Thermostability

Deutsche Ausgabe: DOI: 10.1002/ange.201510054
Internationale Ausgabe: DOI: 10.1002/anie.201510054

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

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

Abstract: *The prediction of mutation-induced free-energy changes in protein thermostability or protein–protein binding is of particular interest in the fields of protein design, biotechnology, and bioengineering. Herein, we achieve remarkable accuracy in a scan of 762 mutations estimating changes in protein thermostability based on the first principles of statistical mechanics. The remaining error in the free-energy estimates appears to be due to three sources in approximately equal parts, namely sampling, force-field inaccuracies, and experimental uncertainty. We propose a consensus force-field approach, which, together with an increased sampling time, leads to a free-energy prediction accuracy that matches those reached in experiments. This versatile approach enables accurate free-energy estimates for diverse proteins, including the prediction of changes in the melting temperature of the membrane protein neurotensin receptor 1.*

Evolution has optimized proteins to perform their specific functions in the environmental conditions native to the host organism. Altering certain thermodynamic properties of a protein is often sought after by the pharmaceutical and chemical industries,[1] for example, enhancing the thermal stability of a molecule or altering the strength of a specific protein–protein interaction. Such modifications may be achieved by means of amino acid mutations, and the prediction of free-energy changes upon mutation is thus of key interest.

For an ideal free-energy prediction method, the predictive accuracy should be of the same range as that reached in experiments. A perfect method ought to be system-independent, and hence not require fitting to experimental data. It should be able to robustly predict thermostabilities (or binding affinities) for different mutation types in the core of a protein as well as in the solvent-exposed regions, which requires that solute–solvent interactions are taken into account. The ability to change the environmental conditions, for example, the temperature, pressure, pH, or salt concentration, is another necessary requirement.

Alchemical free-energy calculations have the potential to fulfill these requirements. The approach relies on molecular dynamics (MD) simulations, where both the solute and solvent are modeled atomistically. MD simulations are not restricted to any particular protein class and allow for precise control over the simulation conditions. The estimation of the free-energy differences is based on rigorous theories.[2–4] The major bottlenecks to a routine employment of these methods are high computational costs (subsequently leading to the related undersampling problem), the complex simulation setup, and the dependence of the results on the chosen molecular mechanics force field. Whereas the former two aspects are merely technical caveats, the force-field development is an active field requiring constant updates and benchmarks.[5,6]

Herein, we utilized a state-of-the-art setup[7] for alchemical free-energy calculations to carry out a large-scale protein thermostability scan, in total comprising 762 mutations employing six contemporary force fields. We provide insight into the remarkable prediction accuracy that can be reached with these rigorous alchemical methods and subsequently propose a method to reduce the inherent force-field bias by adopting a consensus approach. An in-depth analysis of the error sources related to the free-energy estimates revealed that the alchemical approaches are able to reach the experimental level of uncertainty. In addition, a simple "rule of thumb" was found, stating that the error is due to the force field, the sampling, and the experimental uncertainty with comparable contributions. Furthermore, we investigated the versatility of the methods in staphylococcal nuclease and a membrane protein, namely neurotensin receptor 1. Finally, we provide insight into the application of the approach to the calculation of changes in the relative protein–protein binding free energies upon an amino acid mutation.

A set of 119 mutations in the enzyme barnase was subjected to alchemical free-energy calculations (Figure 1 a). We estimated the relative unfolding free energies for all of the mutations in six contemporary MD force fields (see the Supporting Information). The overall best-performing force field, Charmm36H, achieved a remarkable prediction accuracy of 3.8 kJ mol^{-1} in terms of the averaged unsigned error (AUE) with respect to the experimental measurements (Figure 1 b). In general, predicting thermostability changes

[*] Dr. V. Gapsys, Dr. S. Michielssens, Dr. B. L. de Groot
Computational Biomolecular Dynamics Group
Max Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Göttingen (Germany)
E-mail: vgapsys@gwdg.de
bgroot@gwdg.de
Dr. D. Seeliger
Lead Identification and Optimization Support
Boehringer Ingelheim Pharma GmbH & Co. KG
Birkendorfer Strasse 65, 88397 Biberach (Germany)

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201510054.

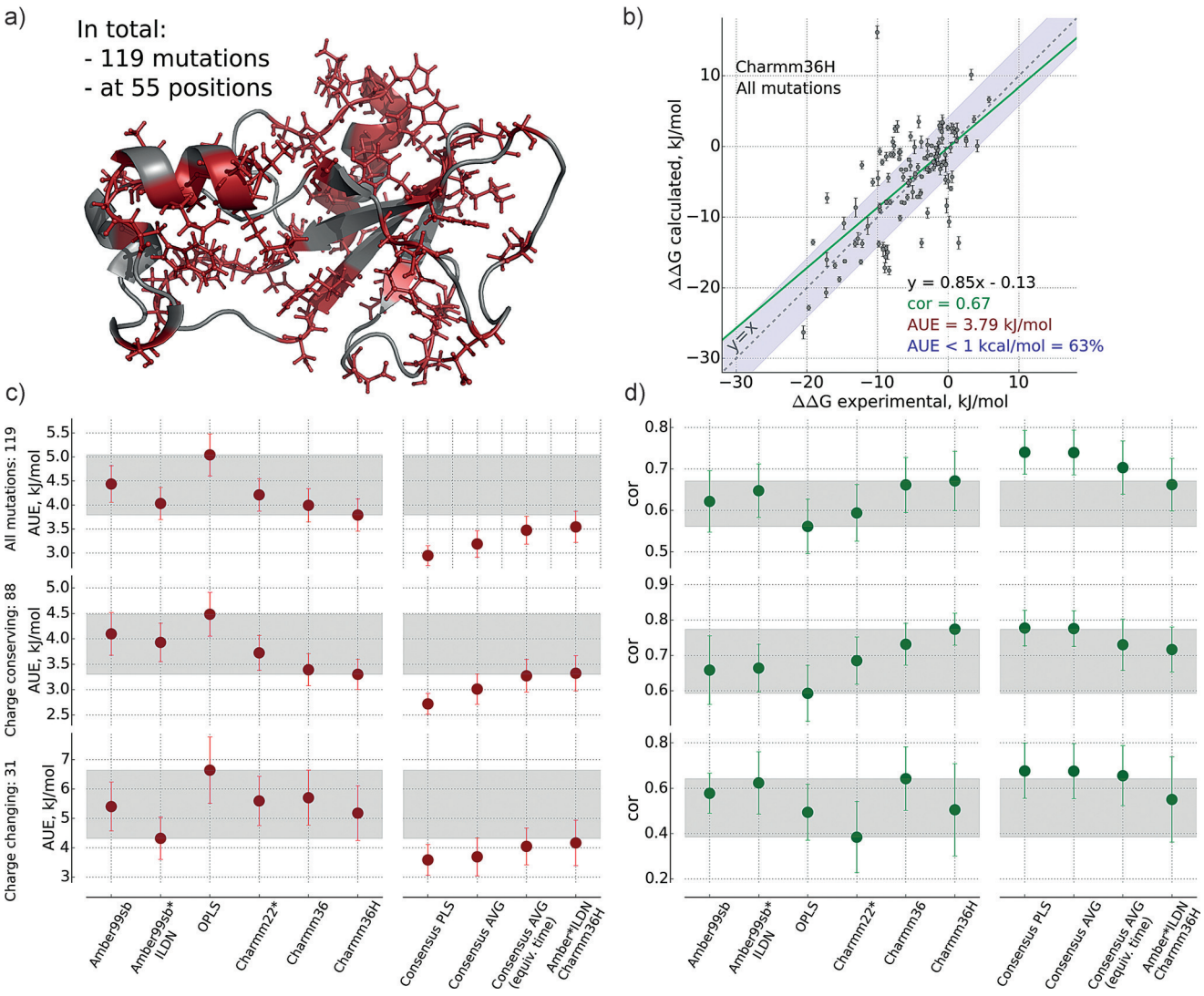

Figure 1. Thermodynamic force-field benchmark. a) Structure of barnase with the mutated residues marked in red (1BNI).[14] b) Experimentally measured double free-energy differences plotted against the $\Delta\Delta G$ values calculated with the Charmm36H force field. The shaded area marks the ± 1 kcal mol^{-1} region around the $\Delta\Delta G_{calc} = \Delta\Delta G_{exp}$ line. c, d) Comparison of the calculated and experimental thermostabilities in terms of the average unsigned error and the correlation, respectively. The comparison considers all mutations (top row), charge-conserving mutations (middle row), and charge-changing mutations (bottom row). The gray areas in (c) and (d) mark the ranges of the AUE and correlation, respectively, between the best and worst performing force fields.

for the charge-conserving mutations proved to be an easier task: The best results achieved an AUE of 3.3 kJ mol^{-1} and a correlation of 0.77.

Charge-changing mutations, however, were more challenging (Figure 1 c, d, bottom row). The best simulation, Amber99sb*ILDN, on average deviates from experiment by 4.32 kJ mol^{-1} with a correlation of 0.62. For an independent reference to the alchemical free-energy calculations, we estimated the relative thermostability changes by means of specialized Rosetta protocols.[8] For the charge-conserving mutations, Rosetta performed comparably to the alchemical calculations, yielding correlations ranging from 0.36 to 0.71 depending on the protocol and the crystal structure used. However, the statistical Rosetta potential was not able to capture the trends in the charge-changing mutations (correlations from −0.04 to −0.26; see Table S1 in the Supporting Information).

Having obtained free-energy estimates for a large number of force fields, we were able to construct a set of consensus $\Delta\Delta G$ values with the aim to minimize the force-field bias. A partial least squares (PLS) regression based consensus model (see the Supporting Information) presents a significant improvement over any single force field taken separately (Figure 1 c, d) with an overall AUE of 2.94 kJ mol^{-1} and a correlation of 0.74. Interestingly, further investigations revealed that a simple averaging of the $\Delta\Delta G$ values obtained from the simulations with different force fields yielded a comparable result to the PLS model (with an AUE of 3.19 kJ mol^{-1} and a correlation of 0.74). However, by training a regression model or averaging over the calculated values, we not only addressed the force-field-related artefacts, but also implicitly combined the sampling times of individual trajectories. Thus, the sampling-related error was effectively attenuated as well. Another consensus model was therefore

constructed by averaging over the thermostability estimates from different force fields and considering a fraction of the sampling time from each trajectory such that the combined simulation time was equivalent to the trajectory of a single force field. This approach still outperformed estimates by single force fields. In fact, a consensus of only two estimates obtained by the Amber and Charmm family force fields overall performed better than the best single force field in terms of the AUE (Figure 1c and Figure S12). This observation also holds deeper implications regarding the force-field evolution. It appears that in spite of the continuous force-field development and fine-tuning of individual parameters, the overall free-energy gradients generated by different force fields do not necessarily point in the same direction. An investigation of the discrepancies between the force fields may therefore provide a novel path for an improved amino acid parameterization (Figures S15–S17).

The consensus approaches provide a way to decrease both the force-field- and the sampling-related artefacts that may affect the free-energy estimates. We quantified the magnitude of both effects. The diagonal elements of the matrix (Figure 2a) denote intra-force-field errors and correlations, hence reporting solely on the sampling-related inaccuracies: The obtained effective sampling error for a single trajectory is approximately 1.5 kJ mol^{-1} (cor > 0.8). The off-diagonal elements in the matrix combine the insufficient sampling artefacts and force-field errors: The maximal AUE for any two force fields with the sampling error discarded reaches approximately 2 kJ mol^{-1}.

Another source of error contributing to the disagreement in Figure 1c,d comes from the experimental side. For an experimental error estimate, we parsed the thermodynamic database ProTherm,[9] extracting those free-energy values for any protein that have been reported multiple times (Figure 2b). A remarkable deviation from perfect correlation (cor = 0.71) and a substantial AUE (3.38 kJ mol^{-1}) were observed. In part, this can be explained by inaccuracies in database annotation: An analysis of some of the outliers in Figure 2b identified a number of mismatches with respect to the original publications. Hence, the experimental error has a dual character in this case: It comprises the actual disagreement between the experiments and the potential error from the subsequent processing. A similar pairwise comparison of experimental values for the mutations restricted to barnase shows a better, though not an ideal agreement between multiple experimental observations with a correlation of 0.93 and an AUE of 1.29 kJ mol^{-1} (Figure 2c). Thus the experimental error ranges from 1 to 3 kJ mol^{-1}, which is of comparable magnitude to the computational error.

The generality of the findings from the large mutation scan in barnase was assessed by application to other proteins. A set of 24 charge-conserving mutations with experimentally measured $\Delta\Delta G$ values in staphylococcal nuclease was collected (Figure 3a). The thermostabilities were calculated using the Amber99sb*ILDN and Charmm36H force fields. The achieved accuracy was similar to that of the barnase mutations (Figure 3c). For the analyzed mutations in staphylococcal nuclease, Amber99sb*ILDN performed better than Charmm36H in terms of the AUE (3.52 kJ mol^{-1} and 4.82 kJ mol^{-1}, respectively). However, the AUE appears to be biased by a few mutations where Charmm36H predicted a significantly stronger destabilizing effect than measured experimentally (Figure S14). For both force fields, correlations of the results with the experimental $\Delta\Delta G$ values show a similar level of accuracy. An average consensus model constructed similarly to the one for the barnase mutations yielded a more accurate prediction than either of the force fields taken separately (Figure 3c).

MD-based free-energy calculations are not limited to soluble proteins. Therefore, we set out to investigate whether

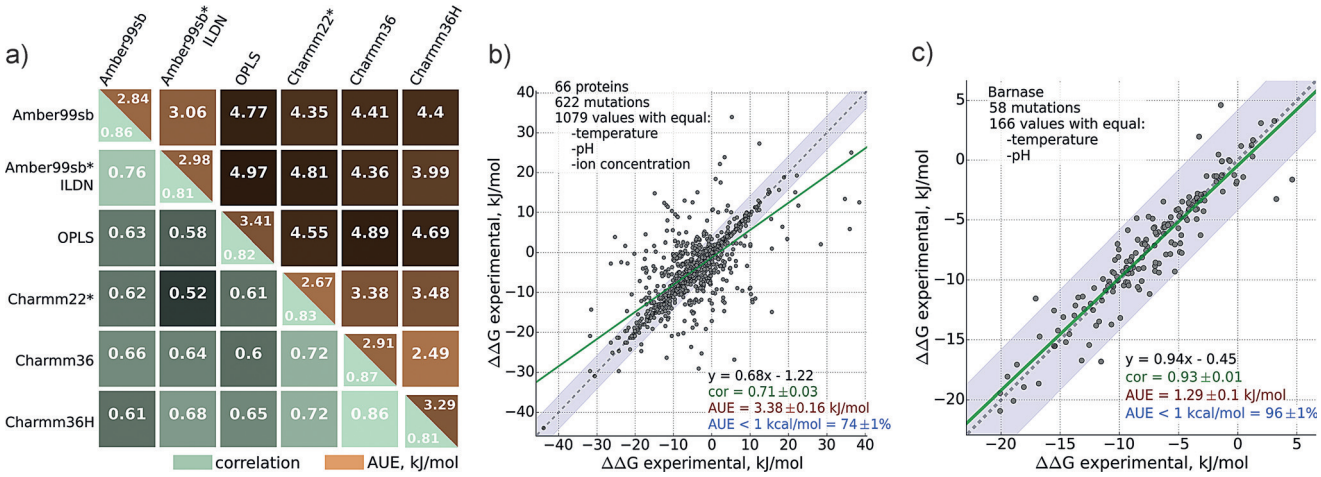

Figure 2. Error assessment and quantification. a) Matrix of the AUE values (top right) and correlations (bottom left) between the $\Delta\Delta G$ values for different force fields. The diagonal elements compare intra-force-field calculations, hence reporting solely on the sampling error. The off-diagonal elements combine the sampling and force-field errors. b) Experimentally measured double free-energy differences for mutations with multiple entries in 66 proteins extracted from the ProTherm database.[9] Experiments reporting on the same mutation were performed at identical temperatures, pH, and ion concentration (if reported in ProTherm). c) Experimentally measured $\Delta\Delta G$ values for barnase mutations with multiple entries extracted from the ProTherm database. Experiments reporting on the same mutation were performed at identical temperatures and pH.

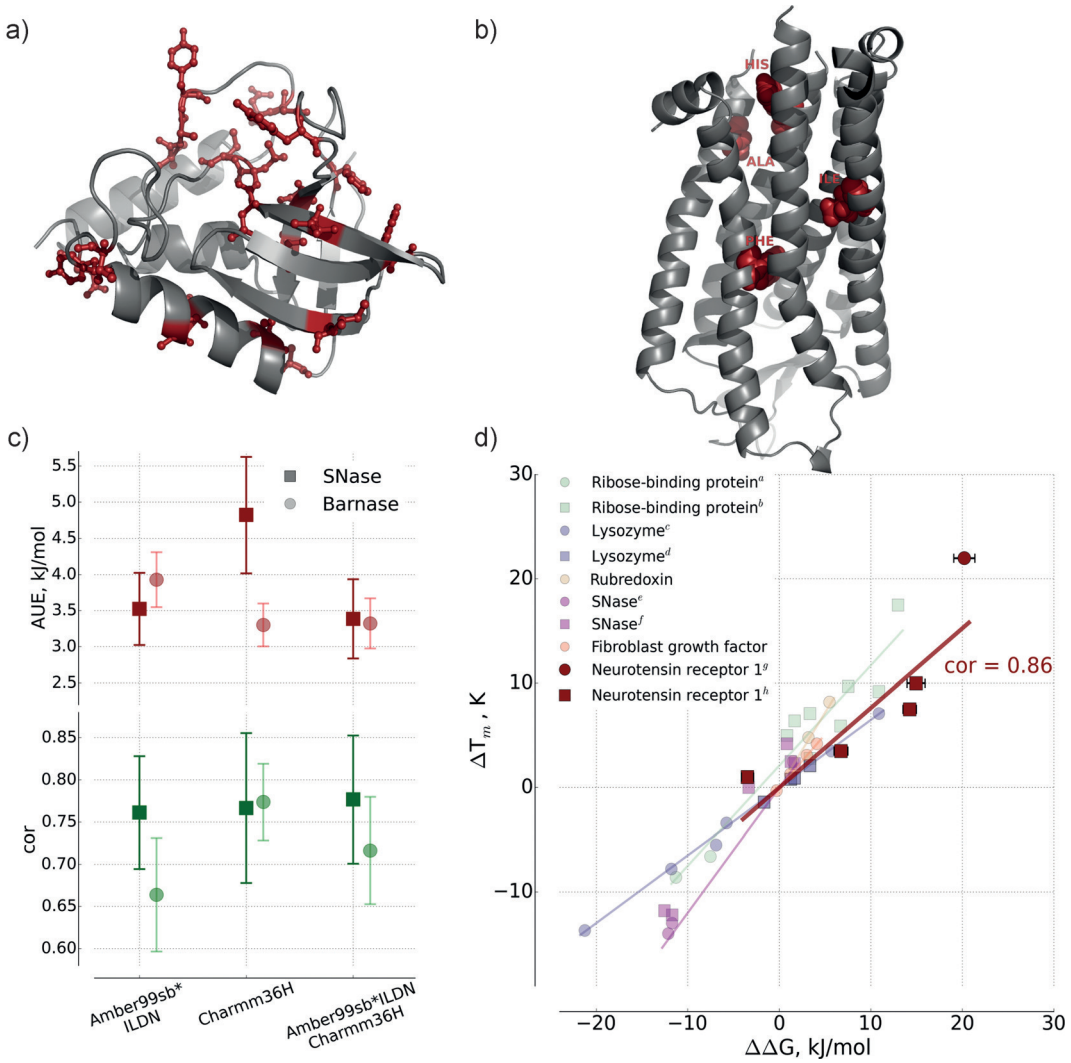

Figure 3. Thermostability estimates for staphylococcal nuclease and neurotensin receptor 1. a) Structure of staphylococcal nuclease with the mutated residues shown in red (1STN).[15] b) Structure of NTR1 with the mutated residues shown in red (4BUO).[13] c) AUE values (top row) and correlations (bottom row) for the calculated $\Delta\Delta G$ values and experimental measurements for staphylococcal nuclease (squares). Results for barnase are shown in circles for comparison. d) Changes in the experimentally measured melting temperatures upon amino acid mutations plotted against the double free-energy differences, also measured experimentally, for five proteins (partially transparent symbols). The $\Delta\Delta G$ values for NTR1 were calculated using the alchemical approach. Literature sources for the experimental ΔT_m and $\Delta\Delta G$ values are provided in the Supporting Information.

the alchemical free-energy calculations would be able to capture the effects of mutations in a membrane protein. The G-protein-coupled neurotensin 1 receptor (NTR1) has been the subject of a number of studies aiming at identifying thermally stabilizing mutations.[10,11] Discovering such mutations culminated in the successful crystallization of NTR1 and the subsequent resolution of its structure.[12,13] We selected five mutants that are known to achieve higher NTR1 stability (Table S4) to be evaluated. For the NTR1 case, a direct comparison of the computed free-energy differences to the experimental measurements was not possible as the experiments only reported on changes in the protein melting temperature (ΔT_m).[11,13] Therefore, we used experimental data to estimate the correlation that ought to be expected between the $\Delta\Delta G$ and ΔT_m values for a number of proteins (Figure 3d and Table S5). Subsequently, the experimentally measured melting temperatures for NTR1 were compared with the calculated $\Delta\Delta G$ values. A remarkable agreement with a correlation of 0.86 was observed (Figure 3d).

The presented approach is not limited to thermostability calculations and can be applied to estimate free-energy changes in protein–protein binding upon amino acid mutation. This challenge proved to be more complex than a thermostability calculation as accurate free-energy estimates are required for the residues exposed to the solvent (for unbound proteins) and the buried amino acids (for proteins forming a complex). We calculated the relative free-energy differences in protein binding for 12 mutations in the turkey ovomucoid third domain complexed with α-chymotrypsin, and 13 mutations in an antibody HyHEL-10 Fv with hen egg lysozyme. The free-energy estimates (summarized in Table S6) demonstrate that the alchemical approaches per-

form on par with or better than the empirical methods in the relative free-energy calculations of several protein complexes.

To summarize, alchemical free-energy calculations are readily available for large-scale protein thermostability estimates. A consensus approach based on predictions in multiple force fields has been introduced. The approach increases the prediction accuracy, bringing the prediction well within the accuracy range of experimental measurements. Such predictions are expected to aid protein engineering and design as well as protein stabilization for crystallization, for example.

Acknowledgements

We are grateful to Elizabeth H. Kellogg and Colin Smith for sharing their experience with the Rosetta software, Rainer Böckmann for constructive comments and suggestions, and Rodolfo Briones for the help with the table of contents figure. We thank the Leibniz Computing Centre and the Flemish Supercomputer Center for computational time. V.G. acknowledges financial support from Boehringer Ingelheim Pharma GmbH.

Keywords: force field · free-energy calculations · proteins · thermostability

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 7364 – 7368
Angew. Chem. **2016**, *128*, 7490 – 7494

[1] M. J. Liszka, M. E. Clark, E. Schneider, D. S. Clark, *Annu. Rev. Chem. Biomol. Eng.* **2012**, *3*, 77 – 102.
[2] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3*, 300 – 313.
[3] R. W. Zwanzig, *J. Chem. Phys.* **1954**, *22*, 1420 – 1426.
[4] G. E. Crooks, *Phys. Rev. E* **1999**, *60*, 2721 – 2726.
[5] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O'Dror, D. E. Shaw, *PloS ONE* **2012**, *7*, e32131.
[6] K. A. Beauchamp, Y.-S. Lin, R. Das, V. S. Pande, *J. Chem. Theory Comput.* **2012**, *8*, 1409 – 1414.
[7] V. Gapsys, S. Michielssens, D. Seeliger, B. L. de Groot, *J. Comput. Chem.* **2015**, *36*, 348 – 354.
[8] E. H. Kellogg, A. Leaver-Fay, D. Baker, *Proteins Struct. Funct. Bioinf.* **2011**, *79*, 830 – 838.
[9] M. D. Shaji Kumar, K. Abdulla Bava, M. Michael Gromiha, P. Prabakaran, K. Kitajima, H. Uedaira, A. Sarai, *Nucleic Acids Res.* **2006**, *34*, D204 – D206.
[10] C. A. Sarkar, I. Dodevski, M. Kenig, S. Dudli, A. Mohr, E. Hermans, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 14808 – 14813.
[11] Y. Shibata, J. F. White, M. J. Serrano-Vega, F. Magnani, A. L. Aloia, R. Grisshammer, C. G. Tate, *J. Mol. Biol.* **2009**, *390*, 262 – 277.
[12] J. F. White, N. Noinaj, Y. Shibata, J. Love, B. Kloss, F. Xu, J. Gvozdenovic-Jeremic, P. Shah, J. Shiloach, C. G. Tate, R. Grisshammer, *Nature* **2012**, *490*, 508 – 513.
[13] P. Egloff, M. Hillenbrand, C. Klenk, A. Batyuk, P. Heine, S. Balada, K. M. Schlinkmann, D. J. Scott, M. Schütz, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E655 – E662.
[14] A. M. Buckle, K. Henrick, A. R. Fersht, *J. Mol. Biol.* **1993**, *234*, 847 – 860.
[15] T. R. Hynes, R. O. Fox, *Proteins Struct. Funct. Genet.* **1991**, *10*, 92 – 105.

Received: October 30, 2015
Revised: January 19, 2016
Published online: April 28, 2016

