SUPPLEMENTAL INFORMATION

Mechanisms of lipid scrambling by the G protein-coupled receptor opsin

Giulia Morra[1,2], Asghar M. Razavi[1], Kalpana Pandey[3], Harel Weinstein[1,4], Anant K. Menon[3], George Khelashvili[1,4]

[1] Department of Physiology and Biophysics, Weill Cornell Medical College, New York, NY, 10065, USA

[2] Istituto di Chimica del Riconoscimento Molecolare, CNR, Via Mario Bianco 9, 20131 Milano Italy

[3] Department of Biochemistry, Weill Cornell Medical College, New York, NY, 10065, USA

[4] Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY 10065, USA

SUPPLEMENTAL FIGURE TITLES AND LEGENDS

Figure S1:

Title of Figure S1: Iterative protocol used for unbiased atomistic MD simulations of opsin. Related to Table 1.

Legend of Figure S1: The simulations were run in four stages, each informed by the output from the previous stage and carried out in multiple replicates. The four stages of the simulations are denoted by shaded rectangles, and the simulations within each stage are demarcated by the vertical dashed lines. For each stage, the trajectory from which a frame was chosen to initiate the next set of ensemble simulations is shown as red rectangle.

Figure S2:

Title of Figure S2: tIC1 describes the opening of the intracellular gate. Related to Figure 6.

Legend of Figure S2: (A) Projections of z-positions of the phosphate atom of the three penetrating lipids (upper panel, Lipid 1, 2, and 3 are represented as in the main text by red, green, and blue traces, respectively) and of the E249-K311 distance (lower panel) onto the tIC 1 vector. (B-C) Structural representation of the protein (in cartoon) corresponding to the minimum (B) and maximum (C) tIC 1 projection values. Also shown are positions of the phosphate atoms of the three lipids (using the same color code as in panel A) in the respective states and top ten pair distances (in orange lines) contributing the most to the projection.

Figure S3:

Title of Figure S3: Implied timescales for the MSM relaxation modes. Related to Figure 7.

Legend of Figure S3: Implied timescales for relaxation modes of opsin obtained from the MSM illustrating Markovian behavior at 40 ns lag-time (vertical dashed line). The first two MSM relaxation modes are characterized by 24.5 and 8.8 μs time-scales, respectively.

Figure S4:

Title of Figure S4: Coarse-graining of the tICA landscape. Related to Figure 8.

Legend of Figure S4: Converting microstate level MSM to macrostate level MSM. All 100 microstates (shown in panel A as yellow circles on the 2D tIC1 vs tIC2 landscape) were lumped into 12 macrostates (color coded and numbered on the 2D tIC1 vs tIC2 landscape in panel B) using the PCCA+ algorithm (see Methods).

Figure S5:

Title of Figure S5: Macrostates with relatively large tIC2 values describe lipid dynamics along the translocation pathway. Related to Figure 8.

Legend of Figure S5: Histograms of the z position of Lipid 1 (red), Lipid 2 (green) and Lipid 3 (blue) phosphate atoms shown for trajectory frames in separate macrostates (1-12).

Figure S6:

Title of Figure S6: Dynamics of the gates at Site 1 and Site 2. Related to Figure 8.

Legend of Figure S6: Distribution of E249-K311 (A), C264-S298 (B), Y306-L76 (C), and Y306-A260 (D) distances in each macrostate. The traces in dotted lines represent the macrostates with opsin conformations preceding lipid insertion, whereas the traces in solid lines correspond to macrostates in which the opsin structure has one or more lipids in the translocation pathway.

Figure S7:

Title of Figure S7: Lipid tail insertion interferes with the translocation pathway. Related to Figure 8.

Legend of Figure S7: Representative snapshots of the system from Macrostate 1 (left) and Macrotate 6 (right) highlighting a mode of lipid-protein interactions where the lipid tail is inserted in the translocation pathway. Opsin protein is represented and color-coded as in Figs. 4B-F and 8 of the main text. The E2496.32, K3117.58, and Y3067.53 residues as well as inserted lipids are drawn in space fill representation.

Figure S8:

Title of Figure S8: Identification of the relevant eigenvectors for the MSM construction. Related to Figure 6.

Legend of Figure S8: Contribution of all 2016 tICA eigenvectors to the total fluctuations. The inset shows a detailed view of the same plot for the first 3 eigenvectors.

Figure S9:

Title of Figure S9. The first three eigenvector are sufficient to describe the dynamic process of lipid translocation. Related to Figure 6.

Legend of Figure S9: Projection of MD simulation data on pairs of tICA eigenvectors. Different columns (from left to right) correspond to Stage 1, 2, 3, and 4 trajectories. Dark red areas corresponding to higher populations and light blue areas correspond to lower populations. As the MD simulation progresses to different stages, the only populations moving along tICA eigenvectors are the ones that correspond to tICs 1, 2, and 3. Because the different stages of MD simulations are designed to capture different steps of lipid traversal along opsin, only the first 3 tICA reaction coordinates contain relevant information and therefore only these reaction coordinates were used for subsequent analysis and MSM construction.

Figure S1

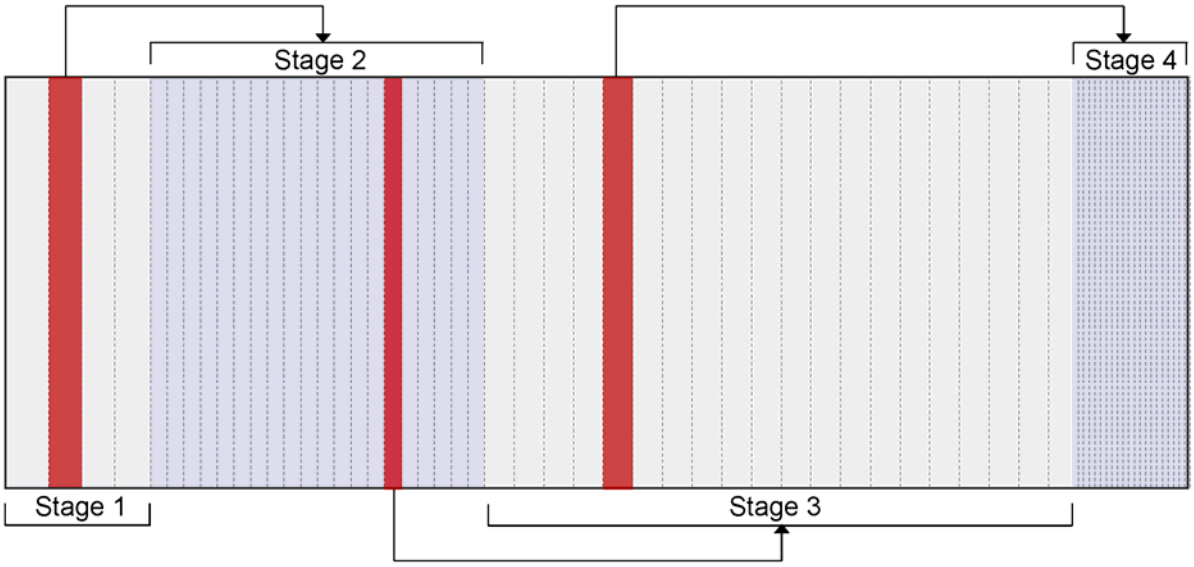


Figure S2

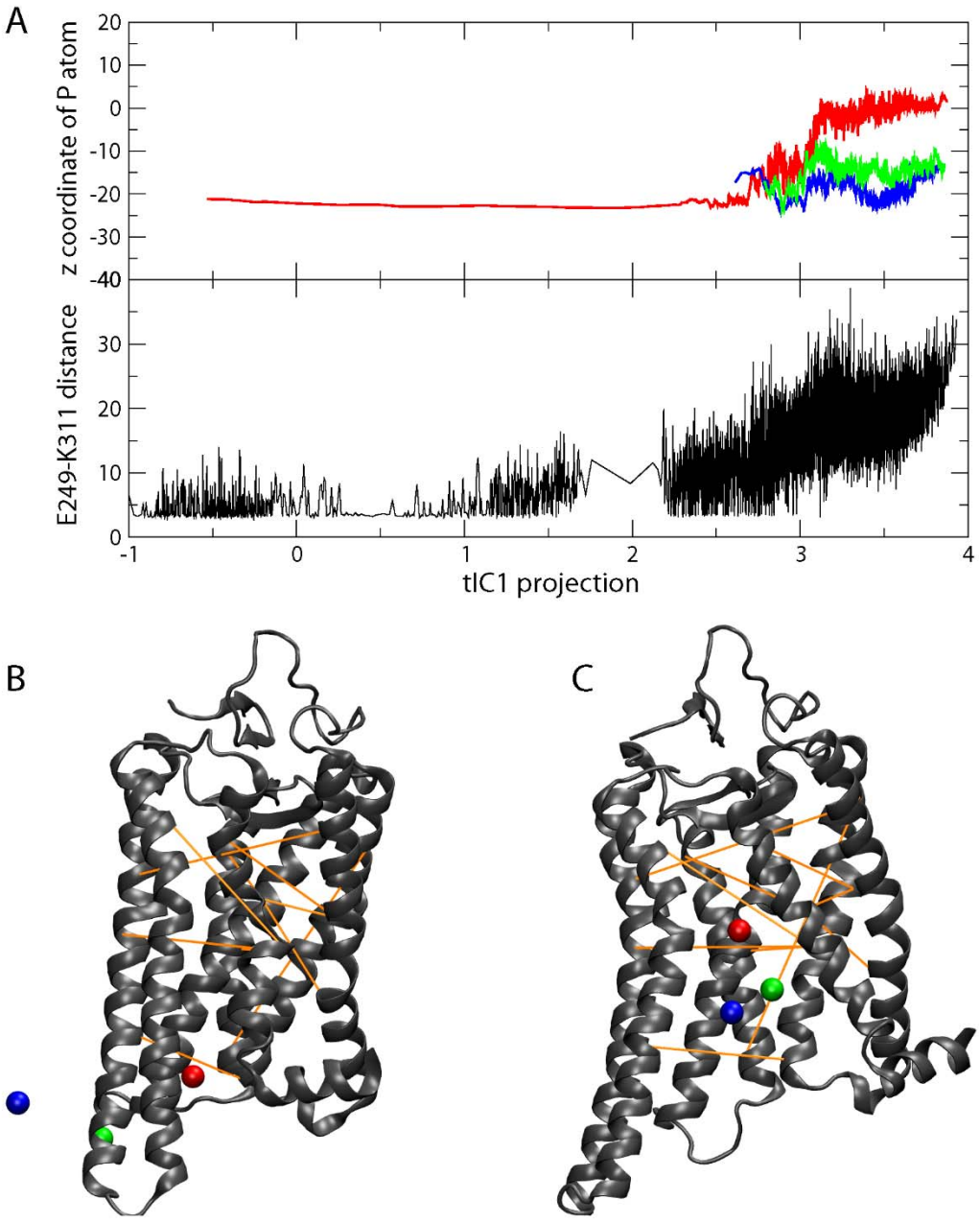

A
z coordinate of P atom
E249-K311 distance
tIC1 projection
B
C

Figure S3

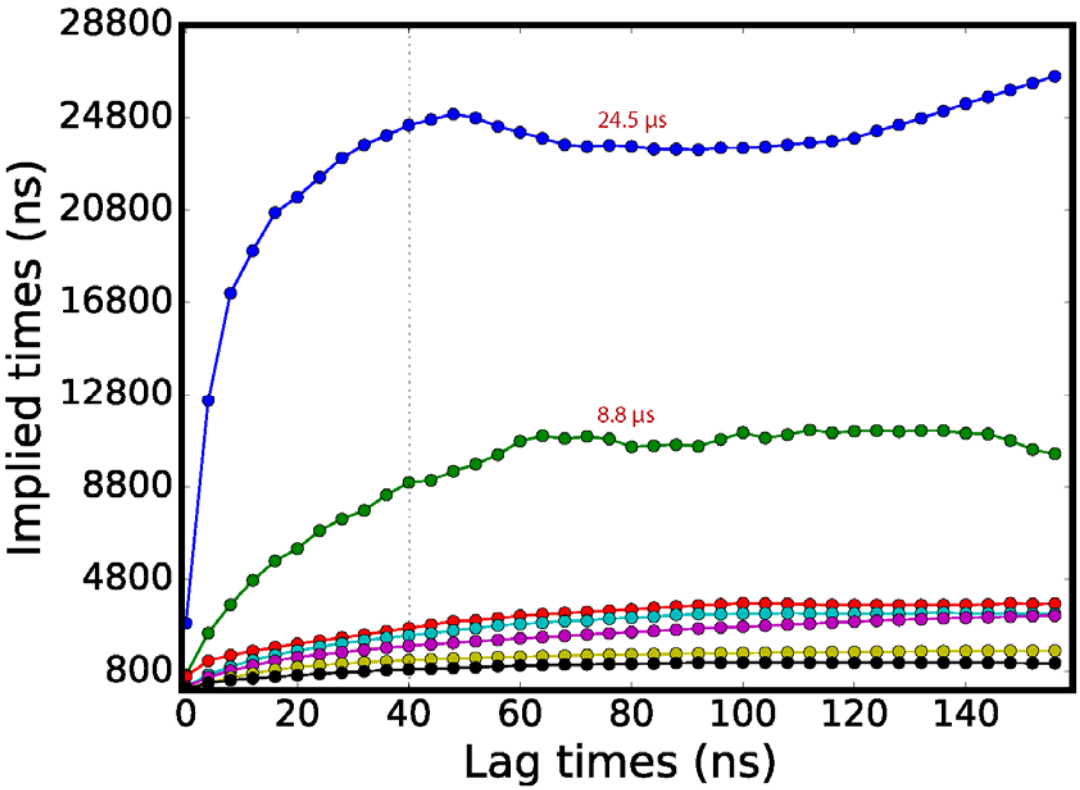

28800
24800
20800
16800
12800
8800
4800
800
Implied times (ns)
24.5 µs
8.8 µs
0
20
40
60
80
100
120
140
Lag times (ns)

Figure S4

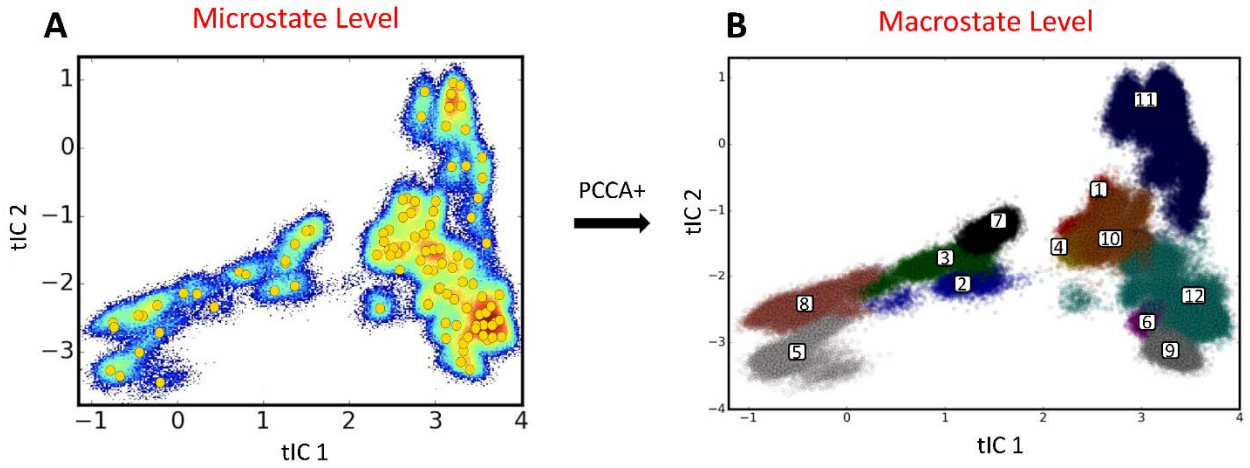

A
Microstate Level
tIC 2
tIC 1
PCCA+
B
Macrostate Level
tIC 2
tIC 1
11
1
7
4
10
3
2
8
12
6
5
9

Figure S5

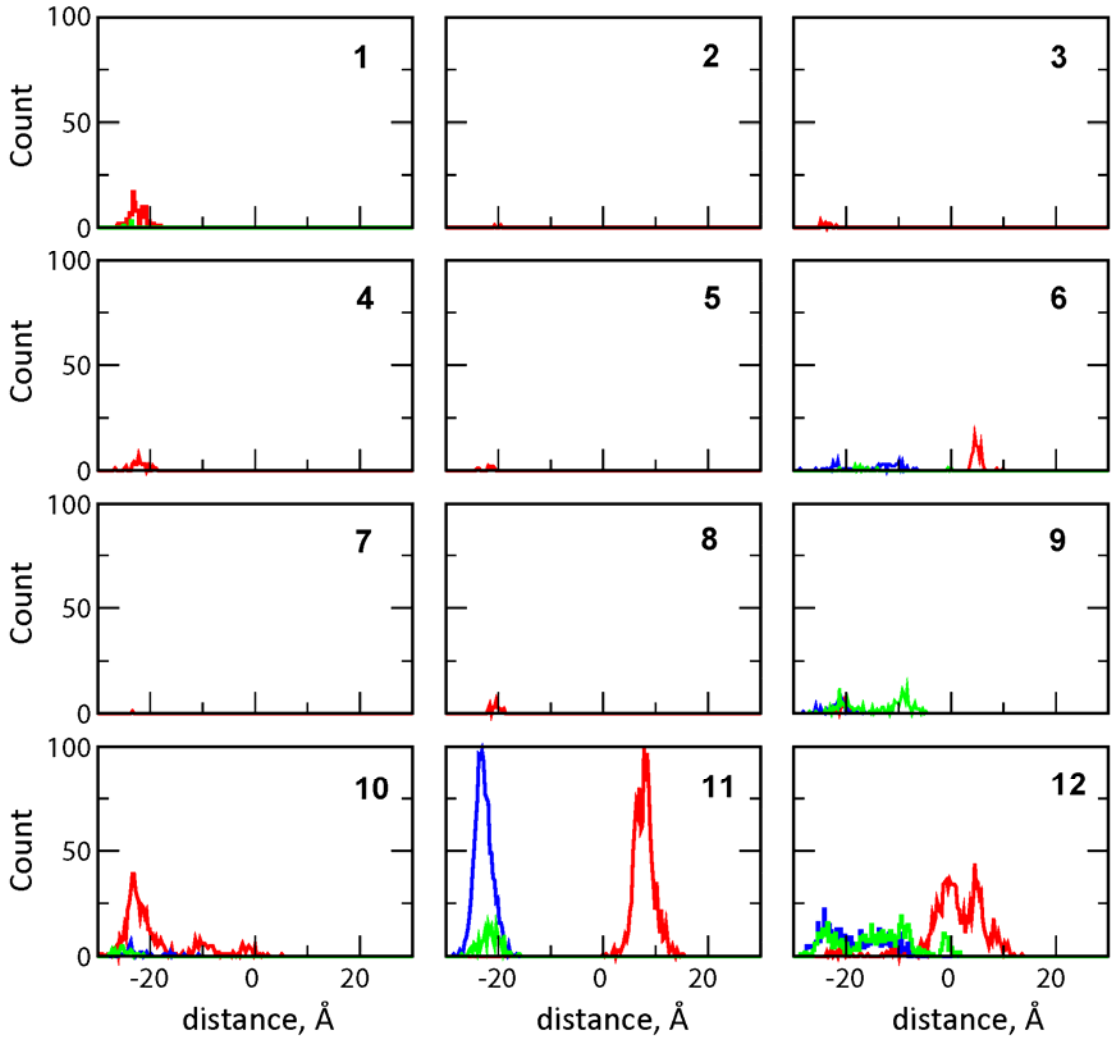


Figure S6

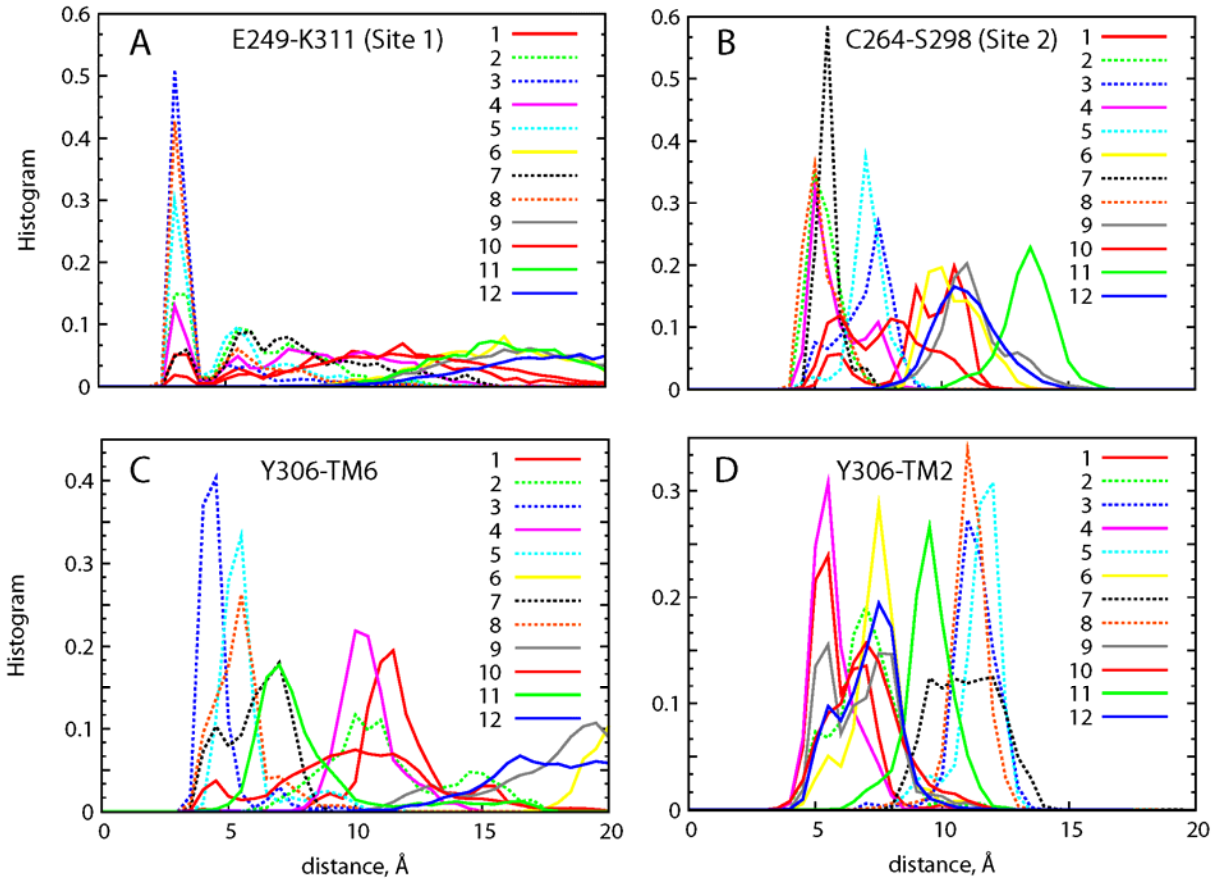

A
E249-K311 (Site 1)
B
C264-S298 (Site 2)
C
Y306-TM6
D
Y306-TM2
Histogram
distance, Å
1
2
3
4
5
6
7
8
9
10
11
12

Figure S7

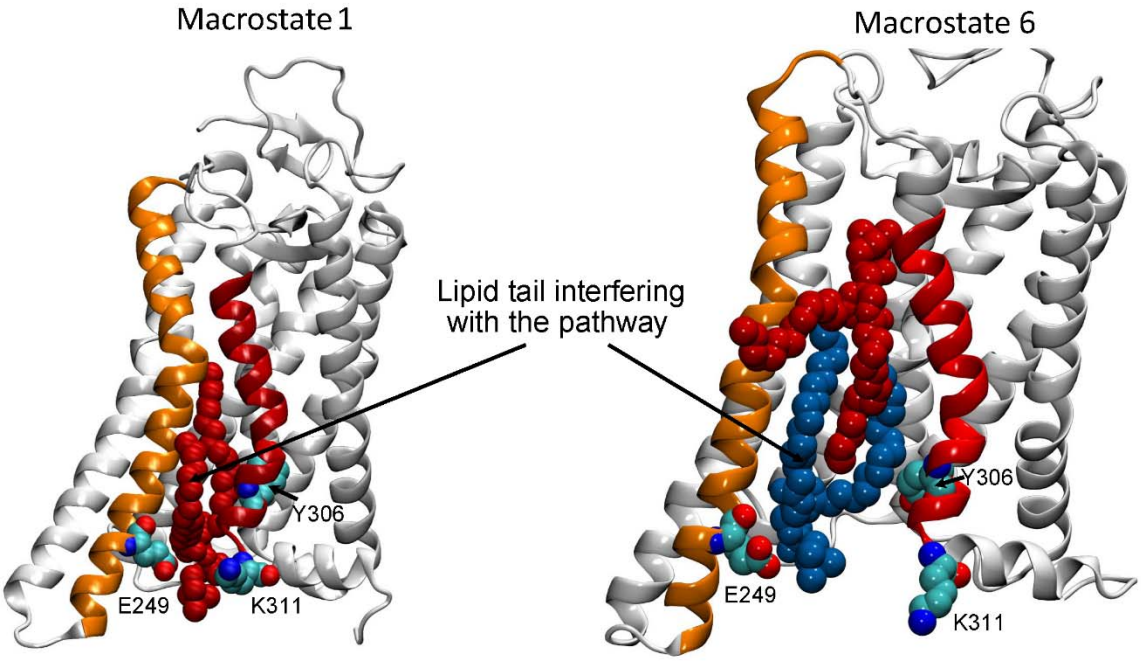

Macrostate 1
Macrostate 6
Lipid tail interfering
with the pathway
Y306
E249
K311
Y306
E249
K311

Figure S8

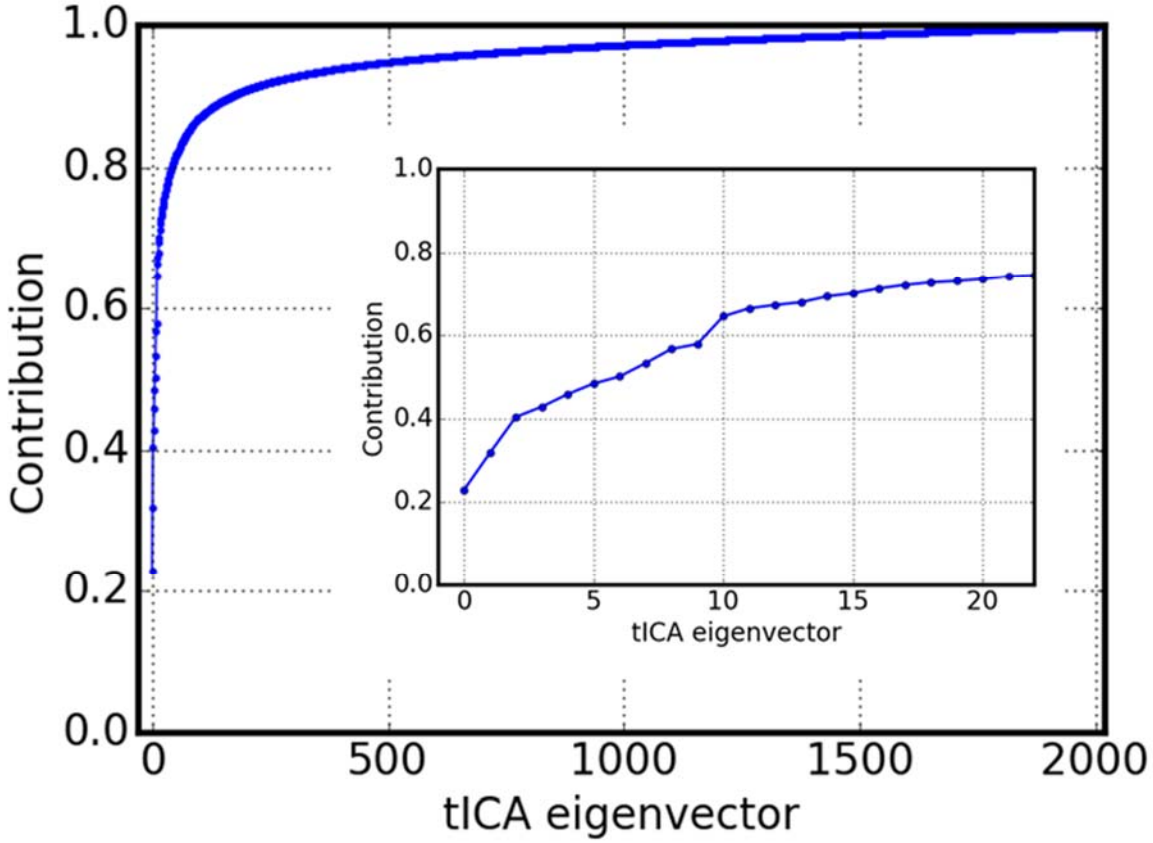
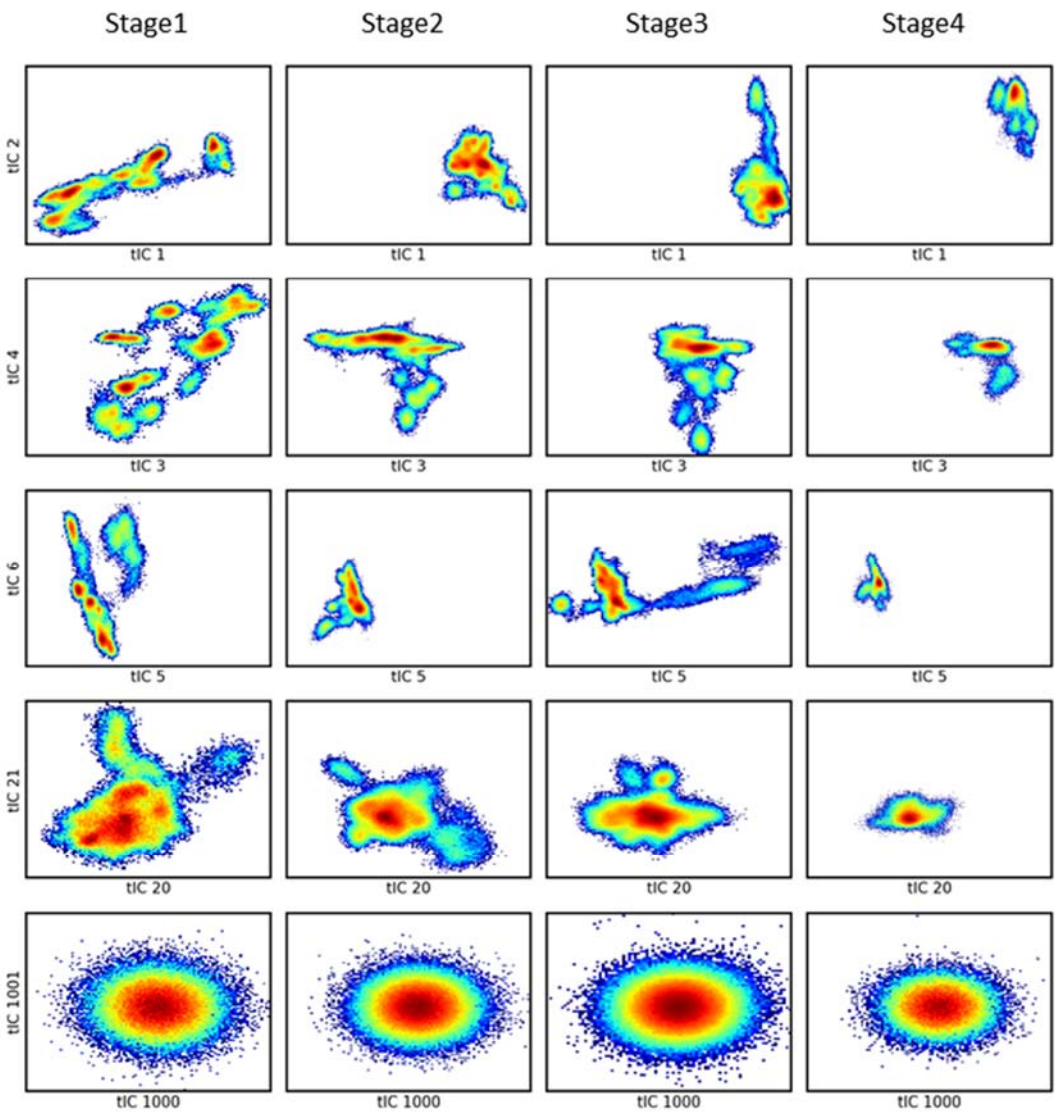

1.0
0.8
0.6
0.4
0.2
0.0
Contribution
0
500
1000
1500
2000
tICA eigenvector
1.0
0.8
0.6
0.4
0.2
0.0
Contribution
0
5
10
15
20
tICA eigenvector

Figure S9


Stage1
Stage2
Stage3
Stage4
tIC 2
tIC 4
tIC 6
tIC 21
tIC 1001
tIC 1
tIC 3
tIC 5
tIC 20
tIC 1000