Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

Distinct Binding Interactions Trigger Opposite Conformational Modulations on Pathogenic and Wildtype Huntingtin Exon 1 Proteins

Jiaming Guan[1,2], Zhijian Song[1,2,3], Guanghong Wei[4] and Qin Qiao[2,3*]

[1] Academy for Engineering and Technology, Fudan University, Shanghai 200433, China

[2] Shanghai Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention, Fudan University, Shanghai 200032, China

[3] Digital Medical Research Center of School of Basic Medical Sciences, Fudan University, Shanghai 200032, China

[4] Department of Physics, State Key Laboratory of Surface Physics, Key Laboratory for Computational Physical Science (Ministry of Education), Fudan University, Shanghai 200438, China

Email: qinqiao@fudan.edu.cn

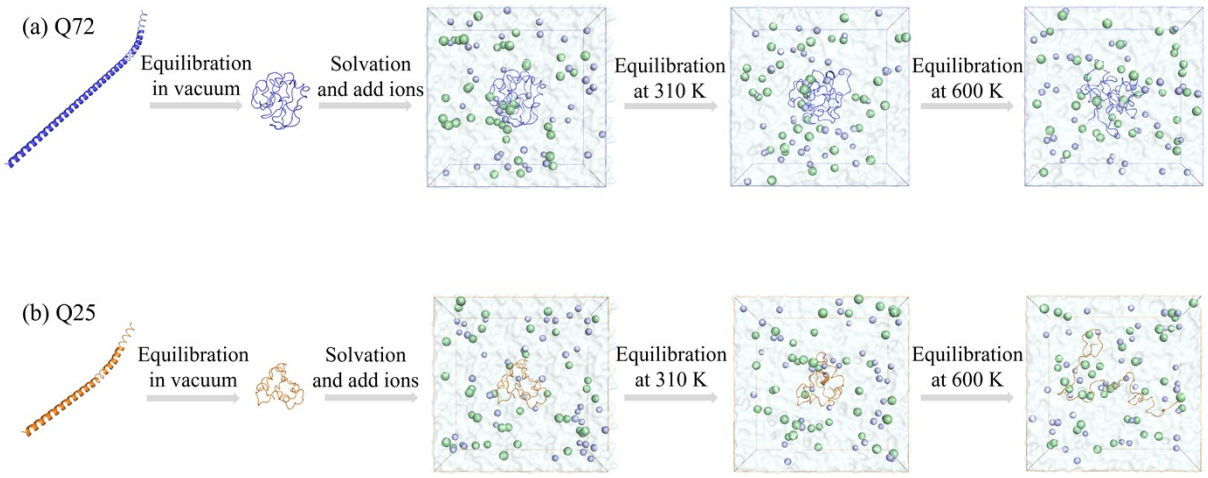

Figure S1. Initial structure preparation and 3-step equilibration process. From left to right are the very initial α-helical structure built via Pymol, the conformation after the 1st step equilibration in vacuum at 700 K, the solvated conformation, the conformation after the 2nd equilibration at 310 K, and the conformation after the 3rd equilibration at 600 K. (a) Q72 system. (b) Q25 system. In each plot, the protein is shown in cartoon, ions are in sphere, and solvent is in surface representation.

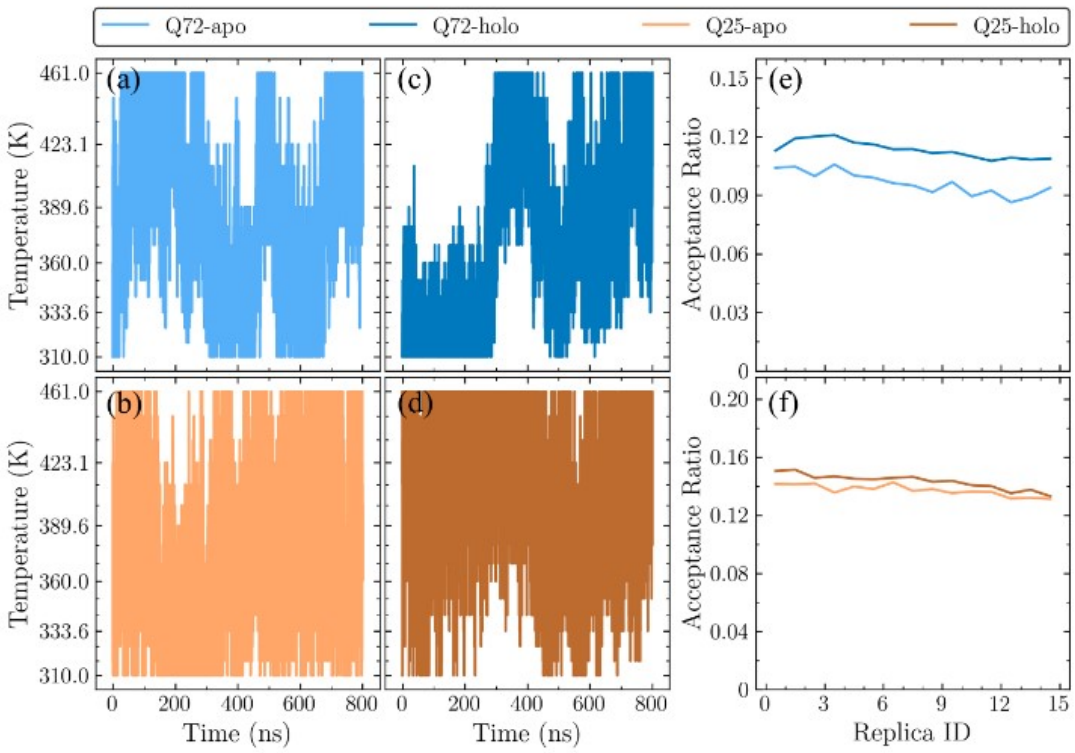

Figure S2. Simulation convergence assessments of Htt-exon-1 systems. In both plots, the light lines correspond to the apo-systems and dark lines to the holo-systems. (a-d) Time evolution of the temperature at 0^{th} replica. (a) Q72-apo. (b) Q72-holo. (c) Q25-apo. (d) Q25-holo. (e-f) Average acceptance ratio between adjacent replicas of Htt-exon-1 systems. (e) Q72 systems. (f) Q25 systems.

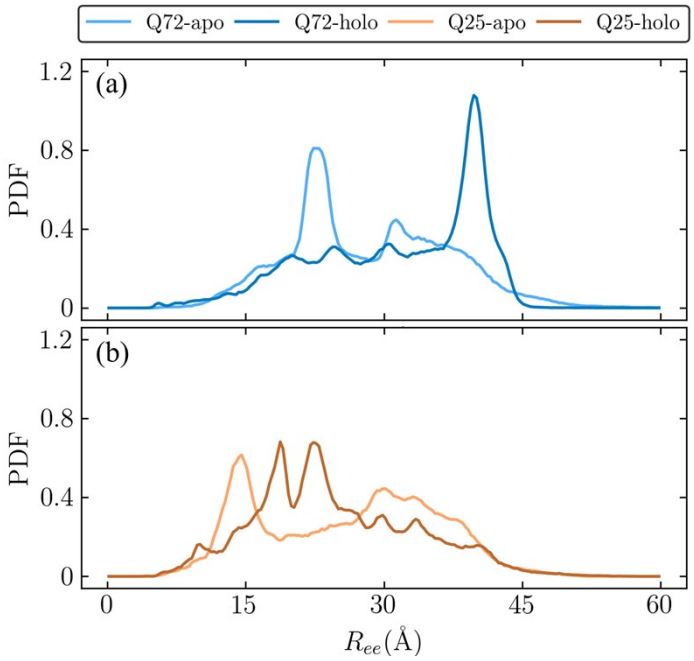

Figure S3. Probability density function (PDF) of the end-to-end distance (R_{ee}) in Htt-exon-1 systems. (a) Q72 systems. (b) Q25 systems. In both plots, the light line corresponds to the apo-system and dark line to the holo-system.

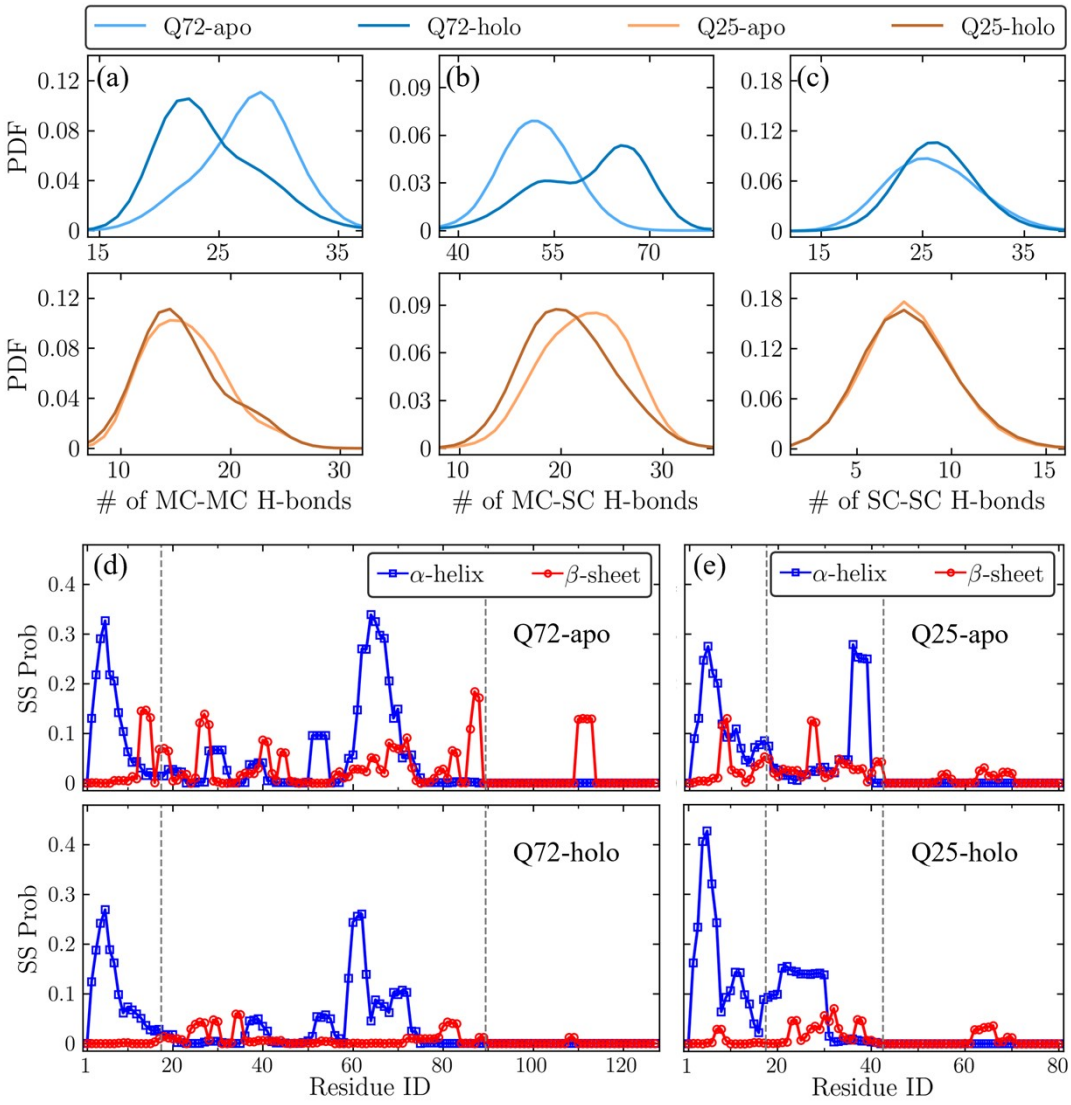

Figure S4. Number of intra-protein hydrogen bonds (H-bonds) and residue-based secondary structures of Htt-exon-1 systems. (a-c) Probability density function (PDF) of H-bond number. The blue lines correspond to Q72 systems and orange lines to Q25 systems. The dark lines are holo-systems and light lines are apo-systems. (a) Number of mainchain-mainchain (MC-MC) H-bonds. (b) Number of H-bonds between mainchain and sidechain (MC-SC). (c) Number of sidechain-sidechain (SC-SC) H-bonds. (d-e) Residue-based secondary structure distributions. In each plot, the upper plots are apo systems and lower plots are holo systems. The blue lines with squares correspond to α-helix structures and the red lines with circles to β-sheet structures. (d) pathogenic Q72. (e) wildtype Q25.

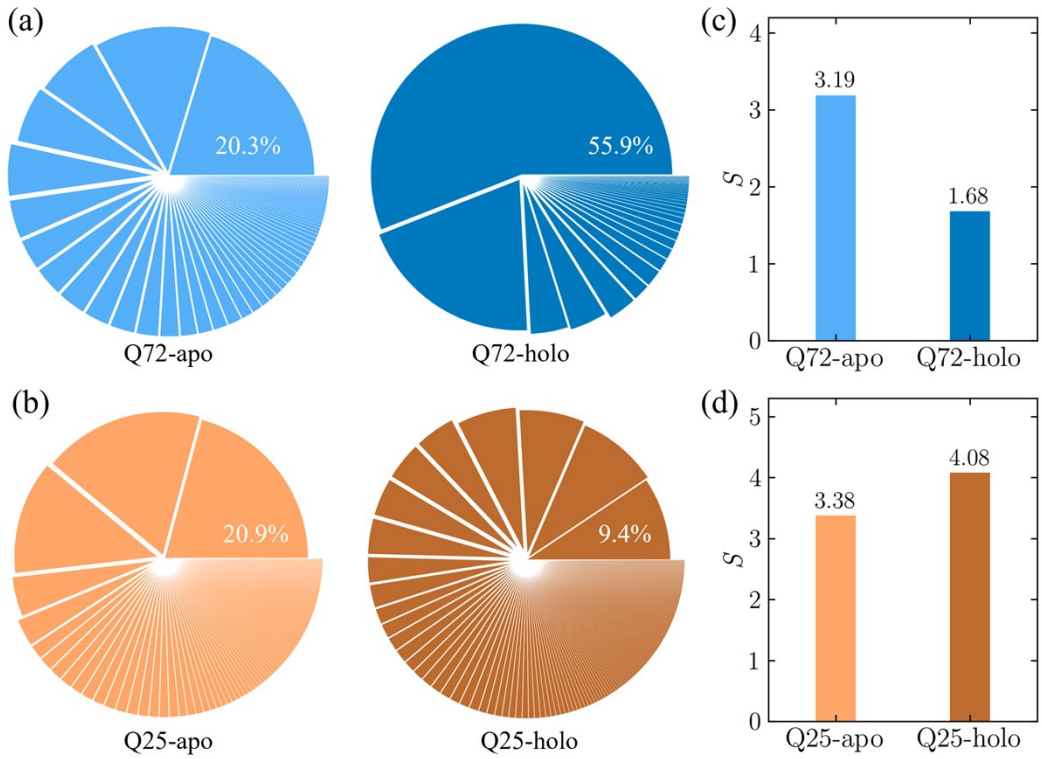

Figure S5. Population distribution of Htt-exon-1 conformational clusters and the corresponding entropy *S* of structural ensemble. (a-b) Pie plots of the cluster population distribution. In (a-b), the left light-colored pies are apo systems, and the right dark-colored pies are holo systems. (a) Q72 systems. (b) Q25 systems. (c-d) Structural ensemble entropy *S* of Htt-exon-1 systems. In (c-d), the light-colored bars are apo systems and dark-colored bars are holo systems. (c) Q72 systems. (d) Q25 systems.

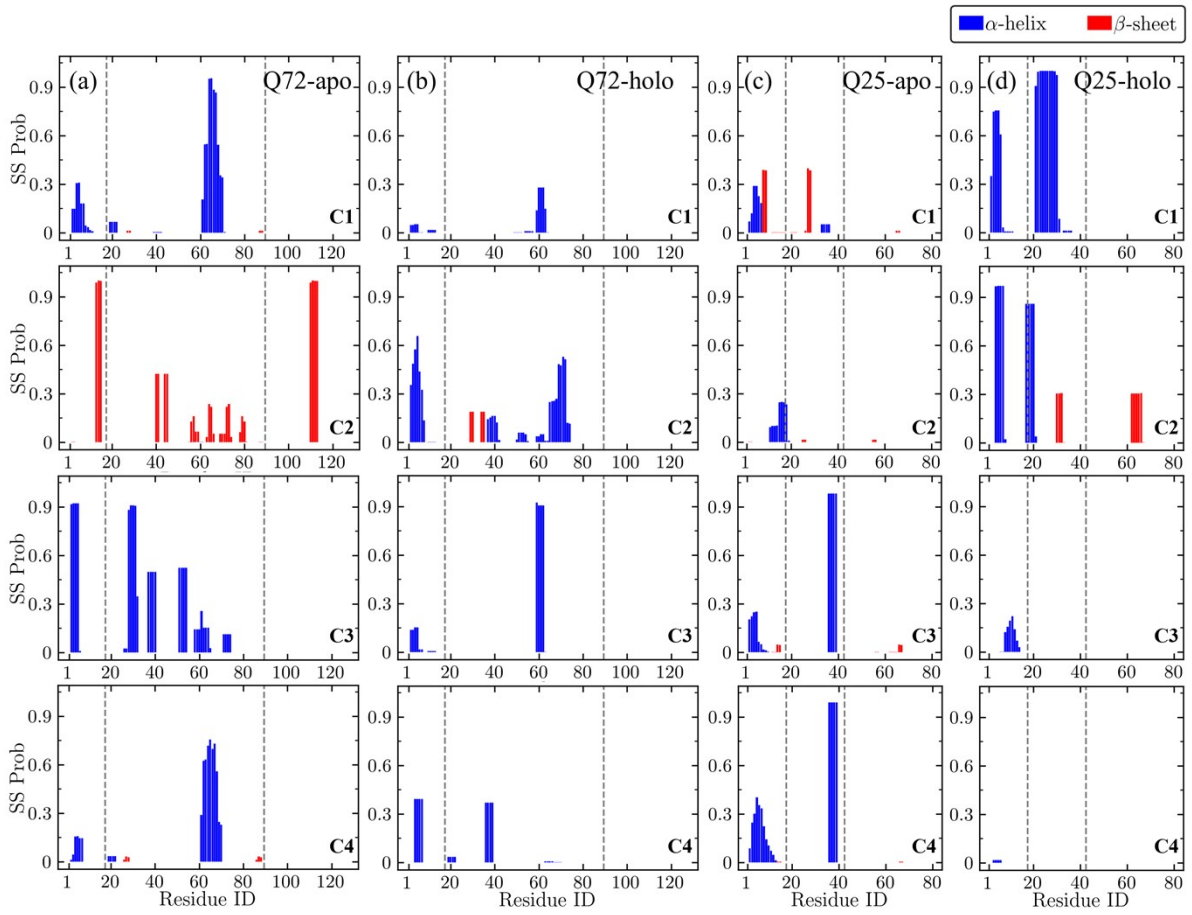

Figure S6. Residue-based secondary structures in the top 4 clusters of each Htt-exon-1 system. In each plot, the subplots from top to bottom correspond to the most populated cluster (C1) to the 4th populated cluster (C4) in each Htt-exon-1 system. The blue bars indicate α-helix structures and the red bars indicate β-sheet structures. (a) Q72-apo. (b) Q72-holo. (c) Q25-apo. (d) Q25-holo.

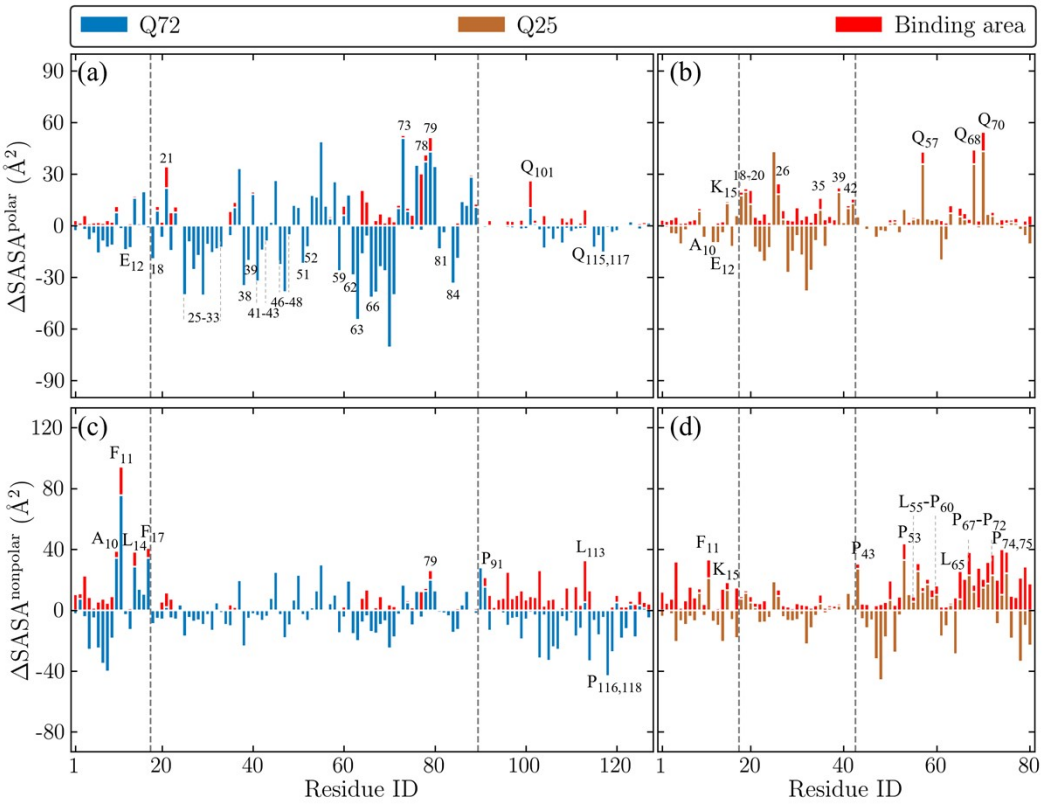

Figure S7. Residue-based SASA difference, $\Delta SASA = SASA_{holo} - SASA_{apo}$, between the holo and apo systems, and ispinesib binding area. The red bars correspond to the ispinesib binding area. The upper labels indicate the residues which became more solvent-exposed with directly binding with ispinesib, while the lower labels indicate the residues which became less solvent-exposed without directly binding with ispinesib. (a-b) Polar ΔSASA. (c-d) Nonpolar ΔSASA.

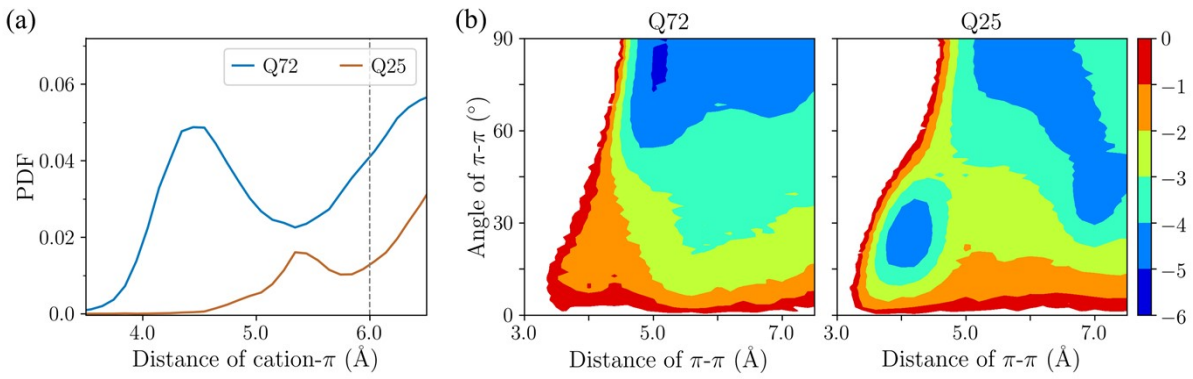

Figure S8. The cation-π interactions and π-π stacking between ispinesib and Htt-exon-1 systems. (a) Distribution of COM distances between aromatic rings in ispinesib and the positively-charged groups NH_3^+ in sidechains of LYS residues. The blue and orange lines correspond to the Q72 and Q25 systems, respectively. (b) Free energy projection on the center of mass (COM) distances and angles between aromatic rings in ispinesib and aryl rings in sidechains of PHE residues.

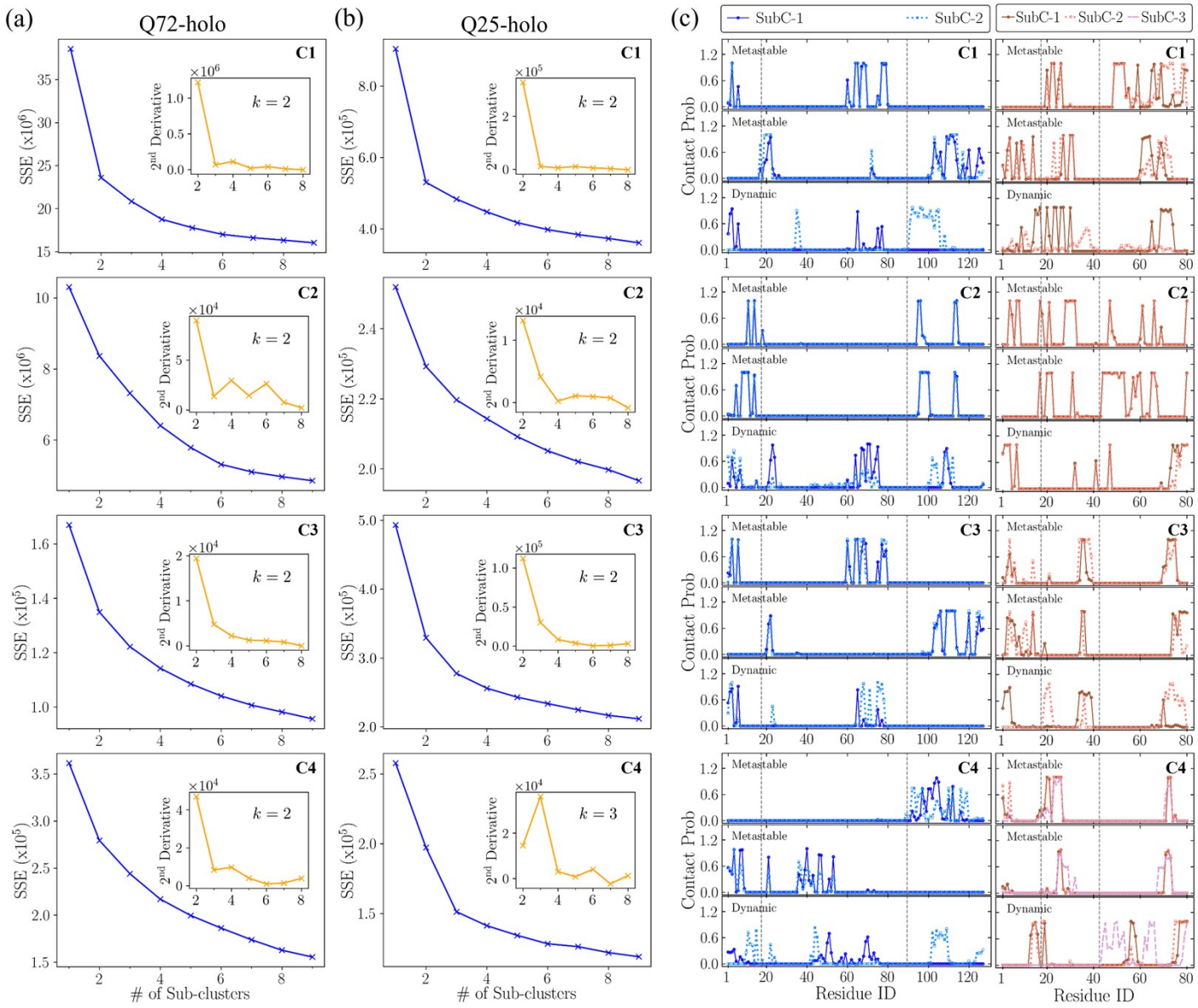

Figure S9. The *k*-means sub-clustering on ispinesib contact patterns in the top 4 conformational clusters. (a-b) Sum of square error (SSE) and its 2nd derivative. In each plot, the subplots from top to bottom correspond to the most populated cluster (C1) to the 4th populated cluster (C4) in each Htt-exon-1 system. In each subplot, the blue line is the SSE changing with the number of sub-clusters *k*, and the inset orange line is its 2nd derivative. (a) Q72-holo. (b) Q25-holo. (c) Comparison of residue-based contact patterns among sub-clusters. In Q72 (subplots on the left), the solid dark blue lines are the contact patterns in the 1st sub-cluster (SubC-1) and the dotted light blue lines are those in the 2nd sub-cluster (SubC-2). In Q25 (subplots on the right), the solid brown, dotted orange and dashed pink lines are the contact patterns of SubC-1, SubC-2, and SubC-3, respectively. In each subplot, one row is for one ispinesib molecule and the up-left label is the corresponding binding mode. The metastable binding corresponds to the consistent binding pattern among the sub-clusters, while the dynamic binding corresponds to the changing binding patterns.

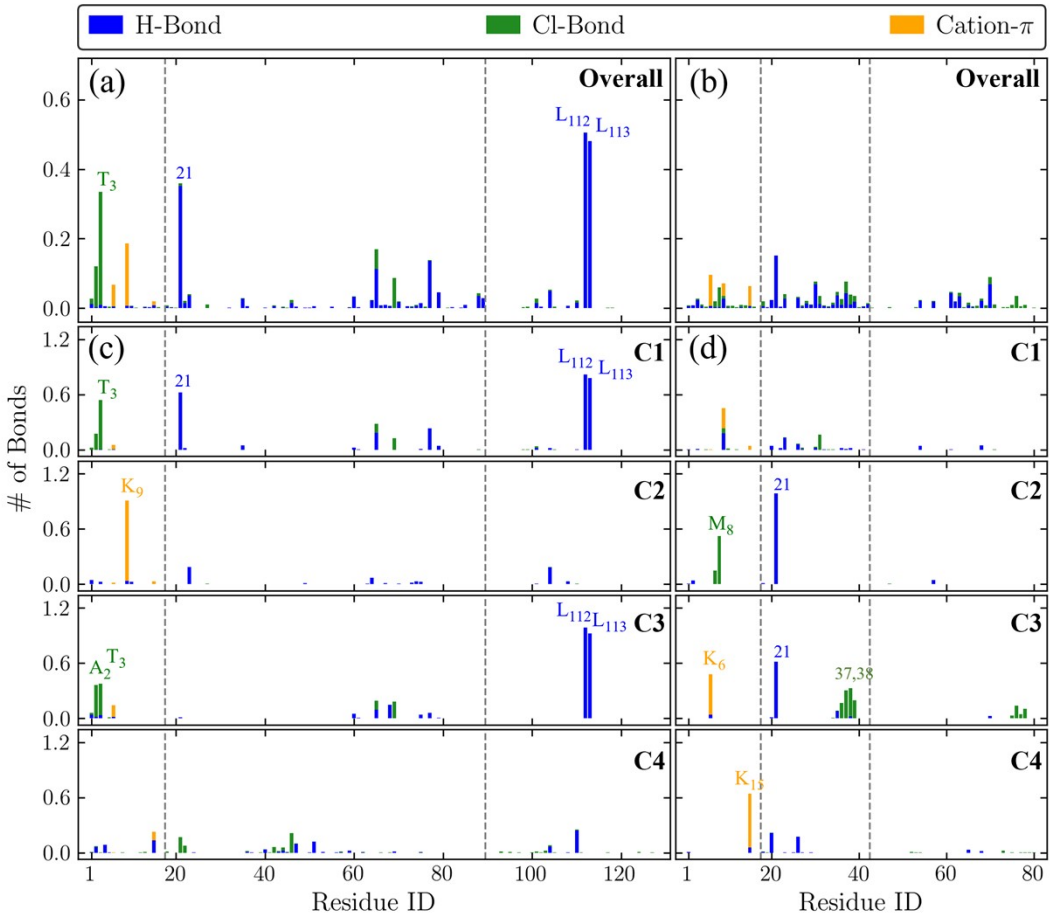

Figure S10. Distribution of residue-based number of polar interactions between ispinesib and Htt-exon-1 systems. The blue/green/orange bars correspond to the hydrogen bond/halogen bond/cation-π interactions respectively. (a-b) The overall distribution. (a) Q72-holo. (b) Q25-holo. (c-d) The distribution in the top 4 clusters. In each plot, the subplots from top to bottom correspond to the most populated cluster (C1) to the 4th populated cluster (C4) in each Htt-exon-1 holo system. (c) Q72-holo. (d) Q25-holo.

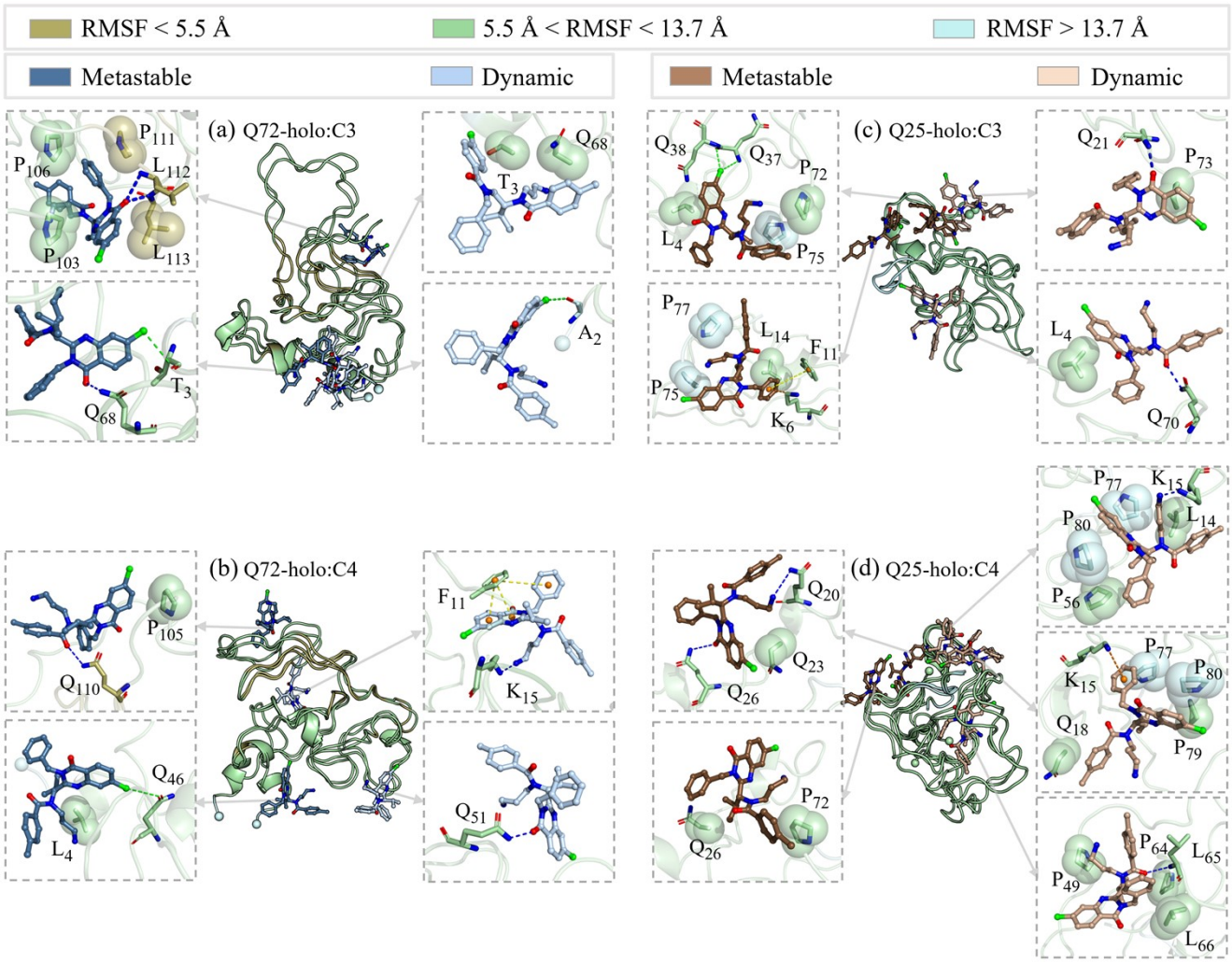

Figure S11. Ispinesib binding sites in the 3rd and 4th cluster of each Htt-exon-1 holo-system. (a-b) Q72-holo system. (a) The 3rd cluster (C3) of Q72-holo. (b) The 4th cluster (C4) of Q72-holo. (c-d) Q25-holo system. (c) The 3rd cluster (C3) of Q25-holo. (d) The 4th cluster (C4) of Q25-holo. In each plot, protein is shown in cartoon representation, and its rigid residues with RMSF < 5.5 Å are in dark khaki, while its flexible residues with RMSF > 13.7 Å are in light cyan, and the rest residues are shown in light green. Ispinesib is shown in stick representation, and the details of binding sites are enlarged around with arrows directed. In the enlarged binding sites, the metastable binding modes are shown in dark color at the left side of complex, while the dynamic binding modes are shown in light color at the right side. The protein residues interacting with ispinesib are also shown in sticks. The detailed interactions are indicated with blue dashes for hydrogen bonds, green dashes for halogen bonds, yellow dashes for π-π stacking, and orange dashes for cation-π interactions. The side-chains with hydrophobic interactions are also shown in sphere.

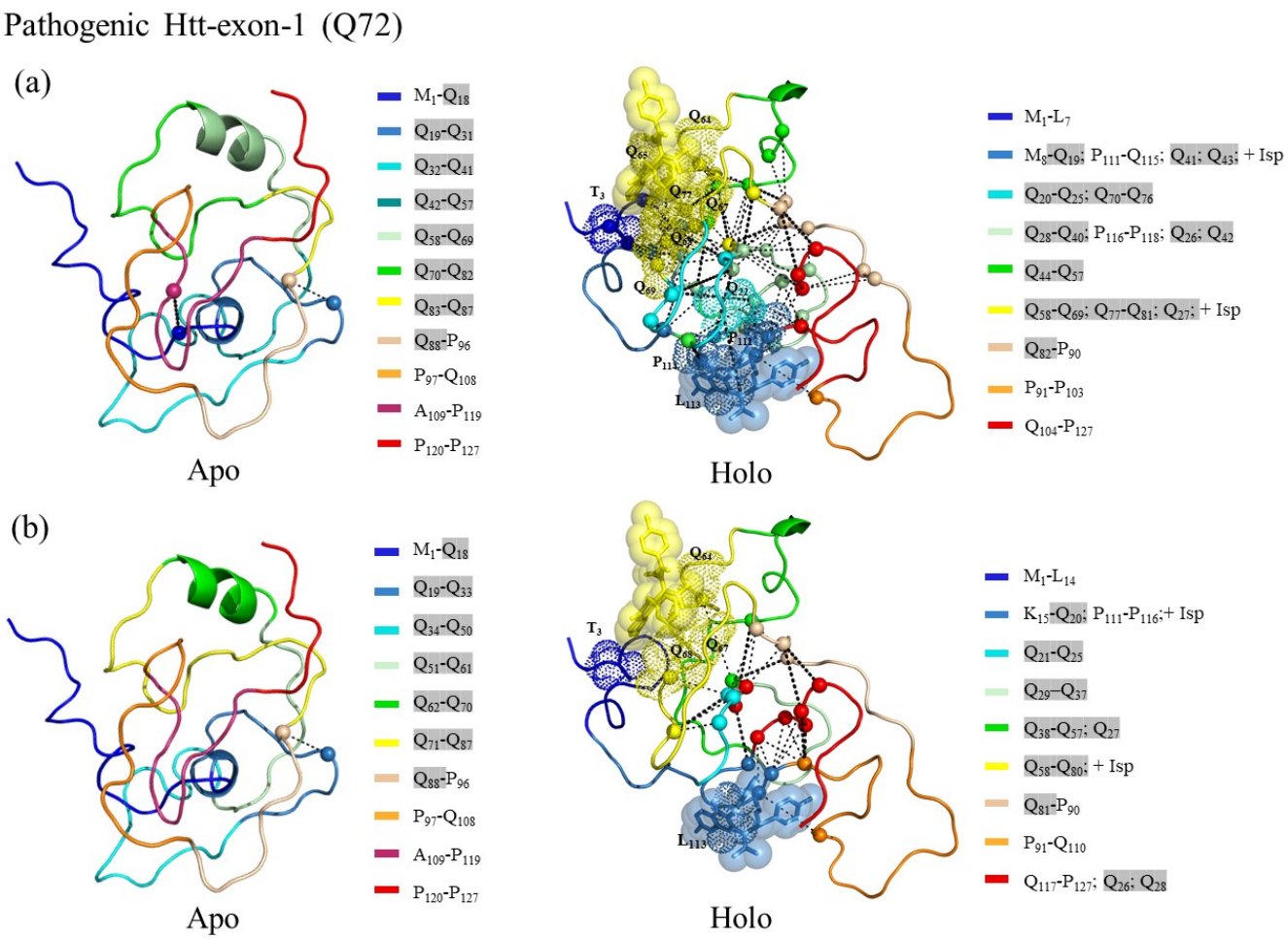

Figure S12. Dynamical community networks in pathogenic Htt-exon-1 (Q72) systems under different thresholds of residue-residue contact probability. The left plots correspond to the apo-forms, and right plots to the holo-forms. For each system, the representative conformation is selected from its most populated cluster and colored according to its community membership. From N-terminus to C-terminus, the communities are colored from blue to red. Dashed gray lines indicate inter-community edges, with solid spheres for the corresponding residues connected through the edges. The width of line is proportional to the betweenness of the edge. The back-bone of Htt-exon-1 systems is in cartoon representation. The residues connected with ispinesib are shown in dots, and ispinesib molecules are in stick with transparent spheres. The corresponding residue indexes of each community are also listed nearby, with the gray shaded area indicating the polyQ region. (a) Threshold of contact probability equal to 60%. The residues connecting with ispinesib are T_3, Q_{22}, Q_{64}, Q_{65}, Q_{67}, Q_{68}, Q_{69}, Q_{77}, P_{111}, L_{113} and P_{114}. (b) Threshold equal to 65%. The residues connecting with ispinesib are T_3, Q_{64}, Q_{67}, Q_{68} and L_{113}.

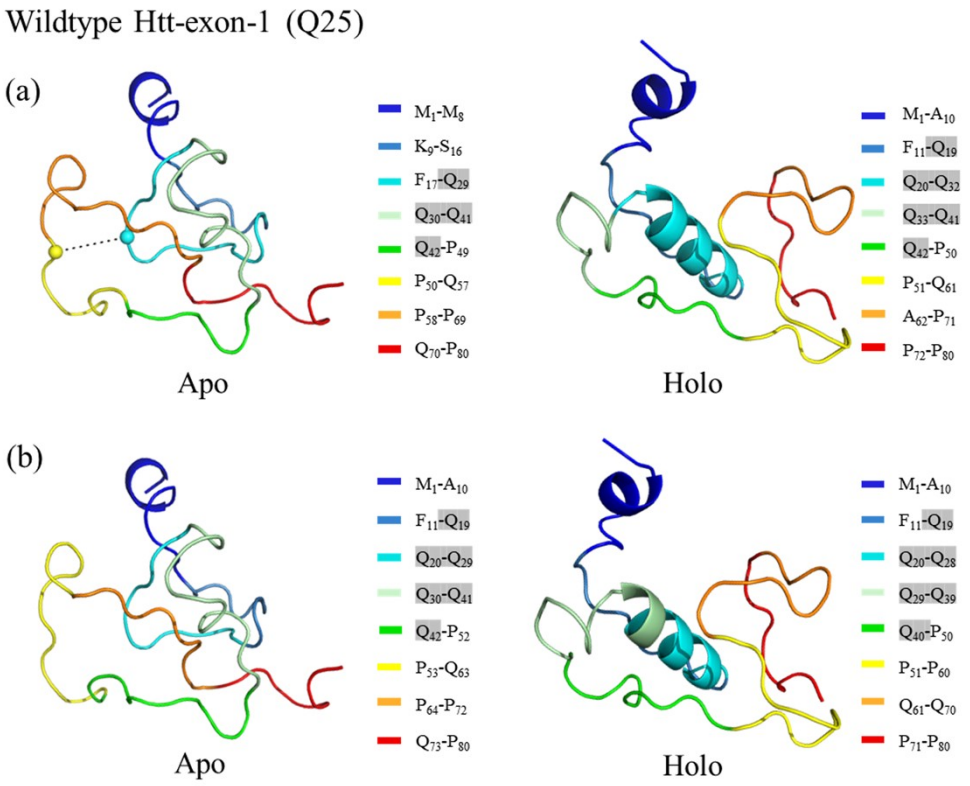

Figure S13. Dynamical community networks in wildtype Htt-exon-1 (Q25) systems under different thresholds of residue-residue contact probability. The left plots correspond to the apo-forms, and right plots for the holo-forms. For each system, the representative conformation is selected from its most populated cluster and colored according to its community membership. From N-terminus to C-terminus, the communities are colored from blue to red. Dashed gray lines indicate inter-community edges, with solid spheres indicating the corresponding residues connected through the edges. The width of line is proportional to the weight of the edge. The back-bone of Htt-exon-1 systems is in cartoon. The residues connected with ispinesib are shown in dots, and ispinesib molecules are in stick with transparent spheres. The corresponding residue indexes of each community are also listed nearby, with the gray shaded area indicating the poly-Q region. (a) Threshold equal to 60%. (b) Threshold equal to 65%. There are no residues connecting with ispinesib in Q25 systems, regardless of the threshold.