JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Letter

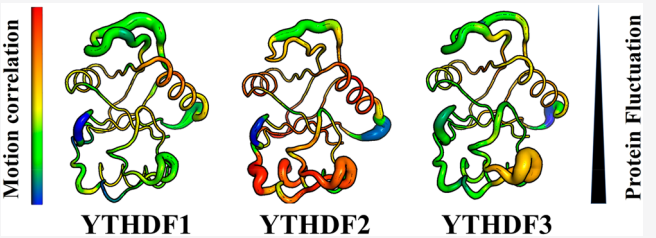
Structural and Dynamic Insights into Redundant Function of YTHDF Proteins

Yaozong Li,[§] Rajiv K. Bedi,[§] Elena V. Moroz-Omori,* and Amedeo Caflisch*

Cite This: https://dx.doi.org/10.1021/acs.jcim.0c01029 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Three YTH-domain family proteins (YTHDF1, YTHDF2, and YTHDF3) recognize the N^6-methyladenosine (m^6A) modification of mRNA in cells. However, the redundancy of their cellular functions has been disputed. We investigate their interactions with m^6A-containing RNA using X-ray crystallography and molecular dynamics (MD). The new X-ray structures and MD simulations show that the three proteins share identical interactions with the m^6A-containing RNA and have similar intrinsic plasticity, thus evidencing the redundant roles of the three proteins in cellular functions.



The finely organized network of gene expression comprising RNA transcription, splicing, transport, translation, and degradation is often perturbed in cancer and other diseases.[1,2] In addition to previously known regulators of gene expression, the recently discovered layer of regulation based on cotranscriptional RNA modifications gave rise to a new field named epitranscriptomics.[2,3] While there are over 160 different RNA modifications discovered to date, one of the most abundant modifications, N^6-methyladenosine or m^6A, is involved in most of the aspects of messenger RNA (mRNA) regulation, i.e., alternative polyadenylation, splicing, nuclear export, degradation, and translation initiation.[2,3]

Most of these diverse functions are realized via the m^6A binding proteins containing the YT521-B homology (YTH) domain, so-called m^6A "readers".[4−7] Among those, YTHDC1 is found in the nucleus, while YTHDC2, YTHDF1, YTHDF2, and YTHDF3 are primarily cytoplasmic in mammalian cells (abbreviated hereafter as DC1, DC2, DF1, DF2, and DF3, respectively).[1] DC1 is involved in mRNA export from nucleus to cytoplasm, as well as in regulation of mRNA splicing and alternative polyadenylation.[1,8] DC2 plays an essential role in spermatogenesis through regulating the transition from mitosis to meiosis.[1]

The three human DF proteins consist of a YTH domain of 145 residues located in the sequence between unstructured N-terminal and C-terminal segments of about 400 and 35 residues, respectively.[9] The sequence identity of the YTH domain is 86%. Microscale thermophoresis[10] and a time-resolved Förster resonance energy transfer assay[11] show that all three YTHDF proteins have similar binding affinities toward the preferred RNA motif. Nevertheless, their functions have been largely disputed.[9] Some studies claim distinct functions of the DF proteins where DF1 enhances translation by interacting with ribosomes and translation initiating factors, DF2 promotes targeted m^6A-marked mRNA degradation through recruitment of RNA decay machinery, and DF3 assists with regulation of mRNA fate via interaction with both DF1 and DF2.[1,12,13] In contrast, recent papers point out the redundancy of the DF proteins.[10,14,15] High-throughput *in vitro* selection methodology provided evidence that DF1 and DF2 have similar sequence binding preferences for m^6A-containing RNA.[16] All three DF readers are thought to mediate mRNA degradation either by direct recruitment of the CCR4-NOT deadenylase complex[10,14] or by transporting its m^6A-modified cargo to P-bodies due to the presence of a disordered proline-glutamine-rich region capable of liquid−liquid phase separation.[15,17−19] On one hand, the sequence and structure similarity of the YTH domains of DF1 and DF2 point toward redundancy or cumulative action; on the other hand, the small variability in sequence, difference in the expression levels in different tissues, and possibility of post-translational modifications could provide the basis for the distinct functions of DF readers (Figure S1). In particular, the overall sequence identity of only 57% does not allow one to draw conclusions on the functional redundancy.

Here, we focus on the YTH domain of the three human DF readers. The protein dynamics and motions are compared between three proteins to check if minor variability in amino acid sequence causes significant difference on them. In addition, the crystal structure of the YTH domain of the

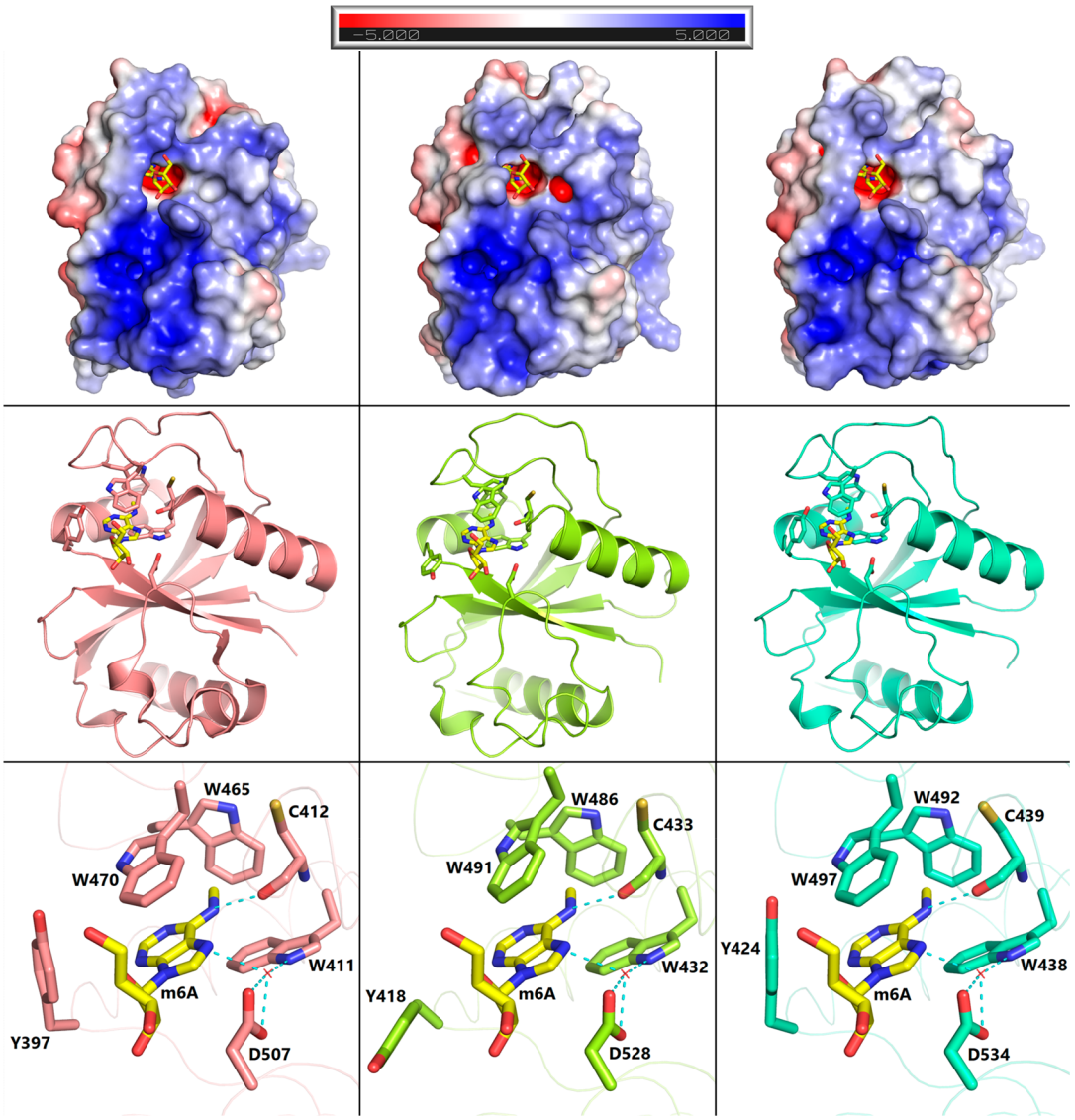


Figure 1. YTH domains of the human DF proteins show similar electrostatic potential at the RNA recognition surface (top), overall fold (middle), and interactions with m^{6}A (bottom). (Left) DF1, PDB code 4RCJ; (middle) DF2, PDB code 4RDN; (right) DF3, PDB code 6ZOT (this work). Electrostatic surface potentials were calculated using the APBS 2.1 PyMOL plugin.[21] The potentials are on a red−white−blue [± 5 kJ/mol/e] color map. Parameters used are 0.15 M ionic strength in monovalent salt, 310 K, protein dielectric 2, and solvent dielectric 78.

DF3 reader has not been reported so far, which should provide additional mechanistic insight into m^{6}A recognition at the atomic level of detail and whether it is similar or different from the YTH domains of DF1 and DF2.

To investigate the binding of m^{6}A-containing RNA (more specifically, the GG(m^{6}A)CU) to the YTH domain of the DF3 protein, we solved the crystal structure of the complex at a 2.7 Å resolution. The overall structure, electrostatic potential on the RNA recognition surface, and binding motif of m^{6}A are essentially identical as in the previously disclosed holo structures of DF1 and DF2 (Figure 1). In detail, the adenine ring is nestled in an aromatic cage formed by the side chains of Tyr424, Trp438, Trp492, and Trp497 (residue numbering of DF3, Figure 1, bottom, and Figure S2). The methyl group of m^{6}A optimally fills the cavity lined by the three indole rings, while the nitrogen in position 6 is involved as the donor in a hydrogen bond with the carbonyl oxygen of Cys439. Furthermore, a conserved water molecule mediates hydrogen bonds between the adenine nitrogen in position 7 and the side chains of Asp534 and Trp438. One minor difference is observed for the orientation of the Tyr418 side chain in DF2 (Figure 1, bottom). This difference is probably due to the smaller ligand as the only available holo structure of DF2 is a complex with the m^{6}A mononucleotide, while the structures of DF1 and DF3 were solved in complex with pentanucleotide GG(m^{6}A)CU. Overall, the essentially identical interactions observed in the crystal structures of the three YTH domains are congruent with the similar affinity for m^{6}A containing RNA oligonucleotides.[20]

As the cellular phenotype might be related to protein dynamics, we decided to study the unbound state of the YTH domains of the DF proteins by atomistic simulations. To interrogate the dynamics of the apo YTH domain of the three DF proteins, we carried out explicit solvent MD simulations for a total sampling of 5 μs for each of the three domains (see Materials and Methods in the Supporting Information). The MD simulations show similar flexibility and correlated motion for the three YTH domains (Figure 2 and Figures S3 and S4).

https://dx.doi.org/10.1021/acs.jcim.0c01029
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

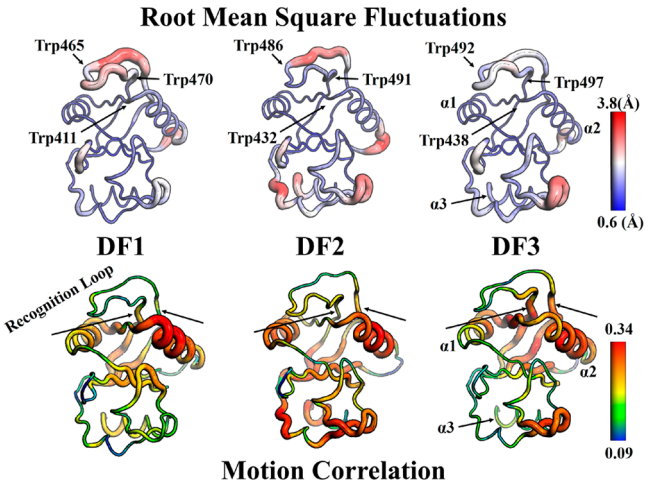


Figure 2. YTH domains of the DF proteins show similar flexibility and correlated motion. (Top) Structural mapping of root mean square fluctuation (RMSF) profiles along the sequence. The RMSF values were calculated for each residue by averaging over the nonhydrogen atoms of the backbone. RMSF values are rendered from low to high by tube thickness (small to large) and color (blue to red). (Bottom) Structural mapping of positional covariance of backbone atoms. The amount of motion correlation for each residue is rendered from low to high by tube thickness (small to large) and color (blue to red).

The recognition loop (residues 480−500 in DF3) shows pronounced flexibility, which contributes two of the four aromatic residues of the m^{6}A binding site (Trp492 and Trp497 in DF3, Figure 2, top). The slightly larger fluctuations of the recognition loop of DF1 (residues 453−473) most likely originate from the intrinsic flexibility of Gly459 as DF2 and DF3 have an asparagine at the corresponding position. Other loop segments that connect regular elements of the secondary structure also show enhanced flexibility in the microsecond time scale. The major differences in flexibility between the three YTH domains are observed for the C-terminal loop and α3 helix (residues 510−540 and 541−550, respectively, in DF3), which are located far away from the m^{6}A recognition pocket.

The recognition of m^{6}A-containing RNA by the YTH domains is related to the protein motions, especially to the recognition loop. We first wondered if the correlated motions of individual segments of the fold are similar in the three YTH domains in the unbound state. To this end, we carried out an analysis of positional covariance (Figures S5 and S6). The (anti)correlation values were mapped to the structures in a residue-based manner.[22] The α1 and α2 helices and most of the β-sheet undergo correlated displacement (red structural segments in Figure 2, bottom). These regular elements of secondary structure form the central core of the YTH fold. In contrast, the large fluctuations of the recognition loop and most of the C-terminal loop are not significantly correlated with other parts of the domain. As expected intuitively, the two types of analyses show that the rigid structural components (blue in Figure 2, top) tend to be more correlated in their displacement (red in Figure 2, bottom). The principal component analysis reveals the pronounced slow motions of the recognition loop, thus further evidencing its dynamic feature for the recognition of m^{6}A-containing RNA (Figure S7). These simulation results indicate that the core region stabilizes the fold of the protein including part (but not all) of the m^{6}A binding pocket, while the recognition loop easily adopts different conformations, which may facilitate the binding of m^{6}A-containing RNAs. The similar plasticity of the three YTHDF proteins and comparable protein motions of most of their regular elements of secondary structure provide further support to their functional redundancy.

In conclusion, our crystal structure of the DF3/GG(m^{6}A)CU complex and atomistic simulation studies have provided support for the view of a common (and cumulative) function of the three human DF proteins. It remains to be investigated if post-translational modifications, differential binding to non-m^{6}A modifications of RNA, and/or sequence differences in the disordered N- and C-terminal segments have a substantial influence on the role of individual DF proteins. To answer these questions, more biochemical, structural, and *in silico* studies on the DF proteins are required. For example, the binding mechanisms of modified RNAs to certain post-translationally decorated DF proteins could be studied by combing ITC and X-ray crystallography, and global conformational changes of the full-length proteins should be analyzed by enhanced sampling MD simulations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c01029.

Electron density map, RMSF and correlation profiles, covariance matrixes, PCA analysis, and materials and methods (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Elena V. Moroz-Omori − *Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland;* orcid.org/0000-0001-6485-3481; Phone: +41 44 635 5521; Email: elena.omori@uzh.ch

Amedeo Caflisch − *Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland;* orcid.org/0000-0002-2317-6792; Phone: +41 44 635 5521; Email: caflisch@bioc.uzh.ch

Authors

Yaozong Li − *Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland; Department of Chemistry, Umeå University, SE-90187 Umeå, Sweden;* orcid.org/0000-0002-5796-2644

Rajiv K. Bedi − *Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.0c01029

Author Contributions

[§]Y. Li and R. V. Bedi contributed equally.

Notes

The PDB code of the crystal structure of the DF3/GG(m^{6}A)CU complex is 6ZOT. The structure will be released upon article publication.

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to Beat Blattmann for his help with setting up protein crystallization trials. We thank the staff at the

https://dx.doi.org/10.1021/acs.jcim.0c01029
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Swiss Light Source (PSI) for their support with data collection. We thank the Swiss National Supercomputing Centre in Lugano for providing the computational resources. We also thank Dr. Paweł Śledź, Dr. Xiang Wang, Dr. Lars Wiedmer, Johannes Widter, and Tong Chen for valuable discussions. This work was supported financially by the SNSF (E.M.-O., Grant CRSK-3_190825; A.C., Grant 310030B-189363) and the Swedish Research Council (Y.L., Grant VR 2019-00608).

REFERENCES

(1) Vu, L. P.; Cheng, Y.; Kharas, M. G. The Biology of m6A RNA Methylation in Normal and Malignant Hematopoiesis. *Cancer Discovery* **2019**, *9*, 25−33.
(2) Kadumuri, R. V.; Janga, S. C. Epitranscriptomic Code and Its Alterations in Human Disease. *Trends Mol. Med.* **2018**, *24*, 886−903.
(3) Fu, Y.; Dominissini, D.; Rechavi, G.; He, C. Gene Expression Regulation Mediated Through Reversible m6A RNA Methylation. *Nat. Rev. Genet.* **2014**, *15*, 293−306.
(4) Dominissini, D.; Moshitch-Moshkovitz, S.; Schwartz, S.; Salmon-Divon, M.; Ungar, L.; Osenberg, S.; Cesarkas, K.; Jacob-Hirsch, J.; Amariglio, N.; Kupiec, M.; Sorek, R.; Rechavi, G. Topology of the Human and Mouse m6A RNA Methylomes Revealed by m6A-Seq. *Nature* **2012**, *485*, 201−206.
(5) 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. N6-Methyladenosine-Dependent Regulation of Messenger RNA Stability. *Nature* **2014**, *505*, 117.
(6) Li, Y.; Bedi, R. K.; Wiedmer, L.; Huang, D.; Śledź, P.; Caflisch, A. Flexible Binding of m6A Reader Protein YTHDC1 to its Preferred RNA Motif. *J. Chem. Theory Comput.* **2019**, *15*, 7004−7014.
(7) Bedi, R. K.; Huang, D.; Wiedmer, L.; Li, Y.; Dolbois, A.; Wojdyla, J. A.; Sharpe, M. E.; Caflisch, A.; Sledz, P. Selectively Disrupting m6A-Dependent Protein-RNA Interactions with Fragments. *ACS Chem. Biol.* **2020**, *15*, 618−625.
(8) Ke, S.; Pandya-Jones, A.; Saito, Y.; Fak, J. J.; Vågbø, C. B.; Geula, S.; Hanna, J. H.; Black, D. L.; Darnell, J. E.; Darnell, R. B. m6A mRNA Modifications are Deposited in Nascent Pre-mRNA and are not Required for Splicing but do Specify Cytoplasmic Turnover. *Genes Dev.* **2017**, *31*, 990−1006.
(9) Patil, D. P.; Pickering, B. F.; Jaffrey, S. R. Reading m6A in the Transcriptome: m6A-Binding Proteins. *Trends Cell Biol.* **2018**, *28*, 113−127.
(10) Zaccara, S.; Jaffrey, S. R. A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. *Cell* **2020**, *181*, 1582−1595 e18.
(11) Wiedmer, L.; Eberle, S. A.; Bedi, R. K.; Śledź, P.; Caflisch, A. A Reader-Based Assay for m6A Writers and Erasers. *Anal. Chem.* **2019**, *91*, 3078−3084.
(12) Shi, H.; Wang, X.; Lu, Z.; Zhao, B. S.; Ma, H.; Hsu, P. J.; Liu, C.; He, C. YTHDF3 Facilitates Translation and Decay of N6-Methyladenosine-Modified RNA. *Cell Res.* **2017**, *27*, 315−328.
(13) Shi, H.; Zhang, X.; Weng, Y.-L.; Lu, Z.; Liu, Y.; Lu, Z.; Li, J.; Hao, P.; Zhang, Y.; Zhang, F.; Wu, Y.; Delgado, J. Y.; Su, Y.; Patel, M. J.; Cao, X.; Shen, B.; Huang, X.; Ming, G.-l.; Zhuang, X.; Song, H.; He, C.; Zhou, T. m6A Facilitates Hippocampus-Dependent Learning and Memory Through YTHDF1. *Nature* **2018**, *563*, 249.
(14) Du, H.; Zhao, Y.; He, J.; Zhang, Y.; Xi, H.; Liu, M.; Ma, J.; Wu, L. YTHDF2 Destabilizes m6A-Containing RNA Through Direct Recruitment of the CCR4-NOT Deadenylase Complex. *Nat. Commun.* **2016**, *7*, na.
(15) Yang, P.; Mathieu, C.; Kolaitis, R. M.; Zhang, P.; Messing, J.; Yurtsever, U.; Yang, Z.; Wu, J.; Li, Y.; Pan, Q.; Yu, J.; Martin, E. W.; Mittag, T.; Kim, H. J.; Taylor, J. P. G3BP1 is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. *Cell* **2020**, *181*, 325−345.
(16) Arguello, A. E.; Leach, R. W.; Kleiner, R. E. In Vitro Selection with a Site-Specifically Modified RNA Library Reveals the Binding Preferences of N6-Methyladenosine Reader Proteins. *Biochemistry* **2019**, *58*, 3386−3395.
(17) Gao, Y.; Pei, G.; Li, D.; Li, R.; Shao, Y.; Zhang, Q. C.; Li, P. Multivalent m6A Motifs Promote Phase Separation of YTHDF Proteins. *Cell Res.* **2019**, *29*, 767−769.
(18) Wang, J.; Wang, L.; Diao, J.; Shi, Y. G.; Shi, Y.; Ma, H.; Shen, H. Binding to m6A RNA Promotes YTHDF2-Mediated Phase Separation. *Protein Cell* **2020**, *11*, 304−307.
(19) Ries, R. J.; Zaccara, S.; Klein, P.; Olarerin-George, A.; Namkoong, S.; Pickering, B. F.; Patil, D. P.; Kwak, H.; Lee, J. H.; Jaffrey, S. R. m6A Enhances the Phase Separation Potential of mRNA. *Nature* **2019**, *571*, 424−428.
(20) 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. Biol. Chem.* **2015**, *290*, 24902−24913.
(21) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037−10041.
(22) Li, Y.; Nam, K. Dynamic, Structural and Thermodynamic Basis of Insulin-Like Growth Factor 1 Kinase Allostery Mediated by Activation Loop Phosphorylation. *Chem. Sci.* **2017**, *8*, 3453−3464.

https://dx.doi.org/10.1021/acs.jcim.0c01029
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX