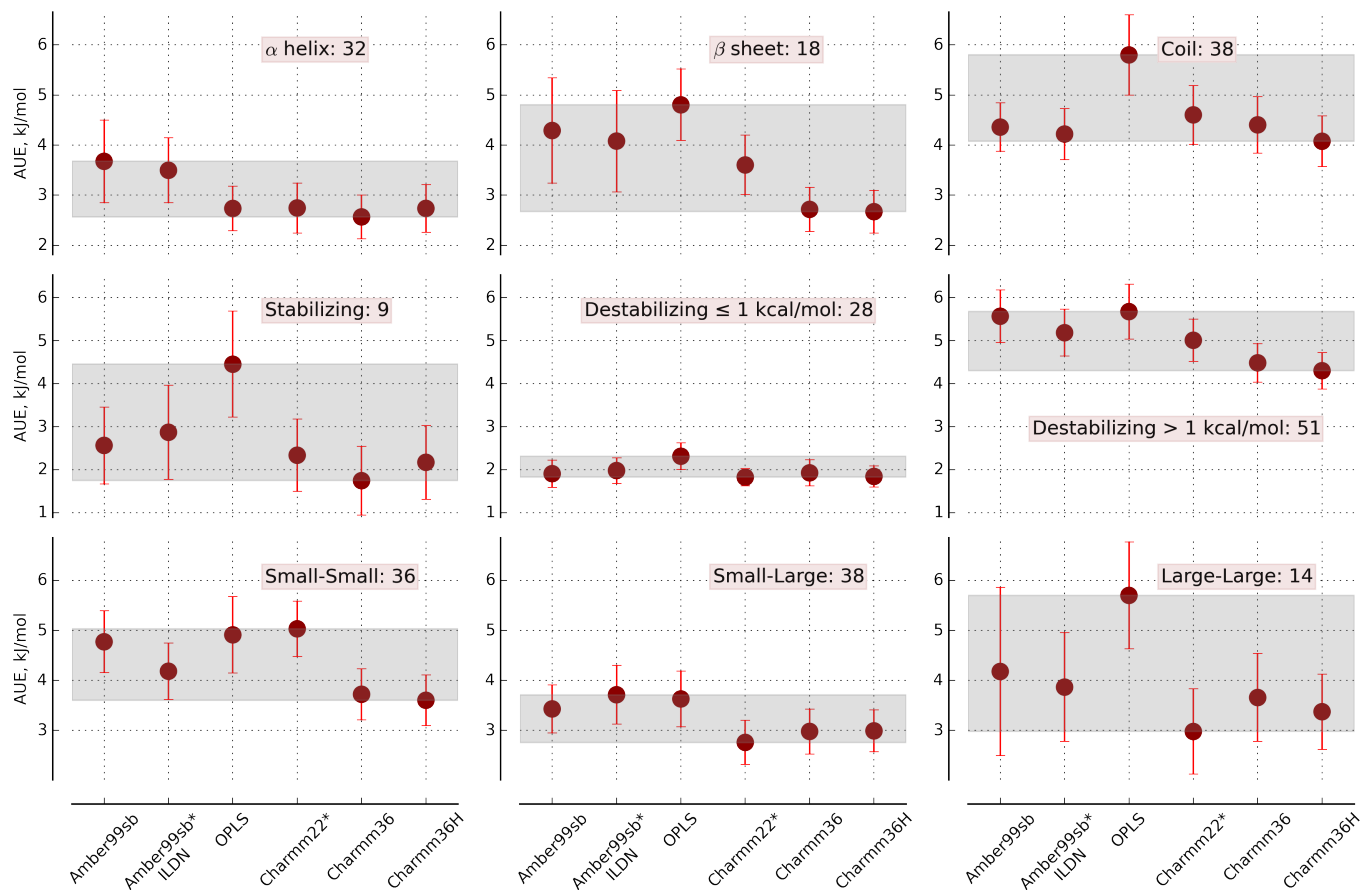


Figure S1: Average unsigned errors of the calculated $\Delta\Delta G$ values from experiment for mutation subsets grouped by secondary structure (top row), stabilizing effect (middle row), amino acid size (bottom row)

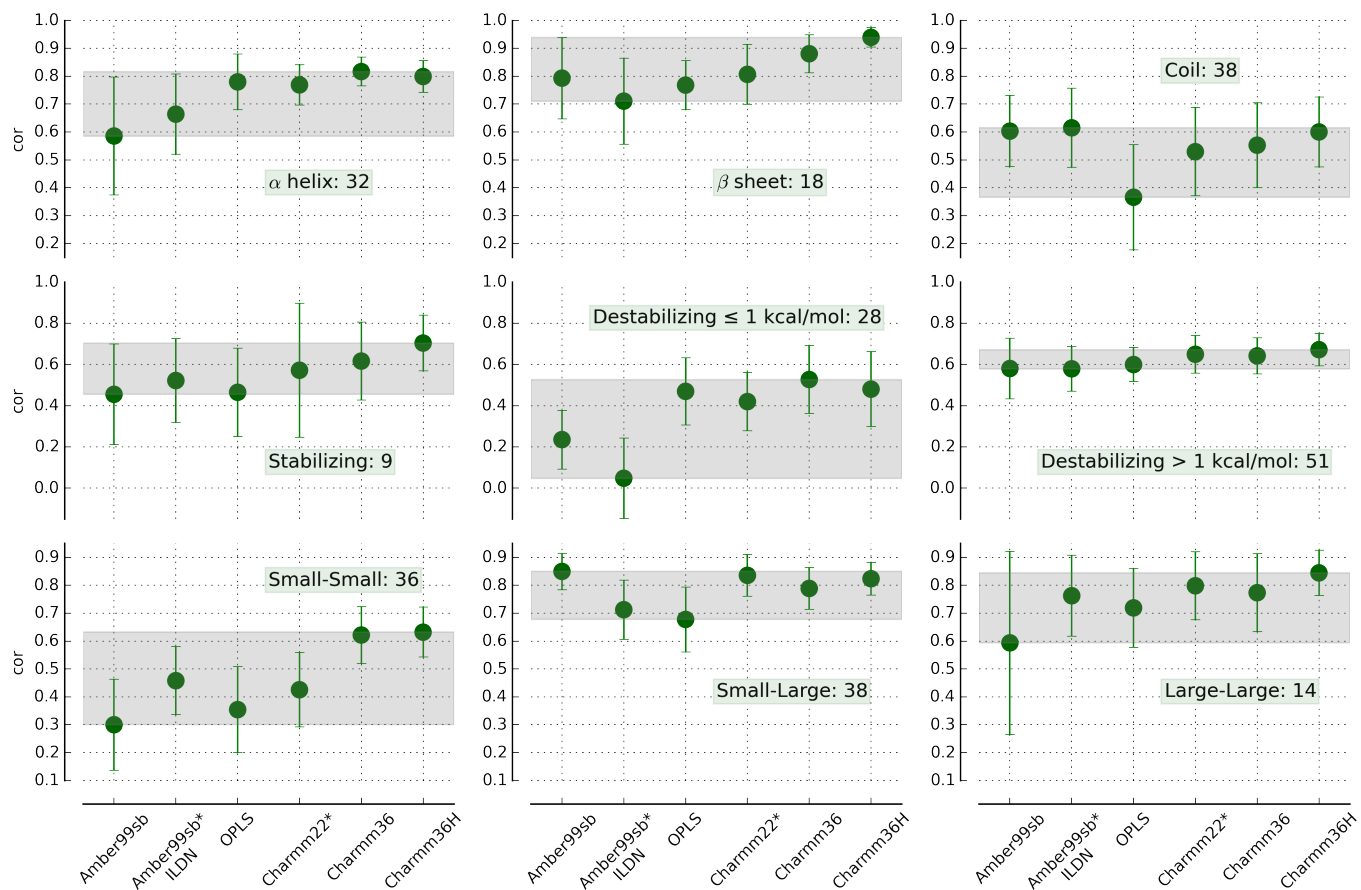


Figure S2: Correlations of the calculated $\Delta\Delta G$ values to experiment for mutation subsets grouped by secondary structure (top row), stabilizing effect (middle row), amino acid size (bottom row)

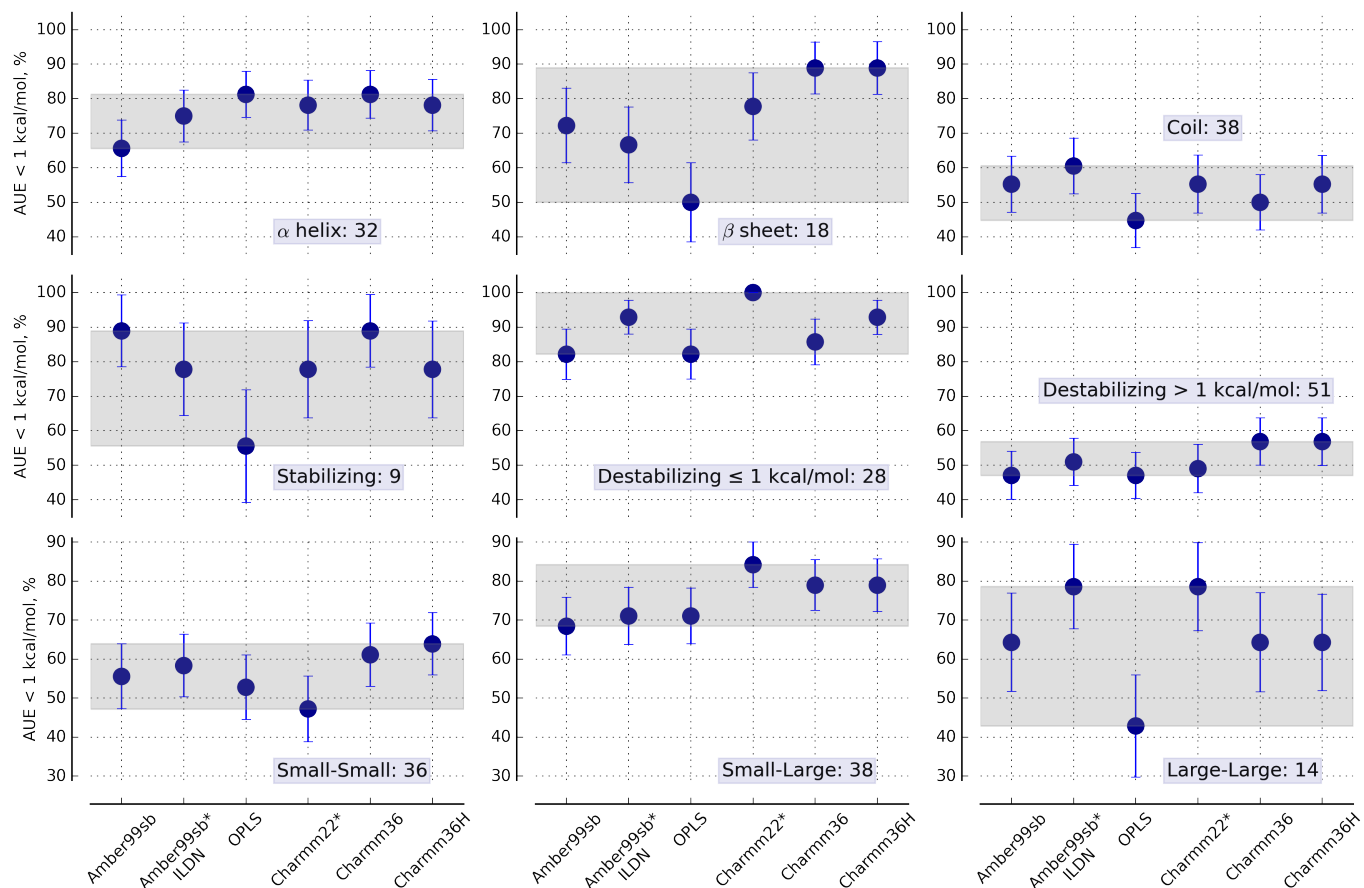


Figure S3: Percentage of the $\Delta\Delta G$ values predicted with 1 kcal/mol accuracy with respect to experiment for mutation subsets grouped by secondary structure (top row), stabilizing effect (middle row), amino acid size (bottom row)

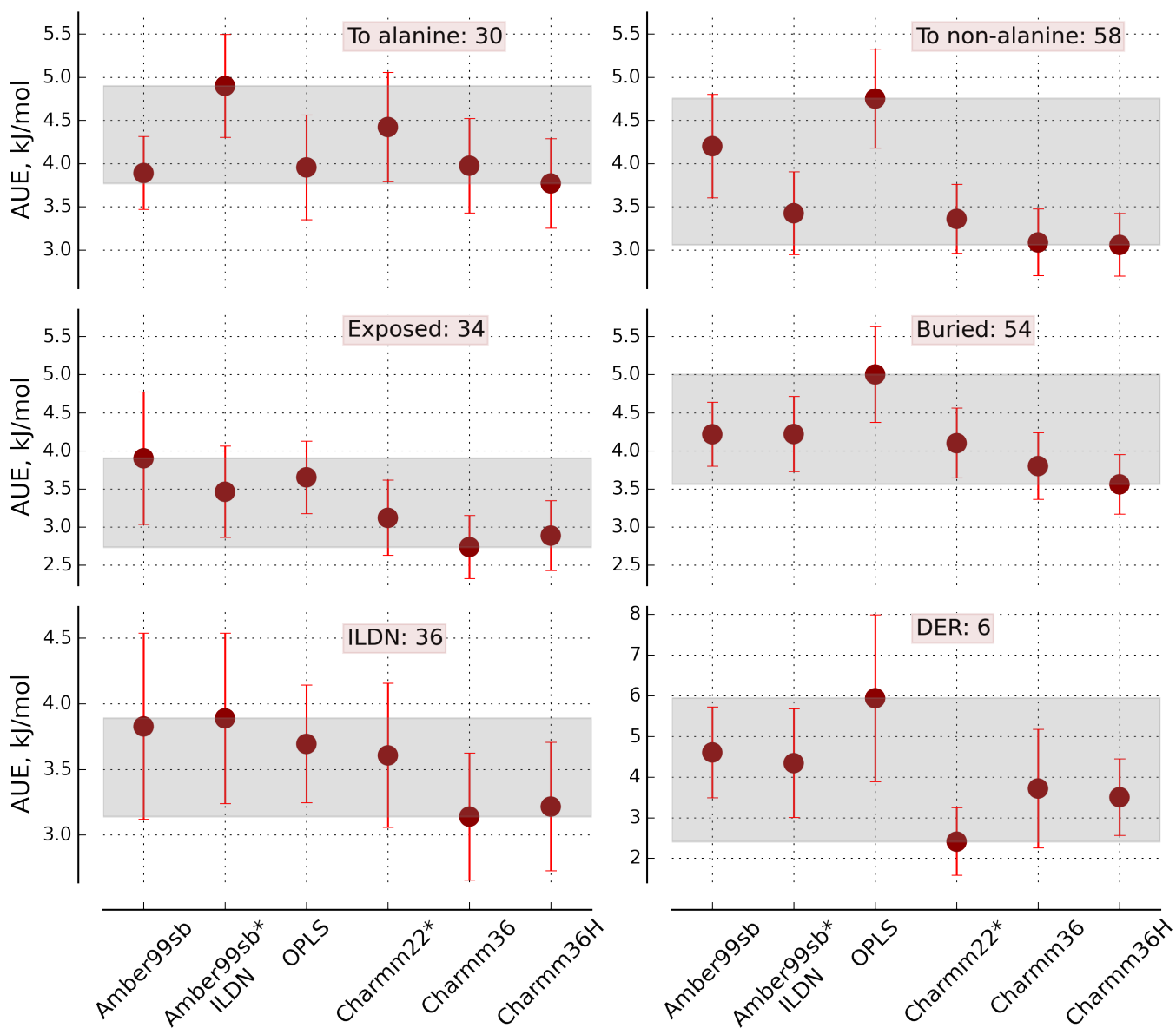


Figure S4: Average unsigned errors of the calculated $\Delta\Delta G$ values from experiment for mutation subsets grouped by mutating to alanine/non-alanine (top row), exposure to solvent (middle row), involved residues ILDN/DER (bottom row)

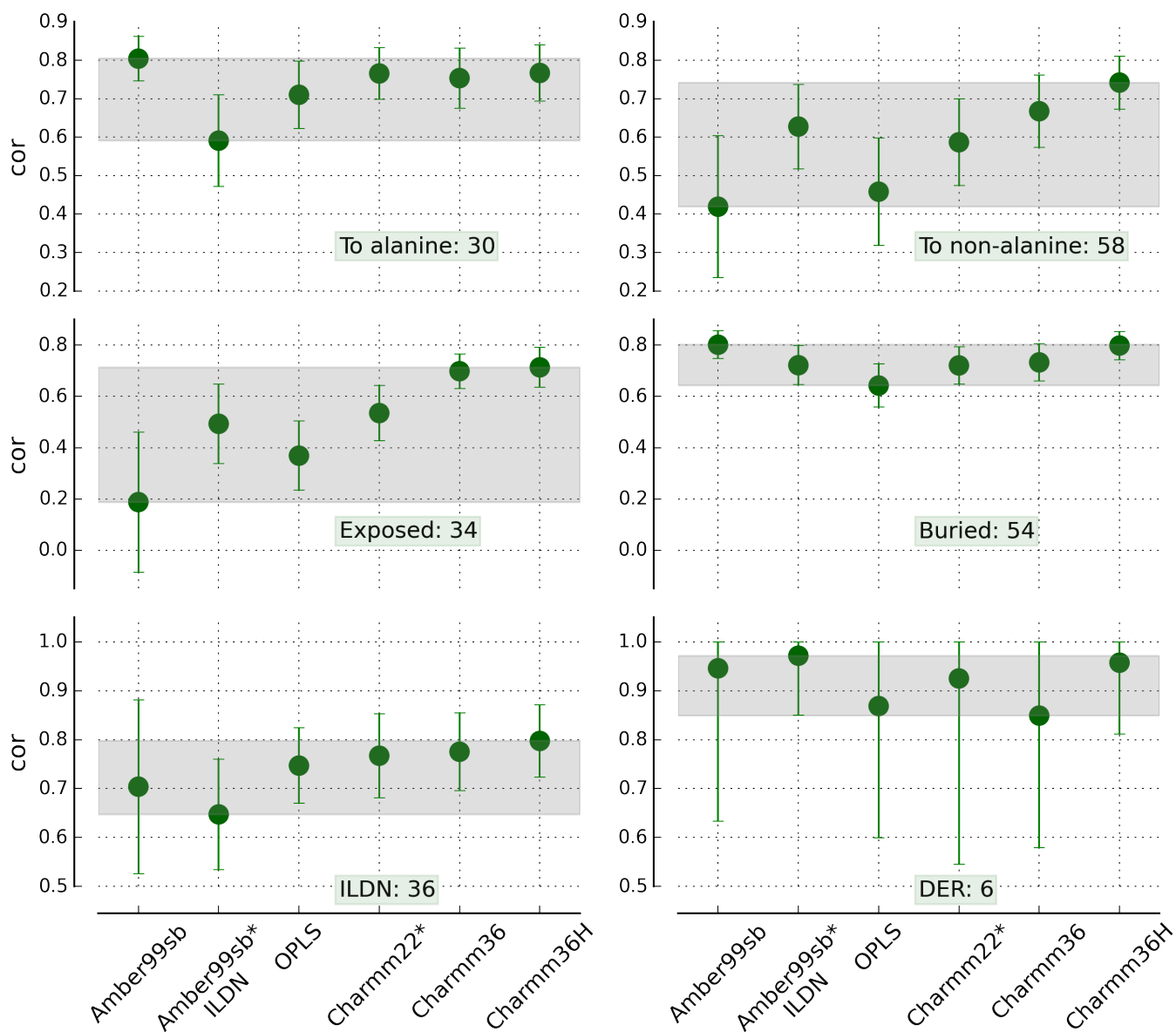


Figure S5: Correlations of the calculated $\Delta\Delta G$ values to experiment for mutation subsets grouped by mutating to alanine/non-alanine (top row), exposure to solvent (middle row), involved residues ILDN/DER (bottom row)

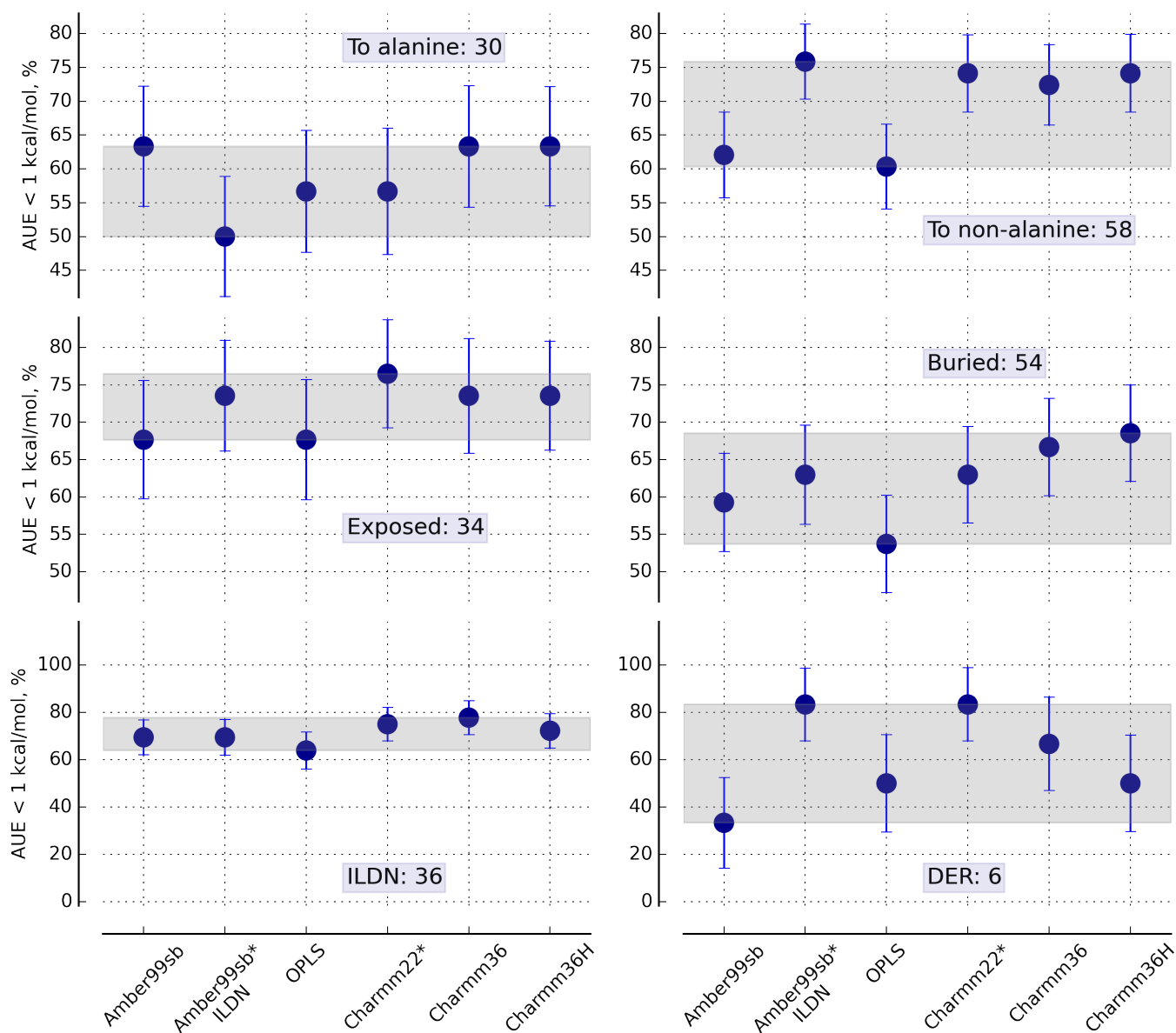


Figure S6: Percentage of the $\Delta\Delta G$ values predicted with 1 kcal/mol accuracy with respect to experiment for the mutation subsets grouped by mutating to alanine/non-alanine (top row), exposure to solvent (middle row), involved residues ILDN/DER (bottom row)

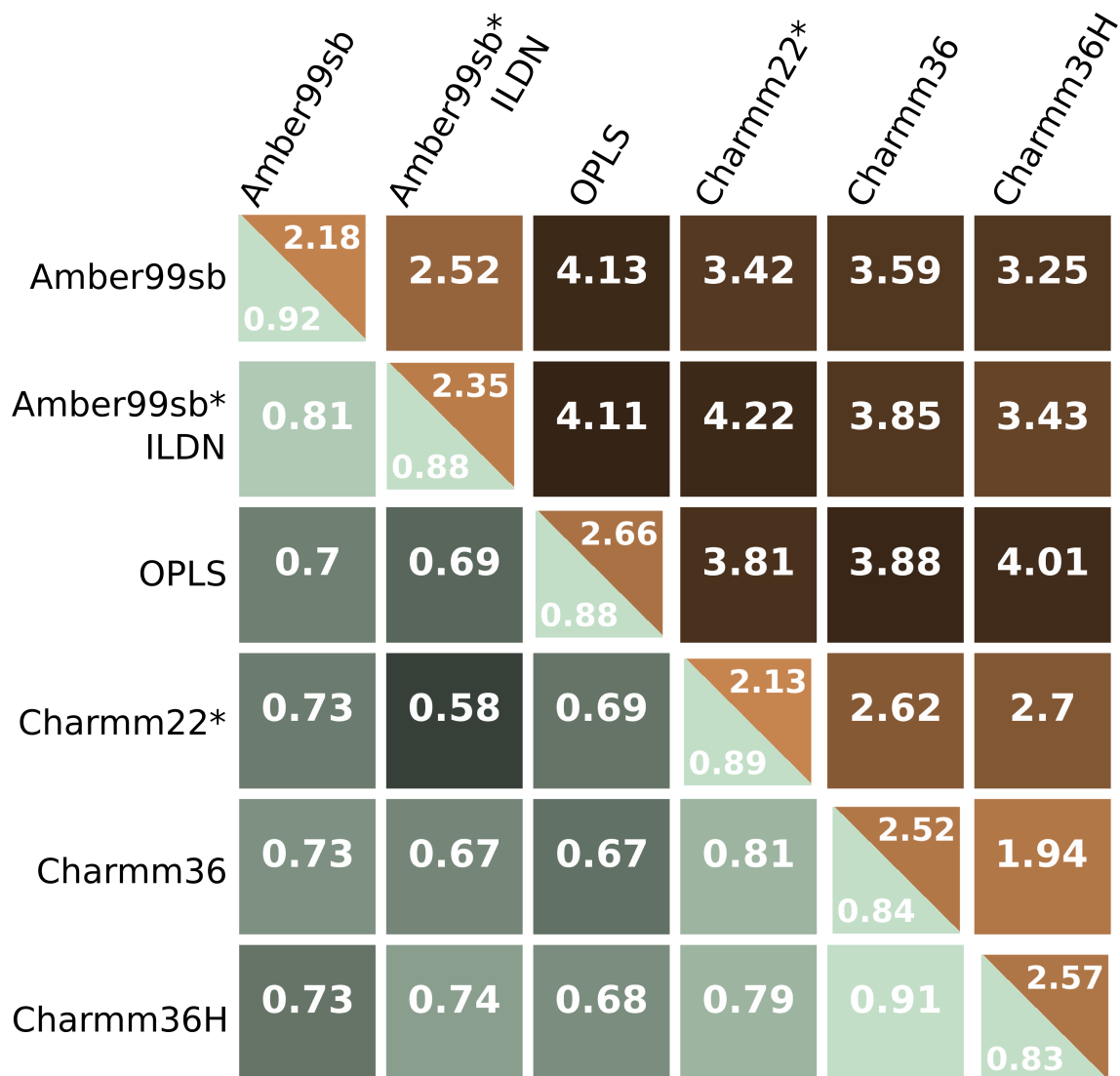


Figure S7: Average unsigned errors and correlations between the force fields (off-diagonal elements) and within the force fields (diagonal elements) for the charge conserving mutations.

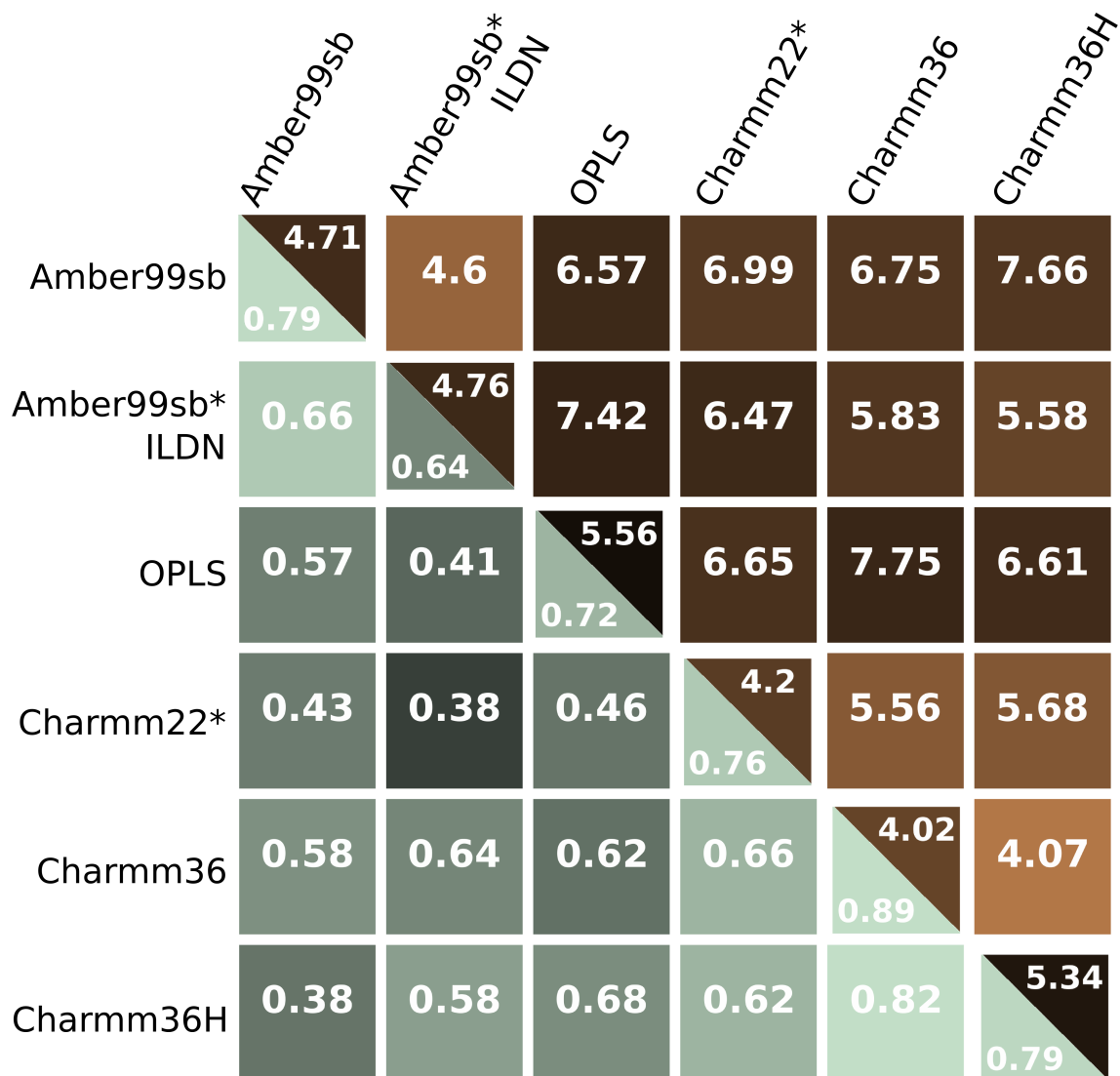


Figure S8: Average unsigned errors and correlations between the force fields (off-diagonal elements) and within the force fields (diagonal elements) for the charge changing mutations.

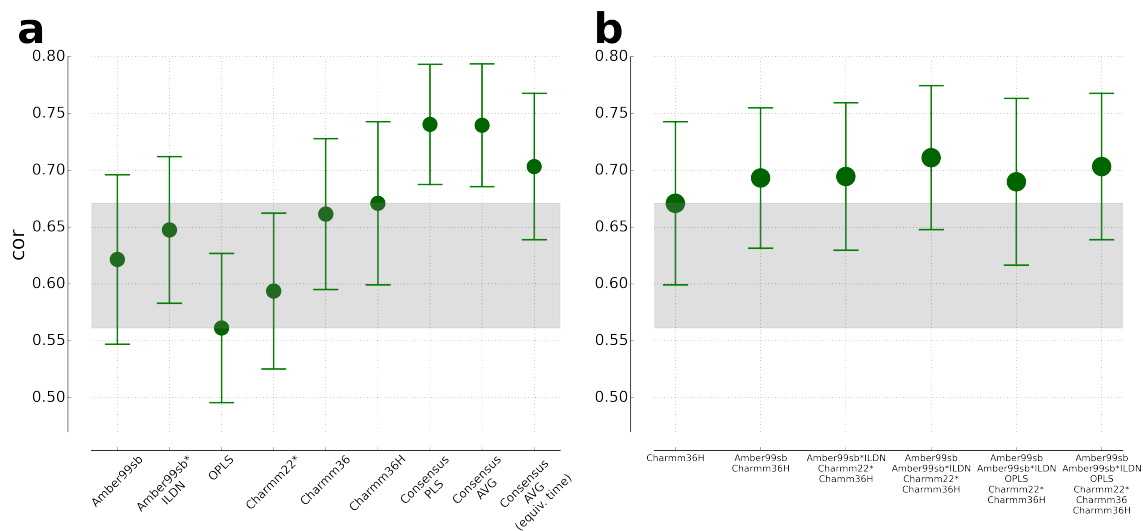


Figure S9: Correlations of the calculated $\Delta\Delta G$ values to experiment for individual force fields and their combinations. **(a)** Values for individual force fields, consensus model built as a PLS regression, consensus model built by averaging and consensus model built by averaging considering simulation time equivalent to a trajectory of a single force field. **(b)** The best performing individual force field and the other best performing combinations.

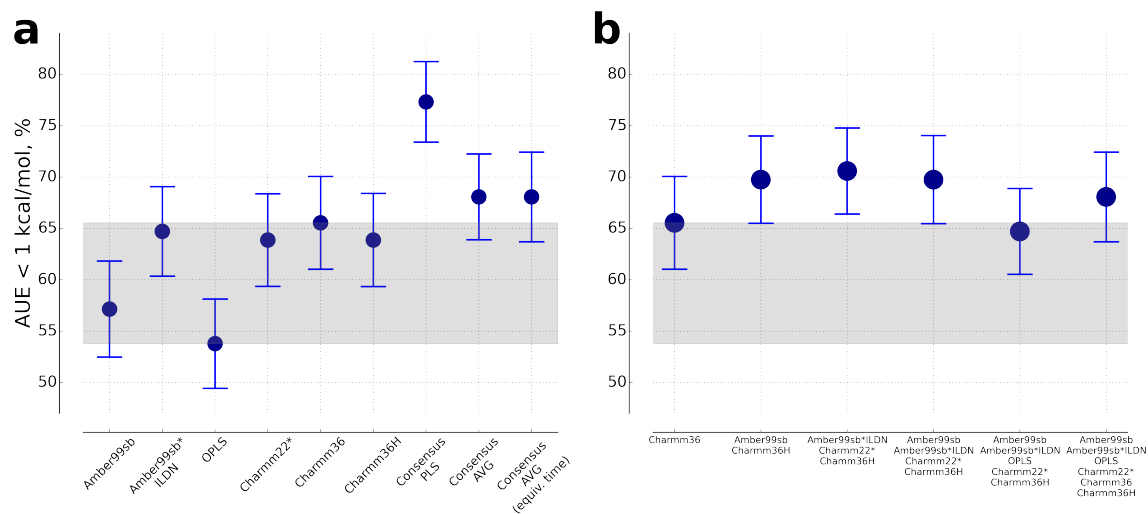


Figure S10: Percentage of the $\Delta\Delta G$ values predicted with 1 kcal/mol accuracy with respect to experiment for individual force fields and their combinations. (a) Values for individual force fields, consensus model built as a PLS regression, consensus model built by averaging and consensus model built by averaging considering simulation time equivalent to a trajectory of a single force field. (b) The best performing individual force field and the other best performing combinations.

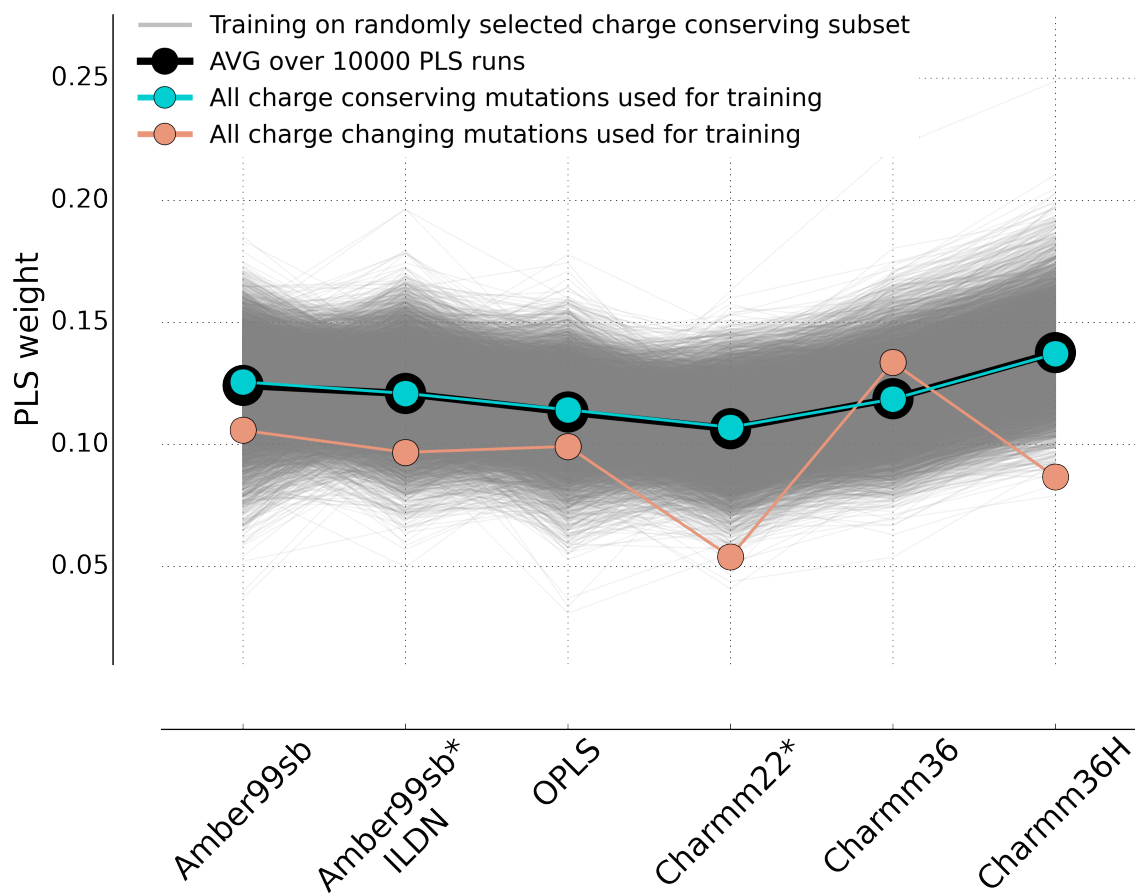


Figure S11: Weights assigned to the force fields by the partial least squares (PLS) regression model. Gray: 10000 PLS models built using by randomly selecting 44 amino acids from the charge conserving mutation set for training at a time. Black: average over the 10000 PLS models. Turquoise: PLS model built using all 88 charge conserving mutations. Salmon: PLS model built using all 31 charge changing mutations.

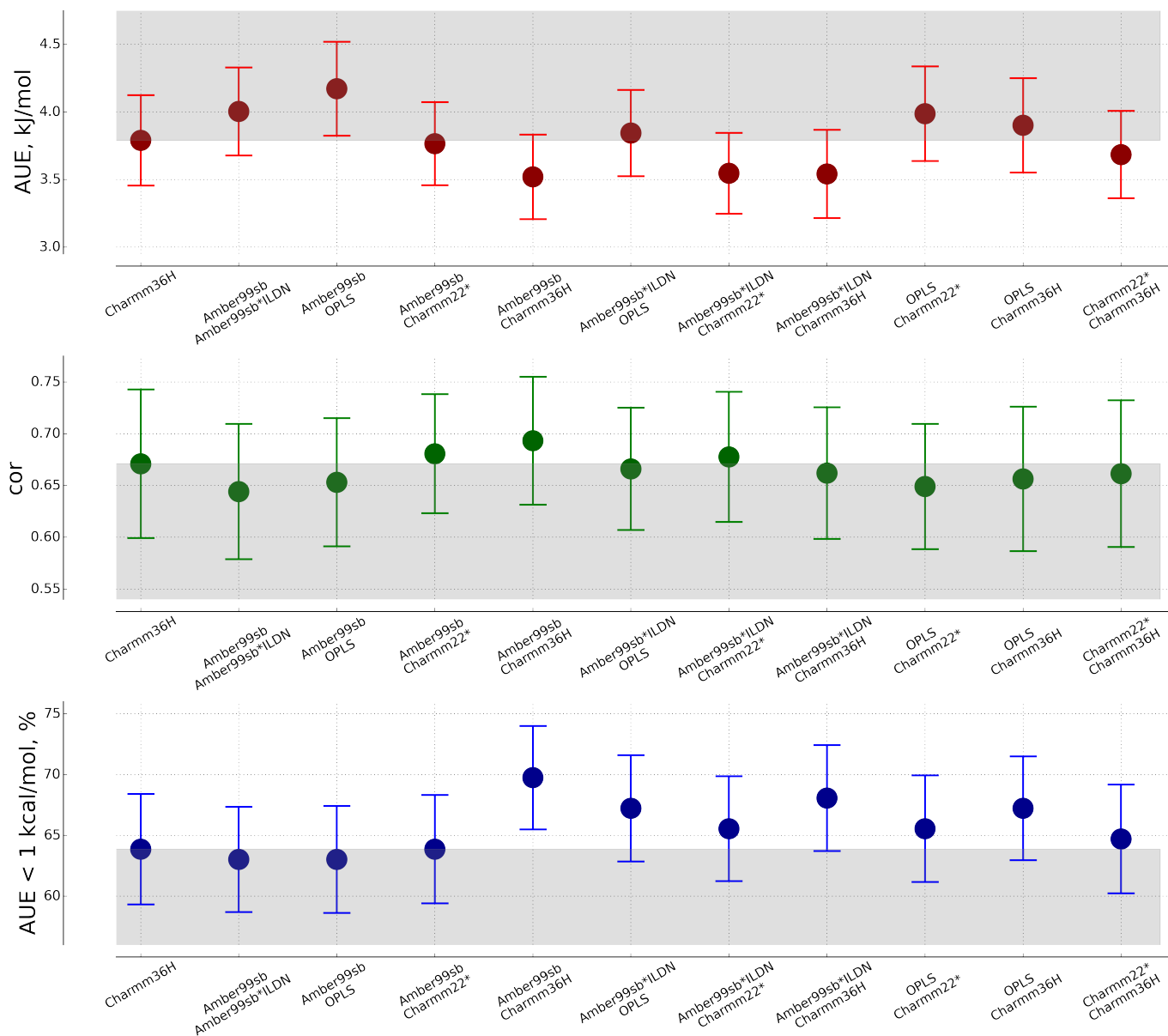


Figure S12: Consensus average models constructed from two force field combinations considering sampling time equivalent to that of a simulation with a single force field.

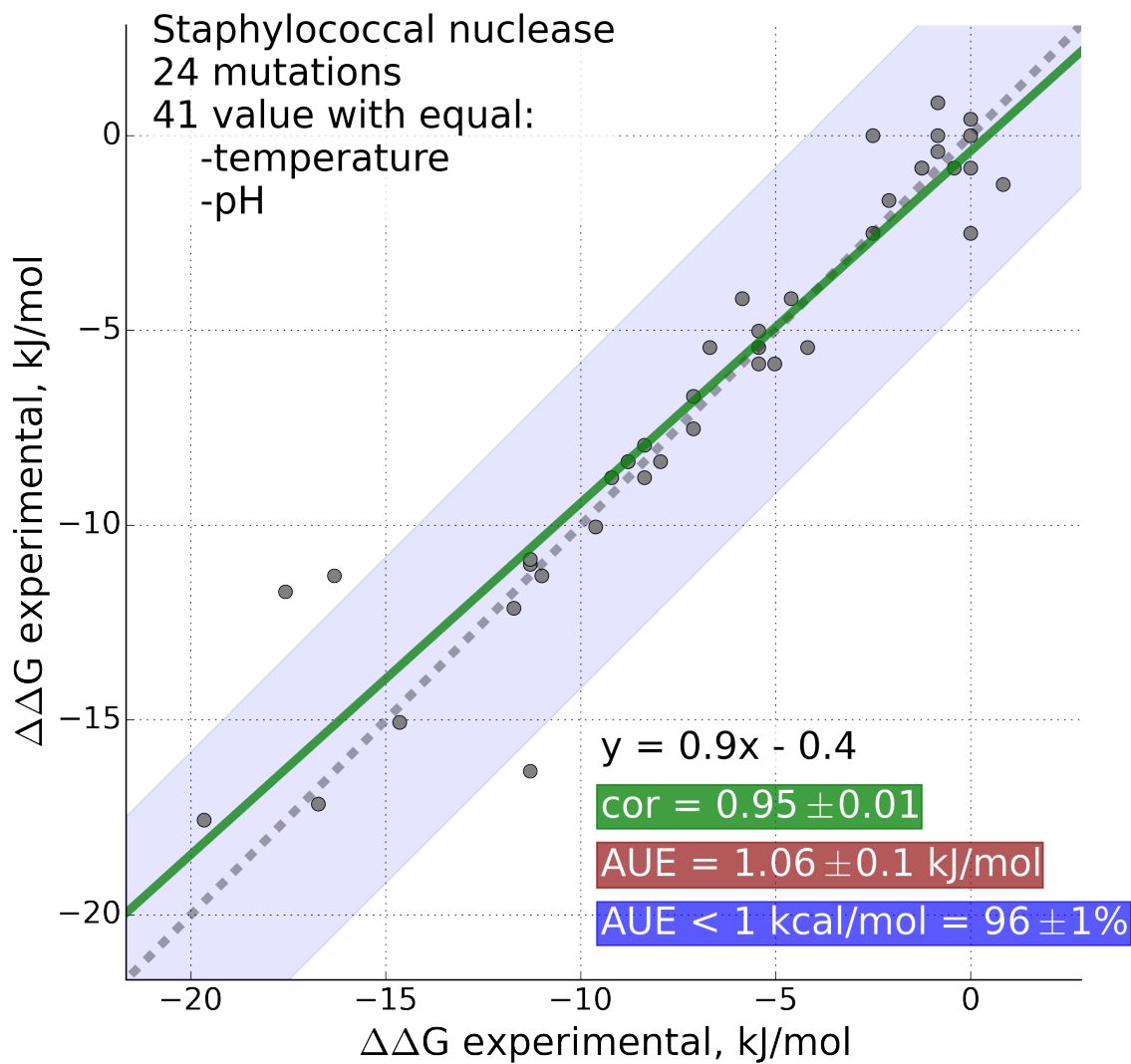


Figure S13: Experimentally measured double free energy differences for staphylococcal nuclease that were used in the current study. The values that were reported multiple times for the same mutation are plotted against one another. Experiments reporting on the same mutation were performed at identical temperatures and pH.

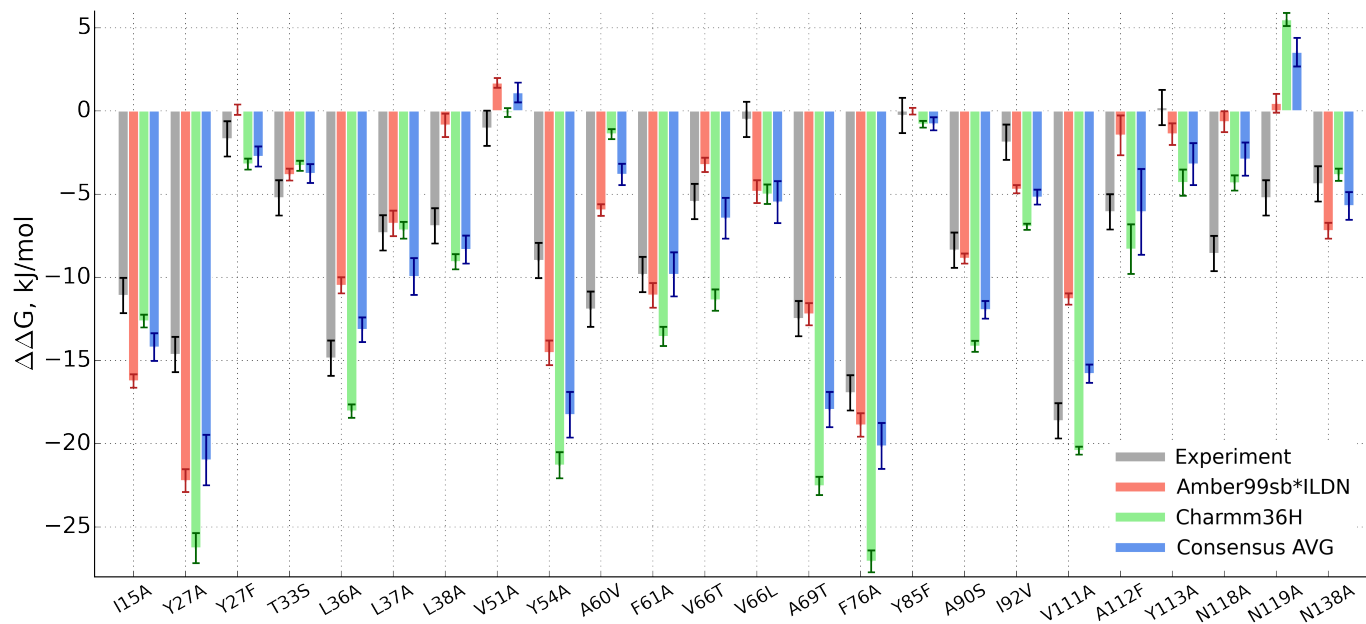


Figure S14: $\Delta\Delta G$ values for 24 charge conserving mutations in staphylococcal nuclease. Consensus AVG was built from the Amber99sb*ILDN and Charmm36H results considering the combined sampling time equivalent to that of a simulation with a single force field. Experimental values are depicted with the 1.06 kJ/mol errorbars dictated by the AUE in Fig. S13 in the Supporting Information.

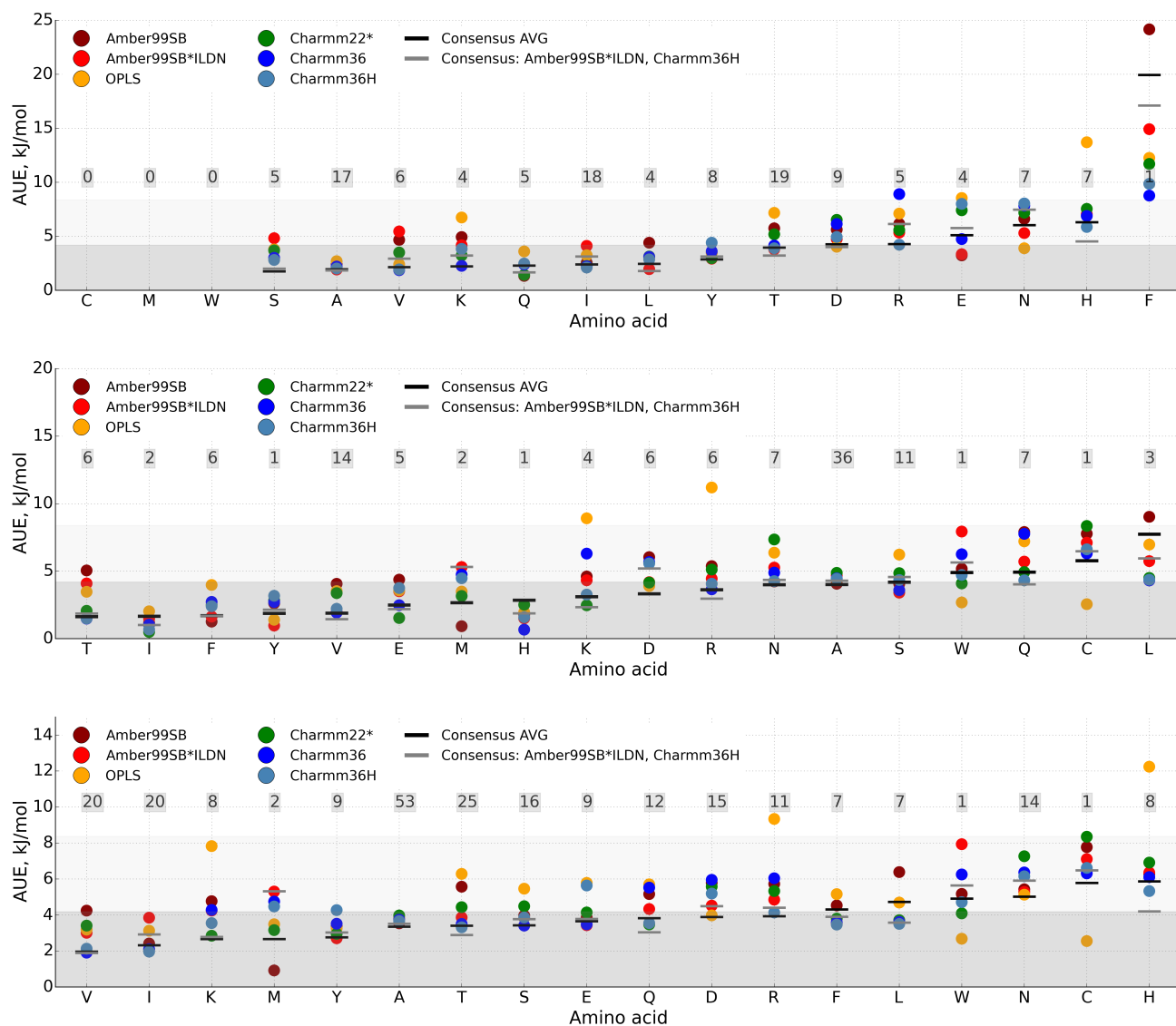


Figure S15: AUE values averaged over every amino acid extracted from the barnase free energy calculations using all 119 mutations for 6 force fields and 2 consensus models. Top panel: averaging over AUE for the amino acids that serve as a source for mutations, e.g. the Ala case comprises 17 mutations from alanine to any residue. Middle panel: averaging over AUE for the amino acids that serve as a target for mutations, e.g. the Ala case comprises 36 mutations from any residue to alanine. Bottom panel: averaging over all the instances of every amino acid, e.g. the Ala case comprises 53 mutations where either alanine is mutated to any residue, or any residue is mutated to alanine. Numbers in the panels depict the number of occurrences for a given amino acid in the mutation scan. Amino acids are sorted by the Consensus AVG AUE values in ascending order.

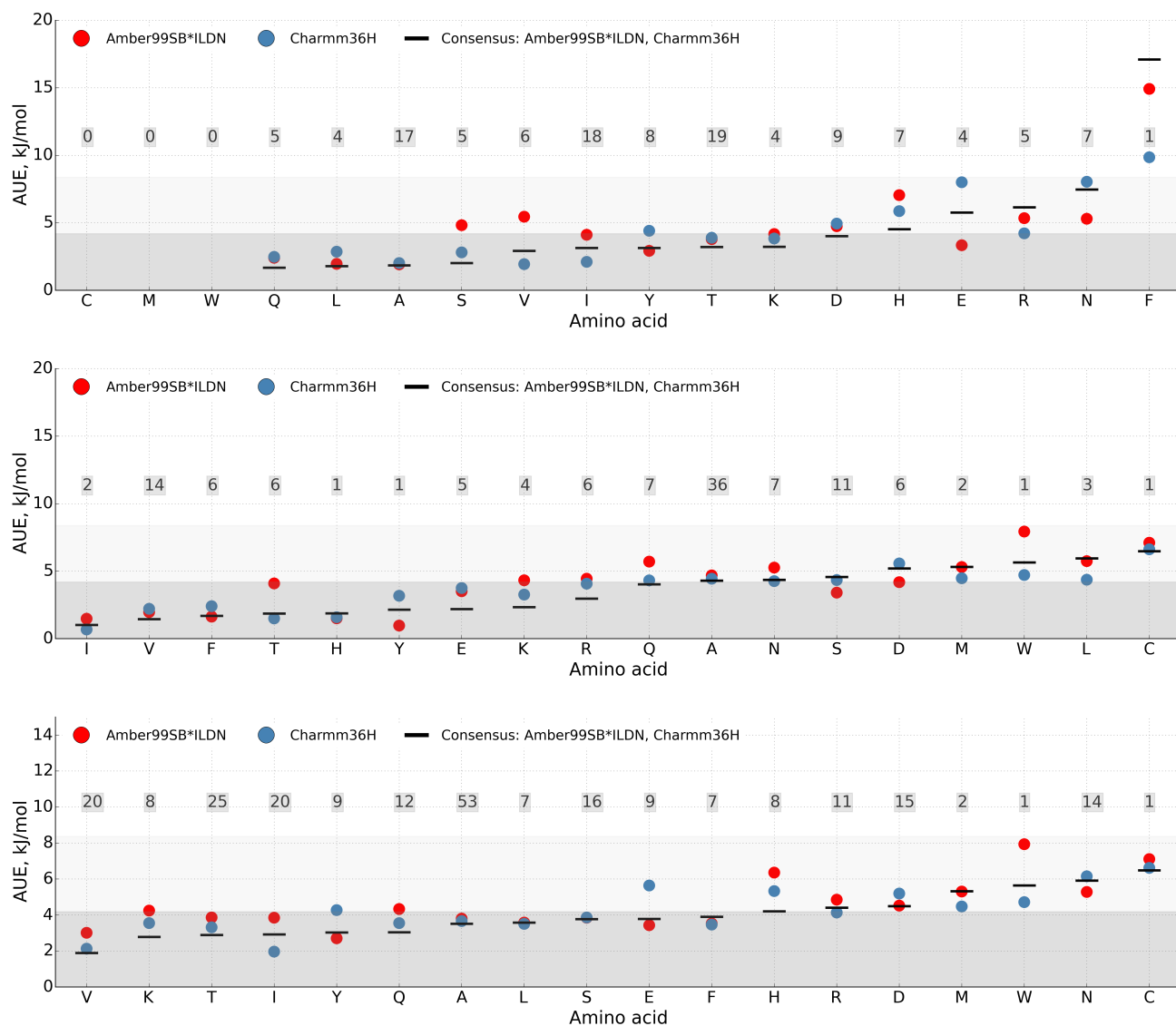


Figure S16: AUE values averaged over every amino acid extracted from the barnase free energy calculations using all 119 mutations for 2 force fields and 1 consensus model. Top panel: averaging over AUE for the amino acids that serve as a source for mutations, e.g. the Ala case comprises 17 mutations from alanine to any residue. Middle panel: averaging over AUE for the amino acids that serve as a target for mutations, e.g. the Ala case comprises 36 mutations from any residue to alanine. Bottom panel: averaging over all the instances of every amino acid, e.g. the Ala case comprises 53 mutations where either alanine is mutated to any residue, or any residue is mutated to alanine. Numbers in the panels depict the number of occurrences for a given amino acid in the mutation scan. Amino acids are sorted by the two force field consensus AUE values in ascending order.

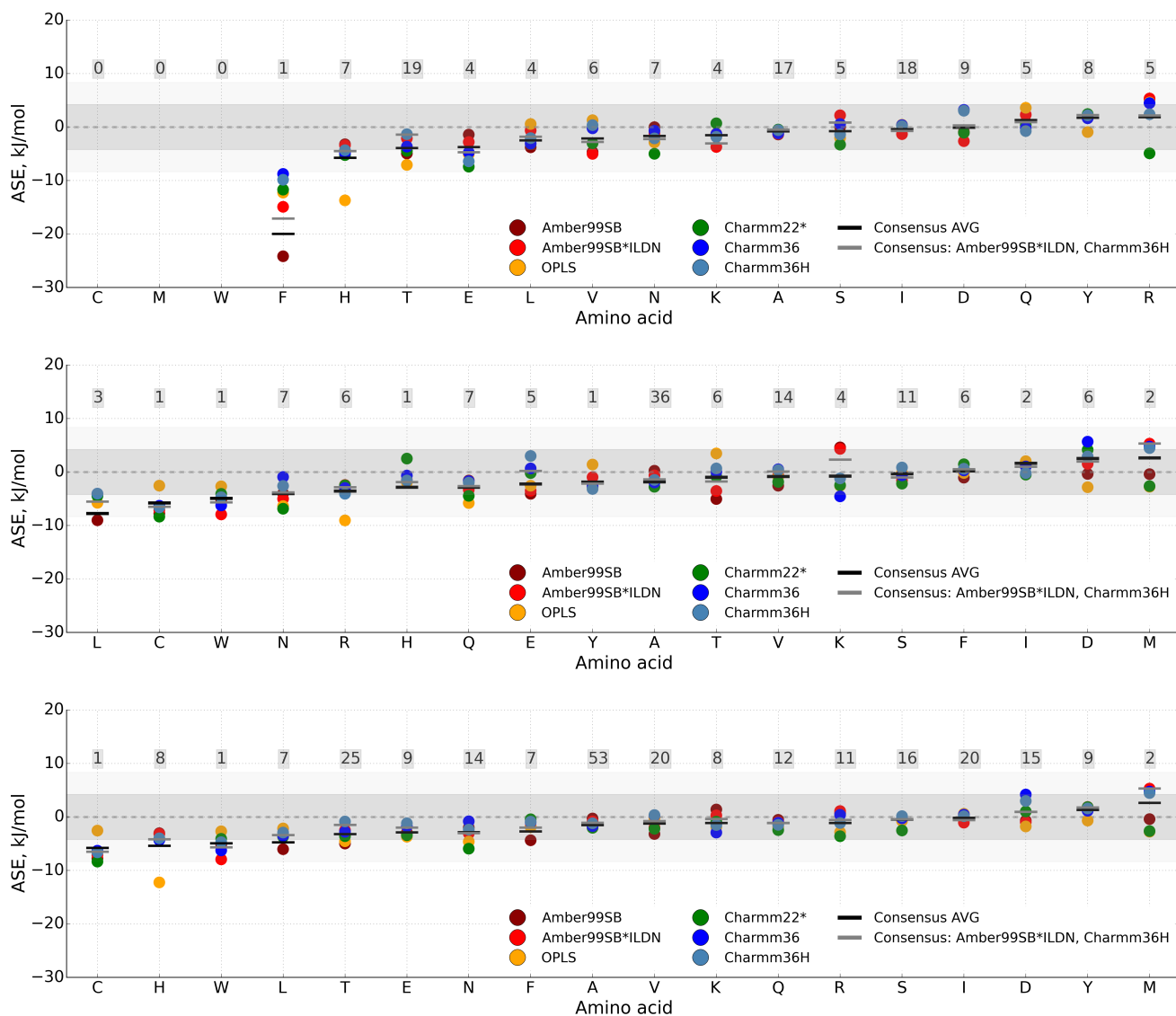


Figure S17: Average signed error ($ASE = \Delta\Delta G_{exp} - \Delta\Delta G_{calc}$) values averaged over every amino acid extracted from the barnase free energy calculations using all 119 mutations for 6 force fields and 2 consensus models. Top panel: averaging over ASE for the amino acids that serve as a source for mutations, e.g. the Ala case comprises 17 mutations from alanine to any residue. Middle panel: averaging over ASE for the amino acids that serve as a target for mutations, e.g. the Ala case comprises 36 mutations from any residue to alanine. Bottom panel: averaging over all the instances of every amino acid, e.g. the Ala case comprises 53 mutations where either alanine is mutated to any residue, or any residue is mutated to alanine. Numbers in the panels depict the number of occurrences for a given amino acid in the mutation scan. Amino acids are sorted by the Consensus AVG ASE values in ascending order.

Mutation	Expt	Amber99sb	Amber99sb*	OPLS	Charmm22*	Charmm36	Charmm36H
ILDN							
I4A	-4.30	-2.53	-1.24	-6.90	-3.69	-6.30	-2.90
I4V	-3.00	2.25	0.80	-5.28	-4.85	-2.59	-3.24
N5A	-7.67	-4.68	-2.20	-1.93	-0.53	0.13	0.12
T6A	-9.34	-3.59	-3.93	-0.61	-1.80	2.86	-2.22
T6D	0.46	0.26	-0.94	3.26	-2.58	-0.40	2.58
T6E	-1.13	5.40	5.57	5.65	-1.77	-1.36	-3.20
T6N	-5.31	1.33	1.38	4.13	-0.08	-4.00	1.45
T6Q	-7.82	1.27	-0.02	6.66	-3.42	5.83	0.67
T6S	-0.92	0.85	0.75	0.07	2.71	4.23	2.12
F7L	-17.15	7.01	-2.24	-4.89	-5.45	-8.39	-7.30
D8A	-3.95	-12.28	-6.81	-5.56	-13.75	-6.72	-6.71
D8S	-2.93	-11.10	-7.94	-3.51	-6.20	-8.60	-9.40
V10A	-14.64	-11.59	-10.35	-15.19	-5.51	-14.11	-16.23
V10T	-9.83	-1.19	2.98	-15.62	-11.26	-11.99	-13.80
D12A	0.12	-3.53	-0.61	-4.11	-8.96	-12.97	-10.65
D12S	-3.77	-4.89	-2.70	-10.13	-6.22	-10.45	-13.66
Y13A	-14.76	-22.95	-21.67	-12.60	-14.74	-12.56	-10.89
Y13F	-2.64	-2.91	-3.09	-4.29	-2.28	-2.67	-3.37
L14A	-19.14	-20.38	-20.78	-20.27	-13.24	-11.19	-13.56
Q15A	-0.84	-1.70	-4.73	-2.87	-0.79	1.28	1.13
Q15I	5.77	3.66	3.05	3.13	6.11	5.07	6.67
T16A	-2.09	-6.92	-5.28	2.43	0.67	-1.20	-2.18
T16R	1.27	12.25	4.79	21.61	7.52	4.12	2.35
T16S	-6.91	1.03	1.79	2.14	-2.68	-1.45	-1.76
Y17A	-8.96	-13.91	-12.67	-12.02	-13.63	-17.81	-17.16
Y17F	-2.43	-1.93	-1.24	-2.03	-0.64	-1.00	-1.17
Y17S	-8.37	-16.41	-12.10	-5.29	-13.49	-11.50	-17.55
H18A	-8.47	-1.24	-0.79	7.05	-2.29	-4.59	-4.43
H18D	-9.20	-2.60	-7.69	-2.62	-17.24	-16.33	-14.57
H18K	-3.81	-10.86	-10.49	14.36	4.67	3.99	0.66
H18N	-7.53	-8.33	-4.08	-1.09	2.04	3.74	2.53
H18Q	-6.36	2.20	3.02	9.61	0.08	-1.80	-4.32
H18R	-4.18	-9.23	-10.22	7.22	5.17	3.46	3.52
H18S	-8.79	4.28	5.80	13.08	-4.07	-2.89	-1.41
K19R	3.68	1.74	4.70	-2.77	1.57	2.55	3.89
N23A	-9.69	-1.62	-2.37	-10.34	-2.77	-0.85	-0.68
Y24F	0.38	0.04	-0.84	-5.14	-6.37	-7.70	-5.95
I25A	-15.40	-22.58	-24.33	-17.71	-18.26	-17.36	-18.79
I25V	-4.37	-8.42	-7.01	-8.36	-8.55	-7.68	-7.12
T26A	-8.21	-4.28	-6.63	0.73	-1.22	-2.65	-7.88
T26D	0.00	6.60	-7.57	6.11	-3.20	-1.50	-4.90
T26E	-0.21	-0.95	1.02	3.39	2.48	3.01	-8.39
T26N	-5.40	-3.77	-2.59	1.63	-2.08	-1.46	-4.88
T26Q	-7.20	2.20	-0.65	2.11	8.25	-0.44	2.85
T26S	-2.34	-0.94	-2.33	-2.55	1.07	2.49	0.12
T26V	-9.67	-6.14	-7.39	-3.25	-3.72	-11.69	-9.02

Mutation	Expt	Amber99sb	Amber99sb*	OPLS	Charmm22*	Charmm36	Charmm36H
ILDN							
S28A	2.51	-0.27	-0.32	0.22	2.16	0.99	1.03
S28E	2.55	9.06	9.12	7.54	1.93	-3.05	0.78
E29A	-5.36	-7.93	-4.75	-4.80	0.30	-4.37	-2.73
E29Q	-0.92	1.93	1.79	0.14	3.55	2.96	-2.67
E29S	-5.02	-6.08	-6.11	-11.94	-1.67	-4.84	-6.44
Q31A	-0.77	0.20	-2.60	-2.80	1.54	2.93	3.36
Q31S	0.91	0.86	1.15	-5.01	-2.42	1.51	1.98
A32C	-7.74	0.03	-0.63	-5.19	0.61	-1.43	-1.12
A32D	-1.72	4.27	-1.00	-1.91	-4.72	-1.91	-1.44
A32E	0.54	2.74	0.78	-1.37	-1.52	-1.48	-4.30
A32F	-1.80	-0.85	-1.07	-6.00	-0.85	-0.20	-1.41
A32H	-2.93	-1.16	-1.42	-0.97	-5.44	-2.26	-1.34
A32I	-2.85	-3.08	-2.61	-4.25	-2.20	-4.21	-3.31
A32K	-0.42	-5.44	-3.55	4.38	-0.18	3.48	-1.17
A32L	0.33	2.19	0.24	-1.48	0.28	-0.06	-0.15
A32M	-0.33	-0.85	-0.36	-1.07	-0.86	-1.03	-2.54
A32N	-1.38	0.06	-0.12	0.07	-3.05	-0.18	-1.34
A32Q	-1.46	-0.91	1.73	-0.06	-0.49	0.68	-0.49
A32R	-0.84	-0.05	0.46	12.24	2.56	3.04	3.47
A32S	-0.67	0.30	0.38	-2.50	-1.22	-1.54	-2.24
A32T	-2.05	-1.77	-1.11	-4.51	-4.28	-3.45	-2.99
A32V	-1.42	-4.45	-3.71	-3.35	-2.41	-3.20	-2.39
A32W	-4.77	0.40	3.16	-2.09	-0.68	1.48	-0.06
A32Y	-2.93	-0.31	-1.96	-4.33	-0.09	-0.21	0.25
L33Q	-5.67	-2.92	-3.71	-2.36	-4.56	-3.49	-5.15
V36A	-5.98	0.87	-0.86	-4.39	-2.94	-1.87	-3.28
V36T	-4.67	-1.25	-2.94	-8.62	-3.63	-5.66	-3.89
N41D	-10.67	-21.96	-4.95	-5.94	-12.81	-17.16	-5.07
D44E	-1.34	4.54	1.47	-1.94	0.28	-0.04	0.52
V45A	-6.58	-6.45	-7.89	-4.82	-2.74	-8.32	-7.90
V45T	-9.73	-3.84	-2.30	-10.50	-7.22	-8.08	-8.48
I51A	-19.75	-22.63	-4.24	-14.10	-22.10	-23.23	-22.82
I51V	-6.67	-7.33	-6.40	-4.10	-5.97	-9.65	-3.25
D54A	-12.79	-8.49	-5.16	-10.09	-11.92	-3.56	-13.05
D54N	-9.93	-7.07	8.05	-9.13	1.77	-7.29	-4.48
I55A	-5.77	-6.06	-9.85	-10.37	-4.91	-4.00	-7.19
I55T	-4.69	-2.67	-5.32	-6.57	-2.06	-3.80	-4.24
I55V	-1.74	0.04	-2.22	-3.76	0.45	-0.90	-0.17
N58A	-8.84	-15.08	-15.16	-5.04	-1.26	-9.46	-14.44
N58D	1.49	-4.07	-6.70	-1.49	-4.02	-16.38	-13.66
K62R	-1.26	3.00	2.89	1.99	2.59	4.30	2.99
K66A	4.10	-0.20	3.21	-0.50	5.30	5.63	0.09
R69K	-13.10	-17.57	-16.82	-24.19	-12.78	-3.08	-8.65
R69M	-8.87	-7.55	-19.45	-2.63	-3.08	-17.67	-15.61
R69S	-11.38	-10.36	-11.69	-23.07	7.93	-10.30	-11.28
E73A	-10.04	-3.60	-1.10	15.61	6.12	3.87	16.18

Mutation	Expt	Amber99sb	Amber99sb*	OPLS	Charmm22*	Charmm36	Charmm36H
ILDN							
D75N	-20.50	-5.92	-16.75	-4.28	-8.40	-30.97	-26.30
I76A	-6.78	-4.52	4.98	-0.77	-3.38	-1.42	-0.76
I76V	-3.74	-3.00	-3.34	-0.78	-1.63	-3.85	-4.26
N77A	-6.96	-1.00	-4.66	-2.05	7.24	0.23	-0.83
Y78F	-5.56	-0.07	-3.96	1.41	-8.16	-1.30	-1.84
T79V	0.75	-1.02	-1.39	0.04	3.08	2.33	2.61
R83K	-17.28	-19.15	-21.04	-18.87	-16.44	-20.77	-20.67
R83Q	-8.58	-30.65	-16.93	-13.49	-10.24	-29.73	-14.98
N84A	-8.16	-1.92	-6.37	-3.73	-1.37	-2.13	-1.14
S85A	1.17	-1.83	-3.79	-0.88	1.14	2.02	0.84
I88A	-16.02	-19.34	-8.91	-15.38	-19.64	-12.69	-16.83
I88L	-0.08	0.97	-2.30	6.75	1.65	3.68	2.67
I88V	-6.21	-7.22	-4.95	-10.08	-9.77	-9.81	-7.99
L89T	-12.18	-2.11	-13.16	-18.15	-9.64	-10.39	-13.76
L89V	-1.14	2.41	2.10	0.37	1.82	-0.71	2.61
S91A	-8.76	-12.44	-16.62	-0.19	0.44	-10.57	-7.90
S92A	-12.22	-9.26	-14.13	-11.21	-3.97	-6.97	-2.62
D93N	-17.20	-15.53	-18.09	-14.01	-9.38	-20.63	-16.00
I96A	-13.17	-16.97	-17.89	-19.18	-17.34	-15.69	-13.81
I96V	-3.87	0.41	-4.07	-7.39	-3.63	-5.86	-6.54
T99V	-12.36	5.66	-13.84	-4.02	-3.78	-6.85	-16.32
Y103F	-0.38	-0.36	-4.97	4.74	-2.83	-1.32	-2.36
Q104A	-0.38	-3.04	-3.74	-5.76	-1.52	-5.16	-4.69
T105V	-9.26	-1.03	-7.98	-0.80	-0.42	-8.09	-14.37
K108R	3.26	12.49	13.88	15.93	-2.49	2.43	10.15
I109A	-6.82	-9.48	-8.04	-6.78	-6.03	-6.41	-10.15
I109V	-3.33	-2.38	2.43	-2.62	-0.59	-4.58	-1.63

Table S2: Experimental and calculated $\Delta\Delta G$ values in kJ/mol for the mutations in barnase.

Mutation	Expt	Amber99sb*	Charmm36H
		ILDN	
I10A	-11.09	-14.95	-11.88
Y22A	-14.64	-23.06	-23.77
Y22F	-1.67	0.44	-3.06
T28S	-5.23	-3.86	-3.40
L31A	-14.85	-11.86	-16.46
L32A	-7.32	-5.11	-5.59
L33A	-6.90	-0.69	-7.49
V46A	-1.05	0.55	0.55
Y49A	-9.00	-6.82	-18.78
A55V	-11.92	-4.01	-2.04
F56A	-9.83	-10.98	-12.72
V61L	-0.50	-6.95	-6.58
V61T	-5.44	-2.33	-11.69
A64T	-12.48	-12.94	-23.14
F71A	-16.95	-18.67	-26.24
Y80F	-0.28	-0.26	-0.67
A85S	-8.37	-8.86	-15.13
I87V	-1.88	-3.36	-6.19
V106A	-18.62	-11.93	-19.77
A107F	-6.07	-2.75	-9.15
Y108A	0.21	0.15	-1.80
N113A	-8.58	-0.64	-4.05
N114A	-5.23	0.51	5.77
N133A	-4.39	-7.93	-3.55

Table S3: Experimental and calculated $\Delta\Delta G$ values in kJ/mol for the mutations in staphylococcal nuclease.

Mutation	$\Delta\Delta G_{\text{calc}}$, kJ/mol	$\Delta T_{m\text{exp}}$, K
H103A	-3.48	1
A86L	7.13	3
F358A	13.99	8
A86L/F358A	14.87	10
A86L/I253A/F358V	20.24	22

Table S4: Calculated $\Delta\Delta G$ values in kJ/mol together with the experimentally measured changes in the melting temperature for neurotensin receptor 1.

Protein	$\text{cor}(\Delta T_m, \Delta\Delta G)$
Ribose-binding protein	0.95
Lysozyme	1.00
Rubredoxin	1.00
SNase	0.98
Fibroblast growth factor	1.00

Table S5: Correlations between the experimentally measured ΔT_m and $\Delta\Delta G$ values extracted from literature. Literature sources are listed in the Supplementary Information text on page S4.

	1CHO cor	1CHO AUE, kJ/mol	2DQJ cor	2DQJ AUE, kJ/mol
CC/PBSA	0.89	6.32	0.07	3.03
Rosetta (row 3)	0.66	-	0.26	-
Rosetta (row 6)	0.89	-	0.45	-
Rosetta (row 16)	0.91	-	0.62	-
Amber99sb*ILDN	0.35	10.27	0.51	6.83
Charmm36H	0.80	6.28	0.46	6.07
ConsensusFF	0.74	7.97	0.51	5.82

Table S6: Correlations and AUE for the calculated $\Delta\Delta G$ values and the experimental measurements. 1CHO: chymotrypsin and turkey ovomucoid third domain complex. 2DQJ: hyhel-10 FV and lysozyme complex. ConsensusFF is calculated as an average of the results of two force fields effectively reducing force field and sampling related artefacts. CC/PBSA refers to the Concoord/Poisson-Boltzmann surface area approach [41].

References

- [1] D. Seeliger, B. L. de Groot, *Biophys. J.* **2010**, *98*, 2309–2316.
- [2] V. Gapsys, S. Michielssens, J. H. Peters, B. L. de Groot, H. Leonov in *Molecular Modeling of Proteins*, Springer, **2015**, pp. 173–209.
- [3] G. E. Crooks, *Phys. Rev. E* **1999**, *60*, 2721–2726.
- [4] M. R. Shirts, E. Bair, G. Hooker, V. S. Pande, *Phys. Rev. Lett.* **2003**, *91*, 140601.
- [5] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Bioinf.* **2006**, *65*, 712–725.
- [6] R. B. Best, G. Hummer, *J. Phys. Chem. B* **2009**, *113*, 9004–9015.
- [7] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins Struct. Funct. Bioinf.* **2010**, *78*, 1950–1958.
- [8] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- [9] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- [10] S. Piana, K. Lindorff-Larsen, D. E. Shaw, *Biophys. J.* **2011**, *100*, L47–L49.
- [11] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, A. D. MacKerell Jr, *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- [12] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [13] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. Shirts, J. Smith, P. Kasson, D. van der Spoel, B. Hess, E. Lindahl, *Bioinformatics* **2013**, *29*, 845–854.
- [14] V. Gapsys, S. Michielssens, D. Seeliger, B. L. de Groot, *J. Comput. Chem.* **2015**, *36*, 348–354.
- [15] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- [16] V. Gapsys, D. Seeliger, B. L. de Groot, *J. Chem. Theory Comput.* **2012**, *8*, 2373–2382.
- [17] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- [18] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.
- [19] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- [20] A. M. Buckle, K. Henrick, A. R. Fersht, *J. Mol. Biol.* **1993**, *234*, 847–860.
- [21] T. R. Hynes, R. O. Fox, *Proteins Struct. Funct. Bioinf* **1991**, *10*, 92–105.
- [22] P. Egloff, M. Hillenbrand, C. Klenk, A. Batyuk, P. Heine, S. Balada, K. M. Schlinkmann, D. J. Scott, M. Schütz, A. Plückthun, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, E655–E662.

- [23] M. Fujinaga, A. R. Sielecki, R. J. Read, W. Ardelt, M. Laskowski, M. N. James, *J. Mol. Biol.* **1987**, *195*, 397–418.
- [24] M. Shiroishi, K. Tsumoto, Y. Tanaka, A. Yokota, T. Nakanishi, H. Kondo, I. Kumagai, *J. Biol. Chem.* **2007**, *282*, 6783–6791.
- [25] E. H. Kellogg, A. Leaver-Fay, D. Baker, *Proteins Struct. Funct. Bioinf.* **2011**, *79*, 830–838.
- [26] C. Martin, V. Richard, M. Salem, R. Hartley, Y. Mauguen, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **1999**, *55*, 386–398.
- [27] B.-H. Mevik, R. Wehrens, *Journal of Statistical Software* **2007**, *18*, 1–24.
- [28] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, **2013**.
- [29] D. G. Isom, E. Vardy, T. G. Oas, H. W. Hellinga, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4908–4913.
- [30] Y. Tian, PhD thesis, Duke University, **2007**.
- [31] W. J. Becktel, J. A. Schellman, *Biopolymers* **1987**, *26*, 1859–1877.
- [32] B. W. Matthews, H. Nicholson, W. J. Becktel, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6663–6667.
- [33] D. M. LeMaster, G. Hernández, *Biophys. Chem.* **2007**, *125*, 483–489.
- [34] D. Shortle, A. K. Meeker, E. Freire, *Biochemistry* **1988**, *27*, 4761–4768.
- [35] J. Lee, M. Blaber, *J. Mol. Biol.* **2009**, *393*, 113–127.
- [36] Y. Shibata, J. F. White, M. J. Serrano-Vega, F. Magnani, A. L. Aloia, R. Grisshammer, C. G. Tate, *J. Mol. Biol.* **2009**, *390*, 262–277.
- [37] W. Lu, I. Apostol, M. Qasim, N. Warne, R. Wynn, W. L. Zhang, S. Anderson, Y. W. Chiang, E. Ogin, I. Rothberg, et al., *J. Mol. Biol.* **1997**, *266*, 441–461.
- [38] K. Tsumoto, K. Ogasahara, Y. Ueda, K. Watanabe, K. Yutani, I. Kumagai, *J. Biol. Chem.* **1995**, *270*, 18551–18557.
- [39] K. Tsumoto, K. Ogasahara, Y. Ueda, K. Watanabe, K. Yutani, I. Kumagai, *J. Biol. Chem.* **1996**, *271*, 32612–32616.
- [40] A. Yokota, K. Tsumoto, M. Shiroishi, H. Kondo, I. Kumagai, *J. Biol. Chem.* **2003**, *278*, 5410–5418.
- [41] A. Benedix, C. M. Becker, B. L. de Groot, A. Caffisch, R. A. Böckmann, *Nat. Methods* **2009**, *6*, 3–4.