JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2019, 15, 7004−7014

pubs.acs.org/JCTC

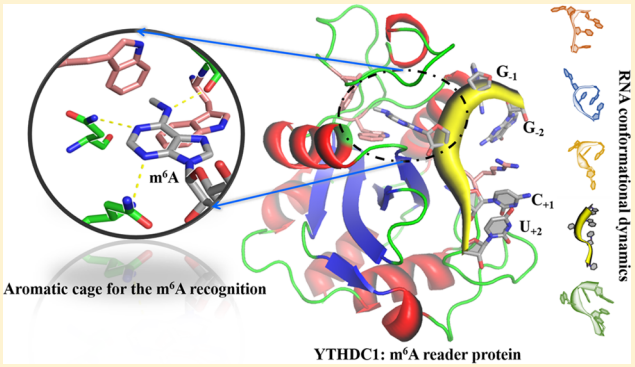
Flexible Binding of m^{6}A Reader Protein YTHDC1 to Its Preferred RNA Motif

Yaozong Li, Rajiv Kumar Bedi, Lars Wiedmer, Danzhi Huang, Paweł Śledź,* and Amedeo Caflisch*

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Supporting Information

ABSTRACT: N^6-Methyladenosine (m^{6}A) is the most prevalent chemical modification in human mRNAs. Its recognition by reader proteins enables many cellular functions, including splicing and translation of mRNAs. However, the binding mechanisms of m^{6}A-containing RNAs to their readers are still elusive due to the unclear roles of m^{6}A-flanking ribonucleotides. Here, we use a model system, YTHDC1 with its RNA motif 5′-G$_{-2}$G$_{-1}$(m^{6}A)C$_{+1}$U$_{+2}$-3′, to investigate the binding mechanisms by atomistic simulations, X-ray crystallography, and isothermal titration calorimetry. The experimental data and simulation results show that m^{6}A is captured by an aromatic cage of YTHDC1 and the 3′ terminus nucleotides are stabilized by cation−π−π interactions, while the 5′ terminus remains flexible. Notably, simulations of unbound RNA motifs reveal that the methyl group of m^{6}A and the 5′ terminus shift the conformational preferences of the oligoribonucleotide to the bound-like conformation, thereby facilitating the association process. The binding mechanisms may help in the discovery of chemical probes against m^{6}A reader proteins.



INTRODUCTION

Methylation of position N6 in adenines, also known as m^{6}A modification, is the most prevalent internal modification of messenger RNA (mRNA) with approximately one in 200 adenine bases modified, and as such, it makes significant contributions to central events in biology.[1,2] On the molecular level, the introduction of m^{6}A affects structures of RNAs and their ability to form protein-RNA interactions and thus modulates processing,[3,4] translation,[5,6] and stability of the cellular transcripts.[7] As a consequence, m^{6}A is implicated in controlling embryonic development processes and stem cell differentiation,[8−11] regulating the mammalian circadian clock,[12] and modulating stress response, for example, heat shock.[13] The malfunctions of the cellular machinery regulating the m^{6}A modification have in turn been linked to pathologies like obesity,[14] cancer,[15] and neurodegeneration.[16] Therefore, the m^{6}A field has recently attracted broad attention of drug discovery community.[17−20]

While the m^{6}A modification has been known for several decades, it was only recently that it has been proven reversible due to the discovery of the selective demethylases (erasers) that remove it from mRNA[21,22] and balance the actions of the writer complex.[23−27] Similar to the epigenetic modifications of histone proteins, the reversibility of the mRNA methylation gives it the potential to constitute an additional layer of epitranscriptomic regulation of gene expression.[28−30] The reader proteins, that is, proteins that recognize the modification and elicit the response upon it, play the key role at the executive end of such a regulation mechanism. A number of proteins have been identified as direct m^{6}A readers with different structures, subcellular localizations, and mechanisms of the recognition of the epitranscriptomic mark. There are five YTH-domain-containing m^{6}A readers that adopt highly similar open α/β folds and specifically recognize the m^{6}A mark through interactions of the aromatic cage.[31−34] Cytoplasmic readers YTHDF1 and YTHDF3 play roles in translation initiation,[6] while the highly similar YTHDF2 reader controls destabilization of cellular transcripts by targeting them to P-bodies.[7] The nuclear reader YTHDC1 plays a role in splicing regulation,[3] while YTHDC2 contains an additional RNAse helicase domain and is important in meiosis.[35]

Specific recognition of m^{6}A-containing transcripts by the YTH-domain readers has been studied in detail by the means of protein crystallography, pointing out at a rigid and conserved recognition of the consensus sequence GG(m^{6}A)-CU. Beyond YTH-domain m^{6}A readers, there are also indirect readers (i.e., proteins not interacting with m^{6}A but binding to m^{6}A-containing transcripts) including HBRNPC[36] and anti-readers like G3BP1 that bind preferably to unmodified transcripts.[37] The crucial role of m^{6}A in the molecular recognition has been confirmed where an aromatic cage consisting of a number of key tryptophan residues enable binding. However, the exact role of m^{6}A-flanking nucleotides

Received: October 2, 2019
Published: October 31, 2019

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

that make hydrogen bonds and salt bridges to the protein[33,38] remains elusive. Further studies on these interactions are crucial not only for understanding the mechanism and timescale of epitranscriptomic regulation but also for informing the development of therapeutic agents.

It is not easy to infer the importance of the interactions between the m^{6}A-flanking nucleotides and reader proteins, partially because they are observed on protein surfaces and thus may be stabilized by crystal packing. In addition, unbound oligoribonucleotides are highly flexible and may adopt many meta-stable conformations in aqueous solutions, which is difficult to study by experiments. To this end, molecular dynamics (MD) may provide complementary information for understanding the binding mechanisms.[39−41] However, the field of MD simulations on RNA and its complexes with proteins, for example, m^{6}A reader proteins and its ligand RNAs, is less mature than that on proteins (or with organic molecules). In particular, three potential issues hinder the use of RNA-related force fields:[42−45] (i) the inaccuracy of ribonucleotide force fields themselves,[43,44] (ii) the lack of parameters for modified nucleotides,[46,47] and (iii) the difficulties in balancing interactions between proteins and RNA molecules.[44,45]

The first issue has been moderately alleviated by the recent improvements of common RNA force fields, for example, CHARMM,[48−50] AMBER,[51] and OPLS.[52] For this purpose, different methodologies have been explored, for example, refining bonded and nonbonded potentials[49,51,52] and adding external hydrogen terms.[53] The improved force fields have led to successful implementations of MD simulations on some RNA-containing systems, including conformational dynamics of folded RNAs[54−60] and biological catalysis.[61−63] Concerning the second issue, naturally occurring modified ribonucleotides' parameters have now been developed and are consistent with commonly used force fields.[46,47] The last issue may be most critical for the successful simulation of a protein−RNA complex. For balancing protein−RNA interactions, one strategy is to modify pair-specific corrections to nonbonded interactions by targeting related osmotic pressure.[64−67] Besides, polarizable force fields may capture essential physical effects for protein−RNA interactions, but their developments are not fully mature yet.[68−70]

Here, we use the YTHDC1 protein as a model system to investigate the binding of the preferred RNA motif to m^{6}A reader proteins by protein crystallography, isothermal titration calorimetry (ITC), and explicit solvent MD simulations with a force field suitable for protein−RNA complexes.

MATERIALS AND METHODS

Protein Purification. The plasmid expressing the N-terminally hexahistidine-tagged YTH domain (residues 345−509) of the human YTHDC1 protein was obtained as a gift from Cheryl Arrowsmith (Addgene ID: 64652). The recombinant protein was purified to homogeneity in two chromatographic steps. The protein was overexpressed for 16 h at 24 °C in *Escherichia coli* BL21 (DE3) cells upon induction with 0.4 mM IPTG. The cells were harvested and resuspended in the lysis buffer containing 100 mM Tris-HCl at pH 8.0, 500 mM NaCl, and 10 mM imidazole. The cells were lysed by sonication, and the cell lysate was clarified by centrifugation at 48,000*g* for 1 h and loaded onto a Ni-NTA affinity column (5 mL HisTrap FF from GE Healthcare). After extensive washing with the wash buffer containing 100 mM Tris-HCl at pH 8.0, 500 mM NaCl, and 50 mM imidazole, the target protein was eluted with the elution buffer containing 100 mM Tris-HCl at pH 8.0, 500 mM NaCl, and 250 mM imidazole. The N-terminal hexahistidine tag was removed by cleavage with tobacco etch virus (TEV) protease at a 1:50 ratio. The excess imidazole was removed by overnight dialysis, and the sample was subjected to a secondary subtractive Ni-NTA affinity chromatography step to remove the protease and uncleaved protein. Finally, the protein was subjected to a gel filtration step using a Superdex 75 16/60 column in a buffer containing 10 mM Tris-HCl at pH 7.5, 150 mM NaCl, and 1 mM DTT. The protein was concentrated to 10 mg/mL, flash-frozen in liquid nitrogen, and stored at −80 °C for future experiments.

Crystallography. The crystals of the YTHDC1 YTH domain were obtained by mixing 1 μL of the protein solution at 10 mg/mL with the mother liquor containing 0.1 M Bis-Tris at pH 6.5, 0.2 M ammonium sulfate, and 25% PEG 3350 at 22 °C in a hanging drop vapor diffusion setup. To obtain crystals of the protein complexed with oligoribonucleotides (custom-synthesized and HPLC-purified by Dharmacon, now Horizon Discovery), the crystals were transferred to a 1 μL drop containing 10 mM oligoribonucleotide directly dissolved in 0.1 M Bis-Tris at pH 6.5, 0.2 M ammonium sulfate, and 30% PEG 3350, soaked overnight at 22 °C, harvested, and frozen in liquid nitrogen without additional cryoprotection.

Diffraction data were collected at the Swiss Light Source (Villigen, Switzerland) using the beamline X06DA (PXIII) and processed using XDS.[71] The structures were solved by molecular replacement using the Phaser program[72] from the Phenix package.[73] The unliganded structure of YTHDC1 (PDB ID: 4R3H) was used as a search model. The model building and refinements were performed using COOT[74] and phenix.refine.[73] Data collection and refinement statistics are summarized in Table S1.

Coordinates. The atomic coordinates and structure factors for the four determined complexes have been deposited in the Protein Data Bank (PDB) with the accession codes 6RT4, 6RT5, 6RT6, and 6RT7.

Binding Assays. Isothermal titration calorimetry (ITC) experiments were carried out at 18 °C using MicroCal ITC200 (GE Healthcare). For the ITC experiments, the final gel filtration step of the purification of the YTH domain of YTHDC1 was performed in the ITC buffer (20 mM HEPES, pH 7.5, and 150 mM NaCl). Oligoribonucleotides were directly dissolved at 500−800 μM in the ITC buffer and titrated into the sample cell containing the protein at 50−60 μM. After an initial injection of 0.4 μL, 19 injections of 2.0 μL each were performed. The raw data were integrated, normalized for concentration, and analyzed using a single-binding site model, provided in the MicroCal Origin software package.

Molecular Dynamics Simulations. All the systems for MD simulations in this study were constructed based on the structure of the YTH domain of YTHDC1 complexed with a modified pentaribonucleotide reported previously by others (PDB ID: 4R3I).[33] The protein YTHDC1, the RNA oligomer GG(m^{6}A)CU, and crystal water molecules were kept as in the initial structure. Based on this structure, six different systems were generated, that is, three complex systems and three structures of unbound modified oligoribonucleotides. For the three complex systems, the protein coordinates were kept unchanged, while the oligoribonucleotides were truncated to obtain the tetrameric G(m^{6}A)CU and GG(m^{6}A)C from the

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

pentameric GG(m^{6}A)CU whose coordinates were available in the crystal structure. The hydrogen atoms were added by the CHARMM program,[75] and the protonation states were determined at neutral pH conditions. The complex systems were solvated in a 67 Å rhombic dodecahedron (RHDO) TIP3P water box[76] to ensure a 10 Å buffer space between the macromolecular atoms and the boundary of the water box. To neutralize the system and mimic the physiological conditions, Na^+ and Cl^- ions at a 0.15 M concentration were added to the solvated systems. Similarly, the three free oligoribonucleotides were solvated in a 47 Å RHDO water box with 0.15 M NaCl. To explore the role of the m^{6}A's methyl group in the regulation of RNA's conformations, the unmethylated pentamer GGACU was also simulated in its free state.

Each simulation system was initially minimized for 10,000 steps under a series of restraints and constraints on the solute atoms to release its bad contacts and poor geometry. The minimized structure was heated to 300 K and equilibrated in an *NVT* condition (constant volume and temperature). Finally, the structure was further equilibrated in an *NPT* condition (constant pressure and temperature). All the equilibration phases lasted for 1 ns using the CHARMM program (version 42b2).[75] Production runs of 500 ns each were carried out in *NPT* conditions using the NAMD program (version 2.12).[77] The pressure was controlled by the Nosé−Hoover Langevin piston method with a 200 ps piston period and 100 ps piston decay time.[78,79] The temperature was maintained at 300 K using the Langevin thermostat with a 5 ps friction coefficient. The integration time step was set to 2 fs by constraining all the bonds involving hydrogen atoms by the SHAKE algorithm. van der Waals energies were calculated using a switching function with a switching distance from 9 to 11 Å,[80] and electrostatic interactions were evaluated using the particle mesh Ewald summation (PME) method.[81]

The CHARMM36 force field was used for protein[82] and RNA[49] molecules. For the parameterization of m^{6}A, the force field for naturally occurring modified ribonucleotides was used.[46] Five independent runs with random initial velocities were carried out for each system for a total of 2.5 μs. MD snapshots were saved every 20 ps along the MD trajectories for further analysis. Geometric measurements, for example, distance, dihedrals, root mean square deviation (RMSD), and contact map analysis, were performed with CHARMM routines. All statistical figures were plotted by MATLAB (version 2018a MathWorks, Inc.), and structural figures were generated with PyMOL graphic software (version 2.2 Schrödinger, LLC).

RESULTS

Flexible Binding of the $G_{-2}G_{-1}$ Segment of the GG(m^{6}A)CU Ligand to the YTH Domain. The crystal structure of the complex between the pentamer GG(m^{6}A)CU and YTHDC1 YTH domain (PDB ID: 4R3I) shows that m^{6}A is fully buried in an aromatic cage consisting of Trp377 and Trp428 (Figure 1).[33] On the other hand, the contribution to binding of the nonmodified nucleotides is less clear. The two guanosines of GG(m^{6}A)CU are visible in the structure 4R3I, interacting with the surface residues of YTHDC1 (Figure 1A). The guanine bases of G_{-1} and G_{-2} form hydrogen bonds (HB) with the backbone of Val382 and Asp476, respectively (Figure 2A). Furthermore, there is a HB between the two guanines.[33] To test the stability of these interactions, we conducted MD simulations on the complex system of YTHDC1 with

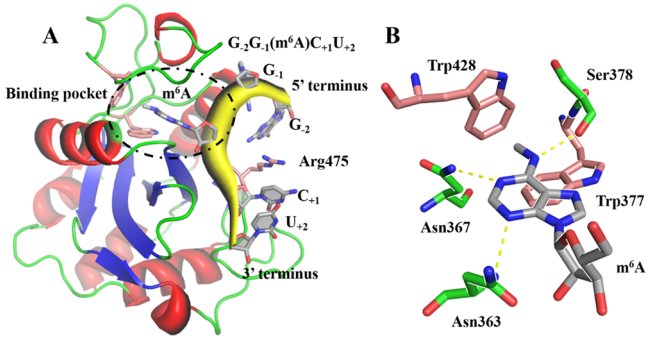


Figure 1. Binding mode of the RNA motif GG(m^{6}A)CU to the YTHDC1 reader domain (PDB ID: 4R3I). (A) The backbones of RNA (yellow) and YTHDC1 (red, blue, and green for α helices, β strands, and loops, respectively) are shown in cartoons, and the two tryptophans of the aromatic cage are shown in sticks. The binding pocket of m^{6}A is highlighted (black dashed circle). (B) Detailed interactions between m^{6}A (carbon atoms in gray) and the pocket of YTHDC1 (tryptophan and polar residues in salmon and green, respectively). The importance of the two tryptophans for the binding has been confirmed by site-directed mutagenesis experiments.[33]

GG(m^{6}A)CU. The interaction analysis shows that the HB interactions are not well maintained during the simulations with a large deviation from the distances in the crystal structure (Figure S1).

We checked whether crystal packing plays a role in stabilizing the HB interactions between the nucleotides $G_{-1}G_{-2}$ and YTHDC1. In particular, several interactions are found at the interface between molecules related by a crystallographic symmetry (Figure 2B). For example, toward the oligoribonucleotide's 5′ end, the sugar group of G_{-2} interacts with the side chain of Glu405 of the monomer related by translation and the 2_1 screw axis. Similarly, the phosphate linker of G_{-1} in the gray monomer interacts with the sugar group of U_{+2} in the blue monomer and is further stabilized by a salt bridge contributed by Lys472. However, C_{+1} does not show similar contacts between any image monomers. Such contacts of oligoribonucleotides from two YTHDC1-RNA complexes are unlikely populated in an aqueous environment, and therefore they are probably a consequence of crystal packing. Similar packing modes are also found in other structures of reader proteins (Figure S2).

We hypothesized that truncated variants of GG(m^{6}A)CU, that is, without terminal nucleotides, would not permit the crystal contacts mentioned above. To test this, four different RNAs were synthesized for crystallographic study, that is, GG(m^{6}A)C, G(m^{6}A)CU, G(m^{6}A)C, and (m^{6}A)CU. We solved the crystal structures of the YTH domain of YTHDC1 in the complex with these four RNA oligonucleotides at resolutions of 1.5, 1.7, 2.3, and 1.5 Å, respectively (Table S1). The new structures mostly agree well with the structure of the complex with the pentaribonucleotide ligand (Figure 2C). However, the coordinates of the two guanine nucleotides, including bases, sugars, and phosphates, at positions −2 and −1 with respect to m^{6}A cannot be modeled in our structures due to missing electron densities. By contrast, the cytosine and uridine at positions +1 and +2, respectively, show well-defined electron densities. Thus, our co-crystal structures suggest that the two guanine ribonucleotides upstream of m^{6}A do not form stable interactions with the protein, while the downstream motif CU forms conserved

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

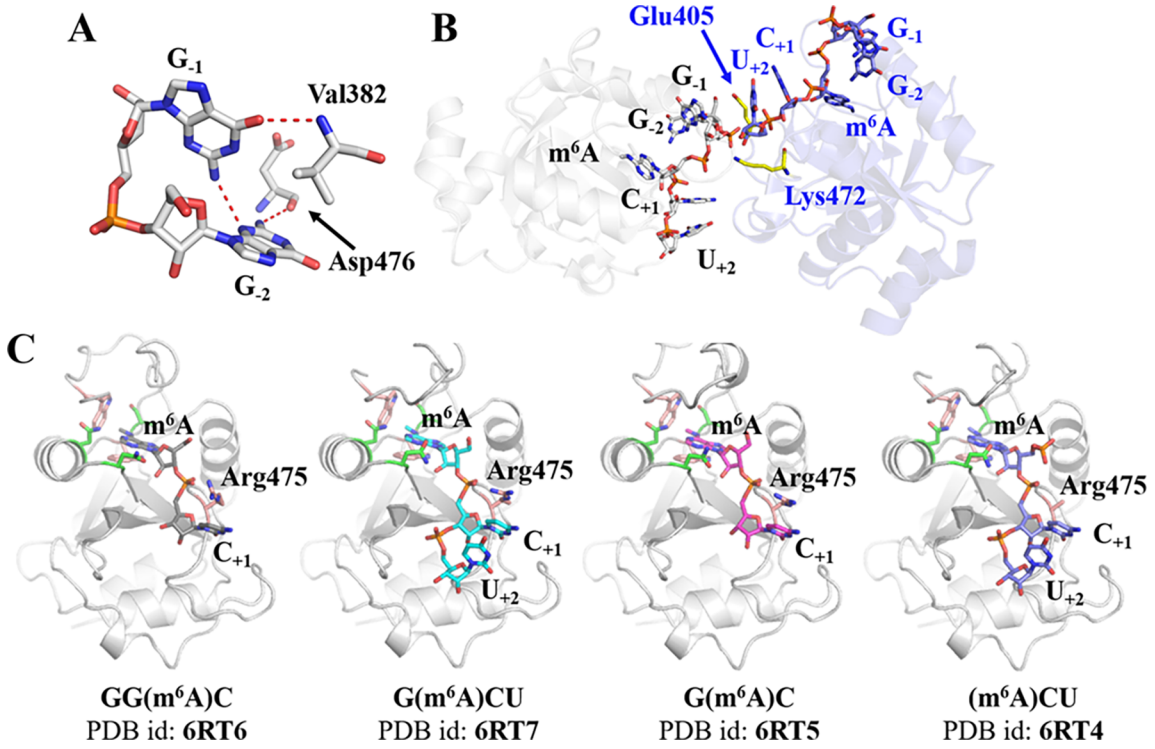


Figure 2. Interactions stabilized by crystal packing. (A) Hydrogen bonds between nucleotides (G-1 and G-2) and YTHDC1. The interaction mode was extracted from the structure 4R3I. The plausible hydrogen bonds are indicated in red dotted lines. (B) Two m^{6}A-containing RNA oligomer-YTHDC1 complexes related by crystallographic symmetry, shown and labeled in gray and blue for clarity. Their RNA structures are presented in sticks. The protein residues that contribute the crystal packing interactions are colored in yellow. (C) New complex structures solved in the present study. The proteins of the complex structures are shown in a gray cartoon, and the bound RNAs are shown in sticks with different colors. The residues in the m^{6}A binding site are shown in sticks (same colors as in Figure 1B).

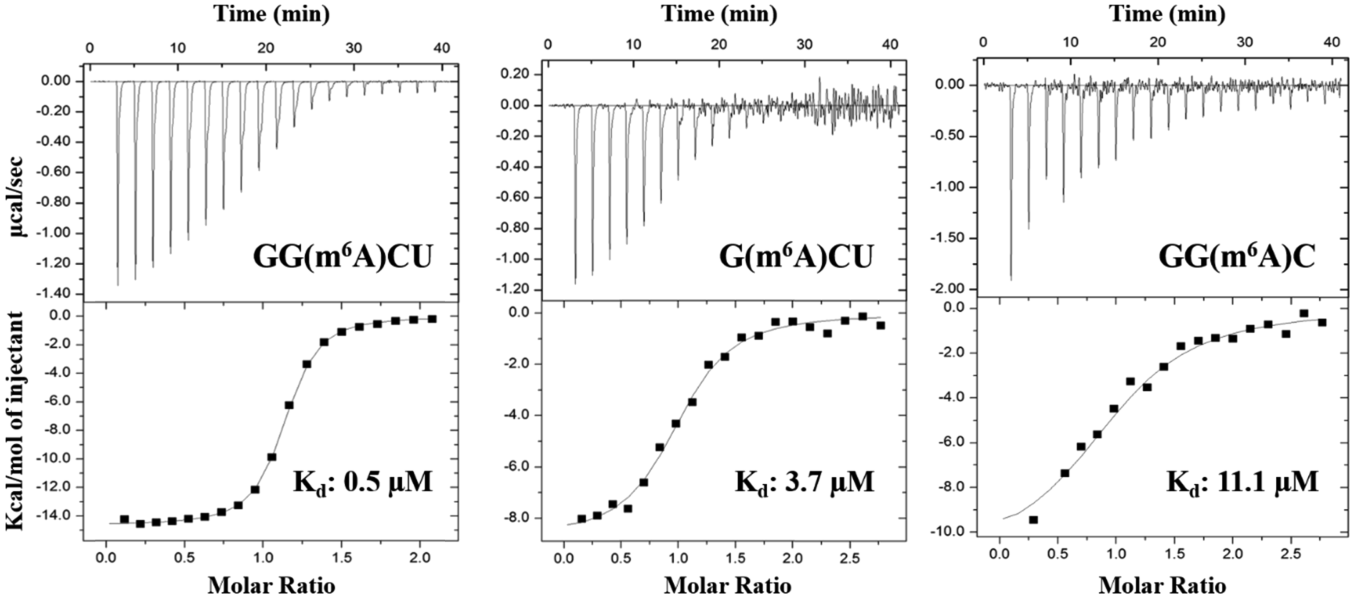


Figure 3. Isothermal titration calorimetry measurements of the binding of the RNA oligomers GG(m^{6}A)CU, G(m^{6}A)CU, and GG(m^{6}A)C to YTHDC1. The raw ITC curves and binding isotherms with the best fit are shown in the top and bottom panels, respectively. The truncation of one nucleotide at either the 5′- or 3′-terminus of GG(m^{6}A)CU deteriorates the affinity.

interactions with the protein. These findings are consistent with our hypothesis on the stabilization of the untruncated GG(m^{6}A)CU due to crystal packing in the YTHDC1 structure (PDB id: 4R3I).

5′- and 3′-Terminal Nucleotides Contribute to the Affinity of GG(m^{6}A)CU to YTHDC1. The four crystal structures have provided evidence that the $G_{-2}G_{-1}$ dinucleotide segment does not form stable interactions with YTHDC1, while the segment $C_{+1}U_{+2}$ shows conserved contacts with the protein. However, it is not clear if and how much these terminal nucleotides contribute to the binding of GG(m^{6}A)-CU to YTHDC1. To address this question, we measured the

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

dissociation constant (K_d) values of three representative oligomers, namely, GG(m^{6}A)CU, G(m^{6}A)CU, and GG(m^{6}A)C, using isothermal titration calorimetry (ITC) (Figure 3 and Table S2). We obtained a K_d value of 0.5 μM for GG(m^{6}A)CU, which is similar to a value reported previously (2.0 μM).[33] The U_{+2}-deficient tetramer, that is, GG(m^{6}A)C, is less potent by a factor of approximately 20 than the pentamer GG(m^{6}A)CU. Such a binding reduction can be intuitively understood from the crystal structures, namely, U_{+2} forms $\pi-\pi$ stacking with C_{+1} and thus stabilizes interactions of the 3′ end with YTHDC1. Surprisingly, the G_{-2}-deficient oligomer, that is, G(m^{6}A)CU, also shows significant reduction of affinity (~7 times compared to the pentamer) although G_{-2} does not form stable interactions with YTHDC1 in our crystal structures and MD simulations (Figure 2C and Figure S1). Thus, further investigation is needed to understand how G_{-2} impacts the binding affinity (see below).

MD Simulations of Liganded YTHDC1 Reveal Flexibility of the Nucleotides Upstream of m^{6}A. Three complexes with ligands GG(m^{6}A)CU, G(m^{6}A)CU, and GG(m^{6}A)C were investigated by MD simulations to analyze their relative flexibility. In all simulations, the two guanine nucleotides upstream of m^{6}A show large deviations compared to their crystal poses (Figure 4 and Figure S3). This result is consistent with the lack of electron density for this RNA segment. During the MD simulations, the two guanine nucleotides do not show stable and specific interactions with YTHDC1 in contrast with the crystal structure (PDB code: 4R3I) (Figure 2A). Instead, the two nucleotides are dynamically coupled for a long time by the $\pi-\pi$ interaction between their bases. This is illustrated by the contact map analysis in which G_{-1} and G_{-2} interact with each other in approximately half of the samplings (lower triangles of Figure 5).

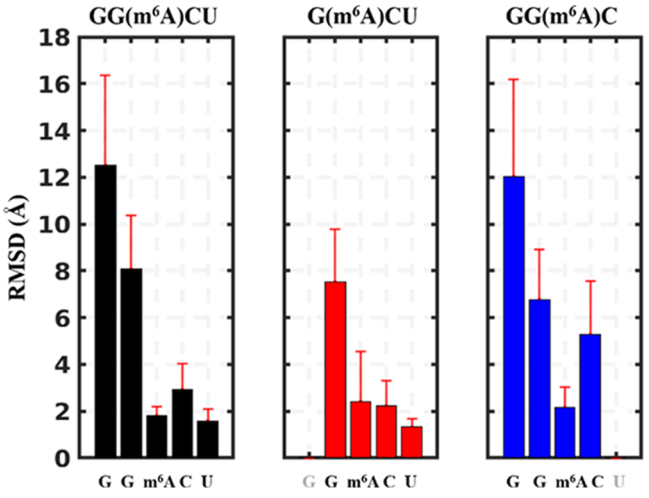


Figure 4. RMSD profile of RNA oligomers in the bound state. The plots show average RMSD values and standard deviations calculated over the cumulative sampling of five trajectories for each simulation system. To calculate the RMSD value, heavy atoms of nucleotides in their crystal structure (PDB ID: 4R3I) were used as reference. All snapshots from MD simulations were superimposed to the protein backbone of the crystal structure excluding C- and N-terminal segments.

The tetraribonucleotide ligands allow us to assess the role of the terminal G_{-2} and U_{+2}; without which, m^{6}A-containing RNA oligomers show a reduced binding affinity compared to the pentamer. Although the nucleotide G_{-2} was flexible during the simulations of the complex, it helps in stabilizing the pose of m^{6}A in the binding pocket of YTHDC1. Without the terminal guanine nucleotide at position −2, m^{6}A experiences slightly larger deviations and fluctuations than in G_{-2}-containing ligands (Figure 4 and Figure S4). The larger RMSD values are due to the partial dissociation of m^{6}A from the YTHDC1 binding pocket in two of the five trajectories of G(m^{6}A)CU, which can be monitored by three key hydrogen bonds between m^{6}A and YTHDC1 (Figure S4). Noteworthy, if the dissociated snapshots are neglected in the system G(m^{6}A)CU, the RMSD value of m^{6}A is nearly unchanged compared to other two systems (Figure S5). Therefore, the reduced affinity of the G_{-2}-deficient ligand may not be plainly explained based on simulation results of the bound states. Other factors causing the affinity reduction will be investigated in the later sections.

Concerning the nucleotides downstream of m^{6}A, our crystal structures and MD simulations provide evidence that C_{+1} and U_{+2} form more stable interactions with YTHDC1 than G_{-1} and G_{-2}. The downstream C_{+1} establishes cation$-\pi$ and $\pi-\pi$ interactions with Arg475 and U_{+2}, respectively. A previous study has confirmed the importance of the cation$-\pi$ interaction for the molecular recognition between YTHDC1 and GG(m^{6}A)CU by mutating Arg475 to phenylalanine or alanine with 10-fold and 100-fold reduction, respectively, compared to the wild type.[33] However, the role of U_{+2} has not be confirmed as it is stabilized by the crystal packing in the structure 4R3I (Figure 2B). The simulations for the tetrameric ligand GG(m^{6}A)C show that the C_{+1}−Arg475 interaction becomes less stable than in the case of the ligand with U_{+2}. Consistently, the RMSD of C_{+1} becomes much larger and fluctuates more than that of the U_{+2}-containing oligomers (blue bars in Figure 4). This observation indicates that U_{+2} indeed stabilizes the interaction between C_{+1} and Arg475. As shown in Figure S6, the contact distance for C_{+1} and Arg475 is locally distributed at 3.8 Å for U_{+2}-containing ligands (black and red lines in Figure S6), while it becomes loosely distributed in the U_{+2}-deficient ligand (the blue line). These structural and dynamic analyses rationalize the role of U_{+2} and the reduced binding of the tetrameric ligand GG(m^{6}A)C (shown by ITC data).

The $C_{+1}U_{+2}$ segment specifically interact with a positively charged shallow pocket on the protein surface and may play a comparably important role in the protein−RNA recognition compared to m^{6}A. An equivalent mutation to Arg475Ala was also reported for a homologous protein YTHDF2, that is, Arg527Ala. This mutation showed approximately 25-fold reduction of the binding affinity for its RNA oligoribonucleotides compared to the wild type.[32] Moreover, the $C_{+1}U_{+2}$ segment binds to a well-defined pocket on the YTHDC1 surface, consisting of four positively charged residues, that is, Lys361, Arg404, Lys472, and Arg475 (Figure 1A). The binding data and structural analysis suggest that the interaction between the $C_{+1}U_{+2}$ segment and the reader protein is specific rather than an undirected charge−charge interaction. In addition, Zhu et al. suggested that a two-step model fits the data from a fluorescence polarization assay better than the one-step model for the binding of YTHDF2 to an m^{6}A-containing oligoribonucleotide.[32] They proposed that the binding of m^{6}A-flanking nucleotides facilitates the recognition of m^{6}A. Integrating this two-step model with our results from ITC

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

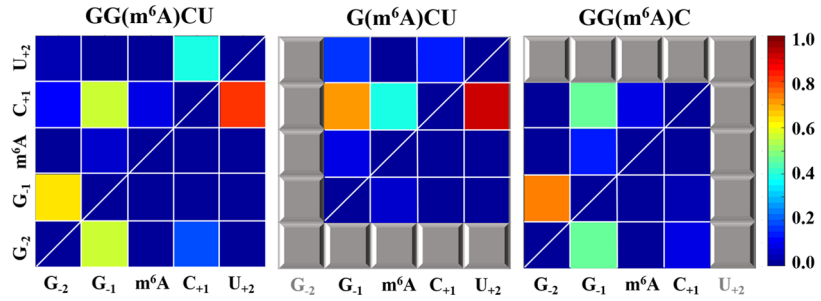


Figure 5. Contact maps of m^{6}A-containing RNA oligomers in their free state (upper, left triangle) and in the complex with YTHDC1 (lower, right triangle). Pairwise contact frequencies are evaluated by the distance of the center of mass between two bases and a cutoff value of 5 Å. The color bar represents the frequency of contact between two nucleotides, ranging from dark blue (no interaction) to dark red (persistent interaction). The nucleotide pairs, which are not included in the systems, are presented in gray squares.

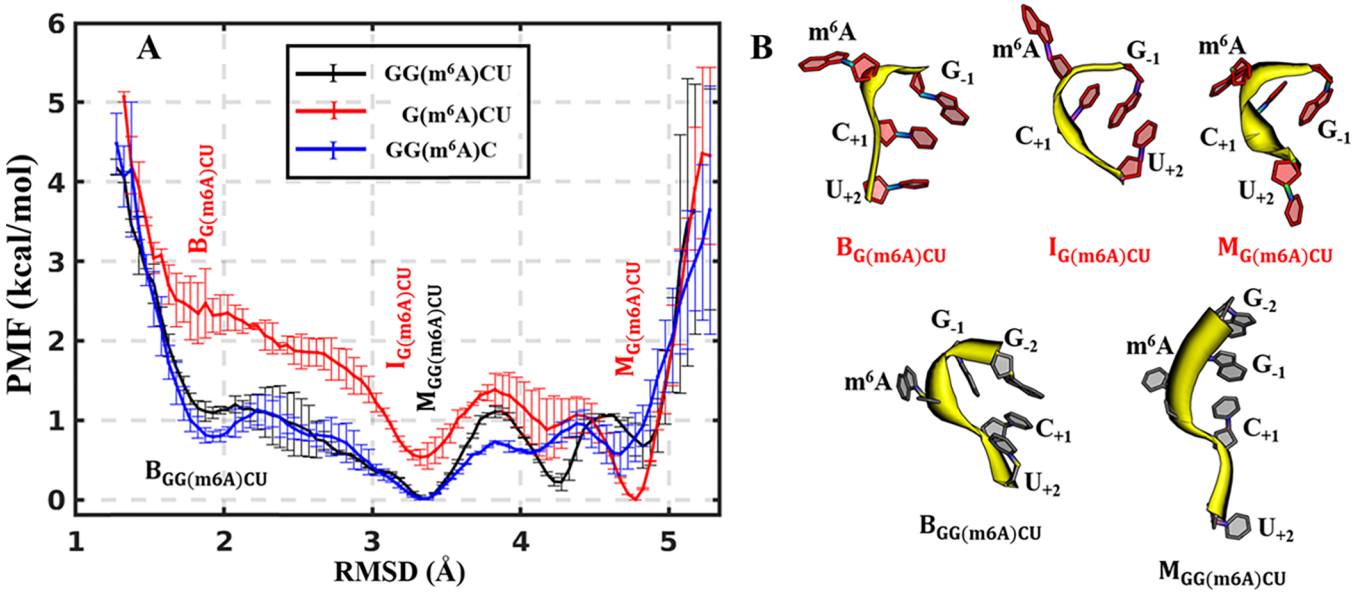


Figure 6. Conformational landscape of m^{6}A-containing RNA oligomers in the unbound state. (A) Potential of mean force (PMF) along the RMSD from the structure in the bound state. The RNA segment (m^{6}A)C in the complex with YTHDC1 (PDB code: 4R3I) is selected as the reference for the bound state, and its heavy atoms are included in the RMSD calculation. The 5′ GG segment is neglected because of its flexibility in the bound state. All MD snapshots of each simulated oligonucleotide are superposed to the (m^{6}A)C reference segment, and the corresponding RMSD values are calculated. The PMF is calculated as $-k_BT \ln \frac{p_i}{p_0}$ where k_B is Boltzmann factor, T is temperature (300 K), p_i is the population in bin i, and p_0 is the probability of the most populated bin (with a 0.05 Å bin width). To estimate the statistical error, the five trajectories for each system were divided into two blocks, namely, 2.5 trajectories for each block. The standard deviation of the two blocks is shown (error bars). The labels M, I, and B indicate mostly populated, intermediate, and bound-like conformational states, respectively. (B) Representative structures of different conformational states of free G(m^{6}A)CU (top) and GG(m^{6}A)CU (bottom).

and MD simulations, we infer that the $C_{+1}U_{+2}$ segment plays an essential role in the association process of GG(m^{6}A)CU to DC1 comparable to that of m^{6}A. However, it is still not clear which is the first motif to be recognized by the reader protein, m^{6}A or $C_{+1}U_{+2}$ segment, and this question requires further studies.

Conformational Heterogeneity of m^{6}A-Containing RNAs in Bound and Free States. To rationalize relative binding affinities, we analyzed conformations of three RNA oligomers in their bound and free states. Significant differences are found in the intramolecular contact pattern between the bound and free states (Figure 5). The bound RNAs show reduced contacts between nucleotides compared to their free states. For example, three pairs of nucleotides in GG(m^{6}A)-CU's free state are in contact for more than 40% of all snapshots. Its bound state, by contrast, loses one stable interaction pair, that is, the G_{-1}−C_{+1} interaction pair (Figure 5). That is because the centrally located m^{6}A of the bound state firmly anchors in the aromatic cage of YTHDC1. Thus, the pentameric GG(m^{6}A)CU is divided into two independent segments, that is, G_{-1}−G_{-2} and C_{+1}−U_{+2}, separated by m^{6}A. Two nucleotides belonging to the same segment are frequently coupled, but nucleotides from different segments interact much less (lower triangles in Figure 5).

The purine base of m^{6}A fully interacts with the aromatic cage of YTHDC1. Therefore, it is reasonable to expect that m^{6}A should be sufficiently exposed to the solvent before the association with the protein so that its binding pocket can effectively recognize m^{6}A. The contact maps show that m^{6}A in the G_{-2}-containing ligands, that is, GG(m^{6}A)CU and GG-(m^{6}A)C, is not in frequent contact with other nucleotides in the unbound state. By contrast, m^{6}A in G(m^{6}A)CU does interact with C_{+1} (upper triangles in Figure 5). The distribution of the m^{6}A−C_{+1} distance for G(m^{6}A)CU is

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

more populated for values approximately 5 Å than for the other two ligands (Figure S7).

To detect how the exposure of m^6A varies in different ligands, we measured the solvent accessible surface area (SASA) of m^6A in the free state (Figure S8A). The two G_{-2}-containing ligands show a similar distribution of the SASA, which is mostly populated around 180 $Å^2$ with a second peak at 200 $Å^2$. By contrast, G(m^6A)CU populates its main peak at 150 $Å^2$. Such a difference in the SASA is caused by the frequent contact between m^6A and C_{+1} (the upper triangle for G(m^6A)CU in Figure 5). Taken together, the G_{-2} nucleotide plays a role in the exposure of m^6A from the RNAs in the unbound state and thereby benefits the recognition of m^6A by its reader proteins. To study if the methyl group of m^6A facilitates the base exposure, we also compared the oligomers GG(m^6A)CU and GGACU in their unbound states. Again, a significant difference was found in the distribution of the SASA with the unmethylated oligomer showing a less exposed adenine than its N6-methylated counterpart (Figure S8B).

Conformational Free-Energy Profiles of m^6A-Containing RNAs in the Unbound State. Multiple MD simulations of the unbound state of GG(m^6A)CU, G(m^6A)CU, and GG(m^6A)C were carried out to analyze potential differences in the most populated conformers. The simulations were used to construct free-energy profiles with the RMSD of (m^6A)C relative to its crystal pose as the progress variable. The free-energy profiles show that oligomers in aqueous solutions adopt multiple meta-stable conformations, which appear as local minima of the free-energy profiles (Figure 6). The ligands GG(m^6A)CU and GG(m^6A)C show a similar conformational propensity. There is approximately a 1 kcal/mol free-energy difference between their global minimum and the bound-like local minimum (black and blue lines in Figure 6). By contrast, G(m^6A)CU experiences an ∼2 kcal/mol of energy penalty from the unbound to the bound state. The conformational free-energy difference indicates that the former two ligands have less energy penalty for redistributing their conformations to the bound-state than G(m^6A)CU. It is useful to introduce a two-state approximation in which all conformations of an RNA oligomer are divided into either bound or unbound states. The two-state model shows again a significant energy difference between G_{-2}-containing oligomers and G(m^6A)CU irrespective of the RMSD threshold (Figure S9). A similar difference was also found by comparing the PMF profiles of unbound GG(m^6A)CU and GGACU (Figure S8C), indicating that the methyl group of m^6A helps in adopting the bound conformation.

The analysis on meta-stable conformations suggests an important role of G_{-2} in maintaining the bound-like conformations and exposing m^6A to bulk water. For example, for the oligomer GG(m^6A)CU, the global minimum conformation along the free-energy profile not only maintains a bound-like pose but also exposes m^6A to the solvent ($M_{GG(m6A)CU}$ in Figure 6B). In this conformation, G_{-2} interacts with G_{-1} in a $\pi-\pi$ interaction and thereby frees m^6A to expose it to the bulk water. In the most stable conformation of G(m^6A)CU, G_{-1}, m^6A, and C_{+1} stack together, and thus m^6A is less exposed to the solvent (the upper panel in Figure 6B). Although the conformational change from $M_{G(m6A)CU}$ to $I_{G(m6A)CU}$ increases the solvent exposure of m^6A, C_{+1} and G_{-1} become packed with each other, resulting in a detrimental pose for its binding to YTHDC1. Finally, to transform this intermediate state to the bound-like one, the ligand G(m^6A)CU has to pay a higher free-energy cost compared to GG(m^6A)CU. Thus, G_{-2} is a key nucleotide for modulating the conformational preference of m^6A-containing oligoribonucleotides in the unbound state.

DISCUSSION

Crystal Packing Artificially Stabilizes the Poses of m^6A-Flanking Oligoribonucleotides Bound to Reader Proteins. The crystal packing is likely to stabilize the binding mode observed in the crystal. For testing this hypothesis, we soaked varied m^6A-containing oligoribonucleotides into apo crystals of YTHDC1. Its space group does not permit artificial intramolecular contacts of oligoribonucleotides from different asymmetric units but still leave enough space for their natural interactions with the protein in the same unit. As a result, the $G_{-2}G_{-1}$ segment is disordered in all soaking structures, indicating weak interactions between them and the protein surface. The MD simulations further support this observation as the $G_{-2}G_{-1}$ segment significantly deviates from the crystal pose and shows substantial flexibility (Figure 4 and Figure S1). Our evidence suggests a dynamical model for the $G_{-2}G_{-1}$ segment binding to YTHDC1 against the static model, which seems to emerge from the previously published structure (4R3I).[33]

Structural artifacts due to crystal packing are observed not only for YTHDC1, but also for other reader domains, leading to well-defined binding poses of m^6A-flanking nucleosides, which otherwise are flexible in solution. We found in the PDB Data Bank another two m^6A reader protein structures that are complexed with m^6A-containing nucleotides, and they show multiple packing contacts between their m^6A-flanking nucleotides and the protein surface (Figure S2). For example, the $G_{-2}G_{-1}$ segment in the structure 4RCJ (YTHDF1) interacts with the protein residue Ser461 and Ala462 of another asymmetric unit by multiple hydrogen bonds. Moreover, in the structure of the YTH domain of the MRB1 protein from Z. roux (PDB code: 4U8T), the m^6A-flanking nucleotides from different protein copies are even tangled, resulting in substantial artificial contacts.[38] In addition, in an in-house YTHDF3 structure (to be published), the $G_{-2}G_{-1}$ segment is also stabilized by its own copies from another asymmetric unit.

Conformational Analysis of Oligoribonucleotides in Their Unbound State Is Essential for Understanding the Recognition Mechanism. The importance of the terminal nucleotide G_{-2} for the affinity is supported by our ITC measurements, which show a 7-fold reduction of the G(m^6A)CU motif with respect to the pentameric GG(m^6A)CU (Figure 3 and Table S2). The attenuated binding affinity is not due to the missing interactions between the 5′ terminus and YTHDC1's protein surface. It is rationalized by MD simulations of different oligoribonucleotides in aqueous solutions. The corresponding free energy landscape analysis demonstrates that the 5′-terminal nucleotide G_{-2} contributes to the binding by facilitating the adoption of bound-like conformations of the RNA oligomers (Figure 6), thereby reducing the free energy expense for the association process. These results show that information from both complex structures and conformations of oligoribonucleotides in aqueous solutions are indispensable to study the recognition mechanisms of reader proteins and their cognate RNA binders. This is because an oligoribonucleotide ligand has much more degrees of freedom compared to a normal-size small-molecule ligand in aqueous solutions.

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

The simulation analysis suggests a conformational selection model.[83] Those oligoribonucleotide motifs that mainly populate the bound state in aqueous solution will pay less penalty for the conformational change than those predominating bound-resistant conformations. This model explains why GG(m^{6}A)CU binds stronger than G(m^{6}A)CU, thus elucidating the role of 5′ terminus G_{-2} in the binding. Our analysis also shows that m^{6}A is mostly solvent exposed in the bound-like conformations (Figure S8), which is prerequisite for recognition by the aromatic cage. In fact, the adenine also tends to be more solvent-exposed in the methylated GGACU than in the unmethylated counterpart (Figure S8B). This suggests that the methyl group not only contributes the interactions of oligoribonucleotides with reader proteins but also facilities the nucleotides adopting a bound-like conformation. The capacity of the methyl group for conformational regulation is not only limited to explain the recognition mechanism for reader proteins. The exposed conformation of methylated adenine might be also required to accommodate it in the active site of m^{6}A-specific demethylase FTO.[84]

Differential and Shared Recognition Mechanisms between YTHDC1 and Other Reader Proteins. The five human m^{6}A reader proteins can be classified into two classes according to their binding pocket features with YTHDC1 and YTHDC2 in one class and YTHDF1, YTHDF2, and YTHDF3 in the other (Figure S10). Here, we select YTHDC1 and YTHDF1 as representatives in the two classes. The first difference is in the side chain that interacts with the N1 position of m^{6}A (Figure S10). Asn367 of YTHDC1 interacts with the N1 atom by donating a hydrogen atom, while the corresponding side chain in YTHDF1 is Asp401. The interaction between the N1 atom and the asparagine side chain is more favorable than that with aspartic acid.[31] Second, the two reader proteins have different gatekeeper residues at the border of their binding pockets, that is, Asn363 and Tyr397 in YTHDC1 and YTHDF1, respectively. For the latter, the Tyr397Ala mutation abolishes its binding to GG(m^{6}A)CU.[31] The X-ray structure of holo YTHDF1 shows that Tyr397 forms $\pi-\pi$ stacking with the 5′-terminal nucleotides.[31] In contrast to YTHDC1, such a specific interaction probably restrains the 5′-terminus dynamics on YTHDF1.

The two classes of reader proteins share two structural features, that is, the aromatic cage for recognizing m^{6}A and a positively charged shallow pocket for interactions with the 3′-terminal nucleotides. The two conserved features suggest a common binding mechanism. First, the aromatic cage permits the selective recognition of m^{6}A. The indole side chain of one of the tryptophans provides $\pi-\pi$ stacking with the modified adenine, while the other indole is at an optimal van der Waals distance with the methyl group. Our simulations show that GG(m^{6}A)CU tends to expose its m^{6}A to the solvent in its free state more than its unmethylated counterpart (Figure S8B). The m^{6}A exposure provides a specific geometry for fitting the aromatic cage in all reader proteins, thereby distinguishing unmethylated RNAs. Second, the surface pocket adjacent to the aromatic cage conservatively contains two charged residues in all five reader proteins, that is, Arg475 and Lys361 for YTHDC1 (Figure S10). Their importance for the binding of m^{6}A-containing RNAs has been verified in several reader proteins.[31,33,38] The 3′-terminal nucleotides contribute to the protein−RNA binding by specific association to the positively charged pocket on the YTHDC1 surface.[33] The structural features suggest a two-step binding event in which the electrostatically driven association of the 3′-terminal nucleotides facilitates the further recognition and formation of optimal contacts between m^{6}A and the aromatic cage.[32] Such a mechanism might be common to the five human m^{6}A reader proteins as they share a positive electrostatic potential on the surface that recognizes the 3′-terminus nucleotides (Figure S11). The inverse sequence of events is less probable considering that electrostatic attraction has a longer range than van der Waals interactions.

CONCLUSIONS

We have investigated the molecular recognition between YTHDC1 and its RNA motif GG(m^{6}A)CU by atomistic simulations, X-ray crystallography, and isothermal titration calorimetry. The selected model oligoribonucleotide is biologically relevant because the sequence is preferentially recognized by all five human m^{6}A reader proteins.[33,34,84−86] To investigate the flexibility of terminal nucleotides, the structures of the YTH domain of YTHDC1 in a complex with GG(m^{6}A)C, G(m^{6}A)CU, G(m^{6}A)C, and (m^{6}A)CU were solved at high resolution. ITC binding experiments were utilized to examine the contribution of 5′- and 3′-terminal nucleotides to the affinity of GG(m^{6}A)CU for YTHDC1. MD simulations were carried out for bound and unbound states of three representative RNA ligands, that is, GG(m^{6}A)CU, G(m^{6}A)CU, and GG(m^{6}A)C.

The MD simulations of the bound states have revealed the flexibility of the 5′-terminal nucleotides and the stabilizing role of the cation$-\pi-\pi$ interactions by the 3′-terminal nucleotides. The unbound-state simulations suggest that the methyl group of m^{6}A and 5′-terminus GG motif shift the free-energy landscape of the m^{6}A-containing oligonucleotide into the bound-like state as observed in the structure of the complex. Taken together, the crystal structures, calorimetry data, and simulation results of bound and free oligoribonucleotides provide an atomistic explanation for the roles of m^{6}A-flanking ribonucleotides in the recognition by the reader domain of YTHDC1. It remains to be investigated if the binding mechanism is similar in other m^{6}A reader domains.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00987.

X-ray crystallographic statistical data, thermodynamic parameters for ITC measurements, trajectory analyses of MD simulations, more examples of artificial contacts caused by crystal packing, structural comparison between YTHDC1 and YTHDF1, and electrostatic potential surfaces of five m^{6}A-reader proteins (PDF)

Accession Codes

PDB IDs: 6RT4, 6RT5, 6RT6, and 6RT7.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: p.sledz@bioc.uzh.ch. Phone: +41 44 63 55587 (P.Ś.).
*E-mail: caflisch@bioc.uzh.ch. Phone: +41 44 635 5521. Fax: +41 44 635 6862 (A.C.).

ORCID

Yaozong Li: 0000-0002-5796-2644
Paweł Śledź: 0000-0002-4440-3253

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

Amedeo Caflisch: 0000-0002-2317-6792

Funding

This work was supported by the Swiss National Science Foundation (to A.C.), the International Postdoc Grant funded by the Swedish Research Council (VR 2019-00608 to Y.L), and the grant from the Innosuisse (Swiss Innovation Agency) as well as the Entrepreneur Fellowship in Biotechnology funded by University of Zurich (to P.Ś).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The use of beamlines and user support at the Swiss Light Source are gratefully acknowledged. We thank the Swiss National Supercomputing Centre (CSCS) in Lugano for providing the computational resources. We also thank Dr. Xiang Wang for the valuable discussion.

■ REFERENCES

(1) Dominissini, D.; Moshitch-Moshkovitz, S.; Salmon-Divon, M.; Amariglio, N.; Rechavi, G. Transcriptome-wide mapping of N-6-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. *Nat. Protoc.* **2013**, *8*, 176.

(2) Meyer, K. D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C. E.; Jaffrey, S. R. Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3 ' UTRs and near Stop Codons. *Cell* **2012**, *149*, 1635−1646.

(3) Xiao, W.; Adhikari, S.; Dahal, U.; Chen, Y. S.; Hao, Y. J.; Sun, B. F.; Sun, H. Y.; Li, A.; Ping, X. L.; Lai, W. Y.; Wang, X.; Ma, H. L.; Huang, C. M.; Yang, Y.; Huang, N.; Jiang, G. B.; Wang, H. L.; Zhou, Q.; Wang, X. J.; Zhao, Y. L.; Yang, Y. G. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol. Cell* **2016**, *61*, 507−519.

(4) Alarcon, C. R.; Lee, H.; Goodarzi, H.; Halberg, N.; Tavazoie, S. F. N6-methyladenosine marks primary microRNAs for processing. *Nature* **2015**, *519*, 482.

(5) Lin, S.; Choe, J.; Du, P.; Triboulet, R.; Gregory, R. I. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol. Cell* **2016**, *62*, 335−345.

(6) Wang, X.; Zhao, B. S.; Roundtree, I. A.; Lu, Z.; Han, D.; Ma, H.; Weng, X.; Chen, K.; Shi, H.; He, C. N-6-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **2015**, *161*, 1388−1399.

(7) Wang, X.; Lu, Z.; Gomez, A.; Hon, G. C.; Yue, Y.; Han, D.; Fu, Y.; Parisien, M.; Dai, Q.; Jia, G.; Ren, B.; Pan, T.; He, C. N-6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **2014**, *505*, 117−120.

(8) Qi, S. T.; Ma, J. Y.; Wang, Z. B.; Guo, L.; Hou, Y.; Sun, Q. Y. N-6-Methyladenosine Sequencing Highlights the Involvement of mRNA Methylation in Oocyte Meiotic Maturation and Embryo Development by Regulating Translation in Xenopus laevis. *J. Biol. Chem.* **2016**, *291*, 23020−23026.

(9) Geula, S.; Moshitch-Moshkovitz, S.; Dominissini, D.; Mansour, A. A.; Kol, N.; Salmon-Divon, M.; Hershkovitz, V.; Peer, E.; Mor, N.; Manor, Y. S.; Ben-Haim, M. S.; Eyal, E.; Yunger, S.; Pinto, Y.; Jaitin, D. A.; Viukov, S.; Rais, Y.; Krupalnik, V.; Chomsky, E.; Zerbib, M.; Maza, I.; Rechavi, Y.; Massarwa, R.; Hanna, S.; Amit, I.; Levanon, E. Y.; Amariglio, N.; Stern-Ginossar, N.; Novershtern, N.; Rechavi, G.; Hanna, J. H. m(6)A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science* **2015**, *347*, 1002−1006.

(10) Wang, Y.; Li, Y.; Toth, J. I.; Petroski, M. D.; Zhang, Z.; Zhao, J. C. N-6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* **2014**, *16*, 191−198.

(11) Batista, P. J.; Molinie, B.; Wang, J. K.; Qu, K.; Zhang, J. J.; Li, L. J.; Bouley, D. M.; Lujan, E.; Haddad, B.; Daneshvar, K.; Carter, A. C.; Flynn, R. A.; Zhou, C.; Lim, K. S.; Dedon, P.; Wernig, M.; Mullen, A. C.; Xing, Y.; Giallourakis, C. C.; Chang, H. Y. m(6)A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells. *Cell Stem Cell* **2014**, *15*, 707−719.

(12) Fustin, J. M.; Doi, M.; Yamaguchi, Y.; Hida, H.; Nishimura, S.; Yoshida, M.; Isagawa, T.; Morioka, M. S.; Kakeya, H.; Manabe, I.; Okamura, H. RNA-Methylation-Dependent RNA Processing Controls the Speed of the Circadian Clock. *Cell* **2013**, *155*, 793−806.

(13) Zhou, J.; Wan, J.; Gao, X. W.; Zhang, X. Q.; Jaffrey, S. R.; Qian, S. B. Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* **2015**, *526*, 591−594.

(14) Ben-Haim, M. S.; Moshitch-Moshkovitz, S.; Rechavi, G. FTO: linking m(6)A demethylation to adipogenesis. *Cell Res.* **2015**, *25*, 3.

(15) Peng, L.; Yuan, X.; Jiang, B.; Tang, Z.; Li, G. C. LncRNAs: key players and novel insights into cervical cancer. *Tumor Biol.* **2016**, *37*, 2779−2788.

(16) Satterlee, J. S.; Basanta-Sanchez, M.; Blanco, S.; Li, J. B.; Meyer, K.; Pollock, J.; Sadri-Vakili, G.; Rybak-Wolf, A. Novel RNA Modifications in the Nervous System: Form and Function. *J. Neurosci.* **2014**, *34*, 15170−15177.

(17) Delaunay, S.; Frye, M. RNA modifications regulating cell fate in cancer. *Nat. Cell Biol.* **2019**, *21*, 552−559.

(18) Dai, D.; Wang, H.; Zhu, L.; Jin, H.; Wang, X. N6-methyladenosine links RNA metabolism to cancer progression. *Cell Death Dis.* **2018**, *9*, 124.

(19) Boriack-Sjodin, P. A.; Ribich, S.; Copeland, R. A. RNA-modifying proteins as anticancer drug targets. *Nat. Rev. Drug Discovery* **2018**, *17*, 435−453.

(20) Hsu, P. J.; Shi, H. L.; He, C. Epitranscriptomic influences on development and disease. *Genome Biol.* **2017**, *18*, 197.

(21) Zheng, G. Q.; Dahl, J. A.; Niu, Y. M.; Fedorcsak, P.; Huang, C. M.; Li, C. J.; Vagbo, C. B.; Shi, Y.; Wang, W. L.; Song, S. H.; Lu, Z. K.; Bosmans, R. P. G.; Dai, Q.; Hao, Y. J.; Yang, X.; Zhao, W. M.; Tong, W. M.; Wang, X. J.; Bogdan, F.; Furu, K.; Fu, Y.; Jia, G. F.; Zhao, X.; Liu, J.; Krokan, H. E.; Klungland, A.; Yang, Y. G.; He, C. ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Mol. Cell* **2013**, *49*, 18−29.

(22) Jia, G.; Fu, Y.; Zhao, X.; Dai, Q.; Zheng, G.; Yang, Y.; Yi, C.; Lindahl, T.; Pan, T.; Yang, Y. G.; He, C. N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **2011**, *7*, 885−887.

(23) Wiedmer, L.; Eberle, S. A.; Bedi, R. K.; Śledź, P.; Caflisch, A. A Reader-Based Assay for m(6)A Writers and Erasers. *Anal. Chem.* **2019**, *91*, 3078−3084.

(24) Wang, X.; Feng, J.; Xue, Y.; Guan, Z. Y.; Zhang, D. L.; Liu, Z.; Gong, Z.; Wang, Q.; Huang, J. B.; Tang, C.; Zou, T. T.; Yin, P. Structural basis of N6-adenosine methylation by the METTL3-METTL14 complex. *Nature* **2016**, *534*, 575−578.

(25) Wang, P.; Doxtader, K. A.; Nam, Y. Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol. Cell* **2016**, *63*, 306−317.

(26) Sledz, P.; Jinek, M. Structural insights into the molecular mechanism of the m(6)A writer complex. *Elife* **2016**, *5*, No. e18434.

(27) Liu, J.; Yue, Y.; Han, D.; Wang, X.; Fu, Y.; Zhang, L.; Jia, G.; Yu, M.; Lu, Z.; Deng, X.; Dai, Q.; Chen, W.; He, C. A METTL3-METTL14 complex mediates mammalian nuclear RNA N-6-adenosine methylation. *Nat. Chem. Biol.* **2014**, *10*, 93−95.

(28) Zhang, X.; Jia, G. F. RNA epigenetic modification: N6-methyladenosine. *Yi Chuan* **2016**, *38*, 275−288.

(29) Liu, N.; Pan, T. N-6-methyladenosine-encoded epitranscriptomics. *Nat. Struct. Mol. Biol.* **2016**, *23*, 98−102.

(30) Meyer, K. D.; Jaffrey, S. R. The dynamic epitranscriptome: N-6-methyladenosine and gene expression control. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 313−326.

(31) Xu, C.; Liu, K.; Ahmed, H.; Loppnau, P.; Schapira, M.; Min, J. R. Structural Basis for the Discriminative Recognition of N-6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. *J. Biol. Chem.* **2015**, *290*, 24902−24913.

(32) Zhu, T.; Roundtree, I. A.; Wang, P.; Wang, X.; Wang, L.; Sun, C.; Tian, Y.; Li, J.; He, C.; Xu, Y. Crystal structure of the YTH

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. *Cell Res.* **2014**, *24*, 1493−1496.
(33) Xu, C.; Wang, X.; Liu, K.; Roundtree, I. A.; Tempel, W.; Li, Y.; Lu, Z.; He, C.; Min, J. Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* **2014**, *10*, 927.
(34) Li, F.; Zhao, D.; Wu, J.; Shi, Y. Structure of the YTH domain of human YTHDF2 in complex with an m(6)A mononucleotide reveals an aromatic cage for m(6)A recognition. *Cell Res.* **2014**, *24*, 1490−1492.
(35) Wojtas, M. N.; Pandey, R. R.; Mendel, M.; Homolka, D.; Sachidanandam, R.; Pillai, R. S. Regulation of m(6) A Transcripts by the 3′ -> 5′ RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. *Mol. Cell* **2017**, *68*, 374.
(36) Liu, N.; Dai, Q.; Zheng, G. Q.; He, C.; Parisien, M.; Pan, T. N-6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **2015**, *518*, 560−564.
(37) Edupuganti, R. R.; Geiger, S.; Lindeboom, R. G.; Shi, H.; Hsu, P. J.; Lu, Z.; Wang, S. Y.; Baltissen, M. P.; Jansen, P. W.; Rossa, M.; Müller, M. N6-methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. *Mol. Cell. Proteomics* **2017**, *16*, S42−S42.
(38) Luo, S.; Tong, L. Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. *Proc. Natl. Acad. Sci.* **2014**, *111*, 13834−13839.
(39) Śledź, P.; Caflisch, A. Protein structure-based drug design: from docking to molecular dynamics. *Curr. Opin. Struct. Biol.* **2018**, *48*, 93−102.
(40) Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.* **2012**, *41*, 429−452.
(41) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646.
(42) Vangaveti, S.; Ranganathan, S. V.; Chen, A. A. Advances in RNA molecular dynamics: a simulator's guide to RNA force fields. *Wiley Interdiscip. Rev.: RNA* **2017**, *8*, No. e1396.
(43) Šponer, J.; Krepl, M.; Banáš, P.; Kührová, P.; Zgarbová, M.; Jurečka, P.; Havrila, M.; Otyepka, M. How to understand atomistic molecular dynamics simulations of RNA and protein-RNA complexes? *Wiley Interdiscip. Rev.: RNA* **2017**, *8*, No. e1405.
(44) Krepl, M.; Havrila, M.; Stadlbauer, P.; Banas, P.; Otyepka, M.; Pasulka, J.; Stefl, R.; Sponer, J. Can We Execute Stable Microsecond-Scale Atomistic Simulations of Protein-RNA Complexes? *J. Chem. Theory Comput.* **2015**, *11*, 1220−1243.
(45) MacKerell, A. D., Jr; Nilsson, L. Molecular dynamics simulations of nucleic acid-protein complexes. *Curr. Opin. Struct. Biol.* **2008**, *18*, 194−199.
(46) Xu, Y.; Vanommeslaeghe, K.; Aleksandrov, A.; MacKerell, A. D., Jr; Nilsson, L. Additive CHARMM Force Field for Naturally Occurring Modified Ribonucleotides. *J. Comput. Chem.* **2016**, *37*, 896−912.
(47) Aduri, R.; Psciuk, B. T.; Saro, P.; Taniga, H.; Schlegel, H. B.; SantaLucia, J. AMBER force field parameters for the naturally occurring modified nucleosides in RNA. *J. Chem. Theory Comput.* **2007**, *3*, 1464−1475.
(48) Denning, E. J.; MacKerell, A. D., Jr Intrinsic contribution of the 2′-hydroxyl to RNA conformational heterogeneity. *J. Am. Chem. Soc.* **2012**, *134*, 2800−2806.
(49) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr Impact of 2 ′-Hydroxyl Sampling on the Conformational Properties of RNA: Update of the CHARMM All-Atom Additive Force Field for RNA. *J. Comput. Chem.* **2011**, *32*, 1929−1943.
(50) MacKerell, A. D., Jr Contribution of the intrinsic mechanical energy of the phosphodiester linkage to the relative stability of the A, BI, and BII forms of duplex DNA. *J. Phys. Chem. B* **2009**, *113*, 3235−3244.
(51) Tan, D.; Piana, S.; Dirks, R. M.; Shaw, D. E. RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc. Natl. Acad. Sci.* **2018**, *115*, E1346−E1355.
(52) Robertson, M. J.; Qian, Y.; Robinson, M. C.; Tirado-Rives, J.; Jorgensen, W. L. Development and Testing of the OPLS-AA/M Force Field for RNA. *J. Chem. Theory Comput.* **2019**, *15*, 2734−2742.
(53) Kuhrova, P.; Mlynsky, V.; Zgarbova, M.; Krepl, M.; Bussi, G.; Best, R. B.; Otyepka, M.; Sponer, J.; Banas, P. Improving the Performance of the Amber RNA Force Field by Tuning the Hydrogen-Bonding Interactions. *J. Chem. Theory Comput.* **2019**, *15*, 3288−3305.
(54) Templeton, C.; Elber, R. Why Does RNA Collapse? The Importance of Water in a Simulation Study of Helix-Junction-Helix Systems. *J. Am. Chem. Soc.* **2018**, *140*, 16948−16951.
(55) Xu, Y.; MacKerell, A. D., Jr.; Nilsson, L. Structural effects of modified ribonucleotides and magnesium in transfer RNAs. *Bioorg. Med. Chem.* **2016**, *24*, 4826−4834.
(56) Kuhrova, P.; Best, R. B.; Bottaro, S.; Bussi, G.; Sponer, J.; Otyepka, M.; Banas, P. Computer Folding of RNA Tetraloops: Identification of Key Force Field Deficiencies. *J. Chem. Theory Comput.* **2016**, *12*, 4534−4548.
(57) Krepl, M.; Cléry, A.; Blatter, M.; Allain, F. H.; Sponer, J. Synergy between NMR measurements and MD simulations of protein/RNA complexes: application to the RRMs, the most common RNA recognition motifs. *Nucleic Acids Res.* **2016**, *44*, 6452−6470.
(58) Musiani, F.; Rossetti, G.; Capece, L.; Gerger, T. M.; Micheletti, C.; Varani, G.; Carloni, P. Molecular Dynamics Simulations Identify Time Scale of Conformational Changes Responsible for Conformational Selection in Molecular Recognition of HIV-1 Transactivation Responsive RNA. *J. Am. Chem. Soc.* **2014**, *136*, 15631−15637.
(59) Salmon, L.; Bascom, G.; Andricioaei, I.; Al-Hashimi, H. M. A General Method for Constructing Atomic-Resolution RNA Ensembles using NMR Residual Dipolar Couplings: The Basis for Interhelical Motions Revealed. *J. Am. Chem. Soc.* **2013**, *135*, 5457−5466.
(60) Banás, P.; Hollas, D.; Zgarbová, M.; Jurecka, P.; Orozco, M.; Cheatham, T. E., III; Sponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6*, 3836−3849.
(61) Banáš, P.; Jurečka, P.; Walter, N. G.; Šponer, J.; Otyepka, M. Theoretical studies of RNA catalysis: Hybrid QM/MM methods and their comparison with MD and QM. *Methods* **2009**, *49*, 202−216.
(62) Nam, K.; Gao, J.; York, D. M. Quantum mechanical/molecular mechanical simulation study of the mechanism of hairpin ribozyme catalysis. *J. Am. Chem. Soc.* **2008**, *130*, 4680−4691.
(63) Nam, K.; Gao, J.; York, D. M. Electrostatic interactions in the hairpin ribozyme account for the majority of the rate acceleration without chemical participation by nucleobases. *RNA* **2008**, *14*, 1501−1507.
(64) Yoo, J.; Aksimentiev, A. New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8432−8449.
(65) Yoo, J.; Aksimentiev, A. Improved Parameterization of Amine-Carboxylate and Amine-Phosphate Interactions for Molecular Dynamics Simulations Using the CHARMM and AMBER Force Fields. *J. Chem. Theory Comput.* **2016**, *12*, 430−443.
(66) Yoo, J. J.; Aksimentiev, A. Improved Parametrization of Li+, Na +, K+, and Mg2+ Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. *J. Phys. Chem. Lett.* **2012**, *3*, 45−50.
(67) Luo, Y.; Roux, B. Simulation of Osmotic Pressure in Concentrated Aqueous Salt Solutions. *J. Phys. Chem. Lett.* **2010**, *1*, 183−189.
(68) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J. P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput.* **2018**, *14*, 2084−2108.
(69) Yuan, Y.; Zhang, Z.; Mills, M. J.; Hu, R.; Zhang, R. Assessing Force Field Potential Energy Function Accuracy via a Multipolar Description of Atomic Electrostatic Interactions in RNA. *J. Chem. Inf. Model.* **2018**, *58*, 2239−2254.

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014

(70) Lemkul, J. A.; MacKerell, A. D., Jr Polarizable force field for RNA based on the classical drude oscillator. *J. Comput. Chem.* **2018**, *39*, 2624−2646.
(71) Kabsch, W. Xds. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125−132.
(72) Mccoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.
(73) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352−367.
(74) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.
(75) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(76) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(77) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(78) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant-Pressure Molecular-Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(79) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. *J. Chem. Phys.* **1995**, *103*, 4613−4621.
(80) Steinbach, P. J.; Brooks, B. R. New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. *J. Comput. Chem.* **1994**, *15*, 667−683.
(81) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(82) Huang, J.; MacKerell, A. D., Jr CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135−2145.
(83) Changeux, J. P.; Edelstein, S. Conformational selection or induced fit? 50 years of debate resolved. *F1000Prime Rep.* **2011**, *3*, 19.
(84) Zhang, X.; Wei, L. H.; Wang, Y.; Xiao, Y.; Liu, J.; Zhang, W.; Yan, N.; Amu, G.; Tang, X.; Zhang, L.; Jia, G. Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 2919−2924.
(85) Shi, H.; Wang, X.; Lu, Z.; Zhao, B. S.; Ma, H.; Hsu, P. J.; Liu, C.; He, C. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. *Cell Res.* **2017**, *27*, 315.
(86) Li, A.; Chen, Y. S.; Ping, X. L.; Yang, X.; Xiao, W.; Yang, Y.; Sun, H. Y.; Zhu, Q.; Baidya, P.; Wang, X.; Bhattarai, D. P.; Zhao, Y. L.; Sun, B. F.; Yang, Y. G. Cytoplasmic m(6)A reader YTHDF3 promotes mRNA translation. *Cell Res.* **2017**, *27*, 444−447.

DOI: 10.1021/acs.jctc.9b00987
J. Chem. Theory Comput. 2019, 15, 7004−7014