

# Inferring Structural Ensembles from Noisy Experiments: Application to Trialanine

Kyle A. Beauchamp,<sup>†</sup> Rhiju Das<sup>†,‡</sup> and Vijay S. Pande<sup>†\*,¶</sup>

*Biophysics Program, Biochemistry Department, Stanford University, Stanford, CA, and Chemistry  
Department, Stanford University, Stanford, CA*

E-mail: rhiju@stanford.edu, pande@stanford.edu

## Abstract

Inferring biomolecular conformation from experiment is the fundamental task of structural biology. Due to limited experimental resolution, structure determination often requires a combination of modeling and experiment. The vast majority of algorithms for structure determination, however, are limited to modeling a single conformation and provide only limited uncertainty information. Here we describe Bayesian Energy Landscape Tilting (BELT), a scheme that combines simulation and experiment to infer solution ensembles. BELT leverages the machinery of Bayesian statistics to give rigorous uncertainty estimates for structural features and equilibrium properties, given simple, well-defined assumptions. Using trialanine as a test system, we show that BELT corrects force field bias, as measured by  $\chi^2$  agreement with trialanine measurements *not used* to fit the models. BELT modeling with just six NMR measurements estimates  $\alpha$ ,  $\beta$ , and  $PP_{II}$  conformational populations that are independent of force field.

---

\*To whom correspondence should be addressed

<sup>†</sup>Biophysics Program

<sup>‡</sup>Biochemistry Department

<sup>¶</sup>Chemistry Department

# Introduction

Over the past forty years, structural biologists have solved “ground-state” structures of countless biological macromolecules.<sup>1</sup> Modern biology, however, presents many systems that do not fit the single-structure paradigm. Excited states of nucleic acids,<sup>2</sup> natively disordered proteins,<sup>3</sup> and protein folding intermediates<sup>4</sup> alike are poorly described by single conformation models. For such systems, models of conformational ensembles are required to understand and predict structural and equilibrium properties.

A growing body of research has sought to understand conformational ensembles. Much of it has focused on incorporating dynamical information during NMR structure determination<sup>5,6</sup> or the extraction of multiple conformers from X-ray diffraction data.<sup>7,8</sup> While these techniques are powerful, they share difficulties in data collection and the unified treatment of heterogeneous experimental data. At the same time, molecular dynamics studies have demonstrated heterogeneity in the atomic-scale structure of macromolecules.<sup>9</sup> However, simulation studies suffer from possible inaccuracies in molecular dynamics force fields, as revealed by benchmarks and optimizations of molecular dynamics force fields.<sup>10–12</sup> In the present work, we connect these two extremes by combining equilibrium experimental measurements with molecular simulation.

Here we introduce a statistical approach to modeling solution ensembles of biological macromolecules. The algorithm, Bayesian Energy Landscape Tilting (BELT), uses solution experiments to reweight a collection of atomistic models. BELT extends a recent maximum entropy method for restraining simulations<sup>13</sup> to *reweight* existing simulations. Furthermore, BELT leverages Markov Chain Monte Carlo to transform experimental ambiguity into error bars on arbitrary structural features.

To validate BELT, we investigate the conformational propensities of trialanine using NMR measurements<sup>14</sup> and MD simulations performed in five different force fields. Although the raw simulations show wide variations in their conformational preferences, BELT corrects force field errors to provide self-consistent estimates of the  $\alpha$ ,  $\beta$ , and  $PP_{II}$  populations. The ability to correct the biases of all tested forcefields suggests that BELT is a powerful technique for connecting

simulation and equilibrium measurements.

## Theory: Bayesian Energy Landscape Tilting

### Model Inputs

To model an ensemble using BELT requires three components (Fig. 1). First, we need conformations  $(x_j)_{j=1}^m$  sampled from the approximate equilibrium distribution of our system. In the present work, such conformations will be generated from molecular dynamics simulations. Second, we require equilibrium experimental measurements  $(F_i)_{i=1}^n$  and their associated uncertainties  $(\sigma_i)_{i=1}^n$ . Third, it is necessary to have a direct connection between simulation and experiment. This connection is achieved by predicting each experimental observable at each conformation:  $f_i(x_j)$  is the predicted value of experiment  $i$  at conformation  $x_j$ .

### Reweighting

The next step in constructing an ensemble is to calculate the population of each conformation. For convenience, we work with the log populations (i.e. free energies). Inspired by a previous method for restraining simulations<sup>13</sup> (see Appx. S1), we reweight individual conformations by a biasing potential that is a linear combination of the predicted observables:

$$\Delta U(x; \alpha) = \sum_i^n \alpha_i f_i(x)$$

In  $\Delta U(x; \alpha)$ , the parameters  $\alpha_i$  determine how strongly each experiment contributes to the biasing potential. One way to think about  $\alpha$  is via “tilting” the energy landscape along the order parameters  $f_i(x)$ . Given the biasing potential, the population of each conformation can be calculated using exponential averaging (see Appx. S2):

$$\pi_j(\alpha) = \frac{1}{\sum_k \exp[-\Delta U(x_k; \alpha)]} \exp[-\Delta U(x_j; \alpha)]$$

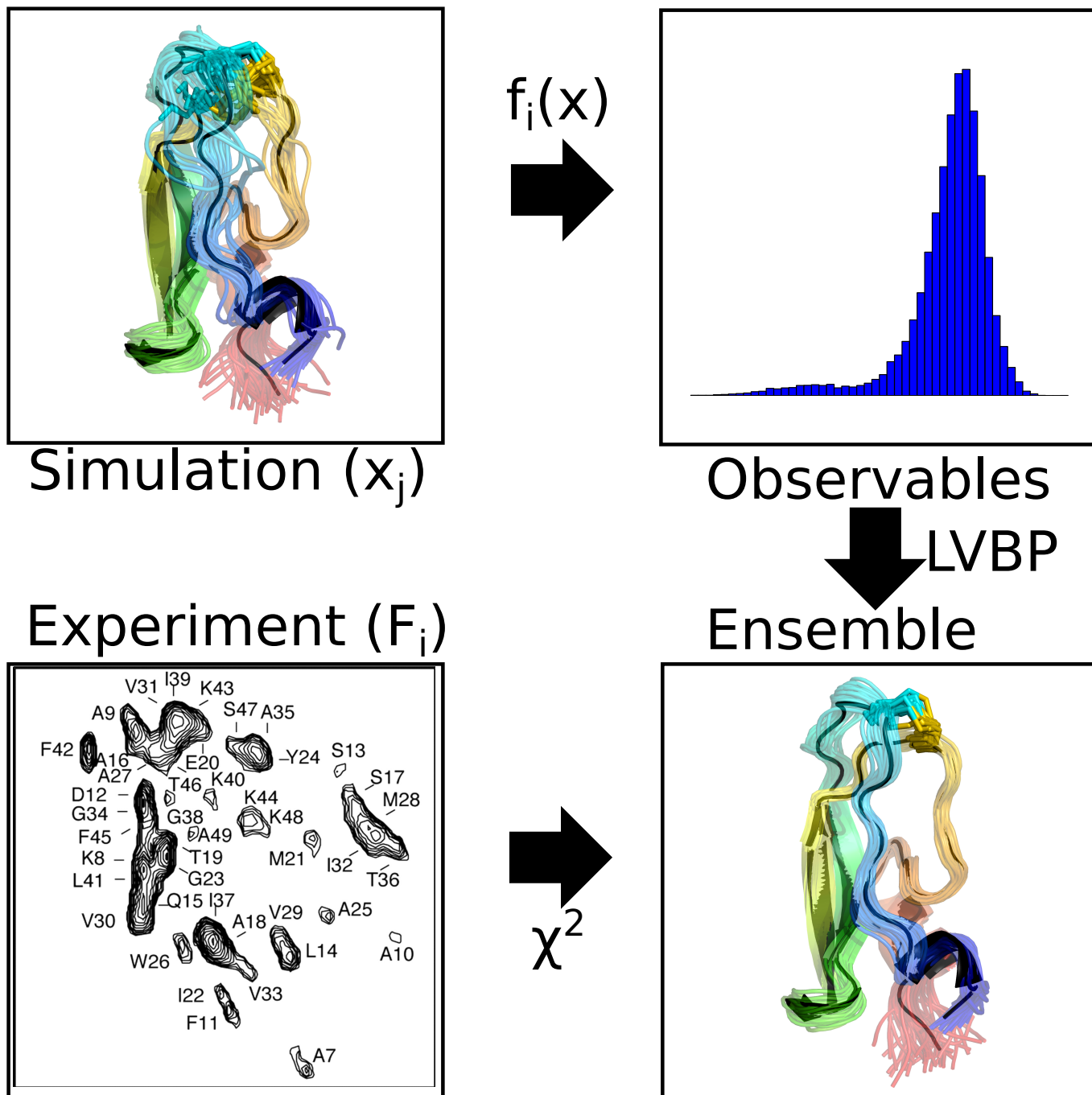


Figure 1: General scheme for BELT modeling.

It is informative to consider the case of a single observable  $f(x)$ . Suppose the molecule of interest shows a bimodal observable. If we let  $\alpha = 0$ , then the biasing potential is 0 everywhere and our reweighted ensemble simply returns the results of the MD simulation (Fig. 2b). If we let  $\alpha = -1$ , conformations with large values of  $f(x)$  are upweighted, while conformations with lower values of  $f(x)$  are downweighted (Fig. 2a). Finally, if  $\alpha = 1$ , the ensemble shifts in the opposite direction (Fig. 2c).

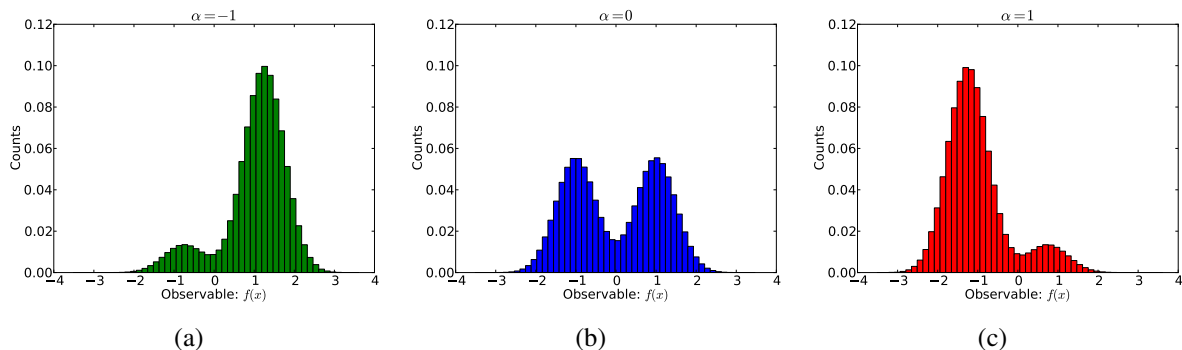


Figure 2: Raw ( $\alpha = 0$ ) and reweighted (e.g. tilted) histograms of a one dimensional observable.

With the equilibrium populations, we can calculate the equilibrium expectations of an arbitrary observable  $h(x)$ :

$$\langle h(x) \rangle_\alpha = \sum_j h(x_j) \pi_j(\alpha)$$

In the above bracket notation,  $\langle h(x) \rangle_\alpha$  is the ensemble average of  $h(x)$  in an ensemble that is perturbed by a biasing potential  $\Delta U(x; \alpha)$ . At this point, the determination of  $\alpha$  has not yet been discussed. The key idea, however, is that the  $\alpha$  reweighted ensemble  $\langle \rangle_\alpha$  should recapitulate the experimental measurements:

$$\langle f_i(x) \rangle_\alpha \approx F_i$$

## A likelihood framework

We now derive a likelihood framework for determining the coefficients  $\alpha$  used in the biasing potential. Given adequate sampling and self-consistent experiments, there should exist some *true* value of  $\alpha$  whose ensemble matches the experimental data. However, experimental uncertainty ( $\sigma_i$ ) prevents exact agreement between the measurements and the true ensemble. For the models in the current work, we model  $\sigma_i$  as the uncertainty associated with predicting chemical shifts and scalar couplings from structures; this dominant error is quantified by the RMS uncertainty estimated during the parameterization of chemical shift and scalar coupling models. We use an independent normal approximation (see Appx. S3) to model the agreement between the  $\alpha$  ensemble and the measurements:

$$P(F_i|\alpha) \sim N(\langle f_i(x) \rangle_\alpha, \sigma_i^2)$$

Using Bayes Theorem, we can calculate the posterior distribution of  $\alpha$ :

$$P(\alpha|F_1, \dots, F_n) \propto P(F_1, \dots, F_n|\alpha)P(\alpha)$$

Now we let  $LP(\alpha)$  denote the log posterior of  $\alpha$  and simplify, dropping terms that are independent of  $\alpha$ :

$$LP(\alpha) = \log[P(\alpha|F_1, \dots, F_n)] = -\sum_i^n \frac{1}{2\sigma_i^2} (\langle f_i(x) \rangle_\alpha - F_i)^2 + \log P(\alpha)$$

Note the simple form of the log posterior. The first term (i.e. the likelihood) measures the  $\chi^2$  agreement between the reweighted ensemble and measurements. The second term is the log of the prior distribution on  $\alpha$ . In the present work, we evaluate two different choices of prior, finding similar results for each. The first is the maximum entropy prior, which penalizes ensembles as they deviate from the raw simulation results:

$$\log P_{\maxent}(\alpha) = \lambda \sum_j^m \pi_j(\alpha) \log \pi_j(\alpha)$$

The second prior we consider is a multivariate normal prior:  $P_{MVN}(\alpha) \sim N(0, \Sigma)$ . The value of  $\Sigma$  is given by  $\Sigma_{ij} = \lambda \text{Cov}(f_i(x), f_j(x))$ , as derived in Appx. S4.

Both priors can be used to achieve regularization, which is a powerful technique to reduce overfitting. With large values of  $\lambda$ , both priors favor the raw simulation results (i.e. uniform conformational populations):  $\pi_i \approx \frac{1}{n}$ . The value of  $\lambda$  can be chosen via cross-validation or other methods (see Appx. S5).

## Bayesian Modeling of Structural Ensembles

Because ensemble inference often presents many plausible solutions,<sup>15,16</sup> we avoid statistical methods that return a single solution (e.g. maximum likelihood). We therefore use Markov chain Monte Carlo (MCMC), as implemented in PyMC,<sup>17</sup> to sample the distribution of structural ensembles consistent with experiment. The result is an ensemble of ensembles—a statistical ensemble of conformational ensembles. Averaging all MCMC samples provides posterior mean estimates of arbitrary structural features. Similarly, examining the MCMC variances provides statistical uncertainties of equilibrium or structural features. A Bayesian bootstrapping procedure<sup>18</sup> can also be used to model the statistical uncertainty of the MD simulations (see Appx. S6).

## Results

### Conformational Propensities of Trialanine

Short peptides provide crucial tests for evaluating and optimizing molecular dynamics force fields.<sup>14,19–22</sup> Such peptides offer a window into the intrinsic conformational propensities of amino acids, free from the secondary structure bias found in statistical surveys of protein structures.<sup>23</sup> Here, we use BELT to infer the conformational populations of trialanine from chemical shift and scalar coupling

measurements.<sup>14</sup>

Trialanine was simulated (see Methods) in five different force fields; six experimental measurements (three scalar couplings, three chemical shifts) probing the central alanine residue were used to construct a BELT ensemble. The five force fields show considerable variation in their agreement with experiment (Fig. 3). The amber96, amber99, charmm27, and oplsaa force fields, for example, initially show significant deviation from the experimental measurements. Upon reweighting, however, all five force fields agree with experiment—including experiments that were *not* used to fit the model (Fig. 3b). Additionally, these results are robust to differences in the prior placed on  $\alpha$ .

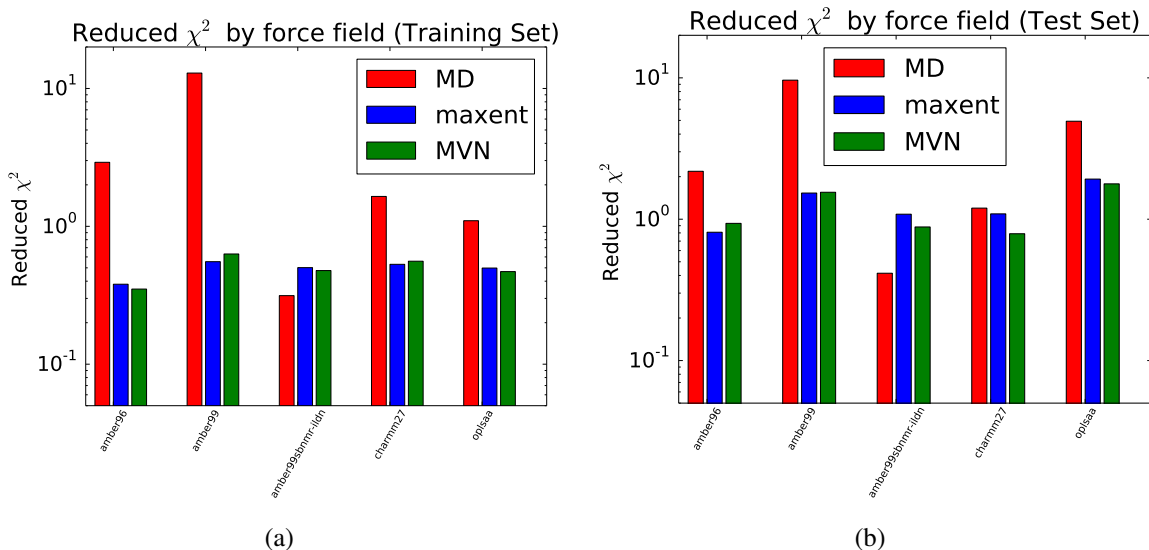
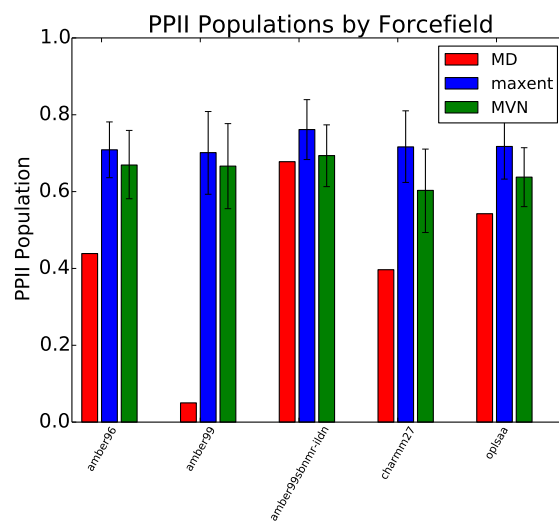


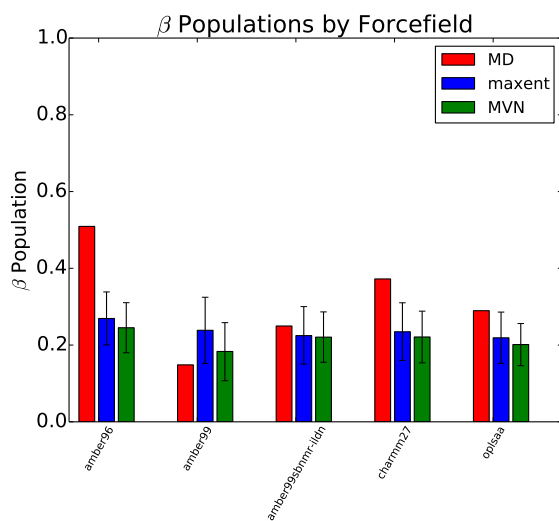
Figure 3: The reduced  $\chi^2$  error (e.g.  $\frac{\chi^2}{n}$ ) for MD and BELT (maxent and MVN priors) models. The BELT reduced  $\chi^2$  is estimated as the mean reduced  $\chi^2$  over all MCMC samples. (a). Calculated using the six measurements used to fit the BELT model. (b). Calculated using four measurements *not* used to fit the BELT model. See Methods for the definition of training and test sets.

Several independent experimental studies have shown that short alanine peptides prefer the polyproline type helix (PPII) in solution.<sup>14,22,24</sup> Most molecular dynamics force fields, however, are known to underpopulate the PPII state.<sup>14,19–21</sup> Our trialanine simulations recapitulate this known deficiency (Fig. 4; red), with amber96 showing a strong  $\beta$  bias (Fig. 4b) and amber99 showing a strong  $\alpha$  bias (Fig. 4c). However, combining simulation and experiment leads to con-

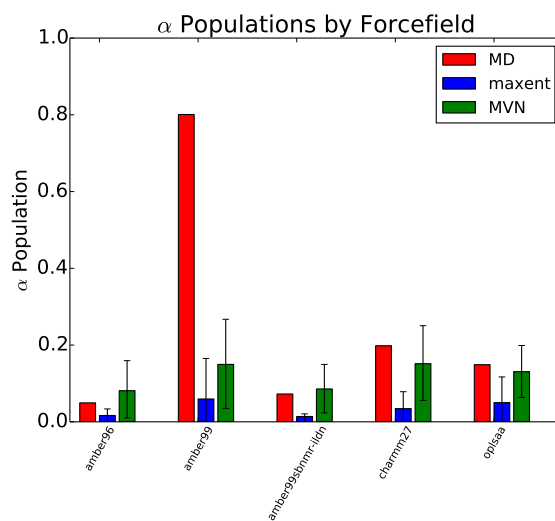




(a)



(b)



(c)

Figure 4: MD and BELT (maxent and MVN priors) conformational propensities for each force field.

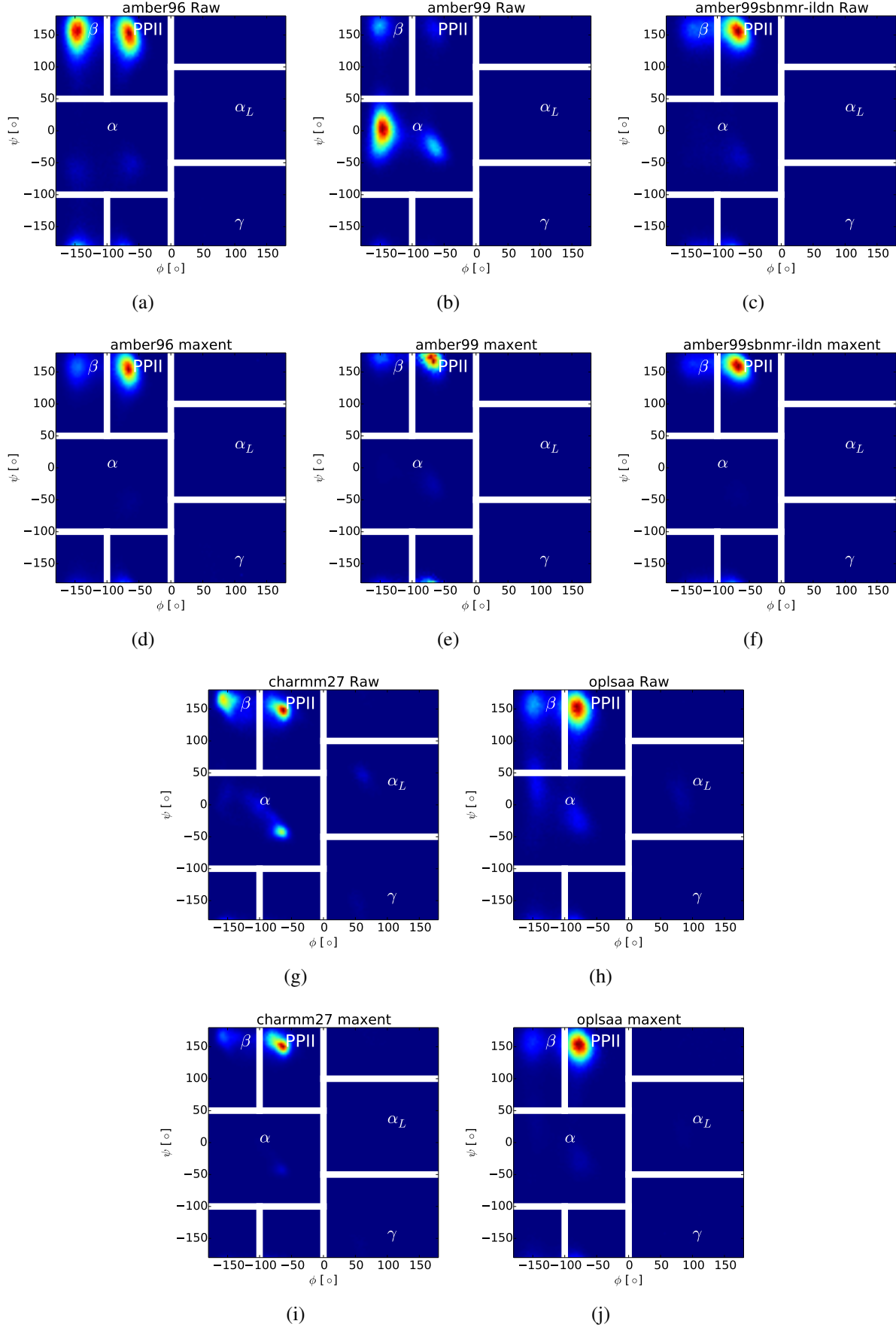


Figure 5: Ramachandran plot of MD and BELT (maxent prior) ensembles for each force field. Comparison of BELT models with different priors (maxent and MVN) are given in Fig. S1.

formational ensembles that are robust to differences in force field and prior (Fig. 4; green and blue). Full Ramachandran plots of the  $(\phi, \psi)$  preferences (Fig. 5) reveal similar features in all five BELT models.

## Bayesian Inference allows Ambiguous Experiments

BELT allows the full characterization of posterior distributions using MCMC. Most previous approaches, however, have focused on obtaining point estimates of conformational ensembles.<sup>14,25</sup> In many cases, however, ambiguous experimental data precludes a point-estimate of the conformational ensemble. To illustrate this point, we plot the measured<sup>14</sup> value of  ${}^3J(H_N H_\alpha)$  in the context of the Karplus equation relating  $\phi$  to  ${}^3J(H_N H_\alpha)$ . The measured coupling corresponds to four *different* values of  $\phi$  (Fig. 6a), suggesting that a point estimate is inappropriate for modeling conformations or ensembles. As a concrete example of bimodality, we calculated the posterior distribution of  $PP_{II}$  for the oplsa force field reweighted with two NMR measurements (Fig. 6b). The observed bimodal distribution indicates the need to fully characterize the posterior distribution.

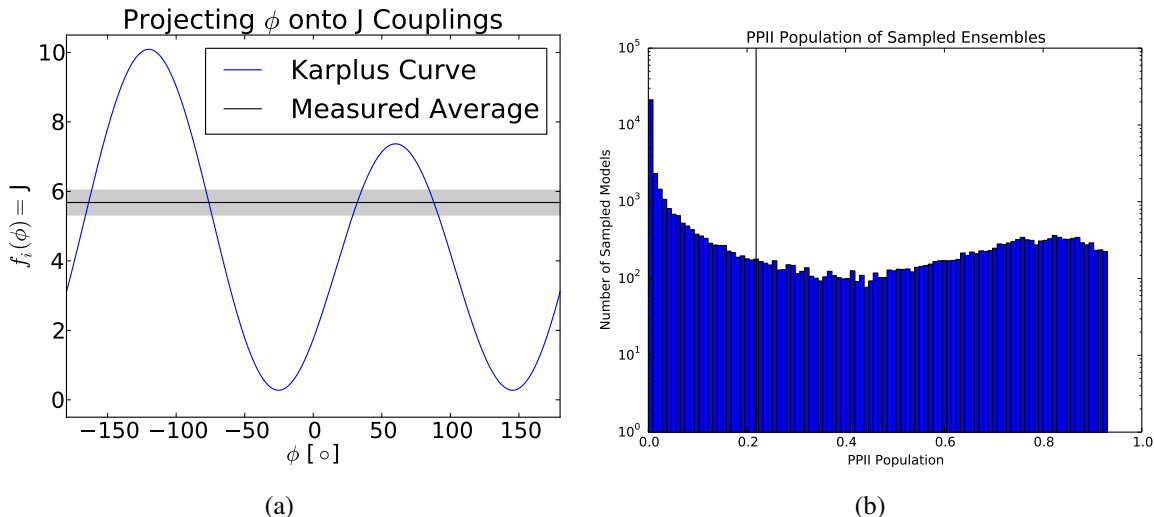


Figure 6: (a). The Karplus equation connecting the backbone torsion  $\phi$  to  ${}^3J(H_N H_\alpha)$  is ambiguous; observed values of  ${}^3J(H_N H_\alpha)$  are consistent multiple conformations. (b). BELT modeling (oplsaa force field) with two measurements ( ${}^3J(H_N C\beta)$ ,  $C\beta$  chemical shift) shows a bimodal posterior.

## Discussion

### Structural (Ensemble?) Biology

Why model structural ensembles, rather than just structures? At least three compelling reasons favor ensembles. First, biological molecules are multi-state machines that fold, unfold, bind ligands, aggregate, and change conformation. Biology is controlled by the relative populations of these states. Ensembles capture aspects of these phenomena by encoding equilibrium populations with structures. A second argument for ensemble modeling is fidelity to experiment. Most solution experiments measure ensemble average equilibrium properties—chemical shifts, scalar couplings, NOEs, SAXS, and FRET are reasonably well-described as equilibrium properties. A truly quantitative connection to these measurements requires modeling the equilibrium ensemble. Finally, recent advances in atomistic simulation,<sup>26–29</sup> special-purpose hardware,<sup>30</sup> and distributed computing analysis<sup>31,32</sup> have enabled atomistic simulations to reach the millisecond timescale;<sup>9,33–35</sup> the computational cost of ensemble modeling is quickly becoming manageable.

One might argue that structural ensembles are unnecessary because many proteins occupy a single state under physiological conditions. For such proteins, it is probably safe to enforce single state behavior, as is done in current modeling approaches. However, we suggest that the number of states be *inferred*—not *assumed*.

### Resolving Fine Structure may Require Improved Observables

The five BELT ensembles show quantitative agreement in  $\alpha$ ,  $\beta$ , and  $PP_{II}$  populations (Fig. 4). However, the finer details of each Ramachandran plot (Fig. 5) suggest subtle differences between the five models. Because all five BELT ensembles show excellent agreement with experiment (Fig. 3), we conclude that current predictors of chemical shifts and scalar couplings are insufficiently precise to resolve (and falsify) subtle force field differences. The most obvious such difference is the width, shape, and orientation of the  $PP_{II}$  basin. Most strikingly, amber96 and oplsa have  $PP_{II}$  basins that are vertically oriented, while amber99, amber99sbnmr-ildn, and charmm27 show

diagonally oriented  $PP_{II}$  basins. This highlights the need for more sensitive connections between simulation and experiment, such as chemical shift and scalar coupling models.

## Comparison to Previous Ensemble Methods

BELT offers several key advantages over previous ensemble estimation techniques. Most previous ensemble modeling efforts involve a protocol with three key ingredients: clustering, a  $\chi^2$  objective function, and population inference on the clusters. For example, this recipe describes the approach used in previous analyses of homopeptides,<sup>14</sup> the EROS technique for SAXS modeling,<sup>25</sup> and the Bayesian Weighting formalism.<sup>15</sup> In our opinion, the primary disadvantage of these techniques is the need for a clustering step. Clustering can introduce two different errors, depending on whether the number of clusters is small or large. In the limit of few states, clustering can overly coarsen the system of interest, possibly preventing the model from reproducing multiple experimental observables. At the other extreme, too many states could lead to over-fitting when estimating the many populations required to parameterize the model. This will lead to poor generalization performance and large errors when predicting experiments *not* used to train the model.

In contrast, BELT avoids clustering by projecting simulations onto a basis defined by the predicted experimental observables. The advantage of working in this basis are fourfold. First, in BELT, one estimates a single parameter ( $\alpha_i$ ) for each experimental observable. If the number of experiments is small, as is often the case, the inference problem involves only a few parameters. Second, the predicted observables are a *natural* basis for biophysical calculations. This basis allows direct connection to experiment and often provides insight into the molecular interactions driving biophysical phenomena. For example, the projection onto observables could be used to rationally infer force field parameters, similar to the ForceBalance method.<sup>36,37</sup> Third, suppose two conformations have the same predicted experimental observables. In BELT, these conformations will have the same population. In clustering based methods, the populations of these conformations could vary simply due to random fluctuations. Finally, in the limit of exact measurements, BELT reduces to a previous<sup>13</sup> maximum entropy approach (see Appx. S1).

We also point out some surprising differences between BELT and BW-like methods. BW-like methods have the interesting property that the in-state means of features are preserved. More precisely, suppose that  $\chi_s(x)$  is the indicator function of a conformational state  $s$ . Then in-state averages of the form  $\langle h(x)\chi_s(x) \rangle$  do *not* depend on the reweighted populations. BELT, however, does not preserve the in-state averages; in fact, this property is the direct result of BELT’s connection to maximum entropy modeling (see Appx. S1 and ref.<sup>13</sup>). The effect of this property is that the “peaks” of reweighted histograms are slightly shifted relative to the raw MD results, as observed in Fig. 2.

## Quantitative Comparison of Ensemble Inference Techniques

The above reasons provide theoretical arguments supporting the BELT method. However, it is possible to evaluate the method *quantitatively*. One hallmark of a good statistical model is its predictive ability on set-aside test data.<sup>38</sup> We therefore evaluated BW (see Methods) and BELT models on set-aside test data.

For all five force fields, BELT produces models that give statistical agreement with experiment on the test data. For three out of five force fields, BELT gives the *best* model. The BELT prior appears relatively unimportant. In cases where BELT outperforms BW, it does so significantly.

For the two remaining cases, BW outperforms BELT. For these force fields (amber99sbnmr-ildn and charmm27), however, all four models (BW, BELT-MVN, BELT-maxent), and MD) are consistent ( $\chi^2 \lesssim 1$ ) with experiment. We conclude that while both BW and BELT provide reasonable models, BELT gives better models in situations with poor starting force fields.

	amber96	amber99	amber99sbnmr-ildn	charmm27	oplsaa
Method (Prior)					
BW	0.67	1.73	<b>0.46</b>	<b>0.47</b>	2.75
BELT (MVN)	0.62	0.90	0.66	0.76	<b>0.91</b>
BELT (maxent)	<b>0.50</b>	<b>0.80</b>	0.61	0.76	1.11
MD	2.18	9.63	0.41	1.20	4.93

## Future Work

The BELT method can be extended in several ways. We have already worked out some of these extensions. In App. S3, we derive an approximate correction for working with dependent data. Another obvious extension is the use of non-Normal error models. These models can be directly inserted into the current framework by replacing the  $\chi^2$  term in the likelihood with some other loss function. More sophisticated models could separately treat the uncertainties associated with predicting observables and the uncertainties of conformations. This would replace the regularization and Bayesian Bootstrapping (App. S6) approaches used herein. Another promising avenue is to combine Bayesian modeling of structural ensembles with Bayesian models for experimental observables. A Bayesian formalism for NMR experiments has previously been developed;<sup>16,39</sup> connecting BELT to these methods may be straightforward.

## Conclusion

Bayesian Energy Landscape Tilting allows the simultaneous characterization of structural and equilibrium properties. Through its use of MCMC, BELT is robust to ambiguous experiments and provides rigorous uncertainty estimates. BELT models constructed with a handful of NMR measurements correct significant force field bias and provide generalizable, *force field independent* trialanine ensembles.

# Acknowledgements

We thank John Chodera, TJ Lane, Frank Cochran, Pehr Harbury, Xuesong Shi, and Dan Herschlag for helpful discussions.

# Methods

## Molecular Dynamics Simulations

Trialanine was simulated in the amber96,<sup>40</sup> amber99,<sup>41</sup> amber99sbnmr-ildn,<sup>42</sup> charmm27,<sup>43,44</sup> and oplsa<sup>45</sup> force fields, as previously reported.<sup>19</sup> Simulations were performed using Gromacs 4.5<sup>26</sup> and run at constant temperature (300 K) and pressure (1.01 atm). Each simulation was at least 225 ns long. Conformations were stored every 1 ps.

## Chemical Shift and Scalar Couplings

All NMR measurements in this work refer to experiments<sup>14</sup> probing the central residue of trialanine. We chose to focus on this residue because it will be most robust to pH dependent effects, which may be difficult to model with current force fields.

Chemical shifts (H, HA, CA, CB) for each frame were calculated using the average prediction of ShiftX2,<sup>46</sup> SPARTA+,<sup>47</sup> and PPM;<sup>48</sup> uncertainties for each model were estimated using their reported RMS prediction errors. Overall uncertainties were estimated as  $\sqrt{\sum w_i \sigma_i^2}$ , where  $w_i \propto \frac{1}{\sigma_i^2}$  is the weight (e.g  $\sum_i w_i = 1$ ) of each chemical shift model and  $\sigma_i$  is the uncertainty of each chemical shift model. The J couplings were calculated using the following Karplus relations:  $^3J(H^N C')$ ,<sup>49</sup>  $^3J(H^N H^\alpha)$ ,<sup>50</sup>  $^2J(NC^\alpha)$ ,<sup>14</sup>  $^3J(H^\alpha C')$ ,<sup>49</sup>  $^1J(NC^\alpha)$ ,<sup>14</sup>  $^3J(H^N C^\beta)$ .<sup>50</sup> J coupling uncertainties were approximated as the RMS errors reported when fitting the Karplus coefficients.

We have divided the available experimental measurements into a training and test set, with the training set consisting of  $^3J(H^N C')$ ,  $^2J(NC^\alpha)$ ,  $^3J(H^N C^\beta)$  scalar couplings and the  $C_\alpha$ , the  $H$ , and the  $C_\beta$  chemical shifts. The test set consists of  $^3J(H^N H^\alpha)$ ,  $^3J(H^\alpha C')$ , and  $^1J(NC^\alpha)$ . The division



into training and test sets serves three purposes. First, it provides a test of overfitting. Second, it allows us to reduce the computational cost of BELT calculations. Third, it allows us to train on data that is approximately uncorrelated; BELT is best suited for working with uncorrelated data. Additional suggestions for data curation are provided in Appx. S7.

## BELT

All BELT calculations were performed using the FitEnsemble package (<https://github.com/kyleabeauchamp/FitEnsemble>). The online FitEnsemble tutorial demonstrates the use of BELT with a single experimental measurement ( $^3J(H^N H^\alpha)$ ). Source code for calculations in this work will be made available at <https://github.com/kyleabeauchamp/EnsemblePaper>.

Regularization strength was determined via cross validation as described in Appx. S5. MCMC sampling for each of the current models the current work took approximately 12 hours. These datasets used included six measurements and approximately 250,000 conformations. For each model, we used PyMC to sample 5,000,000 values of  $\alpha$ ; sampled values of  $\alpha$  were thinned 100-fold to reduce correlation. The first 5,000 samples (before thinning) are discarded as burn-in. MCMC traces are shown in Fig. S2 and discussed in Appx. S8.

## Bayesian Weighting

Bayesian weighting models were constructed with four conformational states consisting of  $\alpha$ ,  $\beta$ ,  $PP_{II}$ , and other. State definitions were taken from.<sup>51</sup> 1,000,000 MCMC samples were generated. The first 5,000 were discarded as burn-in and the remainder were thinned 100-fold. A dirichlet prior (with 1 count per state) was used to model the prior populations of each conformational state. We point out that there are several varieties of BW-like models; our implementation is one of several possible choices of prior. BW calculations were performed using the bayesian weighting module within FitEnsemble.

## References

- (1) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (2) Dethoff, E. A.; Petzold, K.; Chugh, J.; Casiano-Negroni, A.; Al-Hashimi, H. M. *Nature* **2012**,
- (3) Fink, A. L. *Current opinion in structural biology* **2005**, *15*, 35–41.
- (4) Korzhnev, D. M.; Salvatella, X.; Vendruscolo, M.; Di Nardo, A. A.; Davidson, A. R.; Dobson, C. M.; Kay, L. E. *Nature* **2004**, *430*, 586–590.
- (5) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. *Nature* **2005**, *433*, 128–132.
- (6) Lange, O. F.; Lakomek, N.-A.; Farès, C.; Schröder, G. F.; Walter, K. F.; Becker, S.; Meiler, J.; Grubmüller, H.; Griesinger, C.; de Groot, B. L. *science* **2008**, *320*, 1471–1475.
- (7) DePristo, M. A.; de Bakker, P. I.; Blundell, T. L. *Structure* **2004**, *12*, 831–838.
- (8) Lang, P. T.; Ng, H.-L.; Fraser, J. S.; Corn, J. E.; Echols, N.; Sales, M.; Holton, J. M.; Alber, T. *Protein Science* **2010**, *19*, 1420–1431.
- (9) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341–346.
- (10) Li, D.-W.; Bruschweiler, R. *J. Chem. Theory Comput.* **2011**, *7*, 1773.
- (11) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell, A. D. *J. Chem. Theory Comput.* **2012**,
- (12) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M.; Dror, R.; Shaw, D. *PloS one* **2012**, *7*, e32131.

- (13) Pitera, J.; Chodera, J. *J. Chem. Theory Comput.* **2012**,
- (14) Graf, J.; Nguyen, P.; Stock, G.; Schwalbe, H. *J. Am. Chem. Soc.* **2007**, *129*, 1179–1189.
- (15) Fisher, C. K.; Huang, A.; Stultz, C. M. *J. Am. Chem. Soc.* **2010**, *132*, 14919.
- (16) Rieping, W.; Habeck, M.; Nilges, M. *Science* **2005**, *309*, 303–306.
- (17) Patil, A.; Huard, D.; Fonnesbeck, C. J. *Journal of statistical software* **2010**, *35*, 1.
- (18) Rubin, D. *The annals of statistics* **1981**, *9*, 130–134.
- (19) Beauchamp, K.; Lin, Y.; Das, R.; Pande, V. *J. Chem. Theory Comput.* **2012**, *8*, 1409.
- (20) Nerenberg, P.; Head-Gordon, T. *J. Chem. Theory Comput.* **2011**, *7*, 1220–1230.
- (21) Best, R.; Buchete, N.; Hummer, G. *Biophys. J.* **2008**, *95*, L07–L09.
- (22) Grdadolnik, J.; Mohacek-Grosec, V.; Baldwin, R.; Avbelj, F. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1794.
- (23) Jha, A. K.; Colubri, A.; Zaman, M. H.; Koide, S.; Sosnick, T. R.; Freed, K. F. *Biochemistry* **2005**, *44*, 9691–9702.
- (24) Avbelj, F.; Grdadolnik, S.; Grdadolnik, J.; Baldwin, R. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 1272.
- (25) Rozycki, B.; Kim, Y. C.; Hummer, G. *Structure* **2011**, *19*, 109–116.
- (26) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (27) Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D. *Bioinformatics* **2013**,

- (28) Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L.-P.; Shukla, D.; Pande, V. S. *J. Chem. Theory Comput.* **2012**, *9*, 461–469.
- (29) Eastman, P.; Pande, V. *Comp. in Sci. Eng.* **2010**, *12*, 34–39.
- (30) Shaw, D.; Deneroff, M.; Dror, R.; Kuskin, J.; Larson, R.; Salmon, J.; Young, C.; Batson, B.; Bowers, K.; Chao, J. *Commun. ACM* **2008**, *51*, 91–97.
- (31) Senne, M.; Trendelkamp-Schroer, B.; Mey, A. S. J. S.; Schütte, C.; Noé, F. *J. Chem. Theory Comput.* **2012**,
- (32) Beauchamp, K.; Bowman, G.; Lane, T.; Maibaum, L.; Haque, I.; Pande, V. *J. Chem. Theory Comput.* **2011**, *7*, 3412–3419.
- (33) Voelz, V.; Bowman, G.; Beauchamp, K.; Pande, V. *J. Am. Chem. Soc.* **2010**, *132*, 1526–1528.
- (34) Bowman, G. R.; Voelz, V. A.; Pande, V. S. *J. Am. Chem. Soc.* **2011**, *133*, 664.
- (35) Lindorff-Larsen, K.; Piana, S.; Dror, R.; Shaw, D. *Science* **2011**, *334*, 517–520.
- (36) Wang, L.-P.; Chen, J.; Van Voorhis, T. *J. Chem. Theory Comput.* **2012**, *9*, 452–460.
- (37) Wang, L.-P.; Head-Gordon, T. L.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martínez, T. J.; Pande, V. S. *J. Phys. Chem. B* **2013**,
- (38) Friedman, J.; Hastie, T.; Tibshirani, R. *The elements of statistical learning*; Springer Series in Statistics, 2001; Vol. 1.
- (39) Habeck, M.; Rieping, W.; Nilges, M. *Proc. Natl. Acad. Sci.* **2006**, *103*, 1756–1761.
- (40) Kollman, P. *Acc. Chem. Res.* **1996**, *29*, 461–469.
- (41) Wang, J.; Cieplak, P.; Kollman, P. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
- (42) Li, D.; Bruschweiler, R. *Angew. Chem.* **2010**, *122*, 6930–6932.

- (43) Mackerell Jr, A.; Feig, M.; Brooks III, C. *J. Comput. Chem.* **2004**, *25*, 1400–1415.
- (44) Bjelkmar, P.; Larsson, P.; Cuendet, M.; Hess, B.; Lindahl, E. *J. Chem. Theory Comput.* **2010**, *6*, 459–466.
- (45) Kaminski, G.; Friesner, R.; Tirado-Rives, J.; Jorgensen, W. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (46) Han, B.; Liu, Y.; Ginzinger, S.; Wishart, D. *J. Biomol. NMR* **2011**, 1–15.
- (47) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 13–22.
- (48) Li, D.-W.; Brüschweiler, R. *Journal of Biomolecular NMR* **2012**, 1–9.
- (49) Schmidt, J.; Blümel, M.; Löhr, F.; Rüterjans, H. *J. Biomol. NMR* **1999**, *14*, 1–12.
- (50) Vögeli, B.; Ying, J.; Grishaev, A.; Bax, A. *J. Am. Chem. Soc.* **2007**, *129*, 9377–9385.
- (51) Jha, A. K.; Colubri, A.; Zaman, M. H.; Koide, S.; Sosnick, T. R.; Freed, K. F. *Biochemistry* **2005**, *44*, 9691–9702.