

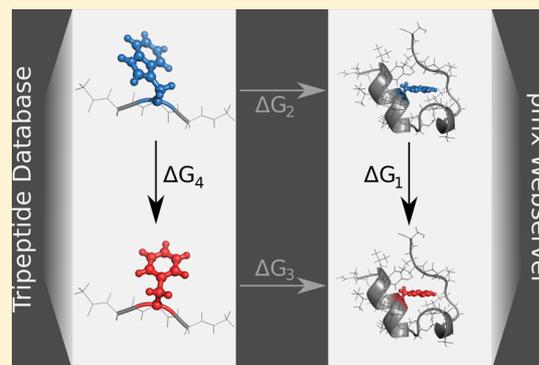
# pmx Webserver: A User Friendly Interface for Alchemy

Vytautas Gapsys\*<sup>1</sup> and Bert L. de Groot\*

Computational Biomolecular Dynamics Group, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, Göttingen, 37077, Germany

**S** Supporting Information

**ABSTRACT:** With the increase of available computational power and improvements in simulation algorithms, alchemical molecular dynamics based free energy calculations have developed into routine usage. To further facilitate the usability of alchemical methods for amino acid mutations, we have developed a web based infrastructure for obtaining hybrid protein structures and topologies. The presented webserver allows amino acid mutation selection in five contemporary molecular mechanics force fields. In addition, a complete mutation scan with a user defined amino acid is supported. The output generated by the webserver is directly compatible with the Gromacs molecular dynamics engine and can be used with any of the alchemical free energy calculation setup. Furthermore, we present a database of input files and precalculated free energy differences for tripeptides approximating a disordered state of a protein, of particular use for protein stability studies. Finally, the usage of the webserver and its output is exemplified by performing an alanine scan and investigating thermodynamic stability of the Trp cage mini protein. The webserver is accessible at <http://pmx.mpibpc.mpg.de>



## INTRODUCTION

Molecular dynamics (MD) based alchemical free energy calculations have become a method of choice to obtain high accuracy estimates of free energy differences from biomolecular simulations. The increased usability of alchemical approaches was in large part enabled by the automation of the otherwise technically complex simulation setup procedure.<sup>1–4</sup> Preparation of the alchemical amino acid mutation simulations requires the generation of hybrid protein structures and topologies where both physical states—wild type and mutant—are represented. Recently, the software pmx was presented<sup>1</sup> providing a command-line based toolkit allowing generation of hybrid structures/topologies for amino acids in a number of contemporary molecular mechanics force fields. The output of pmx is compatible with the Gromacs<sup>5</sup> simulation package and can be directly used with any state-of-art free energy calculation protocol.<sup>6,7</sup> pmx based alchemical amino acid mutations have proven to yield accurate results in calculating protein thermodynamic stabilities<sup>8,9</sup> changes in protein–protein interaction free energies<sup>9</sup> and shifting equilibrium of the protein conformational substates.<sup>10</sup>

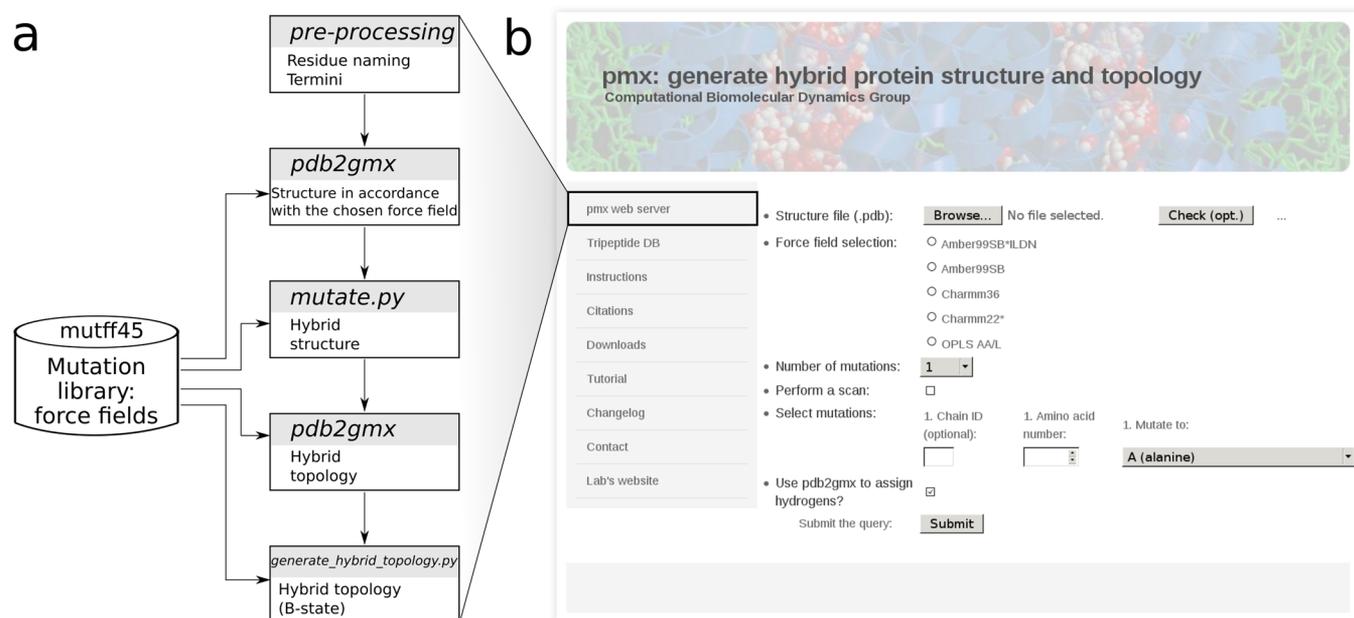
For the calculation of protein thermodynamic stabilities a proper representation of an unfolded state is required. Obtaining a converged and reliable sampling of a protein in a disordered state is a computationally challenging task, the results of which strongly depend on the choice of a force field and water model.<sup>11–14</sup> For alchemical thermodynamic stability calculations, however, it has been observed that a simplified representation of the unfolded state suffices to obtain accurate free energy estimates matching experimental measure-

ments.<sup>8,9,15–18</sup> Numerous variants of the unfolded state representations have been used, ranging from a single capped amino acid<sup>15</sup> to a heptapeptide fragment of the protein of interest.<sup>15–18</sup> Here we follow the procedure by Seeliger and de Groot,<sup>8</sup> who demonstrated that quite accurate thermodynamic stability estimates can be obtained when the unfolded state is represented as a capped GXG (where X stands for an amino acid to be mutated) tripeptide, thus without requiring an amino acid context from the folded protein. Such a construct allows precalculating the whole library of unfolded state reference values for subsequent routine protein scans in search for thermodynamically stabilizing mutations. Naturally, the reference free energy values are force field and mutation library dependent. Therefore, in the context of current work we have calculated reference free energy estimates for five MD force fields.

Here, we report on the progress in development of the software further facilitating usability of the alchemical MD based approaches for amino acid mutations. A webserver has been created as a user-friendly platform to generate amino acid hybrid structures and topologies for alchemical free energy calculations. Running pmx<sup>1</sup> in the back-end, the webserver acts as a front-end solution for a quick and convenient preparation of the input that can be directly used in the alchemical simulations with the Gromacs MD engine. In this work we describe the specifics of the webserver usage, provide a web-based access to the pregenerated files for the free energy

**Received:** August 24, 2016

**Published:** February 9, 2017



**Figure 1.** pmx webserver. (a) Schematic representation of the back-end pmx based framework that lies at the basis of the webserver. (b) Webserver's front page.

calculations of the unfolded reference state, as well as precalculated free energy estimates for the reference states. Finally, we provide several illustrative applications of MD based alchemical thermodynamic protein stability calculations in the Trp cage protein, illustrating the strengths of the alchemical approaches for amino acid mutations. For an in depth investigation of the performance of different force fields and various mutation effects in a large scale mutation scan by means of Gromacs and pmx we refer the reader to our previous publication.<sup>9</sup>

## RESULTS AND DISCUSSION

**pmx and the Webserver.** The core functionality of pmx has been described in detail in our previous work.<sup>1</sup> The brief summary of the pmx based hybrid structure/topology generation framework that lies at the core of the webserver is provided in Figure 1a. pmx utilizes pregenerated force field specific mutation libraries. The pmx command-line tool provides utilities for the generation of new libraries, whereas five prepared mutation libraries are readily supported by the webserver: Amber99sb,<sup>19</sup> Amber99sb\*ILDN,<sup>20,21</sup> OPLSAA/L,<sup>22,23</sup> Charmm22\*,<sup>24</sup> and Charmm36.<sup>25</sup> New force fields will be supported by pmx as soon as their official ports appear in Gromacs. The mutation libraries follow a single topology approach, thus minimizing the perturbation to the system upon an amino acid mutation and allowing faster convergence of the free energy estimates.<sup>26,27</sup>

Prior to starting the main structure/topology generation, the webserver runs several routines processing a user submitted structure. The specifics of these structure manipulations are described in the following section.

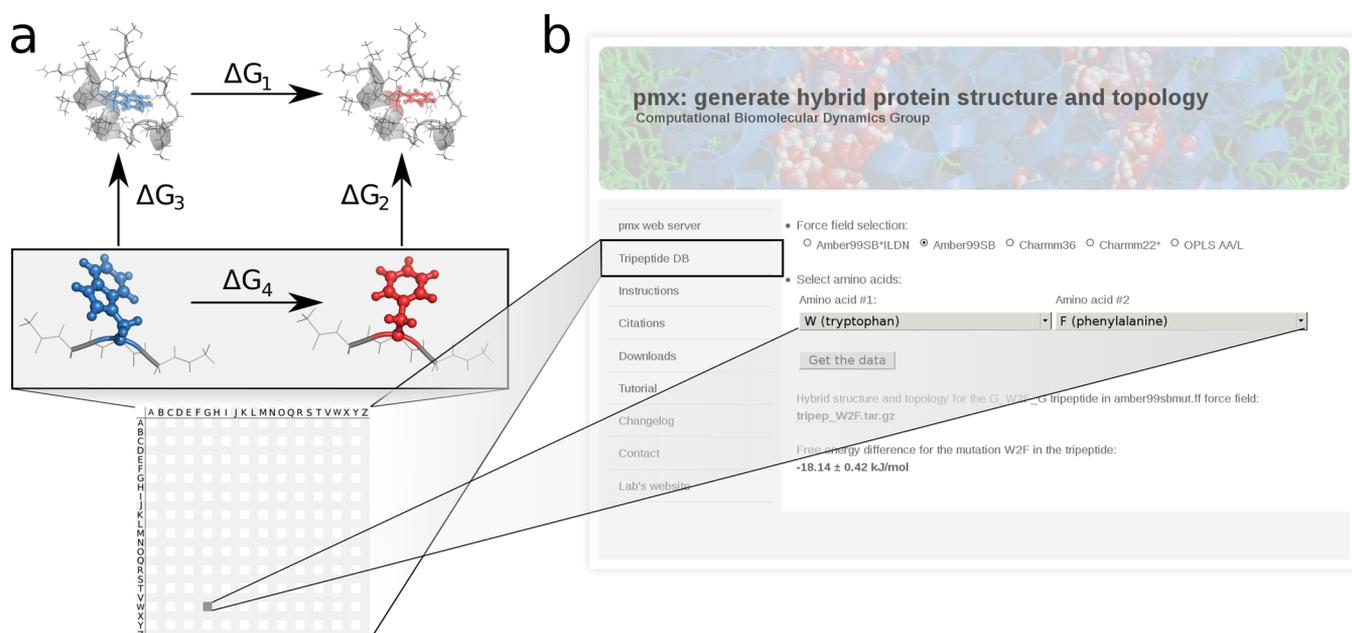
**Webserver Specifics. Structure Files.** The webserver accepts a structure in .pdb format. If the provided structure file contains more than one model, only the first one will be considered for further processing. In the preprocessing step all the moieties not recognized as a protein will be discarded. The purpose of the pmx webserver is the generation of protein hybrid structures and topologies, thus the other components of

a simulation system, e.g. water, ions, lipids, ligands etc., need to be added by the user to the output from the webserver. If any atoms or entire side-chains are missing the webserver will not correct for them and the run will result in an error. Numerous specialized packages exist designed to correct an initial protein structure, e.g. Modeller<sup>28</sup> or WhatIf.<sup>29</sup> The nonstandard residues that are not readily defined in the force fields will not be recognized by the webserver. For such cases the users are encouraged to use a command-line based pmx version after porting the required changes into the selected mutation force field.

The chains and terminal residues in the structure are determined by the TER entries or changes in the chain IDs in the input .pdb file. Hydrogen atoms can be generated by means of the *pdb2gmx* program. If hydrogen generation is selected, the disulfide bridges between the cysteines are also created by *pdb2gmx*. Otherwise, if the user decides to keep the hydrogens in the submitted structure, the atoms and residues will be only renamed for compatibility with the force field conventions and all the protonation states will remain intact. Note that this step may lead to errors if hydrogen naming in the user specified file cannot be recognized. In this case, execution of the query will be terminated and the user will be notified about an error at the structure preprocessing step.

**Force Fields.** Mutation libraries have been generated for five aforementioned force fields and mutations in these force fields are supported by the webserver. To prepare molecular dynamics simulation input using the structures/topologies containing hybrid residues it is necessary to download these modified force field files (termed *mutff45*, as the topology organization here follows that of the Gromacs 4.5 and higher versions). Subsequently, the GMXLIB variable needs to be set to the modified force field directory and the further simulation preparation ought to follow a user selected setup for the free energy calculations.<sup>6,7</sup>

All the standard residue mutations are supported by pmx with only a few exceptions, namely, prolines, disulfide bridge forming cysteines and terminal residues. If a terminal residue is



**Figure 2.** Tripeptide database. (a) Thermodynamic cycle to calculate protein stability. The disordered state is approximated by a capped GXG tripeptide. The simulation input files and  $\Delta G$  values for all the supported mutations in pmx for the tripeptide were tabulated and made accessible as a searchable database. (b) Web-based tripeptide database.

to be mutated, the user is advised to add a capping group to the terminal residue, thus technically making the amino acid not terminal any more. Glycine involving mutations are not supported in the Charmm family force fields (Charmm36 and Charmm22\*), because performing an alchemical transition for such mutations requires morphing the CMAP correction term<sup>30,31</sup> between the physical states. However, this feature is not yet available in Gromacs. Proline and disulfide bridge forming cysteine mutations require bond breaking, which ought to be avoided in alchemical free energy calculations.<sup>32</sup>

Naming conventions of the amino acids in the supported force fields are provided in Table S1.

**Mutations.** To specify an amino acid to be mutated the user needs to provide a residue number and optionally a chain identifier. If no chain ID is given, the webserver will seek for the first residue number in the structure matching the one entered by the user. An arbitrary number of mutations can be introduced into a protein at once. Up to 3 mutations can be selected interactively on the webserver. If more than 3 mutations are to be introduced, the user can enter them in a web form following a simple format: *chainID(opt.) amino\_acid\_number mutation\_1 letter\_code*. An example of the mutation input for the web form is provided in Table S2.

**Scan.** The webserver also allows performing a mutational scan over a protein with a user defined amino acid. If the scan option is selected, the same steps in the run workflow as outlined for a single mutation (Figure 1a) will be iterated over all the residues in the protein or over the residues of a user defined chain. The webserver will discard unsupported or irrelevant mutations, e.g. proline mutations, glycine mutations in the Charmm family force fields or mutations to the same amino acid.

**Output.** The generated hybrid structure is provided in a *hybrid.pdb* file. The hybrid topology is stored in a *hybrid.itp* file. For a multichain protein several .itp files are generated: one for each chain. The .itp files are included in the *hybrid.top* file. The obtained structure and topology files can readily be used in the

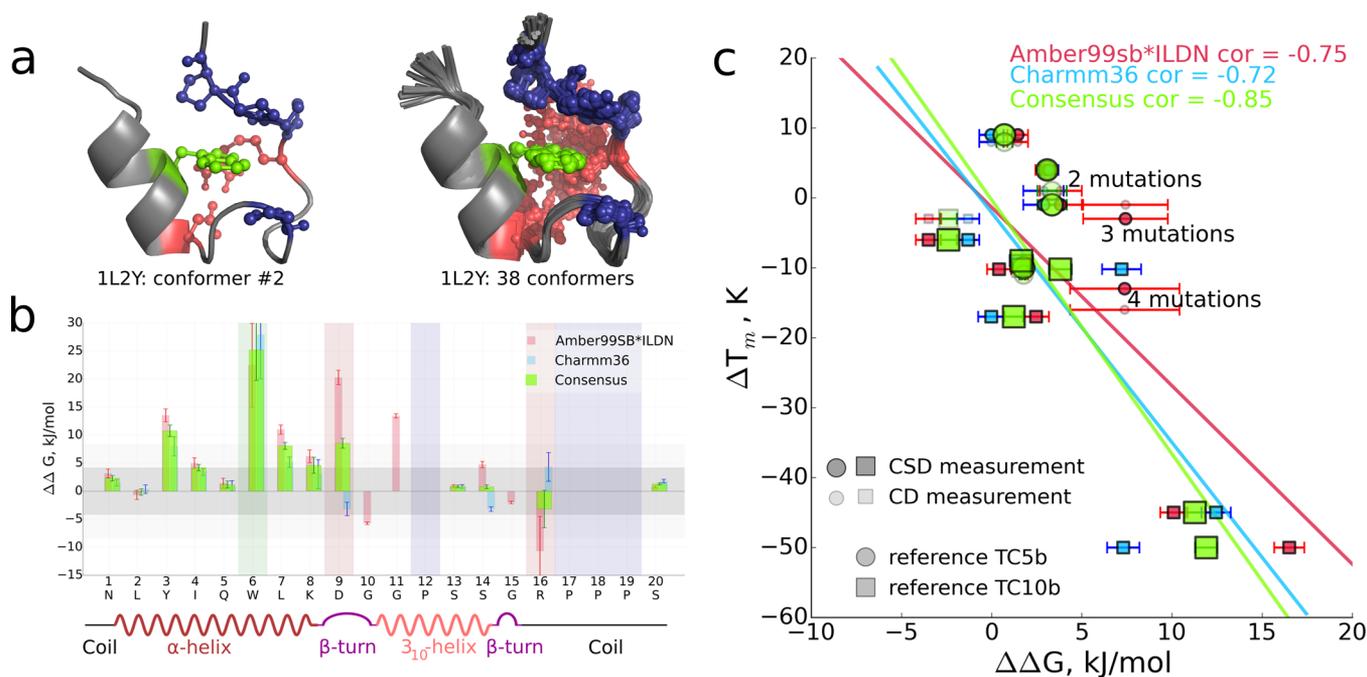
subsequent simulation system preparation following a protocol of the user's choice.

**Terms of Use.** The webserver requires no log in and can be used anonymously. The mutation libraries and generated hybrid structures/topologies can be used without any restrictions.

**Unfolded Reference State.** Representing the disordered protein state by a capped GXG tripeptide allows preparing the hybrid structures/topologies for all possible mutations in the unfolded state: schematically represented by a matrix in Figure 2a. We have generated all the mutation combinations supported by pmx for the GXG tripeptide in five force fields for which the mutation libraries are available. The webserver allows for a convenient search and retrieval of the pregenerated tripeptide hybrid structures/topologies.

We have also calculated the  $\Delta G$  values in the GXG tripeptide for all the charge conserving mutations supported by pmx. The tabulated values for five force fields are provided in the Tables S4–S8 and can also be accessed via the webserver interface. When using the precalculated  $\Delta G$  values, it is important to bear in mind that these estimates depend on a number of conditions: thermodynamic variables (temperature, pressure), salt concentration, water model used. In addition, some dependence may be expected on the various MD simulation specific parameters, e.g. the nonbonded interaction cutoff distances. The exact details of the tripeptide simulation setup used in the current work are provided in the Supporting Information.

**Trp Cage Mini Protein Alanine Scan.** To illustrate the potential venues of the pmx webserver usage we performed several alchemical free energy based investigations of the thermodynamic stability of the Trp cage mini protein.<sup>33</sup> Trp cage is a 20 amino acid long peptide that adopts a stable fold by burying a tryptophan residue in a hydrophobic core (Figure 3a). A number of Trp cage variants have been designed aiming to optimize stability of the protein.<sup>33,34</sup>



**Figure 3.** Trp cage mini protein mutation analysis. (a) Trp cage NMR ensemble (PDB ID 1L2Y<sup>33</sup>) with the crucial features for retaining the fold marked in color: green buried Trp6 residue; blue conserved proline residues; red Asp9 and Arg16 forming a salt bridge. (b) Results of an alanine scan in two force fields: Amber99sb\*ILDN (red bars) and Charmm36 (blue bars). The consensus results (green bars) were obtained by averaging the results of the two force fields. The gray shading indicates 1 kcal/mol region. The colored shadings correspond to the Trp cage structural features marked in color in panel a. (c) Calculated double free energy differences plotted against the experimentally measured changes in the melting temperature for the Trp cage variants (Table S3). The  $\Delta\Delta G$  values were calculated in Amber99sb\*ILDN (red) and Charmm36 (blue) force fields. The consensus result (green) was calculated as an average over the two force fields. Transparent symbols correspond to the experimental values obtained from the circular dichroism (CD) measurements. Nontransparent symbols depict  $\Delta T_m$  values estimated from the chemical shift deviations (CSD). The correlations were calculated using  $\Delta T_m$  values from the CSD experiments. Circles denote the set of mutations where the TC5b variant was used as a reference, whereas squares mark those mutations where TC10b served as a reference.

For the first computational study we utilized the webserver's mutational scanning feature to perform an alanine scan of the TC5b<sup>33</sup> sequence (Table S3) of the Trp cage. TC5b is an optimized version of the protein, where, besides the buried tryptophan (marked in green in Figure 3a) and conserved proline residues (blue in Figure 3a), a salt bridge between Asp9 and Arg16 residues is formed (red in Figure 3a), thus stabilizing the fold.

We started by generating the hybrid structures and topologies in two force fields: Amber99sb\*ILDN and Charmm36. The termini of the protein were capped, thus allowing to mutate the first and last residues as well. Subsequently, the pregenerated hybrid structures and topologies for the corresponding mutations in a GXG tripeptide were retrieved. For this investigation we did not use the library of the precalculated values but performed simulations of the disordered state by placing a tripeptide in the same simulation box with the folded protein. This setup allows calculating the double free energy differences also for charge changing mutations, while preserving the neutral charge of the simulation box during an alchemical transition.<sup>7</sup> Finally, the changes in the Trp cage thermodynamic stabilities were obtained using the nonequilibrium free energy calculation procedure (see Methods).

The calculated  $\Delta\Delta G$  values are depicted in Figure 3b: red bars denote the estimates in the Amber99sb\*ILDN force field, while blue bars are for Charmm36. The consensus results (green bars in Figure 3b) were obtained by averaging the results from two force fields. Most of the mutations were

predicted to have a destabilizing effect: positive  $\Delta\Delta G$  estimates. This is an expected outcome for a protein that has already been designed to exhibit high thermal stability. Substitution of the crucial Trp6 residue with alanine clearly results in the strongest destabilization. Similarly, replacing another aromatic residue known to be important for the Trp cage fold (Tyr3), is also estimated to have a significant destabilizing effect.

Calculated results for the glycine mutations also correspond well to the experimental observations. Gly11 is known to be conserved across the different Trp cage variants, thus making it important for retaining the protein's fold.<sup>33</sup> This is well confirmed by our calculations: a large positive  $\Delta\Delta G$  value is estimated for this mutation. The glycine residue at the 10th position, on the other hand, has been shown to be replacable.<sup>35</sup> In accordance, the calculated free energy difference predicts stabilization of the protein by mutating Gly10 to alanine. Note that glycine mutation results come only from the Amber99sb\*ILDN force field, since alchemical mutations of glycine are not available in the Charmm family force fields due to the CMAP correction term.

The force field results differ for the salt bridge forming residues (Asp9, Arg16). The consensus results indicate that replacement of Asp9 is destabilizing. For the Arg16 mutation the effect is mildly stabilizing, however, a large error is associated with this estimate (Figure 3b). In general, free energy estimates for the charge changing mutations have been shown to have a worse agreement with the experimentally measured values.<sup>9</sup>

**Trp Cage Stability Optimization.** In our second investigation of the Trp cage thermodynamic stability, we aimed for a quantitative comparison of the calculated free energy values with the experimental measurements. For that purpose we used the melting temperature ( $T_m$ ) values measured for a number of Trp cage variants (Table S3).<sup>34,36</sup> It has been shown that correlation between the  $\Delta T_m$  and  $\Delta\Delta G$  values is to be expected.<sup>9</sup>

Two sets of mutants were generated by considering the variants TCSb and TC10b as the WT references to introduce the mutations. In the first set, where TCSb served as a reference, up to four mutations were introduced at once. In the set with the TC10b as a reference, only single amino acid mutations were introduced.

The calculated  $\Delta\Delta G$  values are plotted against the  $\Delta T_m$  measurements in Figure 3c. Both force fields, as well as the consensus approach, were able to discriminate between the strongly destabilizing mutations and those having a mild effect on the stability. In the representation in Figure 3c, the agreement between the calculations and experiment manifests as a negative correlation: destabilization causes decrease in the melting temperature and increase in the double free energy difference. For the cases where more than one mutation was introduced, the correct trend was picked up by the alchemical calculations as well. Notably, with the introduction of more mutations the uncertainty associated with the estimated value increases, as indicated by the larger error bars for the marked mutations in Figure 3c.

## CONCLUSIONS AND OUTLOOK

In this work we presented a web based server for the amino acid hybrid structure and topology generation in five molecular mechanics force fields facilitating high accuracy alchemical free energy calculations. Furthermore, a database of pregenerated structures/topologies for an approximation of a disordered protein state has been assembled and made available via the webserver utilities. In addition, the  $\Delta G$  values for the disordered state were precalculated and tabulated for all the pmx supported charge conserving mutations.

The presented webserver is aimed at lowering the barrier to correctly preparing input for the alchemical MD simulations that otherwise may be technically demanding. Naturally, the next step in the development could be a fully automated web based service for the MD based estimation of the free energy changes upon an amino acid mutation. Similar webserver already exist for the computationally less demanding statistical approaches.<sup>37–39</sup> Recent advances in accelerating molecular dynamics simulations<sup>40,41</sup> together with the approach presented in this work bring the idea of a fully automated alchemical MD based mutation scan, requiring no user intervention, one step closer to realization.

## METHODS

For the tripeptide and subsequent Trp cage simulations a nonequilibrium free energy calculation protocol was employed. First, equilibrium ensembles of the physical end states were obtained. Afterward, short alchemical transitions between the states were performed, evaluating the work done by the system during the transitions. Finally, we used the Crooks Fluctuation Theorem<sup>42</sup> to obtain the free energy values and associated errors by means of the maximum likelihood estimator.<sup>43</sup>

All the molecular dynamics simulations were performed at constant temperature and pressure. The systems were solvated with TIP3P water molecules<sup>44</sup> and ions were added to keep the simulation box neutral and reach 150 mM salt concentration. Long range electrostatic interactions were treated by means of the particle mesh Ewald (PME) method.<sup>45</sup> For the alchemical transitions soft-core function was used for the nonbonded interactions.<sup>46</sup> Further simulation details are provided in the Supporting Information.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00498.

Simulation details, a table of the amino acid naming conventions in the mutation libraries, an example of an input for more than three mutations, a table of the Trp cage sequence variants, five tables with the precalculated free energy differences for the GXG tripeptides in five force fields, and additional references (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: [vgapsys@gwdg.de](mailto:vgapsys@gwdg.de)

\*E-mail: [bgroot@gwdg.de](mailto:bgroot@gwdg.de)

### ORCID

Vytautas Gapsys: 0000-0002-6761-7780

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work has been done as part of the BioExcel CoE ([www.bioexcel.eu](http://www.bioexcel.eu)), a project funded by the European Union contract H2020-EINFRA-2015-1-675728. V.G. acknowledges support by Boehringer Ingelheim Pharma GmbH. We are grateful to Vance Jaeger and Wojciech Kopec for discussions, suggestions, and webserver testing.

## REFERENCES

- (1) Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. pmx: Automated Protein Structure and Topology Generation for Alchemical Perturbations. *J. Comput. Chem.* **2015**, *36*, 348–354.
- (2) Loeffler, H. H.; Michel, J.; Woods, C. FESetup: Automating Setup for Alchemical Free Energy Simulations. *J. Chem. Inf. Model.* **2015**, *55*, 2485–2490.
- (3) Klimovich, P. V.; Mobley, D. L. A Python Tool to Set Up Relative Free Energy Calculations in GROMACS. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 1007–1014.
- (4) Ramadoss, V.; Dehez, F.; Chipot, C. AlaScan: A Graphical User Interface for Alanine Scanning Free-Energy Calculations. *J. Chem. Inf. Model.* **2016**, *56*, 1122–1126.
- (5) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845–854.
- (6) Shirts, M. R.; Mobley, D. L. An Introduction to Best Practices in Free Energy Calculations. *Methods Mol. Biol.* **2013**, *924*, 271–311.
- (7) Gapsys, V.; Michielssens, S.; Peters, J. H.; de Groot, B. L.; Leonov, H. Calculation of Binding Free Energies. *Methods Mol. Biol.* **2015**, *1215*, 173–209.
- (8) Seeliger, D.; de Groot, B. L. Protein Thermostability Calculations Using Alchemical Free Energy Simulations. *Biophys. J.* **2010**, *98*, 2309–2316.

- (9) Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. Accurate and Rigorous Prediction of the Changes in Protein Free Energies in a Large-Scale Mutation Scan. *Angew. Chem.* **2016**, *128*, 7490–7494.
- (10) Michielssens, S.; Peters, J. H.; Ban, D.; Pratihari, S.; Seeliger, D.; Sharma, M.; Giller, K.; Sabo, T. M.; Becker, S.; Lee, D.; Griesinger, C.; de Groot, B. L. A Designed Conformational Shift To Control Protein Binding Specificity. *Angew. Chem.* **2014**, *126*, 10535–10539.
- (11) Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; Shaw, D. E. Structure and Dynamics of an Unfolded Protein Examined by Molecular Dynamics Simulation. *J. Am. Chem. Soc.* **2012**, *134*, 3787–3791.
- (12) Best, R. B.; Zheng, W.; Mittal, J. Balanced Protein–Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124.
- (13) Rauscher, S.; Gapsys, V.; Gajda, M. J.; Zweckstetter, M.; de Groot, B. L.; Grubmüller, H. Structural Ensembles of Intrinsically Disordered Proteins Depend Strongly on Force Field: A Comparison to Experiment. *J. Chem. Theory Comput.* **2015**, *11*, 5513–5524.
- (14) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B* **2015**, *119*, 5113–5123.
- (15) Sneddon, S. F.; Tobias, D. J. The Role of Packing Interactions in Stabilizing Folded Proteins. *Biochemistry* **1992**, *31*, 2842–2846.
- (16) Prevost, M.; Wodak, S. J.; Tidor, B.; Karplus, M. Contribution of the Hydrophobic Effect to Protein Stability: Analysis Based on Simulations of the Ile-96→Ala Mutation in Barnase. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 10880–10884.
- (17) Dang, L. X.; Merz, K. M., Jr; Kollman, P. A. Free Energy Calculations on Protein Stability: Thr-157 → Val-157 Mutation of T4 Lysozyme. *J. Am. Chem. Soc.* **1989**, *111*, 8505–8508.
- (18) Pitera, J. W.; Kollman, P. A. Exhaustive Mutagenesis In Silico: Multicoordinate Free Energy Calculations on Proteins and Peptides. *Proteins: Struct., Funct., Genet.* **2000**, *41*, 385–397.
- (19) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712–725.
- (20) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.
- (21) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950–1958.
- (22) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (23) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (24) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* **2011**, *100*, L47–L49.
- (25) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone  $\phi$ ,  $\psi$  and Side-Chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (26) Pearlman, D. A. A Comparison of Alternative Approaches to Free Energy Calculations. *J. Phys. Chem.* **1994**, *98*, 1487–1493.
- (27) Michel, J.; Essex, J. W. Prediction of Protein–Ligand Binding Affinity by Free Energy Simulations: Assumptions, Pitfalls and Expectations. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 639–658.
- (28) Šali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
- (29) Vriend, G. WHAT IF: A Molecular Modeling and Drug Design Program. *J. Mol. Graphics* **1990**, *8*, 52–56.
- (30) MacKerell, A. D., Jr; Feig, M.; Brooks, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004**, *126*, 698–699.
- (31) MacKerell, A. D., Jr; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400–1415.
- (32) Liu, S.; Wang, L.; Mobley, D. L. Is Ring Breaking Feasible in Relative Binding Free Energy Calculations? *J. Chem. Inf. Model.* **2015**, *55*, 727–735.
- (33) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-Residue Protein. *Nat. Struct. Biol.* **2002**, *9*, 425–430.
- (34) Barua, B.; Lin, J. C.; Williams, V. D.; Kummeler, P.; Neidigh, J. W.; Andersen, N. H. The Trp-cage: Optimizing the Stability of a Globular Miniprotein. *Protein Eng., Des. Sel.* **2008**, *21*, 171–185.
- (35) Herman, R. E.; Badders, D.; Fuller, M.; Makienko, E. G.; Houston, M. E.; Quay, S. C.; Johnson, P. H. The Trp Cage Motif as a Scaffold for the Display of a Randomized Peptide Library on Bacteriophage T7. *J. Biol. Chem.* **2007**, *282*, 9813–9824.
- (36) Culik, R. M.; Serrano, A. L.; Bunagan, M. R.; Gai, F. Achieving Secondary Structural Resolution in Kinetic Measurements of Protein Folding: A Case Study of the Folding Mechanism of Trp-cage. *Angew. Chem., Int. Ed.* **2011**, *50*, 10884–10887.
- (37) Capriotti, E.; Fariselli, P.; Casadio, R. I-Mutant2.0: Predicting Stability Changes Upon Mutation from the Protein Sequence or Structure. *Nucleic Acids Res.* **2005**, *33*, W306–W310.
- (38) Dehouck, Y.; Kwasigroch, J. M.; Gilis, D.; Rooman, M. PoPMuSiC 2.1: A Web Server for the Estimation of Protein Stability Changes upon Mutation and Sequence Optimality. *BMC Bioinf.* **2011**, *12*, 151.
- (39) Laimer, J.; Hofer, H.; Fritz, M.; Wegenkittl, S.; Lackner, P. MAESTRO – Multi Agent Stability Prediction upon Point Mutations. *BMC Bioinf.* **2015**, *16*, 116.
- (40) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1*, 19–25.
- (41) Kutzner, C.; Páll, S.; Fechner, M.; Esztermann, A.; de Groot, B. L.; Grubmüller, H. Best Bang for your Buck: GPU Nodes for GROMACS Biomolecular Simulations. *J. Comput. Chem.* **2015**, *36*, 1990–2008.
- (42) Crooks, G. E. Entropy Production Fluctuation Theorem and the Nonequilibrium Work Relation for Free Energy Differences. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1999**, *60*, 2721–2726.
- (43) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. Equilibrium Free Energies From Nonequilibrium Measurements Using Maximum-Likelihood Methods. *Phys. Rev. Lett.* **2003**, *91*, 140601.
- (44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (45) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (46) Gapsys, V.; Seeliger, D.; de Groot, B. L. New Soft-Core Potential Function for Molecular Dynamics Based Alchemical Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 2373–2382.