Conformational Entropies and Order Parameters: Convergence, Reproducibility, and Transferability

Samuel Genheden,[†] Mikael Akke,[‡] and Ulf Ryde*,[†]

[†]Department of Theoretical Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

[‡]Department of Biophysical Chemistry, Center for Molecular Protein Science, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

S *Supporting Information*

ABSTRACT: Conformational entropy provides major contributions to protein folding and functions, such as ligand binding, making it a potentially important driver of biologically relevant processes. NMR spectroscopy is a unique technique to estimate conformational entropy changes at atomic resolution, an approach that can be favorably augmented by comparisons with results from molecular dynamics (MD) simulations, for example, by generating an order-parameter-to-entropy dictionary. Here, we address critical issues pertaining to such an approach, including reproducibility, convergence, and transferability by analyzing long (380 ns −1 ms) MD trajectories obtained for five different proteins. We observe that order parameters and conformational entropies calculated over 10−100 ns windows are typically well converged among individual MD trajectories and reproducible between pairs of independent trajectories, when calculated on a per-residue level. However, significant discrepancies sometimes arise for the total conformational entropy evaluated as the sum of the residue-specific entropies, especially in cases that involve rare transitions to alternative conformational states. Furthermore, we find that the order-parameter-to-entropy dictionary depends strongly on the protein and the sampling frequency, but much less so on the molecular dynamics force field. Thus, the transferability of the dictionary is poor between proteins but relatively good between different states (e.g., different ligand-bound complexes) of the same protein, provided that a protein-specific dictionary has been derived.

NMR order-parameter → entropy dictionary depends on protein & sampling frequency

$$S = R\,n\,[intercept + slope\,f(1 − O^2)]$$
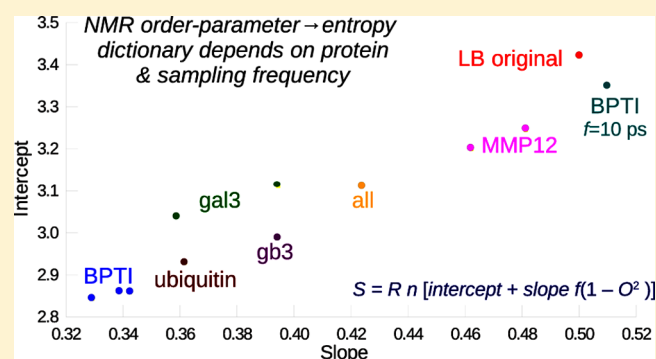
■ INTRODUCTION

Changes in conformational entropy have been shown to play a major role in several cases of biomolecular recognition and therefore much effort has been invested in method development to quantify such entropies.[1−4] NMR relaxation provides a unique experimental source of information on the motional amplitudes of individual bond vectors, quantified in terms of order parameters, which can be related to conformational entropy via various analytical relationships,[5−8] although a direct one-to-one relationship does not exist in general.[9] However, a limitation of this approach is that NMR relaxation experiments typically only give information on a restricted number of bond vectors and sample fluctuation amplitudes of motions occurring on time scales shorter than the overall rotational correlation time. By measuring order parameters from residual dipolar couplings (RDC), it is possible to access longer time scales, which are limited in this case by the inverse of the residual dipolar coupling constant. An attractive approach is to derive the total conformational entropy from molecular dynamics (MD) simulations that have been validated against experimental order parameters, thereby lifting the limitations of the accessible degrees of freedom in NMR experiments.

Recently, Li and Brüschweiler[7] suggested an approach to interpret restricted sets of NMR order parameters in terms of conformational entropies of entire amino-acid residues. Their main result is a dictionary of linear or logarithmic relationships between order parameters and conformational entropies, derived by calculating both quantities from MD simulations of three small and highly stable proteins, namely bovine pancreatic trypsin inhibitor (BPTI), ubiquitin (ubq), and calbindin D9k. However, this elegant approach requires further testing to ensure that it is robust and generally applicable to proteins and to assess whether the dictionary is transferable between proteins. We recently showed that it is extremely hard to reach convergence of entropies in MD simulations,[10] an observation that questions also the dictionary-based approach, which necessarily needs to be based on converged entropies. A related issue is the convergence of the order parameters derived in long MD simulations, which so far has been addressed only sparingly.[11] In this paper, we address these issues by analyzing previously published MD trajectories ranging in length from 380 ns to 1 ms.

Table 1. Summary of the Analyzed Trajectories

protein	ligand	length (μs)	sampling interval (ps)	force field protein/ligand	solvent	ref
mmp12	cn1h	0.38	10	Amber99SB/GAFF	TIP4P-Ew	10
mmp12	cn2h	0.38	10	Amber99SB/GAFF	TIP4P-Ew	10
galectin-3	lactose	0.50	10	Amber99SB/GAFF	TIP4P-Ew	10
galectin-3	l02	0.50	10	Amber99SB/GAFF	TIP4P-Ew	10
BPTI A		0.5	250	Amber99SB	TIP4P-Ew	10
BPTI B		0.5	250	Amber99SB	TIP4P-Ew	10
BPTI C		1000	250	Amber99SB-ILDN	TIP4P-Ew	12
gb3		10	200	Amber99SB-ILDN	TIP3P	13
ubiquitin		10	200	Amber99SB-ILDN	TIP3P	13

Table 2. Convergence and Reproducibility of Order Parameters and Entropies[a]

protein	set 1	set 2	O^2 MAD	O^2 MAX	dictionary entropy MAD	dictionary entropy MAX	dictionary entropy SUM	histogram entropy MAD	histogram entropy MAX	histogram entropy SUM
					Backbone					
BPTI	500 ns A	500 ns B	0.01	0.08	0.1	0.7	−3.8	0.1	0.4	−3.0
	500 ns A	500 ns A								
BPTI	f = 10 ps	f = 250 ps	0.00	0.01	0.0	0.1	0.5	0.8	3.5	40.6
BPTI	500 ns C	1 ms C	0.03	0.22	0.3	2.2	10.0	0.2	1.0	6.0
BPTI	10 μs C	1 ms C	0.02	0.22	0.3	2.2	10.2	0.2	1.1	5.6
BPTI	100 μs C	1 ms C	0.01	0.13	0.1	1.2	0.4	0.1	0.7	1.0
gb3	500 ns	10 μs	0.01	0.10	0.1	0.7	−0.4	0.1	0.2	−1.1
ubiquitin	500 ns	10 μs	0.01	0.04	0.1	0.5	−2.3	0.1	0.4	−2.8
					Side Chains					
BPTI	500 ns A	500 ns B	0.04	0.21	0.3	2.0	−2.4	0.3	2.2	−1.8
	500 ns A	500 ns A								
BPTI	f = 10 ps	f = 250 ps	0.00	0.01	0.0	0.1	2.5	0.4	1.7	17.8
BPTI	500 ns C	1 ms C	0.04	0.12	0.2	0.9	2.0	0.2	1.1	2.2
BPTI	10 μs C	1 ms C	0.03	0.12	0.2	1.1	4.4	0.2	1.0	4.1
BPTI	100 μs C	1 ms C	0.02	0.07	0.1	0.6	−0.2	0.1	0.5	0.2
gb3	500 ns	10 μs	0.02	0.11	0.1	0.7	1.2	0.1	1.0	0.9
ubiquitin	500 ns	10 μs	0.04	0.27	0.3	2.1	−8.1	0.3	1.4	−9.3

[a]The mean absolute deviation (MAD) and the maximum absolute deviation (MAX) are reported between two sets of simulations (two independent simulations or the full and truncated simulation, as indicated). In one case, two different sampling intervals (f) are instead compared. Entropies are in kJ/mol at 300 K and for these, the accumulated sum of the differences for the whole protein is also given (SUM).

METHODS

Trajectories. We have analyzed nine previously published trajectories: Matrix metalloproteinase 12 (mmp12) was simulated for 380 ns with two different ligands bound, denoted cn1h and cn2h (mmp12-cn1h and mmp12-cn2h).[10] Galectin-3 was simulated for 500 ns with either lactose or a synthetic sugar derivative denoted l02 (gal3-lac and gal3-l02).[10] Bovine pancreatic trypsin inhibitor (BPTI) was simulated for 500 ns in two separate simulations started with different initial velocities.[10] We also used a 1-ms BPTI simulation[12] and 10-μs simulations of the third IgG-binding domain of protein G (gb3) and ubiquitin (ubq).[13] The latter three trajectories were generously provided by D. E. Shaw Research. Table 1 outlines the MD simulation protocols used to generate the various trajectories analyzed herein. It should be noted that the simulations used to create the original order-parameter-to-entropy dictionary relied on two different solvent models; explicit SPC/E water was used for calbindin D9k and ubq, but an implicit generalized Born model for BPTI.[14] The simulations analyzed in the present paper were performed using the explicit TIP3P or TIP4P-Ew water models (Table 1). Coordinates were extracted every 250 ps, unless otherwise noted, to be consistent with the trajectories provided by D. E. Shaw Research.

Order Parameters and Entropies. Order parameters were calculated from the MD simulations by the isotropic reorientational eigenmode dynamics (iRED) approach[15] for those bond vectors that were used in the dictionary by Li and Brüschweiler (LB).[7] These vectors will be divided into backbone N−H vectors and side-chain vectors. Entropies were calculated using the same method as by LB,[7] that is, a histogram approach based on the von Mises kernel estimation of the dihedral angle distribution.[16] The side-chain entropies were evaluated as a sum of one-dimensional histograms, whereas the backbone entropies were evaluated as two-dimensional histograms over the φ and ψ angles. Many other methods to compute entropies exist,[17−19] for example, methods that consider correlations between different dihedrals, but as we and others have compared them previously,[10,16] we have chosen to concentrate on the method that was used to build the LB dictionary.

Entropies (S) were also computed from the order parameters (O) based on the LB dictionary

$$S = Rn[m + kf(1 − O^2)] \quad (1)$$

where R is the gas constant, n is the number of dihedral angles that were probed (n is 2 for the backbone and between 1 and 4 for the side chains),[7] m and k are dictionary parameters, taken

from ref 7. $f(x) = \ln(x)$ for the backbone and the side chains of Asp and Glu, whereas $f(x) = x$ for all other side chains.[7] Throughout this paper, entropies are reported in energy units, that is, TS in kJ/mol, at $T = 300$ K.

For a direct comparison with order parameters determined by NMR relaxation methods, the order parameter should be calculated over window lengths matching the overall rotational correlation time or, in the case of sufficiently long trajectories, by directly fitting against the full correlation function calculated from the entire trajectory.[11] However, here, we are mainly concerned with comparisons between entropies calculated directly from the trajectories or indirectly via the original LB dictionary, which used a window size of 100 ns.[7] Thus, in order to be consistent with previous work, we choose to calculate order parameters and entropies using a window size of 100 ns (76 ns for mmp12).

■ RESULTS AND DISCUSSION

In this article, we investigate the precision, convergence and reproducibility of order parameters and conformational entropies in long MD simulations. Furthermore, we test the accuracy and transferability of the recently suggested LB dictionary for the translation of NMR order parameters into residue-specific conformational entropies. To this end, we analyze our own 0.38−0.5 μs MD simulations of gal3, mmp12, and BPTI,[10] as well as the 10−1000 μs MD simulations of BPTI, ubq, and gb3 from D. E. Shaw Research.[12,13] Table 1 provides an outline of the various simulations. We address these questions in separate sections below.

Reproducibility, Precision, and Convergence of Calculated Entropies and Order Parameters. The LB dictionary relates NMR order parameters measured for selected vectors in the protein to total residue entropies calculated from MD simulations with a dihedral histogram approach. For such an approach to meaningful, it is of course essential that the calculated entropies are converged and reproducible. Therefore, we first assess the reproducibility and precision of calculated order parameters and entropies by monitoring the mean absolute pairwise difference (MAD) and the maximum deviation (MAX) between independent trajectories of the same protein and between segments of the same trajectory. These metrics are good indicators of the average and worst-case performance. We have tested several other metrics, but they typically convey similar information in a less lucid way. For example, Pearson's correlation coefficient is >0.90 for all 42 entries in Table 2 and generally shows a strong inverse correlation to both MAD and MAX (cf. Figure S1 in the Supporting Information).

We start by analyzing the differences in calculated parameters between the two 500 ns simulations of BPTI (A and B), as this gives a good indication of the reproducibility of order parameters and entropies. Table 2 shows that the two simulations give quite similar backbone order parameters with a MAD of 0.01 and a MAX of 0.08. The differences of the side-chain order parameters are larger, with a MAD of 0.04 and a MAX of 0.21. Translating these order parameters into entropies using the LB dictionary,[7] we find that the MAD and MAX of the backbone entropies between the two simulations are only 0.1 and 0.7 kJ/mol (all entropies in this article are presented in energy units at 300 K). However, these small residue-wise differences accumulate to a difference of 4 kJ/mol for the entire protein. Again, the differences for the side-chain dictionary entropies are somewhat larger, with MAD and MAX values of

0.3 and 2.0 kJ/mol, but in this case, the residue-wise differences cancel to a greater extent and the accumulated difference for the whole protein is only 2 kJ/mol. We also computed the entropies directly from the MD trajectories using the histogram method. Table 2 shows that the differences in entropy between the two simulations are similar for the two approaches: The histogram method yields MAD values identical to those obtained with the LB dictionary and MAX values of 0.4 and 2.2 kJ/mol for the backbone and side-chain entropies, respectively.

The trajectories from D. E. Shaw Research were obtained with a sampling interval (f) of 250 ps, and therefore, we have extracted coordinates with this interval for the other trajectories as well. However, this interval is much larger than what we typically use.[10] Therefore, we calculated order parameters and entropies for the BPTI simulations with $f = 10$ ps as well. The differences between the two sampling frequencies for one of the 500 ns BPTI simulations are shown in Table 2 (the results for the other simulation are similar). For the order parameters, the differences are negligible for both backbone and side-chain vectors. Consequently, this is also true for the dictionary entropy, which shows maximum differences of 0.1 kJ/mol for both backbone and side-chain. However, the discrepancy is much greater for the histogram entropy. The MAD is 0.8 and 0.4 kJ/mol for the backbone and side-chain, respectively, on par with the difference between the two 500 ns trajectories. A few residues display a much greater dependence on the sampling frequency, giving MAX of 3.5 and 1.7 kJ/mol for the backbone and side-chain, respectively, and this accumulates to differences of 41 and 18 kJ/mol for the backbone and side-chain, respectively. This is a serious issue that has not previously been discussed with respect to entropies, indicating that conformational fluctuations might be misrepresented if the sampling frequency is low. It should be noted that it is not possible a priori to determine which sampling interval should be used, although it can be argued that it is better to base the analysis on a larger number of samples (i.e., a shorter sampling interval). However, to be consistent with the D. E. Shaw Research trajectories, we have used $f = 250$ ps in the following analyses.

We assess convergence of the calculated order parameters and entropies by monitoring the agreement between successively extended segments of the 1 ms long simulation of BPTI (C).[12] First, we compare the results obtained using only the first 500 ns of the simulation to those obtained for the full trajectory, as this gives an indication of the convergence of the shortest simulations in this project. The resulting MAD and MAX values for the backbone parameters (shown in Table 2) increase by a factor of 2−3 relative to those for the comparison of simulations A and B. In particular, the difference in the total backbone entropy is 6−10 kJ/mol. On the other hand, the differences in the side-chain order parameters and entropies are on par with those between simulations A and B. Similar results are obtained if we instead compare our own 500 ns simulations (A and B) to the full trajectory of simulation C. Such results are expected from the general observations presented previously that the backbone fluctuations in BPTI are characterized by rare, large-amplitude transitions occurring on a time-scale of 10 μs, but with limited fluctuations within each conformational state.[12] In contrast, the side chains undergo large fluctuations on the nanosecond time-scale that are almost identical in all conformational states). Thus, for the backbone, the degree of similarity of the results obtained from trajectories of different

Table 3. Difference between the Dictionary and Histogram Entropies[a]

	backbone					side chains		
	MAD	MAX	SUM	SUM[b]	SUM[c]	MAD/n	MAX/n	SUM
mmp12-cn1h	0.8	3.1	123.0	−18.5	−109.6	0.8	3.1	122.5
mmp12-cn2h	0.8	2.2	110.4	−22.2	−103.7	0.8	3.7	152.3
mmp12-Δ[d]	0.3	2.6	−12.7	−3.7	5.9	0.3	1.4	29.8
gal3-Lac	0.6	2.7	67.4	−37.6	−100.9	0.7	3.4	75.0
gal3-l02	0.6	3.1	69.0	−33.2	−93.8	0.7	3.3	58.5
gal3-Δ[d]	0.2	1.5	−0.7	−2.2	−3.7	0.2	2.0	16.5
BPTI A, f = 250 ps	1.2	3.8	63.0	21.9	−2.2	0.3	2.2	10.9
BPTI A, f = 10 ps	0.5	2.5	22.9	−16.9	−41.1	0.3	2.1	−6.8
BPTI B	1.2	4.0	63.9	22.1	−2.6	0.3	2.0	11.4
BPTI C	1.4	2.6	71.8	27.3	0.0	0.2	2.1	12.0
BPTI 500 ns of C	1.3	3.4	81.7	38.7	13.0	0.2	1.6	12.2
BPTI 10 μs of C	1.3	3.5	81.1	38.1	12.5	0.2	1.9	11.7
BPTI 100 μs of C	1.4	3.0	72.5	28.0	0.8	0.8	6.9	12.4
gb3	1.1	2.4	60.8	17.7	−7.3	0.2	0.8	8.2
ubiquitin	1.1	3.1	81.5	23.0	−12.0	0.3	2.1	8.2
average[e]	1.0	3.0	79.0	24.8	48.0	0.5	2.5	51.0
median[e]	1.1	3.1	69.0	17.7	−12.0	0.3	2.2	12.0

[a]All differences are given in kJ/mol. The mean absolute deviation (MAD), the maximum absolute deviation (MAX), and the accumulated difference for all residues (SUM) are reported. For side chains, MAD and MAX are divided by the number of dihedral angles probed (MAD/n and MAX/n). Unless otherwise stated, the original parametrization[7] was used. [b]Calculated with a reparametrisation of the dictionary based on all the nine trajectories in Table 1. [c]Calculated with a reparametrisation of the dictionary based on the 1-ms BPTI simulation. [d]The difference in the entropy from the simulations with two ligands, i.e., mmp12-cn2h − mmp12-cn1h or gal3-l02 − gal3-lac. [e]The absolute average and median over all nine trajectories, i.e., over all rows in the table, except mmp12-Δ, gal3-Δ, and the three truncated BPTI C simulations.

lengths depends critically on the extent of averaging between conformational states that takes place in a given trajectory. On the other hand, such averaging does not significantly affect the side-chain fluctuations in the case of BPTI.

If we increase the length of the initial, shorter segment from the C trajectory to 10 or 100 μs, the agreement with the results obtained with the full trajectory gradually improves. There is not much difference between 500 ns and 10 μs, presumably because the sampling of different conformational states is similar for these two trajectory lengths. However, after 100 μs the MAD for the order parameters decrease to 0.01 and 0.02 for the backbone and side chains, respectively, and the MAX decreases to 0.13 and 0.07. Likewise, the MAD of the entropies decreases to 0.1 kJ/mol, whereas the maximum deviation is 0.5−1.2 kJ/mol, with cumulative differences of only 0.2−1.0 kJ/mol. This is explained by the fact that the most highly populated conformational substate in the 1 ms trajectory becomes dominant only after 200 μs (an analysis based on the various conformational substates of BPTI is presented in the Supporting Information). It is clear from these results that rare events can have a significant effect on the calculated properties if they lead to basins that are characterized by a different level of fluctuations.[20]

Finally, we also looked at the 10 μs simulations of gb3 and ubq by comparing properties calculated from the first 500 ns or the full trajectories. From the results in Table 2, it can be seen that most of the properties appear to be converged with MADs of 0.01−0.04 for order parameters and 0.1−0.3 kJ/mol for entropies. The entropies for gb3 show accumulated differences of 1 kJ/mol for both backbone and side-chain vectors, and similar deviations are observed between the results obtained from the LB-dictionary and histogram approaches. However, ubq shows slightly worse convergence, with an accumulated difference of 2−3 kJ/mol for the backbone and 8−9 kJ/mol for side chains. This shows that the convergence depends

somewhat on the protein. In general, it is clear that 500-ns simulations cannot be expected to be converged to better than 10 kJ/mol in the total entropy, and if several conformational states are visited as the trajectory is extended, the uncertainty is even larger.

These results contrast with our recent report that absolute and relative entropies do not converge to better than 31 kJ/mol for the same BPTI C trajectory.[10] This apparent discrepancy is explained by the different protocols used. In this article, we calculate average entropies over 100 ns windows, whereas in our previous study we accumulated histograms over the entire trajectory. Hence, it seems that window-averaging is an effective approach to improve the convergence of entropies, because it ignores transitions between conformations taking place on a time-scale slower than the time windows, which otherwise give a slowly increasing entropy.[10] However, it is not evident that such an approach yields correct results, because window-averaging downplays the entropic contributions from these rare transitions between different conformational states of some groups in the protein. On the other hand, the agreement with entropies calculated from NMR-relaxation order parameters should improve using window-averaging, because the order parameters are sensitive to fluctuations on time-scales shorter than the overall rotational correlation time, which typically is on the order of 5−10 ns for the proteins considered here.

Accuracy of Dictionary. Although it appears to take at least several microseconds to fully converge entropies, it is of interest to study how well the LB-dictionary entropies reproduce the histogram entropies for the various proteins and simulations. Such a comparison is presented in Table 3. For the backbone, the MAD (per residue) is 1 kJ/mol for all proteins and the maximum deviation is 2−4 kJ/mol. However, even if the individual deviations are quite small, the total deviations for all residues accumulate to 61−123 kJ/mol for the 11 simulations studied. Thus, the dictionary entropies are not

Table 4. Reparametrization of the Dictionary for the Backbone Vectors Based on the Various Simulations[a]

simulation	slope	intercept	R	MAD	MAX
mmp12-cn1h	0.46 ± 0.01	3.20 ± 0.03	0.90 ± 0.03	0.4	2.7
mmp12-cn2h	0.48 ± 0.01	3.25 ± 0.02	0.93 ± 0.02	0.4	2.4
gal3-Lac	0.39 ± 0.01	3.11 ± 0.03	0.82 ± 0.03	0.4	1.9
gal3-l02	0.36 ± 0.02	3.04 ± 0.03	0.81 ± 0.03	0.5	1.6
BPTI A, f = 250 ps	0.34 ± 0.01	2.86 ± 0.02	0.89 ± 0.02	0.3	1.4
BPTI A, f = 10 ps	0.51 ± 0.01	3.35 ± 0.03	0.90 ± 0.03	0.4	2.1
BPTI B	0.33 ± 0.02	2.84 ± 0.04	0.88 ± 0.04	0.3	1.5
BPTI C	0.34 ± 0.00	2.86 ± 0.00	0.93 ± 0.00	0.3	0.8
gb3	0.39 ± 0.00	2.99 ± 0.01	0.92 ± 0.01	0.3	1.3
ubiquitin	0.36 ± 0.00	2.93 ± 0.01	0.90 ± 0.00	0.4	1.5
all nine	0.42 ± 0.01	3.11 ± 0.00	0.87 ± 0.01		
LB original[7]	0.50 ± 0.02	3.42 ± 0.03	0.88		

[a]The slope and intercept (i.e., k and m in eq 1) as well as the correlation coefficient (R) is given. Mean absolute deviation (MAD) and maximum absolute deviation (MAX) between the dictionary and histogram entropy is also given in kJ/mol. Uncertainties of k, m, and R was estimated using bootstrapping.

fully transferable to other simulations, even for proteins involved in the original parametrization (BPTI and ubq) and the deviations are systematic, leading to sizable errors in total entropies.

For the side-chain vectors, the MAD per dihedral angle is below 1 kJ/mol for all proteins, and the maximum deviation per dihedral angle is 1−7 kJ/mol. However, in this case, the deviations are less systematic, so they accumulate to more acceptable deviations (8−12 kJ/mol) for the small proteins BPTI, gb3, and ubq but to unacceptably large errors (59−152 kJ/mol) for the two larger proteins gal3 and mmp12. This shows that although there is a rather small deviation per bond vector as also reported previously,[7] the deviations accumulate over the entire protein chain and sometimes results in very large deviations.

For mmp12 and gal3, we also considered the entropy difference between simulations with two different ligands. From Table 3, it can be seen that this reduces the MAD by a factor of 2−3. For gal3 the deviations accumulate to 1 and 17 kJ/mol for the backbone and side chain, respectively. The corresponding accumulated differences for mmp12 are 13 and 30 kJ/mol, respectively. This indicates that for differences between two states, the accuracy of the dictionary is improved by cancellation of errors (in particular, the intercept, m in eq 1, cancels exactly). However, the deviations are still much greater than expected for a reliable method.

Transferability of dictionary. In an attempt to improve the result and to study the transferability of the dictionary, we performed a reparametrisation of the dictionary, based on the new simulations. This was attempted only for the backbone vectors, because for some of the side chains there are very little data available (the dictionary provides a separate equation for each type of side chain, so the number of data points is the number of residues of each type in the proteins; in fact, the original LB-dictionary is based on only four data points for Met[7]). The reparametrization was performed by linear regression, using eq 1 with n = 2 and $f(x)$ = ln(x). Separate parametrizations were performed for each of the nine simulations in Table 1, but also for the pooled data of all nine simulations. The results are presented in Table 4.

In the original LB-dictionary, the slope (k in eq 1) was 0.50, but for the simulations presented in this paper, it is always smaller, 0.46−0.48 for the two mmp12 simulations, but 0.33−0.39 for the other four proteins (with uncertainties of 0.02 or

less). The intercept (m in eq 1) also shows a systematic variation between the various proteins, 3.20−3.25 for mmp12, 3.04−3.11 for gal3, 2.84−2.86 for BPTI, and 2.93−2.99 for the other two proteins (with uncertainties of 0.04 or less). All these values are significantly smaller than in the original LB-dictionary (3.42).[7] If we instead pool the data from all trajectories, we obtain slopes and intercepts that are of intermediate values, as expected: k = 0.42 and m = 3.11. The reparametrization reduces the MAD in the various simulations to 0.3−0.5 kJ/mol and the MAX error to 0.8−2.7 kJ/mol. Notably, the variation between different simulations of the same protein is much smaller than that between different proteins, even though both the mmp12 and gal3 simulations involve different bound ligands. This suggests that the dictionary depends quite strongly on the protein.

On the other hand, the force field seems to be less important, because the three BPTI simulations give similar results although they use slightly different force fields (cf. Table 1). Calculations based on truncated segments of the 1-ms BPTI simulation show that the parametrization depends on the length of the simulations, with accumulated errors of up to 13 kJ/mol for the two shorter trajectories (500 ns and 10 μs), but only 0.8 kJ/mol for the 100 μs simulation (Table 3).

The deviations between the histogram and dictionary entropies using the reparametrisations based on the BPTI C simulation or on all nine simulations are included in Table 3. Using the reparametrized dictionary based on all nine simulations, the average deviation over all simulations decreases from 79 to 25 kJ/mol for the backbone, a significant improvement. However, the improvement is essentially caused only by a shift of the systematic error: In the original parametrization, all proteins gave positive deviations (SUM = 61−123 kJ/mol), whereas with the reparametrization mmp12 and galectin-3 have negative deviations (−19 to −38 kJ/mol) and the other four proteins have positive deviations (18−27 kJ/mol). Note that the range of the errors is nearly the same for both parametrizations, 62 and 65 kJ/mol and that no protein has a SUM close to zero in any of the parametrizations. If we instead look at the result of the parametrization based on only the BPTI C simulation, the average error is 48 kJ/mol, which is better than the original parametrization, but worse than that based on all trajectories. This is in accordance with the results in Table 4, which show that the optimal parameters are significantly different for the different proteins. As expected, the

BPTI parametrization gives good results for gb3 and ubq, which both have similar best-fit parameters to those obtained for BPTI but appreciably worse results for gal3 and mmp12, which deviate more from BPTI. This strongly indicates that a single dictionary is not generally applicable to all proteins, instead each protein should be parametrized separately.

Our reparametrization indicates that the original parametrization is not reproducible in the case of BPTI and ubq. However, our parametrization is based on snapshots extracted every 250 ps, an interval that is much longer than typically used (it is unclear what frequency was used in the original paper). Using a sampling interval of 10 ps for one of the 500 ns simulations of BPTI (A), we obtain a parametrization that is much closer to the original LB dictionary. The slope is 0.51 and the intercept is 3.35 (see Table 4), which is significantly different from the parametrization employing f = 250 ps. Still, the difference between the dictionary and histogram entropies is only slightly lower for f = 10 ps (MAD = 0.5 kJ/mol and MAX 2.5 kJ/mol for the backbone; Table 3) and the accumulated differences are still sizable (23 and 7 kJ/mol for the backbone and side-chain vectors, respectively). This shows that the dictionary strongly depends on the sampling frequency and that even if we use a smaller sampling interval, we still obtain a large deviation in the total entropy.

CONCLUSIONS

We have analyzed nine previously published MD trajectories with lengths of 380 ns to 1 ms, to investigate the reproducibility and convergence of calculated order parameters and conformational entropies. Moreover, we investigate the accuracy and transferability of the dictionary developed by Li and Brüschweiler, which relates a specified subset of order parameters to residue-specific entropies.[7] We find that the order parameters are rather well converged in general, with MADs of 0.03–0.04 after 500 ns simulation, which decrease further to 0.01–0.02 after 100 μs. However, some bond vectors show much slower convergence, with errors of up to 0.27 after 500 ns simulation, 0.22 after 10 μs, and 0.13 after 100 μs. These bond vectors are primarily located in flexible loops and consequently have low order parameters with relatively high uncertainty (there is a fair anticorrelation between MAX and O^2, e.g. r^2 = 0.48 for the 100 μs/1 ms BPTI C simulation). This uncertainty is of the same magnitude as the variability caused by different sizes of the averaging window used in the iRED procedure.[21]

Likewise, entropies calculated either from the order parameters using the LB-dictionary or by the histogram approach, show a reasonable convergence with MADs of 0.1–0.3 kJ/mol per bond vector after 500 ns and 0.1 kJ/mol after 100 μs, and maximum errors of up to 2 kJ/mol after 500 ns and 1 kJ/mol after 100 μs. However, these uncertainties apply to each dihedral in the protein, so the errors can accumulate to sizable uncertainties in the entropy of the entire protein. For all proteins, we observe uncertainties in the total conformational entropy of up to 10 kJ/mol, and these remain even after 10 μs for BPTI, but they are reduced to 1 kJ/mol after 100 μs.

From this, we can conclude that individual calculated order parameters and entropies are converged to an acceptable level already after 500 ns. However, owing to the large number of degrees of freedom in the protein, these can add up to large uncertainties for the total conformational entropy, so large that it is questionable whether they are useful (10 kJ/mol). After

100 μs simulation, the uncertainty seems to be acceptable, but it is currently not known if this is by chance or if it is caused by the fact that we look at 10% of the entire simulation (i.e., that the uncertainty would increase if we elongate the simulation further).

Interestingly, window-averaged entropies show a much better convergence than the entropies calculated over the entire trajectory. This might suggest that window-averaging can offer a solution to obtaining converged entropies from MD simulations.[10] However, window-averaging means that entropic contributions from conformational fluctuations on time-scales longer than the windows become downweighted. On the other hand, the window-averaged entropy should capture those contributions that correspond to the fluctuations governing the NMR order parameters, provided that any rare transitions do not involve conformational states with significantly different fast time-scale behavior.

Finally, the accuracy of the dictionary-based entropies is reasonable for individual residues, with MADs (with respect to the entropies obtained by the histogram approach) of 1 kJ/mol and maximum errors of 2–4 kJ/mol. Unfortunately, the errors are systematic, so that the deviation accumulates to prohibitively high levels for the entire protein, 61–123 kJ/mol for the backbone and 8–152 kJ/mol for the side chains. The reason for this is that the optimized dictionary parameters (k and m in eq 1) vary significantly between different proteins, as shown by our protein-specific reparametrizations of the dictionary (Table 4). Thus, to reach an acceptable accuracy of the total entropy, a reparametrization of the dictionary is needed for each protein of interest.

Unfortunately, the parametrization is also quite sensitive to the length of the simulations and convergence is not reached until 100 μs for BPTI. Moreover, it strongly depends on the sampling frequency. However, it appears that the accuracy of the dictionary is significantly improved when relative entropies are considered, for example, when two different ligand-bound states of the same protein are compared. For example, we obtain accumulated differences of 2–4 kJ/mol for the reparametrized backbone data, and 17–30 kJ/mol for the total side-chain entropy calculated using the original LB dictionary. This type of comparative analysis arguably represents the most useful approach for assessing contributions from conformational entropy to the binding free energy. As such the LB dictionary is a promising approach.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Convergence of substate entropies and transferability of dictionary for the various substates. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Ulf.Ryde@teokem.lu.se. Tel: +46 − 46 2224915. Fax: +46 − 46 2228648.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

■ REFERENCES

(1) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007**, *448*, 325−330.

(2) Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U.; Akke, M. *J. Am. Chem. Soc.* **2010**, *132*, 14577−14589.

(3) Diehl, C.; Genheden, S.; Modig, K.; Ryde, U.; Akke, M. *J. Biomol. NMR* **2009**, *45*, 157−169.

(4) Tzeng, S-R; Kalodimos, C. G. *Nature* **2012**, *488*, 236−240.

(5) Akke, M.; Brüschweiler, R.; Palmer, A. G. *J. Am. Chem. Soc.* **1993**, *115*, 9832−9833.

(6) Yang, D.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 369−382.

(7) Li, D. W.; Brüschweiler, R. *J. Am. Chem. Soc.* **2009**, *131*, 7226−7227.

(8) Li, Z.; Raychaudhuri, S.; Wand, A. J. *Protein Sci.* **1996**, *5*, 2647−2650.

(9) Dhulesia, A.; Bodenhausen, G.; Abergel, D. *J. Chem. Phys.* **2008**, *129*, 095107.

(10) Genheden, S.; Ryde, U. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8662−8677.

(11) Maragakis, P.; Lindorff-Larsen, K.; Eastwood, M. P.; Dror, R. O.; Kepeis, J. L.; Arkin, I. T.; Jensen, M. O.; Xu, H.; Trbovic, N.; Friesner, R. A.; Palmer, A. G., III; Shaw, D. E. *J. Phys. Chem. B* **2008**, *112*, 6155−6158.

(12) 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.

(13) Lindorff-Larsen, K.; Maragakis, P.; Plana, S.; Eastwood, M. P.; Dror, R. O.; Shaw., D. E. *PLoS One* **2012**, *7*, e32131.

(14) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins* **2004**, *55*, 383.

(15) Prompers, J.J.; Brüschweiler, R. *J. Phys. Chem. B* **2000**, *104*, 11416−11424.

(16) Li, D.-A.; Brüschweiler, R. *Phys. Rev. Lett.* **2009**, *102*, 118108.

(17) Carlsson, J.; Åqvist, J. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5385−5395.

(18) Zhou, H.-X.; Gilson, M. K. *Chem. Rev.* **2009**, *109*, 4092−4107.

(19) Wereszczynski, J.; McCammon, J. A. *Q. Rev. Biophys.* **2012**, *45*, 1−25.

(20) Zeiske, T.; Stafford, K. A.; Friesner, R. A.; Palmer, A. G. *Proteins* **2012**, *81*, 499−509.

(21) Genheden, S.; Diehl, C.; Akke, M.; Ryde, U. *J. Chem. Theory Comput.* **2010**, *6*, 2176−2190.