This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

Metadynamic Metainference: convergence towards force field independent structural ensembles of a disordered peptide

Thomas Löhr,[1] Alexander Jussupow,[1] and Carlo Camilloni[1, a)]

Department of Chemistry and Institute for Advanced Study, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany

(Dated: 4 April 2017)

Metadynamic Metainference has been recently introduced as a theoretical framework to determine structural ensembles by combining and weighting by their noise multiple sources of experimental data with molecular mechanics force fields and Metadynamics simulations. Here we build upon these initial developments to further extend and streamline the computational approach and to show that Metadynamic Metainference can actually determine a structural ensemble for a disordered peptide that is essentially independent from the employed force field. We also show that it is possible to use a very computationally efficient implicit solvent force field in place of very expensive state-of-the-art explicit solvent ones without a significant loss in accuracy.

a)Electronic mail: carlo.camilloni@ch.tum.de

I. INTRODUCTION

Structural ensembles of proteins are becoming an invaluable tool to understand biological mechanisms at the molecular level[1–5]. Accurate ensembles can represent the conformational fluctuations of proteins and enable the observation of multiple substates populated by these molecules in given experimental conditions, thus providing a link between structure and function[6,7]. This is of particular importance when considering systems that are intrinsically dynamic, like multi-domain proteins with disordered linkers and, more in general, intrinsically disordered proteins and regions[8–10].

Determining structural ensembles for such dynamic systems is, however, a challenging task that requires at the same time an accurate modelling and a thorough sampling of the system's conformational space[1,11,12]. Recently, to address these challenges, we have introduced Metadynamic Metainference[13,14] (M&M). Metainference[13] is a Bayesian framework that allows integrating multiple sources of information about a system and optimally balancing them. In this approach, one modifies the *a priori* knowledge about a system (i.e. its physicochemical properties as described by molecular mechanics force fields) using data acquired from experimental measurements, and balances those by sampling on-the-fly a statistical distribution of noises that can effectively take into account all the sources of errors (i.e. ensemble averaging, statistical errors, systematic errors and experimental data modelling errors). By combining Metainference with Metadynamics[15,16] one can then enhance the sampling of the Metainference model and explore conformational states that can be separated by significant high free-energy barriers on the time scale of standard molecular dynamics[14].

While Metainference, as well as other statistical methods to determine ensembles[1,11], can update prior knowledge to take the available experimental knowledge into account, the question is open about the possibility of obtaining ensembles that do not depend on the specific prior knowledge employed. From a theoretical point of view, the farther the prior is from providing a good description of a system, the more abundant and better the data must be to obtain a good representation of it[13]. In practice, one could be interested in how close the ensembles determined for a disordered system are when employing state-of-the-art molecular mechanics force fields and different sources of experimental data[1].

In the following we have studied the disordered peptide EGAAWAASS[17] making use of

M&M, two state-of-the-art force fields from different families that have given promising results in describing folded and unfolded proteins (CHARMM22*[18] and AMBER99SB with TIP4P-D[19]) and integrated three different sources of commonly available experimental data, NMR chemical shifts (CS), ^{3}J-couplings and Residual Dipolar Couplings (RDCs). This model system has been employed recently, due to the availability of multiple accurate NMR measures[17,20], to highlight the deficiencies of force fields in describing disordered systems[20,21]. M&M allowed us to determine structural ensembles for this peptide that are essentially indistinguishable from the point of view of a number of independent parameters. These include actual experimental observables, probability distributions of multiple global degrees of freedom and secondary structures contents.

Prompted by this observation we challenged M&M to determine an ensemble using CHARMM36[22] with the EEF1-SB[23,24] implicit solvent model as an extremely computationally inexpensive prior. In the case of the current peptide we could determine an ensemble of comparable quality with those obtained from the explicit solvent force fields but at a fraction of the computational cost. This approach could alleviate the computational cost of studying systems that require the use of large simulation boxes and huge amounts of water molecules.

II. THEORY AND METHODS

Metainference[13] employs Bayesian statistics to allow updating a prior distribution by considering some new additional information. In particular, Metainference is derived to take into account informations that are the result of ensemble averaging, i.e. averaging over a full probability distribution. In computational structural biology the prior is usually a mechanistic force field that describes more or less accurately the interactions of the atoms. Additional informations are structural equilibrium measures like those obtained by NMR spectroscopy. Equilibrium observables are always the result of time and ensemble averaging and as such should be employed to update an ensemble and not the probability of observing a single structure. This latter case is a good approximation when the single structure represents by far the most populated state of the system. Notably, Metainference is equivalent to Inferential Structural Determination if ensemble averaging is not taken into account[25].

In M&M[14] multiple simulations, replicas, are performed in parallel for the same system,

in the same conditions and using the same force field. These independent simulations are coupled with an energy defined as a function of the difference between the average over the replicas of a backcalculated quantity and a reference value for the same quantity (e.g. the difference between a chemical shift calculated for all the replicas and averaged and the experimental value for that same chemical shift). The sampling of each replica is enhanced by Metadynamics[26,27] (in this case by Parallel Bias Metadynamics[28]), which adds a history dependent bias as a function of a set of collective variables (CVs). In M&M Metainference and Metadynamics are coupled by the calculation of the average over the replicas. While in standard Metainference each replica contributes with the same weight to the average of backcalculated experimental observables, in M&M the weighted average provides a better estimator. This is particularly feasible in the case where the same bias is applied to all replicas in such a way that in the limit of a quasi-static bias the weight of a replica can be approximated on the fly as $w_r \propto \exp\left(+V^{\mathrm{MetaD}}(\mathrm{CV}(X_r))/k_{\mathrm{B}}T\right)$.

In the following we first reprise Metainference theory, then we introduce a simple on-the-fly estimate for the only parameter of Metainference and introduce a correction to take into account the effect of the weighted average on the distribution of the forces when using Metadynamics. Finally we extend Metainference to work with data defined but for a multiplicative constant. All the methods described in this work are implemented and available in PLUMED[29], a working and complete input file that allows reproducing our results and gain a better understanding of M&M is provided in the Supporting Information.

A. Metainference.

The energy of a Metainference simulation is defined as $-k_{\mathrm{B}}T\ln(P)$ where k_{B} is the Boltzmann constant, T the temperature and P the Metainference posterior distribution, calculated over a finite number of replicas, N_r. Its general definition is:

$$P(\tilde{\boldsymbol{f}}, \boldsymbol{\sigma}^{\mathbf{B}}, \boldsymbol{X}, \boldsymbol{\sigma}^{\mathbf{SEM}} | \boldsymbol{d}) = \prod_{r=1}^{N_r} p(X_r) \prod_{i=1}^{N_d} p(d_i | \tilde{f}_i, \sigma^{\mathrm{B}}_{r,i}) p(\tilde{f}_i | \boldsymbol{X}, \sigma^{\mathrm{SEM}}_i) p(\sigma^{\mathrm{B}}_{r,i}) p(\sigma^{\mathrm{SEM}}_i), \qquad (1)$$

where $p(d_i|\tilde{f}_i, \sigma^{\mathrm{B}}_{r,i})$ is the likelihood of the experimental data d_i given $\tilde{f}_i$ and a vector of $\sigma^{\mathrm{B}}_{r,i}$. $\tilde{f}_i$ is the average of the forward model f_i used to predict the experimental observable i from a model calculated on an infinite number of replicas and $\sigma^{\mathrm{B}}_{r,i}$ is an uncertainty parameter

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

that describes random and systematic errors in the experimental data as well as errors in the forward model. $p(\tilde{f}_i|\boldsymbol{X},\sigma_i^{\mathrm{SEM}})$ is the likelihood of observing $\tilde{f}_i$ given an estimate $\langle f_i\rangle_{N_r}$ over a finite number of replicas, where $\langle f_i(\boldsymbol{X})\rangle_{N_r} = \frac{1}{N_r}\sum_{r=1}^{N_r} f_i(X_r)$ and σ_i^{SEM} is an uncertainty parameter. $p(X_r)$, $p(\sigma_{r,i}^{\mathrm{B}})$ and $p(\sigma_i^{\mathrm{SEM}})$ are the priors on the conformations X_r (i.e. the force field), on the uncertainty σ^{B} in the experimental and backcalculated data, and on the uncertainty σ^{SEM} in the estimate of the true ensemble average, respectively.

A simple form can be obtained by choosing a Gaussian form for the likelihood $p(d_i|\tilde{f}_i,\sigma_{r,i}^{\mathrm{B}})$ given that $p(\tilde{f}_i|\boldsymbol{X},\sigma_i^{\mathrm{SEM}})$ is a Gaussian for the central limit theorem. In this case it is possible to write the Metainference energy as:

$$E_{\mathrm{MI}}(\boldsymbol{X},\boldsymbol{\sigma}|\boldsymbol{d}) = \sum_{r=1}^{N_r}\Big\{E_{\mathrm{ff}}(X_r) + k_{\mathrm{B}}T\sum_{i=1}^{N_d}\Big[\frac{(\langle f_i(\boldsymbol{X})\rangle - d_i)^2}{2\sigma_{r,i}^2} + \frac{1}{2}\ln 2\pi\sigma_{r,i}^2 + \frac{1}{2}\ln\frac{\sigma_{r,i}^2}{2}\Big]\Big\}, \quad (2)$$

where $E_{\mathrm{ff}}(X_r)$ is the energy of the force field for the conformation X_r, $\sigma_{r,i}$ is the total uncertainty defined as $\sigma_{r,i}^2 = (\sigma_{r,i}^{\mathrm{B}})^2 + (\sigma_i^{\mathrm{SEM}})^2$ and the two logarithmic terms are the normalisation and the Jeffreys' prior for $\sigma_{r,i}$, respectively.

1. *Estimate of* σ^{SEM}.

The Metainference energy in Gaussian form immediately shows the similarities between Metainference and the replica-averaged simulations based on the maximum entropy[30] principle. Indeed Metainference reduces to the maximum entropy replica-averaged modelling[6,30,31] in the case that the only source of error is the ensemble averaging, σ_i^{SEM} is equivalent to the force constant employed there. In Metainference as well as in replica-averaged restrained simulations $1/(\sigma_i^{\mathrm{SEM}})^2$ is the only parameter to be set and it was shown that this must be chosen to be the largest possible force constant that can be integrated correctly by the system. Furthermore it was observed that it should scale more than linearly with the number of replicas[31]. In Metainference σ^{SEM} is related to the standard error of the mean and as such, in absence of other sources of errors, the force constant actually scales $\propto N_r^2$.

In principle one should set a σ_i^{SEM} for each experimental data used as restraint. A practical solution to this problem that was often employed is that of selecting one value common for all the data in a dataset. Here we introduce an alternative solution. We estimate it as $\sigma_i^{\mathrm{SEM}} = \sqrt{\max(\mathrm{Var}[f_i](t))/N_r}$, that is the square root of the maximum over the simulation

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

ing of the variance of the forward model for the observable i divided by the number of replicas. This guarantees the correct scaling of σ_i^{SEM} with the number of replicas, a weak time dependence and a value proportional to the variance of the backcalculated observable and thus its dynamics. In the Supporting Information, Figure S1 and S2 show how the algorithm allows to quickly reach a stable estimate for σ^{SEM} in the first few nanoseconds of simulations.

2. *Estimate of the weighted* σ^{SEM}.

In M&M the arithmetic average is substituted by a weighted average to consider for the effect of the bias, $\langle f_i(\boldsymbol{X})\rangle_{N_r} = \sum_{r=1}^{N_r} \frac{w_r}{N_w} f_i(X_r)$, where $N_w = \sum_{r=1}^{N_r} w_r$. In this case the forces resulting from Metainference are not equally distributed among the replicas, but are distributed proportionally to the weight of each single replica at each time step. In this case the standard error of the mean should take the variance of the weights into account. This is done by implementing it as[32]:

$$(\sigma_i^{\mathrm{SEM}})^2 = \frac{N_r}{(N_r-1)N_w^2}\left[\sum_{r=1}^{N_r}(w_r f_i(X_r) - \langle w\rangle\langle f_i(\boldsymbol{X})\rangle)^2 - \right.$$
$$\left. +2\langle f_i(\boldsymbol{X})\rangle \sum_{r=1}^{N_r}(w_r - \langle w\rangle)(w_r f_i(X_r) - \langle w\rangle\langle f_i(\boldsymbol{X})\rangle) + \langle f_i(\boldsymbol{X})\rangle^2 \sum_{r=1}^{N_r}(w_r - \langle w\rangle)^2\right] \tag{3}$$

and as for the unweighted case using the square root of the maximum value sampled along the simulation.

3. *Generalisation for observables defined but for a scaling factor.*

Experimental observables can be defined modulo a multiplicative constant, this is the case with RDCs as their intensity is proportional to the fraction of aligned molecules. In these cases it is not possible to directly compare back-calculated and experimental data. One possible solution is that of considering an energy term proportional to the correlation between the experimental and the back-calculated data[33]. Alternatively one can extend the Metainference formalism to take additional parameters into account, for example a scaling factor. In this case the Metainference energy becomes:

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

$$E_{\mathrm{MI}}(\boldsymbol{X},\boldsymbol{\sigma},\lambda|\boldsymbol{d}) = \sum_{r=1}^{N_r}\left\{E_{\mathrm{ff}}(X_r) + k_{\mathrm{B}}T\sum_{i=1}^{N_d}\left[\frac{(\lambda\langle f_i(\boldsymbol{X})\rangle - d_i)^2}{2\sigma_{r,i}^2} + \frac{1}{2}\ln 2\pi\sigma_{r,i}^2 + \frac{1}{2}\ln\frac{\sigma_{r,i}^2}{2}\right]\right\} \tag{4}$$

where λ is the scaling factor and $\sigma_{r,i}^2 = (\sigma_{r,i}^{\mathrm{B}})^2 + \lambda^2(\sigma_i^{\mathrm{SEM}})^2$ and the two logarithmic terms are the normalisation and the Jeffrey prior, respectively.

4. Restraint correction for high forces.

The restraint intensity dependent on σ^{B} and σ^{SEM} can occasionally lead to unrealistic forces, causing instability in particular in the transient time at the beginning of the simulation when σ^{SEM} is still under estimated. To decrease the probability of this occurring a correction factor is introduced in such a way as to temporarily decrease the applied forces. This is defined as:

$$s_t = \begin{cases} s_{t-1} - \frac{\Delta s}{100}\ln(\frac{s_{t-1}}{s_{\min}}) & \text{if } n_{F_{\mathrm{MD}}>F_{\max}} = 0 \\ s_{t-1} + \Delta s\ln(n_{F_{\mathrm{MD}}>F_{\max}} + 1) & \text{if } n_{F_{\mathrm{MD}}>F_{\max}} > 0 \\ s_{\max} & \text{if } n_{F_{\mathrm{MD}}>F_{\max}} > 0 \text{ and } s_t > s_{\max} \end{cases} \tag{5}$$

where s_t is the correction factor at timestep t to be multiplied with σ^{SEM}, Δs is the step size for the correction factor, $s_{\min}$ and $s_{\max}$ are the respective minimum and maximum possible correction values and $n_{F_{\mathrm{MD}}>F_{\max}}$ is the number of molecular dynamics forces above a certain threshold force $F_{\max}$. This update rule has the effect of immediately relaxing the restraint in the case of excessively high forces, followed by a slow annealing to the $s_{\min}$ value in the case of no high force events. By specifying a $s_{\min}$ value different from 1, one is able to bias the restraint intensity towards lower or higher values.

B. Simulation Details.

All simulations (Table I) are carried out with Gromacs 5.1.4[34] and a development version of PLUMED 2.3[29]. The peptide with sequence EGAAWAASS is created in VMD[35] and is solvated in a rhombic dodecahedron box with side lengths of 4.5, 4.5 and 3.2 nm using 2118

Simulation	Force field	Water model	Performances	Convergence	Free Energies	Back-calculations
Unrestrained	CHARMM22*	TIP3P	14.8	Fig. S5	Fig. S13	Fig. S21
	AMBER99SB	TIP4P-D	10.5	Fig. S6	Fig. S14	Fig. S22
CS, JC	CHARMM22*	TIP3P	14.6	Fig. S7	Fig. S15	Fig. S23
	AMBER99SB	TIP4P-D	10.3	Fig. S8	Fig. S16	Fig. S24
CS, JC, RDCs	CHARMM22*	TIP3P	14.4	Fig. S9	Fig. S17	Fig. S25
	AMBER99SB	TIP4P-D	10.2	Fig. S10	Fig. S18	Fig. S26
Unrestrained	CHARMM36	EEF1-SB	389.4	Fig. S11	Fig. S19	Fig. S27
CS, JC, RDCs	CHARMM36	EEF1-SB	244.9	Fig. S12	Fig. S20	Fig. S28

TABLE I. All simulations performed. Force-field, water model, performances (ns/day/replica) and experimental data used as restraints are reported. Performances were estimated on an Intel E5-2660 2.4 GHz using one thread per replica.

water molecules. The system is neutralized by addition of 3 Na^+ and 2 Cl^- ions. Minimization of the system is performed with the steepest descent algorithm to a maximum force of less than 100 kJ/mol/nm. Equilibration is performed over a time range of 500 ps in the NVT ensemble using the Bussi thermostat[36] and for 500 ps in the NPT ensemble using Parrinello-Rahman[37] pressure coupling while applying a position restraint on all heavy atoms. Production simulations are carried out with AMBER99SB[38] with TIP4P-D[19] water model and CHARMM22*[18,39] with TIP3P[40] water model with a time step of 2 fs at a temperature of $T = 300$ K in the NPT ensemble. Van der Waals and electrostatic interactions are modelled using the Particle-Mesh-Ewald[41,42] approach and a cutoff for the short-range interactions of 0.9 nm. Constraints are applied on all bonds with the LINCS algorithm[43] using a matrix expansion to the order of 6 and 2 iterations per step.

Metadynamics[15] is performed with the Well-Tempered[44], Parallel-Bias[28] and multiple-walkers[45] protocols, using a Gaussian deposition stride of 500 steps (1 ps), a bias factor of

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

and a Gaussian height of 0.3 kJ/mol for 14 replicas. The following collective variables are biased, corresponding sigma values are given in parentheses: All backbone ψ and ϕ dihedral angles ($\sigma = 0.6$) as well as the E1-S9 C^{α}-C^{α} distance ($\sigma = 0.3$ nm^{-1}), W5 χ^1, W5 χ^2 ($\sigma = 0.6$), similarities between ϕ_3 and ϕ_6 as well as ψ_3 and ψ_6 dihedral angles ($\sigma = 0.3$). Each replica is run for 100 ns for a total of 1.4 μs nominal simulation time per ensemble.

1. *M&M simulation including ^3J couplings and chemical shifts.*

Metainference calculations are performed using ensemble averages weighted according to the Metadynamics bias potential. Experimental data is provided by Dames et al[17]. Chemical shifts were calculated using CamShift[46,47] for NH, HN, Hα, Cα, Cβ and C' backbone atoms while excluding the first and last residues. Hα-N, Hα-HN, W5 C-Cγ and W5 N-Cγ ^{3}J-coupling constants were calculated using the Karplus equation[48]:

$$^3J(\theta) = A\cos^2(\theta + \Delta\theta) + B\cos(\theta + \Delta\theta) + C \tag{6}$$

where $^3J(\theta)$ is the coupling in Hz, A, B and C are the Karplus parameters dependent on the type of coupling, θ is a dihedral angle and $\Delta\theta$ is a constant shift added on to the angle. The Karplus parameters and shift $\Delta\theta$ are taken from ref[49,50]. The noise is sampled independently for each datapoint through brownian motion (flat prior) with a stepsize of 0.5 and hard limits at 0.001 and 25 respectively. The restraint correction for high forces was applied with $s_{\min} = 1.0$, $s_{\max} = 2.0$, $\Delta s = 0.001$ and $F_{\max} = 3500$ kJ/mol/nm.

2. *M&M simulation including RDCs.*

Simulations are performed as described above with the addition of residual dipolar couplings for N-H and Cα-Hα bonds[17]. RDCs are calculated using the θ-method[33], each coupling is calculated independently using the dipolar coupling definition:

$$D_i = -\frac{\mu_0\gamma_1\gamma_2\hbar}{8\pi^3}\left(\frac{3\cos^2\vartheta_i - 1}{r_i^3}\right) \tag{7}$$

where r_i is the bond length, μ_0 is the magnetic constant, γ_1 and γ_2 are the gyromagnetic ratios for the two atoms, $\hbar$ is the Planck constant and ϑ_i is the angle between the bond and the z-axis. The coupling is then averaged and compared modulo a scaling factor λ with

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

ing experimental data. This allows to simultaneously account for the conformational and rotational averaging measured by RDCs[11,33,51,52]. The scaling factor is sampled during the simulation using an Ornstein-Uhlenbeck process (Gaussian prior):

$$\mathrm{d}\lambda_t = \frac{1}{2}(\mu - \lambda_t) + \Delta\lambda \frac{e}{\pi} \mathrm{d}W_t \tag{8}$$

where $\mathrm{d}\lambda_t$ is the step taken, μ is the specified mean of the stationary Gaussian distribution, λ_t is the scaling value at time t, $\Delta\lambda$ is the standard deviation of the stationary Bayesian distribution and $\mathrm{d}W_t$ denotes the Wiener process. The values chosen for N-H and Cα-Hα RDCs are $\mu = 8$ and $\mu = 9$ respectively and $\Delta\lambda = 0.5$. In Figure S3 and S4 the sampling of the scaling factor for N-H and Cα-Hα RDCs is shown. The sampling converges quickly after a few steps, of notice is that the average value for the scaling factor found is different depending on the force field used as a prior, this is due to the differences in the bond lengths of those bonds in the two force fields. In CHARMM22* an N-H bond is 0.0997 nm long and a Cα-Hα is 0.1080 nm long , while in AMBER99SB the same bonds are 0.1010 and 0.1090 nm long.

3. M&M using the EEF1-SB implicit solvent model.

Simulations in implicit solvent are performed using the EEF1 model originally developed by Lazaridis and Karplus[23] and subsequently optimised by Bottaro *et al.* (EEF1-SB)[24] in combination with CHARMM36[22]. EEF1-SB is a solvent-accessible surface area based model, where the free energy of solvation is computed using a pairwise interaction term for non-hydrogen atoms:

$$\Delta G_i^{\mathrm{solv}} = \Delta G_i^{\mathrm{ref}} - \sum_{j \neq i} f_i(r_{ij}) V_j \tag{9}$$

where $\Delta G_i^{\mathrm{solv}}$ is the free energy of solvation, $\Delta G_i^{\mathrm{ref}}$ is the reference solvation free energy, V_j is the volume of atom j and

$$f_i(r) 4\pi r^2 = \frac{2}{\sqrt{\pi}} \frac{\Delta G_i^{\mathrm{free}}}{\lambda_i} \exp\left\{ -\frac{(r - R_i)^2}{\lambda_i^2} \right\} \tag{10}$$

where $\Delta G_i^{\mathrm{free}}$ is the solvation free energy of the isolated group, λ_i is the correlation length equal to the width of the first solvation shell and R_i is the van der Waals radius of atom i. The implicit solvation model is implemented in PLUMED. In our implementation interactions are cut off after a range of $3\lambda_i$. In addition, electrostatic interactions are further screened with a

position dependent dielectric constant of the form $\epsilon = 1/(\alpha r)$. Bottaro *et al.* optimised α to 15 nm^{-1} and added an energy correction for backbone dihedrals on the $N-C'-C\alpha-C\beta$. All these parameters and corrections are designed to be used with the CHARMM36 force field[22]. Charged amino acids are neutralised by adjusting the partial charges, leaving a completely neutral molecule. Minimization is performed as for the explicit solvent simulations. The system is evolved by a Langevin dynamic integrator with a friction coefficient of 1 ps^{-1} at $T = 300$ K. Coulomb interactions are tabulated with a distance dependent dielectric constant of $\epsilon = 15r$ and a cut off at 0.9 nm, while van der Waals interactions are switched off smoothly between 0.7 and 0.9 nm. All pairwise interactions are computed using a neighbourlist with a buffer of 0.2 nm with respect to the cut-off, which is updated every 10 simulation steps. Constraints are applied on all bonds with the LINCS algorithm, as described above. Metadynamics and M&M simulations are performed as already mentioned for explicit solvent simulations using the same collective variables, parameters and experimental data.

4. Analysis.

In well-tempered Metadynamics the time-dependent bias converges to a quasi-static distribution, as a consequence a signature of convergence can be obtained by a block comparison of the sampling after a transient time. If the simulations are converged, the histograms obtained for non-overlapping blocks of simulations should result in comparable effective free energies.

In the present case convergence is assessed by comparing the free energies calculated from the histograms of each biased collective variable for the last two 45 ns segments of the simulation (i.e. from 10 to 55 ns and from 55 to 100 ns). The free energies represent the effective potential felt by the system as a sum of the force field, the Metainference potential if present, and Metadynamics. In Figure S5 to S12 the comparisons of the effective free energies obtained along the 21 collective variables employed are shown, with differences that are limited to few high-energy regions and an average root-mean-square deviation of 0.50 kJ/mol. The converged free energies for all the collective variables are shown in Figure S13 to S18.

Equilibrium distributions are then recovered by reweighting the ensembles according to the final deposited Metadynamics bias[53]. The weight of each sampled conformation is given

by $w_i = \exp\left(+V^{\mathrm{MetaD}}(\mathrm{CV}(X))/k_{\mathrm{B}}T\right)/Z$, where $V^{\mathrm{MetaD}}(\mathrm{CV}(X))$ is the Metadynamics bias calculated for conformation X at the end of the simulation and Z is the normalisation. Chemical shifts are backcalculated using CamShift[46], residual dipolar couplings are computed using the single-value-decomposition method. ^{3}J-couplings are back-calculated using the Karplus equation. All experimental observables are calculated as weighted ensemble averages.

In addition to the conformational ensemble, the result of a Metainference calculation also includes an estimate of the errors, $\sigma^{\mathrm{B}}_{r,i}$, for all the experimental data added[13]. These errors incorporate in a single number an independent estimate of the experimental random and systematic errors as well as the errors in the forward-model. Indeed, while σ_i^{SEM} is an error that accounts for the use of a limited number of replicas on the fly, $\sigma^{\mathrm{B}}_{r,i}$ is a useful additional source of information that results from the use of Metainference.

Finally, in order to further compare the ensembles not only in terms of their agreement with experimental data but also with respect to finer properties, similarities between probability distributions are computed using the Jensen-Shannon divergence. Given two probability distributions P and Q obtained by two ensembles, their difference is

$$D_{\mathrm{JS}}(P||Q) = \frac{1}{2}D_{\mathrm{KL}}(P||M) + \frac{1}{2}D_{\mathrm{KL}}(Q||M) \tag{11}$$

where $M = \frac{1}{2}(P+Q)$ and D_{KL} is the Kullback-Leibler divergence: $D_{\mathrm{KL}} = \sum_i P(i)\ln(P(i)/Q(i))$.

III. RESULTS

In the following we present the results of eight ensembles for the EGAAWAASS[17] peptide (cf. Table I) obtained by running for each case 14 replicas for 100 ns per replica, either using only Metadynamics[15], i.e. without the addition of any experimental restraint, or by coupling the replicas using Metadynamic Metainference[14] and multiple experimental data. We have tested two state-of-the-art force fields in explicit solvent, CHARMM22* in TIP3P[18] and AMBER99SB in TIP4P-D[19], and the CHARMM36 EEF1-SB[24] implicit solvent force field recently optimised to study disordered systems. The addition of experimental data modifies the ensembles towards a result where both local and non-local properties are comparable, irrespective of the original force field employed. The implicit solvent scheme supplemented by the experimental data allows us to obtain results comparable to those obtained in explicit

solvent at a fraction of the computational cost. A PLUMED[29] input file is provided in the Supporting Information to reproduce all the simulations.

A. Comparison with the experimental data.

First, we assessed the quality of the force fields and the ability of Metainference to successfully improve them through the weighted incorporation of experimental informations. Root-mean-square deviations (RMSDs) from the experimental data, shown in Figure 1, show a clear decrease, and hence an increase in the agreement with experimental data, with the addition of more information into the system. Both force fields, CHARMM22* in blue and AMBER99SB in red show a comparably good agreement with chemical shifts, with comparable trends in the per-residue deviations of NH, Hα, Cα and Cβ and more marked differences in the case of C' and HN chemical shifts (cf. Figures S21-S22). Both force fields show a very good agreement with $^3J_{H_\alpha-N}$, CHARMM22* describes the χ_2 angle of W5 well, (i.e $^3J_{C-C\gamma}$), while both agree less well with the χ_1 angle of the same residue as well as the $^3J_{H_\alpha-HN}$ and the RDCs. For the case of chemical shifts, the per-residue comparison shows comparable trends for some observables (i.e. $^3J_{C-C\gamma}$, $^3J_{N-C\gamma}$ and N-H RDCs) and more marked differences for others (i.e. $^3J_{H_\alpha-N}$, $^3J_{H_\alpha-HN}$ and Cα-Hα RDCs). These differences suggest that the two force fields are not giving an equivalent description of the peptide and that the addition of experimental information could actually improve them.

M&M ensembles including chemical shifts (that are not expected to contribute particularly given the already good agreement) and ^{3}J-couplings have indeed a positive effect on the RMSDs of all data, including in particular RDCs (cf. Fig. 1), where the improvement is more pronounced for AMBER99SB than for CHARMM22*. Interestingly, the per-residue trends are now also more comparable, with AMBER99SB showing an overall better agreement with all the available data (cf. Figures S23-S24). Finally, the M&M ensembles also including the RDCs showed comparable, good, agreements (Fig. 1), and comparable trends for all data (cf. Figures S25-S26). This suggests that while the original force fields were providing two alternative and not completely satisfactory descriptions of the dynamics of the peptide under study, the M&M ensembles could instead provide ensembles that are indistinguishable from the point of view of the available experimental observables.

The effect of Metainference on the experimental data can be also observed at finer detail in

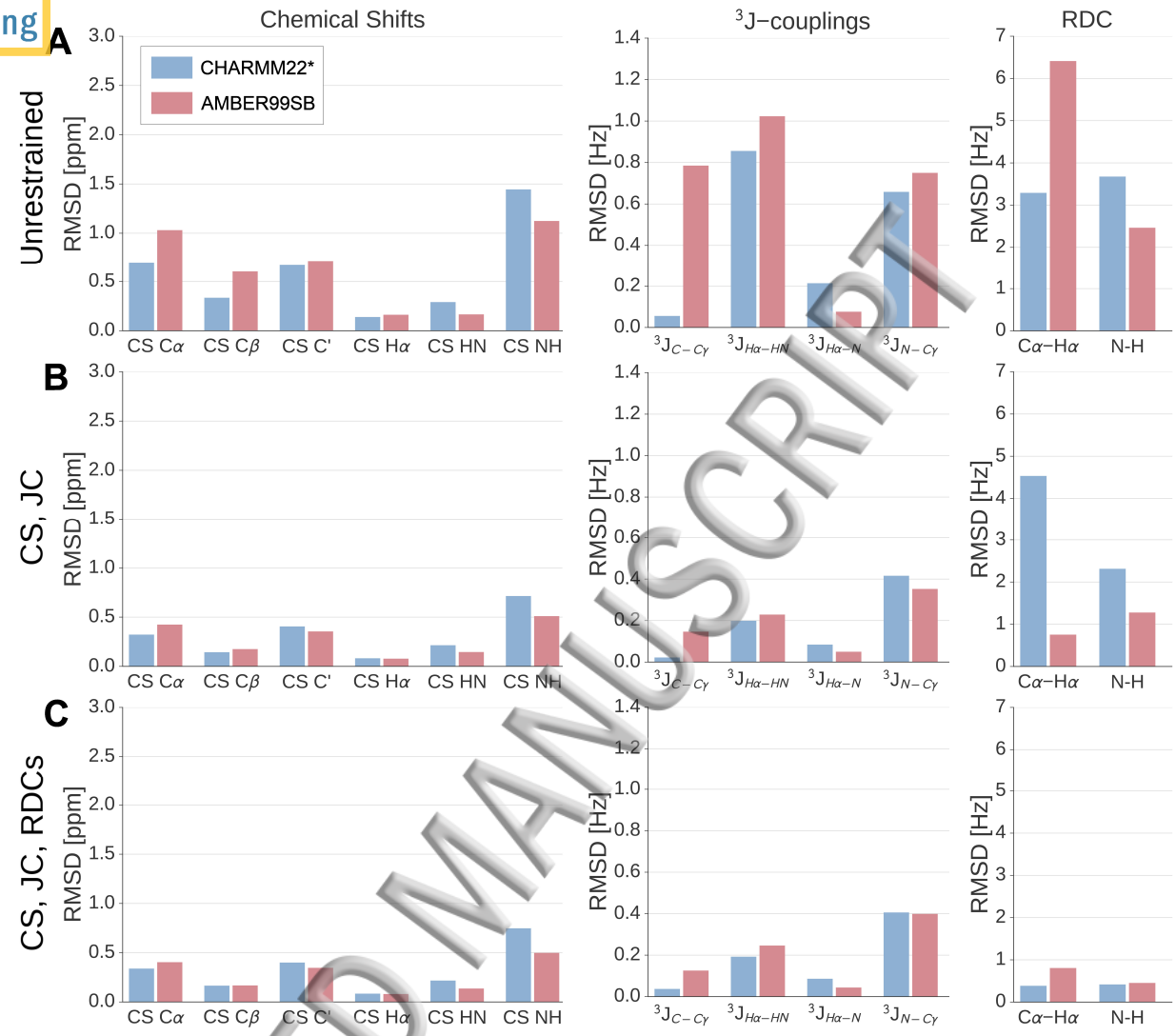


FIG. 1. Root-mean-square deviations between simulated and experimental data for all simulations performed in explicit solvent. Transparent blue bars represent ensembles based on the CHARMM22* prior, while transparent red bars represent ensembles based on the AMBER99SB in TIP4P-D water prior. Each row indicates successive addition of data to the simulation, A) Unrestrained, B) addition of Chemical shifts and ^{3}J-couplings, C) further addition of RDCs. Fully restrained simulations (C) show consistent improvement when compared to unrestrained simulations (A). Addition of RDC restraints has little additional impact on the quality of other experimental observables, while addition of chemical shifts and ^{3}J-couplings has a positive impact on the quality of the RDCs. Per residue comparisons for all data and ensembles can be found in Figures S21-S26.

Figure 2 where the distributions of the Cα-Hα RDCs and those for the Cα carbon chemical shifts are compared. In the case of chemical shifts the distributions were already similar between the two unrestrained force fields, with the exception of A6. Upon restraining the chemical shifts and ^{3}J-couplings, the distributions are translated closer to the reference experimental values but their overall shape is unchanged. The further addition of RDCs does not have any additional effect on the chemical shifts. For RDCs the unrestrained simulations show very broad distributions with average values far from the experimental data. Furthermore, the overall shape of the distribution can be very different for the two force fields as is the case for W5, A6 and A7. Upon restraining with chemical shifts and ^{3}J-couplings the RDCs for W5, A6 and A7 showed an improvement in the agreement with the experimental data and an improved similarity between the two ensembles. The final addition of RDCs restraints shrunk the distributions and translated them closer to the reference experimental values making the overall shape comparable between the two ensembles as visible from the quartiles. Once again the final distributions are not only similar in the average value but also in their quartiles, suggesting that the two final ensembles obtained by restraining CHARMM22* and AMBER99SB in TIP4P-D are not only in good agreement with the experimental values but also very similar to each other.

B. Convergence towards a common ensemble.

While the comparison with the experimental data suggests that it is possible to use M&M to generate ensembles, starting from two alternative priors encoded in the two employed force fields – which are in remarkable good and similar agreement with the experimental data – it is still possible at least in principle that the ensembles could give different results if observed through other techniques. In order to test the hypothesis that M&M can provide at least in principle a unique ensemble, we analyzed the similarities of the ensembles with respect to other independent properties.

In Figure 3 the ensembles are compared using two alternative similarity metrics. In the left panel the probability distributions of the radius of gyration for all the pairs of ensembles are compared, their dissimilarity is measured by the Jensen-Shannon divergence (see Analysis). None of the employed experimental data is a direct measure of the radius of gyration, which makes this a good candidate for an observable that can reveal differences between the

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

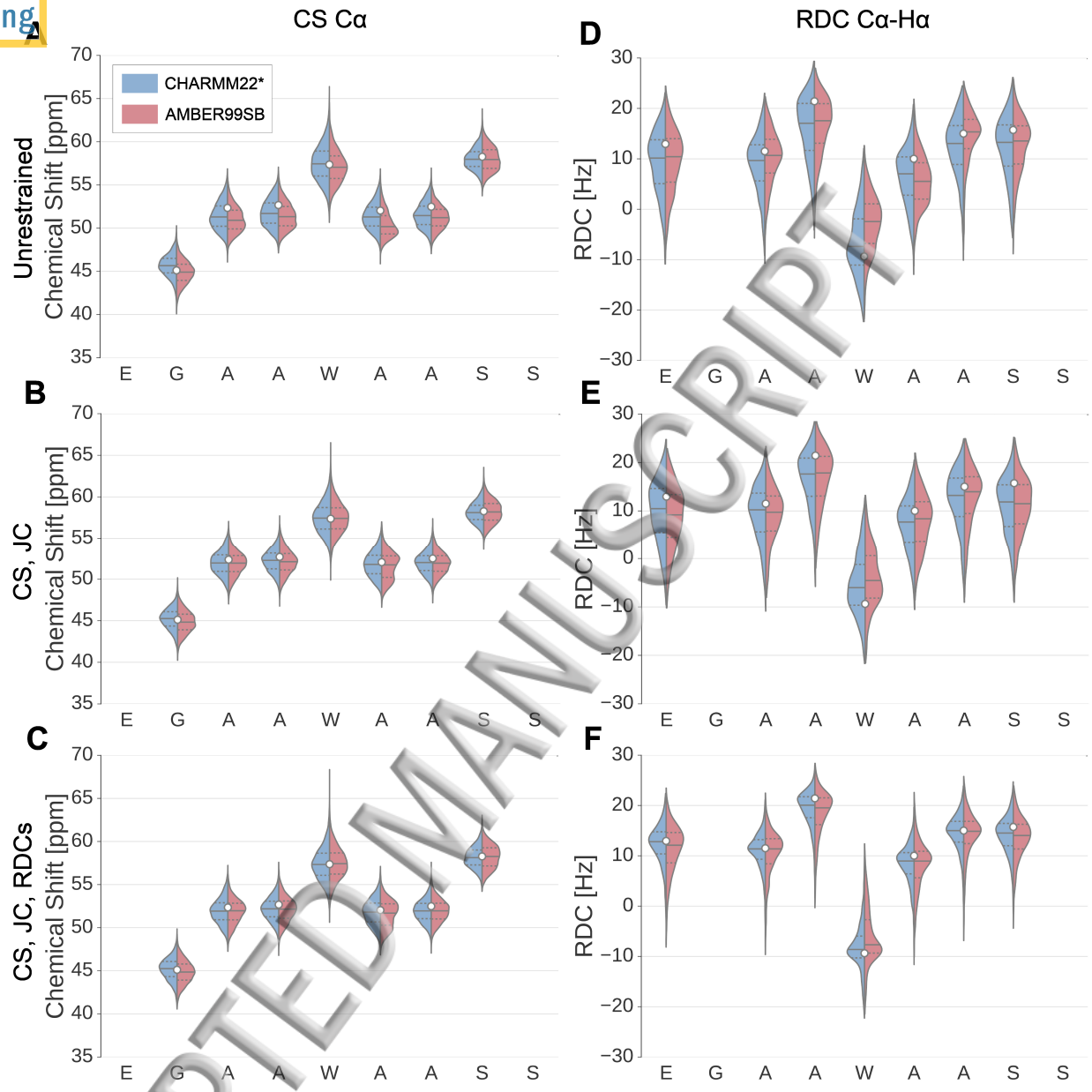


FIG. 2. Violinplots showing the probability distributions of the restrained experimental data per residue without the addition of data (top row, A and D) and with addition of chemical shifts, ^{3}J-couplings (middle row, B and E) and chemical shifts, ^{3}J-couplings and RDCs (bottom row, C and F). Transparent blue distributions represent ensembles based on the CHARMM22* prior, while transparent red distributions represent ensembles based on the AMBER99SB in TIP4P-D water prior. Means and quartiles are indicated by full and dashed lines respectively, while the true experimental value is shown as a dot. Left panels (A, B, C) show the comparison of Cα carbons chemical shifts, right panels (D, E, F) show the comparison of Cα-Hα RDCs .

unrestrained ensembles. With the addition of information into the system, convergence towards a common distribution is remarkably visible (bottom right). While the unrestrained simulations show a remarkably different behaviour, with the CHARMM22* ensemble being more compact than the AMBER99SB one, both start developing a pronounced peak at about 0.85 nm with the introduction of additional information. This is visible in the form of an increased overlap between the two distributions. The Jensen-Shannon divergence confirms the visual suggestion. Of notice is that the probability distribution for the AMBER99SB ensemble, once updated with chemical shifts and ^{3}J-couplings, seems to be already converged, in line with the good agreement of this ensemble with RDCs (cf. Fig. 1). A posteriori one can speculate that the experimental data employed, even if reporting about local quantities, includes indirect information about the extended state of the peptide. Since AMBER99SB with TIP4P-D prior already provides an extended ensemble, less data is needed to converge the distribution overall.

The distance matrices (Figure 3B) also support the notion of convergence of the ensembles towards a unique indistinguishable one. In this case we compared the average distances among all residues calculated using the centre of mass of the residues. Again none of the experimental data report on such high-resolution information. Here the deviation decreases in a fashion corresponding with the Jensen-Shannon divergence in the left panel. To further stress the differences between the original ensembles with respect to the similarities of the final ensemble, the distributions for the overall backbone dihedral similarities and the end-to-end distance are shown in the Supporting Information (Fig. S29).

It is of notice that unrestrained CHARMM22* and fully restrained AMBER99SB are closer in similarity than the two unrestrained simulations and vice versa, as shown by both analysis reported in Figure 3. This again supports the notion of a funnelled picture towards a common unique ensemble that is independent from the prior knowledge, and also suggests how state-of-the-art force fields seem to be converging to such a unique ensemble from different starting points.

C. Ensemble determination in implicit solvent.

While the simulations discussed above show strong differences when not restrained, both use priors of similar quality with respect to experimental measures. An interesting question

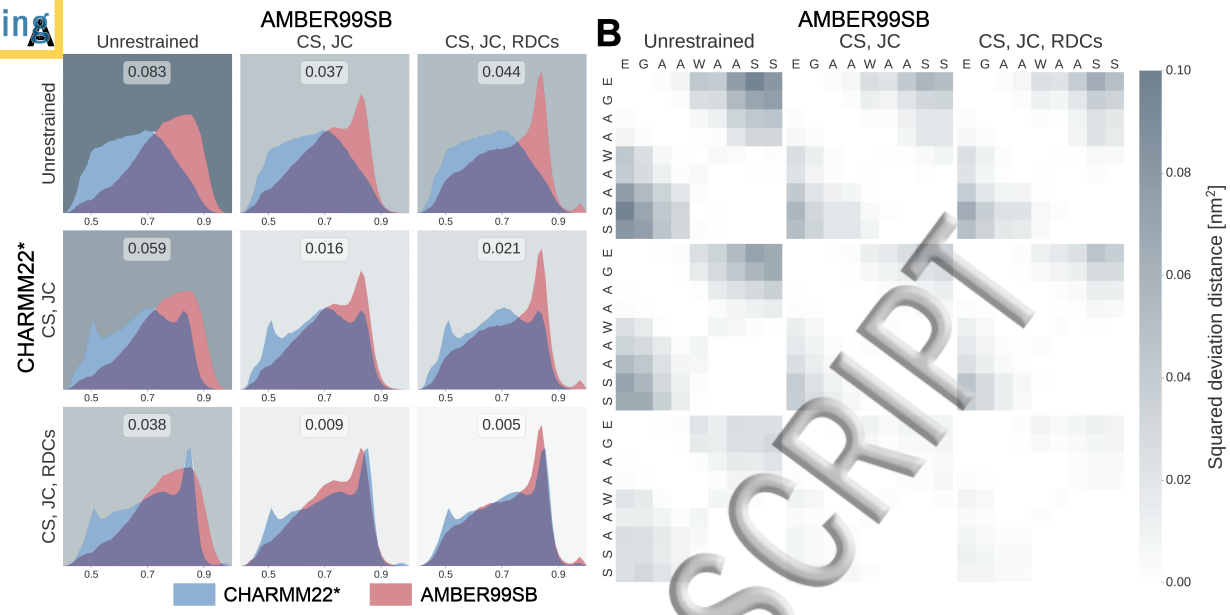


FIG. 3. Convergence towards a unique ensemble. The two panels show two measures of ensembles-dissimilarity. A) Ensemble convergence shown as Jensen-Shannon divergences between the probability distributions of the radius of gyration for all pairs of simulations in explicit solvent given in nat (natural unit of information entropy). Transparent blue distributions represent ensembles based on the CHARMM22* prior, while transparent red distributions represent ensembles based on the AMBER99SB in TIP4P-D water prior. Lighter backgrounds indicate a lower divergence and thus higher similarity. The improvement is subjectively noticeable by comparing the underlying probability distributions. B) Squared-deviation inter-residue distance matrices between each pair of simulations. Lower distance deviations correspond to a higher degree of similarity.

is therefore how a restrained simplified prior fares with respect to more conventional and more accurate priors. To this end, we performed M&M simulations using the computationally very efficient implicit solvent model EEF1-SB (cf. Table I). The combined results can be seen in Figure 4. The unrestrained ensemble is in relative good agreement with RDCs while showing a worse agreement with the other data than the explicit solvent unrestrained ensembles, suggesting that RDCs are better captured by the extended description of the peptide resulting from this prior (cf. Figure 4B). As expected the root-mean-square deviations showed a marked decrease with the addition of experimental data. The restrained ensemble is then, when compared with Figure 1, at least on par with the unrestrained explicit solvent simulations. The probability distribution of the radius of gyration (Figure 4B)

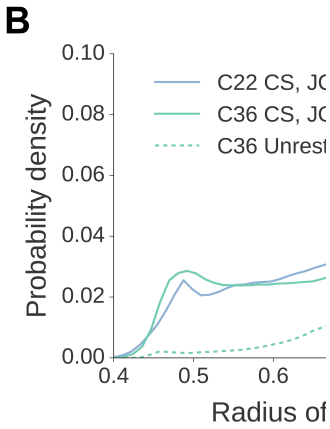


FIG. 4. Results of M&M simulations using the EEF1-SB implicit solvent model. A) Root-mean-square deviations of the unrestrained and fully restrained implicit solvent simulations. B) Probability distributions of the radius of gyration for unrestrained and restrained implicit solvent as well as fully restrained CHARMM22* simulations. The Jensen-Shannon divergences of the probability distributions of the radius of gyration between CHARMM36 and CHARMM22* with full restraints and CHARMM36 with restraint and unrestrained CHARMM36 are 0.015 nat and 0.203 nat (natural unit of information entropy) respectively. C) Squared-deviation inter-residue distance matrix between fully restrained CHARMM22* and fully restrained CHARMM36 EEF1-SB simulations.

shows the dramatic effect of the restraint. The prior is clearly biased towards very open states while the explicit solvent simulations show a more balanced picture (Figure 3A). This bias is not surprising given that EEF1-SB was explicitly optimised for disordered systems[24]. The restrained implicit solvent simulation is able to alleviate the over-extended description provided by the prior even if without reproducing quantitatively the pronounced peak at 0.8 nm present in CHARMM22*. Nonetheless, the Jensen-Shannon divergence of 0.015

between fully restrained CHARMM22* and fully restrained CHARMM36 with EEF1-SB support their overall similarity, indeed this divergence is very similar to the one observed between the partially restrained explicit solvent simulations. Finally we compared the alternative similarity metric in the form of inter-residue distances (Figure 4C), which are practically indistinguishable.

Disordered systems are often characterised by transient secondary structures and elongated conformations. In order to further test the similarity of the ensembles in explicit and implicit solvent we calculated the secondary structure populations over the ensembles using STRIDE[54] with errors estimated using the block standard error approach[55]. While all ensembles show poor helical and hairpin content (cf. Fig. S30-S31), a convergence of the conformational space as a function of the addition of experimental information towards a common ensemble for both the explicit and implicit solvent simulations can be appreciated for the turn and coil content (cf. Fig. S32 and S33, respectively). The unrestrained AMBER99SB as well as the unrestrained CHARMM36 EEF1-SB ensembles show a lower turn content than the CHARMM22* unrestrained ensemble and an opposite behavior for the coil content. The successive addition of experimental data brings the AMBER99SB, CHARMM22* and CHARMM36 EEF1-SB fully restrained ensembles to show essentially the same turn content and the same coil content.

These results further enforce the notion that M&M allows a radical reshaping of the prior ensemble to a common solution that is consistent across vastly different priors.

IV. CONCLUSIONS

The last few years have seen a large increase in the assessment of force fields. While there is a clear trend in the improvement of force field quality, their transferability between disordered and ordered systems and the robustness of the resulting structural ensembles for disordered systems is often questioned[21,56–59]. To circumvent force field limitations as well as limitations in the resolution of experimental techniques, hybrid methods based on the integration of experimental data and molecular dynamics simulations have seen a huge growth[1,11]. Here, we first simplified the setup of Metadynamic Metainference simulations to make them essentially parameter free, and extended the formalism to account for experimental data that are defined modulo a constant. Then we studied a disordered peptide

understand two concepts: First, to which extent, given enough experimental data, it is possible to obtain ensembles of structures that do not depend explicitly on the molecular mechanics force field employed, and second if it possible to obtain results of comparable quality at a fraction of the computational cost. By comparing two state-of-the-art explicit-solvent force fields and integrating them with multiple sources of experimental data we determined two ensembles that are essentially indistinguishable from each other and different from those obtained using the force fields alone. Furthermore, results of comparable quality have been obtained using M&M, the same data and a very inexpensive implicit solvent force field.

SUPPLEMENTARY MATERIALS

See supplementary material for convergence tests, the free energy profiles as a function of the collective variables employed in metadynamics and more comparisons with experimental data.

ACKNOWLEDGMENTS

The authors acknowledge the support by the Technische Universität München - Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement n. 291763. We gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de). We thank Massimiliano Bonomi, Giovanni Bussi, Bernd Reif and Michele Vendruscolo for useful discussions and suggestions.

REFERENCES

[1]M. Bonomi, G. T. Heller, C. Camilloni, and M. Vendruscolo, Curr. Opin. Struct. Biol. **42**, 106 (2017).

[2]G. Wei, W. Xi, R. Nussinov, and B. Ma, Chem. Rev. **116**, 6516 (2016).

[3]V. Venditti, T. K. Egner, and G. M. Clore, Chem. Rev. **116**, 6305 (2016).

[4]L. D. Cabrita, A. M. E. Cassaignau, H. M. M. Launay, C. A. Waudby, T. Wlodarski, C. Camilloni, M.-E. Karyadi, A. L. Robertson, X. Wang, A. S. Wentink, L. S. Goodsell,

G. A. Woolhead, M. Vendruscolo, C. M. Dobson, and J. Christodoulou, Nat. Struct. Mol. Biol. **23**, 278 (2016).

[5]S. Milles, D. Mercadante, I. V. Aramburu, M. R. Jensen, N. Banterle, C. Koehler, S. Tyagi, J. Clarke, S. L. Shammas, M. Blackledge, F. Gräter, and E. A. Lemke, Cell **163**, 734 (2015).

[6]K. Lindorff-Larsen, R. B. Best, M. A. Depristo, C. M. Dobson, and M. Vendruscolo, Nature **433**, 128 (2005).

[7]A. Ramanathan, A. Savol, V. Burger, C. S. Chennubhotla, and P. K. Agarwal, Acc Chem Res **47**, 149 (2014).

[8]H. J. Dyson and P. E. Wright, Nat Rev Mol Cell Biol **6**, 197 (2005).

[9]J. Habchi, P. Tompa, S. Longhi, and V. N. Uversky, Chem. Rev. **114**, 6561 (2014).

[10]E. Papaleo, G. Saladino, M. Lambrughi, K. Lindorff-Larsen, F. L. Gervasio, and R. Nussinov, Chem. Rev. **116**, 6391 (2016).

[11]E. Ravera, L. Sgheri, G. Parigi, and C. Luchinat, Phys. Chem. Chem. Phys. **18**, 5686 (2016).

[12]M. Schor, A. S. J. S. Mey, and C. E. MacPhee, Biophys Rev **8**, 429 (2016).

[13]M. Bonomi, C. Camilloni, A. Cavalli, and M. Vendruscolo, Science Advances **2**, e1501177 (2016).

[14]M. Bonomi, C. Camilloni, and M. Vendruscolo, Sci. Rep. **6**, 31232 (2016).

[15]A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002).

[16]O. Valsson, P. Tiwary, and M. Parrinello, Ann. Rev. Phys. Chem. **67**, 159 (2016).

[17]S. A. Dames, R. Aregger, N. Vajpai, P. Bernado, M. Blackledge, and S. Grzesiek, J. Am. Chem. Soc. **128**, 13508 (2006).

[18]S. Piana, K. Lindorff-Larsen, and D. E. Shaw, Biophys. J. **100**, L47 (2011).

[19]S. Piana, A. G. Donchev, P. Robustelli, and D. E. Shaw, J. Phys. Chem. B **119**, 5113 (2015).

[20]H. T. A. Leung, O. Bignucolo, R. Aregger, S. A. Dames, A. Mazur, S. Bernèche, and S. Grzesiek, J. Chem. Theory Comput. **12**, 383 (2016).

[21]F. Palazzesi, M. K. Prakash, M. Bonomi, and A. Barducci, J. Chem. Theory Comput. **11**, 2 (2015).

[22]R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, J. Chem. Theory Comput. **8**, 3257 (2012).

[23]T. Lazaridis and M. Karplus, Proteins: Structure, Function, and Genetics **35**, 133 (1999).

[24]S. Bottaro, K. Lindorff-Larsen, and R. B. Best, J. Chem. Theory Comput. **9**, 5641 (2013).

[25]W. Rieping, M. Habeck, and M. Nilges, Science **309**, 303 (2005).

[26]C. Camilloni, A. Cavalli, and M. Vendruscolo, J. Chem. Theory Comput. **9**, 5610 (2013).

[27]C. Camilloni and M. Vendruscolo, J. Am. Chem. Soc. **136**, 8982 (2014).

[28]J. Pfaendtner and M. Bonomi, J. Chem. Theory Comput. **11**, 5062 (2015).

[29]G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, Comput. Phys. Commun. **185**, 604 (2014).

[30]A. Cavalli, C. Camilloni, and M. Vendruscolo, J. Chem. Phys. **138**, 094112 (2013).

[31]B. Roux and J. Weare, J. Chem. Phys. **138**, 084107 (2013).

[32]D. F. Gatz and L. Smith, Atmospheric Environment **29**, 1195 (1995).

[33]C. Camilloni and M. Vendruscolo, J. Phys. Chem. B **119**, 653 (2015).

[34]M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX **1-2**, 19 (2015).

[35]W. Humphrey, A. Dalke, and K. Schulten, J Mol Graph **14**, 33 (1996).

[36]G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[37]M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).

[38]V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, Proteins **65**, 712 (2006).

[39]A. D. Mackerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998).

[40]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[41]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).

[42]C. L. Wennberg, T. Murtola, B. Hess, and E. Lindahl, J. Chem. Theory Comput. **9**, 3527 (2013).

[43]B. Hess, J. Chem. Theory Comput. **4**, 116 (2008).

[44]A. Barducci, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **100**, 20603 (2008).

[45]P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, J. Phys. Chem. B **110**, 3533 (2006).

[46]K. Kohlhoff, P. Robustelli, A. Cavalli, X. Salvatella, and M. Vendruscolo, J. Am. Chem. Soc. **131**, 13894 (2009).

[47]C. Camilloni, P. Robustelli, A. De Simone, A. Cavalli, and M. Vendruscolo, J. Am. Chem. Soc. **134**, 3968 (2012).

[48]M. Karplus, J. Chem. Phys. **30**, 11 (1959).

[49]C. Pérez, F. Löhr, H. Rüterjans, and J. M. Schmidt, J. Am. Chem. Soc. **123**, 7081 (2001).

[50]B. Vögeli, J. Ying, A. Grishaev, and A. Bax, J. Am. Chem. Soc. **129**, 9377 (2007).

[51]A. N. Borkar, M. F. Bardaro Jr., C. Camilloni, F. A. Aprile, G. Varani, and M. Vendruscolo, Proc. Natl. Acad. Sci. U.S.A. **113**, 7171 (2016).

[52]A. N. Borkar, P. Vallurupalli, C. Camilloni, L. E. Kay, and M. Vendruscolo, Phys. Chem. Chem. Phys. **19**, 2797 (2017).

[53]D. Branduardi, G. Bussi, and M. Parrinello, J. Chem. Theory Comput. **8**, 2247 (2012).

[54]D. Frishman and P. Argos, Proteins **23**, 566 (1995).

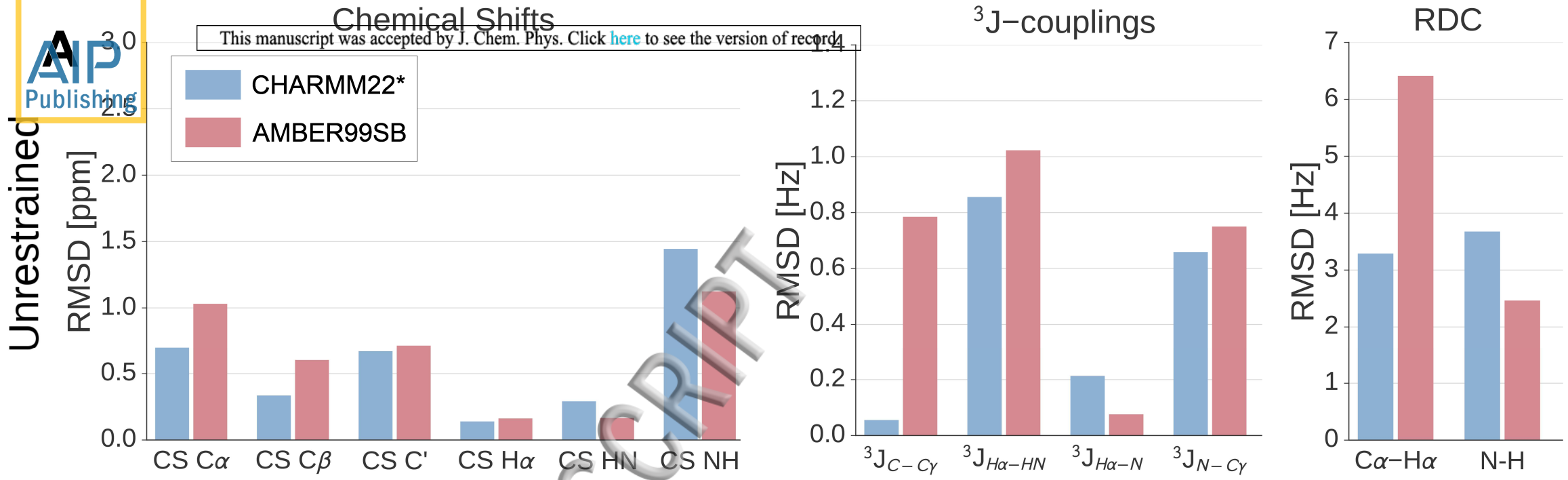
[55]H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).

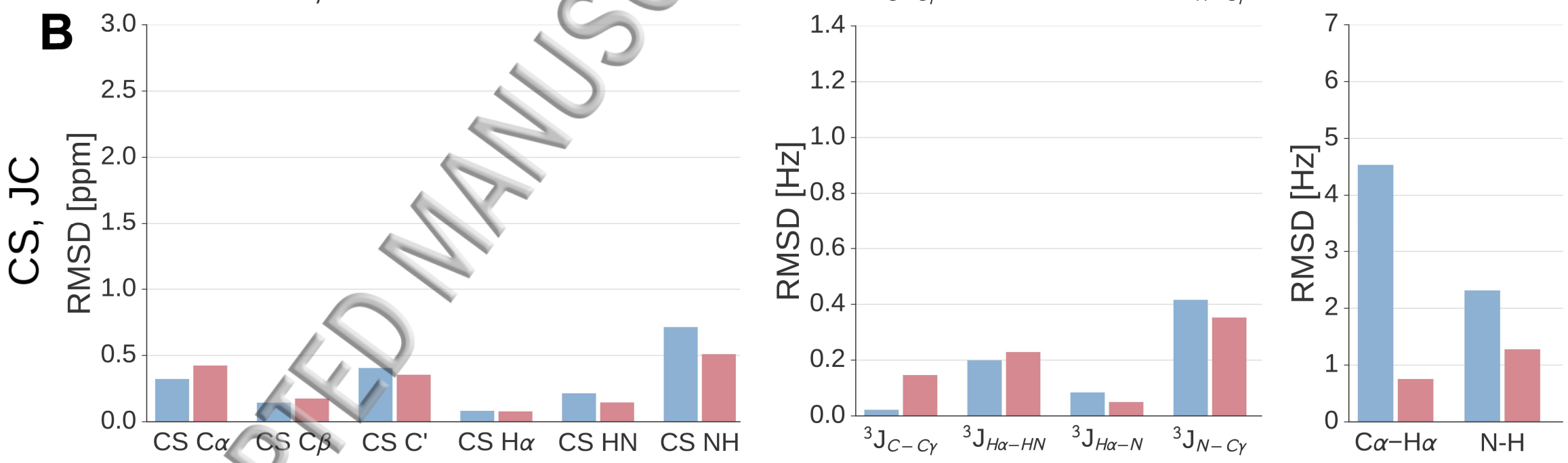
[56]S. Rauscher, V. Gapsys, M. J. Gajda, M. Zweckstetter, B. L. de Groot, and H. Grubmüller, J. Chem. Theory Comput. **11**, 5513 (2015).

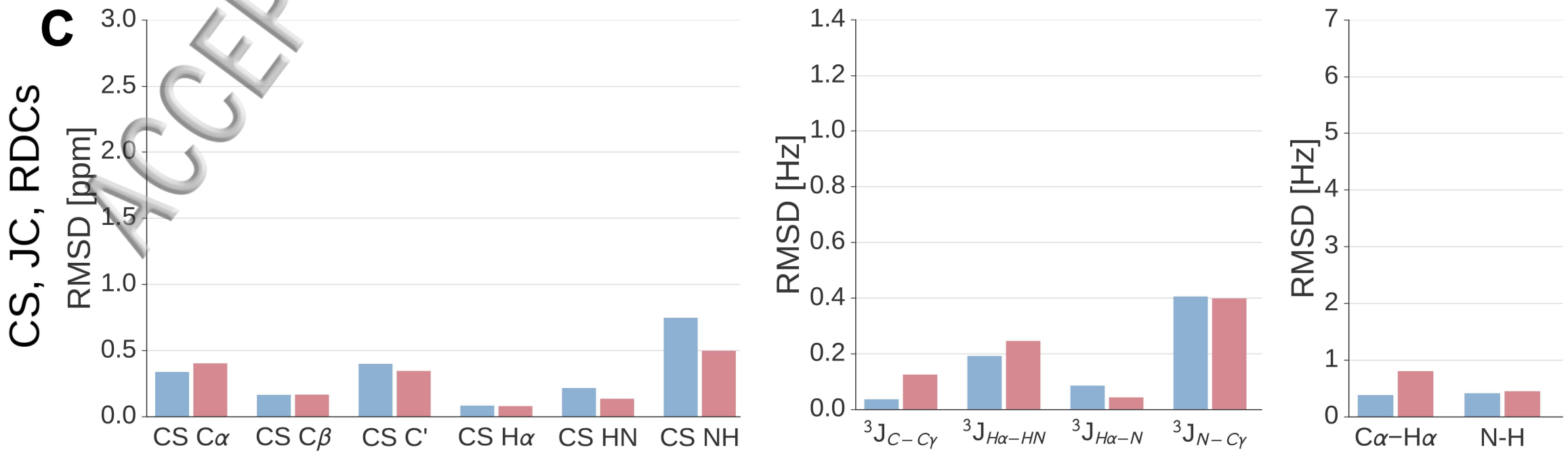
[57]F. Martín-García, E. Papaleo, P. Gomez-Puertas, W. Boomsma, and K. Lindorff-Larsen, PLoS ONE **10**, e0121114 (2015).

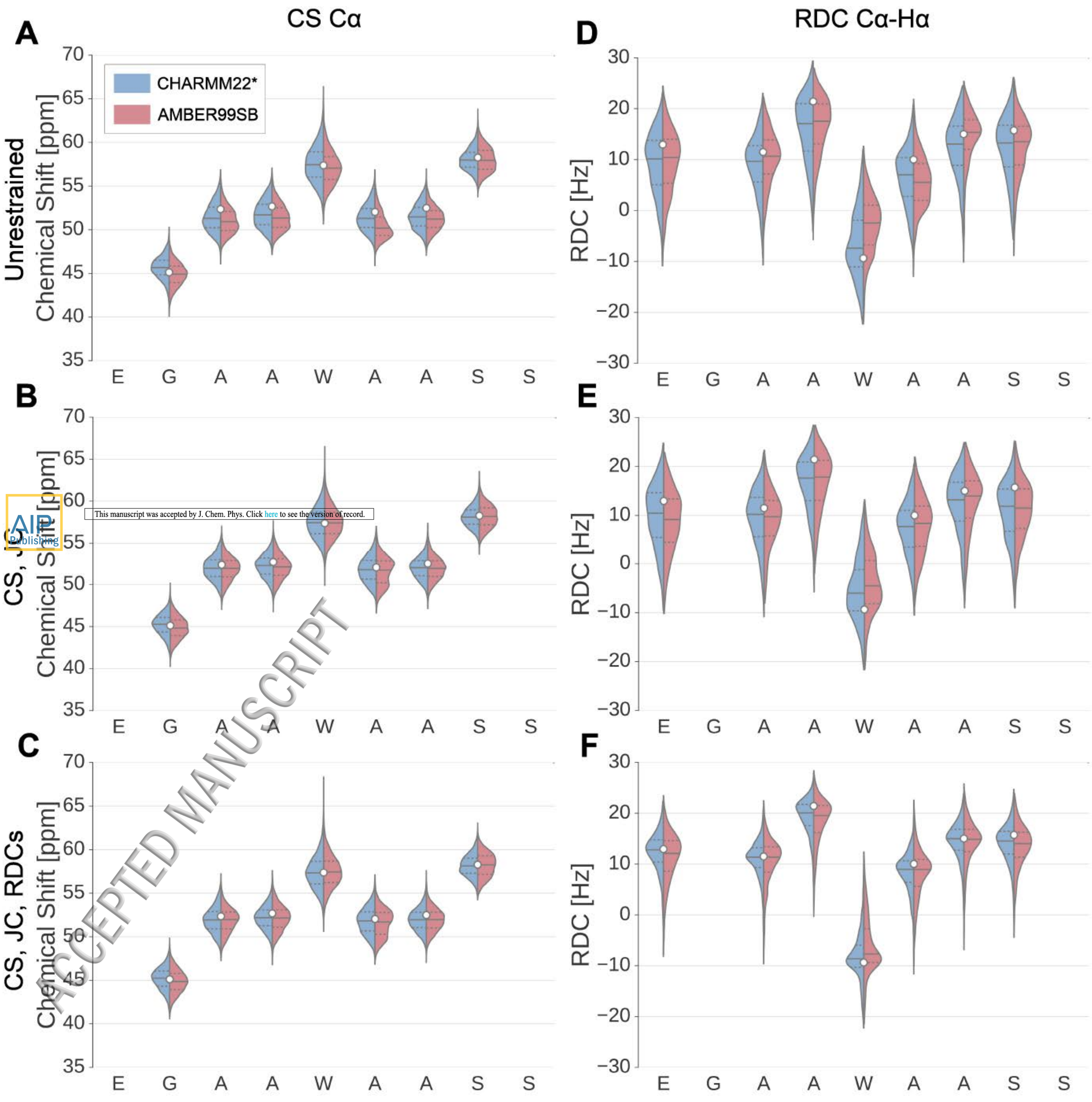
[58]F. Vitalini, A. S. J. S. Mey, F. Noé, and B. G. Keller, J. Chem. Phys. **142**, 084101 (2015).

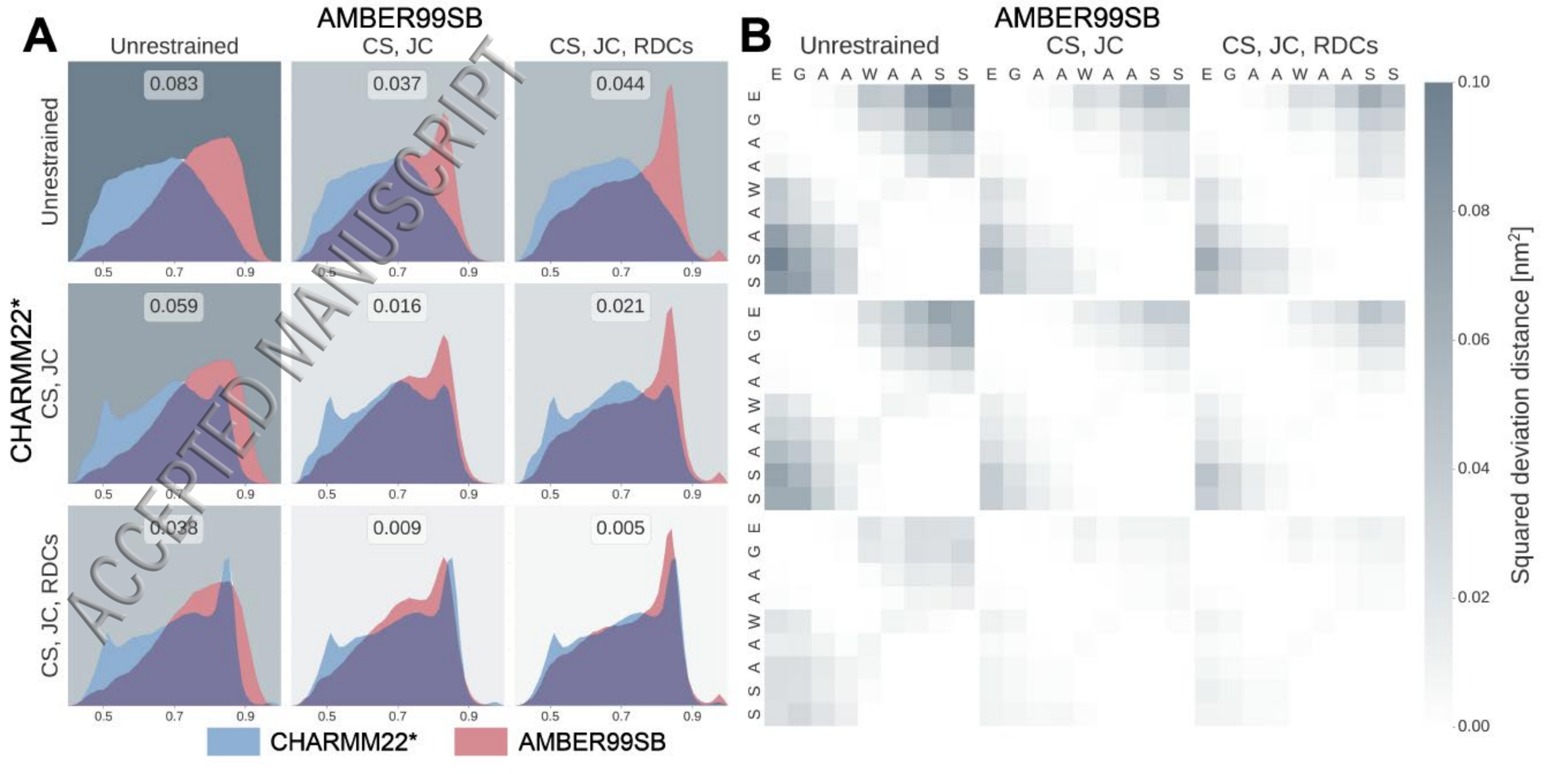
[59]S. Piana, J. L. Klepeis, and D. E. Shaw, Curr. Opin. Struct. Biol. **24**, 98 (2014).



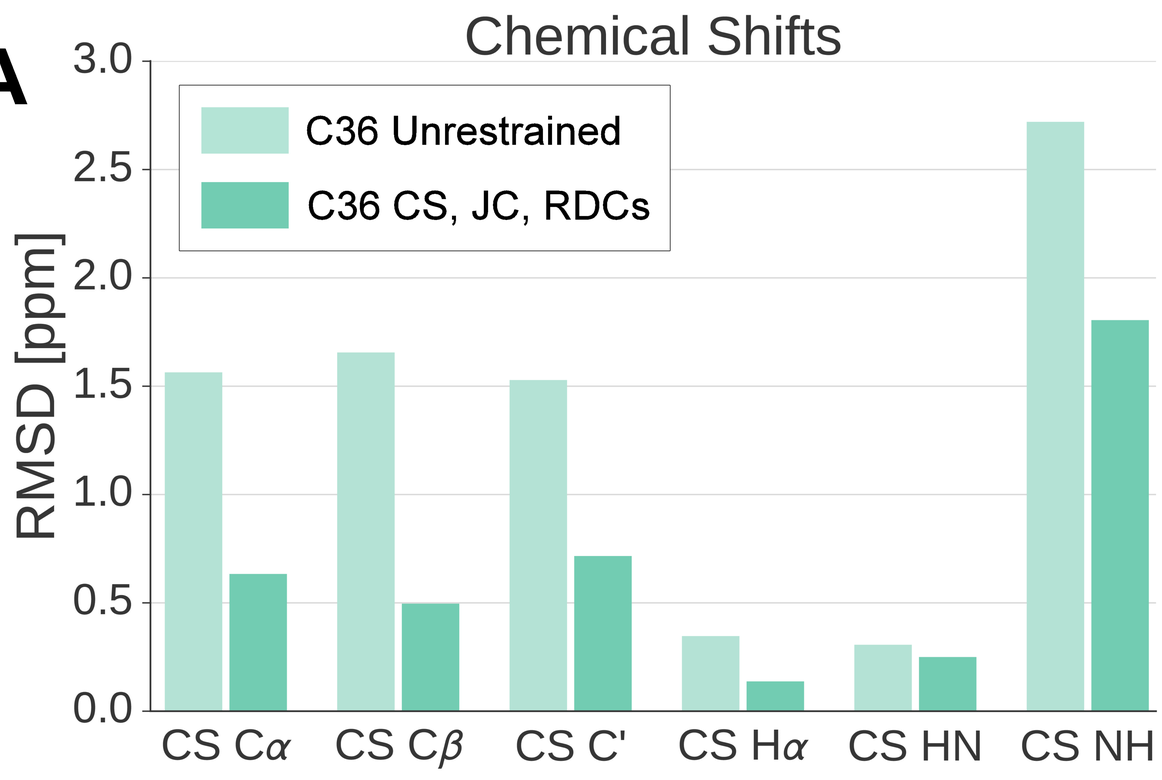


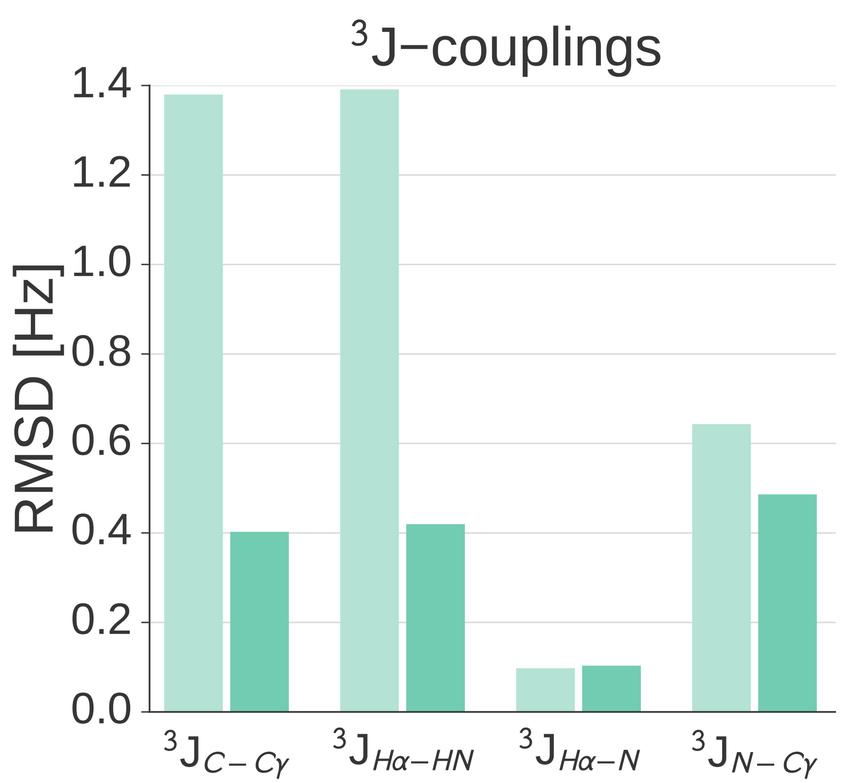


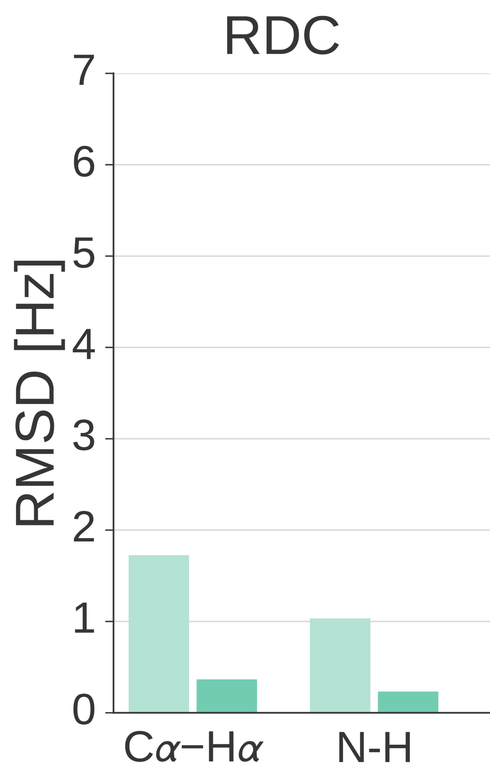




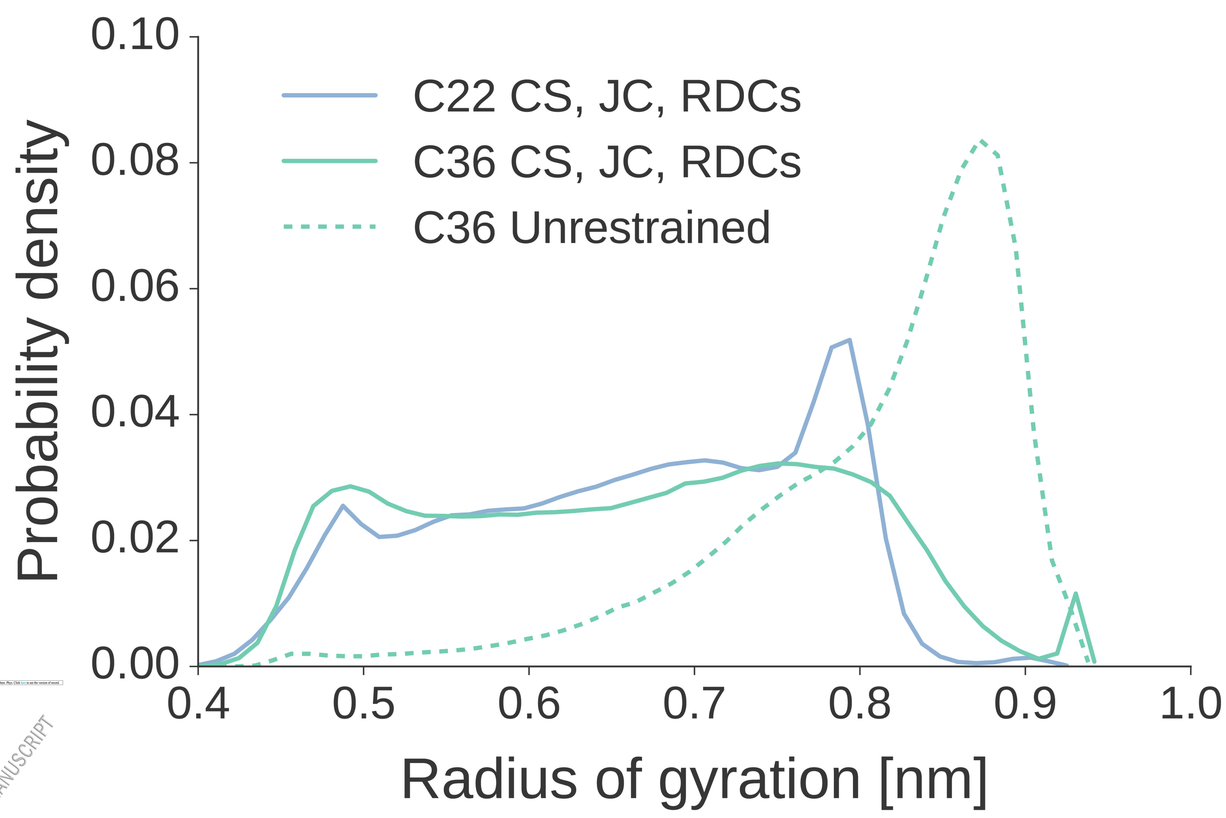
A







B



C