# Inferring Structural Ensembles from Noisy Experiments: Application to Trialanine

Kyle A. Beauchamp,[†] Rhiju Das[†,‡] and Vijay S. Pande[†*,¶]

*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 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. Current 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 method that combines simulation and experiment to infer solution ensembles. By using the machinery of Bayesian statistics, BELT gives rigorous uncertainty estimates for structural features and equilibrium properties. Using trialanine as a test system, we show that BELT corrects forcefield bias and outperforms previous analysis methods.

*To whom correspondence should be addressed
†Biophysics Program
‡Biochemistry Department
¶Chemistry Department

# Introduction

Over the past forty years, structural biologists have solved "ground-state" structures of countless biological macromolecules.[1] Modern biology, however, presents many systems that do not fit the single-structure paradigm–excited states of nucleic acids,[2] natively disordered proteins,[3] and protein folding intermediates[4] alike are poorly described by single conformation models. For such systems, conformational ensembles provide a rigorous framework for understanding structural and equilibrium properties.

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. In BELT, Bayesian inference transforms experimental ambiguity into error bars on arbitrary structural features.

To validate BELT, we investigate the conformational propensities of trialanine using NMR measurements.[5] We use BELT to correct MD simulations performed in five different forcefields. Although the raw simulations show wide variations in their conformational preferences, BELT corrects forcefield errors to provide self-consistent estimates of the $\alpha$, $\beta$, and $PP_{II}$ populations.

# 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 $x_j$.

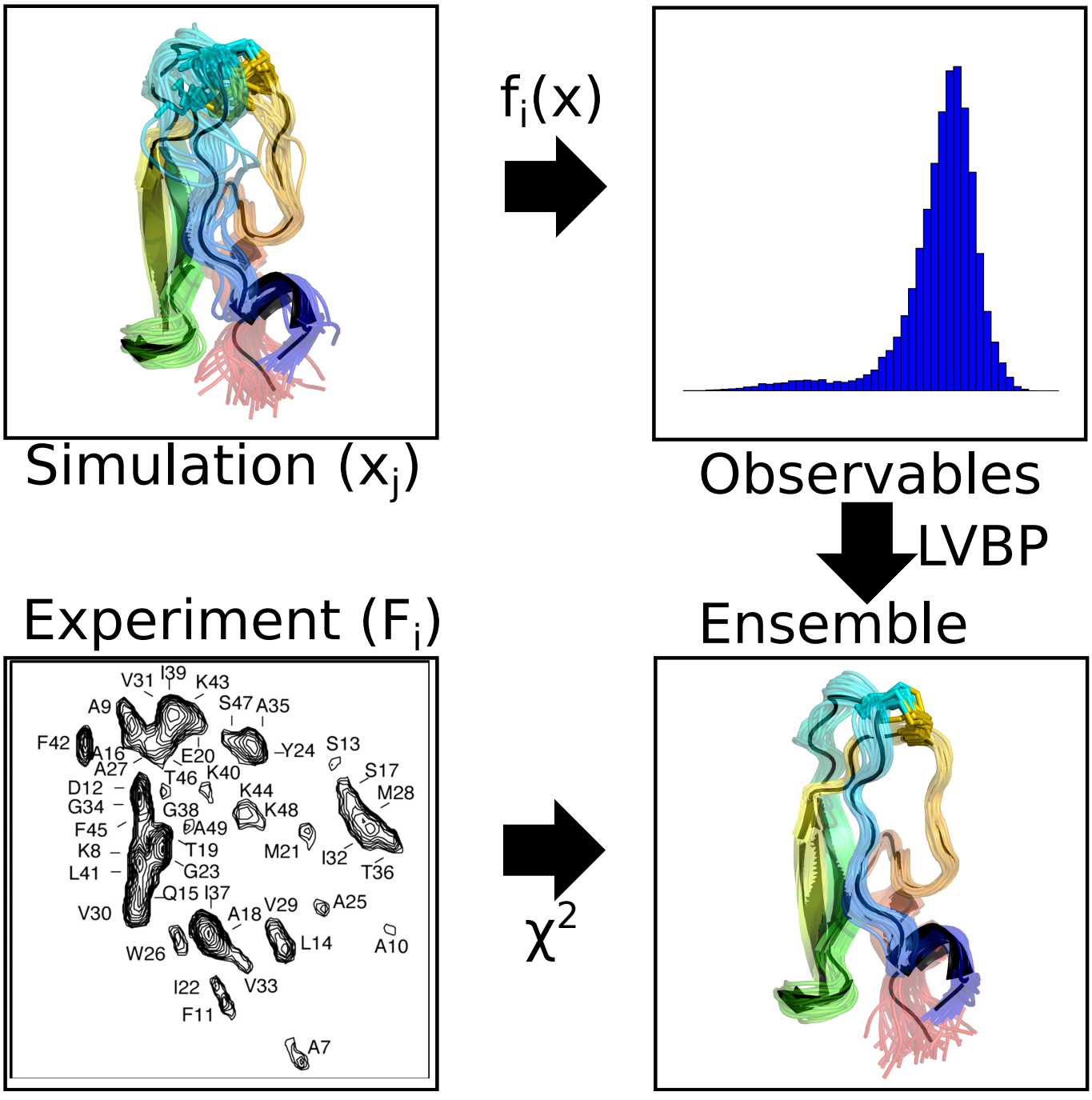



Figure 1: General scheme for BELT modeling.

## Reweighting

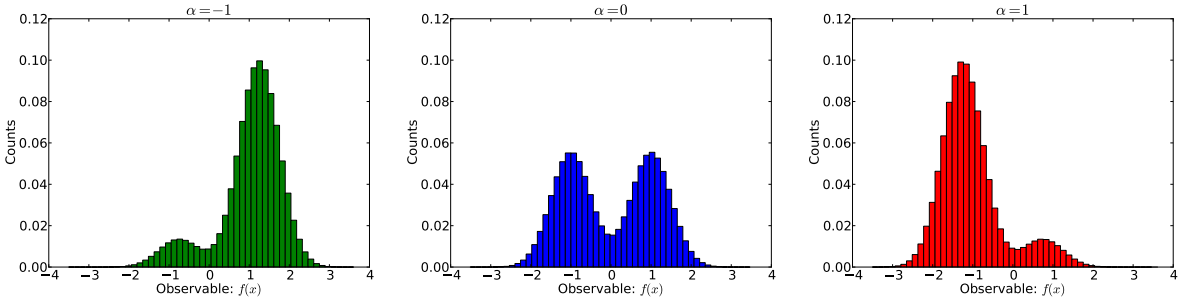
The next step in constructing an ensemble is to calculate the population of each conformation. For convenience, we work with the log populations (i.e. free energies). Inspired by a previous method for restraining simulations,[6] we reweight individual conformations by a biasing potential that is a linear 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. S1):

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

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).




(a) (b) (c)

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. S2) 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 likelihood of $\alpha$ given the measurements:

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

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

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

Note the simple form of the log likelihood. The first term measures the $\chi^2$ agreement between the reweighted ensemble and measurements. The second term is the log of the prior distribution on $\alpha$. Although other priors are possible (see Appxs. S3, S4), we recommend a maximum entropy prior, which penalizes ensembles as they deviate from the raw simulation results:

$$\log P(\alpha) = \lambda \sum_j^m \pi_j(\alpha) \log \pi_j(\alpha)$$

With large values of $\lambda$, the maximum entropy prior favors 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,[7,8] 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,[9] 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[10] can also be used to model the statistical uncertainty of the MD simulations (see Appx. S6).

# Results

## Conformational Propensities of Trialanine

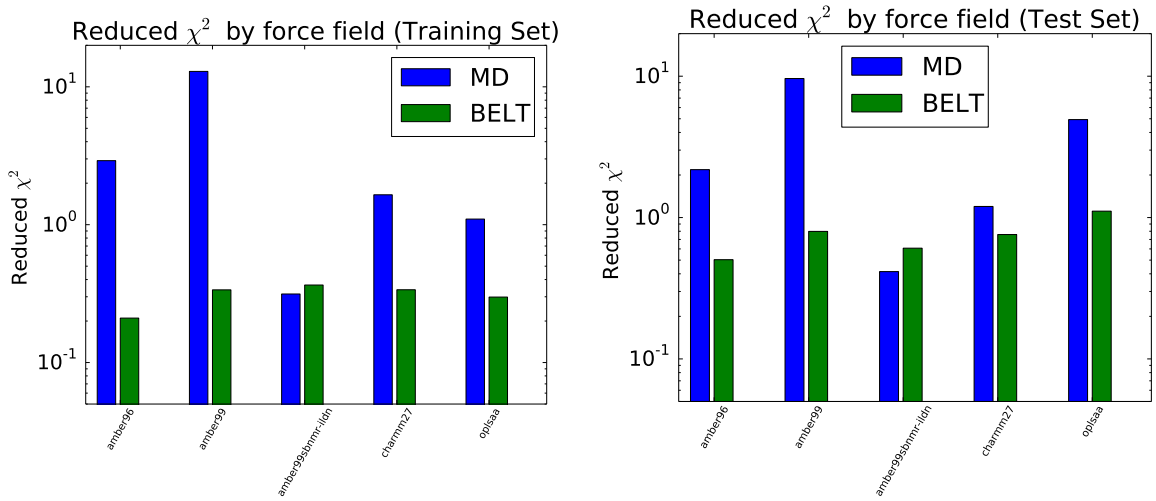
Short peptides provide crucial tests for evaluating and optimizing molecular dynamics forcefields.[5,11–14] Such peptides offer a window into the intrinsic conformational propensities of amino acids, free from the secondary structure bias found in statistical surveys of protein structures.[15] Here, we use BELT to infer the conformational populations of trialanine from chemical shift and scalar coupling measurements.[5]

Trialanine was simulated (see Methods) in five different force fields; six experimental measurements (three scalar couplings, three chemical shifts) probing the central ALA residue were used to construct an BELT ensemble. The five force fields show considerable variation in their agreement with experiment (Fig. 3). The amber96, amber99, charmm27, and oplsaa forcefields, for example, initially show significant deviation from the experimental measurements. Upon reweighting, however, all five forcefields agree with experiment–including experiments that were *not* used to fit the model (Fig. 3b).

Previous experimental studies have shown that short alanine peptides prefer the polyproline type helix (PPII) in solution.[5,14,16] Most molecular dynamics forcefields, however, are known to underpopulate the PPII state.[5,11–13] Our trialanine simulations recapitulate this known deficiency (Fig. 4; blue), 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 conformational ensembles that are robust to forcefield differences (Fig. 4; green). 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.[5,17] In many cases, however, ambiguous experimental data precludes a point-estimate of the conforma-




(a) (b)

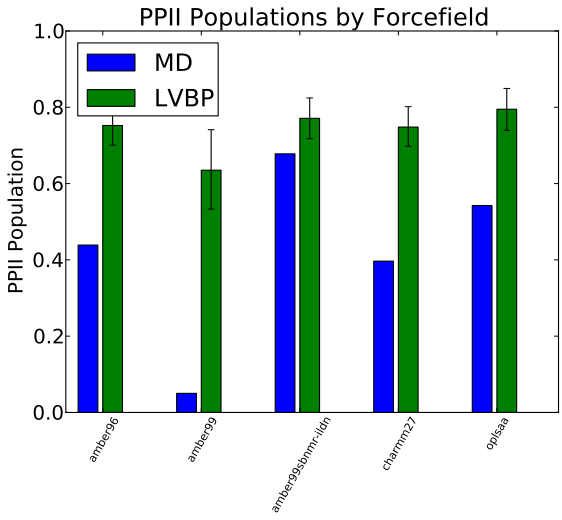
Figure 3: (a). The reduced $\chi^2$ error (e.g. $\frac{\chi^2}{n}$) for raw and reweighted simulations. The BELT reduced $\chi^2$ is estimated as the mean reduced $\chi^2$ over all MCMC samples. (a). Calculated using the six measurements used to fit the BELT model. (b). Calculated using four measurements *not* used to fit the BELT model.

tional ensemble. To illustrate this point, we plot the measured[5] 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 amber99 forcefield reweighted with two (${}^3J(H_N H_\alpha)$, ${}^3J(H_N C')$) NMR measurements (Fig. 6b). The observed bimodal distribution indicates the need to fully characterize the posterior distribution.

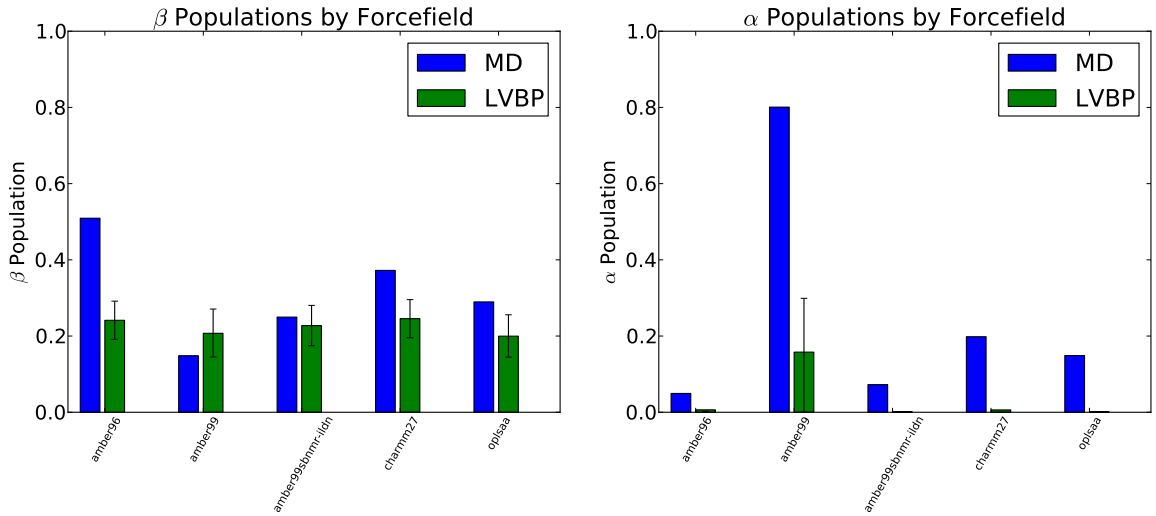
# Discussion

## 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




(a)




(b) (c)

Figure 4: Raw and reweighted conformational propensities for each force field. Note that amber99 shows larger uncertainties due to the extremely limited amount of $PP_{II}$ sampling in the amber99 forcefield.

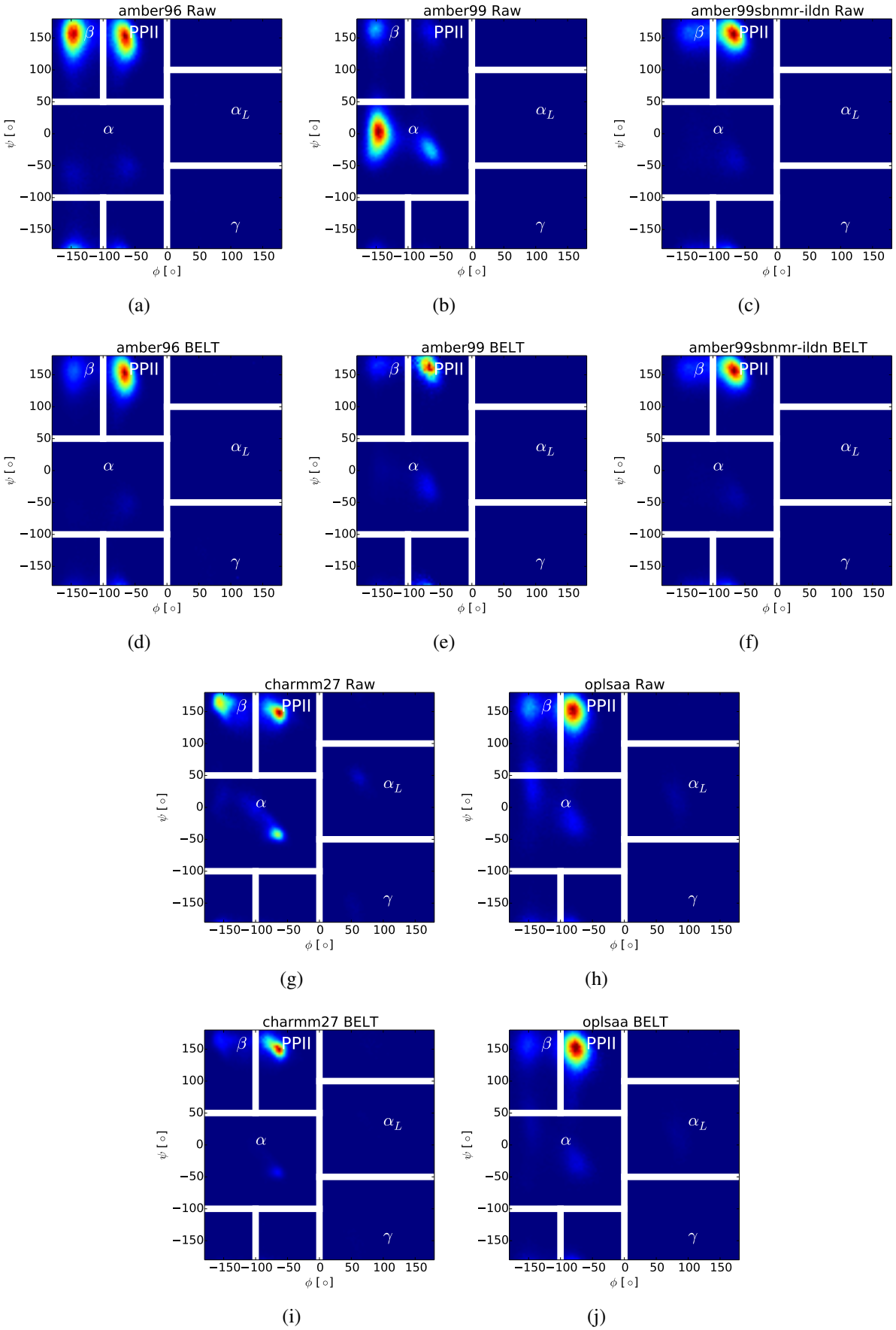
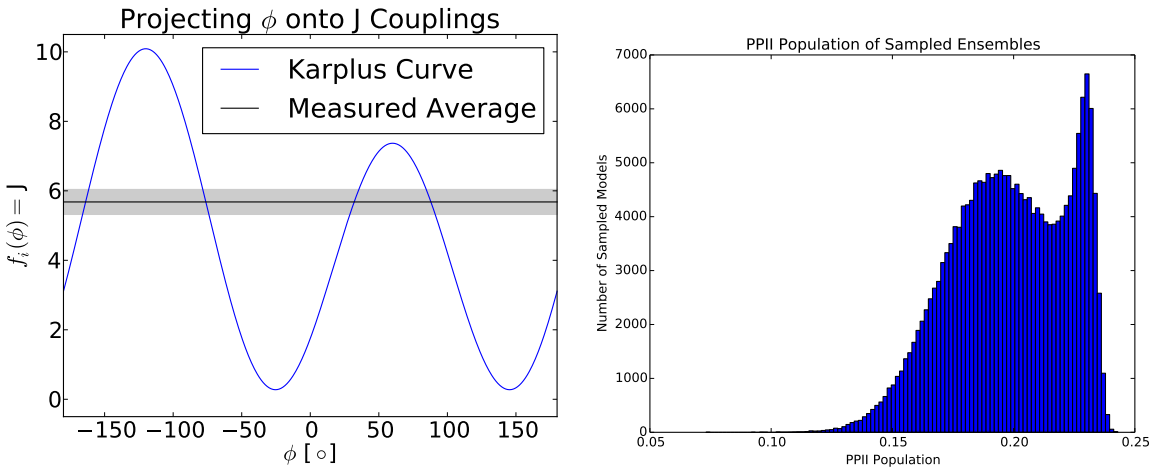





Figure 5: Ramachandran plot of MD and BELT ensembles for each forcefield.




(a) (b)

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 (amber99 forcefield) with two measurements (${}^3J(H_N H_\alpha)$, ${}^3J(H_N C')$) shows a bimodal posterior.

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,[18–21] special-purpose hardware,[22] and distributed computing analysis[23,24] have enabled atomistic simulations to reach the millisecond timescale;[25–28] 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 forcefield differences. The most obvious such difference is the width, shape, and orientation of the $PP_{II}$ basin. Most strikingly, amber96 and oplsaa 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 that involves 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,[5] the EROS inference technique,[17] and the Bayesian Weighting formalism.[7] In our opinion, the primary disadvantage of these techniques is the need for a clustering step. The clustering step can introduce two forms of error. First, clustering can overly coarsen the system of interest. For example, too few states could prevent the model from reproducing multiple experimental observables. At the other extreme, too many states could lead to over-fitting–this is because one needs to estimate a large number of populations. This will lead to poor generalization performance–high 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 threefold.

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 phenenoma. For example, the projection onto observables could be used to rationally infer forcefield parameters, similar to ForceBalance method.[29,30]

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 gives the same answer as a previous[6] maximum entropy approach (see Appx. SX).

We also point out some possibly surprising features of BELT as compared to 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 $< h(x)\chi_s(x) >$ 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. SX and ref.[6]). 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.[31] We therefore evaluated BW (see Appx. SX for details) and BELT models on set-aside test data.

For all five forcefields, BELT produces models that give statistical agreement with experiment on the test data. For three out of five forcefields, BELT gives the *best* model. The BELT prior is relatively unimportant, but the maximum entropy prior appears marginally better than the MVN

prior.

For the two remaining cases, Bayesian Weighting outperforms BELT. For these forcefields (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 provide reasonable models, BELT gives better models in situations with poor starting forcefields.

| | amber96 | amber99 | amber99sbnmr-ildn | charmm27 | oplsaa |
|---|---|---|---|---|---|
| Method (Prior) | | | | | |
| BW | 0.67 | 1.73 | **0.46** | **0.47** | 2.75 |
| BELT (MVN) | 0.62 | 0.90 | 0.66 | 0.76 | **0.91** |
| BELT (maxent) | **0.50** | **0.80** | 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. SX, 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. SX) approaches used herein. In such a model, it is likely that both the prediction and conformational uncertainties are non-normal; future work will be useful for exploring such models.

Another promising avenue is to combine Bayesian modeling of structural ensembles with Bayesian models for experimental observables. A Bayesian formalism for NMR experients has previously been developed;[8,32] 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 forcefield bias and provide generalizable, *forcefield-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,[33] amber99,[34] amber99sbnmr-ildn,[35] charmm27,[36,37] and oplsaa[38] force fields, as previously reported.[11] Simulations were performed using Gromacs 4.5[18] and run at constant temperature (300 K) and pressure (1.01 atm). Each simulation was at least 225 ns long. Conformations were stored every 5 ps.

## Chemical Shift and Scalar Couplings

Chemical shifts (H, HA, CA, CB) for each frame were calculated using the average prediction of ShiftX2,[39] SPARTA+,[40] and PPM;[41] uncertainties were estimated as the mean uncertainty reported by SPARTA+. J couplings were calculated using the following Karplus relations: ${}^3J(H^NC')$,[42] ${}^3J(H^NH^\alpha)$,[43] ${}^2J(NC^\alpha)$,[5] ${}^3J(H^\alpha C')$,[42] ${}^1J(NC^\alpha)$,[5] ${}^3J(H^NC^\beta)$.[43] J coupling uncertainties were approximated as the RMS errors reported when fitting the Karplus coefficients.

## 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`. Unless otherwise noted, we used BELT with the maximum entropy prior; see Appx. SX for results with the MVN prior. Regularization strength was determined via cross validation as described in Appx. SX.

## 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) Graf, J.; Nguyen, P.; Stock, G.; Schwalbe, H. *J. Am. Chem. Soc.* **2007**, *129*, 1179–1189.

(6) Pitera, J.; Chodera, J. *J. Chem. Theory Comput.* **2012**,

(7) Fisher, C. K.; Huang, A.; Stultz, C. M. *J. Am. Chem. Soc.* **2010**, *132*, 14919.

(8) Rieping, W.; Habeck, M.; Nilges, M. *Science* **2005**, *309*, 303–306.

(9) Patil, A.; Huard, D.; Fonnesbeck, C. J. *Journal of statistical software* **2010**, *35*, 1.

(10) Rubin, D. *The annals of statistics* **1981**, *9*, 130–134.

(11) Beauchamp, K.; Lin, Y.; Das, R.; Pande, V. *J. Chem. Theory Comput.* **2012**, *8*, 1409.

(12) Nerenberg, P.; Head-Gordon, T. *J. Chem. Theory Comput.* **2011**, *7*, 1220–1230.

(13) Best, R.; Buchete, N.; Hummer, G. *Biophys. J.* **2008**, *95*, L07–L09.

(14) Grdadolnik, J.; Mohacek-Grosev, V.; Baldwin, R.; Avbelj, F. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1794.

(15) Jha, A.; Colubri, A.; Zaman, M.; Koide, S.; Sosnick, T.; Freed, K. *Biochemistry* **2005**, *44*, 9691–9702.

(16) Avbelj, F.; Grdadolnik, S.; Grdadolnik, J.; Baldwin, R. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 1272.

(17) Rozycki, B.; Kim, Y. C.; Hummer, G. *Structure* **2011**, *19*, 109–116.

(18) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(19) 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**,

(20) 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.

(21) Eastman, P.; Pande, V. *Comp. in Sci. Eng.* **2010**, *12*, 34–39.

(22) 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.

(23) Senne, M.; Trendelkamp-Schroer, B.; Mey, A. S. J. S.; Schütte, C.; Noé, F. *J. Chem. Theory Comput.* **2012**,

(24) Beauchamp, K.; Bowman, G.; Lane, T.; Maibaum, L.; Haque, I.; Pande, V. *J. Chem. Theory Comput.* **2011**, *7*, 3412–3419.

(25) Voelz, V.; Bowman, G.; Beauchamp, K.; Pande, V. *J. Am. Chem. Soc.* **2010**, *132*, 1526–1528.

(26) Bowman, G. R.; Voelz, V. A.; Pande, V. S. *J. Am. Chem. Soc.* **2011**, *133*, 664.

(27) 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.

(28) Lindorff-Larsen, K.; Piana, S.; Dror, R.; Shaw, D. *Science* **2011**, *334*, 517–520.

(29) Wang, L.-P.; Chen, J.; Van Voorhis, T. *J. Chem. Theory Comput.* **2012**, *9*, 452–460.

(30) 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**,

(31) Friedman, J.; Hastie, T.; Tibshirani, R. *The elements of statistical learning*; Springer Series in Statistics, 2001; Vol. 1.

(32) Habeck, M.; Rieping, W.; Nilges, M. *Proc. Natl. Acad. Sci.* **2006**, *103*, 1756–1761.

(33) Kollman, P. *Acc. Chem. Res.* **1996**, *29*, 461–469.

(34) Wang, J.; Cieplak, P.; Kollman, P. *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(35) Li, D.; Bruschweiler, R. *Angew. Chem.* **2010**, *122*, 6930–6932.

(36) Mackerell Jr, A.; Feig, M.; Brooks III, C. *J. Comput. Chem.* **2004**, *25*, 1400–1415.

(37) Bjelkmar, P.; Larsson, P.; Cuendet, M.; Hess, B.; Lindahl, E. *J. Chem. Theory Comput.* **2010**, *6*, 459–466.

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

(39) Han, B.; Liu, Y.; Ginzinger, S.; Wishart, D. *J. Biomol. NMR* **2011**, 1–15.

(40) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 13–22.

(41) Li, D.-W.; Brüschweiler, R. *Journal of Biomolecular NMR* **2012**, 1–9.

(42) Schmidt, J.; Blümel, M.; Löhr, F.; Rüterjans, H. *J. Biomol. NMR* **1999**, *14*, 1–12.

(43) Vögeli, B.; Ying, J.; Grishaev, A.; Bax, A. *J. Am. Chem. Soc.* **2007**, *129*, 9377–9385.