QresFEP: an automated protocol for free energy calculations of protein mutations in Q

Willem Jespers[a,1], Geir V. Isaksen[a,b,1], Tor A. H. Andberg[b], Silvana Vasile[a], Amber van Veen[a], Johan Åqvist,[a] Bjørn Olav Brandsdal[b,] and Hugo Gutiérrez-de-Terán[a,*]*

[a]Department of Cell and Molecular Biology, Biomedical Center, Uppsala University, Box 590, S-75124 Uppsala, Sweden

[b]Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Tromsø, N9037 Tromsø, Norway

[1] Authors contributed equally

[*] Corresponding authors: BOB (bjorn-olav.brandsdal@uit.no); HGT (hugo.gutierrez@icm.uu.se)

KEYWORDS. Free Energy Perturbation, Molecular Dynamics, site directed mutagenesis, GPCR, Protein stability, Ligand binding affinity.

Abstract

Predicting the effect of single-point mutations on protein stability or protein-ligand binding is a major challenge in computational biology. Free energy calculations constitute the most rigorous approach to this problem, though the estimation of converged values for amino acid mutations is still challenging. To overcome this limitation, we developed a tailored protocol to calculate free energy shifts associated to single point mutations, which is now revised and implemented in an API framework and the graphic interface QGui for the MD software Q. We herein describe the QligFEP protocol and benchmark it in several model systems, optimizing a number of variables concerning sampling or the performance of Zwanzig's exponential formula and Bennet's acceptance ratio methods. QligFEP shows an excellent performance on estimating the hydration free energies of amino acid side chain mimics, included their charged analogs which were elusive in alternative FEP strategies. We next examined the performance on a protein-ligand binding problem of pharmaceutical relevance, the neuropeptide Y_1 G protein-coupled receptor. Here, the calculations show very good agreement with the experimental effect of 16 mutations on the binding of antagonists BIBP3226, in line with our recent applications in this field. Finally, the characterization of 43 mutations of T4-lysozyme reveals the capacity of our protocol to assess variations of the thermal stability of proteins, achieving a similar performance to alternative FEP approaches. In summary, QresFEP is a robust, versatile and user-friendly computational FEP protocol to examine biochemical effects of single-point mutations with high accuracy.

Introduction

Computational modeling of the effect of point mutations on protein structure and function has become a powerful tool for both ligand and protein design.[1] Two groups of approaches exist to address this problem: The first group comprises statistical and empirical methods, which are fast to compute and do not necessarily consider the tree dimensional structure of the protein.[2,3] This allows the treatment of large datasets, although they require well curated training sets to create the underlying empirical functions,[4] and the physical interpretation of the predictions based on statistical analysis tends to be difficult. The second class of methods are physics-based, meaning that they explicitly consider protein structures and involve the computation of energy functions.[5] While they do not rely on the availability of training sets, the associated simulations are more computationally demanding. This is especially the case for rigorous first principle methods such as free energy perturbation (FEP) and related protocols. In essence, with FEP simulations one can estimate the shifts in biophysical properties, such as equilibrium constants or thermal stability, associated to a single point mutation. This is done by computing the relative changes in free energy between the wild type and the mutant protein, always considering a reference state in the framework of a thermodynamic cycle.[6]

The use of physics-based methods is gaining momentum, partially due to the increasing availability of computational power, which now allows the use of FEP methodologies in a routine way within the scope of pharmaceutical projects.[7] In addition, the increase in experimental structures deposited in the PDB provide the necessary starting points to perform these simulations for a wide variety of protein families.[8] Alternatively, the current structural coverage facilitates the generation of homology models of enough quality to be used in structure based drug design,[9] including traditionally difficult targets such as G protein-

coupled receptors (GPCRs).[10] In this scenario, an accurate evaluation of free energy shifts caused by protein mutations can provide the missing link between experimental mutagenesis data and their 3D structures. Despite the fact that one of the earliest applications of FEP in biochemistry was the estimation of the effect of single point mutants on binding and catalysis of subtilisin,[11] FEP simulations are predominantly used in ligand design, i.e. to explore chemical modifications of lead compounds.[7] The automation of FEP protocols to characterize ligand series has become of the outmost interest for both the academia and industry.[12–16] To this extent we recently implemented a protocol for small molecule FEP simulations in our open-access MD software package Q[17,18] called QligFEP, which showed an excellent performance comparable to existing commercial solutions.[19] This protocol allows the comparison between any pair of ligands within the vast space of drug-like molecules by using a dual topology approach, making it unnecessary to define the atoms that differ between the ligand pair explored.[19]

As opposed to ligand perturbations, the chemical space of amino acid sidechains is finite, while it includes many challenging cases (e.g His to Trp which changes the topology of the ring). While some recent FEP protocols to perform *in silico* mutations approach this problem with a dual topology setting,[20,21] large perturbations (e.g. Trp to Ala) proved challenging to converge.[21] Our alternative to this problem was introduced some years ago, and includes a collection of single-topology FEP recipes, carefully designed to improve convergence based on a single-topology, stepwise annihilation of the natural side chains to alanine.[5] In addition, this method could be generalized to account for any sidechain transformation, by joining two thermodynamic cycles through the common alanine leg.[22] We have extensively used this methodology to calculate ligand-binding free energy shifts due to point mutations of different nature.[5,22–27]

The potential of *in silico* mutagenesis can be extended to estimate properties such as protein stability,[21,28] protein-protein interactions (PPI's),[29] or ligand binding affinities to mutant receptors.[5,30] We herein present QresFEP, a protocol to account for all possible point mutations for natural amino acids (proline is excluded due to differences in the backbone dihedral), implemented in friendly GUI and API interfaces of the Q MD software.[19,31] The sampling parameters used in QresFEP are initially optimized on a data set of hydration free energies of amino acid side chain mimics. Thereafter we show the performance on two situations relevant for biochemistry and molecular biology: the effect of point mutations on ligand binding affinities to a drug target, in this case the human G protein-coupled receptor for neuropeptide Y type 1 (hY_1), and the thermal fluctuations reflecting protein stability of a panel of 43 mutations on T4-lysozyme.

Materials and Methods

QresFEP protocol

QresFEP is a pipeline for the automated setup and analysis of FEP simulations of amino acid sidechain mutations. It is based on the alanine scanning protocol initially proposed by us,[5] where the basic idea is to ensure convergence by designing a smooth and progressive annihilation of the sidechain atoms of the wild type (WT) residue to convert it into alanine (Ala-mutant). This is possible by using the charge-group division of atoms as defined in the OPLS-AA forcefield, and design a series of subperturbations that, starting from the group topologically most distant from the Cβ position, allow each group to undergo three consecutive transformations: *i*) annihilation of partial charges, *ii*) introduction of a soft-core potential for the Lennard-Jones potential and *iii*) annihilation of the soft-core potential.[5] Figure 1 illustrates the protocol for the biggest sidechain annihilation, i.e. Trp to Ala. Here,

the three transformations are applied to each of the six charge groups present in the tryptophan sidechain in a series of 9 subperturbations. Each subperturbation is further divided in a pre-defined number of λ-windows, usually 20 (although the user can modify this number), each of which is sampled by short MD simulations (see specific parameters in the next section). QresFEP allows the estimation of the free energy difference associated with each subperturbation with two different methods. One alternative is Zwanzig's exponential formula:[32]

$$\Delta G = G_B - G_A = -\beta^{-1} \sum_{m=1}^{n-1} \ln\langle \exp\left(-\beta^{-1}(U_{m+1} - U_m)\right)\rangle_A \qquad (1)$$

where U_m denotes the effective potential energy function of a particular FEP window and n is the number of intermediate λ-states ($\beta = 1/kT$). U_m is constructed as a linear combination of the initial (A) and final (B) potentials of the subperturbation:

$$U_m = (1 - \lambda_m)U_A + \lambda_m U_B \qquad (2)$$

where the coupling parameter λ_m is stepwise incremented from 0 to 1.

From the potential energy data collected at a number of such λ-points discussed above, the free energy difference between two adjacent windows may also be alternatively obtained by the Bennett acceptance ratio (BAR) method[33] as

$$\Delta G_i = -\beta^{-1} ln \frac{\langle 1+e^{-\beta(\Delta U \Delta\lambda_i - C_i)}\rangle_{i+1}}{\langle 1+e^{+\beta(\Delta U \Delta\lambda_i - C_i)}\rangle_i} + C_i \qquad (3)$$

where the constants C_i are optimized iteratively so that the two ensemble averages become equal, yielding $\Delta G_i = C_i$. The concatenation of the successive subperturbations yields the calculated free energy of a mutation within a given environment (e.g. aqueous solution, vacuum). Figure 1 illustrates the thermodynamic cycle needed to estimate relative free

energies, in this case the hydration free energy of the Trp → Ala mutation, where two legs need to be computed in water and vacuum, respectively, and the difference in the Gibbs free energy on each of them ($\Delta\Delta G$) equals to the experimental (hydration) free energy difference between the two sidechains.

QresFEP contains a tailored FEP recipe for the mutation of each natural aminoacid to alanine (Pro is excluded due to the different topology of its main chain). This scheme can be extended to calculate the relative free energy associated with non-Ala mutations by combination of two thermodynamic cycles, each of them connecting the WT or mutant sidechain, respectively, to the common Ala intermediate, in which case the mutant sidechain is usually modeled from the WT structure.[22] Additionally, we herein developed an alternative to the combination of two thermodynamic cycles for a number of non-Ala amino acid pairs, which we refer to as 'shortcut' protocols. In these selected cases, a simple chemical transformation within the sidechain could clearly avoid the double annihilation of the two sidechains to alanine. These 'shortcut' protocols are implemented for the perturbations Ile → Val, Thr → Ser, (both involving a methyl annihilation), Tyr → Phe (annihilation of the p-hydroxy on the phenyl sidechain), and the transformation of carboxylic acid into amide for the pairs Asn → Asp and Gln → Glu.

The strategy of smooth annihilation by concatenation of subperturbations ensures high precision, fidelity and a minimum hysteresis, which is the measure of convergent results. The hysteresis is calculated as the difference between the forward and the reverse pathways of the transformation, as

$$Hysteresis = \left|\langle \Delta G_{fwd} \rangle - \langle \Delta G_{rev} \rangle\right|$$

where $\langle \Delta G \rangle$ are the average values of Gibbs free energies over several replicate MD trajectories (i.e., same conditions but different initial velocities) in the forward (fwd) and reverse (rev) directions. The total hysteresis is the sum of the hysteresis values for each subperturbation of the transformation taken into account. The use of different replicate simulations (as a default in QresFEP we recommend 10 replicates) also allows the estimation of the associated errors as the standard error of the mean (s.e.m.).

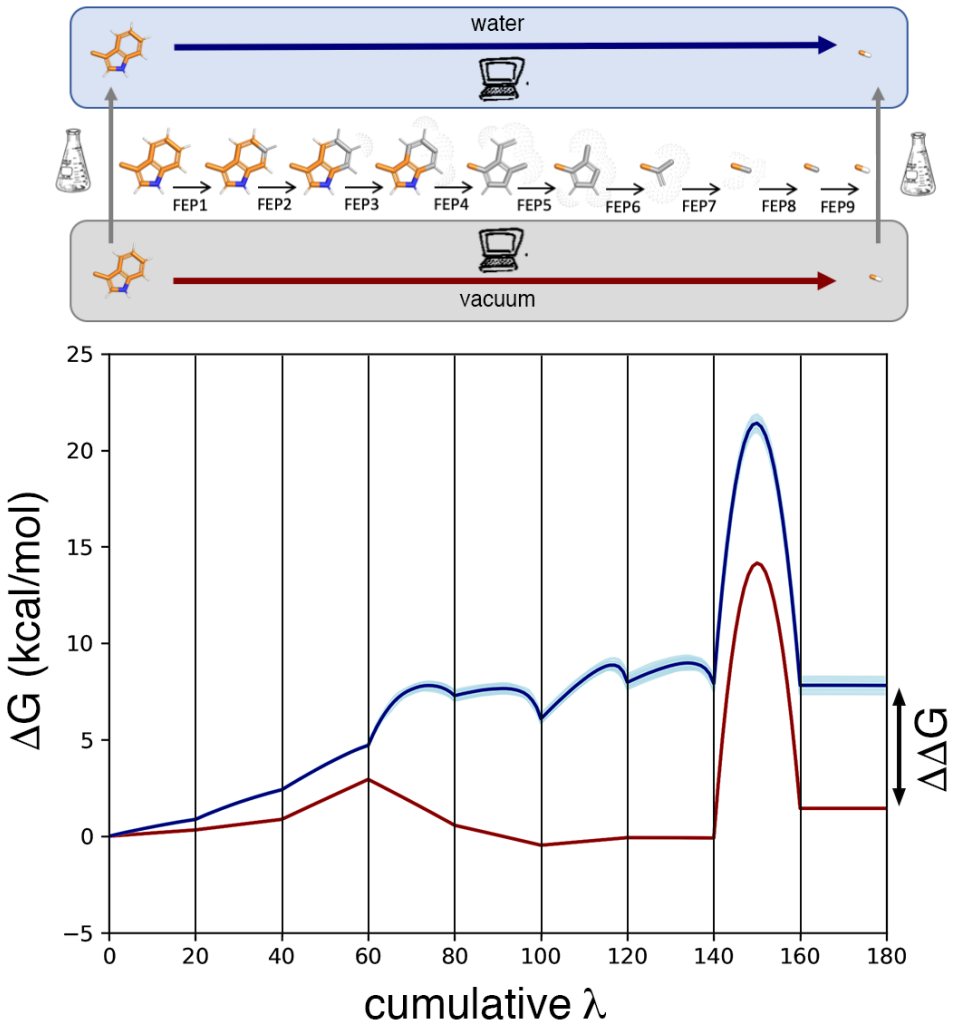


Figure 1: Overview of the FEP scheme for sidechain mutations to alanine applied in this work. The figure shows the annihilation of a given Trp, which in this case is gradually broken down in 9 FEP subperturbations consisting on 20 λ steps each. Annihilation of partial charges and introduction of soft-core potentials are indicated with gray color and dots, respectively,

on the molecular structure. Blue and red lines account for the accumulated free energy (ΔG, kcal/mol) along the transformation in water or vacuum, respectively (dark line is the average of 10 independent simulations, each represented on a light colored line). According to the thermodynamic cycle, the difference between the final values (ΔΔG kcal/mol) is the estimation of the solvation free energy of Trp as compared to Ala.

Molecular Dynamics setup

The MD software package Q[17,18] was used for MD sampling and energy collection for free energy perturbation (FEP) calculations. MD simulations were performed with spherical boundary conditions (SBC) using the surface constrained all-atom solvent (SCAAS)[17,34] where solvent atoms in the boundary are subject to polarization and radial restrains to mimic the properties of bulk water at the sphere surface. The simulation sphere was prepared and solvated with a recently developed API module.[19] Here, the user only needs to define two parameters: (1) the center of the sphere, i.e. the center of geometry of the mutated residue and (2) the size of the sphere, which should encompass all the residues playing a role in the process of interest (i.e. binding of the ligand or stabilization of the mutated residue), while allowing enough solvation patch to ensure sufficient dielectric screening. This module accounts for the protonation of protein residues including the necessary neutralization of ionizable residues in the restrained area and outside the sphere dimensions.[19,35] Atoms lying outside the simulation sphere were tightly constrained (200 kcal/mol/Å^2 force constant) and excluded from the calculation of non-bonded interactions. Long range electrostatic interactions beyond a 10 Å cut off were treated with the local reaction field (LRF) method,[36] except for the atoms undergoing the FEP transformation, where no cutoff is applied. Solvent bond and angles were constrained using the SHAKE algorithm.[37] The OPLS-AA/M force field as implemented in Q was adopted for FEP calculations.[38]

Each system was then subjected to 10 parallel MD replicate simulations, which only differed in their initially assigned random (Maxwell-Boltzman) velocities, and included the following equilibration scheme: *i)* an initial phase of 31 ps where the simulation sphere is heated from 0.1 to 275 K, the bath coupling increased from 0.2 to 1 ps and a positional restraint of 25 kcal/mol/Å^2 initially imposed on all solute heavy atoms slowly released until 5 kcal/mol/Å^2; *ii)* a 100 ps unbiased and unrestrained equilibration, with the temperature set to 298 K and the bath coupling to 10 ps. The same parameters were used during the subsequent production phase, where the sampling along a λ transformation pathway (defining the FEP transformation) was performed with the following parameters in the different datasets (unless indicated the contrary): the transformation was divided into 20 λ windows, evenly distributed (i.e. linear sampling), the corresponding MD sampling for each λ step being 10 ps using a 1 fs time step.

Side chain mimics (solvation free energies)

Experimental solvation free energies were obtained from Wolfenden et al.[39] (neutral residue sidechains) and from Zhang et al.[40] (ionized residue sidechains), in the last case combined with the absolute hydration free energy of the proton retrieved from Tissandier et al.,[41] as implemented in the Minnesota solvation database.[42,43] Sidechains were generated using PyMOL[44] and truncated at the Cβ atom, where a positional restraint of 1 kcal/mol/Å^2 was applied to keep the molecule within the center of the simulation sphere. Each sidechain was mutated to Alanine with the corresponding QresFEP recipe. The annihilation protocol was applied both in a water sphere and in vacuum, to close a thermodynamic cycle that allows calculating relative free energies of solvation for each sidechain as compared to the reference Ala (see Figure 1). In these homogenous environments, we also performed the reverse

protocol (i.e. growing the sidechain from Ala) to yield the hysteresis associated to each transformation.

Neuropeptide Y1 receptor - BIBP3226 antagonist (ligand binding)

The recently published crystal structure of the hY_1 receptor in complex with the antagonist UR-MK299 (PDB entry 5ZBQ)[45] was refined by deleting the co-crystallized T4 lysozyme and the ligand, and adding missing sidechains with the Schrödinger Protein Preparation Wizard.[46] Hydrogen atoms were added and rotamers of Asn/Gln/His sidechains optimized with the Molprobity Web server (http://molprobity.biochem.duke.edu):[47] $Q219^{5.46}$ (Ballesteros-Weinstein numbering in superscript[48]) was flipped; the Asp, Glu, Arg and Lys residues were modeled in their charged form while all His residues were modeled as neutral with the proton on Nδ. The GPCR-ModSim web server (http://open.gpcr-modsim.org) was used for membrane insertion and MD equilibration.[49,50] Briefly, this protocol inserts the receptor structure into a pre-equilibrated bilayer of POPC lipids, with the TM bundle aligned to its vertical axis. The system is then solvated with SPC waters and builds a prism-hexagonal shaped box that is energy minimized with GROMACS 4.6[51] using the OPLS-AA force field [52] combined with the Berger parameters for the lipids.[53] Thereafter a 2.5 ns MD equilibration under periodic boundary conditions (PBC) followed, where initial restraints on protein and ligand atoms were gradually released from 1000 to 200 kJ/mol$Å^2$, followed by a 2.5 ns extension in which the protein was unrestrained except for 22 pairs of distance restraints applied to ensure the TM bundle topology observed in class-A GPCRs, according to the last version of the PyMemDyn protocol implemented in GPCR-ModSim.[49,50] BIBP3226 was manually docked into the equilibrated structure of the hY_1 receptor by superimposing the

common substructure with the co-crystallized analog UR-MK299 (both ligands only differ in the carbamoyl substituent at the guanidine group present in UR-MK299).

A sphere with a radius of 24 Å centered between T97$^{2.61}$ and T212$^{5.39}$ was created in Q from the equilibrated system. Ionizable residues near the sphere boundary were neutralized to avoid artifacts due to missing dielectric screening.[35] The FEP transformation consisted in this case of 50 λ windows for each subperturbation as previously described for this system.[5] In the case of the D287$^{6.59}$A mutation, a chloride ion was simultaneously charged to maintain the overall charge of the system. This ion was placed in a sufficiently solvated region distant from the binding site and any other other charged sidechains, and subject to a positional restrain of 5 kcal/mol/Å. The initial position of the non-Ala mutant F173W$^{4.60}$ was generated with PyMOL.[44]

Protein stability calculations

Atomic coordinates for T4 lysozyme were obtained from the crystallographic structure with PDB code 2LZM.[54] The structure was prepared using Maestro's Protein Preperation Wizard, and the simulation sphere was defined in each case with a radius of 25 Å and centered on the Cβ atom of the residue undergoing the mutation. All ionizable residues within 22 Å of the sphere center were treated in their ionized form, while those on the sphere boundary and beyond were neutralized. The reference calculations to account for the unfolded state were done on the corresponding Ala-X-Ala tripeptide fragment (with a small positional restraint of 1 kcal/mol applied to the Cα atom of the central residue undergoing mutation) with the same sphere size. All other parameters were the same as described in the section 'side chain mimics'.

Results

GUI and API implementation

QresFEP is included in the last release (2.0) of the Q graphical interface Qgui.[31] This software is written in Python using object oriented design principles based on the TkInter GUI library (https://docs.python.org/2/library/tkinter.html). The 2.0 release of Qgui includes two modules accounting for the QresFEP protocol: (i) The "QresFEP setup" module (Figure 2 left panel), allows an intuitive setup for the calculations involved. In a typical case, like the one illustrated in Figure 1, one needs to run two simulations, one for each vertical leg of the thermodynamic cycle. The preparation of each relevant topology file is done with the general tools in Qgui described elsewhere, and in this module the user can load each topology file and select the residue(s) undergoing mutation in each of them. The parameters for the equilibration and MD sampling within the FEP scheme include default values for temperature, lambda step size, number of replicas and general MD settings (see methods), which can of course be modified and, in any case, this module ensures that the same parameters are applied between the independent simulations of the two legs. Qgui2 includes all FEP recipes to mutate each natural amino acid to alanine, with the exception of proline, with separate versions for the different charge states of ionizable residues. In addition, the user can select protocols where a counter ion is charged along the FEP scheme, maintaining equal charge of the simulation sphere over the perturbation. All FEP protocols are editable through the Qgui window, and the user can even add customized protocols. (ii) The "Analyze resFEP" (Figure 2 right panel) is designed to combine the resulting FEP calculations for each residue. This module allows the unified analysis of FEP simulations performed on different positions, which is the most typical case in ligand binding or protein stability. In each case, the difference between the two sets of simulations regarding each horizontal leg in the thermodynamic cycle (see Fig 1) will be calculated, loading them with the appropiate "+" and

"-" sign in the "combine" menu. The result (Figure 2 right panel) summarizes the average free energies, as well as relevant statistical figures of merit such as the s.e.m. The summary for each residue can be exported and collected for ulterior integrated analysis.

The API for QresFEP includes these and other functionalities optimized for the systematic scan of residue mutations within a given system, based on our recently published framework for ligand FEP, QligFEP.[19] It includes three modules, written as command line tools, accounting for: (i) the preparation of the topology (ii) the setup of the FEP calculations and (iii) an analysis module, all of which following the same philosophy as described above.

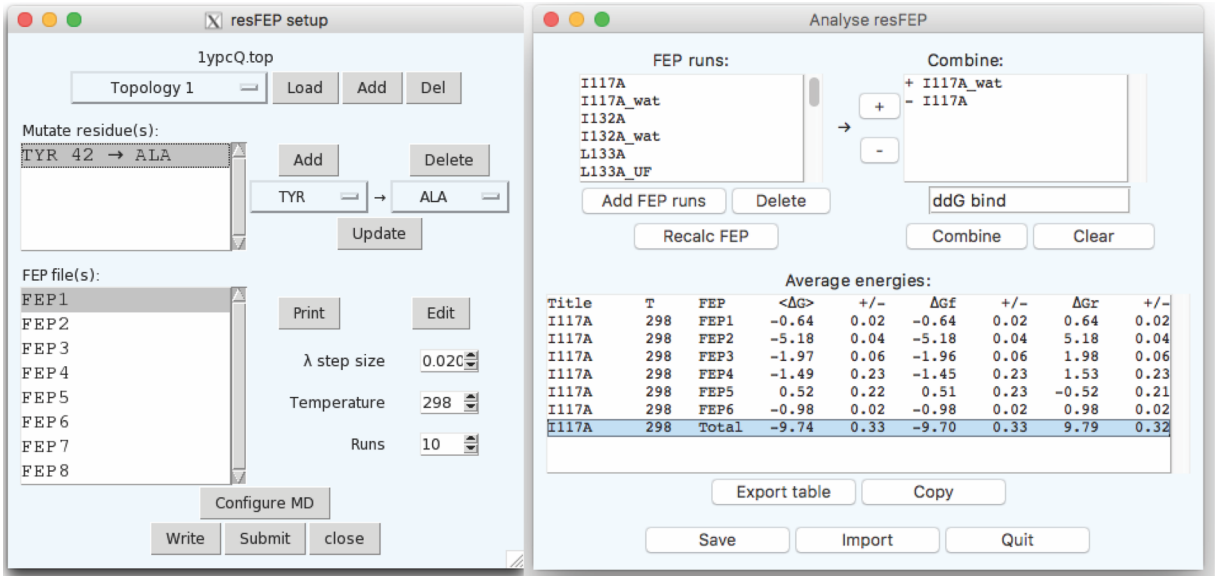


Figure 2: Snapshots of the graphical interface software here reported, Qgui 2.0. The left window shows the resFEP setup, where the user can access several mutations from different topologies. Right window, the Analyze resFEP module, which allows to calculate and combine the result from the different perturbations.

Solvation free energies.

Since the original report of the experimental hydration free energies of aminoacid sidechains by Wolfenden et al. more than three decades ago,[39] this dataset has been regularly used as a benchmark for different force field parameters and free energy calculation methods.[55–60] The implementation of the QresFEP protocol for the estimation of the hydration free energy is illustrated in Figure 1. Before the application to the whole dataset, a preliminary parameter optimization was performed for two extreme cases: one aromatic, large sidechain (Trp) and one small, polar sidechain (Ser). The conclusion of this analysis, which focused on the number of λ windows per subperturbation and the sampling time per subperturbation (Tables S1a-S1f of the SI), was that the optimal balance between accuracy and computation time is obtained when each FEP stage is divided in 20 λ windows, each sampled for 10 ps (i.e. 200 ps per subperturbation). We also explored two alternative methods to calculate free energies, i.e. Zwanzig exponential formula or Bennet Acceptance Ratio (BAR), with no significant difference between the two methods in line with our previous findings for base pair mutations.[61] Finally, we decided to keep the sampling per FEP stage equal, rather than the total sampling time per perturbation constant, since this yields a more robust protocol, i.e. compare the equal convergence achieved for the biggest (Trp) or the smallest (Ser) perturbation to Ala. Thus, as a standard protocol, we use 20 λ windows, 200 ps per FEP stage and Zwanzig's equation, in which case the largest Trp → Ala perturbation consists of 36 ns of sampling simulation time and the smallest Ser → Ala of 16 ns.

Following this line of reasoning, we estimated solvation free energies relative to alanine for all sidechain analogs except Proline and Glicine (we have recipe for the last one, but the experimental value is not available). The results (Table 1) show very good agreement with experimental data, with a total mean averaged error (MAE) of 1.50 kcal/mol. Since the protocol allows the estimation of the reverse pathway (i.e., growing the sidechain from Ala),

besides the accuracy one can estimate the hysteresis, which is as low as 0.45 kcal/mol on average. However, the performance of the annihilating (X → Me) protocols is slightly better (MAE = 1.73 kcal/mol). The results of each case are the average of 10 independent replica simulations, which allows the calculation of the standard error of the mean (s.e.m., indicated as deviation in columns 5 and 6), which is as low as 0.14 kcal/mol on average. Table 2 shows a performance comparison with 14 other combinations of FEP methodology and forcefields published in 6 different studies. Here, the accuracy of the calculated hydration free energies for the subsets of non-charged sidechain mimics allows a more balanced comparison with the remaining methods, which in any case attemp the calculation of the most common charged form of titratable sidechains. The higher MAE of the charged residues is to be expected, since the accuracy of protein force fields in modeling ionized amino acids is largely unknown.[40] Our results on this dataset are comparable with those of other state of the art methods. Of particular interest are the methods implementing the same forcefield in their simulations. Thus, Sandle and Pande obtain MAE of 0.69 and 0.85 kcal/mol, respectively, considering only the neutral sidechains which is comparable to the MAE = 0.85 of our "Neutral" set, while the extended set considered by Tieleman (including neutral form of ionizable residues) registers a MAE of 1.71 kcal/mol, which in our case comes down to 1.19 (subset "all"). The results reported with the Gromos53a6 forcefield show a significantly improved statistical correlation as compared to the rest of the studies.[59] Notably, the partial charges of that forcefield were optimized to better represent the hydration free enthalpies of amino acid side chains, though this comes at the cost of a reduced accuracy in reproducing e.g. the density and heat of vaporization of the pure liquid.[59]

Table 1: Hydration free energies of amino acid side chains.

Amino acid	Side chain mimic	Solvation free energies ($\Delta\Delta G$ kcal/mol) X –> CH_4			
	(neutral)	Exp	X → Me	Me → X	Hysteresis (H)

Asn	Acetamide		11.62	11.19 ± 0.04	-9.92 ± 0.05	1.27
Cys	Methanethiol		3.18	3.16 ± 0.01	-3.42 ± 0.01	0.06
Gln	Propionamide		11.32	9.35 ± 0.06	-10.2 ± 0.06	0.85
Ile	1-butane		-0.21	-0.68 ± 0.47	2.23 ± 0.23	1.55
Leu	Isobutane		-0.34	-0.15 ± 0.23	0.05 ± 0.06	0.73
Met	Methylsulfanylethane		3.42	3.70 ± 0.05	-3.73 ± 0.05	0.03
Phe	Toluene		2.70	3.46 ± 0.05	-3.23 ± 0.08	0.23
Ser	Methanol		7.00	6.93 ± 0.06	-6.99 ± 0.01	0.12
Thr	Ethanol		6.82	4.98 ± 0.09	-4.86 ± 0.10	0.12
Trp	3-methyl-1H-indole		7.82	6.37 ± 0.06	-6.38 ± 0.09	0.01
Tyr	p-cresol		8.05	7.94 ± 0.09	-7.85 ± 0.07	0.09
Val	Propane		-0.05	-2.52 ± 0.02	3.03 ± 0.08	0.51
	Statistical Figures (neutral)			R^2=0.94 MAE=0.85 MCC=1.00	R^2=0.93 MAE=1.07 MCC=1.00	⟨H⟩=0.46
	(ionizable)	Charge form	Exp	X → Me	Me → X	Hysteresis (H)
Arg	n-Propylguanidine	N_2H_3	12.86	11.18 ± 0.29	-10.12 ± 0.09	1.06
		$N_2H_4^+$	69.15	69.80 ± 0.35	-69.81 ± 0.17	0.01
Asp	Acetic acid	COOH	8.64	8.34 ± 0.04	-8.18 ± 0.05	0.16
		COO^-	79.52	82.24 ± 0.14	-83.03 ± 0.11	0.79
Glu	Propionic acid	COOH	8.41	14.56 ± 0.15	-14.92 ± 0.06	0.36
		COO^-	78.04	81.18 ± 0.15	-81.88 ± 0.16	0.70
Lys	Butan-1-amine	NH_2	6.32	3.95 ± 0.25	-4.40 ± 0.11	0.45
		NH_3^+	73.07	78.63 ± 0.29	-79.52 ± 0.19	0.89
His	4-Methylimidazole	Nδ	12.21	11.77 ± 0.06	-11.87 ± 0.06	0.10
		Nε	12.21	12.03 ± 0.09	-12.21 ± 0.06	0.23
		Nδ,ε	64.23	63.78 ± 0.30	-63.75 ± 0.10	0.03
	Statistical Figures (ionizable)			R^2 = 0.99 MAE = 2.21 MCC = 1.00	R^2 = 0.99 MAE = 2.45 MCC = 1.00	⟨H⟩=0.43
	Statistical Figures (all sidechains)			R^2 = 1.00 MAE = 1.50 MCC = 1.00	R^2 = 1.00 MAE = 1.73 MCC = 1.00	⟨H⟩=0.45

[a] Experimental hydration free energies adapted from Wolfenden et al.[58]

Table 2: Performance of this and other methods in the literature on the calculation of solvation free energies of aminoacid sidechain mimics, expressed as the mean absolute error (MAE) compared to experimental values[39]

Group	Forcefield	Subset	(n)	MAE
	OPLS-AA/M	Neutral	12 [a]	0.85

This work	OPLS-AA/M	all	18 [a]	1.19
	OPLS-AA/M	all (ionized)	23 [a]	1.50
Sandler lab[55]	OPLS-AA	Neutral	14 [b,c]	0.69
Pande lab[56]	AMBER(ff94)	Neutral	15 [b]	1.35
	CHARMM22	Neutral	15 [b]	1.31
	OPLS-AA	Neutral	15 [b]	0.85
Tieleman lab[57]	OPLS-AA	all	18 [c,d]	1.71
	Gromos96	all	18 [c,d]	2.41
Mobley lab [58]	Amber(GAFF)	all	16 [d,e]	0.99
van Gunsteren lab.[59]	Gromos43a2	Neutral	15 [b]	2.08
	Gromos45a3	Neutral	15 [b]	2.01
	Gromos53a3	Neutral	15 [b]	1.89
	Gromos53a6	Neutral	15 [b]	0.19
Pardo lab.[60]	OPLS-AA	all	17 [d,f]	0.93
	AMBER99	all	17 [d,f]	0.67
	AMBER03	all	17 [d,f]	1.48

[a] Following the division of sidechain analogs as in Table 1. [b] Ionizable residues excluded in any form. [c] Neutral His only considered as protonated in delta. [d] Ionizable residues only considered in their neutral form. [e] Arg and His mimics excluded. [f] Arg is excluded.

It is worth noting that the hydration free energies here reported with the single-topology QresFEP protocol are very similar to the values obtained for the same ssidechain mimic data set with our dual topology approach for ligands, QligFEP.[19] Finally, to illustrate the utility of the 'shortcut' protocols developed for the 5 pairs of chemically similar sidechains, Supplementary Table 2 compares the relative solvation free energies obtained for each pair with the standard ('two-cycles', see also Supplementary Figure 2) and the 'shortcut' protocols, obtaining a very similar accuracy.

Protein-Ligand binding

One application of the QresFEP protocol is the estimation of ligand binding affinity shifts due to single point mutations. This can be a helpful tool to interpret and design site-directed mutagenesis experiments, as we have shown for recent GPCR-ligand binding studies.[5,22–27] The first of these studies was aimed to characterize antagonist binding to the

hY_1 receptor, examining the effect of 13 single mutants on the binding affinity of the peptide-mimetic antagonist BIBP3226.[5] Parallel QresFEP calculations were done on two alternative docking poses of the ligand, obtained on a homology-based model of the hY_1 receptor, and the results were combined with 7 ligand FEPs to determine the pose that best explained site-directed mutagenesis and SAR experimental data.[5] The binding orientation proposed was recently confirmed by a crystal structure of the hY_1 receptor bound to the antagonist UR-MK299, a derivative of BIBP3226 bearing a carbamoyl substituent on the guanidine moiety.[62] Figure 3B depicts a comparison between the modeled and experimental structures, yielding a root mean square deviation for the Cα trace of the transmembrane (TM) region of 1.74 Å, and an RMSD = 4.59 Å for the common substructure of the two ligands. Despite the overall similar orientation, one can appreciate differences in the detailed intermolecular interactions involved, thus the new crystal structure constitutes an excellent starting point to revisit and expand the original study with QresFEP. Figure 3C and Supplementary Table 3 summarize the results obtained for BIBP3226, finding an agreement with the experimental data comparable with the calculations based on the homology model (see Table 3).[5] This result confirms the applicability of the QresFEP protocol in cases where no crystal structure is available, but where a computational model of the initial complex of sufficient quality can be generated, as previously shown in the case of ligand binding to the GPR139 receptor. [24] In some cases, specific receptor-ligand interactions were more accurately modeled with the crystal structure coordinates: 1) the π-stacking of the biphenylmethyl moiety of BIBP3226 with $Y100^{2.64}$ 2) the charge-π interaction between the guanidine group and $F173^{4.60}$ and 3) the hydrogen bond between $Q219^{5.46}$ and the tyrosyl substituent of BIBP3226, resulting in an improved accuracy of the calculations as compared to the homology model. In addition, the panel of mutations include three loss-of-binding mutants located in TM5: $Y211^{5.38}A$, $N283^{6.55}A$ and $D287^{6.59}A$.[62–64] While this effect is correctly captured by the FEP calculations

for two of them (i.e., binding free energy shift greater than the experimental threshold, see Figure 3C), the effect of Y211$^{5.38}$ is under predicted, possibly since this residue is no longer modeled to be in direct contact with the ligand. As opposed to the previous study and the experimental data, the calculations assign a favorable role for the W276$^{6.48}$A mutation. Located deep in the binding cavity, W276$^{6.48}$ interacts with the carbamoyl substituent that is specific of the co-crystallized ligand UR-MK299, in a conformation and/or solvation pattern that might change when this group is replaced by the unsubstituted guanidinium of BIBP3226 (Figure 3B).

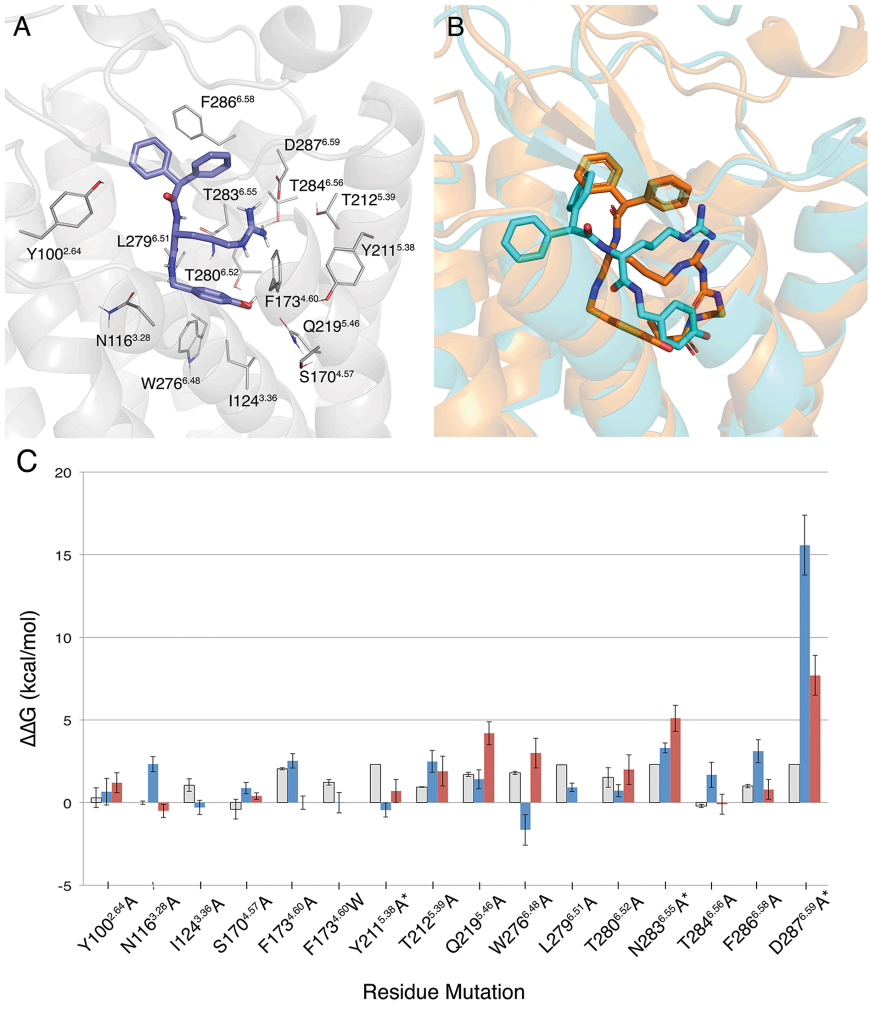


Figure 3: Predictions of site directed mutagenesis effects on ligand binding affinity. (A) Starting configuration used in this work for the simulations of BIBP3226 (violet) as bound to

the hY1 receptor, with the sidechains of the residues undergoing mutation labeled. (B) The same complex as modeled by Boukharta et al.[5] (cyan) as compared to the crystal structure of hY_1 in complex with antagonist UK-M299 (orange). (C) Calculated and experimental relative binding free energies for BIBP3226 to the sixteen hY1 alanine mutants compared to hY1 wt. Blue bars represent $\Delta\Delta G_{bind}^{FEP}$, gray bars $\Delta\Delta G_{bind}^{exp}$ and red bars $\Delta\Delta G_{bind}^{FEP}$ for the 13 positions also calculated by Boukharta *et al.*[5] For mutants marked with an *, $\Delta\Delta G_{bind}^{exp}$ is approximated by the threshold of the experiment (2.3 kcal/mol). The values are reported in SI Supplementary Table 3.

Altogether, the analysis of this panel of 16 mutations reveals that the calculations are converged satisfactorily, with an average s.e.m. of 0.60 kcal/mol which is indeed not so different than the variations within the experimental values, where the discrepancies can be as high as 1.6 kcal/mol in those cases where there exists data from different experiments (see Supplementary Table 4). The correlation with experiments is also quite satisfactory, with a MAE=1.28 kcal/mol. This value is within the same order of magnitude of recent FEP studies of ligand binding performed both us[5,22–27] and others.[30] Table 3 collects the results of our previous studies with the protocols now included in QligFEP. These studies were done on a variety of GPCR systems, such as the A_1 and A_{2A} adenosine receptors in complex with co-crystalized ligands, [22,23,25] and also homology models with docked ligands, as is the case of the neuropeptide Y_1 and Y_2 receptors[5,27] and the orphan GPR139.[24] This gives a global idea of the performance expected with the FEP protocol developed by us, with a MAE in the range of 0.42 – 1.31 kcal/mol and other statistical figures of merit showing the predictive power of these models (see Table 3).

Table 3: Overview of ligand binding calculations with QresFEP.

References	System	Number of mutations[a]	MAE[b]	R^2	MCC[c]
Keranen et al.[22]	$A_{2A}AR$	17 (15)	1.31	0.53	0.58
Keranen et al.[23]	$A_{2A}AR$	26 (13)	1.07	0.28	0.59
Boukharta et al.[5]	Y_1R[d]	13 (10)	0.96	0.34	0.57
Jespers et al. [25]	A_1AR	6 (4)	0.42	0.59	1.00
Nøhr et al.[24]	GPR139[d]	6 (0)	N/A[e]	N/A	1.00
Xu et al. [27]	Y_2R[d]	9 (9)	0.78	0.41	1.00

[a] In parenthesis, the number of mutations with qualitative experimental data, allowing the calculation of MAE and R^2 (in the remaining cases the observed effect was larger than the experimental cutoff, or the experimental data was based on functional data). [b]MAE = Mean absolute error. [c] MCC = Matthew's correlation coefficient. [d] A homology model of the receptor and a docked position of the ligand were used for the simulations. [e] N/A = Not available

Protein stability

The identification of protein folded states from computer simulations is non-trivial, and brute-force MD simulation techniques are associated with convergence problems and limited sampling.[65] In the field of protein engineering and enzyme design, however, the aim is to efficiently predict the effect of potential mutational sites, which can be approached from a free energy calculation perspective. An advantage of this approach is that the definition of a representative unfolded (reference) state for every mutation can be done by simulating simple peptides.[66–71] This has the advantage that the tripeptide model system, representing the unfolded state, only has to be calculated once. We extended these calculations to include a total of 30 replicate simulations, but no significant (< 0.3 kcal/mol) difference was observed from using 10 replicates (see Supplementary figure 5) in all cases. Next, we tested the feasibility of applying the QresFEP protocol in calculating the change in stability of T4-lysozyme (see figure 3A). A total of 43 mutations were selected from the ProTherm database,[72–74] for which thermostabilizing properties have previously been calculated with the proprietary Schrödinger solution FEP+.[21] In contrast to the previous section, since we are now

interested in the effect of point mutations on the overall stability of a system, we centered the sphere on the residue under investigation, rather than the binding site of a ligand. The results obtained with QresFEP are in good agreement with the experimental data (see figure 4B and Supplementary table 5), as represented by an R^2 of 0.70 and Spearman Rho of 0.86. The absolute deviation between calculated and experimental values is slightly higher (MAE=1.66) than observed in the previous section, which can be largely attributed to an overestimation of mutations that have a more drastic effect on stability (i.e. $\Delta\Delta G > 3$ kcal/mol). Still, the method is well capable of classifying whether a given mutation will be (de)stabilizing the receptor (Matthew's Correlation Coefficient (MCC) = 0.53). In addition, we note that previous work has pointed to issues with accuracy of the experimental data in the ProTherm database, where an internal validation showed that errors can accumulate to 0.81 kcal/mol with an R^2 of 0.71.[75]

Finally, the fact that we used the same forcefield allows a direct comparison with the results obtained for the same dataset with FEP+.[21] The main difference between the two methods can thus be attributed to the efficiency of the sampling method applied, which provides insights in the feasibility of using a reduced spherical system with the LRF to calculate protein stability as compared to the more computationally demanding PBC in combination with the particle mesh Ewald summation for longe-range electrostatics used in FEP+.[21] We observe a good agreement between both methods, with a correlation of R^2=0.75 between one another. In addition, the agreement with experimental data is comparable for both methods: MAE=1.12 kcal/mol versus 1.66 kcal/mol, Spearman Rho = 0.81 versus 0.86 and MCC=0.48 versus 0.53 between Steinbrecher et al. and our method respectively. Finally, our calculations fall within the same level of accuracy as reported in the work of Gapsys et al, who used GROMACS and a variety of other force fields.[28]

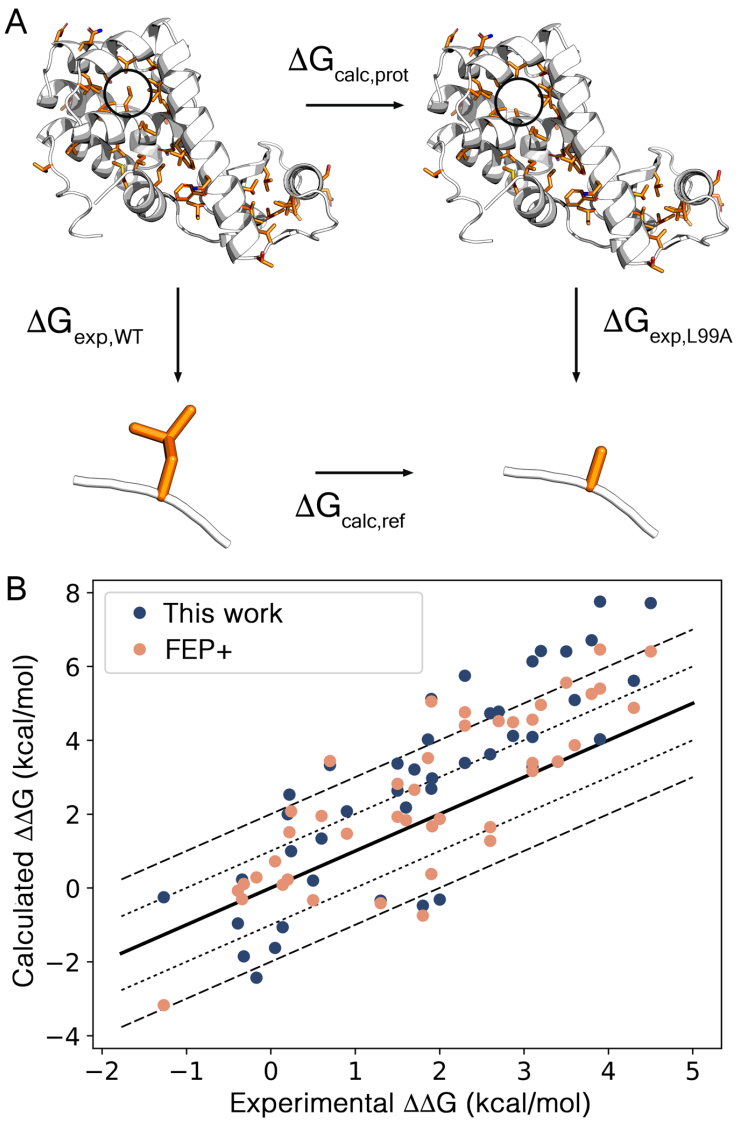


Figure 4: calculations of the relative stability of T4-lysozyme for a total of 43 mutations (iin orange). (A) The top horizontal leg represents the protein in the folded state and the unfolded state is represented by a simple Ala-XXX-Ala tripeptide, and thus in the case of the mutation studied here (L99A) the perturbation consists of Ala-Leu-Ala -> Ala-Ala-Ala. (B) Experimental versus calculated effects on the stability of 43 T4-lysozyme mutations studied in this work (blue dots) and by FEP+ (orange dots). [70] The solid line represents a perfect correlation, and the dotted and dashed lines indicates the thrsholds at ±1 and ±2 kcal/mol respectively

Conclusion

We present QresFEP, an *in silico* mutagenesis tool based on rigorous FEP calculations within the software package Q. We extended our previously established method of gradual annihilation of side-chain atoms to provide all required recipes to perform any given protein mutation, and show that these protocols can be used to address a variety of interesting biological problems, including: solvation free energies, ligand binding and protein stability. In all cases the results obtained show a high level of accuracy compared to both experiment and other available methods. The QGUI, QresFEP and Q code are all freely available on GitHub (https://github.com/qusers).

Acknowledgements

This work was supported by the Research Council of Norway through a Centre of Excellence Grant (Grant No. 179568/V30), the Swedish Research Council (VR) and the Knut and Alice Wallemberg Foundation. Additional support from the Swedish strategic research program eSSENCE is acknowledged. Calculations were performed in the Swedish National Infrastructure for computing (SNIC).

References

(1) Schreiber, G.; Fleishman, S. J. *Current Opinion in Structural Biology*. Elsevier Current Trends December 1, 2013, pp 903–910.

(2) Capriotti, E.; Fariselli, P.; Rossi, I.; Casadio, R. *BMC Bioinformatics* **2008**, *9* (Suppl 2), S6.

(3) Capriotti, E.; Fariselli, P.; Casadio, R. *Nucleic Acids Res.* **2005**, *33* (Web Server), W306–W310.

(4) Lenselink, E. B.; ten Dijke, N.; Bongers, B.; Papadatos, G.; van Vlijmen, H. W. T.; Kowalczyk, W.; IJzerman, A. P.; van Westen, G. J. P. *J. Cheminform.* **2017**, *9* (1), 45.

(5) Boukharta, L.; Gutiérrez-de-Terán, H.; Åqvist, J. *PLoS Comput. Biol.* **2014**, *10* (4), e1003585.

(6) Aqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. *Acc. Chem. Res.* **2002**, *35* (6), 358–365.

(7) Cournia, Z.; Allen, B.; Sherman, W. *Journal of Chemical Information and Modeling*. American Chemical Society December 26, 2017, pp 2911–2937.

(8) Blundell, T. L.; Jhoti, H.; Abell, C. *Nat. Rev. Drug Discov.* **2002**, *1* (1), 45–54.

(9) Schmidt, T.; Bergner, A.; Schwede, T. *Drug Discovery Today*. Elsevier Current Trends July 1, 2014, pp 890–897.

(10) Jazayeri, A.; Andrews, S. P.; Marshall, F. H. *Chem. Rev.* **2017**, *117* (1), 21–37.

(11) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. *Nature* **1987**, *328* (6130), 551–554.

(12) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.;

Berne, B. J.; Friesner, R. A.; Abel, R. *J. Am. Chem. Soc.* **2015**, *137* (7), 2695–2703.

(13) Loeffler, H. H.; Michel, J.; Woods, C. *J. Chem. Inf. Model.* **2015**, *55* (12), 2485–2490.

(14) Liu, S.; Wu, Y.; Lin, T.; Abel, R.; Redmann, J. P.; Summa, C. M.; Jaber, V. R.; Lim, N. M.; Mobley, D. L. *J. Comput. Aided. Mol. Des.* **2013**, *27* (9), 755–770.

(15) Homeyer, N.; Gohlke, H. *J. Comput. Chem.* **2013**, *34* (11), 965–973.

(16) Christ, C. D.; Fox, T. *J. Chem. Inf. Model.* **2014**, *54* (1), 108–120.

(17) Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J.; Aqvist, J. *J. Mol. Graph. Model.* **1998**, *16* (4–6), 213–225.

(18) Bauer, P.; Barrozo, A.; Purg, M.; Amrein, B. A.; Esguerra, M.; Wilson, P. B.; Major, D. T.; Aqvist, J.; Kamerlin, S. C. L. *SoftwareX*. Elsevier January 4, 2018.

(19) Jespers, W.; Esguerra, M.; Åqvist, J.; Gutiérrez-de-Terán, H. *J. Cheminform.* **2019**, *11* (1), 26.

(20) Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. *J. Comput. Chem.* **2015**, *36* (5), 348–354.

(21) Steinbrecher, T.; Zhu, C.; Wang, L.; Abel, R.; Negron, C.; Pearlman, D.; Feyfant, E.; Duan, J.; Sherman, W. *J. Mol. Biol.* **2017**, *429* (7), 948–963.

(22) Keränen, H.; Åqvist, J.; Gutiérrez-de-Terán, H. *Chem. Commun.* **2015**, *51* (17), 3522–3525.

(23) Keränen, H.; Gutiérrez-de-Terán, H.; Åqvist, J. *PLoS One* **2014**, *9* (10), e108492.

(24) Nøhr, A. C.; Jespers, W.; Shehata, M. A.; Floryan, L.; Isberg, V.; Andersen, K. B.; Åqvist, J.; Gutiérrez-de-Terán, H.; Bräuner-Osborne, H.; Gloriam, D. E. *Sci. Rep.* **2017**, *7* (1), 1128.

(25) Jespers, W.; Oliveira, A.; Prieto-Díaz, R.; Majellaro, M.; Åqvist, J.; Sotelo, E.; Gutiérrez-de-Terán, H. *Molecules* **2017**, *22* (11), 1945.

(26) Vasile, S.; Esguerra, M.; Jespers, W.; Oliveira, A.; Sallander, J.; Åqvist, J.; Gutiérrez-

de-Terán, H. In *Methods in Molecular Biology*; Humana Press, New York, NY, 2018; Vol. 1705, pp 23–44.

(27) Xu, B.; Vasile, S.; Østergaard, S.; Paulsson, J. F.; Pruner, J.; Åqvist, J.; Wulff, B. S.; Gutiérrez-de-Terán, H.; Larhammar, D. *Mol. Pharmacol.* **2018**, *93* (4), 323–334.

(28) Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. *Angew. Chemie - Int. Ed.* **2016**, *55* (26), 7364–7368.

(29) Clark, A. J.; Gindin, T.; Zhang, B.; Wang, L.; Abel, R.; Murret, C. S.; Xu, F.; Bao, A.; Lu, N. J.; Zhou, T.; Kwong, P. D.; Shapiro, L.; Honig, B.; Friesner, R. A. *J. Mol. Biol.* **2017**, *429* (7), 930–947.

(30) Aldeghi, M.; Gapsys, V.; de Groot, B. L. *ACS Cent. Sci.* **2018**, *4* (12), acscentsci.8b00717.

(31) Isaksen, G. V.; Andberg, T. A. H.; Åqvist, J.; Brandsdal, B. O. *J. Mol. Graph. Model.* **2015**, *60*, 15–23.

(32) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22* (8), 1420.

(33) Bennett, C. H. *J. Comput. Phys.* **1976**, *22* (2), 245–268.

(34) King, G.; Warshel, A. *J. Chem. Phys.* **1989**, *91* (6), 3647.

(35) Gutiérrez-De-Terán, H.; Åqvist, J. *Methods Mol. Biol.* **2012**, *819*, 305–323.

(36) Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1992**, *97* (5), 3100.

(37) Ryckaert, J.-P. J.; Ciccotti, G.; Berendsen, H. J. . H. *J. Comput. Phys.* **1977**, *23* (3), 327–341.

(38) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2015**, *11* (7), 3499–3509.

(39) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20* (4), 849–855.

(40) Zhang, H.; Jiang, Y.; Yan, H.; Yin, C.; Tan, T.; Van Der Spoel, D. *J. Phys. Chem. Lett.*

2017, *8* (12), 2705–2712.

(41) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. *J. Phys. Chem. A* **1998**, *102* (40), 7787–7794.

(42) Marenich, A.; Kelly, C.; Thompson, J.; Hawkins, G. *Minnesota Solvation Database version 20*.

(43) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110* (32), 16066–16081.

(44) Pymol, v1.3

(45) Yang, Z.; Littmann, T.; Wifling, D.; Huster, D.; Plank, N.; Kögler, L. M.; Yi, C.; Stevens, R. C.; Burkert, K.; Wu, B.; Larhammar, D.; Beck-Sickinger, A. G.; Keller, M.; Bosse, M.; Han, S.; Zhang, R.; Li, B.; Buschauer, A.; Bernhardt, G.; Ye, S.; Schmidt, P.; Zhao, Q.; Kaiser, A.; Meiler, J.; Xu, B.; Bender, B. J. *Nature* **2018**, *556* (7702), 520–524.

(46) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. *J. Comput. Aided. Mol. Des.* **2013**, *27* (3), 221–234.

(47) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21.

(48) Ballesteros, J. A.; Weinstein, H. *Methods Neurosci.* **1995**, *25*, 366–428.

(49) Gutiérrez-de-Terán, H.; Bello, X.; Rodríguez, D. *Biochem. Soc. Trans.* **2013**, *41* (1), 205–212.

(50) Esguerra, M.; Siretskiy, A.; Bello, X.; Sallander, J.; Gutiérrez-de-Terán, H. *Nucleic Acids Res.* **2016**, *44* (W1), W455–W462.

(51) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447.

(52) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105* (28), 6474–6487.

(53) Berger, O.; Edholm, O.; Jähnig, F. *Biophys. J.* **1997**, *72* (5), 2002–2013.

(54) Weaver, L. H.; Matthews, B. W. *J. Mol. Biol.* **1987**, *193* (1), 189–199.

(55) Chang, J.; Lenhoff, A. M.; Sandler, S. I. *J. Phys. Chem. B* **2007**, *111* (8), 2098–2106.

(56) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119* (11), 5740–5761.

(57) Maccallum, J. L.; Peter Tieleman, D. *J. Comput. Chem.* **2003**, *24* (15), 1930–1935.

(58) Duarte Ramos Matos, G.; Kyu, D. Y.; Loeffler, H. H.; Chodera, J. D.; Shirts, M. R.; Mobley, D. L. *Journal of Chemical and Engineering Data*. American Chemical Society May 11, 2017, pp 1559–1569.

(59) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25* (13), 1656–1676.

(60) Cordomí, A.; Caltabiano, G.; Pardo, L. *J. Chem. Theory Comput.* **2012**, *8* (3), 948–958.

(61) Lind, C.; Esguerra, M.; Jespers, W.; Satpati, P.; Gutierrez-de-Terán, H.; Åqvist, J. *Methods* **2019**.

(62) Yang, Z.; Han, S.; Keller, M.; Kaiser, A.; Bender, B. J.; Bosse, M.; Burkert, K.; Kögler, L. M.; Wifling, D.; Bernhardt, G.; Plank, N.; Littmann, T.; Schmidt, P.; Yi, C.; Li, B.; Ye, S.; Zhang, R.; Xu, B.; Larhammar, D.; Stevens, R. C.; Huster, D.; Meiler, J.; Zhao, Q.; Beck-Sickinger, A. G.; Buschauer, A.; Wu, B. *Nature* **2018**, *556* (7702), 520–524.

(63) Sautel, M.; Rudolf, K.; Wittneben, H.; Herzog, H.; Martinez, R.; Munoz, M.; Eberlein, W.; Engel, W.; Walker, P.; Beck-Sickinger, A. G. *Mol. Pharmacol.* **1996**, *50* (2).

(64) Sjödin, P.; Holmberg, S. K. S.; Akerberg, H.; Berglund, M. M.; Mohell, N.; Larhammar, D. *Biochem. J.* **2006**, *393* (Pt 1), 161–169.

(65) Jemth, P.; Gianni, S.; Day, R.; Li, B.; Johnson, C. M.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci.* **2004**, *101* (17), 6450–6455.

(66) Gapsys, V.; de Groot, B. L. *J. Chem. Inf. Model.* **2017**, *57* (2), 109–114.

(67) Peng, Y.; Suryadi, J.; Yang, Y.; Kucukkal, T.; Cao, W.; Alexov, E. *Int. J. Mol. Sci.* **2015**, *16* (11), 27270–27287.

(68) Seeliger, D.; de Groot, B. L. *Biophys. J.* **2010**, *98* (10), 2309–2316.

(69) Pan, Y. P.; Daggett, V. *Biochemistry* **2001**, *40* (9), 2723–2731.

(70) Pitera, J. W.; Kollman, P. A. *Proteins Struct. Funct. Genet.* **2000**, *41* (3), 385–397.

(71) Yamaotsu, N.; Moriguchi, I.; Hirono, S. *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.* **1993**, *1203* (2), 243–250.

(72) Bava, K. A. *Nucleic Acids Res.* **2004**, *32* (90001), 120D–121.

(73) Gromiha, M. M.; An, J.; Kono, H.; Oobatake, M.; Uedaira, H.; Sarai, A. *Nucleic Acids Research*. Oxford University Press January 1, 1999, pp 286–288.

(74) Guerois, R.; Nielsen, J. E.; Serrano, L. *J. Mol. Biol.* **2002**, *320* (2), 369–387.

(75) Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. *Angew. Chemie* **2016**, *128* (26), 7490–7494.

Supporting Information

QresFEP: an automated protocol for free energy calculations of protein mutations in Q

Willem Jespers[a,1], Geir Villy Isaksen[a,b,1], Tor Arne Heim Andberg[b], Silvana Vasile[a], Amber van Veen[a], Johan Åqvist,[a] Bjørn Olav Brandsdal[b,] & Hugo Gutiérrez-de-Terán[a,*]*

[a]Department of Cell and Molecular Biology, Biomedical Center, Uppsala University, Box 590, S-75124 Uppsala, Sweden

[b]Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Tromsø – The Arctic University of Norway, N9037 Tromsø, Norway

[1] Authors contributed equally

Supplementary Table 1: Calculated solvation free energies for the Trp (3-methyl-1H-indole) and Ser (ethanol) to Ala (Methane) side chain mimics. Errors are standard error of the mean (s.e.m.) over 10 replicate simulations. Zwanzig and BAR have been used to obtain the total computed changes in free energy, over a total of 9 (Trp) and 4 (Ser) individual FEP stages, each divided in 10, 20, 50 or 100 lambda windows. The simulation times per FEP stage are 100, 200 and 400 ps.

Table 1a: Trp → Ala 100 ps

λ windows	method	$\Delta\Delta G_{hyd}$	ΔG_{wat}	ΔG_{vac}
10	Zwanzig	6.53 ± 0.08	8.00 ± 0.12	1.47 ± 0.00
	BAR	6.44 ± 0.08	7.85 ± 0.10	1.44 ± 0.00
20	Zwanzig	6.42 ± 0.07	7.86 ± 0.10	1.44 ± 0.00
	BAR	6.41 ± 0.07	7.85 ± 0.10	1.44 ± 0.00
50	Zwanzig	6.45 ± 0.08	7.89 ± 0.12	1.44 ± 0.00
	BAR	6.44 ± 0.08	7.88 ± 0.12	1.44 ± 0.00
100	Zwanzig	6.35 ± 0.10	7.79 ± 0.14	1.44 ± 0.00
	BAR	6.35 ± 0.10	7.79 ± 0.14	1.44 ± 0.00

Table 1b: Ser → Ala 100 ps

λ windows	method	$\Delta\Delta G_{hyd}$	ΔG_{wat}	ΔG_{vac}
10	Zwanzig	6.84 ± 0.08	2.07 ± 0.06	-4.77 ± 0.07
	BAR	6.85 ± 0.08	2.07 ± 0.06	-4.78 ± 0.07
20	Zwanzig	6.74 ± 0.06	1.94 ± 0.05	-4.80 ± 0.05
	BAR	6.74 ± 0.07	1.94 ± 0.04	-4.80 ± 0.05
50	Zwanzig	6.72 ± 0.08	1.98 ± 0.04	-4.74 ± 0.05
	BAR	6.72 ± 0.08	1.98 ± 0.04	-4.74 ± 0.05
100	Zwanzig	6.85 ± 0.07	2.02 ± 0.04	-4.83 ± 0.05
	BAR	6.85 ± 0.07	2.02 ± 0.04	-4.83 ± 0.05

Table 1c: Trp → Ala 200 ps

λ windows	method	$\Delta\Delta G_{hyd}$	ΔG_{wat}	ΔG_{vac}
10	Zwanzig	6.38 ± 0.08	7.85 ± 0.12	1.47 ± 0.00
	BAR	6.34 ± 0.08	7.80 ± 0.11	1.46 ± 0.00
20	Zwanzig	6.37 ± 0.06	7.81 ± 0.08	1.44 ± 0.00
	BAR	6.35 ± 0.06	7.79 ± 0.08	1.44 ± 0.00
50	Zwanzig	6.33 ± 0.10	7.77 ± 0.14	1.44 ± 0.00
	BAR	6.34 ± 0.10	7.78 ± 0.14	1.44 ± 0.00
100	Zwanzig	6.41 ± 0.06	7.85 ± 0.09	1.44 ± 0.00
	BAR	6.41 ± 0.06	7.85 ± 0.09	1.44 ± 0.00

Table 1d: Ser → Ala 200 ps

λ windows	method	$\Delta\Delta G_{hyd}$	ΔG_{wat}	ΔG_{vac}
10	Zwanzig	6.93 ± 0.08	2.11 ± 0.05	-4.82 ± 0.06
	BAR	6.92 ± 0.08	2.10 ± 0.06	-4.82 ± 0.06
20	Zwanzig	6.93 ± 0.06	2.04 ± 0.04	-4.89 ± 0.04
	BAR	6.93 ± 0.06	2.04 ± 0.04	-4.89 ± 0.04
50	Zwanzig	6.89 ± 0.07	2.13 ± 0.05	-4.76 ± 0.05
	BAR	6.89 ± 0.07	2.13 ± 0.05	-4.76 ± 0.05
100	Zwanzig	6.81 ± 0.09	2.09 ± 0.06	-4.72 ± 0.07
	BAR	6.81 ± 0.09	2.09 ± 0.06	-4.72 ± 0.07

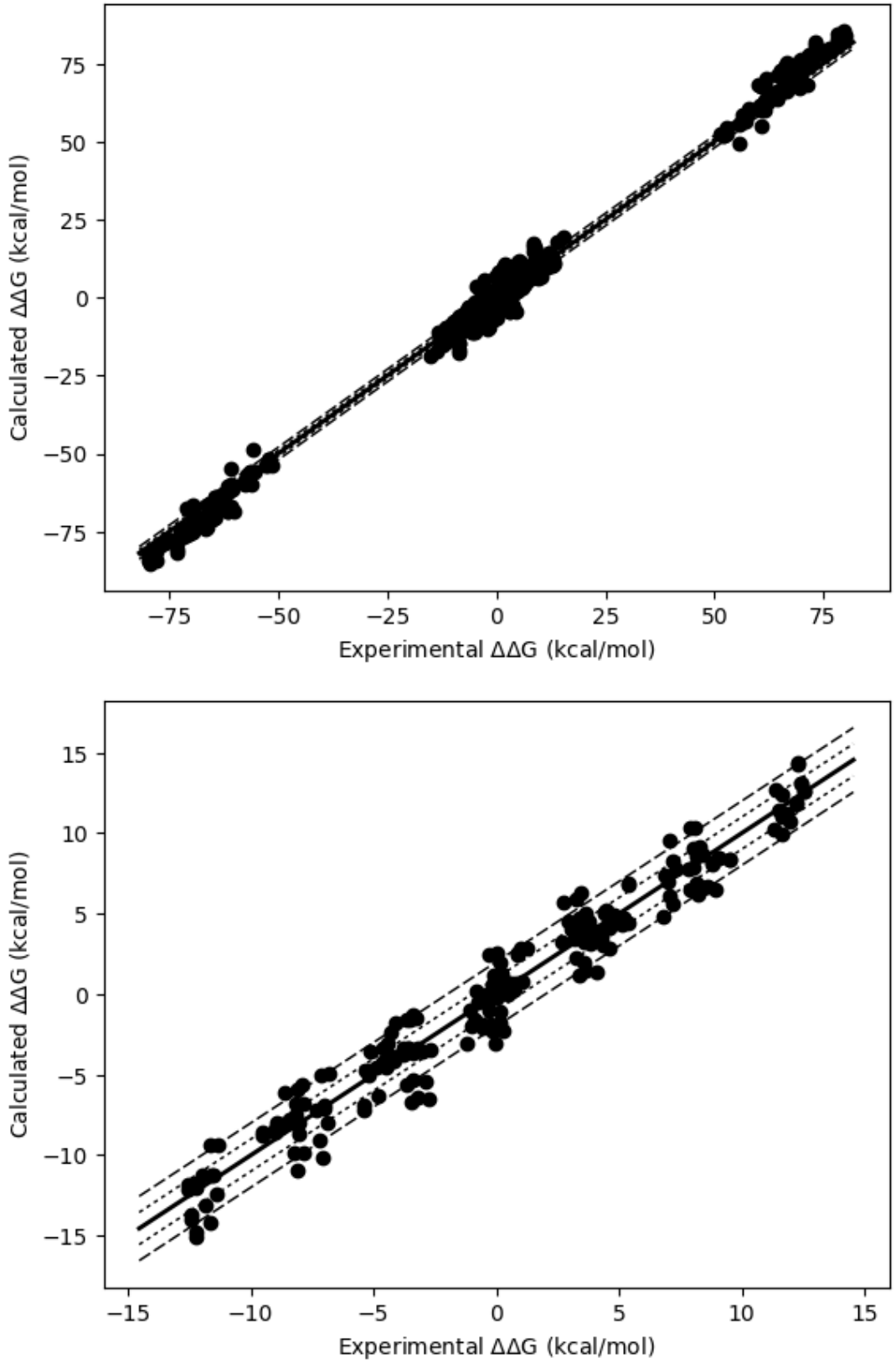
Table 1e: Trp → Ala 400 ps

λ windows	method	$\Delta\Delta G_{hyd}$	ΔG_{wat}	ΔG_{vac}
10	Zwanzig	6.38 ± 0.04	7.85 ± 0.06	1.47 ± 0.00
	BAR	6.35 ± 0.04	7.81 ± 0.06	1.46 ± 0.00
20	Zwanzig	6.50 ± 0.06	7.95 ± 0.08	1.45 ± 0.00
	BAR	6.50 ± 0.06	7.94 ± 0.09	1.44 ± 0.00
50	Zwanzig	6.37 ± 0.04	7.81 ± 0.05	1.44 ± 0.00
	BAR	6.37 ± 0.04	7.81 ± 0.06	1.44 ± 0.00
100	Zwanzig	6.27 ± 0.06	7.72 ± 0.08	1.44 ± 0.00
	BAR	6.27 ± 0.06	7.72 ± 0.08	1.44 ± 0.00

Table 1f: Ser → Ala 400 ps

λ windows	method	$\Delta\Delta G_{hyd}$	ΔG_{wat}	ΔG_{vac}
10	Zwanzig	6.85 ± 0.04	2.06 ± 0.02	-4.79 ± 0.04
	BAR	6.86 ± 0.04	2.07 ± 0.02	-4.79 ± 0.04
20	Zwanzig	6.80 ± 0.06	2.02 ± 0.03	-4.78 ± 0.06
	BAR	6.80 ± 0.06	2.02 ± 0.03	-4.78 ± 0.06
50	Zwanzig	6.91 ± 0.06	2.04 ± 0.03	-4.87 ± 0.06
	BAR	6.91 ± 0.06	2.04 ± 0.03	-4.87 ± 0.06
100	Zwanzig	6.89 ± 0.07	2.01 ± 0.04	-4.88 ± 0.06
	BAR	6.89 ± 0.07	2.01 ± 0.04	-4.88 ± 0.06

Supplementary figure 1: calculated and experimental solvation free energies for all side chain mimics reported in this work. Top: all possible combinations (n=552), bottom: excluding ionizable residues (n=210).



Supplementary Table 2: Calculated solvation free energies using the 'shortcut' protocol, which includes only one FEP stage, compared to the results obtained by using two thermodynamic cycles.

$\Delta\Delta G_{solv}$ (kcal/mol)

	experiment	Two cycles	'shortcut'
Ile-Val	-0.16	1.26 ± 0.08	1.49 ± 0.12
Thr-Ser	-0.18	-2.13 ± 0.11	-1.55 ± 0.06
Tyr-Phe	5.35	4.48 ± 0.10	5.20 ± 0.05
Asn-Asp	-70.62	-71.05 ± 0.13	-71.57 ± 0.06
Gln-Glu	-69.86	-71.83 ± 0.14	-72.71 ± 0.11

Supplementary Table 3: Calculated and experimental BIBP3226 binding free energies for hY_1 mutants in kcal/mol

Position	ΔG_{holo}^{calc}	ΔG_{apo}^{calc}	$\Delta\Delta G_{FEP1}$	$\Delta\Delta G_{FEP2}$	$\Delta\Delta G_{FEP3}$	$\Delta\Delta G_{FEP4}$	$\Delta\Delta G_{FEP5}$	$\Delta\Delta G_{FEP6}$	$\Delta\Delta G_{FEP7}$	$\Delta\Delta G_{FEP8}$	$\Delta\Delta G_{FEP9}$	$\Delta\Delta G_{bind}^{calc}$	$\Delta\Delta G_{bind}^{exp}$
Y2.64A	10.39±0.56	9.73±0.58	-0.45±0.25	-0.17±0.08	0.14±0.14	0.12±0.09	0.06±0.57	0.59±0.39	0.34±0.26	0.02±0.03		0.66±0.8	0.3±0.6[a]
N3.28A	40.63±0.38	38.29±0.26	2.79±0.44	0.25±0.03	-0.4±0.09	0.07±0.09	0.13±0.01					2.33±0.46	0±0.1[b]
I3.36A	-7.81±0.38	-7.52±0.2	-0.09±0.03	0.15±0.04	0.26±0.11	-0.3±0.21	-0.33±0.34	0.01±0.03				-0.29±0.43	1.06±0.38[c]
S4.57A	-3.2±0.25	-4.08±0.22	0.99±0.18	-0.21±0.09	0.1±0.27	0.01±0.01						0.88±0.34	-0.4±0.6[a]
F4.60A	0.38±0.32	-2.14±0.29	-0.31±0.04	-0.1±0.04	0.26±0.08	0.01±0.08	0.53±0.33	1.28±0.19	0.8±0.17	0.06±0.02		2.52±0.43	2.05±0.07[c]
W4.60A	13.24±0.33	10.71±0.29	-0.47±0.05	-0.32±0.06	0.36±0.1	0.07±0.21	0.47±0.22	0.71±0.16	1.01±0.2	0.59±0.12	0.1±0.01	2.53±0.44	NA
F4.60W	-12.86±0.46	-12.85±0.41	0.16±0.06	0.22±0.07	-0.1±0.13	-0.06±0.22	0.06±0.4	0.57±0.25	-0.21±0.26	-0.53±0.12	-0.1±0.01	-0.01±0.62	1.23±0.17[c]
Y5.38A	8.32±0.33	8.76±0.28	-0.45±0.15	-0.02±0.04	-0.09±0.2	-0.07±0.12	-0.1±0.2	-0.01±0.2	0.25±0.15	0.05±0.04		-0.44±0.43	>2.3[a]
T5.39A	-3.64±0.32	-6.13±0.57	1.69±0.54	0±0.07	-0.13±0.17	0.86±0.32	0.06±0.04					2.49±0.66	0.94±0.02[c]
Q5.46A	41.35±0.41	39.93±0.39	0.39±0.3	0.13±0.09	0.33±0.27	0.34±0.23	0.23±0.31	0±0.02				1.41±0.57	1.71±0.13[c]
W6.48A	11.3±0.46	12.95±0.81	-0.17±0.09	-0.18±0.1	-0.31±0.12	-0.77±0.64	-0.38±0.37	-0.78±0.45	0.42±0.25	0.42±0.16	0.12±0.04	-1.65±0.93	1.8±0.1[b]
L6.51A	-0.6±0.14	-1.53±0.22	0.03±0.02	-0.2±0.05	0.32±0.11	0.31±0.17	0.43±0.16	0.03±0.01				0.93±0.26	2.28±0.0[c]
T6.52A	-3.52±0.28	-4.23±0.25	0.11±0.34	0.08±0.01	0.31±0.11	0.25±0.11	-0.03±0.01					0.72±0.37	1.52±0.59[c]
N6.55A	39.94±0.23	36.63±0.18	1.49±0.25	-0.18±0.04	1.03±0.13	0.97±0.08	0±0.02					3.31±0.3	>2.3[a]
T6.56A	-2.19±0.43	-3.87±0.62	1.32±0.69	0.02±0.07	0.15±0.19	0.25±0.21	-0.05±0.03					1.68±0.75	-0.2±0.1[a]
F6.58A	1.86±0.48	-1.25±0.5	-0.17±0.1	-0.11±0.1	-0.03±0.16	0.66±0.23	0.87±0.4	0.88±0.32	0.87±0.35	0.14±0.04		3.11±0.7	1±0.1[a]
D6.59A	30.52±1.59	14.93±0.9	19.41±1.72	-1.3±0.1	-1.65±0.42	-1.07±0.44	0.19±0.06					15.58±1.82	>2.3[a]

[a]: data from Sautel 1996 Mol. Pharmacol. 50:285-292
[b]: data from Sjödin 2006 Biochem. J. 393:161-169
[c]: data from Yang 2018, Nature 556:520-524

Supplementary table 4: Variations in the experimental shifts of the binding free energy of BIBP3226 for mutant versions of Y_1R with respect to the WT. Data from Sautel 1996 Mol. Pharmacol. 50:285-292, Sjödin 2006 393:161-169 and Yang 2018 Nature 556:520-524

	$\Delta\Delta G_{WT\text{-}mut}$ (kcal/mol)		
Y_1R	Sautel et al.	Sjodin et al.	Yang et al.
F4.60A	>2.30	0.70	2.05
Q5.46A	>2.30	n.a.	1.71
T5.39A	0.30	1.70	0.94
T6.52A	-0.70	n.a.	1.52

Supplementary table 5: Experimental and calculated changes in free energy associated with the thermal stability of T4-lysozyme due to single-point mutations.

	ΔΔG (kcal/mol)				
mutation	experiment	FEP+	error	resFEP	error
I3A	0.70	3.44	0.42	3.33	0.35
M6A	1.90	5.05	0.09	5.12	0.30
L7A	2.60	1.28	0.08	4.73	0.46
I17A	2.70	4.52	0.08	4.78	0.30
I27A	3.10	3.39	0.09	3.29	0.25
I29A	2.60	1.65	0.08	3.62	0.35
L33A	3.60	3.87	0.11	5.09	0.36
L39A	0.90	1.47	0.08	2.08	0.18
N40A	-0.32	0.11	0.12	-1.85	0.29
S44A	-0.34	-0.30	0.09	0.23	0.58
L46A	1.86	3.52	0.09	4.02	0.16
I50A	2.00	1.87	0.08	-0.31	0.13
I58A	3.20	4.96	0.10	6.42	0.29
T59A	1.50	1.93	0.49	2.64	0.18
L66A	3.90	5.40	0.07	4.03	0.24
F67A	1.90	0.38	0.12	2.69	0.43
N68A	0.05	0.72	0.10	-1.62	0.37
V71A	1.50	2.82	0.05	3.37	0.19
I78A	1.60	1.84	0.09	2.18	0.24
L84A	3.90	6.46	0.07	7.76	0.24
V87A	1.70	2.66	0.06	3.21	0.17
L91A	3.10	4.56	0.08	6.14	0.34
V94A	1.80	-0.75	0.07	-0.48	0.24
L99A	4.50	6.41	0.07	7.72	0.13
I100A	3.40	3.43	0.09	3.42	0.29
V103A	1.91	1.67	0.07	2.97	0.13
F104A	3.10	3.17	0.08	4.09	0.32
Q105A	0.60	1.95	0.24	1.34	0.35
M106A	2.30	4.76	0.08	3.39	0.30
V111A	1.30	-0.41	0.08	-0.34	0.18
T115A	0.14	0.09	0.11	-1.06	0.31
N116A	-0.17	0.29	0.08	-2.43	0.25
S117A	-1.27	-3.17	0.06	-0.25	0.44
L118A	3.50	5.56	0.08	6.41	0.24
M120A	0.20	0.23	0.08	2.00	0.33
L121A	2.30	4.40	0.07	5.75	0.31
Q122A	0.24	2.08	0.10	1.00	0.14

Q123A	0.22	1.51	0.11	2.53	0.40
V131A	-0.39	-0.07	0.20	-0.96	0.10
L133A	4.30	4.88	0.08	5.61	0.29
V149A	2.87	4.49	0.07	4.12	0.21
F153A	3.80	5.26	0.10	6.71	0.24
T157A	0.50	-0.33	0.12	0.20	0.27

Supplementary figure 2: calculated free energies for every Ala-X-Ala -> Ala-Ala-Ala mutation, as a function of total replicates used. The blue line represents the average calculated energies from 30 replicates, with the s.e.m. indicated as error bars.