Supporting Information:

Integrating Multiple Accelerated Molecular Dynamics to Improve Accuracy of Free Energy Calculation

Xiangda Peng, Yuebin Zhang, Yan Li, QingLong Liu, Huiying Chu, Dinglin Zhang and Guohui Li

Laboratory of Molecular Modeling and Design,

State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics,

Chinese Academy of Sciences, Dalian 116023, P. R. China

Section 1. Use multiple methods to calculate the weight parameters in IaMD for Alanine Dipeptide.

The weight parameters $\{n_i\}$ can be determined by an iterative procedure (histogram flattening approach):

$$\frac{n_j^{l+1}}{n_i^{l+1}} = \frac{\langle e^{-\beta(\Delta V_i(x)-\Delta V(x))}\rangle_l}{\langle e^{-\beta(\Delta V_j(x)-\Delta V(x))}\rangle_l}$$

Here *l* denote the index of update round. After we iterated 20 rounds (and 500 ps each round) using this method, the parameters still have small fluctuations (Figure S2A). In order to promote convergence, the weighted histogram analysis method (WHAM) was introduced after iterating 12 rounds using the histogram flattening method. [1] With the use of preceding all samples, the optimal canonical probability $w(x)$ can be estimated by iteratively solving the WHAM equations:

$$k_B T \ln(w(x)) = -k_B T \ln\left(\sum_l e^{-\beta(\Delta V(x)^l - f^l)}\right) + C$$

$$f^l = -k_B T \ln\left(\sum_x (w(x) \times e^{-\beta \Delta V(x)^l})\right)$$

Here *C* is a normalization constant that meets the condition $\sum_x w(x) = 1.0$. Then, the weight parameters for next round can be obtained by

$$n_i^{l+1} = \sum_x (w(x) \times e^{\beta \Delta V_i(x)})$$

As shown in Figure S2B, the use of WHAM algorithm can significantly improve the convergence of iteration. However, regardless of the convergence, the difference between the weight parameters obtained by the above two methods are very small and are consistent with those obtained by the method of Lu using intersection points. (Figure S2C) The parameters are listed in columns of IaMD-A7^{iter} and IaMD-A7^{wham} in Table S3.

In the original work of Lu[2], the weight parameters were obtained using the midpoint of potential canonical averages. However, in our work, the midpoint of potential canonical averages is replaced by the intersection point of potential distributions. There is a slight difference between the weight parameters calculated by these two schemes. (Figure S2C and the columns of IaMD-A7^{mean} in Table S3) Even so, it does not significantly affect the accuracy and efficiency of the IaMD simulation. In fact, the roughness of the free energy profiles calculated using parameters that are determined by different methods is comparable. (Figure S8B, S8C, and S8D) The accessible φ-ϕ spaces in 300 ns IaMD simulations with parameters IaMD-A7^{iter}, IaMD-A7^{wham}, and IaMD-A7^{mean} are 95.3 %, 95.3 % and 96.5 % respectively, are similar to that of IaMD-A7 (96.0 %).

Section 2. Parameter determination procedure

(1) Run a short cMD simulation, find the lowest potential energy E_{min} and average potential energy V_{mean}, set the value E_0 in IaMD parameters to less than E_{min}, calculate E_i of other subterms based on V_{mean} and empirical rules.

(2) Perform a few short aMD simulations with small α values, find the α_N according to the location of energy distribution peaks.

(3) Obtain $\Delta V_{N,max}$ through the following equation:

$$\Delta V_{N,max} = \frac{(E_N - E_{min})^2}{\alpha_N + E_N - E_{min}}$$

(4) Insert N-2 points at equal intervals between 0 and $\Delta V_{N,max}$ as the value of $\Delta V_{2,max}, \Delta V_{3,max}, \ldots, \Delta V_{N-1,max}$ respectively.

(5) Calculate the values of $\alpha_2, \alpha_3, \ldots, \alpha_{N-1}$ through the following equation:

$$\alpha_i = \frac{(E_i - E_{min})^2}{\Delta V_{i,max}} - E_i + E_{min}$$

α_1 can be set to an arbitrary non-zero number.

(6) Run a series of short aMD simulations with parameters (α_1, E_1), $(\alpha_2, E_2), \ldots, (\alpha_N, E_N)$ respectively. Check the overlap of energy distributions in these simulations. If necessary, change the number of N and repeat steps (4), (5) and (6).

(7) Check if the energy distribution peaks of these aMD simulations are arranged evenly, if yes, go to step (8), otherwise, use interpolation algorithm to get new $\{\alpha'_1, \alpha'_2, \ldots, \alpha'_N\}$:

- Define V_i^p is the position of potential distribution peak in aMD simulation with (E_i, α_i);
- Consider α_i as a function of V_i^p, and use the cubic spline curves to describe this function;
- Find a set of V_i^p which are equidistant and spanning the range from V_0^p to V_N^p;
- Calculate the new α_i value at V_i^p using cubic spline interpolation algorithm.

Then, repeat steps (6) and (7).

(8) Base on the following equation,

$$M_{i+1} = M_i + \frac{\left(E_i - V_{i,i+1}^x\right)^2}{\alpha_i + E_i - V_{i,i+1}^x} - \frac{\left(E_{i+1} - V_{i,i+1}^x\right)^2}{\alpha_{i+1} + E_{i+1} - V_{i,i+1}^x}$$

calculate the value of $M_2, M_3, \ldots, M_N$ recursively. M_1 is always equal to 0.0. $V_{i,i+1}^x$ is the intersection position of two energy distribution curves from aMD simulations with (E_i, α_i) and (E_i, α_{i+1}). It can also be replaced by the midpoint of the adjacent two distribution potential peaks.

If use the midpoint of potential canonical averages to replace the intersection point of potential distributions. There is an added benefit of simplifying the parameter determination process. After performing a series of short aMD simulations in step (6) of above parameter determination procedure, the relationship of $\{\alpha_i\}$ and potential canonical averages $\{V_{mean,i}\}$ can be determined by interpolation algorithm or curve fitting algorithm.

According to the width of the obtained potential energy distribution from aMD simulations, one can choose a new set of equidistant $\{V_{mean,i}\}$, and get a new set of $\{\alpha_i\}$ through the determined relationship. This avoids the possible repeating of steps (4), (5) and (6) to get the well overlap of energy distributions and steps (6) and (7) to keep the energy distributions arranged evenly.

Section 3. Simulation details

All simulations are performed using OpenMM[3-4] on GPUs with mixed precision. We use amber99sb[5] force field for

Alanine Dipeptide, Chignolin, and Trp-cage and amber03[6] force field for Villin Headpiece simulations.

The Alanine Dipeptide was solvated into a cubic box with 630 TIP3P[7] water molecules. The total number of atoms is 1912. The system was initially equilibrated in NVT ensemble for 500 ps and NPT ensemble for 2 ns. After that, the aMD, IaMD, and cMD run with NVT ensemble with 300 K. For Chignolin, Trp-cage, and Villin Headpiece, the initial structures were obtained from the PDB code 1UAO,[8] 2JOF, and1YRF.[9] To get the unfolded structures of Chignolin, Trp-cage, and Villin Headpiece, we first performed 500 ns simulations at 400 K, 500 K, and 500 K respectively. In these simulations, water box extends 12 Å from the solute surface. 0.05 M NaCl was used to balance the charges of the Chignolin and Villin Headpiece system to neutral. The total number of atoms is 7053 for Chignolin, 11468 for Trp-cage and 27144 for Villin Headpiece. Then, the selected frame with the unfolded structure of the three fast-folding proteins is used as the start state for 1ns NVT equilibrium followed by 5 ns NPT equilibrium. Consistent with the Alanine Dipeptide system, simulations with aMD or IaMD for these fast-folding proteins were performed under the NVT ensemble with 300 K.

The Andersen Thermostat[10] was applied to couple all system's temperature with a collision frequency of 1.0 ps^{-1}. In constant pressure simulations, Monte Carlo Barostat was used to keep the pressure to1.0 bar with a trial frequency to change the box every 50 MD steps. To estimate the contribution of long-range non-bonded interactions that beyond the cutoff of 8.0 Å, the Particle-Mesh Ewald (PME)[11] and the dispersion correction algorithm[12] were exploited for electrostatic and van der Waals terms respectively. For all simulations, the Settle Algorithm[13-14] was applied to keep TIP3P water rigid. All hydrogen bond lengths of protein were fixed using CCMA algorithm.[15] Mass of hydrogen in protein (and in water molecules of Alanine Dipeptide simulations) was increased to 4 amu. The added mass is subtracted from the bonded heavy atom. Hence the integration step size of all simulations was set to 4 fs.[16] In the three folding cases, to increase the diffusion motions of waters around protein molecule, we reduce the mass of water oxygen atom to 2 amu and keep water hydrogen atom mass as 1amu. The change does not affect the thermodynamic equilibrium properties.[17-18]

Section 4. The impact of acceleration parameters and data collection frequency on accuracy

We choose another pair of protocols, aMD-A3 and IaMD-A3, to analyze the topic of this section. The standard deviations of aMD-A3 (with 83.5% coverage) and IaMD-A3 (with 88.7% coverage) are the smallest in all aMD and IaMD protocols respectively.

The number of effective sampling points along ϕ−ψ space (denoted as $N_e(\varphi, \phi)$) in single 300ns aMD-A3 and

IaMD-A3 simulation were shown in Figure S5A and S5B respectively. The N_e in aMD-A3 is about one order of magnitude smaller than that in IaMD-A3 at same locations. At the bin of β, N_e are about 2445 and 44175 and the ratio of $N_e$to the sampling number N_s are about 1.2% and 22% for aMD-A3 and IaMD-A3 respectively. Even though the difference with N_e of IaMD is huge, N_e of aMD is large enough to accurately calculate the statistical average, especially at the locations of the minimum.

As shown in Figure S5C and S5D, the $\Delta V_{max}(\varphi, \phi)$ of aMD-A3 changes significantly with the coordinate φ or ϕ change. However, the free energy of IaMD is mainly affected by $N_e(\varphi, \phi)$ due to the fact that the $\Delta V_{max}(\varphi, \phi)$ variation of IaMD-A3 within the range of 0.5 kcal/mol at most of the low free energy regions. Hence, the shape of the effective point distribution along ϕ and ψ is very similar to the free energy profiles and global or local maximums of effective point distribution profiles correspond exactly to the minimums of FEPs. It is beneficial to the sampling efficiency or statistics accuracy of important states which usually have low free energy.

2-D dihedral ϕ-ψ FEPs generated from single 300ns aMD-A3 (Figure S7A) or IaMD-A3 (Figure S7C) simulation are consistent with the result of 5us cMD (Figure S6A) or previous simulation results.[19]

The contour lines in two FEPs are smooth. It indicates that none of them suffered from the big statistical noises in reweighting. Relative to aMD-A3, the sampling efficiency of aMD-A6 was significantly improved, but the contour lines of FEPs from aMD-A6 (Figure S7B) were significantly rough. However, the smoothness of FEPs from IaMD-A7 (Figure S7D) was almost no difference with IaMD-A3.

The FEPs in Figure S7 are generated using the data collected every 0.012 ps in the simulation. Writing data too frequently may significantly affect the computational efficiency, especially on GPU platform. The FEPs in Figure 2A and 2B are calculated from the simulation with 0.1 ps data collection interval. Compare the FEP in Figure S7B, the contour lines of FEP in Figure 2A become very rough and discontinuous. The smoothness of new profile of IaMD-A7 with 0.1 ps (Figure 2B) becomes slightly rougher than that in Figure S7D but still better than that the result of aMD-A6 with 0.012 ps collection interval (Figure S7B). Thanks to sufficient effective sampling points, we do not need to save data with high-frequency in IaMD.

Table S1. The aMD parameters for Alanine Dipeptide.

name	α(kcal/mol)	E(kcal/mol)
aMD-A1	100.00	40.63
aMD-A2	60.00	40.63
aMD-A3	30.00	40.63
aMD-A4	20.00	40.63
aMD-A5	10.00	40.63
aMD-A6	5.00	40.63
aMD-A7	3.00	40.63

Table S2. The parameters of IaMD for Alanine Dipeptide. (The units are kcal/mol)

IaMD-A1			IaMD-A2			IaMD-A3			IaMD-A4		
a	E	M	a	E	M	a	E	M	a	E	M
99.72	2.39	0.00	84.14	2.39	0.00	68.98	2.39	0.00	64.11	2.39	0.00
99.72	40.63	-6.31	84.14	40.63	-7.13	68.98	40.63	-8.16	64.11	40.63	-8.57
43.40	40.63	-10.82	36.86	40.63	-11.79	31.28	40.63	-12.77	29.60	40.63	-13.12
28.69	40.63	-13.30	25.17	40.63	-14.05	21.79	40.63	-14.87	20.64	40.63	-15.19
22.27	40.63	-14.77	19.42	40.63	-15.49	16.40	40.63	-16.38	15.34	40.63	-16.74
18.02	40.63	-15.92	15.34	40.63	-16.71	12.47	40.63	-17.66	11.48	40.63	-18.05
14.72	40.63	-16.94	12.16	40.63	-17.78	9.50	40.63	-18.76	8.59	40.63	-19.15
12.05	40.63	-17.84	9.64	40.63	-18.71	7.23	40.63	-19.67	6.44	40.63	-20.04
									4.82	40.63	-20.76

IaMD-A5			IaMD-A6			IaMD-A7		
a	E	M	a	E	M	a	E	M
47.26	2.39	0.00	47.26	0.24	0.00	48.00	2.39	0.00
47.26	40.63	-10.29	47.26	40.63	-10.29	48.00	40.63	-10.23
23.74	40.63	-14.38	23.74	40.63	-14.38	24.02	40.63	-14.34
16.03	40.63	-16.47	16.03	40.63	-16.47	16.26	40.63	-16.43
11.05	40.63	-18.18	11.05	40.63	-18.18	11.26	40.63	-18.12
7.62	40.63	-19.55	7.62	40.63	-19.55	7.81	40.63	-19.47
5.26	40.63	-20.57	5.26	40.63	-20.57	5.42	40.63	-20.50
3.63	40.63	-21.36	3.63	40.63	-21.36	3.76	40.63	-21.28
			2.51	40.63	-21.87	2.61	40.63	-21.86
			1.73	40.63	-22.27	1.81	40.63	-22.28
			1.19	40.63	-22.54	1.25	40.63	-22.55
						0.87	40.63	-22.76
						0.60	40.63	-23.01

Table S3. The parameters obtained from different methods for Alanine Dipeptide. (The units are kcal/mol)

IaMD-A7^{iter}			IaMD-A7^{wham}			IaMD-A7^{m}		
a	E	M	a	E	M	a	E	M
48.00	2.39	0.00	48.00	2.39	0.00	48.00	2.39	0.00
48.00	40.63	-10.23	48.00	40.63	-10.20	48.00	40.63	-9.77
24.02	40.63	-14.34	24.02	40.63	-14.28	24.02	40.63	-13.84
16.26	40.63	-16.45	16.26	40.63	-16.38	16.26	40.63	-15.95
11.26	40.63	-18.14	11.26	40.63	-18.06	11.26	40.63	-17.65
7.81	40.63	-19.49	7.81	40.63	-19.40	7.81	40.63	-19.04
5.42	40.63	-20.53	5.42	40.63	-20.43	5.42	40.63	-20.13
3.76	40.63	-21.31	3.76	40.63	-21.20	3.76	40.63	-20.97
2.61	40.63	-21.89	2.61	40.63	-21.78	2.61	40.63	-21.61
1.81	40.63	-22.32	1.81	40.63	-22.21	1.81	40.63	-22.09
1.25	40.63	-22.63	1.25	40.63	-22.53	1.25	40.63	-22.46
0.87	40.63	-22.87	0.87	40.63	-22.76	0.87	40.63	-22.74
0.60	40.63	-23.04	0.60	40.63	-22.94	0.60	40.63	-22.95

Table S4. The parameters of ITS for Alanine Dipeptide. T refers to the temperature in Kelvin temperature. The unit of M is kcal/mol.

ITS-A1		ITS-A2		ITS-A3		ITS-A4		ITS-A5	
T	M	T	M	T	M	T	M	T	M
300.00	0.00	300.00	0.00	300.00	0.00	300.00	0.00	300.00	0.00
656.91	6.83	656.91	6.83	656.91	6.83	656.91	6.83	656.91	6.83
1244.18	10.00	1244.18	10.00	1244.18	10.00	1244.18	10.00	1244.18	10.00
		1869.95	11.44	1869.95	11.44	1869.95	11.44	1869.95	11.44
				2084.76	11.79	2084.76	11.79	2084.76	11.79
						2231.11	12.00	2231.11	12.00
								2353.31	12.17

Table S5. The parameters of aMD for Chignolin, Trp-cage and Villin Headpiece. (The unit is kcal/mol)

name	α(kcal/mol)	E(kcal/mol)
aMD-C1	7.17	98.95
aMD-C2	7.17	122.85
aMD-C3	7.17	146.75
aMD-T1	14.34	227.06
aMD-T2	14.34	274.86
aMD-T3	14.34	322.66
aMD-V	32.74	547.32

Table S6. The parameters of IaMD for Chignolin, Trp-cage and Villin Headpiece. (The unit is kcal/mol)

IaMD-C1			IaMD-C2			IaMD-C3		
α	E	M	α	E	M	α	E	M
2000.00	0.00	0.00	2000.00	0.00	0.00	2000.00	0.00	0.00
447.24	122.85	-2.49	1092.02	146.75	-3.01	1384.32	170.65	-4.59
255.18	122.85	-3.84	507.65	146.75	-5.84	641.49	170.65	-9.00
142.82	122.85	-5.67	312.86	146.75	-8.51	393.88	170.65	-13.21
92.53	122.85	-7.20	215.46	146.75	-10.99	270.08	170.65	-17.19
65.21	122.85	-8.44	157.03	146.75	-13.33	195.79	170.65	-20.98
45.79	122.85	-9.65	118.07	146.75	-15.53	146.27	170.65	-24.66
30.99	122.85	-10.82	90.24	146.75	-17.63	110.90	170.65	-28.23
20.11	122.85	-11.86	69.37	146.75	-19.61	84.37	170.65	-31.67
11.70	122.85	-12.91	53.14	146.75	-21.46	63.73	170.65	-34.90
2.39	122.85	-13.76	40.15	146.75	-23.21	47.23	170.65	-38.08
			29.53	146.75	-24.90	33.72	170.65	-41.09
			20.67	146.75	-26.52	22.47	170.65	-43.94
			13.18	146.75	-28.12	12.94	170.65	-46.84
			4.78	146.75	-30.27	4.78	170.65	-49.79

IaMD-T1			IaMD-T2			IaMD-T3			IaMD-V		
α	E	M	α	E	M	α	E	M	α	E	M
822.18	35.85	0.00	1477.06	35.85	0.00	2344.65	35.85	0.00	2919.46	57.36	0.00
822.18	274.86	-3.94	1477.06	322.66	-7.07	2344.65	370.46	-9.48	2919.46	607.08	-17.03
365.08	274.86	-7.23	668.62	322.66	-13.41	1078.51	370.46	-18.23	1322.30	607.08	-32.37
212.71	274.86	-10.03	399.14	322.66	-19.26	656.47	370.46	-26.45	789.91	607.08	-46.22
136.53	274.86	-12.51	264.40	322.66	-24.63	445.45	370.46	-34.30	523.72	607.08	-58.75
90.82	274.86	-14.68	183.56	322.66	-29.56	318.83	370.46	-41.74	364.01	607.08	-70.42
60.35	274.86	-16.60	129.66	322.66	-34.10	234.43	370.46	-48.79	257.53	607.08	-81.14
38.58	274.86	-18.33	91.16	322.66	-38.31	174.13	370.46	-55.43	181.47	607.08	-90.88
22.26	274.86	-19.97	62.29	322.66	-42.28	128.91	370.46	-61.76	124.43	607.08	-99.90
9.56	274.86	-21.57	39.83	322.66	-46.02	93.74	370.46	-67.78	80.07	607.08	-108.36
			21.87	322.66	-49.72	65.61	370.46	-73.47	44.57	607.08	-116.21
			7.17	322.66	-53.24	42.59	370.46	-78.92	15.54	607.08	-123.77
						23.40	370.46	-84.08			
						7.17	370.46	-89.13			

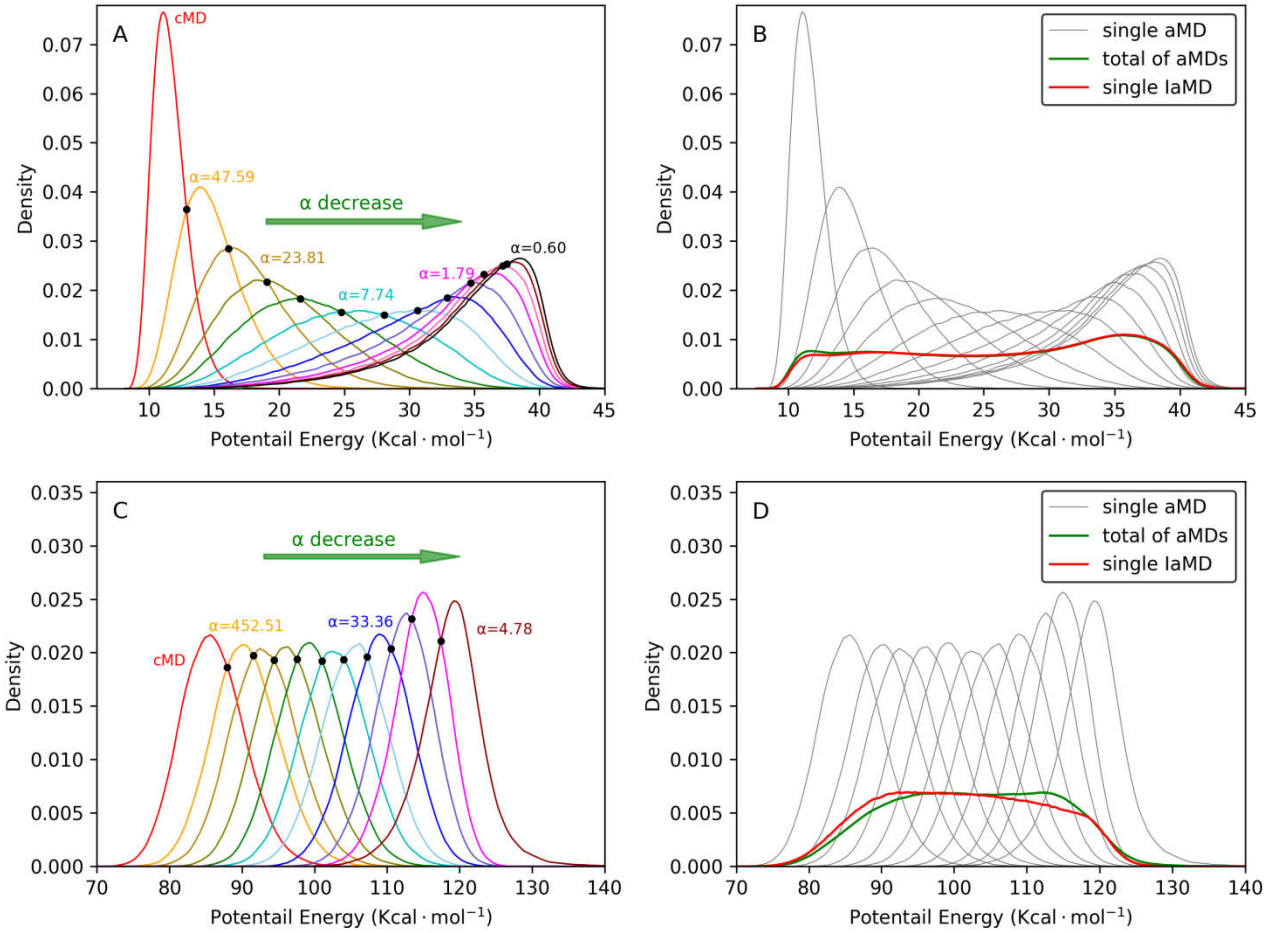

Figure S1. Demonstration of Parameters determining procedure for Alanine Dipeptide (top) and Chignolin (bottom)

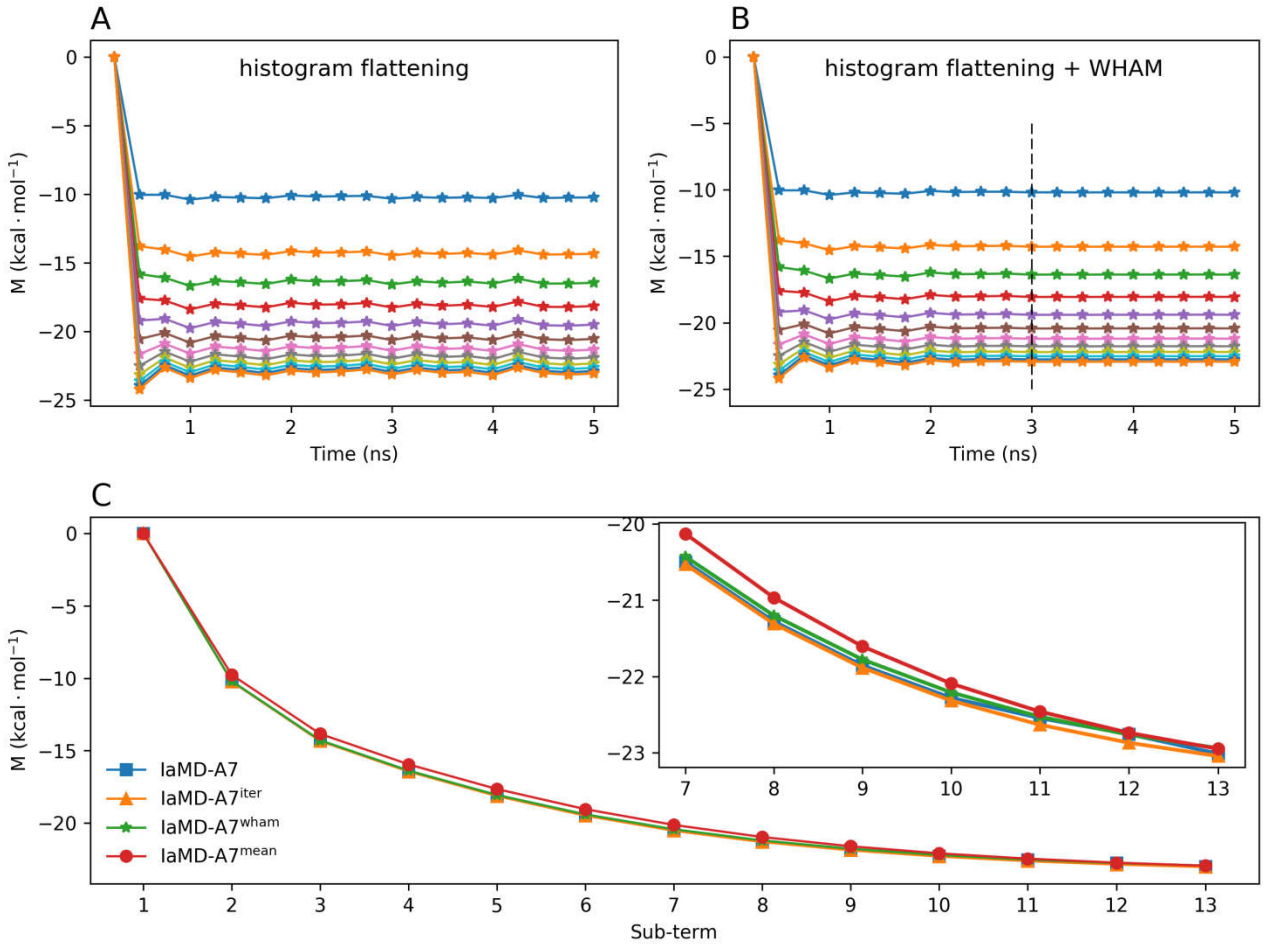

Figure S2. Convergence and comparison of the weight parameters $\{M_i\}$ (that is$-k_BT\ln(n_i)$). (A) the convergence of weight parameters using the histogram flattening; (B) the convergence of weight parameters using the hybrid Approaches (12 rounds histogram flattening + 8 rounds WHAM); (C) the comparison of the weight parameters $\{M_i\}$ obtained from different methods.

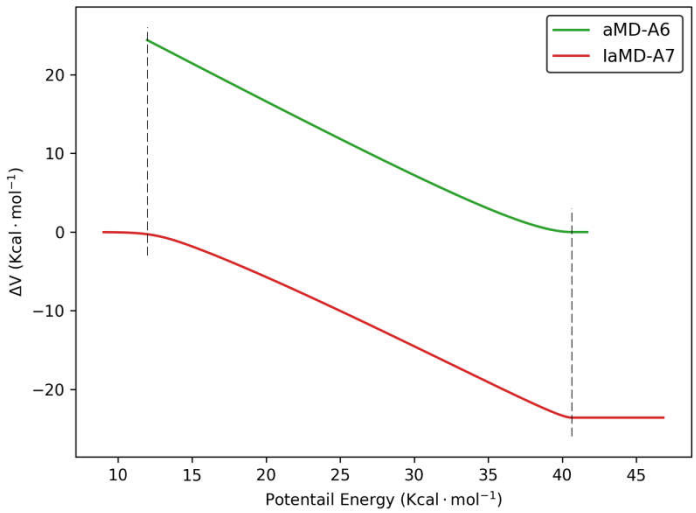

Figure S3. The relationship between torsional potential energy and boost potential of aMD6 and laMD7.

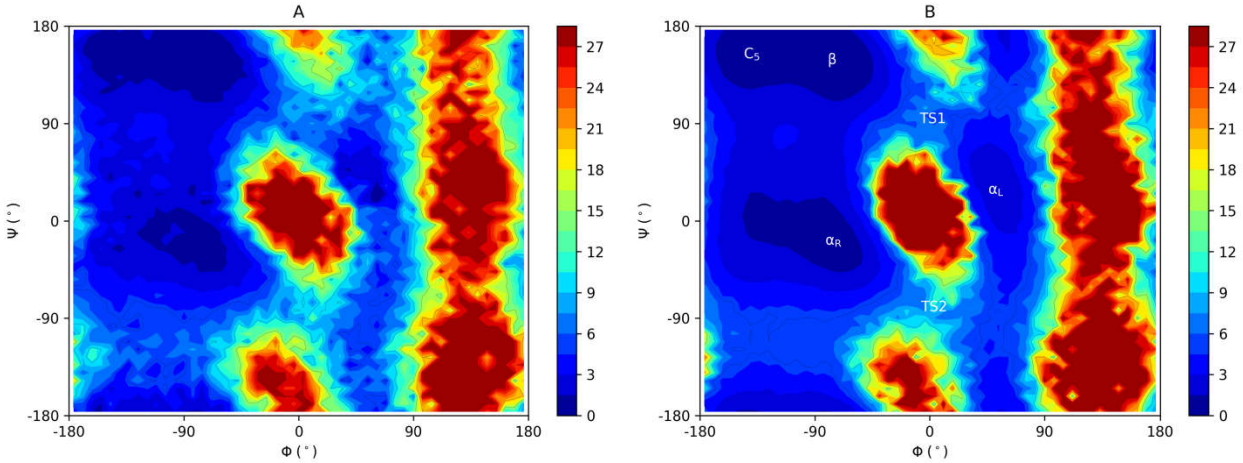

Figure S4. Free energy profiles of Alanine Dipeptide. (A) the result obtained from single aMD-A6 simulation; (B) the result obtained from single IaMD-A7 simulation. The FEPs are calculated from 300 ns simulations with data collected every 0.1 ps. Contour lines are every 1.5 kcal/mol. All φ−ψ plots are sorted in a bin size of 7.2°×7.2°.

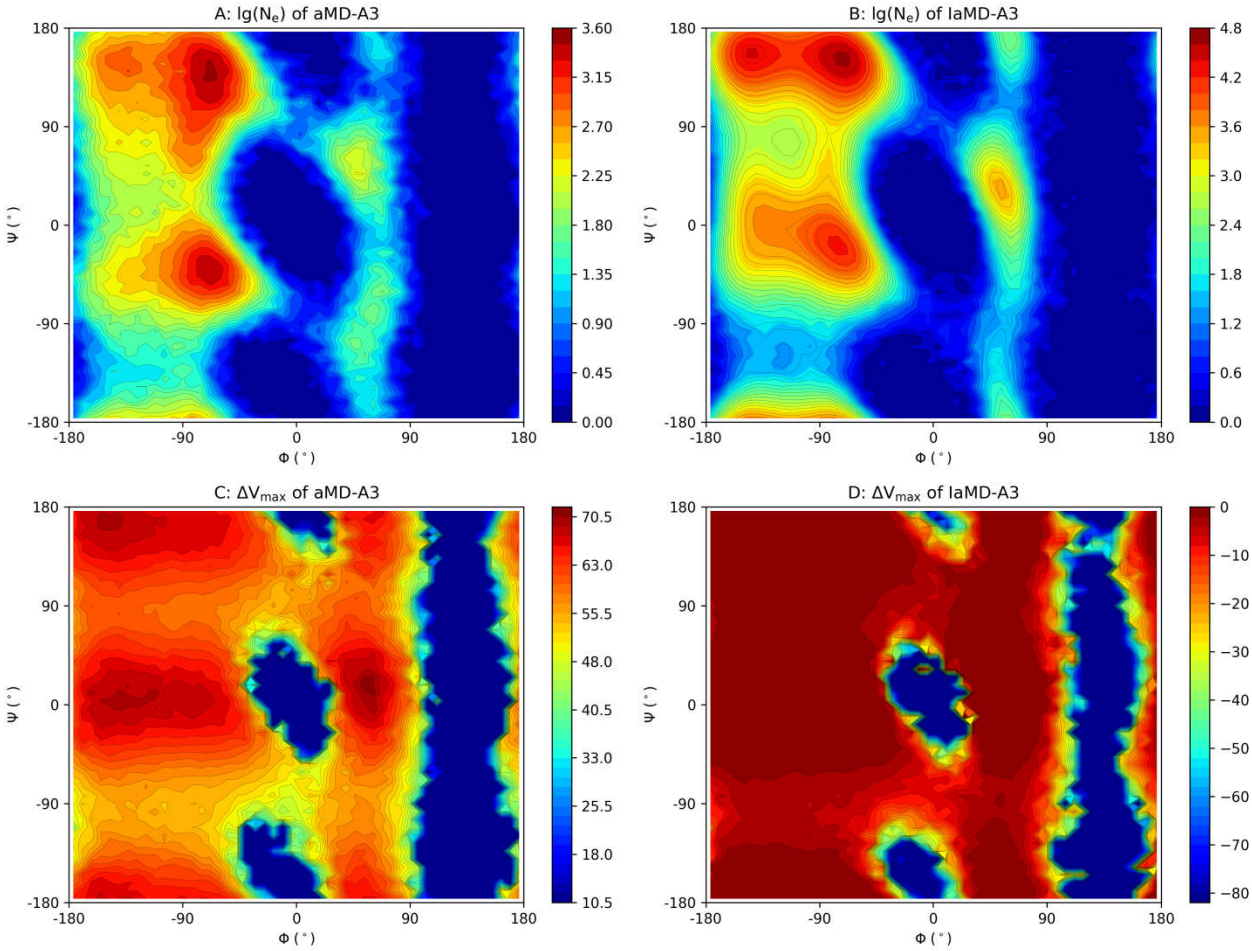

Figure S5. The number of effective sampling points in single 300 ns simulation with (A) aMD-A3 or (B) IaMD-A3 as a function of ϕ and ψ and the maximum boost potential in (C) aMD-A3 or (D) IaMD-A3 simulation as a function of

ϕ and ψ. Color bar of (A) and (B) represents the order of magnitude of the effective sampling point numbers. The unit for the color bar of (C) and (D) is kcal/mol. All ϕ−ψ plots are sorted in a bin size of 7.2°×7.2°.

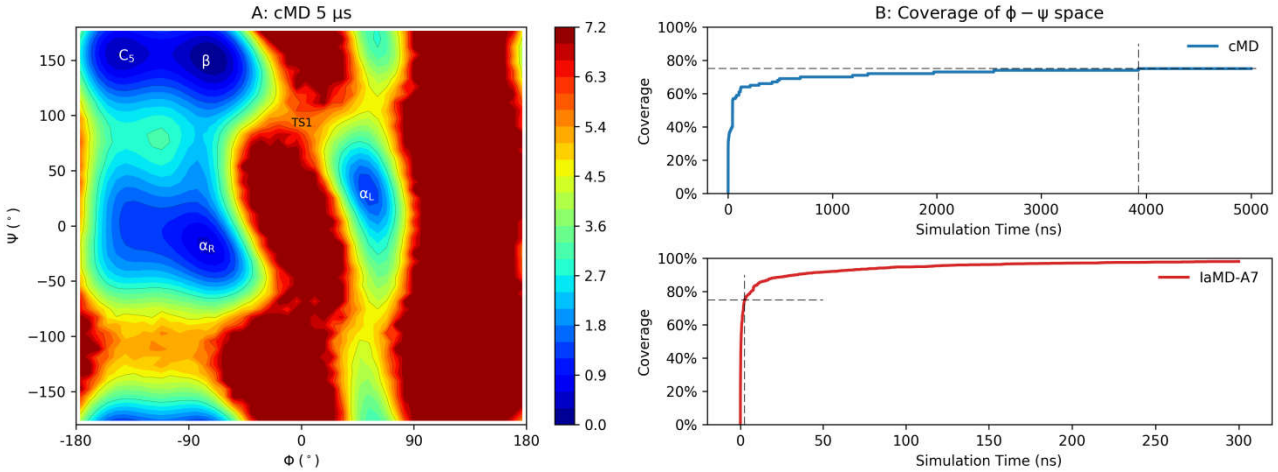

Figure S6. Results of cMD simulation for Alanine Dipeptide. (A) free energy profiles from 5 μs cMD simulation; (B) the coverage over ϕ-ψ space in cMD (top) and IaMD-A7 (bottom) as a function of simulation time.

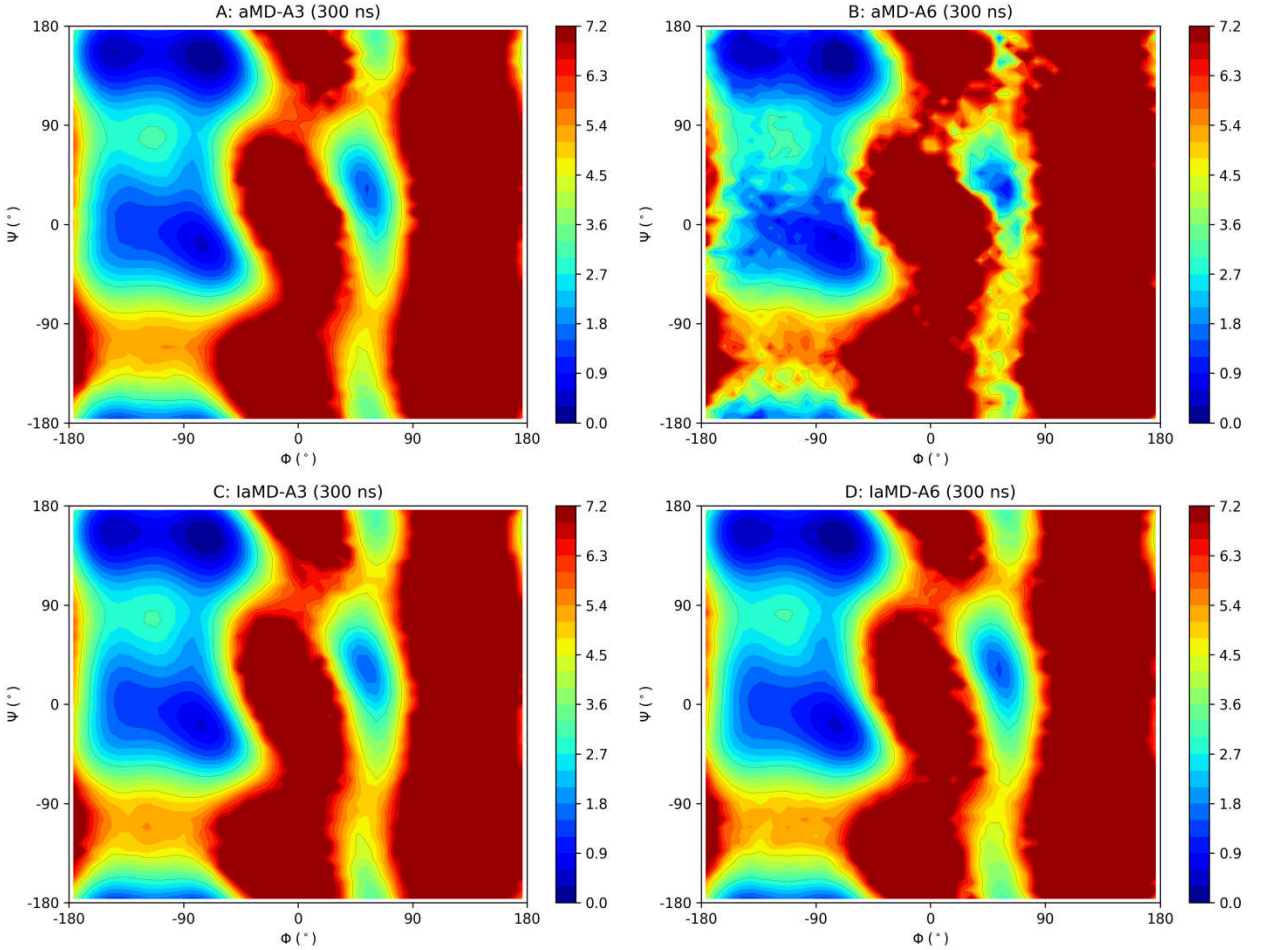

Figure S7. Free energy profiles of Alanine Dipeptide obtained from aMD (top panel) and IaMD (lower panel). (A) aMD-A3 with data collected every 0.012 ps; (B) aMD-A6 with data collected every 0.012 ps; (C) aMD-A6 simulation with data collected every 0.1 ps; (D) IaMD-A3 with data collected every 0.012 ps. All simulation length is 300ns. To

give a clear distinction of free energy for the low-energy regions, the range of free energy represented by different color is 0 to 7.2 kcal/mol and the regions with free energy greater than 7.2 kcal/mol are represented as red. These ϕ−ψ plots are sorted in a bin size of 7.2°×7.2°. Contour lines are every 0.36 kcal/mol.

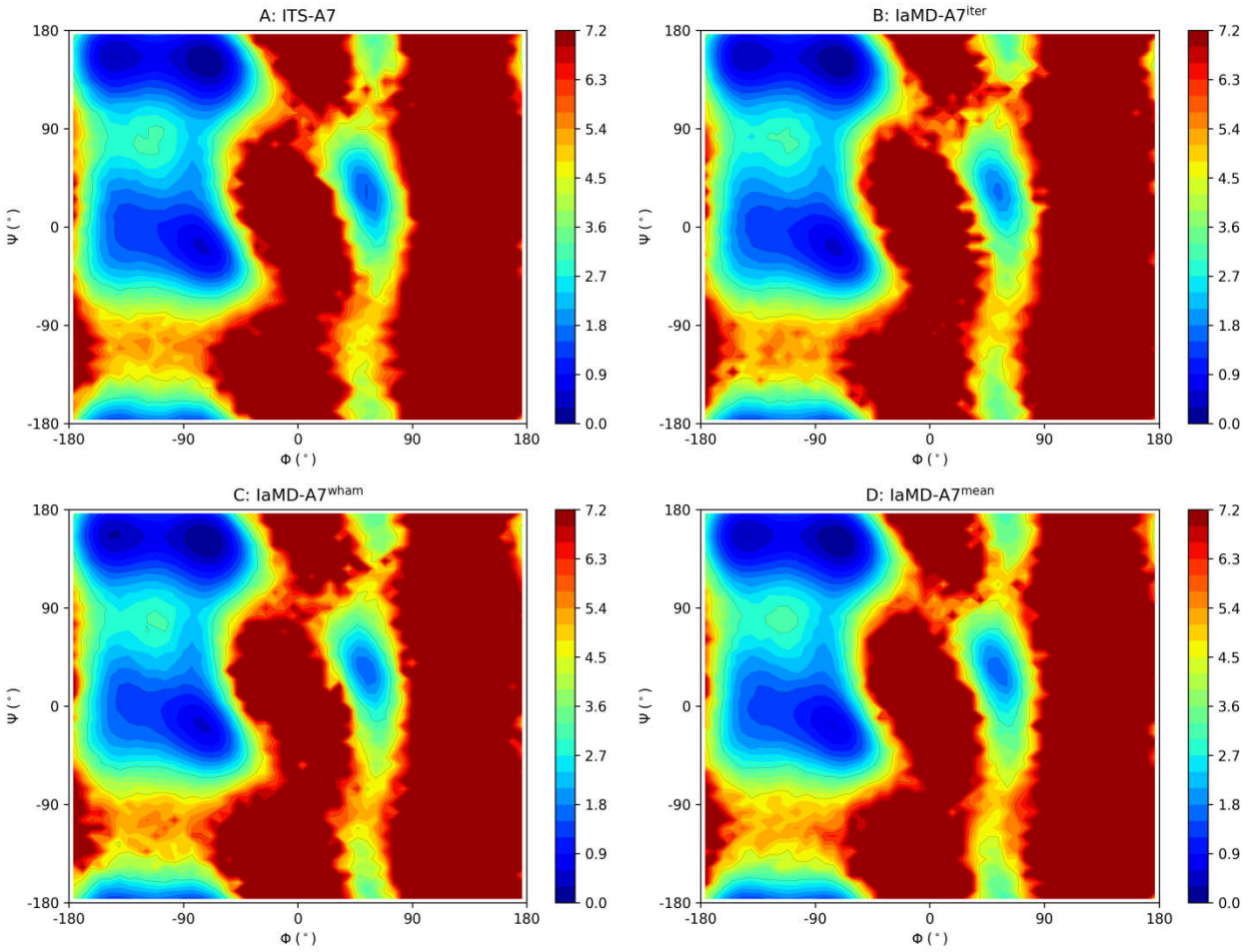

Figure S8. The free energy profiles of single 300 ns simulations with ITS-A8 (left) or IaMD-A7^{m} whose parameters M_i determined using the midpoint of the adjacent two torsional potential distribution peaks of simulations with separate aMD subterms (right). To give a clear distinction of free energy for the low-energy regions, the range of free energy represented by different color is 0 to 7.2 kcal/mol and the regions with free energy greater than 7.2 kcal/mol are represented as red. These ϕ−ψ plots are sorted in a bin size of 7.2°×7.2°. Contour lines are every 0.36 kcal/mol.

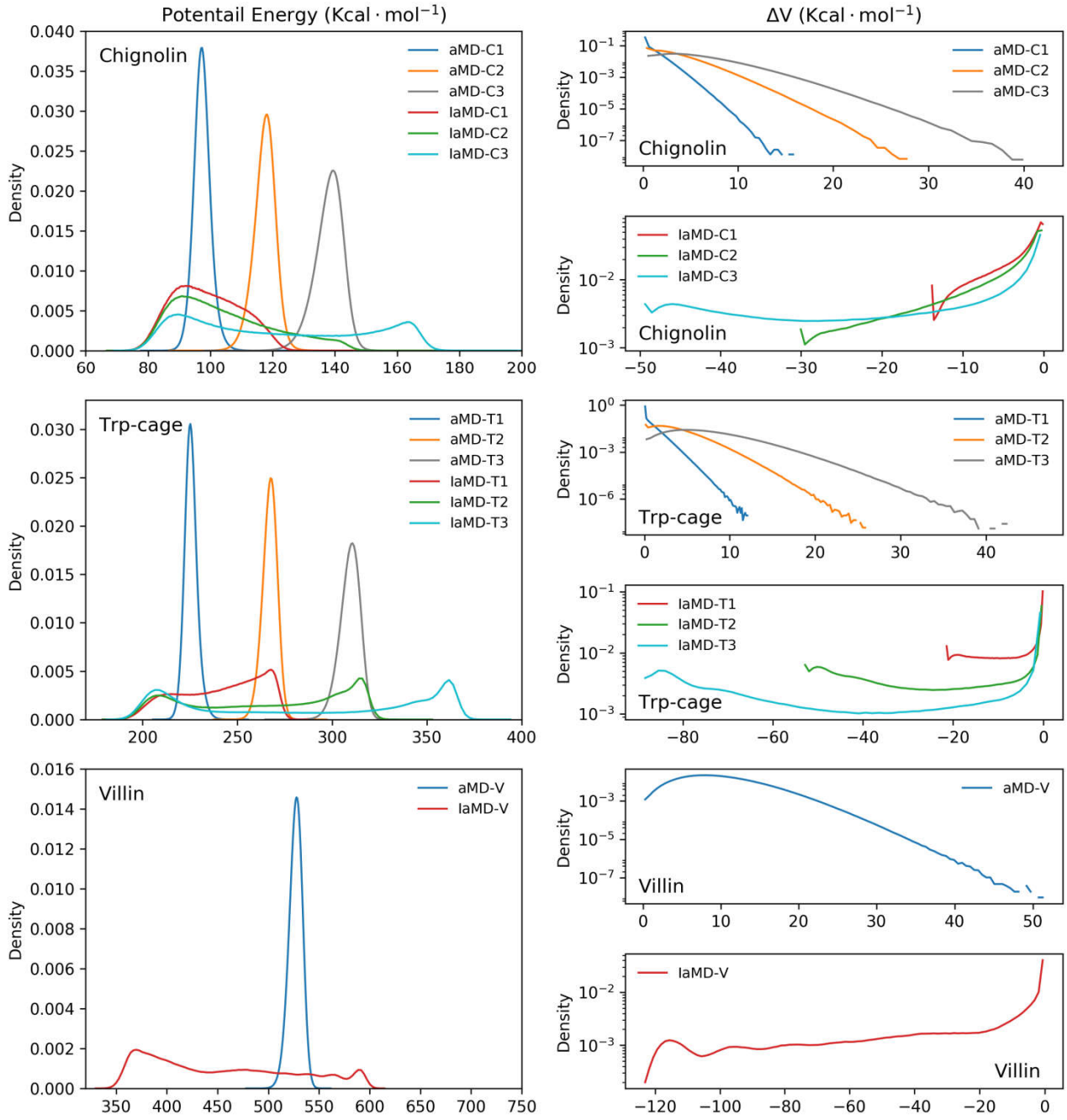

Figure S9. The distributions of torsional potential energies (left column) and boost potentials (right column) in the simulations with aMD or IaMD for Chignolin, Trp-cage, and Villin respectively.

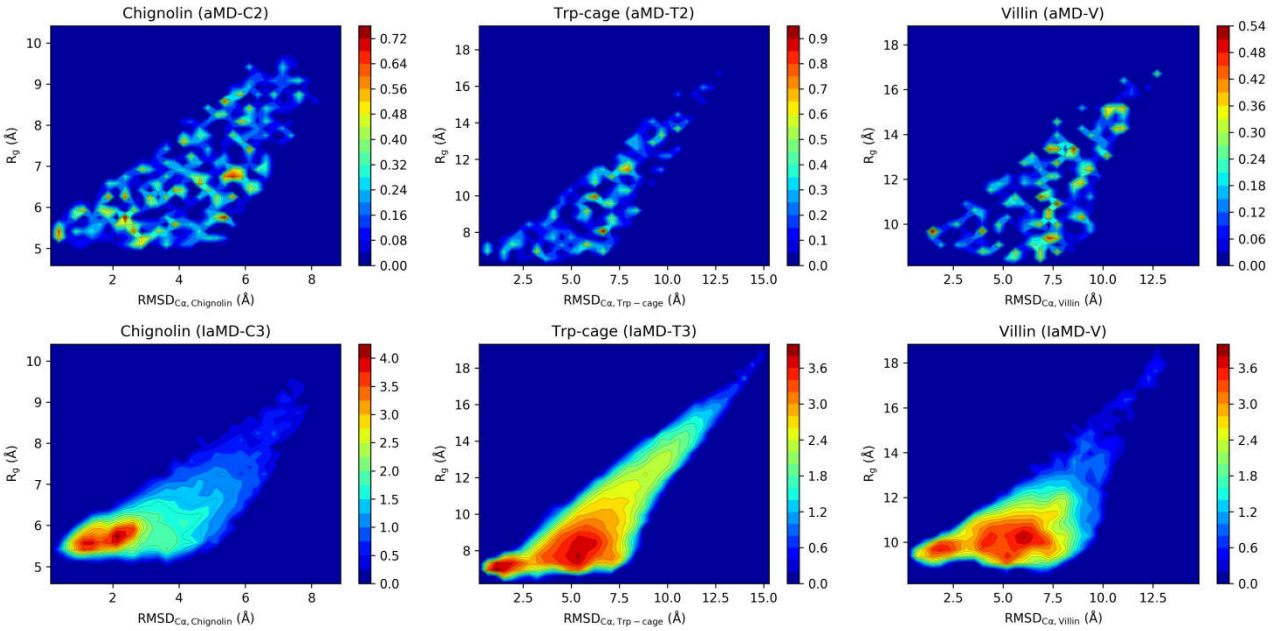

Figure S10. The number of effective sampling points in every 1000 thousand sampling points of simulations with aMD or IaMD for Chignolin, Trp-cage, and Villin Headpiece, as a function of ϕ and ψ. Color bar represents the order of magnitude of the effective sampling point numbers.

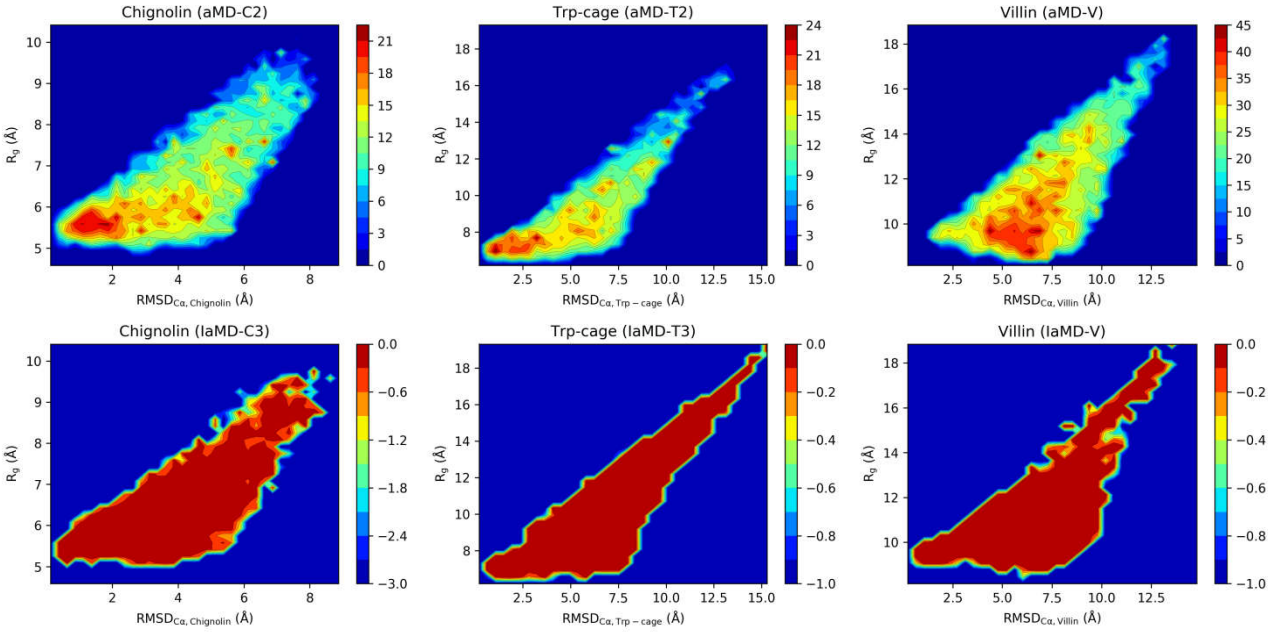

Figure S11. The maximum boost potential of simulations with aMD or IaMD for Chignolin, Trp-cage, and Villin Headpiece, as a function of ϕ and ψ. The unit for color bars is kcal/mol.

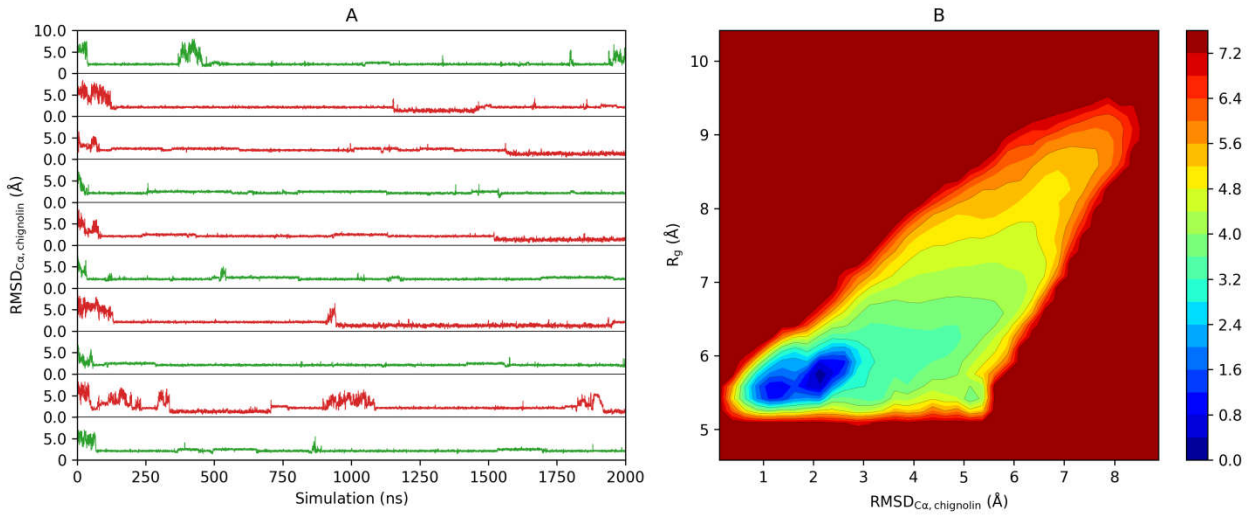

Figure S12. The result of cMD simulations for Chignolin. (A) The RMSD of ten 2 μs simulations with cMD as a function of simulation time. The protein folds into the native in five trajectories (red line) and folds into another state in other trajectories (green line). (B) 2D free energy profile obtained from the trajectories represents the red lines.

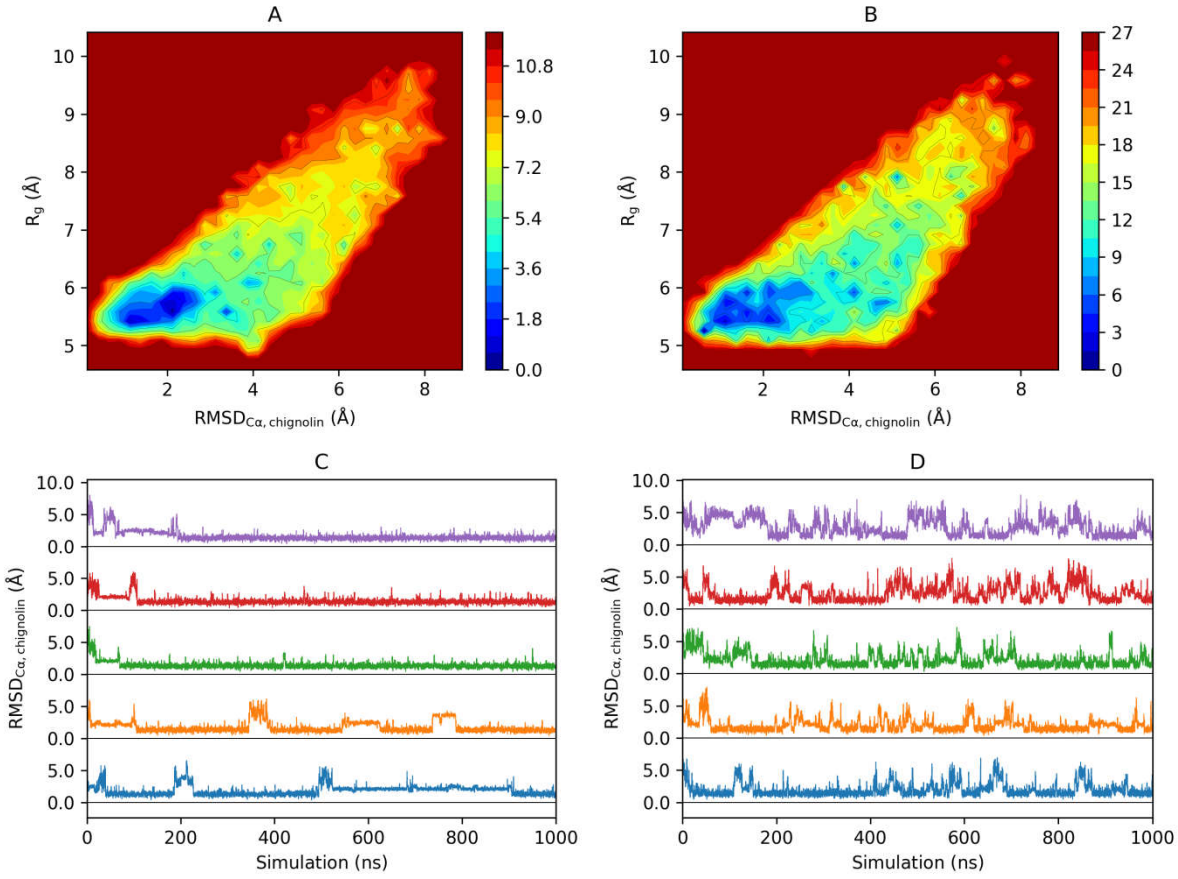

Figure S13. The results of simulations with aMD-C1 and aMD-C3 for the folding of Chignolin. (A) 2D free energy profile of aMD-C1 (in kcal/mol); (B) 2D free energy profile of aMD-C3 (in kcal/mol); (C) RMSD of five 1 μs simulations with aMD-C1 as a function of simulation length; (D) RMSD of five 1 μs simulations with aMD-C3 as a function of simulation length. The free energy profile obtained from total 5 μs simulations for each protocol.

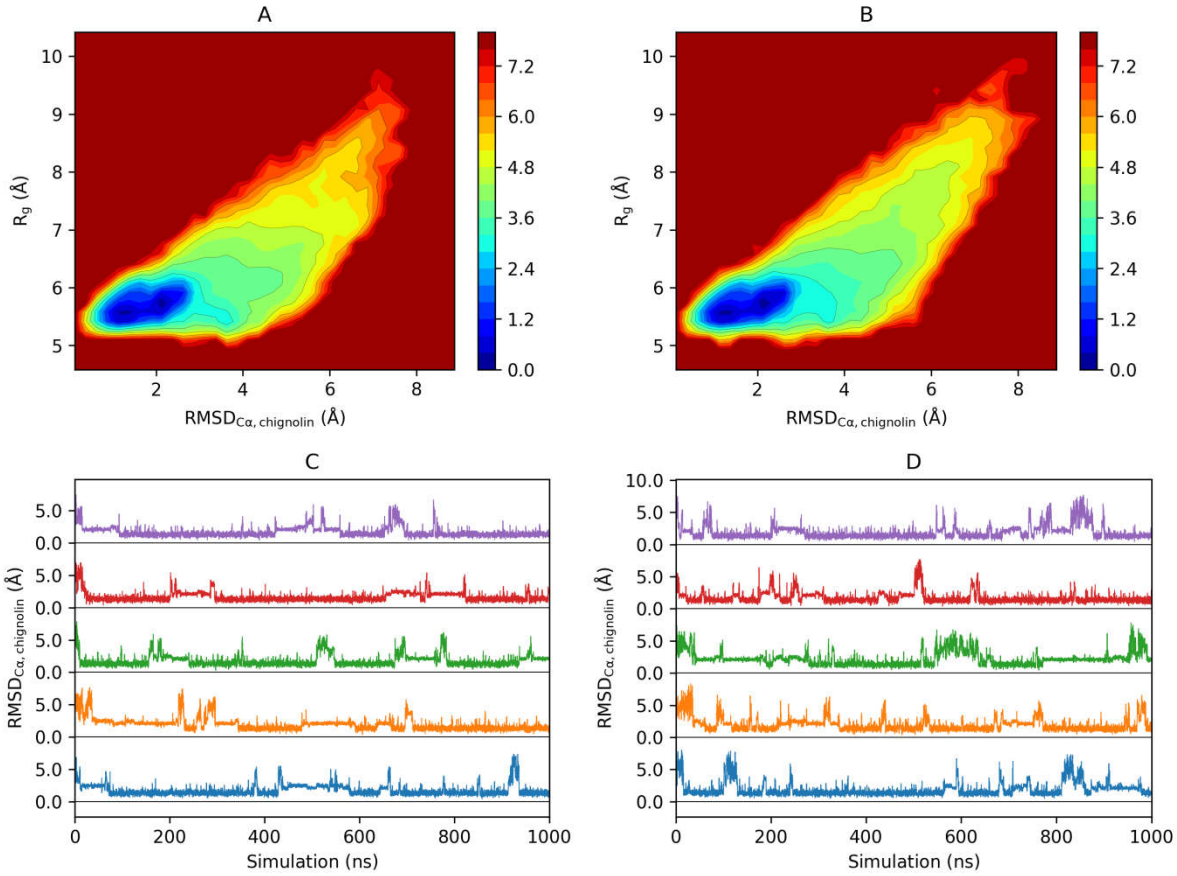

Figure S14. The results of simulations with IaMD-C1 and IaMD-C2 for the folding of Chignolin. (A) 2D free energy profile of IaMD-C1 (in kcal/mol); (B) 2D free energy profile of IaMD-C2 (in kcal/mol); (C) RMSD of five 1 μs simulations with IaMD-C1 as a function of simulation length; (D) RMSD of five 1 μs simulations with IaMD-C2 as a function of simulation length. The free energy profile obtained from total 5 μs simulations for each protocol.

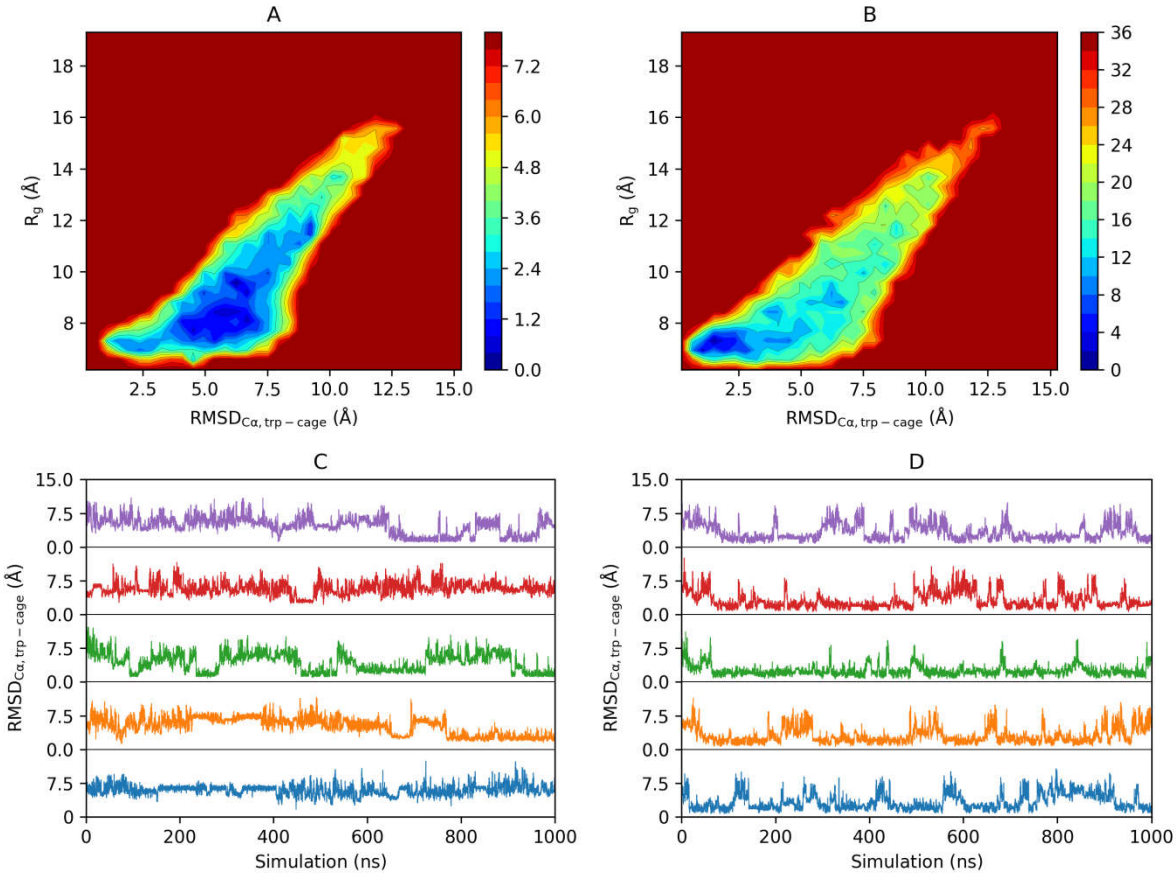

Figure S15. The results of simulations with aMD-T1 and aMD-T3 for the folding of Trp-cage. (A) 2D free energy profile of aMD-T1 (in kcal/mol); (B) 2D free energy profile of aMD-T3 (in kcal/mol); (C) RMSD of five 1 μs simulations with aMD-T1 as a function of simulation length; (D) RMSD of five 1 μs simulations with aMD-T3 as a function of simulation length. The free energy profile obtained from total 5 μs simulations for each protocol.

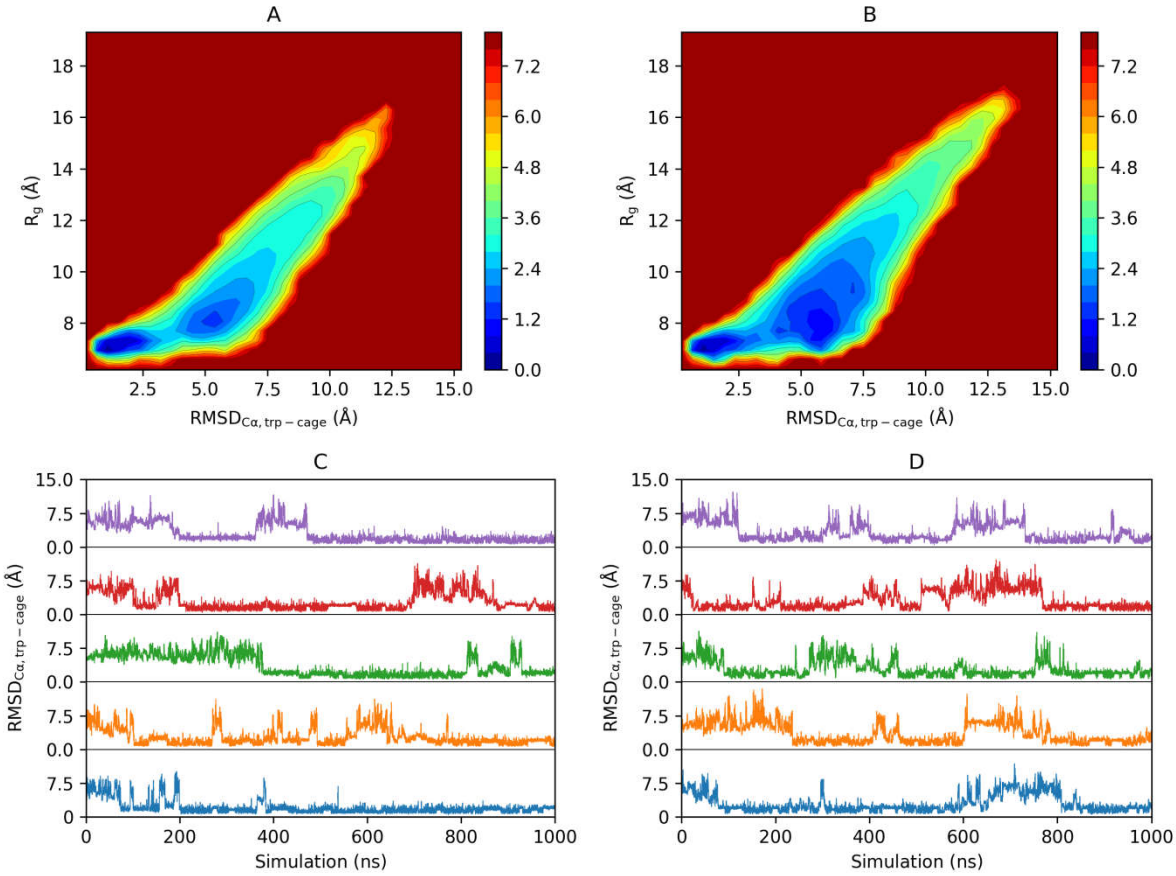

Figure S16. The results of simulations with IaMD-T1 and IaMD-T2 for the folding of Trp-cage. (A) 2D free energy profile of IaMD-T1 (in kcal/mol); (B) 2D free energy profile of IaMD-T2 (in kcal/mol); (C) RMSD of five 1 μs simulations with IaMD-T1 as a function of simulation length; (D) RMSD of five 1 μs simulations with IaMD-T2 as a function of simulation length. The free energy profile obtained from total 5 μs simulations for each protocol.

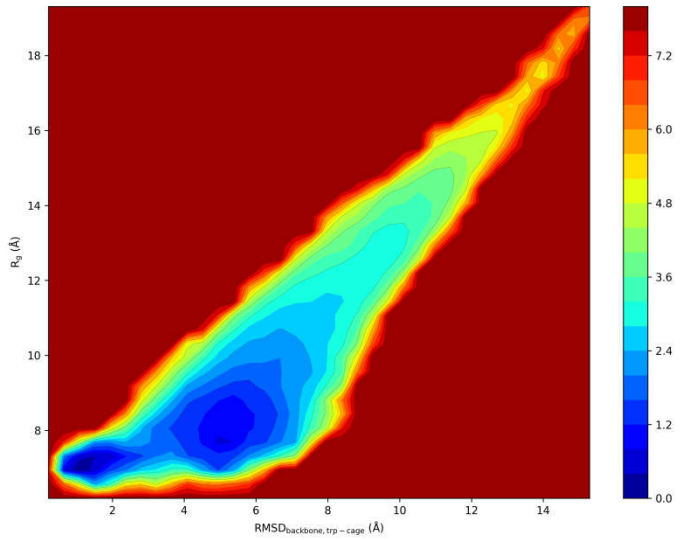

Figure 17. Trp-cage 2D folding free energy profile of IaMD-T3 as a function of backbone RMSD and R_g. The free energy profile obtained from total 5 μs simulations with IaMD-T3. $rmsd_{C\alpha}$ was replaced by $rmsd_{backbone}$ to compare with the result of previous REMD simulation.[20]

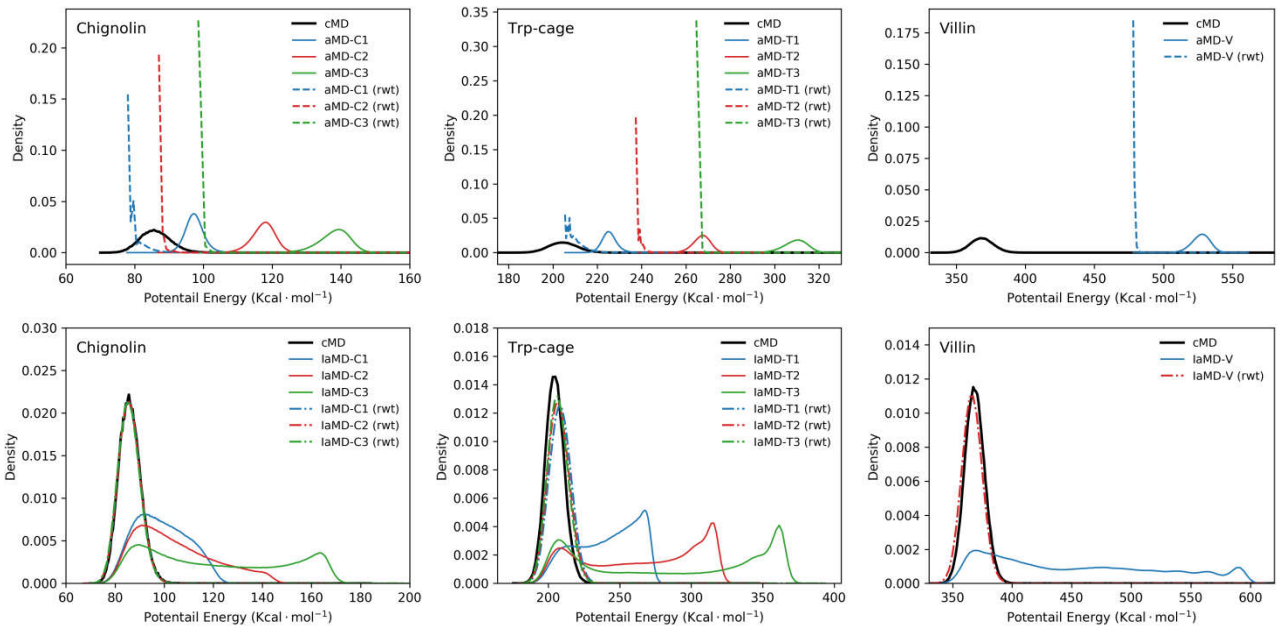

Figure S18. Reweighting results of the torsional potential energy distribution of aMD (top) or IaMD (bottom) simulations for Chignolin, Trp-cage, and Villin Headpiece.

Reference

1. Mori, T.; Hamers, R. J.; Pedersen, J. A.; Cui, Q., Integrated Hamiltonian Sampling: A Simple and Versatile Method for Free Energy Simulations and Conformational Sampling. *J. Phys. Chem. B* **2014,** *118* (28), 8210-8220.

2. Zhao, P.; Yang, L. J.; Gao, Y. Q.; Lu, Z.-Y., Facile implementation of integrated tempering sampling method to enhance the sampling over a broad range of temperatures. *Chem. Phys.* **2013,** *415*, 98-105.

3. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y. T.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S., OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comp. Biol.* **2017,** *13* (7).

4. Eastman, P.; Pande, V. S., OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Comput. Sci. Eng.* **2010,** *12* (4), 34-39.

5. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins-Structure Function and Bioinformatics* **2006,** *65* (3), 712-725.

6. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P., A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003,** *24* (16), 1999-2012.

7. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983,** *79* (2), 926-935.

8. Honda, S.; Yamasaki, K.; Sawada, Y.; Morii, H., 10 residue folded peptide designed by segment statistics. *Structure* **2004,** *12* (8), 1507-1518.

9. Chiu, T. K.; Kubelka, J.; Herbst-Irmer, R.; Eaton, W. A.; Hofrichter, J.; Davies, D. R., High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102* (21),

7517-7522.

10. Andersen, H. C., Molecular-Dynamics Simulations at Constant Pressure and-or Temperature. *J. Chem. Phys.* **1980,** *72* (4), 2384-2393.

11. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995,** *103* (19), 8577-8593.

12. Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Pande, V. S., Accurate and efficient corrections for missing dispersion interactions in molecular Simulations. *J. Phys. Chem. B* **2007,** *111* (45), 13052-13063.

13. Andersen, H. C., Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. *J. Comput. Phys.* **1983,** *52* (1), 24-34.

14. Miyamoto, S.; Kollman, P. A., Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992,** *13* (8), 952-962.

15. Eastman, P.; Pande, V. S., Constant Constraint Matrix Approximation: A Robust, Parallelizable Constraint Method for Molecular Simulations. *J. Chem. Theory Comput.* **2010,** *6* (2), 434-437.

16. Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E., Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015,** *11* (4), 1864-1874.

17. Lin, I. C.; Tuckerman, M. E., Enhanced Conformational Sampling of Peptides via Reduced Side-Chain and Solvent Masses. *J. Phys. Chem. B* **2010,** *114* (48), 15935-15940.

18. Jiang, F.; Wu, Y.-D., Folding of Fourteen Small Proteins with a Residue-Specific Force Field and Replica-Exchange Molecular Dynamics. *J. Am. Chem. Soc.* **2014,** *136* (27), 9536-9539.

19. Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; McCammon, J. A., Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J. Chem. Theory Comput.* **2014,** *10* (7), 2677-2689.

20. Paschek, D.; Day, R.; García, A. E., Influence of water-protein hydrogen bonding on the stability of Trp-cage miniprotein. A comparison between the TIP3P and TIP4P-Ew water models. *Phys. Chem. Chem. Phys.* **2011,** *13* (44), 19840.