VII. SUPPLEMENTARY INFORMATION

A. Validation of flat-bottom restraint implementation

1. Sampled region

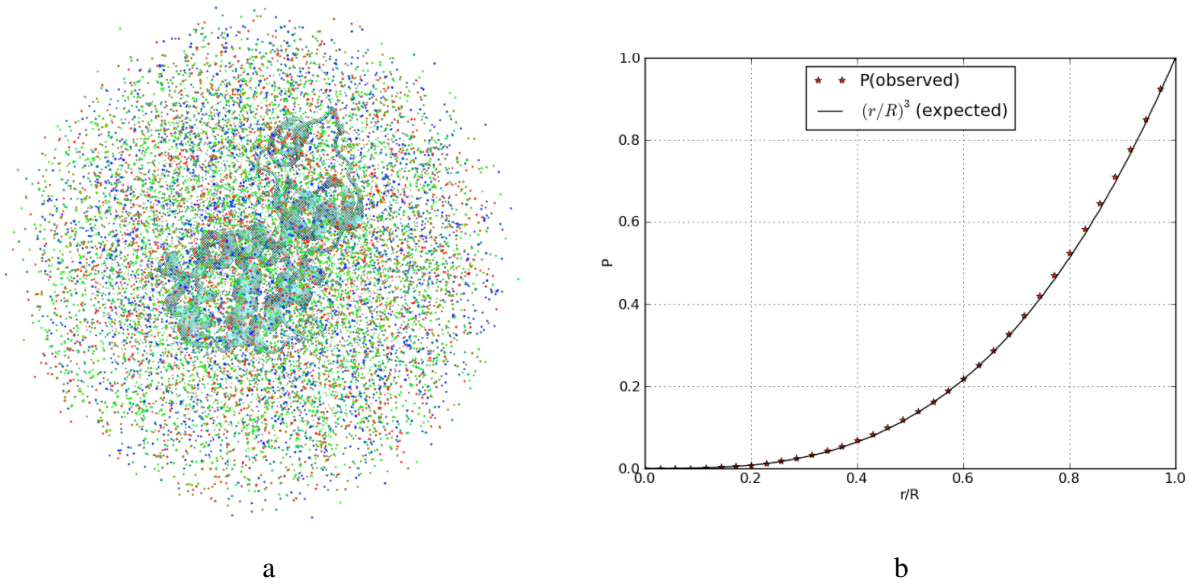
To ensure the flat-bottom harmonic restraint imposed to keep the ligand in the vicinity of the protein was properly implemented, Fig. 8a shows the ligand (1-methylpyrrole) distribution at the fully uncoupled state for a 20000-iteration simulation with only the fully uncoupled state. The dots changing in colors represent the ligand trajectories (one dot for conformation extracted from one iteration). The color changes with iteration following a RGB scale. In other words, red, green and blue dots represent iterations at the early, middle and late stage of the simulation. The protein is shown in the figure for scale purposes alone. As shown in the figure, it qualitatively follows the predicted uniform distribution.

2. Restraint distance distribution

We then examined the distribution in space of the coupled replica to make sure the flat-bottom restraint was correctly implemented within the cutoff radius we defined, we called this the restraint distance distribution. We compared the distribution of the the center of geometry of the ligand relative to that of the protein at each iteration to the uniform distribution within a given cutoff radius r_0.

$$P(r) = \frac{N(r)}{N_{total}} = \left(\frac{r}{r_0}\right)^3 \tag{14}$$

where r is the distance between protein and ligand centers of geometry, $N(r)$ is the number of samples inside a the sphere centered at the protein center of geometry with the radius of r, and N_{total} is total samples generated during simulation. As shown in Fig. 8b, the observed curve matches perfectly with the expected curve out to the beginning of the harmonic wall created by the flat-bottomed potential. The uniform distribution also validates the implementation of the Langevin integrator.



a b

FIG. 8: *(a)* **Qualitative demonstration of correct sampling with the flat bottom potential**. 3D trajectories at the fully uncoupled states of a 20000-iteration simulation with only samples from the fully uncoupled state. The dots changing in colors represent the distance along the ligand trajectories, from red to blue, representing rapid decorrelation within the volume. The protein is shown in the figure to show the scale. It qualitatively follows a uniform distribution; *(b)* Quantitative proof: probability distribution of samples within radius r_0 at the fully uncoupled states. As shown, the observed curve matches perfectly with the expected curve.

B. The tests on determining HREMD simulation key parameters

A series of runs were performed to examine the sensitivity of the simulation efficiency on simulation parameters, including the number of additional fully coupled and fully uncoupled states, the size of the Monte Carlo displacements, and the cutoff radius of the flat bottom sphere using an earlier version of the code. To make this parameter space feasible to search, we fixed the number of states at 24, as this was portable between configurations of 6 or 8 GPUs on the accessible clusters simulations were run on.

We first performed a 2 ns simulation using a reasonable set of parameters as shown in Table III. With this baseline we then performed another 10 simulations changing in four key parameters: the number of fully coupled states (coupled), the number of fully uncoupled states (uncoupled), the size of the Monte Carlo displacements (MC), and the cutoff radius of the flat bottom sphere (cutoff).

Since the goal of this study was to identify binding sites and poses, we first used the ability of the set of parameters to identify the true binding site as criteria, specifically, if the true binding site could be identified (binding site identified), if the binding site cluster the most populated cluster (binding site biggest), the ranking of the binding site cluster in all clusters (ranking), the occupancy fraction of the binding site cluster (occupancy) and the time in terms of iterations it took to find the experimental binding site (time to binding site). To get a sense of the ability and efficiency of the parameter set to sample, we calculated another two parameters. One was the time it took to discover half of the total number of clusters (time half). We also examined the mixing time in state spaces. As discussed in Ref. [38], this mixing time can be calculated from the second eigenvector of the state transition matrix.

Rigorous tests on the parameters were not under the scope of this study, so only simulations with length of only 2 ns were performed, making the results not entirely reliable for large scale validation. For example, as shown in Table III, the true binding site is not identified as the most populated cluster, which is not true for a longer simulation (15 ns), as seen from the main paper. However, we can still see that most of the parameters are not sensitive to the simulation parameters tested.

(a) Monte Carlo displacements (MC). The Monte Carlo displacement has a significant influence on results (by comparing run 1 to 4), since changing it from the baseline displacement 1 nm to either 2 nm, 3 nm or 0 nm leads to a failure in locating the binding site. So 1 nm was chosen.

(b) Number of fully uncoupled states (uncoupled). A smaller or bigger number of uncoupled states compared to the baseline parameter (3) all result in an increase in the time to locate the binding site (by comparing run 1, 5 and 6). Though this was not a criteria to choose the number of fully uncoupled states, but since the test for this category was not sensitive to the parameters, using the baseline parameters was a safe choice.

(c) Number of fully coupled states (coupled). Similarly the test for this category was not sensitive to the parameters, but we do find that increasing it makes it faster to locate the binding site (by comparing run 1, 7 and 8). However, the increase the number of coupled states would either increase the total number of states or decrease the intermediate states with a fixed total number of states, making the simulations more time-consuming and less converged, respectively, combined with the fact that other measures didn't change significantly, we used 6 as the number coupled states.

TABLE V: **Tests to estimate key HREMD simulation parameters.**

run	1	2	3	4	5	6	7	8	9	10	11
uncoupled	3	3	3	3	1	9	3	3	3	3	3
coupled	3	3	3	3	3	3	1	9	3	3	3
MC step size (nm)	1	0	2	3	1	1	1	1	1	1	1
cutoff (Å)	35.34	35.34	35.34	35.34	35.34	35.34	35.34	35.34	32	30	25
binding site identified?	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
binding site largest?	No	No	No	No	No	No	No	No	No	No	Yes
ranking	4/8[a]	11	0/9	0/6	7/13	4/12	4/10	3/12	2/9	4/7	1/8
occupancy (%)	7.6	0	0	0	5.7	5.3	3.8	8.9	16.5	5.1	33.4
first time to binding site[b]	872	N/A	N/A	N/A	1641	1076	1083	88	179	286	376
time half of clusters found	62	66	29	97	24	133	53	46	44	43	57
mixing time	12.8	17.8	12.7	14.7	13.0	26.3	15.6	17.2	14.2	15.3	15.2

[a] x/y represents ranked number x out of a total number of y clusters.

[b] time to binding site, time to half of clusters identified and mixing time are in units of iterations between swaps.

(d) Cutoff radius (cutoff). A decrease in the cutoff radius surprisingly made the results more promising (by comparing run 1, 10 and 11), since for run 11, the binding site cluster jumped to the most populated cluster for the first time for these short simulations. Nevertheless, we found that this cutoff radius resulted area of the protein excluded from the flat-bottom sphere, making this discovery a fluke since the binding pocket was buried inside the protein. Hence, we didn't change our criteria in determining the cutoff radius.