Received: 9 January 2022 | Accepted: 10 May 2022

DOI: 10.1002/prot.26389

RESEARCH ARTICLE

PROTEINS WILEY

Specific recognition between YTHDF3 and m^{6}A-modified RNA: An all-atom molecular dynamics simulation study

Wenxue Zhou | Zhongjie Han | Zhixiang Wu | Weikang Gong | Shuang Yang | Lei Chen | Chunhua Li

Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing, China

Correspondence
Chunhua Li, Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China.
Email: chunhuali@bjut.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 31971180

Abstract

The YTH domain of YTHDF3 belongs to a class of protein "readers" recognizing the N6-methyladenosine (m^{6}A) modification in mRNA. Although static crystal structure reveals m^{6}A recognition by a conserved aromatic cage, the dynamic process in recognition and importance of aromatic cage residues are not completely clear. Here, molecular dynamics (MD) simulations are performed to explore the issues and negative selectivity of YTHDF3 toward unmethylated substrate. Our results reveal that there exist conformation selectivity and induced-fit in YTHDF3 binding with m^{6}A-modified RNA, where recognition loop and loop6 play important roles in the specific recognition. m^{6}A modification enhances the stability of YTHDF3 in complex with RNA. The methyl group of m^{6}A, like a warhead, enters into the aromatic cage of YTHDF3, where Trp492 anchors the methyl group and constraints m^{6}A, making m^{6}A further stabilized by π–π stacking interactions from Trp438 and Trp497. In addition, the methylation enhances the hydrophobicity of adenosine, facilitating water molecules excluded out of the aromatic cage, which is another reason for the specific recognition and stronger intermolecular interaction. Finally, the comparative analyses of hydrogen bonds and binding free energy between the methylated and unmethylated complexes reveal the physical basis for the preferred recognition of m^{6}A-modified RNA by YTHDF3. This study sheds light on the mechanism by which YTHDF3 specifically recognizes m^{6}A-modified RNA and can provide important information for structure-based drug design.

KEYWORDS
dynamic process, N6-methyladenosine, specific recognition and interaction, YTHDF3

1 | INTRODUCTION

N6-methyladenosine (m^{6}A), as a representative of RNA epigenetic modification, plays an essential role in modulating many biological processes such as brain development and immune response to infections.[1] The YT521-B homology (YTH) domain-containing proteins serve as the "Reader" for recognizing m^{6}A in a methylation-dependent manner,[2] which include five types in humans namely YTHDC1, YTHDC2, YTHDF1, YTHDF2, and YTHDF3. YTHDF3, as the last "Reader" crystallized in 2020,[3] harbors a specific m^{6}A-binding surface and aromatic cage similar to other members. Currently, its molecular basis for selective m^{6}A recognition is not well understood. Considering its importance in cancer progression,[4] exploring the binding dynamics of YTHDF3 with m^{6}A-modified RNA is highly significant for understanding their specific recognition mechanism and for developing specific YTHDF3 inhibitors.[5]

In 2020, the crystal structure of YTHDF3 YTH domain in complex with m^{6}A-modified RNA was solved at 2.7 Å resolution,[3] which is

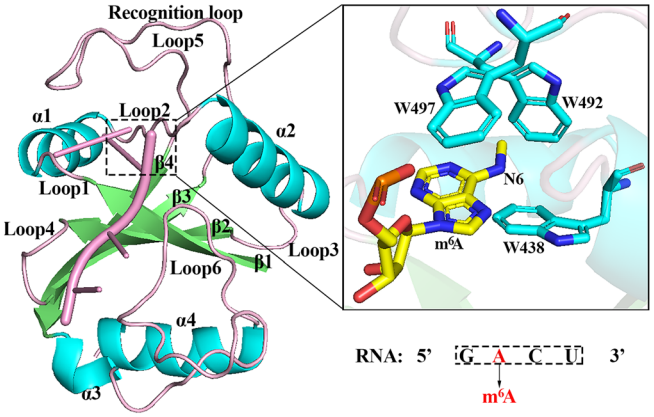

FIGURE 1 Crystal structure of YTHDF3 YTH domain in complex with m^{6}A-modified RNA (PDB ID: 6ZOT). The α-helices, β-sheets, and loops of protein and RNA are labeled. For clarity, the aromatic cage residues and m^{6}A are zoomed in to show more details.

composed of four α-helices and four β-strands, as shown in Figure 1. The β1 strand is parallel with β2 and β4 but antiparallel with β3 to form a β barrel which is wrapped around by the four α-helices to constitute a hydrophobic core.[6] The RNA is tightly locked in a hydrophobic pocket formed by α1, loop2, loop4, loop6, and loop5 also named recognition loop (residues Lys480–Lys500).[3] Specifically, the m^{6}A is positioned in an aromatic cage of three aromatic residues, of which Trp492 from loop5 forms the base, and Trp438 from loop2 and Trp497 from loop5 form the walls. Trp438 and Trp492 are absolutely conserved in YTH domain protein family, whose mutations to alanine completely disrupt protein binding to m^{6}A-modified RNA.[7]

YTH protein family attracts widespread attention from experimental and theoretical researchers. On experimental side, Xu et al. performed isothermal titration calorimetry (ITC) assays and found that both human and yeast YTH domains can recognize m^{6}A-containing RNA, regardless of RNA length, implying YTH domain is an evolutionarily conserved m^{6}A-dependent RNA binding domain.[8] In 2014, Theler et al. used nuclear magnetic resonance (NMR) measurement and mapped the perturbed resonances on YTH domain, revealing that the domain binds the methylated and unmethylated RNA in a similar way.[9] Furtherly, using the liquid chromatography–tandem mass spectrometry (LC–MS/MS) method, Wang et al. found that YTHDF3 has a 5-fold higher binding affinity to the methylated compared to the unmethylated RNA.[10] On theoretical side, in 2020, Li et al. showed that YTHDF3 has similar intrinsic plasticity and protein motions to YTHDF1 and YTHDF2[3] using MD simulations on the three proteins. In the same year, Li et al. studied the importance of the methyl group for RNA-specific recognition by YTHDC1 utilizing the alchemical free energy simulation method.[11] In 2021, Krepl et al. explored the contribution of the methylated substrate's excluding bulk water molecules to YTHDC1's preference for m^{6}A-modified RNA by MD simulations.[12] The above studies give some clues on the specific recognition of m^{6}A-modified RNA by YTH domain. However, few studies have focused on the recognition mechanism of YTHDF3 and m^{6}A-modified RNA. The details on critical residue–nucleotide interactions and the dynamic process involved in their recognition and interactions still need a thorough elucidation.

In this work, MD simulations are performed on YTHDF3 YTH complexes with methylated and unmethylated RNA to dissect specific recognition mechanism of m^{6}A-modified RNA by YTHDF3. Through monitoring the binding dynamics, we explore the contributions of each of aromatic cage residues and methyl group's hydrophobicity to the specific and high affinity binding between m^{6}A-modified RNA and YTHDF3.

2 | MATERIALS AND METHODS

2.1 | Biomolecular systems and molecular dynamics simulation settings

We downloaded the crystal structure of YTHDF3 YTH domain in complex with m^{6}A-modified RNA with PDB ID 6ZOT.[3] Three independent MD simulations were performed on the protein in its apo form (the structure with RNA eliminated from the complex 6ZOT, denoted as apo_pro), in complex with methylated RNA (the structure 6ZOT, denoted as m^{6}A_com), and in complex with unmethylated one (the structure with H atom replacing CH_3 group in complex structure 6ZOT, denoted as A_com). In the above systems, protein core structures (Arg417–Phe553) were remained. All MD simulations were carried out using GROMACS package[13] with CHARMM36 all-atom force field.[14] The structures were solvated in a cubic box (64 Å × 64 Å × 64 Å) full of TIP3P water molecules,[15] whose size was designated to have a minimum distance of 10 Å between any macromolecule atom and the boundary of the water box.[16] Na^+ and Cl^- ions were added to neutralize the solvated systems and mimic the physiological environment (0.15 M). All systems were trimmed similarly with a slightly varied number of water molecules and salt atoms. Afterward, the systems were energy-minimized by the steepest descent method with the heavy atoms harmonically restrained with a spring constant of 1000 kJ·mol^{-1}·nm^{-1}. Next, 1 ns constant temperature and volume (NTV) equilibrations were run with the positions of heavy atoms restricted. And then 500 ns simulations were performed with all position restraints removed under isothermal-isobaric (NPT) condition (300 K and 1 bar), and snapshots were saved every 10 ps. The linear constraint solver (LINCS) algorithm[17] was used to constrain the covalent bonds involving hydrogen atoms. The van der Waals (vdW) interactions were calculated by a switching function with a twin range cutoff at 10 and 12 Å, and electrostatic interactions were evaluated using the particle-mesh Ewald (PME) summation method.[18]

2.2 | Molecular mechanics Poisson–Boltzmann surface area

The molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) method[19] is widely used for binding free energy (ΔG_{bind})

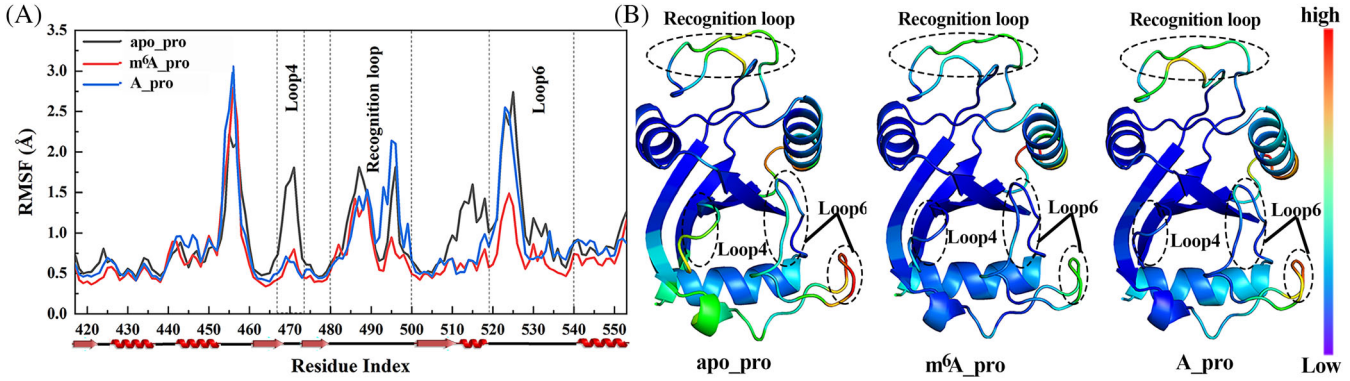

FIGURE 2 (A) Comparison of residue root mean square fluctuation (RMSF) profiles of YTHDF3 (denoted as apo_pro, m^{6}A_pro, and A_pro) in its apo and complexed forms with methylated and unmethylated RNA, respectively, obtained from three independent MD simulations. (B) Structural mappings of the RMSF profiles

calculation from MD trajectory. In brief, ΔG_{bind} can be calculated by the following equations:

$$\Delta G_{bind} = G_{pro+RNA} - (G_{pro} + G_{RNA}) = \Delta E_{MM} + \Delta G_{sol} - T\Delta S, \quad (1)$$

$$\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{ele}, \quad (2)$$

$$\Delta G_{sol} = \Delta G_{PB} + \Delta G_{SA}. \quad (3)$$

In Equation (1), $G_{pro+RNA}$, G_{pro}, and G_{RNA} represent the free energies of protein–RNA complex, protein, and RNA, respectively. ΔG_{bind} can be decomposed into three terms: the gas-phase energy ΔE_{MM}, desolvation free energy ΔG_{sol} and entropy $-T\Delta S$ at temperature T. ΔE_{MM} is composed of van der Waals energy (ΔE_{vdW}) and electrostatic energy (ΔE_{ele}) as shown in Equation (2). In Equation (3), the ΔG_{sol} includes nonpolar (ΔG_{SA}) and polar (ΔG_{PB}) terms. The former is estimated using the solvent accessible surface area (SASA) and the latter is calculated by the Poisson–Boltzmann (PB) model where solvent and solute dielectric constants set to 78.54 and 2, respectively. The entropy ($-T\Delta S$) is calculated by the interaction entropy (IE) method proposed by Zhang et al.[20]

3 | RESULTS AND DISCUSSION

3.1 | RMSDs of the systems and RMSFs of residues from MD trajectories

We performed 500 ns MD simulations on apo_pro, m^{6}A_com, and A_com systems, respectively. Figure S1 shows the changes of root mean square deviations (RMSDs) of the systems as a function of time. From Figure S1, all the systems reach their individual equilibrium states after 50 ns. Thus, all the following analyses were performed on the last 450 ns equilibrium trajectories. At equilibrium stages, the apo protein has a large structural fluctuation with a RMSD of 1.90 ± 0.16 Å, while the proteins in m^{6}A_com and A_com systems are relatively stable with a RMSD of 2.23 ± 0.12 Å and 1.76 ± 0.13 Å, respectively, indicating that both methylated and unmethylated RNA bindings stabilize YTHDF3 structures. In addition, the reason for the higher RMSD of protein in m^{6}A_com than in A_com is partially from that the methylated RNA binding causes a closing conformational change in YTHDF3 binding pocket.

To detect the effect of RNA binding on YTHDF3 flexibility, we calculated root mean square fluctuations (RMSFs) of backbone Cα atoms of YTHDF3 in RNA-bound and -free states, respectively, as shown in Figure 2A, with the corresponding results mapped on protein structure shown in Figure 2B. From Figure 2A, comparing RMSFs of YTHDF3 in apo state and methylated complex, it is found that RNA binding results in a certain loss of flexibility in protein, especially for loop4, recognition loop and loop6, suggesting that the methylated RNA interacts mainly with the three regions. For the protein in unmethylated complex, generally, its flexibility reduction upon RNA binding has a similar pattern to that upon methylated RNA binding with the latter having a greater extent than the former, which on one side is consistent the experimental observation that YTH domain binds the methylated and unmethylated RNAs in a similar way,[9] and in the other side implies that the protein has a stronger interaction with the methylated RNA than with unmethylated one. Furthermore, we note that the main differences in protein flexibility between the two complexes are located in loop6 and recognition loop highlighted in Figure 2B, suggesting that the two loops probably provide the molecular basis for binding pocket reconfiguration to accommodate different RNAs. Interestingly, the flexibilities of both loop6 and recognition loop have a remarkable reduction upon the methylated RNA binding, while upon the unmethylated RNA binding they have a little decrease and an evident increase, respectively. For the recognition loop, the reason for its flexibility difference in the two complexes is likely related to the different RNA recognition modes (see Section 3.2 for detailed analyses). And for loop6, the possible reason, we think, is that compared with the unmethylated RNA, the methylated one has more contacts and interactions (see Section 3.6) with the residues around loop6, which exerts a tighter constraint on loop6, resulting in its larger loss of flexibility.

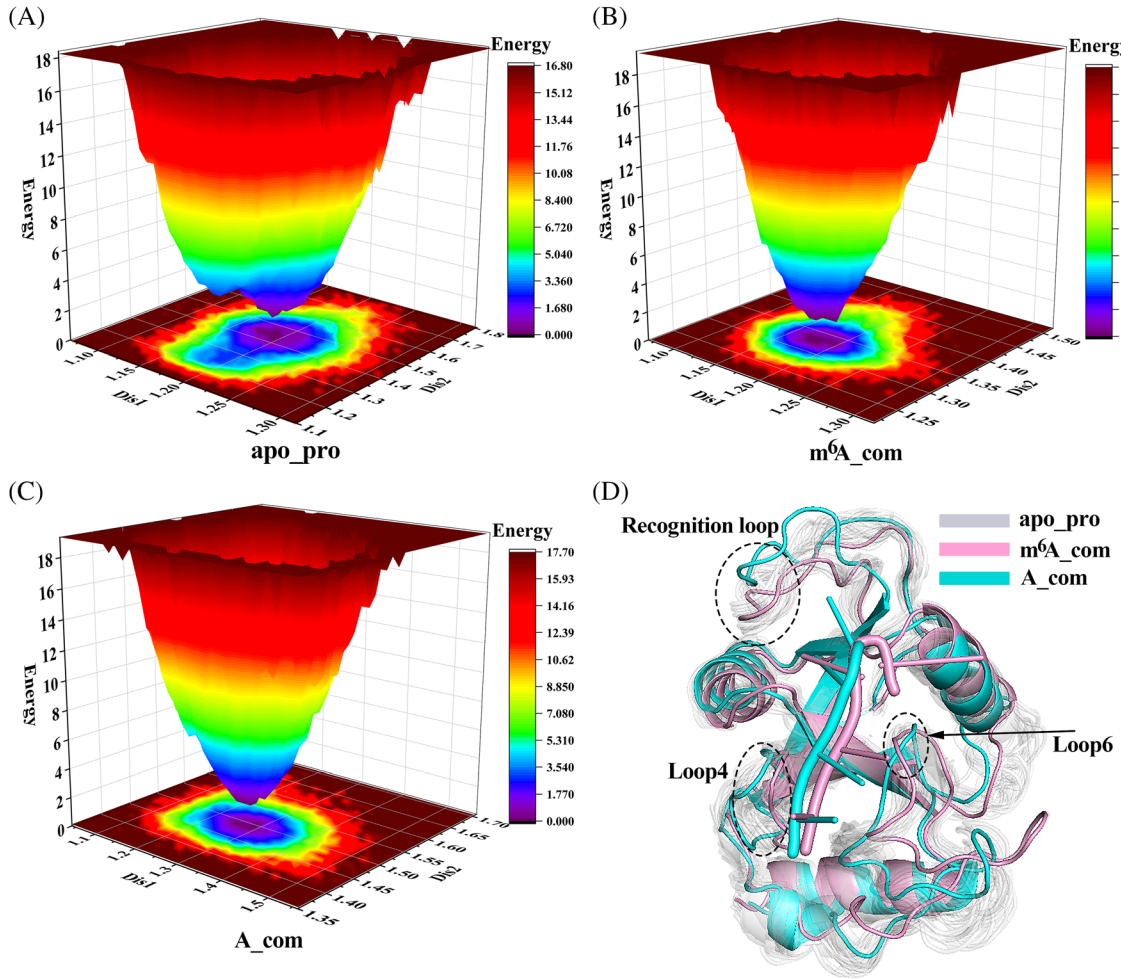

FIGURE 3 Free energy landscapes along the two defined reaction coordinates for YTHDF3 in apo (A), methylated RNA-bound (B) and unmethylated RNA-bound (C) states, respectively, with the representative snapshots from the energy basins shown in (D) where multiple snapshots for apo protein are selected due to its wide energy basins

In summary, the m^{6}A-modified RNA binding results in the loss of YTHDF3 flexibility to a greater extent than the unmethylated RNA binding, especially for loop6 and recognition loop.

3.2 | Motion mode analyses on apo_pro, m^{6}A_com, and A_com systems

To explore the effect of RNA binding on protein collective motions, we performed the principal component analyses (PCAs) on the equilibrium ensembles from apo_pro, m^{6}A_com and A_com systems, with the major structural changes along the first principle component (PC1) illustrated on each of the systems (see Figure S2). From Figure S2, generally the motion amplitudes of protein are attenuated by RNA binding with a larger attenuation observed for protein binding with the methylated RNA, indicating that the methylated RNA exerts a stronger restraint on protein, consistent with the analysis on residue RMSFs. Furthermore, it is noticed that the recognition loop has a different change in motion between protein binding with methylated and with unmethylated RNAs, with a closing trend toward binding pocket for the former and a flipping out one for the latter, which may be related to the fact that two of the three aromatic cage residues involved in specific recognition are located in the recognition loop (illustrated in detail in the following section). Similarly, we illustrated the structural changes along the second and third principle components (PC2 and PC3) for apo_pro, m^{6}A_com and A_com systems as shown in Figure S3. From Figure S3, we observed the similar results to those corresponding to PC1.

To explore the structural difference between the two complexes, we made their free energy landscapes and compared their conformations in local minimal basins. Based on the above analyses, we know that the direct interactions between protein and RNA mainly involve loop4, recognition loop and loop6, which participate in the binding pocket formation. Thus, two distances involved in the three loops were chosen as reaction coordinates: the distance between mass centers of recognition loop Ser493 and Gln494 and of α1 and the distance between mass centers of Loop4 and of Loop6 Ser532-Glu537. Figure 3 shows the free energy landscapes along the reaction

coordinates for the three systems, with the changes of reaction coordinates over time shown in Figure S4. From Figure 3A–C, apo_pro system spans a much larger range along the reaction coordinates and has a much wider energy basin than the two RNA-bound proteins, indicating the former experiences a much larger conformational change. For the latter two, the methylated RNA-bound protein spans a smaller range than the protein binding with unmethylated one. These results are consistent with the analyses on residue RMSFs. Next, the conformations in energy basins were extracted with one conformation selected for complex system and multiple ones for apo protein due to its wide energy basin and then they were superimposed (see Figure 3D) to explore the protein conformational changes upon RNA binding. From Figure 3D, comparing the low-energy conformations from the protein in apo and methylated complex states, it is observed that in apo state, protein loop4, recognition loop and loop6 have an open–close movement (easily observed from Figure S2), during which there exists a conformation very like the low-energy conformation of the protein complexed with the m^{6}A-modified RNA, which probably implies a kind of "conformation selectivity" where the methylated RNA binds selectively to the suitable protein pocket. Furthermore, the narrower binding pocket (formed via loops' closing toward RNA; see Figure S4 (b)) observed in the methylated RNA-bound protein than in apo one (see Figure S4 (a)) suggests a induced-fit occurrence, which helps strengthen intermolecular interactions. The result complies with the current point of view that the "conformation selectivity" and "induced-fit" are not mutually exclusive.[21] For the protein complexed with unmethylated RNA, no significant conformational closing is observed (see Figure 3D and Figure S4), especially for loop4 and recognition loop which on contrary show a little opening movement, which should be not conducive to RNA binding.

In conclusion, the more stable conformation and narrower binding pocket in the protein complexed with methylated RNA than with the unmethylated one validate that YTHDF3 prefers to bind m^{6}A-modified RNA. The evident conformational difference of the recognition loop in the two complexes indicates that it has a great contribution to the specific recognition and interactions between YTHDF3 and m^{6}A-modified RNA.

3.3 | Importance of recognition loop in interacting with methylated RNA

For the role of the recognition loop for the specific recognition and interactions with the methylated RNA, previous studies have reported that its two highly conserved aromatic residues Trp492 and Trp497 and Trp438 in loop4 which form the aromatic cage have a great contribution to YTHDF3-RNA binding.[22,23] However, to what extent the residues contribute to the interactions needs to be illustrated. Thus, we measured the non-bonding interactions including electrostatic and van der Waals (vdW) energies of aromatic cage residues with m^{6}A/adenosine in methylated/unmethylated RNA, with the results shown in Table S1.

From Table S1, the interactions of three aromatic residues with m^{6}A are all stronger than those with adenosine, especially for Trp497 and Trp492. Trp497 and Trp438 form π–π stacking interactions with m^{6}A/adenosine as shown in Figure S5 (local structures from low-energy conformations of the two complexes). From Figure S5, m^{6}A in the methylated complex is stably constrained in the aromatic cage with the dihedral angle of 14.54° ± 7.06° (22.76° ± 10.57°) between the rings from Trp438 (Trp497) and m^{6}A, much smaller than the corresponding value 22.73° ± 12.07° (77.54° ± 21.53°) formed by the residues and adenosine in the unmethylated complex where the π–π stacking's disruption partially attributes to the unmethylated RNA's escaping from the aromatic cage and the flipping of Trp497's aromatic ring (see Figure S5 (b)). The above explains the reduced interactions of Trp438 and Trp497 with adenosine in unmethylated complex. For the interaction of Trp492 and m^{6}A, it is found from Table S1 that the contribution from the methyl group in m^{6}A accounts for 59%, much higher than the corresponding values 34% and 7% for Trp438 and Trp497, respectively, suggesting that Trp492 plays an important role in specifically interacting with m^{6}A and anchoring m^{6}A in the aromatic cage. Thus, we speculate that the absence of methyl group in unmethylated complex reduces the interaction of adenosine with Trp492, which is probably the main reason for RNA's escaping from the aromatic cage.

In a word, the methylation of adenosine enhances its interactions with aromatic cage residues, where Trp492 provides the main specific interaction with methylated RNA, anchoring m^{6}A in the aromatic cage, and Trp497 and Trp438 sandwich m^{6}A through π–π stacking interactions.

3.4 | Water molecules squeezed out of aromatic cage in methylated complex

The methyl group of m^{6}A is of certain hydrophobic property, which is beneficial for its binding to the aromatic cage. The larger buried SASA calculated from the methylated complex (730 ± 44 Å^2) than that from the unmethylated one (605 ± 39 Å^2) can indicate this point to some extent (see Figure S6).

Furtherly, to detect the effect of the methyl group on the water molecules around the aromatic cage, we monitored the radial distributions of water molecules within 8 Å from N6 atom of adenine in the two complexes with the results shown in Figure S7(a). From Figure S7(a), the methylation of adenosine has a significant effect on the radial distribution of water molecules. Evidently, the first hydration shell is more far away from N6 atom, and its peak is much lower in the methylated complex than in the unmethylated one, which is also reflected in the local structures of low-energy conformations of the methylated (with water molecules squeezed out of the aromatic cage) and unmethylated complexes (with the aromatic cage flooded and occupied by water molecules) as shown in Figure S7(b,c), respectively, suggesting that the methyl group resists the proximity of water molecules to some extent. Additionally, from Figure S7(a), the methyl group has a long-range effect,

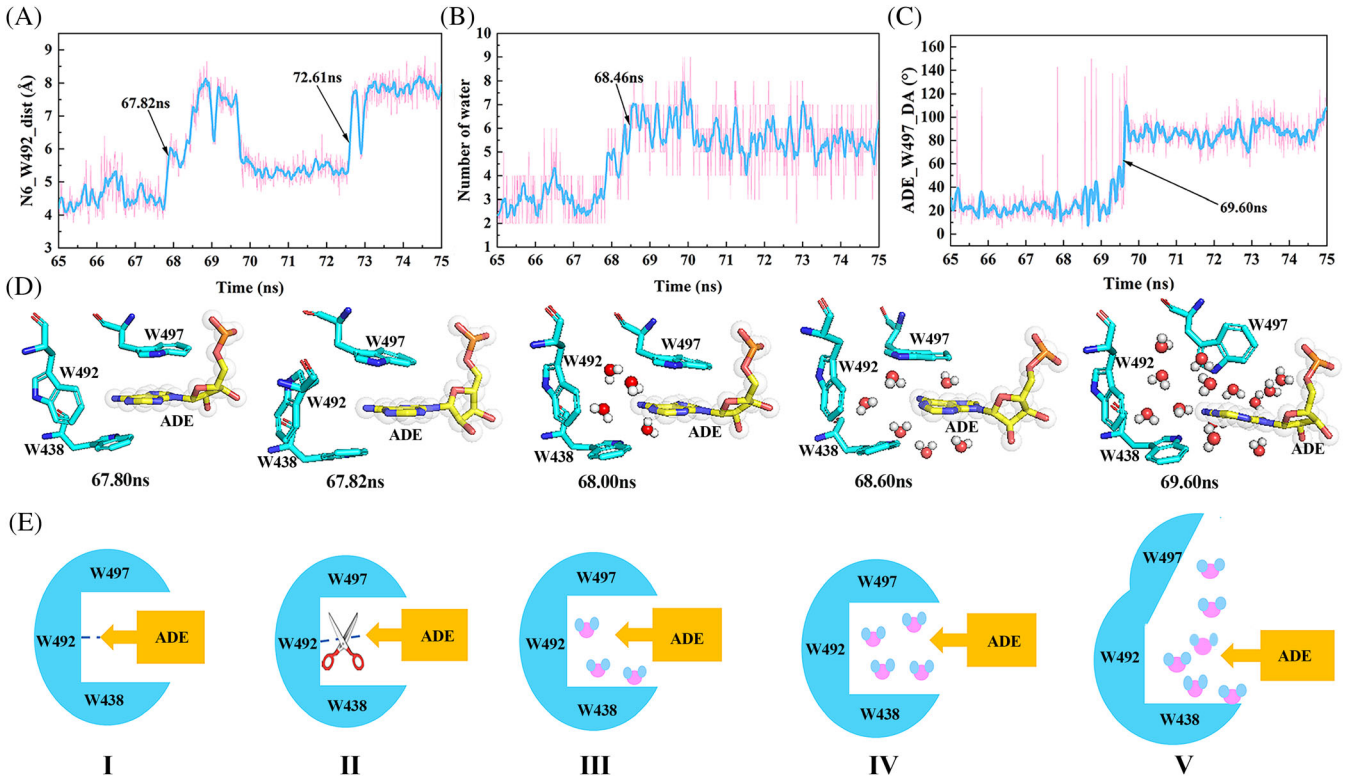

FIGURE 4 Changes of three indicators in the unmethylated complex over time: the distance between adenine N6 atom and Trp492's aromatic ring plane (A), the number of water molecules around adenine N6 atom (B), and the dihedral angle between Trp497's aromatic ring plane and adenine (C), respectively, with the representative snapshots shown in (D). For easy understanding, the process is schematically drawn in (E).

affecting the distribution of water molecules far even beyond 8 Å away from the N6 atom.

3.5 | Hint from the analyses on binding dynamics of YTHDF3 with unmethylated RNA

In fact, for the unmethylated complex system, unbinding between YTHDF3 and RNA to some extent is observed during MD simulation. Through exploring the process, we hope to get some insights into the important factors for stabilizing the methylated complex. Based on the conclusion obtained above that the nonbonding interactions of aromatic cage residues with adenosine and the enhanced hydrophobicity of methyl group play important roles in the specific recognition of m^6A-modified RNA by YTHDF3, we monitored the changes of three indicators over time to explore the issue: the distance between the methylated N6 atom and the plane of Trp492, number of water molecules around N6 atom in adenosine, and dihedral angle between Trp497 and adenosine rings, with the results shown in Figure 4 (showing the changes during 65–75 ns when the unbinding process mainly occurs). From Figure 4A–D, it can be observed that the adenosine first moves away from the aromatic cage at about 67.82 ns, soon the surrounding water molecules rush into the aromatic cage at about 68.46 ns disrupting the interactions of aromatic cage residues with adenosine which further makes adenosine go away from the cage, and later at about 69.60 ns Trp497's aromatic ring rotates outside, losing its π–π stacking interactions with adenosine which makes more water molecules flood into the cage. For easy understanding, we show the process schematically in Figure 4E.

In light of the above analyses and the nonbonding energy between aromatic cage residues and adenosine (Table S1), the methylation enhances not only the interactions of adenosine with aromatic cage residues among which Trp492 plays a critical role in anchoring the methylated RNA in protein binding pocket, but also the aromatic cage's hydrophobicity, which helps prevent the invasion of water molecules, ensuring the stacking interactions of Trp438 and Trp497 with adenine.

3.6 | Comparative analyses of hydrogen bonds and binding free energy between methylated and unmethylated complexes

To reveal the differences in hydrogen bonding pattern in the methylated and unmethylated complexes, we analyzed the intermolecular hydrogen bonds (H-bonds) with an occupancy >40%[24] for the two complex systems, with the results shown in Table S2 and Figure S8 (a schematic diagram). From Table S2 and Figure S8, a total of 12 and 8 H-bonds are observed in the methylated and unmethylated complex systems, respectively. Among the four nucleotides, the adenine has

the biggest difference in the number of formed H-bonds with additional three H-bonds involving Asp534 in loop6 and one involving Lys496 in recognition loop besides the common three H-bonds involving Tyr424, Cys439, and Asn468 formed in the methylated complex compared with the unmethylated one. For cytosine and uracil, they form one less and one more H-bond with YTHDF3, respectively, in the methylated complex compared with the unmethylated one. And guanine does not form H-bonds in both complexes. Generally, for the common H-bonds, their occupancies are higher in the methylated than in unmethylated complexes. The additional H-bonds can partially explain the lower fluctuations of loop6 and recognition loop (see Figure 2A) in methylated complex than in unmethylated one.

Next, we calculated the binding free energies with MM-PBSA method for the two complexes where 45 snapshots were taken at an interval of 10 ns from 50 to 500 ns. The methylated complex has a binding affinity of −128.94 kcal/mol, higher than −106.70 kcal/mol for the unmethylated one, which is qualitatively in agreement with the experiment data.[10] The higher binding affinity attributes to the lower enthalpy change and smaller entropy reduction during the binding process of YTHDF3 with methylated RNA than with unmethylated one. The favorable enthalpy and entropy changes come from the more formed nonbonding interactions and more stable residue–nucleotide interactions due to the introduced methyl group.

4 | CONCLUSION

In this work, we explore the specific recognition and interactions between YTHDF3 and m^{6}A-modified RNA through performing MD simulations on the three systems: protein in apo form and in complexed forms with methylated and unmenthylated RNAs, respectively. The flexibility analyses show that the m^{6}A-modified RNA binding causes a larger loss of protein flexibility, especially for recognition loop and loop6 than the unmethylated RNA binding, suggesting their important roles in the specific recognition. PCA analyses indicate that the methylated RNA exerts a stronger restraint on protein, and reveal a closing (flipping) movement of recognition loop toward (outward) protein binding pocket in methylated (unmethylated) complex. The low-energy conformations from free energy basins reveal that the free YTHDF3 having a variable pocket can recognize and bind m^{6}A-modified RNA using a suitable pocket (conformation selectivity), and further the induced-fit occurs to strengthen their interactions. Based on the nonbonding energy of aromatic cage residues with adenosine and binding dynamics between YTHDF3 and unmethylated RNA, we conclude that Trp492 provides the main specific interaction with methylated RNA, anchoring m^{6}A in the aromatic cage, and Trp497 and Trp438 sandwich m^{6}A through π–π stacking interactions, which are ensured by methyl group as well as the aromatic cage by preventing the invasion of water molecules into the cage. Finally, we carry out H-bond and MMPBSA analyses and the results indicate that the methylation of adenosine enhances the interactions between YTHDF3 and RNA, producing a higher binding affinity and making YTHDF3 prefer to bind with the m^{6}A-modified RNA. In summary, our work is helpful for the understanding of molecular basis for specific binding between YTHDF3 and m^{6}A-modified RNA, and for the related drug design.

AUTHOR CONTRIBUTIONS

Wenxue Zhou and Chunhua Li designed the research. Wenxue Zhou, Zhongjie Han, Zhixiang Wu, Weikang Gong, Lei Chen, and Shuang Yang performed data analyses. Wenxue Zhou and Chunhua Li wrote the manuscript. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31971180).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Chunhua Li https://orcid.org/0000-0002-0895-3506

REFERENCES

1. Zhang C, Fu J, Zhou Y. A review in research progress concerning m6A methylation and immunoregulation[J]. *Front Immunol*. 2019;10:922.
2. Liao S, Sun H, Xu C. YTH domain: a family of N(6)-methyladenosine (m(6)a) readers[J]. *Genom Proteom Bioinf*. 2018;16(2):99-107.
3. Li Y, Bedi RK, Moroz-Omori EV, Caflisch A. Structural and dynamic insights into redundant function of YTHDF proteins[J]. *J Chem Inf Model*. 2020;60(12):5932-5935.
4. Wang Y, Zhang Y, Du Y, et al. Emerging roles of N6-methyladenosine (m(6)a) modification in breast cancer[J]. *Cell Biosci*. 2020;10(1):136.
5. Patil DP, Pickering BF, Jaffrey SR. Reading m(6)a in the transcriptome: m(6)A-binding proteins[J]. *Trends Cell Biol*. 2018;28(2):113-127.
6. Ma C, Liao S, Zhu Z. Crystal structure of human YTHDC2 YTH domain[J]. *Biochem Biophys Res Commun*. 2019;518(4):678-684.
7. Chang G, Shi L, Ye Y, et al. YTHDF3 induces the translation of m(6)A-enriched gene transcripts to promote breast cancer brain metastasis [J]. *Cancer Cell*. 2020;38(6):857-871.
8. Xu C, Liu K, Ahmed H, Loppnau P, Schapira M, Min J. Structural basis for the discriminative recognition of N6-methyladenosine RNA by the human YT521-B homology domain family of proteins[J]. *J Biol Chem*. 2015;290(41):24902-24913.
9. Theler D, Dominguez C, Blatter M, Boudet J, Allain FHT. Solution structure of the YTH domain in complex with N6-methyladenosine RNA: a reader of methylated RNA[J]. *Nucleic Acids Res*. 2014;42(22): 13911-13919.
10. Wang X, Lu Z, Gomez A, et al. N6-methyladenosine-dependent regulation of messenger RNA stability[J]. *Nature*. 2014;505(7481): 117-120.
11. Li Y, Bedi RK, Wiedmer L, Sun X, Huang D, Caflisch A. Atomistic and thermodynamic analysis of N6-Methyladenosine (m(6)a) recognition by the reader domain of YTHDC1[J]. *J Chem Theory Comput*. 2021; 17(2):1240-1249.
12. Krepl M, Damberger FF, von Schroetter C, et al. Recognition of N6-Methyladenosine by the YTHDC1 YTH domain studied by molecular dynamics and NMR spectroscopy: the role of hydration[J]. *J Phys Chem B*. 2021;125(28):7691-7705.
13. Pronk S, Páll S, Schulz R, et al. GROMACS 4.5[J]. *Bioinformatics*. 2013;29(7):845-854.

14. Huang J, MacKerell AJ. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data[J]. *J Comput Chem*. 2013;34(25):2135-2145.
15. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water [J]. *J Chem Phys*. 1983;79(2):926-935.
16. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM[J]. *J Comput Chem*. 2008;29(11):1859-1865.
17. Hess B. P-LINCS: a parallel linear constraint solver for molecular simulation[J]. *J Chem Theory Comput*. 2008;4(1):116-122.
18. Darden T, York D, Pedersen L. Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems[J]. *J Chem Phys*. 1993;98(12):10089-10092.
19. Sheng YJ, Yin YW, Ma YQ, Ding HM. Improving the performance of MM/PBSA in protein–protein interactions via the screening electrostatic energy[J]. *J Chem Inf Model*. 2021;61(5):2454-2462.
20. Duan L, Xiao L, Zhang J. Interaction entropy: a new paradigm for highly efficient and reliable computation of protein–ligand binding free energy[J]. *J Am Chem Soc*. 2016;138(17):5722-5728.
21. Ren W, Dokainish HM, Shinobu A, Oshima H, Sugita Y. Unraveling the coupling between conformational changes and ligand binding in ribose binding protein using multiscale molecular dynamics and free-energy calculations[J]. *J Phys Chem B*. 2021;125(11):2898-2909.
22. Xu C, Wang X, Liu K, et al. Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain[J]. *Nat Chem Biol*. 2014;10(11):927-929.
23. Bedi RK, Huang D, Wiedmer L, et al. Selectively disrupting m6A-dependent protein-RNA interactions with fragments[J]. *ACS Chem Biol*. 2020;15:618-625.
24. Kong R, Xu L, Piao L, Zhang D, Hou TJ, Chang S. Exploring the RNA-bound and RNA-free human Argonaute-2 by molecular dynamics simulation method[J]. *Chem Biol Drug des*. 2017;90(5):753-763.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhou W, Han Z, Wu Z, et al. Specific recognition between YTHDF3 and m^6A-modified RNA: An all-atom molecular dynamics simulation study. *Proteins*. 2022; 1-8. doi:10.1002/prot.26389