Supporting information for: Localization of Millisecond Dynamics: Dihedral Entropy from Accelerated MD

Anna S. Kamenik,[†] Ursula Kahler,[†] Julian E. Fuchs,[*,†,‡] and Klaus R. Liedl[*,†]

†*Institute of General, Inorganic and Theoretical Chemistry, Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria*

‡*Current Address: Medicinal Chemistry/Structural Research, Boehringer Ingelheim RCV GmbH & Co. KG, Vienna, Austria*

E-mail: julian.fuchs@uibk.ac.at; klaus.liedl@uibk.ac.at

Supporting Information Available

Calculation of aMD Parameters

All parameters were calculated according to the formulas below.

$$E_{threshD} = \bar{E}_{\mathrm{DIHED}} + a_1 \cdot N_{\mathrm{RES}} \tag{1}$$

$$\alpha_D = a_2 \cdot \frac{N_{\mathrm{RES}}}{5} \tag{2}$$

$$E_{threshP} = \bar{E}_{\mathrm{TOT}} + b_1 \cdot N_{\mathrm{ATOMS}} \tag{3}$$

$$\alpha_P = b_2 \cdot N_{\mathrm{ATOMS}} \tag{4}$$

$\bar{E}_{\mathrm{DIHED}}$ is the average dihedral and $\bar{E}_{\mathrm{TOT}}$ the total potential energy resulting from previous cMD simulations. N_{RES} and N_{ATOMS} are the number of residues and atoms respectively, in each system. The variables a_1, a_2, b_1 and b_2 were altered systematically for each system to optimize the boosting level. The applied values for the aMD simulations are shown in table S1. For Bet v 1a we systematically tested 5 sets of boosting parameters. Further we set

Table S1: Parameters for aMD simulation of each system.

System	a_1	a_2	E_{threshD}	α_D	b_1	b_2	E_{threshP}	α_P
Di-Ala	3.5	3.5	18.7	1.4	0.175	0.175	-343	532.2
BPTI	4	4	837.8	46.4	0.16	0.16	-40754	2198.6
Bet v 1a	3	3	2189.2	95.4	0.30	0.16	-73135	4333.6

up an aMD simulation in which we only boosted the dihedral potential. We started with a set of boosting parameter that we considered to only cause a slight boosting effect and slowly increased the intensity of the boosting until the protein would unfold (Set 5). We chose to continue with the most aggressive boosting parameters, that would not unfold the protein after 1 μs of aMD simulation. Despite the awareness of the resulting increase in inaccuracy we tried to maximize the effect of the aMD method. We considered this approach to give us the most information on potential slow conformational changes in Bet v 1a, as well as on the robustness of our metric.

Table S2: Parameters tested for aMD simulation of Bet v 1a.

Set	a_1	a_2	E_{threshD}	α_D	b_1	b_2	E_{threshP}	α_P
1	4	4	2348.2	127.2	0.30	0.20	-73135	8125.5
2	4	4	2348.2	127.2	0.20	0.20	-75843	5417
3	3	2	2189.2	95.4	0.20	0.20	-75843	5417
4	3	3	2189.2	95.4	0.30	0.16	-73135	4333.6
dh	3	3	2189.2	95.4	-	-	-	-

Alanine Dipeptide

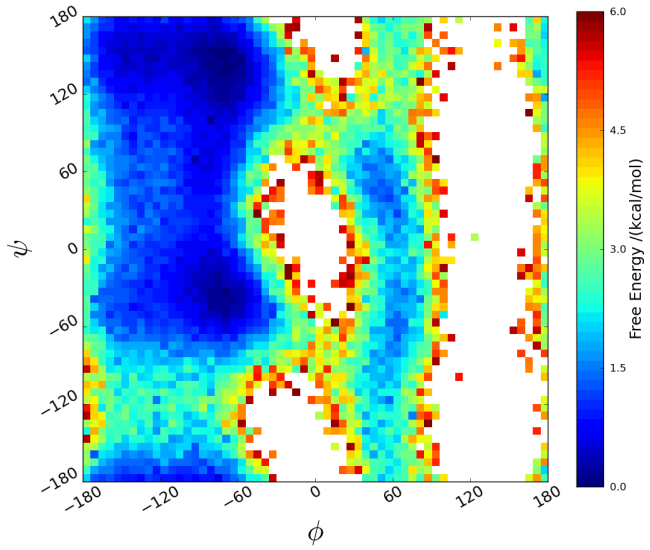

Figure S1: Conformational space sampled in 1 μs of aMD. Blue regions indicate the most favorable states with the lowest energy. Unfavorable torsional states with a free energy higher than 6 kcal/mol (red) are depicted in white.

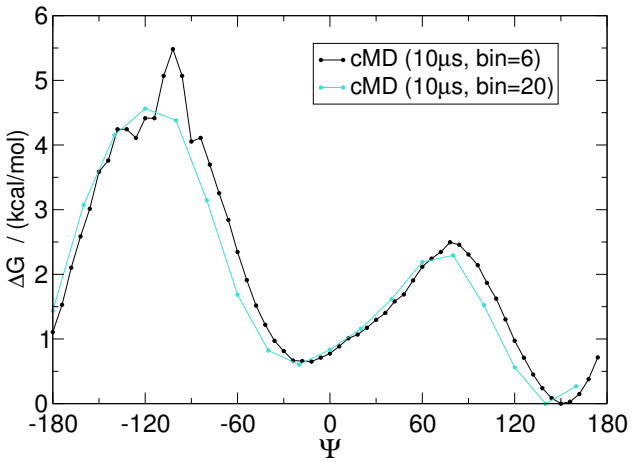

Figure S2: Free energy of Ψ calculated using a bin size of 6°(black) and 20°(turquoise). The increase of bin sizes smoothens the free energy surface but causes a shift of minima. The jaggedness of the profile using a bin width of 6°is most likely due to the limited number of recorded frames (100 000).

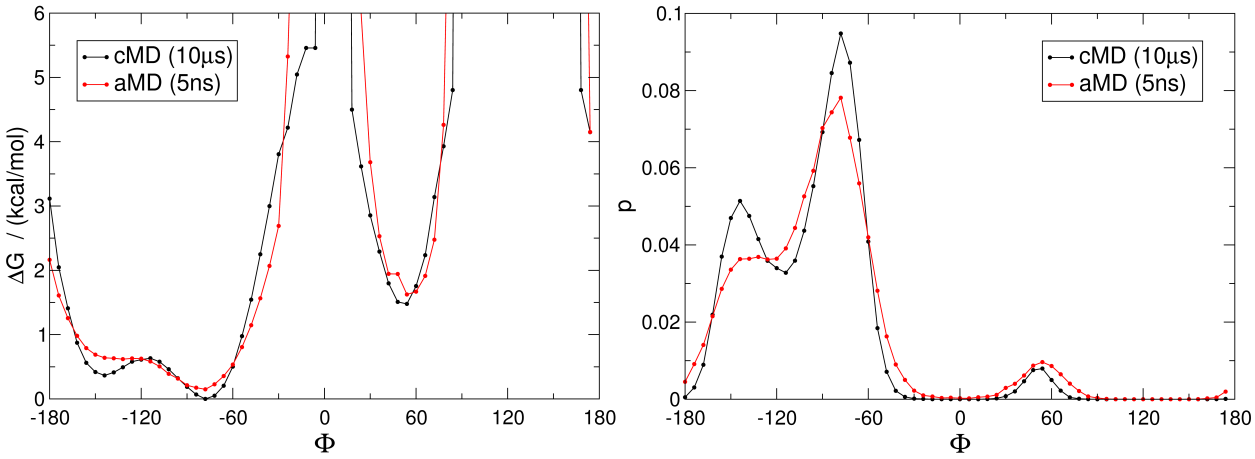

Figure S3: Free energy and state populations of Φ in Di-Ala. Left: Free energy distribution of Φ from a 10 μs cMD (black) and 5 ns aMD (red) simulation of Di-Ala. Rarely or not visited dihedral states showing highly unfavorable free energies were cut off at 6 kcal/mol. Right: State populations calculated from the free energies of Φ as shown in Figure 1

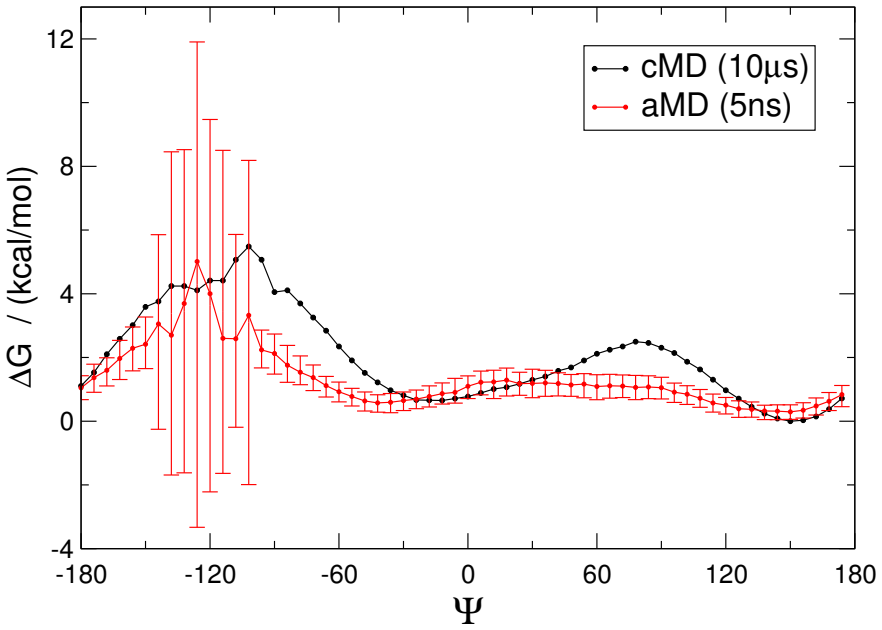

Figure S4: Reweighted free energies of Di-Ala. A 1 μs trajectory was split into 200 segments of 5 ns each (2 500 frames) using the segments for averaging. (red) The free energy landscape of Ψ was reconstructed using reconstructed using Maclaurin series in the reweighting protocol. As a reference the free energy surface of Ψ from a 10 μs cMD simulation (black) is shown. The standard deviation of the aMD trajectory shows the stability of the results in well sampled areas and highlights strong fluctuations for less sampled ones.

BPTI

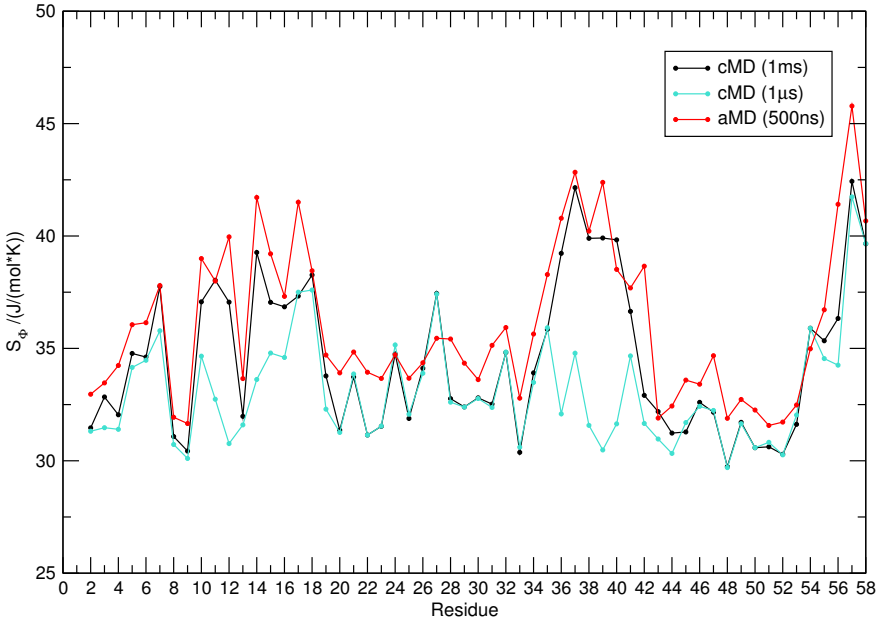

Figure S5: Comparing local flexibility of BPTI captured in cMD and aMD simulations. Residue-wise dihedral entropies S_Φ from a 1 ms cMD simulation (black) and 500 ns aMD simulation of BPTI (red) show remarkable rank correlation. Local flexibility observed in a 1 μs cMD simulation (turquoise) clearly differs from the aMD results.

Comparing dihedral entropies from Φ and Ψ we find a Spearman rank correlation r=0.77 between S_Ψ and S_Φ of BPTI. For Bet v 1a we observe correlation of r=0.86. These results support the assumption of a similar extent of motions captured in both backbone dihedrals phi and psi. When considering the information displayed in Ramachandran plots of single amino acids, the Ψ-axis generally shows a broader distribution than the Φ-axis.[S1] So most amino acids secondary structure elements, such as alpha-helices and beta-sheets, can be distinguished solely by looking at the psi-distribution.[S2] Hence, not the whole backbone dynamics are reflected by the Ψ angle. Φ dihedral distribution were calculated as well. Yet, for the representation of protein dynamics based on dihedral entropies we prioritized Ψ over Φ as it captures the backbone dynamics more comprehensively.

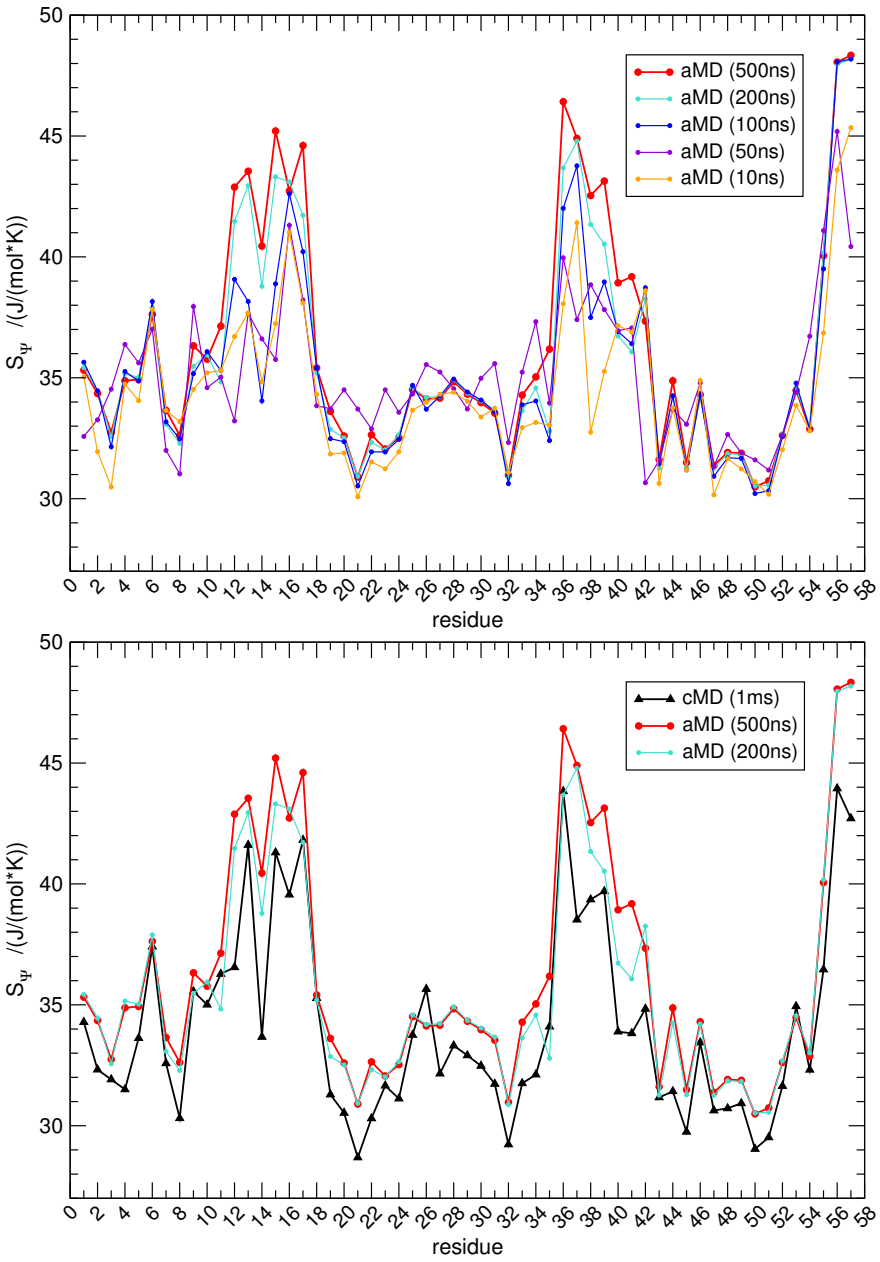

Figure S6: Benchmarking sampling time of BPTI. Dihedral entropies of BPTI were evaluated after 10, 50, 100, 200 and 500 ns aMD simulation time. TOP: 200 ns and 500 ns result in similar flexibility patterns, while shorter sampling runs capture only small increase of flexibility from residues 10–20 and 32–44. BOTTOM: Comparison of dihedral entropies from 200 ns (blue) and 500 ns (red) aMD to 1 ms cMD (black) sampling. In both aMD simulations the same regions are captured as flexible, yet after 500 ns the shape of the 1 ms simulation is reproduced more accurately.

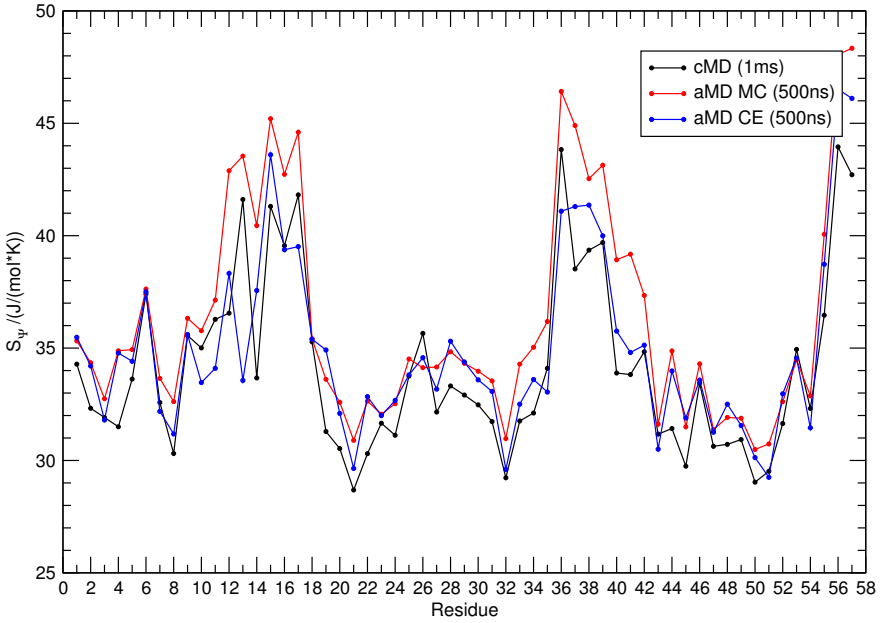

Figure S7: Probing different reweighting protocols. Dihedral entropies S_Ψ were calculated from the same 500 ns trajectory of BPTI using Maclaurin series (red) and cumulant expansion (blue) to approximate the exponential in the reweighting protocol. Using cumulant expansion we find a Spearman rank correlation of r=0.85 between the 1 ms control (black) and 500 ns aMD simulation. Reweighting with Maclaurin series increases the correlation between aMD and cMD results to r=0.90.

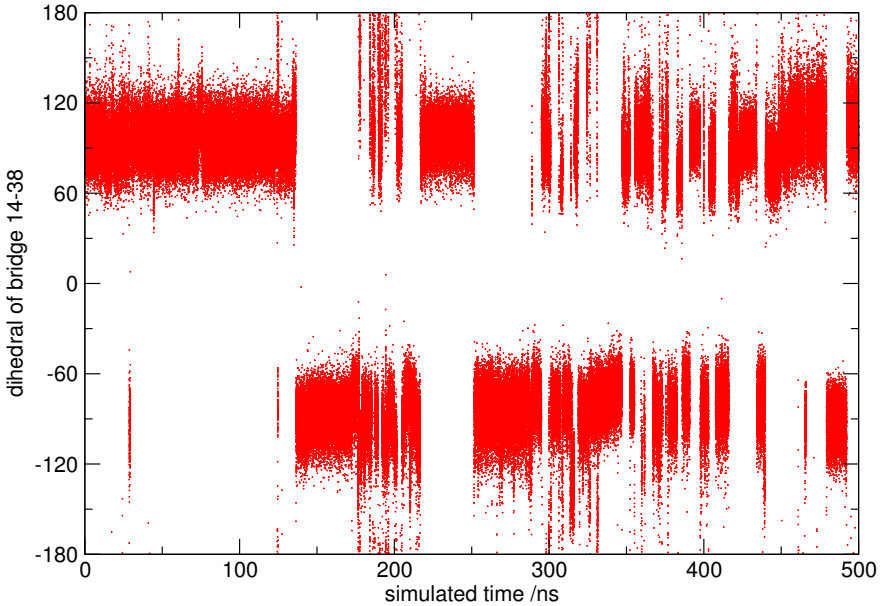

Figure S8: Isomerization of disulfide bridge CYS14-CYS38. During 500 ns aMD simulation the dihedral of the disulfidebridge between CYS14 and CYS38 switches multiple times between values around 100 degrees to -100 degrees. The two populated dihedral states represent the cis- and trans-conformation of the disulfide bridge.

Bet v 1a

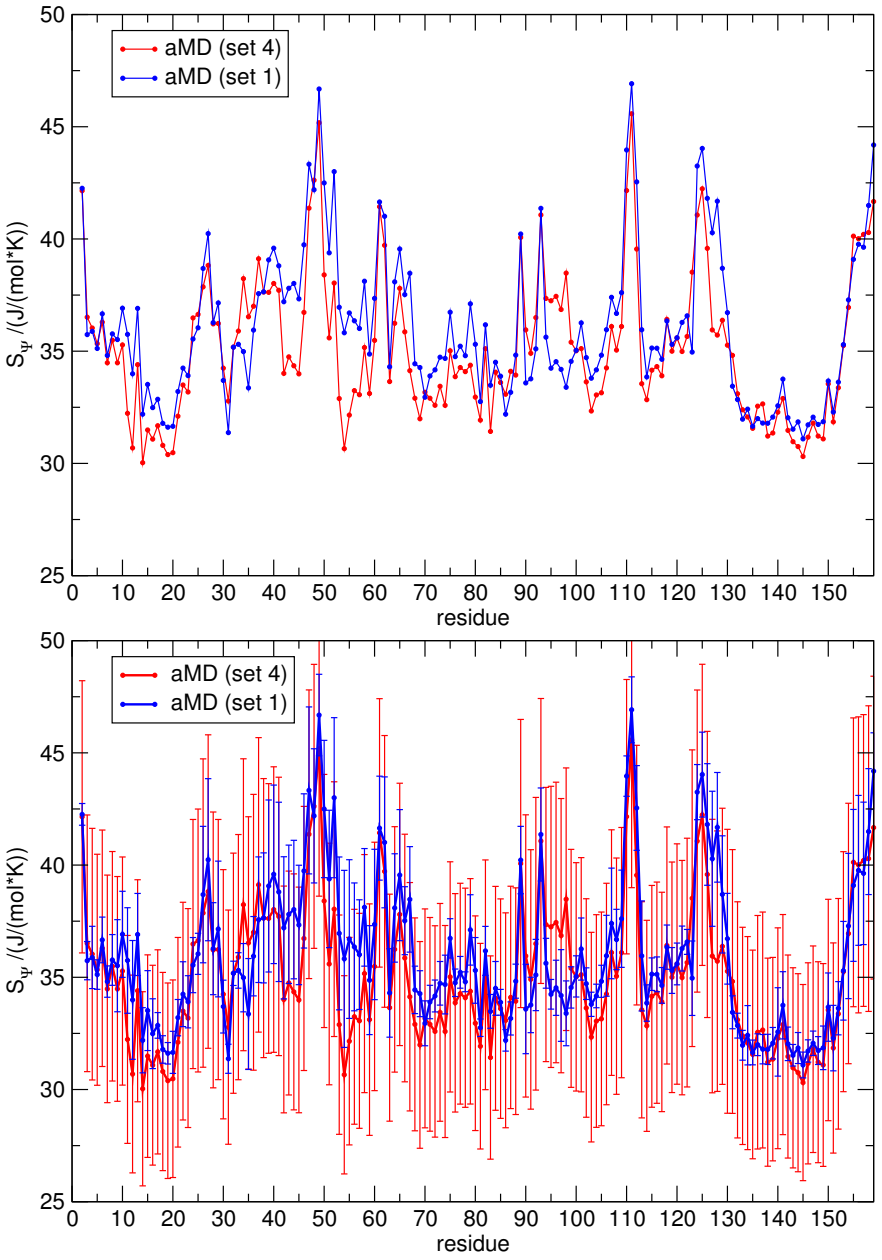

Figure S9: Probing different levels of acceleration in Bet v 1a aMD simulations. Dihedral entropies S_Ψ for two sets of aMD parameter are shown (r=0.80) one with less boosting (blue, set 1) and a more aggressively boosted one (red, set 4). The error shown in the bottom derives from trajectory splitting. The 1 μs trajectories are split into 50 segments, resulting in the shown average and standard deviation representing 20 ns of aMD sampling. It is clear to see that the error of an aMD simulation is strongly dependent on the chosen boosting parameters, i.e. level of acceleration.

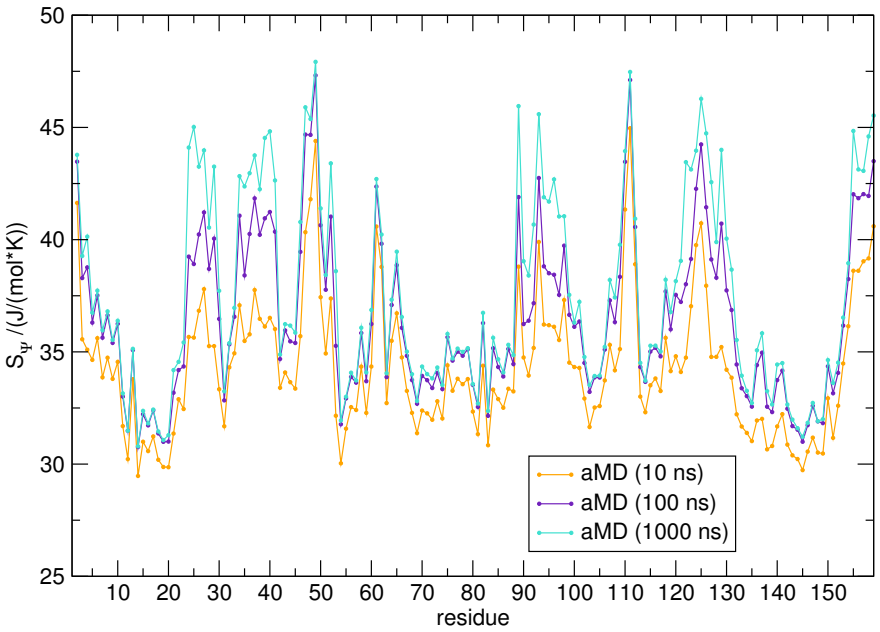

Figure S10: Benchmarking sampling time of Bet v 1a. Evaluation of local flexibility in Bet v 1 a on different time scales of aMD simulation. Most notable differences in the three aMD simulations are found from residue 15-45.

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(S1) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95–99.

(S2) Wood, M. J.; Hirst, J. D. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 476–481.