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

Supporting Information

The Human eIF4E:4E-BP2 Complex Structure for Studying the Hyperphosphorylation

Juan Zeng[1, *], CuiMin Lu[1], Xuan Huang[1], Yang Li [2, *]

[1.] *School of Biomedical Engineering, Guangdong Medical University, Dongguan 523808, China.*

[2.] *Department of Urology, Huaihe Hospital of Henan University, Kaifeng, 475000, Henan, China.*

Table S1. All simulated systems in this work. The second column is the template for constructing the initial structure of the MD simulation. The third column shows the sequence range of eIF4G and 4E-BPs in the work. The residue number of the conserved Tyr residue in the canonical motif is in parentheses. The last column shows the simulation time.

System	Initial Structure	Sequence	Time
eIF4E:4Gtrj1 E608-L641	5T46	E608-L641 (Y612)	1.5 μs
eIF4E:4Gtrj2 E608-L641			1.5 μs
eIF4E:4Gtrj3 E608-L641			1.5 μs
eIF4E:BP1trj1 T50-E88	5BXV	T50-E88(Y54)	1.5 μs
eIF4E:BP1trj2 T50-E88			1.5 μs
eIF4E:BP1trj3 T50-E88			1.5 μs
eIF4E:BP25T46 T50-I88	5T46	T50-I88(Y54)	1.5 μs
eIF4E:BP25BXV T50-I88	5BXV		1.5 μs
eIF4E:BP25BXV^{trj2} T50-I88			1.5 μs
eIF4E:BP35T46 T36-P74	5T46	T36-P74(Y40)	1.5 μs
eIF4E:4E-BP35BXV T36-P74	5BXV		1.5 μs
eIF4E:4E-			1.5 μs

BP35BXV^{trj2} T36-P74			
eIF4E:BP2$_{P18-I88}$	eIF4E:BP25BXV T50-I88	P18-I88(Y54)	1.5 μs

Table S2. The PDB IDs of the experimental structures of eIF4E-binding proteins. The first column is the organism with the abbreviations in parentheses. The second column is the PDB IDs of structures that only include the canonical eIF4E-binding motif (CEBM). The third column is the PDB IDs of structures including both the CEBM and the auxiliary eIF4E-binding motif (CEBM&AEBM).

Organism	PDB IDs of CEBM	PDB IDs of CEBM&AEBM
Homo sapiens	1WKW, 2JGB, 2JGC, 2V8W, 2V8X, 2V8Y, 2W97, 3AM7, 3HXG, 3HXI, 3M93, 3M94, 3U7X, 4AZA, 4BEA, 5EHC, 5EI3, 5EIR, 5EKV, 5XLN, 5ZJY, 5ZJZ, 5ZK5, 5ZK7, 5ZK9, 5ZML, 7EZW	4UED, 5BXV, 5NVK, 5NVL, 5NVM, 5NVN, 5T46
Drosophila melanogaster (D.M.)	4UEC, 5T48	4AXG, 4UE8, 4UE9, 4UEA, 4UEB, 5ABU, 5ABV, 5T47,
Saccharomyces cerevisiae (S.C.)		1RF8, 6FC1, 6FC2, 6FC3
Mus musculus	1EJ4, 1EJH	
Caenorhabditis elegans (C.E.)		5ABX, 5ABY
Thermochaetoides thermophila DSM 1495 (T.T.)		6FBZ, 6FC0
Cucumis melo (C.M.)		5ME5
Leishmania major (L.M.)		5WB5

Figure S1. (A) The sequence alignment of the human 4E-BPs family and the corresponding peptide in eIF4G. The identical residues are colored in blue. The canonical eIF4E-binding motif (CEBM) is colored in red. The auxiliary eIF4E-binding motif (AEBM) is colored in green. The figure is plotted with Jalview software. (B) The structure and function of eIF4E in mRNA translation. eIF4E interacting with eIF4G and m^7GTP initiates mRNA translation. 4E-BPs share the CEBM and compete with eIF4G for binding eIF4E, then inhibit mRNA translation. The hyperphosphorylated 4E-BP2 at T37, T46, S65, T70 and S83 dissociates from eIF4E, then eIF4G binds onto eIF4E. There are three binding pockets at the surface of eIF4E as shown in the structure.

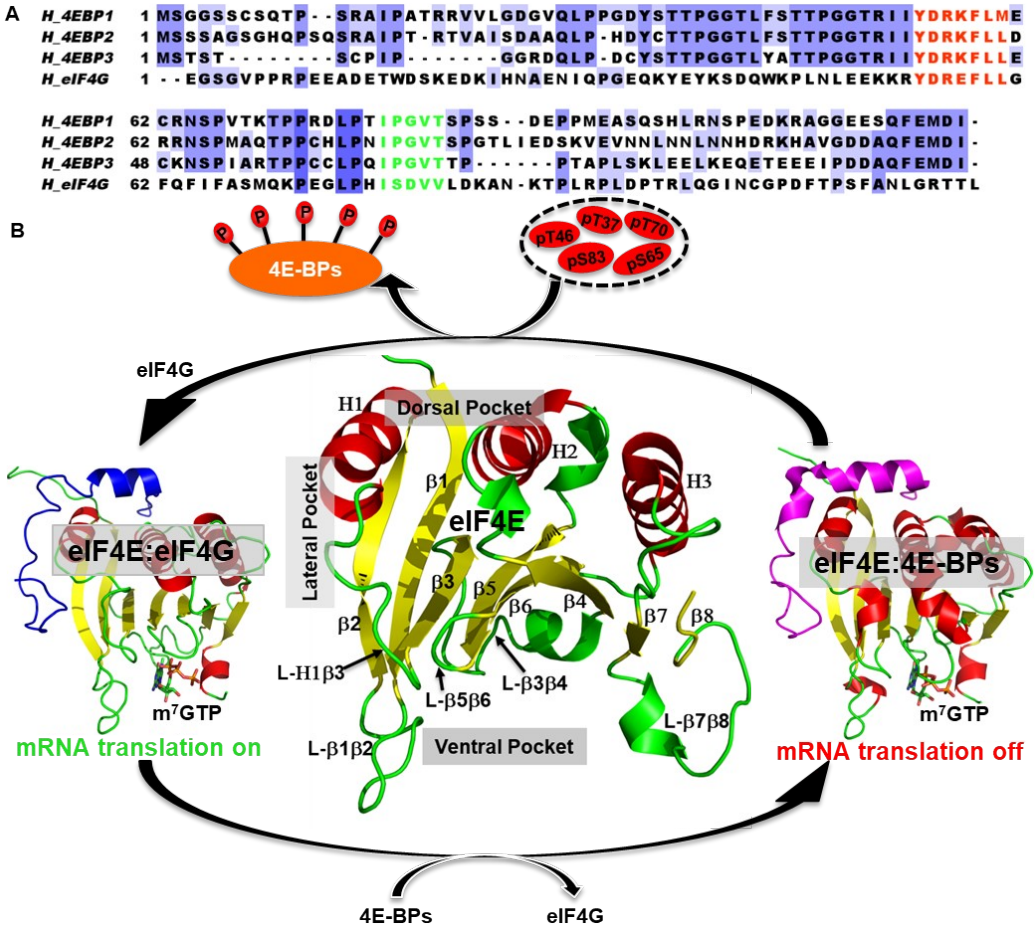

Figure S2. The two-dimensional (2D) root-mean-standard deviations of backbone atoms (RMSD_BB) between PDB ID 5T46 and 5BXV. $5T46_B$ and $5T46_D$ represent the results from chains B and D of eIF4G, respectively. $5BXV_B$ and $5BXV_D$ represent the RMSD_BB values from chains B and D of 4E-BP1, respectively.

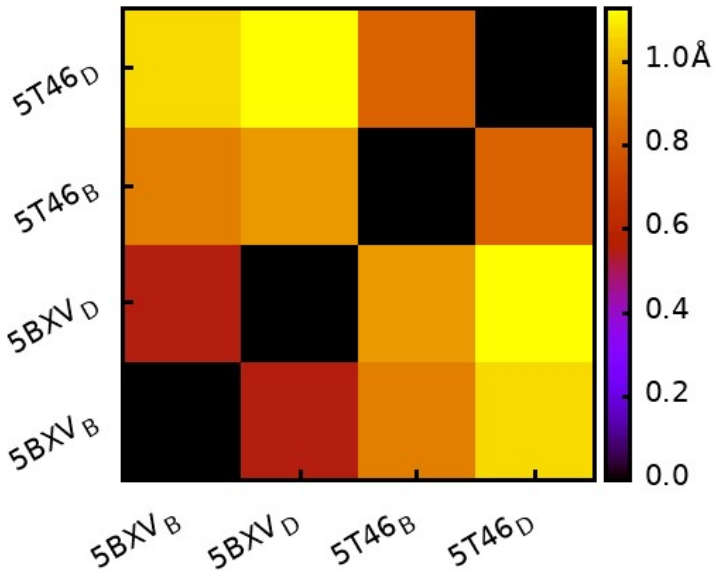

Figure S3. (A) The distribution of the experimental complex structures of eIF4E-binding proteins. The blue bars represent the number of complex structures including only the CEBM. The purple bars represent the number of complex structures including both the CEBM and the AEBM. (B) The 3-dimensional structures of eIF4E-binding proteins with both CEBM and AEBM. The PDB IDs are 5BXV, 5T46, 5ABV, 6FC2, 5ME5, 5ABY, and 6FC0.

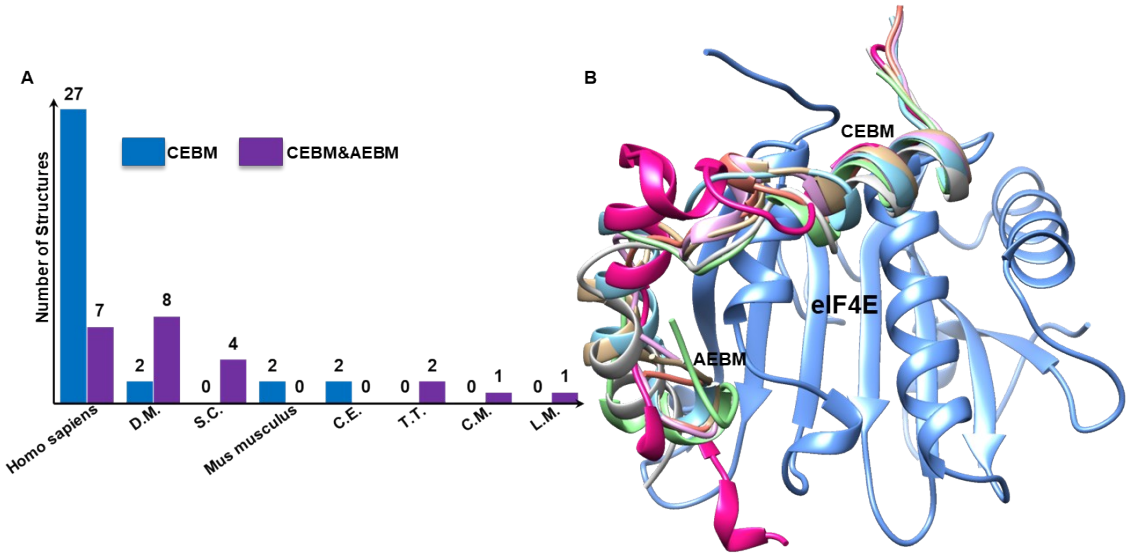

Figure S4. The RMSD_BB values of eIF4E (A1, B1, C1 and D1 excluding the first 13 residues), eIF4G/4E-BPs (A2, B2, C2 and D2), the m^7GTP cap (A3, B3, C3 and D3), and the whole system (A4, B4, C4 and D4) evolve as the simulation time.

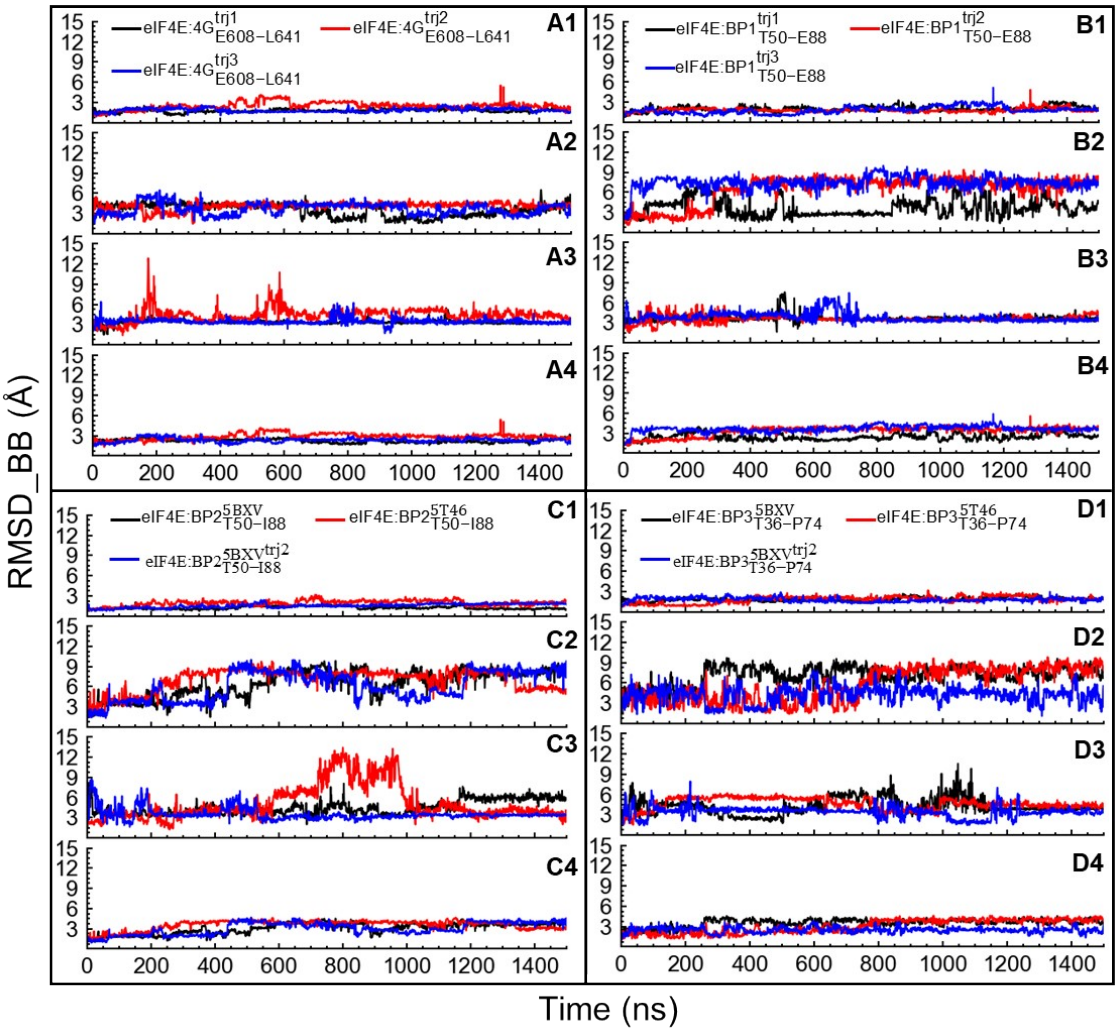

Figure S5. The secondary structure of eIF4Gtrj1 E608-L641 (A1), eIF4Gtrj2 E608-L641 (A2), eIF4Gtrj3 E608-L641 (A3), 4E-BP1trj1 T50-E88 (B1), 4E-BP1trj2 T50-E88 (B2), 4E-BP1trj3 T50-E88 (B3), 4E-BP25BXV T50-I88 (C1), 4E-BP25T46 T50-I88 (C2), 4E-BP25BXV^{trj2} T50-I88 (C3), 4E-BP35BXV T36-P74 (D1), 4E-BP35T46 T36-P74 (D2) and 4E-BP35BXV^{trj2} T36-P74 (D3) functions as the simulation time.

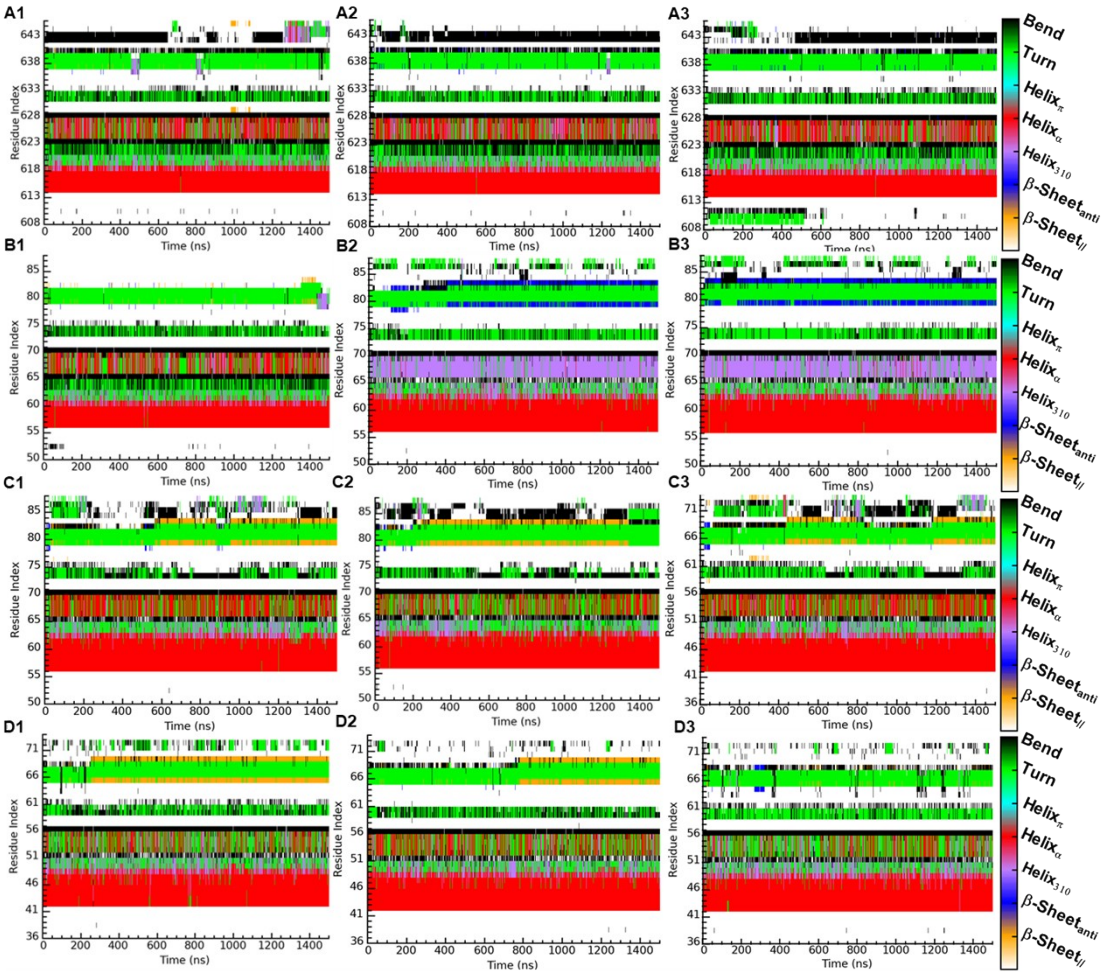

Figure S6. The secondary structure of eIF4E-binding proteins from experiments. The *x*-axis is the sequence of eIF4E-binding proteins, where the highly conserved Tyr residue is numbered 0. The *y*-axis is PDB IDs. The subscript of PDB ID is the chain.

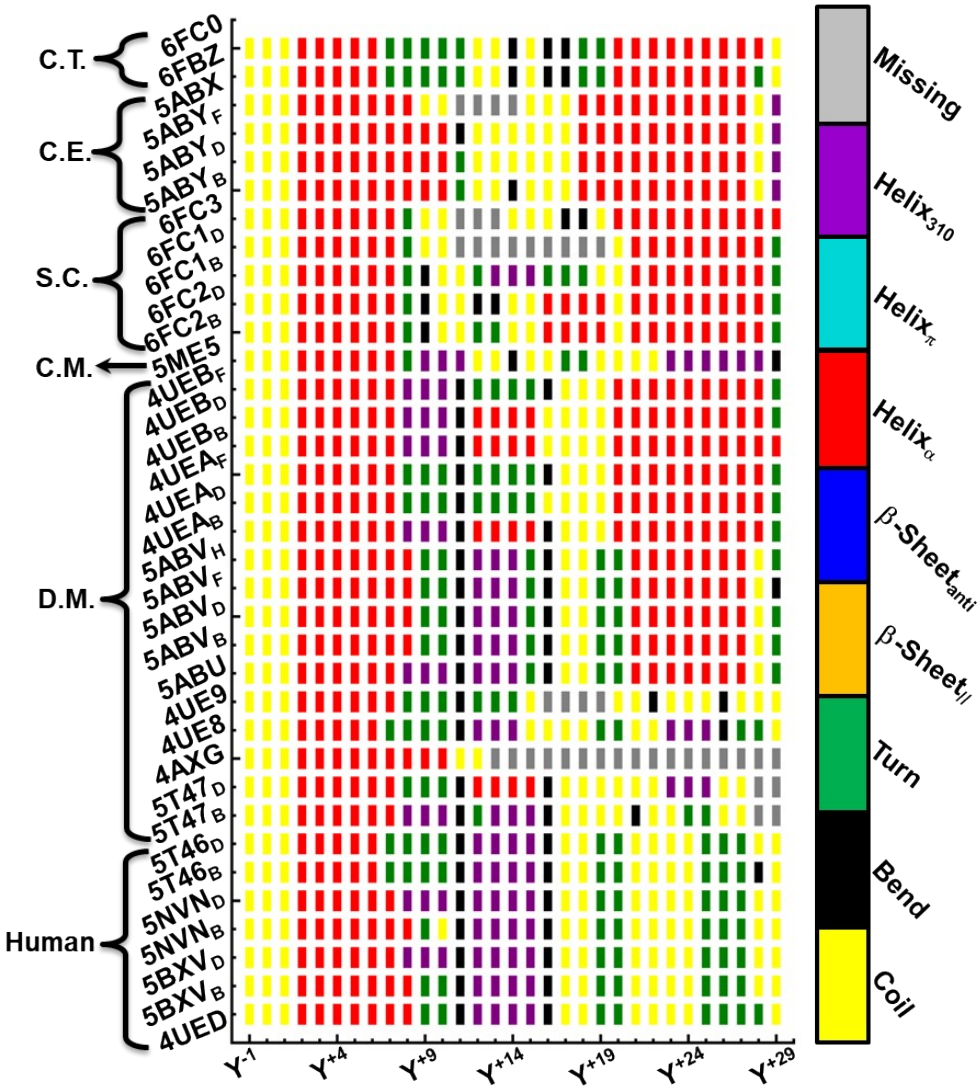

Figure S7. The presentative structure of the 4E-BP2$_{P18\text{-}S65}$ peptide from our previous REMD simulation (JCTC, 2017, 13, 320.) is shown in the cartoon model.

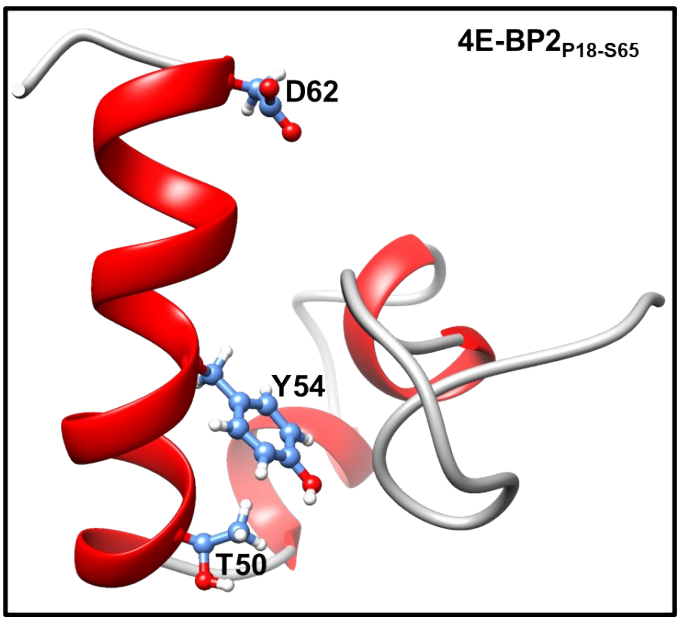

Figure S8. (A) The experimental structures of PDB IDs 5ABU (protein Mextli from D.M.), 4UEA (a designed peptide based on eIF4G), and 5ABX (protein Mextli from C.E.). (B) The experimental structures of PDB ID 6FC3 (yeast caf20p). (C) The experimental structures of PDB ID 6FC0 (C.T. eIF4G). (D) The experimental structures of PDB ID 6FC2 (yeast Eap1p).

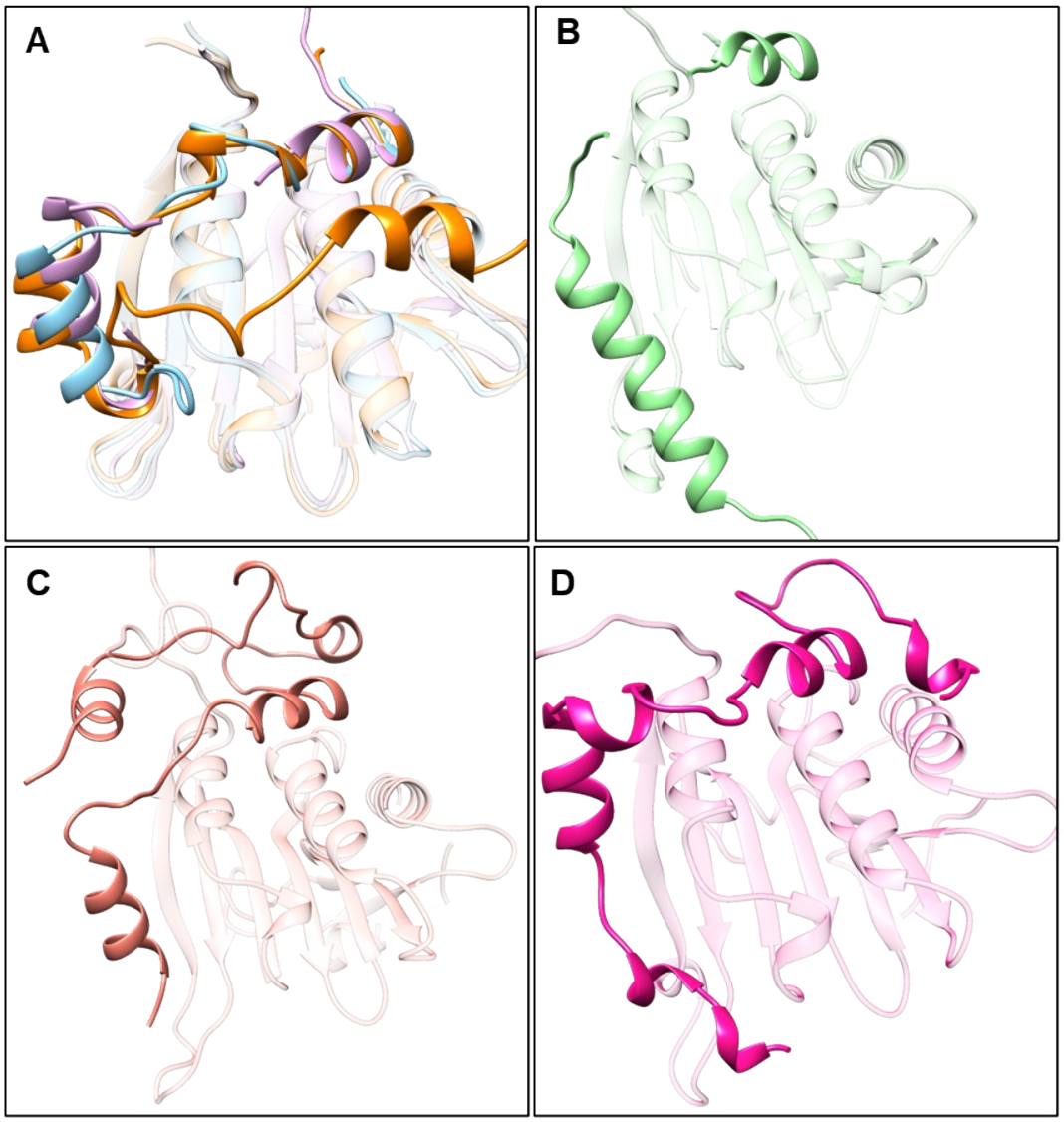

Figure S9. The sequence alignment of eIF4E-binding proteins with experimental structures. This figure is plotted with Jalview software. The MuscleWS method is used to align the sequence.

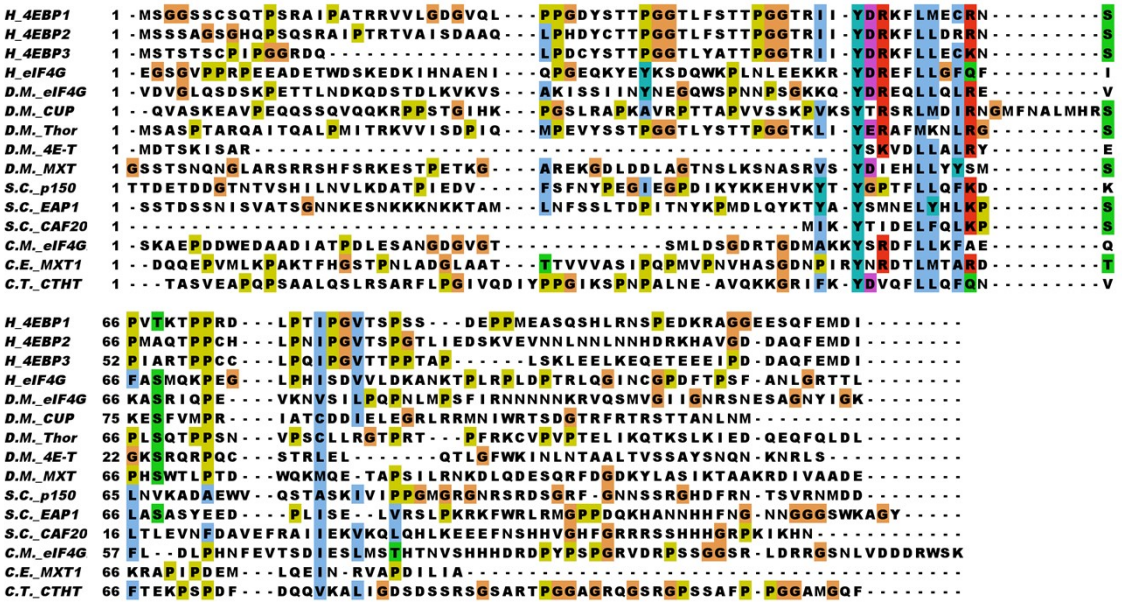

Figure S10. The percentage of HBond between human eIF4E and eIF4G/4E-BPs is calculated based on the last 300-ns trajectory. For defining the HBond, the distance cutoff between the acceptor and donors is set to 3 Å, and the angle cutoff among the acceptor, the hydrogen atom of donor and donor is set to 135^{o}. The percentage of HBond is from the total number of HBonds divided by the total number of frames, then is averaged from three simulations. The black bars show the results from the NTer region. The red bars represent the results from the ECEBM. The blue and green bars are the results of the elbow loop and linker, respectively. The cyan and brown bars represent the results of the AEBM and CTer. The bar in the full filled means the HBond shows in all three simulations. The bar filled with horizontal lines means the HBond is observed in two of three simulations. The bar filled with the slash means the HBond is only observed in one of three simulations.

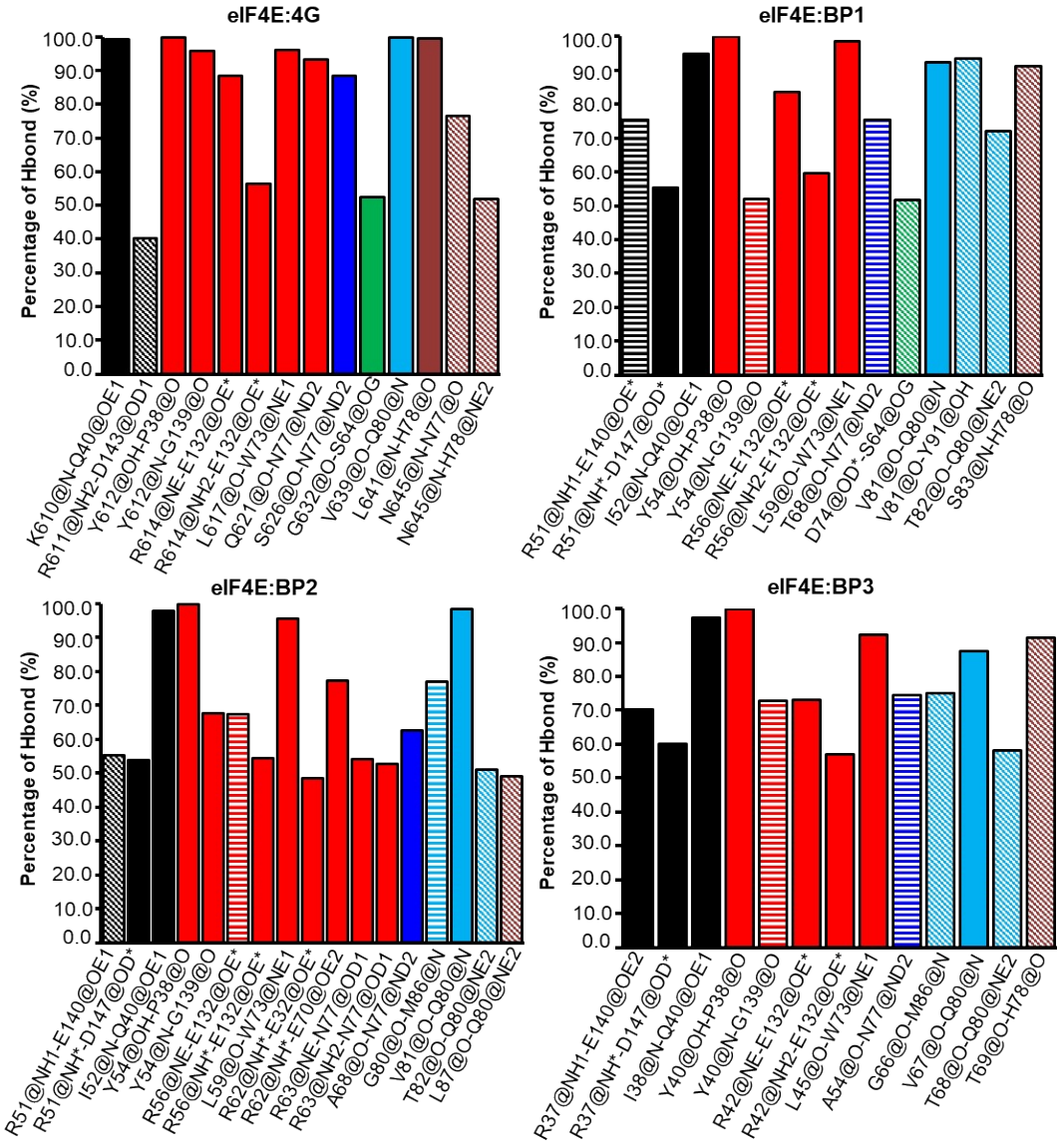

Figure S11. The salt bridges between the R62^{BP2} residue and E32^{4E}, E70^{4E}.

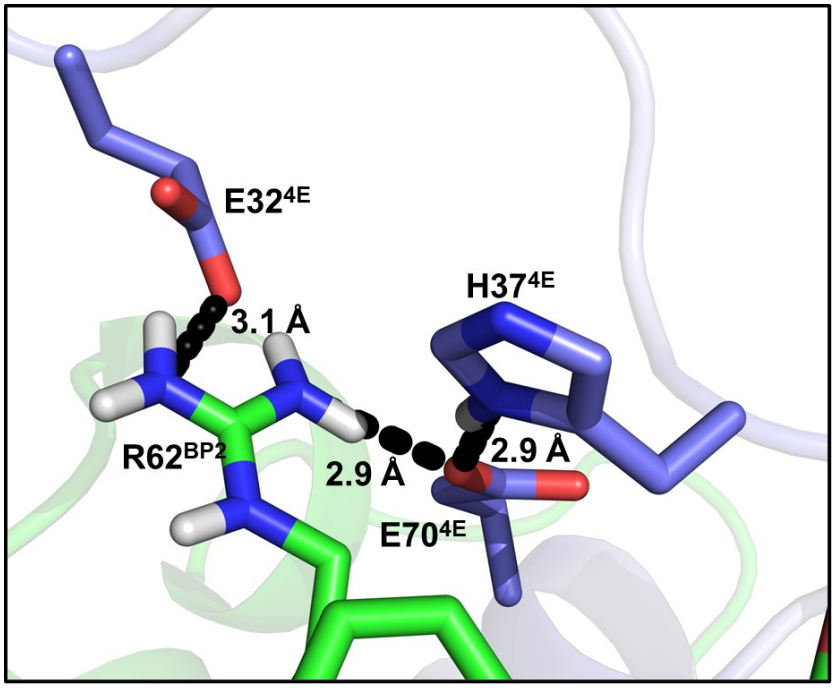

Figure S12. The representative structures of human eIF4E:4Gtrj1 E608-L641 (orange), eIF4E:BP1trj1 T50-E88 (green), eIF4E:BP25BXV T50-I88 (cyan), and eIF4E:BP35BXV T36-P74 (blue) complexes are shown in the cartoon model. The eIF4E is colored in grey.

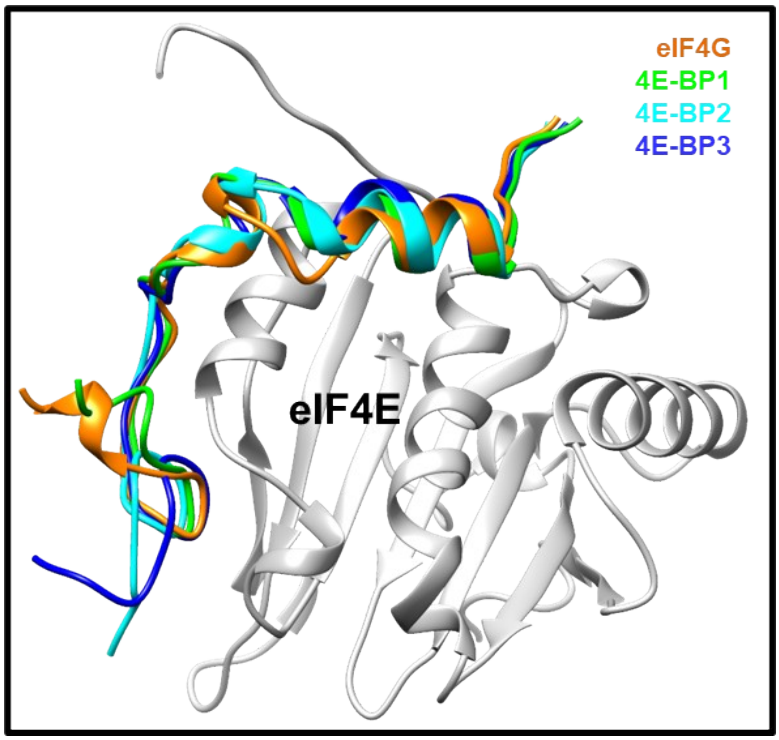

Figure S13. (Left) The RMSD values of the eIF4E:BP2$_{P18\text{-}I88}$ complex function as the simulation time. (Right) The RMSF values of the 4E-BP2$_{P18\text{-}I88}$ peptide are calculated from the last 500-ns trajectory

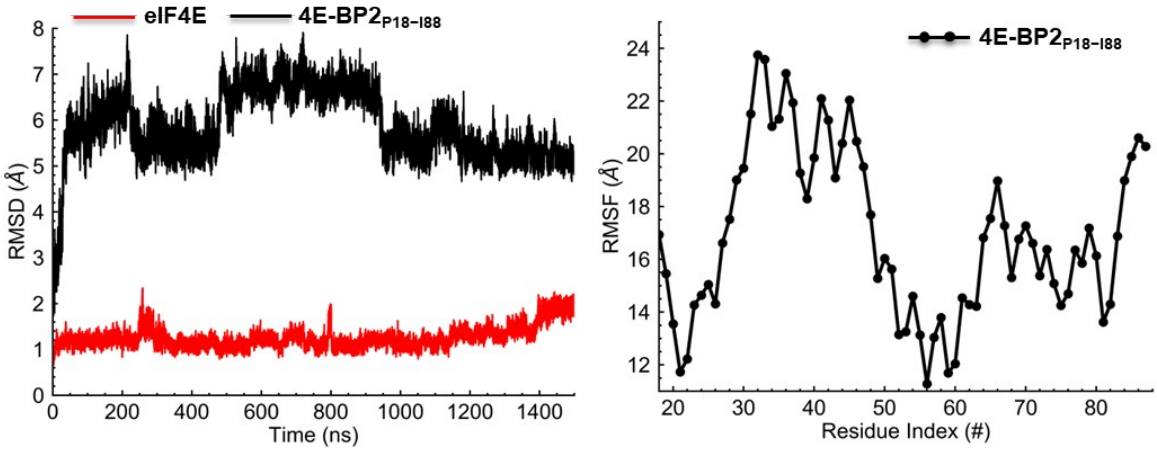

Figure S14. The experimental structures of S.C. eIF4E:eIF4G complex (Left) and C.T. eIF4E:eIF4G complex (Right).