Mixed-probe simulation and probe-derived surface topography map analysis for ligand binding site identification in membrane proteins

*Abdallah Sayyed-Ahmad and Alemayehu A. Gorfe**

University of Texas Health Science Center at Houston, Department of Integrative Biology and Pharmacology, 6431 Fannin St., Houston, Texas 77030

***Corresponding author:** Tel: 713-500-7538; Fax: 713-500-7444; E-mail: Alemayehu.G.Abebe@uth.tmc.edu

Probe binding does not significantly affect K-Ras dynamics or membrane binding

Dynamics is important for identifying ligand-binding sites that might be hidden in an average experimental structure. At the same time, the probes should not drastically alter the structure or internal dynamics of the target protein. Our mixed-probe simulations did not result in major conformational changes on K-Ras. For example, the time evolution of root mean square deviations (Figure S2) and radius of gyrations (Figure S3) show very small deviations from the initial structure. For example, the mean RMSD over all of the simulations in OS2 and OS1 was 0.8 and 0.9 ± 0.2 Å, respectively (Figure 5D in main text). Similarly, the time-averaged root mean square fluctuations (RMSF, Figures 5E and S6) show very small backbone fluctuations except at the flexible switch regions and the C-terminus. Figure 5B (main text) and Figure S5 show the orientational fluctuations with respect to the POPC/POPS bilayer as quantified by z_{lb1} and z_{lb2}, the vertical distances of lobe 1 (residues 1-86) and lobe 2 (residues 87-166) from the lipid bilayer center. For example, compared with the starting orientations $(-45.7, -36.6)$ and

$(-41.0, -47.8)$ the orientation of G12D K-Ras changed little during both the OS1 and OS2 simulations (mean: $(-48 \pm 2, -39 \pm 2)$ and $(-41 \pm 2, -48 \pm 3)$). This shows that K-Ras catalytic domain-lipid interaction remained stable, because the timescale for membrane re-orientation is likely significantly longer (see ref [24] in the main text) than the current simulations (60ns). For the same reason, the simulations did not capture the potential impact of the mutations on membrane orientation.

Since the level of fluctuation observed in the current simulations is also observable in probe-free simulations (e.g., ref [12] in the main text), we conclude that the probes have negligible effect on the dynamics of K-Ras within the current simulation timescales. Therefore, any potential difference in probe binding among the K-Ras mutants or orientations arises from small local differences rather than large-scale conformational changes. As a result, the global propensity of each probe type to interact with the protein is similar among the mutants and orientations (Table S1). However, there are also important differences. For example, pocket p1 has substantially lower affinity for probes in OS2 of each mutant, whereas the probe-binding propensity of p3 is weaker in OS1 of G13D (Table S2). These and other differences are discussed in the main text.

Convergence of probe-protein interaction

Convergence of probe binding in both single- and mixed-probe pMD simulations has been discussed extensively in previous reports (e.g. ref. [8-13] in the main text). In the case of K-Ras, our previous analyses in solution and membrane environments showed that multiple relatively short simulations allow for adequate sampling and yield convergent results (ref. [12, 16] main text). Inspection of the time evolution of probe-protein interactions in each of the current 60ns-long simulations yielded essentially the same conclusion, which is that there is no systematic drift despite sizable fluctuations (see the standard deviations in Tables S1 and S2). In particular,

probe occupancy of the four known allosteric ligand binding pockets p1-p4 does not significantly change during the simulations after about 5-10ns. For example, in figure S9 we show the time evolution of the residue-normalized occupancy of the pockets during the concatenated trajectory of G12D K-Ras in the OS2 orientation (excluding the first 10ns of each run). This very simple measure of probe-binding propensity demonstrates that the probes quickly engage the residues of each pocket and remain in contact throughout the simulations with only small variation among the different trajectories. For example, dividing the concatenated data into two equal blocks and calculating the mean occupancy yields essentially the same values for p1, p3 and p4. Even for the comparatively weakly reactive p2 that exhibits the largest fluctuations, the mean occupancy in the two blocks is 83.4 % and 86.6%.

Table S1: Average total number of probe-protein contacts*.

	OS1			OS2		
	G12D	**G12V**	**G13D**	**G12D**	**G12V**	**G13D**
All probes	160±20	150±20	150±20	150±20	160±20	150±20
Isobutane	34±8	29±9	32±10	32±9	35±10	33±9
Isopropanol	24±8	23±7	21±7	20±7	21±8	21±7
Acetamide	22±8	22±8	22±9	21±8	21±8	22±8
Acetate	22±9	21±8	17±7	21±8	23±9	22±7
Acetone	15±7	15±6	15±7	14±7	16±7	14±6
DMSO	13±6	14±6	13±6	13±6	12±6	12±6
Urea	27±9	26±8	25±9	25±9	30±10	25±9

*, Time-averaged number of contacts between the heavy atoms of probes and the K-Ras catalytic domain using a cutoff of 4.5 Å.

Table S2: Average number of contacts between probes and residues at pockets p1-p4.

	OS1			OS2		
	G12D	**G12V**	**G13D**	**G12D**	**G12V**	**G13D**
p1	17±5	16±5	16±5	8±5	13±5	8±5
p2	4±3	5±3	4±3	4±3	4±3	4±3
p3	13±5	12±5	9±5	13±5	12±5	12±5
p4	14±5	14±5	14±5	11±5	13±5	12±5

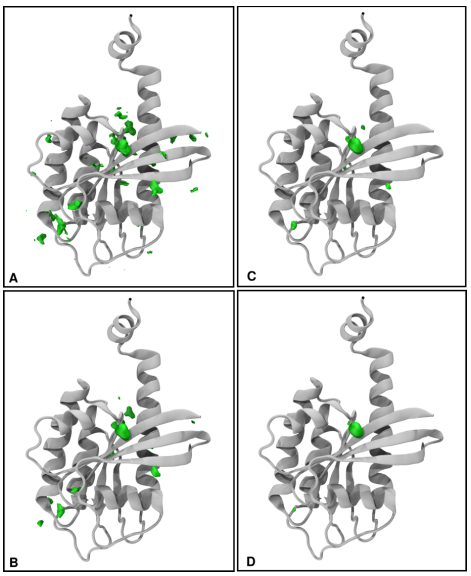

Figure S1: Effect of recursive noise filtering on the iso-surface of the density of urea. The filter was not applied in (A) and was applied once (B), two times (C) and three times (D).

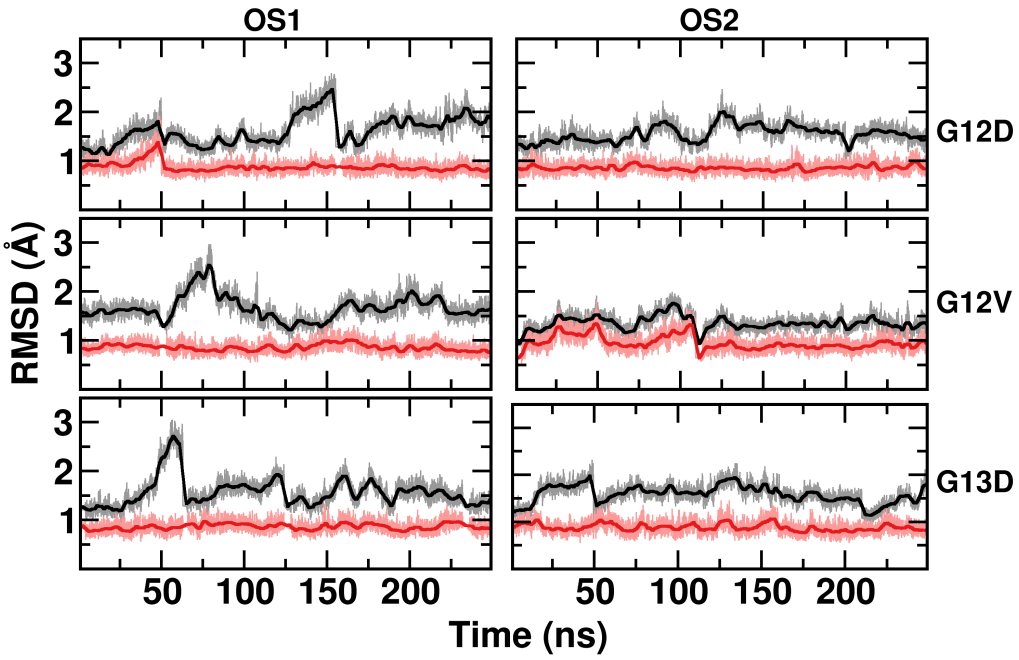

Figure S2: Time evolution of backbone root-mean square deviations (RMSD). RMSD of K-Ras including *black) and excluding SI and SII (red) sampled during the concatenated trajectory of the five simulations (excluding the first 10ns) of each mutant in the two orientation states. The reference was the G12D K-Ras X-Ray structure (PDB: 4DSO). Thick lines represent data sampled every 1 ns.

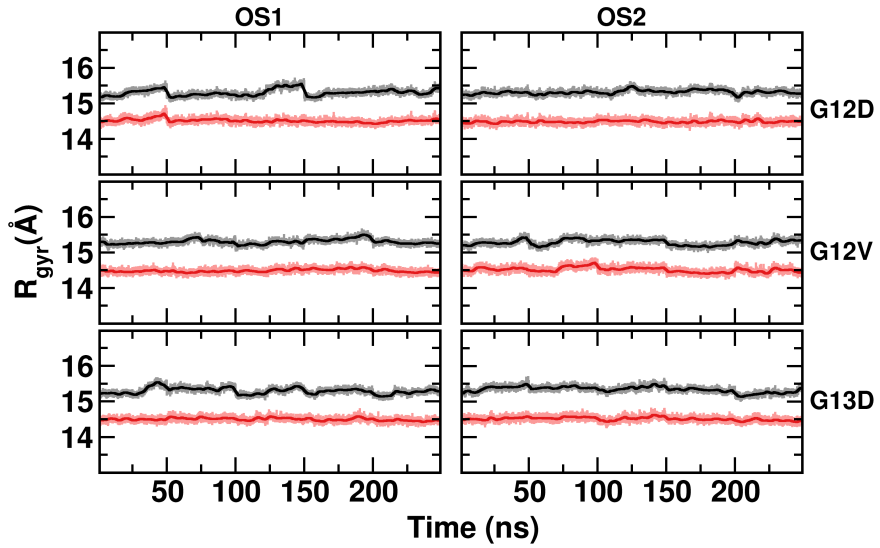

Figure S3: Time evolution of radius of gyration. Radius of gyration of the K-Ras catalytic domain including (black) and excluding SI and SII (red) sampled during the concatenated trajectory of five simulations of each mutant in the two orientation states.

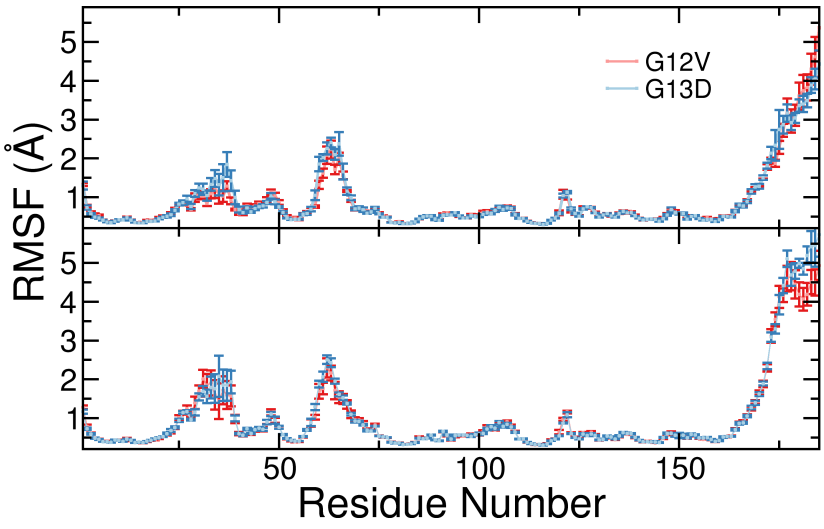

Figure S4: Ensemble-averaged C_α atom root-mean square fluctuations of G12V (red) and G13D (blue) K-Ras in OS1 (bottom) and OS2 (top).

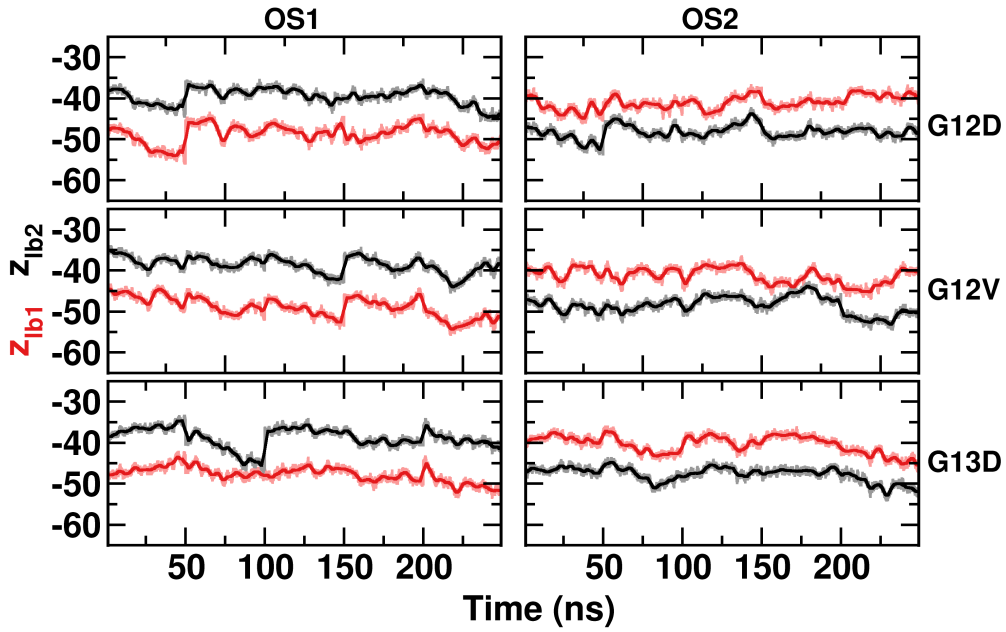

Figure S5: Time evolution of z_{lb1}(red) and z_{lb2} (black) representing the distance of the center of mass of lobe 1 (residues 1-86) and lobe 2 (residues 87-166) of K-Ras from the center of the POPC/POPS bilayer, respectively. The data is from a concatenated trajectory of five independent simulations for each mutant and each membrane orientation state, excluding the first 10ns data of each run.

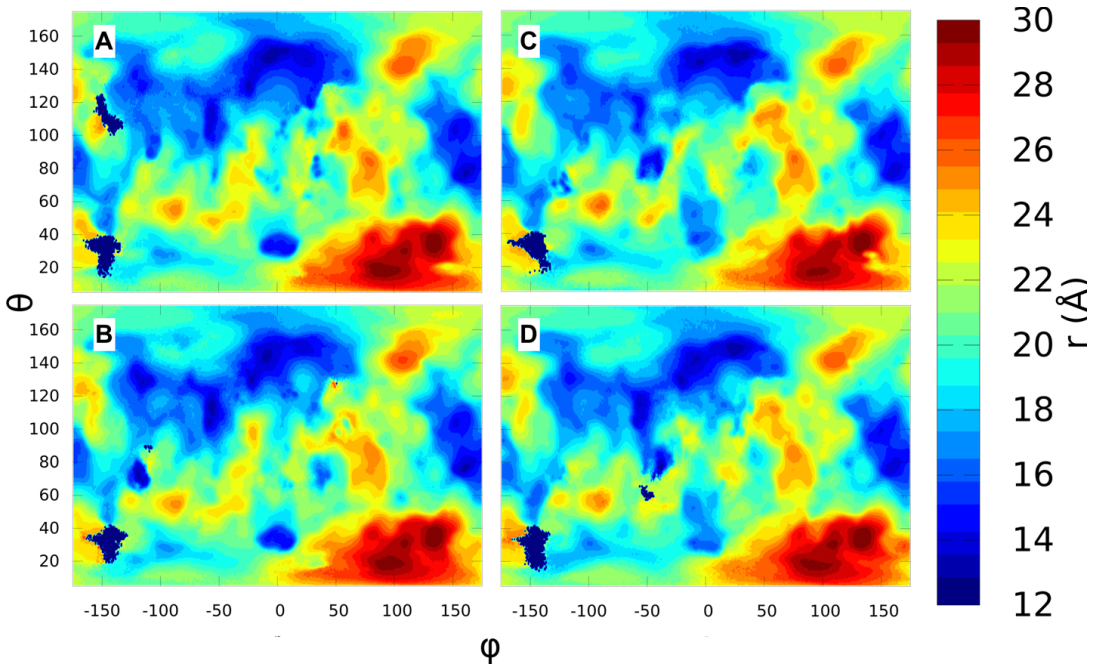

Figure S6: Protein surface topography maps of (A and C) G12V and (B and D) G13D K-Ras in OS1 (left) and OS2 (right) membrane orientations.

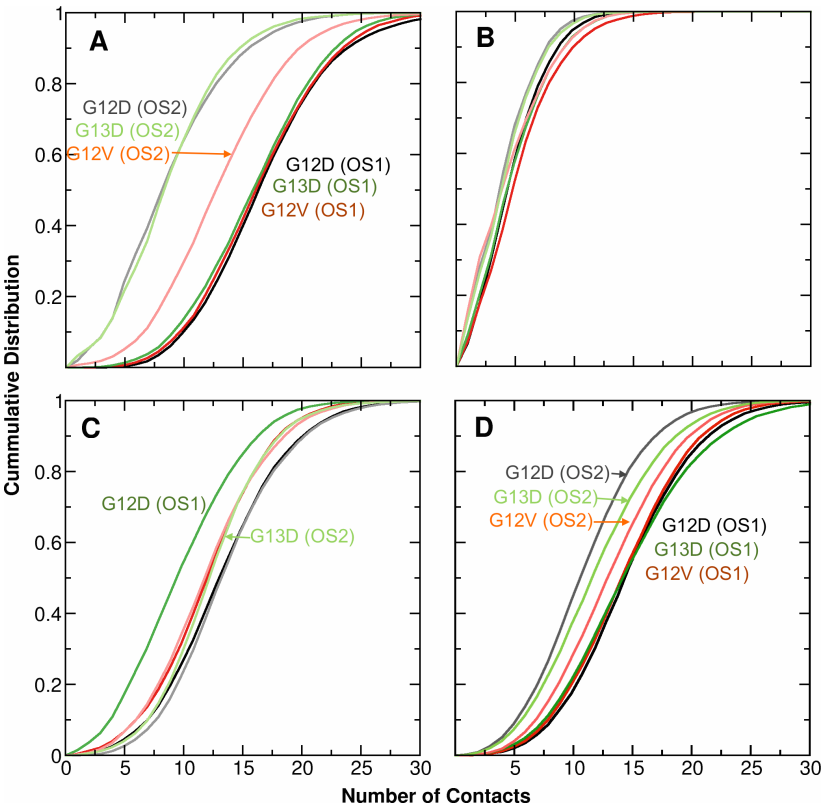

Figure S7: Cumulative distribution of the total number of contacts of probes with residues at pocket p1 (A), p2 (B), p3 (C) and p4 (D). G12D, G12V and G13D are in black, red and green, respectively, with OS1 in darker and OS2 in lighter colors. Where applicable, important differences between the two orientation states are annotated.

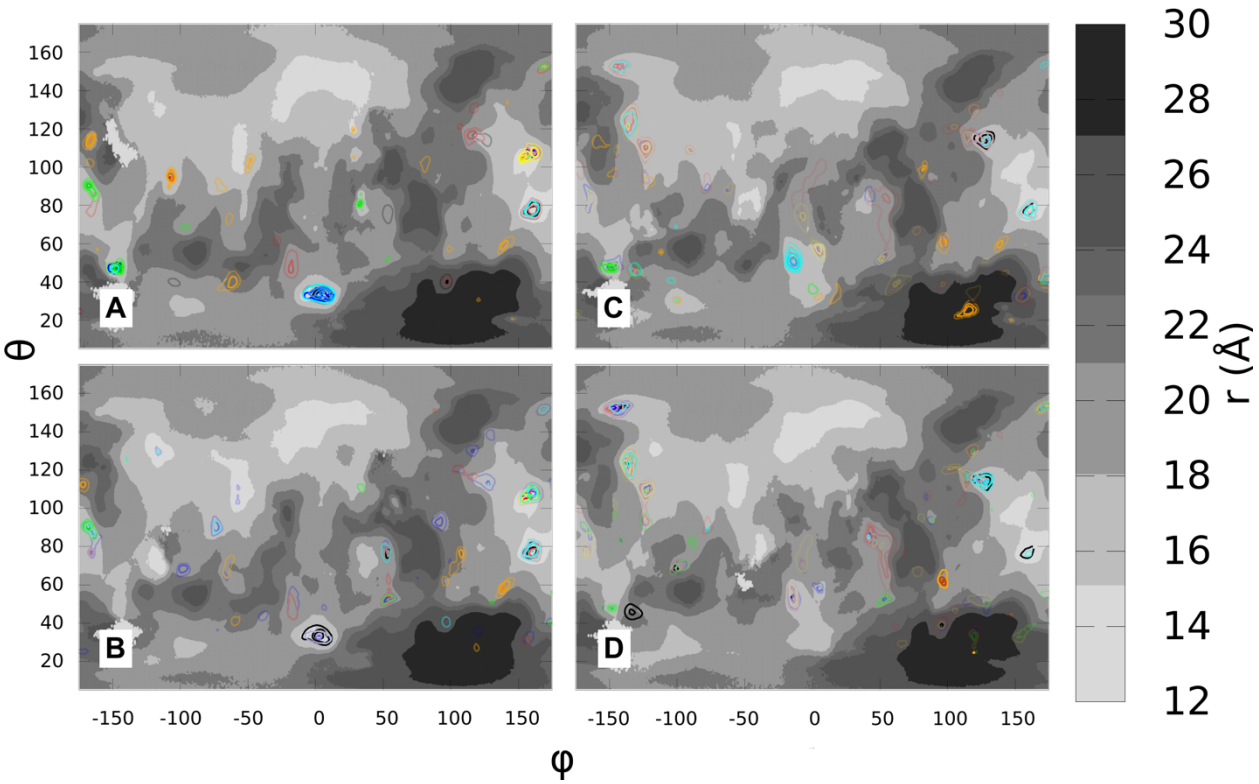

Figure S8: Probe surface densities projected onto a probe-derived protein surface topography map of (A and C) G12V and (B and D) G13D K-Ras in OS1 (left) and OS2 (right). Densities of isobutane, isopropanol, acetamide, acetate, acetone, DMSO and urea are shown in black, cyan, blue, orange, red, yellow and green at a contour level of ≤$-RT \ln \frac{s}{c_0} = -1.8\ kcal/mol$.

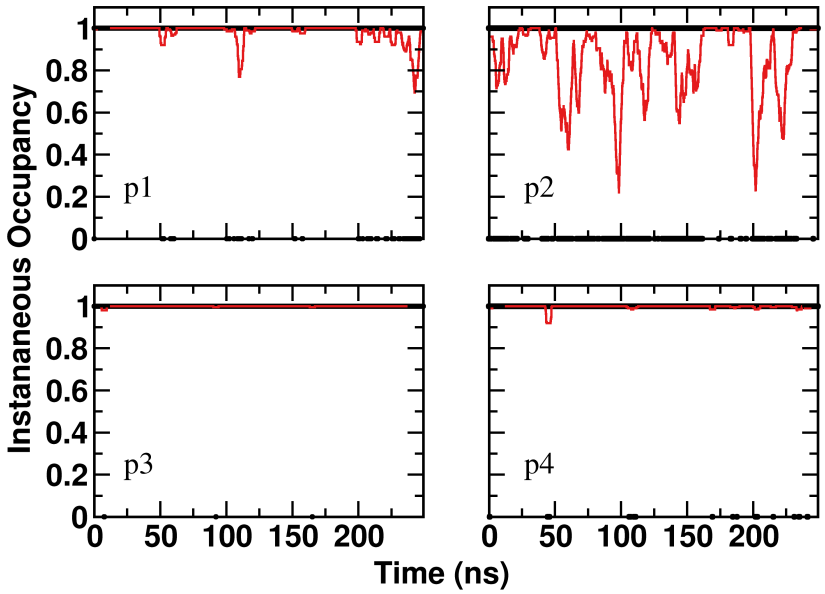

Figure S9: Residue-averaged instantaneous probe occupancy of pockets p1-p4 of G12D in OS2. Black and red symbols represent data sampled every 10ps and 1ns, respectively. The mean occupancies in the first half and second half of the aggregate simulation are 0.987 and 0.973 (p1), 0.834 and 0.866 (p2), 0.999 and 0.999 (p3) and 0.996 and 0.998 (p4). The first 10ns data of each trajectory was excluded as equilibration phase.