Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2019

Supporting Information

Opening dynamics of HIV-1 gp120 upon receptor binding is dictated by a key hydrophobic core

Lin-Tai Da[a,*], Mengna Lin[a]

[a]Key Laboratory of System Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China.

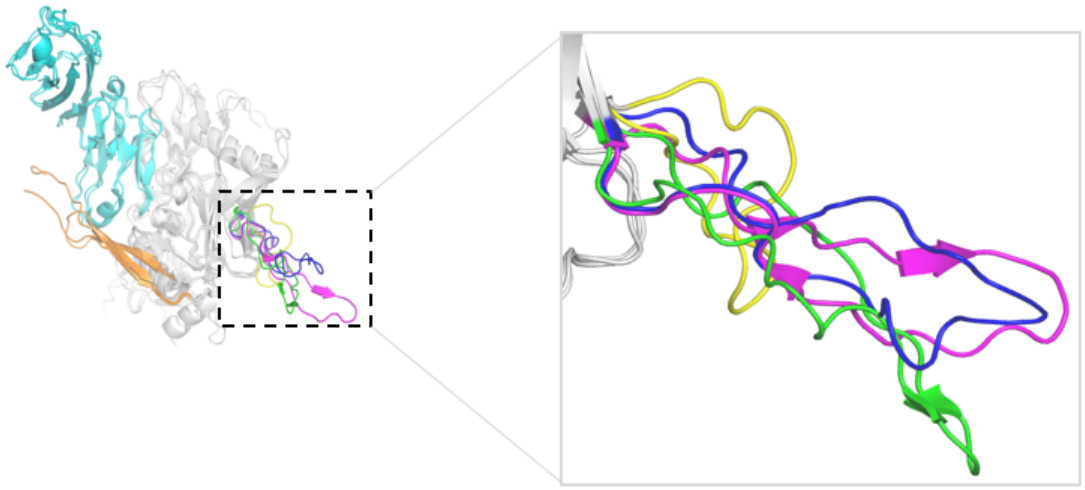

Fig. S1. Crystal structures of the V3 loop in different open conformations captured in various contexts, with the PDB id of 6meo (blue), 3tyg (yellow), 2qad (magenta), and 2b4c (green), respectively.

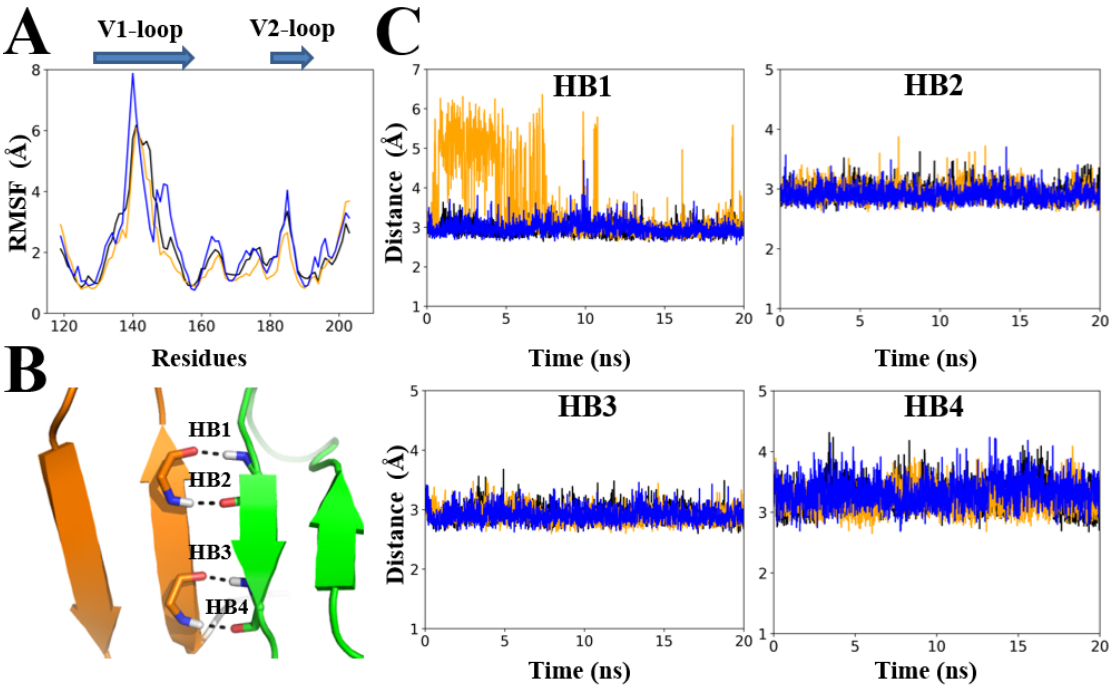

Fig. S2. (A) Root mean square fluctuation (RMSF) for the V1/V2 residues (from Cys119 to Cys205) in the open gp120-CD4 complex for three parallel 20-ns MD simulations, illustrated by different colors. (B) Four hydrogen bonds (HBs) within the bridging sheet domain (HB1-HB4). (C) The donor-acceptor distances for the above defined HBs are plotted as the function of simulation time.

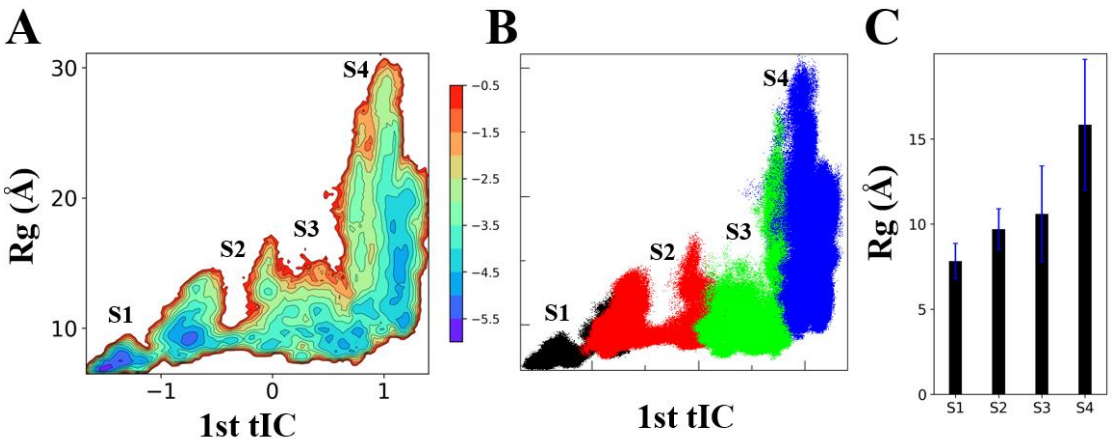

Fig. S3. (A) Free energy projection of MD conformations onto two reaction coordinates: the 1st tIC and radius of gyration (Rg) for the hydrophobic core. (B) Scatter plot with each MD conformation projected onto the 1st tIC and Rg. The conformations that belong to the same macrostate are shown in same color, namely S1 (black), S2 (red), S3 (green), and S4 (blue). (C) Rg of the hydrophobic core calculated for each macrostate. The mean values were averaged over all the microstates that belong to certain macrostate, and the corresponding standard errors were calculated.

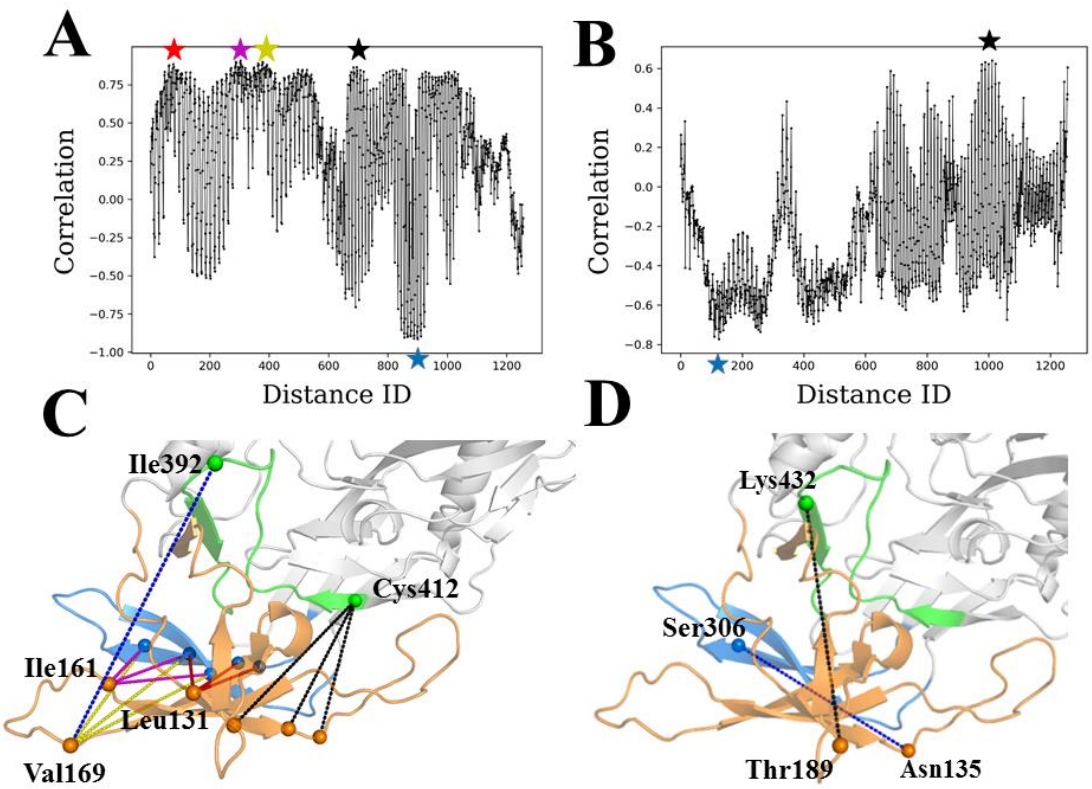

Fig. S4. Correlation analysis for the tIC1 (A) and tIC2 (B) with each of the 1257 distance pairs used for MSM construction. The distance pairs that correlate with the tIC1 the most in a positive manner are indicated by purple, red, yellow, and black stars, with the negative correlation indicated by blue star. The distance pairs that correlate with the tIC2 the most in a positive and negative manner are indicated by black and blue stars, respectively. The corresponding structural illustrations are provided in C and D, in which the identified atom pairs are indicated by colored dashed lines.

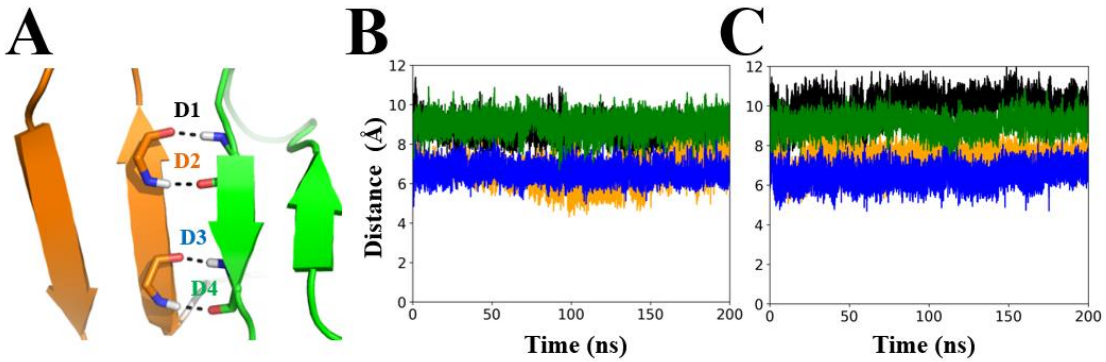

Fig. S5. (A) Four distances within the bridging sheet (D1-D4) were measured for the gp120-CD4-17b (B) and gp120-CD4-21c (C) complexes during the 200-ns MD simulations.

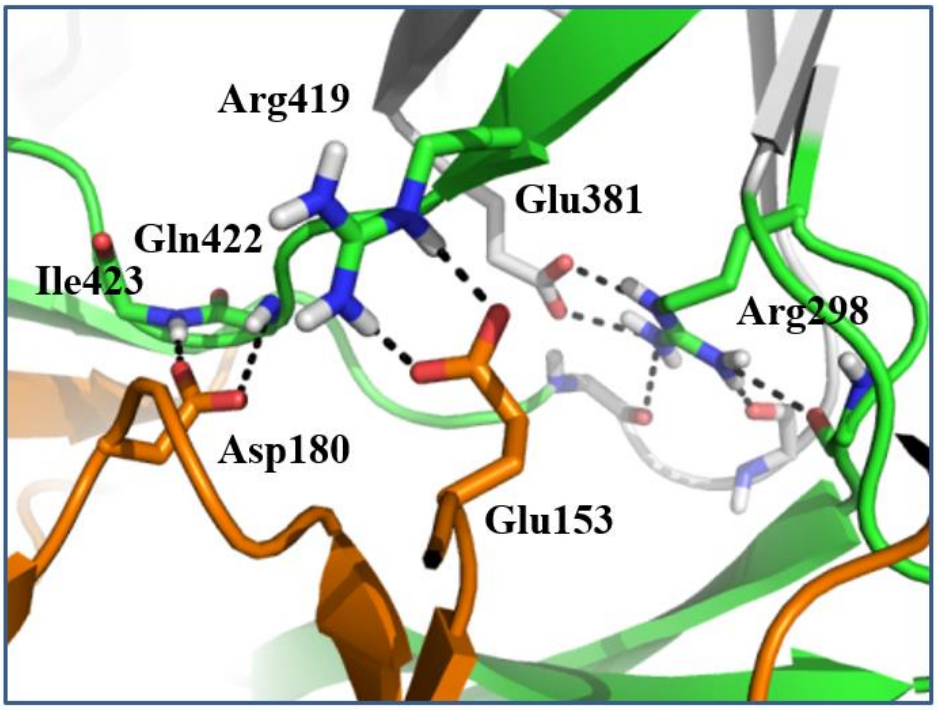

Fig. S6. Several charged residues are buried inside gp120 in the S1 state, including the V1/V2 residues Asp180 and Glu153, and the V3-stem residue Arg298. The surrounding residues that interact with the above residues are highlighted in stick models and the HBs are indicated with black dashed lines.

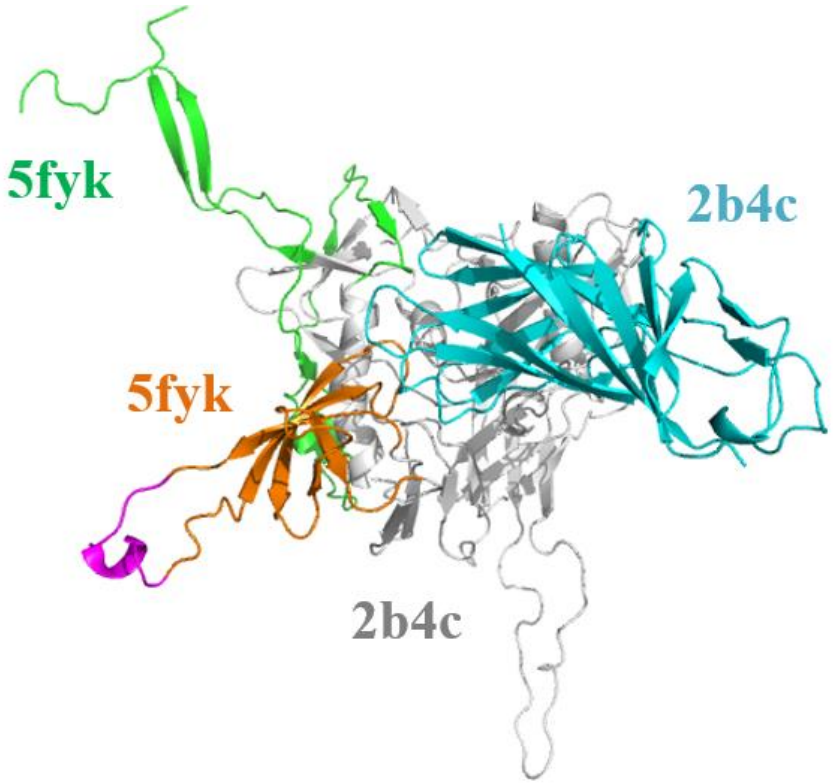

Fig. S7. Construction of the open gp120-CD4 complex. Two crystal structures, 5fyk and 2b4c, were used for the model construction, and the V1 loop was built using the SWISSPDB software. Refer to the "**Materials and methods"** section in the main text for the details.

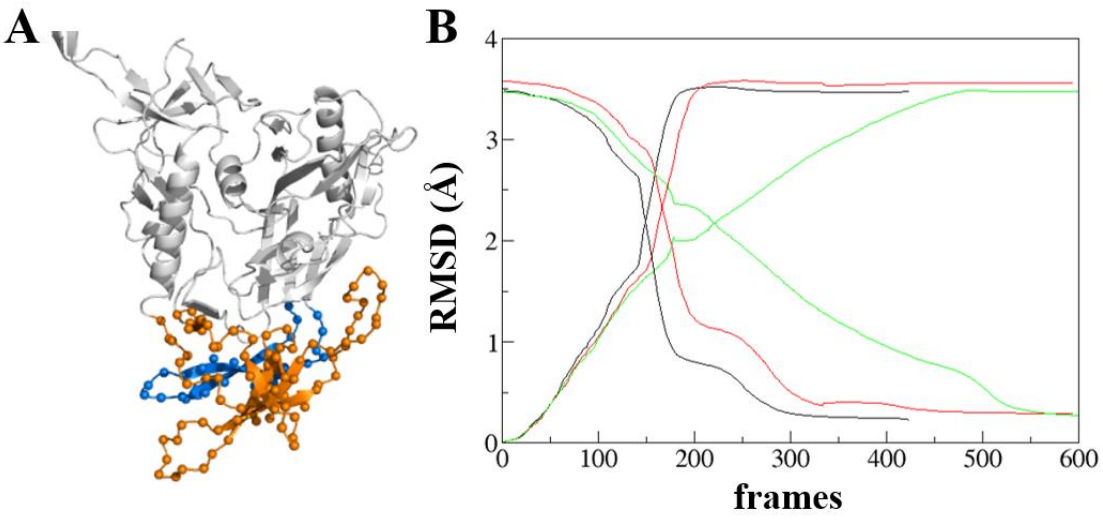

Fig. S8. (A) The atoms that are used for the Climber simulations, including the C_α atom of the residues Cys119-Cys205 from the V1/V2 variable loops (orange) and the residues Arg298-Gln328 from the V3 loop (blue). (B) Three parallel Climber simulations were performed to obtain the initial gp120 opening pathways (shown in different colors), each monitored by the RSMD changes of the targeting regions against the simulation time with respect to the closed and open gp120 structure, respectively.

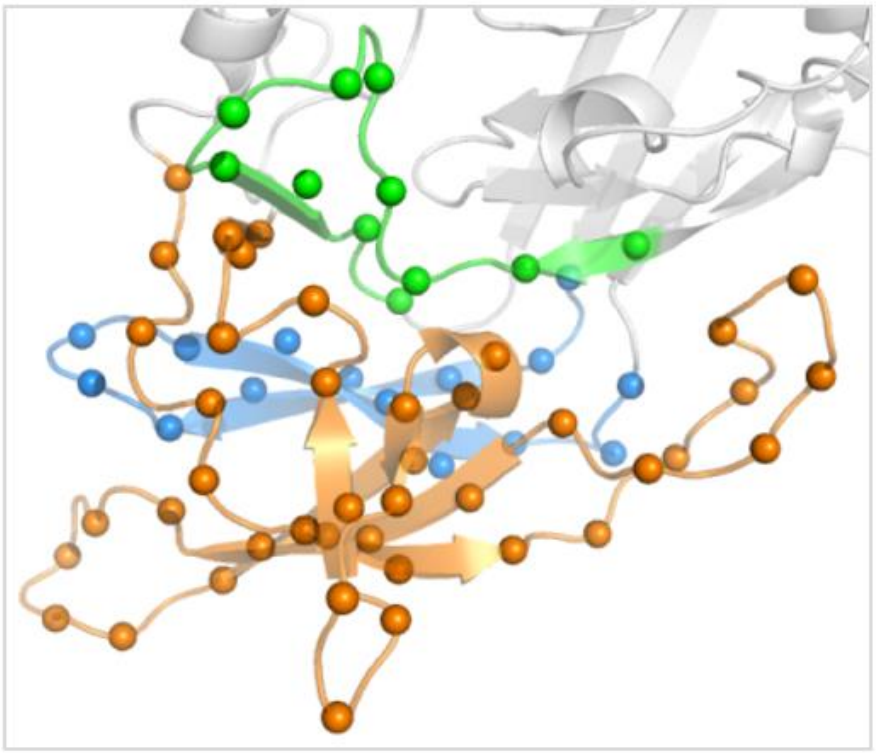

Fig. S9. Atoms used for the tICA decomposition, including the Cα atoms of the V1/V2 residues Cys119 – Cys205 (at the interval of 2 residues, shown in orange), Cα atoms of the V3 residues Arg298-Gln328 (at the interval of 2 residues, shown in blue), and Cα atoms of the β20/21 residues Cys418 – Pro438 (at the interval of 2 residues, shown in green).

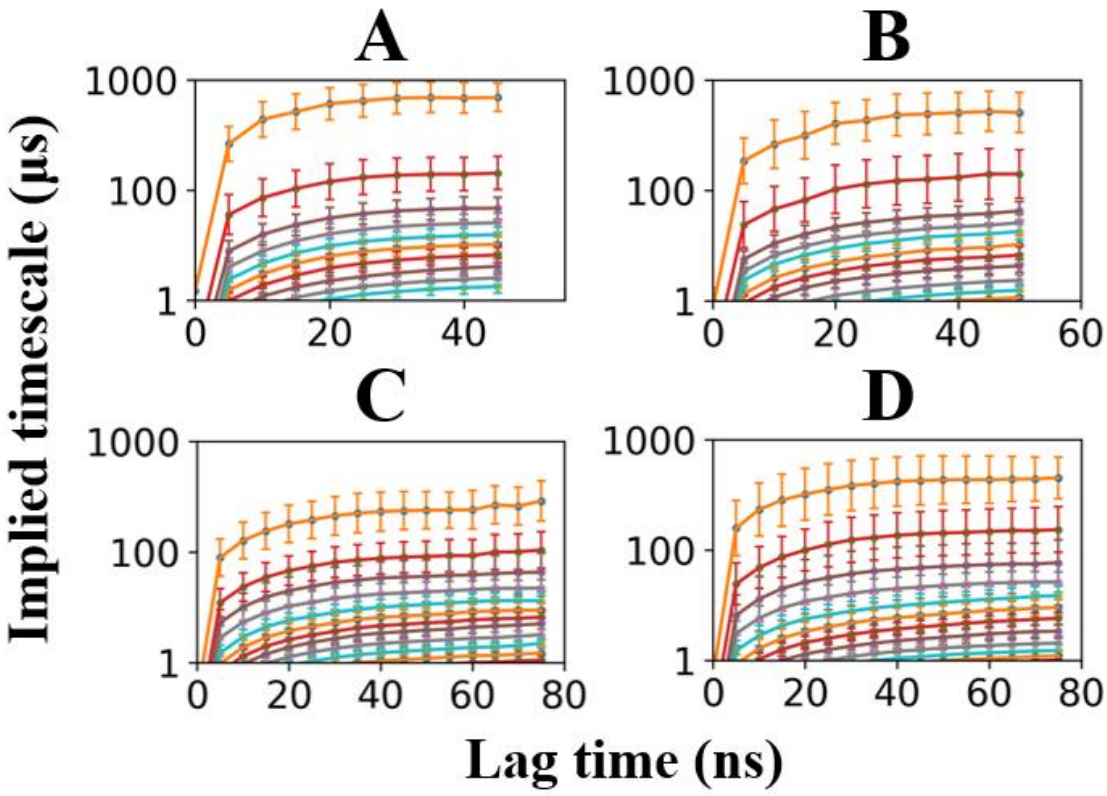

Fig. S10. Implied timescale plots for different datasets with various data sizes: 14 μs (A), 16 μs (B), 18 μs (C), 20 μs (D), respectively. We used the microstate number of 500 to construct each MSM.

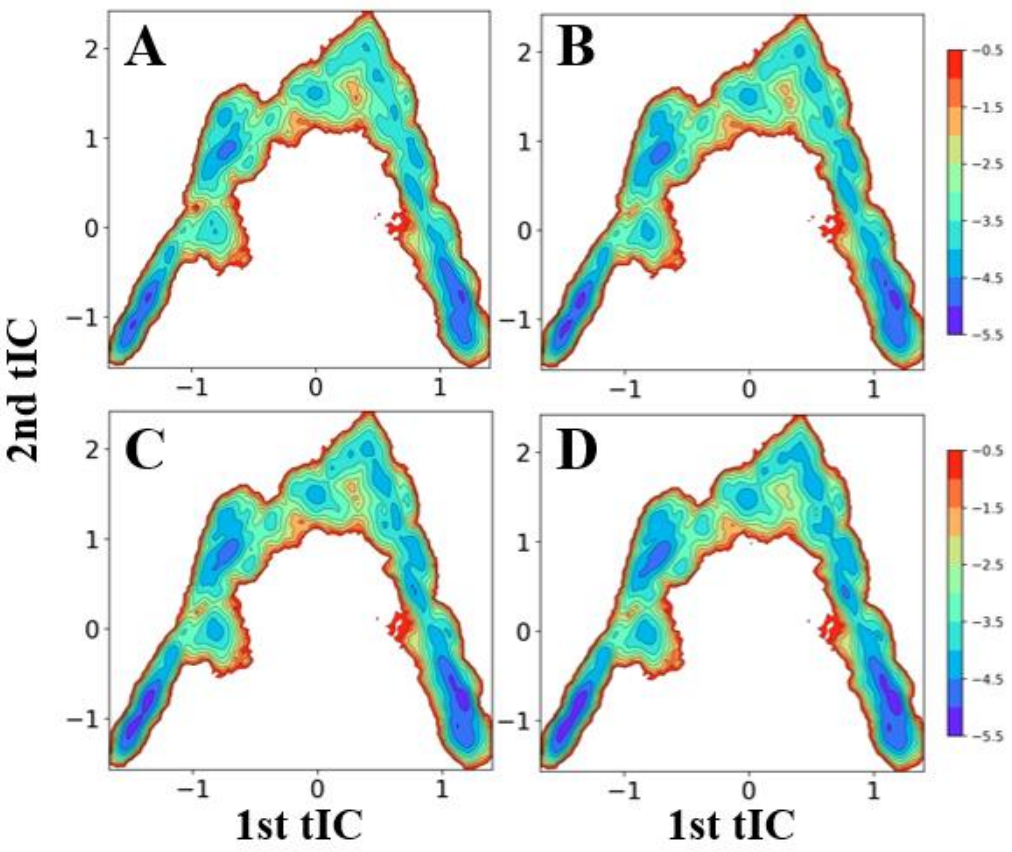

Fig. S11. Projections of different subsets of the complete simulations dataset onto the same top two tICs (unit in *kcal/mol*), with the aggregated simulation time of 14 μs (A), 16 μs (B), 18 μs (C), 20 μs (D).