nature chemical biology

BRIEF COMMUNICATION

PUBLISHED ONLINE: 21 SEPTEMBER 2014 | DOI: 10.1038/NCHEMBIO.1654

Structural basis for selective binding of m^{6}A RNA by the YTHDC1 YTH domain

Chao Xu[1,6*], Xiao Wang[2,3,6], Ke Liu[1,6], Ian A Roundtree[3,4], Wolfram Tempel[1], Yanjun Li[1], Zhike Lu[2,3], Chuan He[2,3*] & Jinrong Min[1,5*]

N^6-methyladenosine (m^{6}A) is the most abundant internal modification of nearly all eukaryotic mRNAs and has recently been reported to be recognized by the YTH domain family proteins. Here we present the crystal structures of the YTH domain of YTHDC1, a member of the YTH domain family, and its complex with an m^{6}A-containing RNA. Our structural studies, together with transcriptome-wide identification of YTHDC1-binding sites and biochemical experiments, not only reveal the specific mode of m^{6}A-YTH binding but also explain the preferential recognition of the GG(m^{6}A)C sequences by YTHDC1.

Methylation of internal adenosines to form m^{6}A is a key processing event during maturation of eukaryotic mRNAs that occurs in concert with 5′ capping, 3′ polyadenylation and splicing[1]. m^{6}A is highly conserved across all eukaryotes from yeast[2] to humans[3] within a G(m^{6}A)C (70%) or A(m^{6}A)C (30%) motif and is physiologically essential to metazoans[4–6]. Transcriptome-wide studies have revealed that m^{6}A is a widespread modification present in over 7,000 human genes[7,8]. Analogous to other chemical codes that overlay primary sequence, such as DNA methylation and histone marks, the m^{6}A modification on mRNA and certain noncoding RNA is reversible and carries regulatory functions[9–11]. The biology of m^{6}A is achieved by three types of proteins[9]: a m^{6}A methyltransferase complex (such as METTL3–METTL14) that installs the methyl group[12,13], demethylases (FTO[10] and ALKBH5 (ref. 11)) that reverse methylation and m^{6}A-specific RNA-binding proteins that recognize m^{6}A ('readers')[7,9,14]. Identification of m^{6}A readers is especially important in determining the cellular function of m^{6}A. YTHDF2, a member of the YTH family proteins, has recently been characterized as the first m^{6}A reader that regulates the cytoplasmic stability of methylated RNA[14], implicating reversible RNA modification as a new layer of gene regulation at the post-transcriptional level[9,14].

The YTH RNA-binding domain is highly conserved in eukaryotes[15], and there are five proteins (YTHDF1–3, YTHDC1 and YTHDC2) that contain the YTH domain in human cells (**Supplementary Results**, **Supplementary Fig. 1**). We and others have previously shown that three of these proteins, YTHDF1, YTHDF2 and YTHDF3, are cytoplasmic m^{6}A-specific binders[7,14]; however, the molecular basis for selective m^{6}A recognition by the YTH domain is unclear. YTHDC1 resides in the nucleus and is particularly interesting. YTHDC1 forms YT bodies at transcriptionally active sites and adjacent to RNA processing speckles[16]. It interacts with other splicing factors and has also been reported to be a potential tumor repressor in endometrial cancer[17,18].

Studies of m^{6}A RNA methylomes in different organisms, including human, mouse and yeast, reveal that m^{6}A often exists in an RRACN (R is G or A) consensus motif[2,7,8] and that the YTH domain of YTHDF2 preferentially binds a conserved motif of G(m^{6}A)C[14]. To corroborate the binding of the YTH domain of YTHDC1 to m^{6}A-containing RNA and determine its sequence preference, we first measured its binding affinity to the previously reported 42-mer RNA oligonucleotides by gel shift assay and found that the YTH domain of YTHDC1 preferentially binds the methylated RNA with a dissociation constant (K_d) of around 0.4 μM (**Supplementary Fig. 2**), similar to that of YTHDF proteins[15]. The selective binding of YTHDC1 to methylated RNA was further confirmed by *in vitro* pull-down assay (**Fig. 1a**). We next applied photoactivatable ribonucleoside crosslinking and immunoprecipitation (PAR-CLIP)[19] to identify the RNA binding sites of YTHDC1. Among the 10,245 identified PAR-CLIP peaks, 51% overlap with the previously reported m^{6}A sites

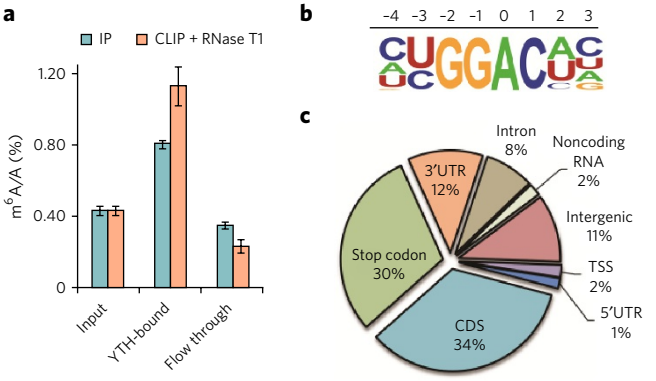

Figure 1 | YTHDC1 is a nuclear m^{6}A reader. (**a**) LC/MS/MS showed that the recombinantly expressed YTH domain of YTHDC1 was able to enrich m^{6}A-containing RNA from poly(A)-tailed RNAs of human HeLa cell. When the YTHDC1 protein was crosslinked with its associated RNA by ultraviolet light and subjected to partial digestion, the enrichment of m^{6}A at the YTH-bound portion was increased compared to that without crosslinking treatment. Blue, immunoprecipitation (IP) of the YTH domain of YTHDC1 with its associated RNA; pink, protein–RNA complexes that are UV-crosslinked and partially RNase T1 digested before IP (CLIP). Error bars represent mean ± s.d., n = 2 technical replicates. (**b**) Binding motif identified by HOMER motif analysis (P = 1.0 × 10^{-246}; 2,778 sites were found under this motif). (**c**) Pie chart depicting the region distribution of YTHDC1-binding sites identified by PAR-CLIP, TSS, 200-bp window from the transcription starting site; CDS, code-determining sequence; stop codon, 400-bp window centered on stop codon.

[1]Structural Genomics Consortium, University of Toronto, Toronto, Ontario, Canada. [2]Department of Chemistry and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois, USA. [3]Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois, USA. [4]Medical Scientist Training Program and Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois, USA. [5]Department of Physiology, University of Toronto, Toronto, Ontario, Canada. [6]These authors contributed equally to the work. *e-mail: chaor.xu@utoronto.ca, chuanhe@uchicago.edu or jr.min@utoronto.ca

Table 1 | Binding affinities of RNA constructs to the YTH domain of wild-type and mutant YTHDC1 measured by ITC.

	RNAs	K_d (μM)
1	GAACCGG(m^{6}A)CUGUCUUA	0.30 ± 0.06
2	GAACCGA(m^{6}A)CUGUCUUA	2.0 ± 0.4
3	GAACCGU(m^{6}A)CUGUCUUA	0.50 ± 0.12
4	GAACCGC(m^{6}A)CUGUCUUA	0.40 ± 0.07
5	GG(m^{6}A)CU	2.0 ± 0.1
6	GA(m^{6}A)CU	15 ± 2
7	GU(m^{6}A)CU	5.7 ± 0.7
8	GC(m^{6}A)CU	5.6 ± 0.9
9	GG(m^{6}A)AU	6.2 ± 0.4
10	GG(m^{6}A)UU	3.3 ± 0.2
11	GG(m^{6}A)GU	2.8 ± 0.3
12	AG(m^{6}A)CU	3.8 ± 0.4
13	UG(m^{6}A)CU	4.3 ± 0.5
14	G(m^{6}A)C	28 ± 5
15	GAACCGGACUGUCUUA	NB
16	GGACU	NB
17	GG(m^{6}A)CU	NB (W377A)
18	GG(m^{6}A)CU	NB (W428A)
19	GG(m^{6}A)CU	18 ± 2 (R475F)
20	GG(m^{6}A)CU	210 ± 20 (R475A)

NB, No detectable binding. The data represent mean values ±s.d., where s.d. values were calculated from the ITC curve fitting by Origin Software (MicroCal Inc.).

in human HeLa cells (**Supplementary Fig. 3a**)[11], and 21% overlap with the YTHDF2 binding sites (**Supplementary Fig. 3b**). RGAC is enriched in the PAR-CLIP peaks of YTHDC1 (**Supplementary Fig. 3c**), and the top binding motif of YTHDC1 is GGAC (**Fig. 1b** and **Supplementary Fig. 3d–f**), which is the same as that of YTHDF2 and consistent with the reported high-resolution mapping of m^{6}A sites[8]. The YTHDC1 PAR-CLIP sites mainly distribute on exons and peaks around stop codons, resembling the subtranscript distribution of the m^{6}A methylome (**Fig. 1c**), and indicating that YTHDC1 is a bona fide m^{6}A reader inside the cell. Functional clustering of the YTHDC1 gene targets reveals a significant portion of transcription regulators (13%, compared with 6% in nontargets) and implies an important role of YTHDC1 in gene expression ($P = 6.0 \times 10^{-15}$; **Supplementary Fig. 4**).

To understand the sequence selectivity of YTHDC1 and search for suitable m^{6}A RNA oligonucleotides for crystallography, we measured the binding affinities of the YTH domain of YTHDC1 to a series of RNA oligos of different lengths by the quantitative isothermal titration calorimetry (ITC) binding assay (**Table 1** and **Supplementary Table 1**). We found that (i) the YTHDC1 YTH domain recognizes RNA in an m^{6}A-dependent manner regardless of RNA length; (ii) the YTH domain of YTHDC1 prefers a G nucleotide at the −1 position relative to the m^{6}A, and an adenosine is least favored at this position; and (iii) the YTH domain of YTHDC1 shows a slight preference for G and C at the −2 and +1 positions, respectively. Most notably, our ITC binding results are consistent with our PAR-CLIP results, supporting the assertion that YTHDC1 preferentially binds the GG(m^{6}A)C sequence (**Table 1**).

To gain insights into the molecular mechanism of specific m^{6}A recognition by the YTH domain, we determined crystal structures of the YTH domain of human YTHDC1 and its complex with a 5-mer m^{6}A RNA (GG(m^{6}A)CU) (**Supplementary Table 2**). Consistent with the previously determined solution structure of apo-YTHDC1 (Protein Data Bank (PDB) code 2YUD), the YTH domain of YTHDC1 adopts an open α/β fold. Part of the loop linking the β4 and β5 strands (called L_{45} henceforth) is disordered in the apo structure (**Supplementary Fig. 5**) but is resolved in the m^{6}A RNA complex structure (**Fig. 2a**). In addition, The YTHDC1 YTH domain contains 3_{10} helices (η1–η3), one following L_{45} (η1) and the other two between the β5 and β6 strands (η2 and η3; **Supplementary Fig. 1b**). Like L_{45}, η1 is missing in the apo structure but is visible in the complex structure. Other than that, the apo and m^{6}A complex structures are almost identical with an r.m.s. deviation of only 0.42 Å between these two structures. In the YTHDC1–GG(m^{6}A)CU complex structure, all five nucleotides are visible (**Fig. 2a**).

The base of m^{6}A is snugly accommodated in a deep pocket formed by residues from β1, the loop between β1 and α1, α1, β2 and L_{45} (**Fig. 2a**,**b**). Notably, the methyl group of m^{6}A is recognized by an aromatic cage consisting of W377, W428 and L439 (**Fig. 2c** and **Supplementary Fig. 6**), and the distances between the N6 atom and the Leu439 and the indole planes of W377 and W428 are similar to those between the N^ε of the methylated lysines of histone H3 and the rings of the aromatic cage residues of HP1 or JARID1 (**Supplementary Fig. 7**)[20,21]. Notably, W377 and W428 are absolutely conserved in YTH domain family proteins (**Supplementary Fig. 1b**). Mutating either W377 or W428 to alanine completely disrupts its binding to m^{6}A RNA (**Table 1**), underscoring the importance of the cage residues in m^{6}A recognition. In addition to the cage accommodation, the adenine moiety of m^{6}A also forms three hydrogen bonds with the main chain carbonyl oxygen of S378 (6-amino group), the side chain NH_2 of N367 (N^1 in purine) and the main chain NH group of N363 (N^3). The C2′-ribosyl hydroxyl oxygen of m^{6}A forms another hydrogen bond with the side chain of N363 (**Supplementary Fig. 7**).

Our complex structure not only sheds light on why the YTH domain specifically recognizes m^{6}A but also provides a structural basis for the preference of a G nucleotide at the −1 position (G–1). The carbonyl oxygen (guanine 6-oxo) of the guanosine G–1

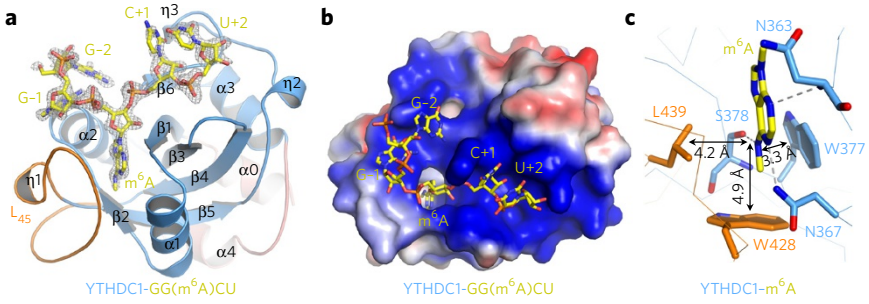

Figure 2 | Structural basis of preferential recognition of the GG(m^{6}A)C consensus motif by the YTH domain of YTHDC1. (**a**) Overall structure of the YTHDC1 YTH domain in complex with RNA GG(m^{6}A)CU. The protein is shown in blue cartoon except loop L_{45} and two α helices (α0 and α4), which are colored in orange and red, respectively. The 5-mer RNA is shown in yellow sticks. The simulated annealing m|F_o| – D|F_c| omit map of the 5-mer RNA is contoured at 2.5σ. (**b**) The electrostatic surface of the YTHDC1 YTH domain in complex with the 5-mer RNA GG(m^{6}A)CU (yellow stick). The m^{6}A RNA resides in a highly positively charged binding groove formed by the C-terminal ends of strands β1, β2 and β3 and subsequent loops, helix α1, L_{45} and the loop preceding β6. (**c**) The m^{6}A binding pocket of the YTHDC1 YTH domain. The protein and m^{6}A are shown in blue ribbon and yellow sticks, respectively. Residues involved in recognizing m^{6}A are shown in blue sticks. The distances between N^6 of m^{6}A and the three residues are labeled in black, and the hydrogen bonds are shown as gray dashes.

forms a hydrogen bond with the main chain NH group of V382 (**Supplementary Figs. 7** and **8a**). Replacing G–1 with any other nucleotide would disrupt this hydrogen bond (**Supplementary Fig. 8b–d**). Furthermore, replacing G–1 with adenosine could introduce steric clashes with V382 between the NH_2 group of adenosine and the main chain NH group of V382 (**Supplementary Fig. 8b**).

The cytidine at the +1 position (C+1) forms a water-mediated hydrogen bond with the side chain of N466 via its base carbonyl oxygen, two hydrogen bonds via its ribosyl hydroxyl oxygen and two hydrogen bonds, one with the side chain of R475 and one with the backbone of D476 via its phosphate group. Most importantly, this cytosine is stacked between the guanidinium group of R475 of YTHDC1 and a uracil at the +2 position, forming cation-π and π-π interactions, respectively (**Supplementary Fig. 6**). The similar stacking mode has been found in other protein–nucleic acid complexes, (**Supplementary Fig. 9**)[22]. The importance of this R475-mediated interaction has also been verified by our mutagenesis studies. Mutating R475 to phenylalanine diminishes binding affinity by 9-fold, whereas mutating R475 to alanine decreases binding affinity over 100-fold (**Table 1**). In contrast, we found that this stacking interaction is not base specific, consistent with our ITC data showing that replacing C+1 with other nucleotides only slightly affects the binding affinity (**Table 1**).

The guanosine at the −2 position (G–2) forms two base-specific hydrogen bonds with the main chain carbonyl oxygen of YTHDC1 D476, one formed directly via its NH_2 group (N^2) and the other water-mediated hydrogen bond via its NH group (N^1; **Supplementary Fig. 6**). Substituting G–2 for other nucleotides will disrupt both hydrogen bonds but will not introduce steric clashes (**Supplementary Fig. 10**). Accordingly, m^{6}A binding to YTHDC1 is only slightly reduced when replacing G–2 with other nucleotides (1.9- to 2.1-fold; **Table 1**). The uridine at the +2 position (U+2) forms five hydrogen bonds with YTHDC1 via phosphate and hydroxyl groups (**Supplementary Fig. 6**), consistent with a lack of selectivity at this position. Although YTHDC1 does not show sequence selectivity at the +2 position, it may also contribute to binding because a 3-mer G(m^{6}A)C RNA displays 14-fold weaker binding affinity than the 5-mer GG(m^{6}A)CU RNA (**Table 1**).

Structural and binding studies reported in this study unveil a general function of the YTH domain as a specific m^{6}A RNA reader, and a guanine nucleotide is favored at the position preceding m^{6}A. Previous m^{6}A methylome profiling shows that m^{6}A is often embedded in a RRACU sequence (R = G or A) in mouse and human[7,8] and in a RGAC motif in yeast[2], which is consistent with earlier biochemical studies showing that the m^{6}A methyltransferases specifically modify adenine within the consensus sequence motifs of GAC (70%) and AAC (30%)[3]. In addition to the YTH domain proteins, a RRM domain protein ELAVL1 has also been identified as an m^{6}A reader from an m^{6}A RNA pulldown assay[7]. We anticipate more m^{6}A readers to be identified in the future, which could shed further insight into the function of m^{6}A.

As the most abundant internal modification in RNA, the reversible m^{6}A modification has regulatory roles analogous to reversible DNA and histone modifications[9]. The m^{6}A mark could exhibit biological functions by disturbing the binding of normal protein factors, altering RNA secondary structure or recruiting specific reader proteins. We show here that the YTH domains have an exquisite pocket for specific recognition of the methyl group of m^{6}A, which consists of two conserved tryptophan residues (W377 and W428 of YTHDC1) and is distinct from that of the 5-methylcytosine by the MBD and SRA domains[22] and that of 7-methyl-GDP by eIF4E[23] (**Supplementary Fig. 9**). Instead, the recognition mode is similar to that of the methyllysine or methylarginine in histones[20,21] (**Supplementary Fig. 7**), which has not been observed for nucleic acid methylations in the past; the evolution of a methyl-binding mode of m^{6}A recognition similar to that of histones suggests functional importance of this reading process in mammalian cells. We also provide the molecular basis for the preference of the GG(m^{6}A)C sequence by the YTH domain proteins. Therefore, the current work serves as a prototype for selective recognition of the new emerging mode of m^{6}A RNA methylation in biological regulation.

Received 27 June 2014; accepted 3 September 2014; published online 21 September 2014; corrected after print 19 August 2015 and 11 November 2015

Methods

Methods and any associated references are available in the online version of the paper.

Accession codes. PDB. The crystal structures of YTHDC1 were deposited under accession codes 4R3H (apo-YTHDC1 YTH domain) and 4R3I (YTHDC1–m6A RNA complex), respectively. Gene Expression Omnibus. The PAR-CLIP sequencing data were deposited under accession number GSE58352.

References

1. Tuck, M.T. *Int. J. Biochem.* **24**, 379–386 (1992).
2. Schwartz, S. *et al. Cell* **155**, 1409–1421 (2013).
3. Wei, C.M. & Moss, B. *Biochemistry* **16**, 1672–1676 (1977).
4. Zhong, S. *et al. Plant Cell* **20**, 1278–1288 (2008).
5. Bokar, J.A. Fine-tuning of RNA functions by modification and editing. Vol. 12 (ed. Grosjean, H.) 141–177 (Springer-Verlag, Berlin Heidelberg, 2005).
6. Hongay, C.F. & Orr-Weaver, T.L. *Proc. Natl. Acad. Sci. USA* **108**, 14855–14860 (2011).
7. Dominissini, D. *et al. Nature* **485**, 201–206 (2012).
8. Meyer, K.D. *et al. Cell* **149**, 1635–1646 (2012).
9. Fu, Y., Dominissini, D., Rechavi, G. & He, C. *Nat. Rev. Genet.* **15**, 293–306 (2014).
10. Jia, G. *et al. Nat. Chem. Biol.* **7**, 885–887 (2011).
11. Zheng, G. *et al. Mol. Cell* **49**, 18–29 (2013).
12. Liu, J. *et al. Nat. Chem. Biol.* **10**, 93–95 (2014).
13. Wang, Y. *et al. Nat. Cell Biol.* **16**, 191–198 (2014).
14. Wang, X. *et al. Nature* **505**, 117–120 (2014).
15. Zhang, Z. *et al. J. Biol. Chem.* **285**, 14701–14710 (2010).
16. Nayler, O., Hartmann, A.M. & Stamm, S. *J. Cell Biol.* **150**, 949–962 (2000).
17. Zhang, B. *et al. Int. J. Gynecol. Cancer* **20**, 492–499 (2010).
18. Hirschfeld, M. *et al. Mol. Carcinog.* doi:10.1002/mc.22045 (13 June 2013).
19. Hafner, M. *et al. Cell* **141**, 129–141 (2010).
20. Jacobs, S.A. & Khorasanizadeh, S. *Science* **295**, 2080–2083 (2002).
21. Wang, G.G. *et al. Nature* **459**, 847–851 (2009).
22. Zou, X., Ma, W., Solov'yov, I.A., Chipot, C. & Schulten, K. *Nucleic Acids Res.* **40**, 2747–2758 (2012).
23. Marcotrigiano, J., Gingras, A.C., Sonenberg, N. & Burley, S.K. *Cell* **89**, 951–961 (1997).

Acknowledgments

We would like to thank J. Walker for assistance in structure determination. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Boehringer Ingelheim, the Canada Foundation for Innovation, the Canadian Institutes for Health Research, Genome Canada through the Ontario Genomics Institute (OGI-055), GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda and the Wellcome Trust (092809/Z/10/Z to J.M.). C.H. is supported by the Howard Hughes Medical Institute as an investigator. The Mass Spectrometry Facility of the University of Chicago is funded by the National Science Foundation (CHE-1048528).

Author contributions

C.X. and J.M. conceived the project; C.X. performed the structural and binding experiments with assistance from K.L., W.T. and Y.L.; X.W. and I.A.R. conducted the PAR-CLIP experiment of YTHDC1; Z.L. and X.W. analyzed the PAR-CLIP data. C.X., X.W., C.H. and J.M. wrote the manuscript. All authors contributed to editing the manuscript. C.H. and J.M. supervised the project.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to C.X., C.H. or J.M.

ONLINE METHODS

Cloning, expression and purification of human YTHDC1 YTH domain for crystallography and ITC. The YTH domains of humanYTHDC1 (residues 345–509) were subcloned into pET28a-MHL vector. The recombinant YTHDC1 YTH domain was overexpressed at 18 °C with N-terminal His tag in *Escherichia coli* BL21 (DE3) Codon plus RIL (Stratagene) in the presence of 1 mM IPTG. The overnight cell cultures were harvested by centrifugation and dissolved in the lysis buffer containing 20 mM Tris-HCl, pH 7.5, 400 mM NaCl, 0.5 mM PMSF. The cells were lysed by sonication. Supernatant was collected after centrifugation at 16,000*g* for 40 min and then applied to Ni-NTA resin (Qiagen). The target protein was washed with lysis buffer plus 30 mM imidazole and then eluted in a buffer containing 20 mM Tris-HCl, pH 7.5, 400 mM NaCl and 0.5 M imidazole. Tobacco etch virus (TEV) protease was added to remove the N-terminal tag of the recombinant protein and dialyzed with lysis buffer overnight. The mixture was applied to another Ni-NTA resin to remove the protease and uncleaved proteins. Then the cleaved recombinant proteins were further purified by Superdex 75 gel filtration (GE Healthcare) in a buffer containing 10 mM Tris-HCl, pH 7.5, 150 mM NaCl and 1 mM DTT. The purified protein was concentrated to 20 mg/ml and stored at −80 °C.

The mutants of YTHDC1 YTH domain were cloned using the site-directed mutagenesis kits (Invitrogen) and were expressed and purified in the same way as the wild type (**Supplementary Figs. 11** and **12**).

Crystallization, data collection and structure determination. For crystallization of the YTHDC1 YTH domain (residues 345–509), 1 µl protein (15 mg/ml) was mixed with 1 µl crystallization buffer using the hanging drop vapor diffusion method at 18 °C. The diffracting crystal was crystallized in a buffer containing 0.1 M Bis-Tris, pH 6.5, 0.2 M ammonium sulfate and 25% PEG 3350. For crystallization of the YTHDC1 (345-509)-GG(m^{6}A)CU complex, 10 mg/ml protein (final concentration) was mixed with the modified 5-mer RNA GG(m^{6}A)CU (Thermo Fisher Scientific, Inc.) in a molecular ratio of 1:2. The mixture was incubated on ice for 0.5 h before crystallization. The RNA complex was crystallized under the same conditions as the apo-YTHDC1.

Diffraction data were collected using a copper rotating anode X-ray source, integrated with XDS[24] and scaled with POINTLESS/AIMLESS[25]. In the case of the YTHDC1-RNA complex, data were also integrated and scaled in HKL-3000 for use in molecular replacement and early refinement steps. Structures were solved by molecular replacement with PHASER[26]. A crystal structure isomorphous to the YTHDC1-apo crystal, below referred to as 'crystal form A', was solved with a search model derived from PDB code 2YUD. Model coordinates of crystal form A served as a search model for the YTHDC1–RNA complex structure. RNA geometry restraints were supplemented with JLIGAND[27]. The current models were iteratively rebuilt, refined and validated with COOT, REFMAC and MOLPROBITY[28,29], respectively. A summary of crystallographic statistics is shown in **Supplementary Table 2**. An m$|F_o|$ – D$|F_c|$ map was calculated within the PHENIX program suite after omission of RNA and unassigned atoms from the complex model and subsequent simulated annealing.

ITC measurements. All ITC measurements were recorded at 25 °C using a MicroCal ITC200 (GE Healthcare). All RNAs used for ITC binding experiments were purchased from Thermo Fisher Scientific except the unmodified GGACU, which was purchased from Integrated DNA Technologies, Inc. The purity of all purchased RNAs was > 90%. All proteins and RNAs are dialyzed or dissolved in the same buffer containing 20 mM Tris, pH 7.5, 150 mM NaCl before use. 10–15 injections were performed by injecting 2 µl 400–700 µM RNAs into a sample well containing 10–60 µM of proteins. The concentration of the proteins and RNAs were estimated with absorbance spectroscopy using the extinction coefficients OD_{280nm} and OD_{260nm}, respectively. Binding isotherms were plotted, analyzed and fitted in a one-site binding model by Origin Software (MicroCal Inc.) after subtraction of respective controls. The K_d, entropy, enthalpy and Gibbs free energy as well as their deviations were also calculated using Origin Software (MicroCal Inc.) during the ITC curve fitting in a one-site model (**Table 1** and **Supplementary Table 1**).

Plasmid construction and protein expression for PAR-CLIP and *in vitro* IP. C-terminal Flag-tagged YTHDC1 was cloned from commercial cDNA (Open Biosystems, clone ID: 5541053) into vector pcDNA 3.0 (EcoRI, XhoI; forward primer: CGTCAGAATTCATGGCGGCTGACAGTC; reverse primer: GGCATCTCGAGTTACTTGTCATCGTCGTCCTTGTAATCTCTTCTATATCGACCTCTCTC). Plasmids with high purity for mammalian cell transfection were prepared with a Maxiprep kit (Qiagen). Recombinant His-tagged YTH domain of YTHDC1 was constructed by subcloning into vector pET-28a (NdeI, XhoI; forward primer: CGTCACATATGCAAACCAGTAAACTCAAATATGTGC; reverse primer: GGCATCTCGAGTCAGTGACGCATTTTATGAATGACCTG). The resulting clones were transfected into the *E. coli* strain BL21 and expression was induced at room temperature with 1 mM IPTG for 20 h. The pellet collected from 2 l bacteria culture was then lysed in 30 ml lysis buffer (20 mM Tris, pH 7.5, 200 mM NaCl, 0.1% (v/v) Triton X-100) and sonicated for 10 min. After removing cell debris by centrifuge at 12,000 r.p.m. for 30 min, the supernatant was purified with a Ni-NTA cartridge (Qiagen, 5 ml) following the manufacturer's instructions. The crude products were further purified by gel-filtration chromatography in GF buffer (10 mM Tris, pH 7.5, 200 mM NaCl, 3 mM DTT). The yield was around 10 mg per liter of bacterial culture.

Gel shift binding assay and *in vitro* RNA IP. Gel shift binding assay and *in vitro* RNA IP were conducted with recombinant His-tagged YTH domain of YTHDC1 following previous reported procedure[14]. The procedure of *in vitro* RNA CLIP is the same as *in vitro* RNA IP, except for the following modifications: (i) before immunoprecipitation with anti-His magnetic beads, the RNA-protein mixture was transferred to a UV-transparent 96-well plate (placed on ice and without cover) and cross-linked three times with 0.15 J/cm^2 of 254 nm UV light each time in a Stratalinker 2400 (Stratagene); (ii) after UV cross-linking, the RNA-protein mixture was quenched on ice and subjected to RNase T1 digestion (1 U/µl RNase T1 for 8 min at 22 °C); (iii) after IP and washing, the RNA was detached from magnetic beads by protease K digestion (1 mg/ml final concentration, 50 °C, 30 min), and the RNA was further recovered by Zymo RNA Clean and Concentrator.

Mammalian cell culture and plasmid transfection. Human HeLa cells used in this study were purchased from ATCC (CCL-2) and cultured in DMEM (Gibco, 11965) medium supplemented with 10% FBS and 1% 100× Pen Strep (Gibco). Transfection was achieved using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol.

PAR-CLIP. PAR-CLIP was performed based on a previously reported procedure[30] with the following modifications. 4 × 15 cm plates of cells were transfected with Lipofectamine 2000 according to manufacturer's protocol. After 6 h, the medium was replaced, and cells were cultured in fresh medium supplemented with 200 µM 4-thiouridine (Sigma) overnight. The first RNase T1 digestion was conducted at 1 U/µl RNase T1 for 8 min. For the second digestion, the concentration of RNase T1 was reduced from 100 U/µl to 20 U/µl. Following dephosphorylation, one-tenth of the sample was partitioned for ^{32}P labeling, and the remaining volume was treated with 1 U/µl T4 PNK at 37 °C for 10 min, followed by addition of ATP to 1 µM for 30 min at 37 °C. This sample was then washed and digested with proteinase K. RNA was purified using Zymo RNA Clean and Concentrator before library construction using the Tru-seq small RNA sample preparation kit (Illumina). The cDNA library was sequenced by Illumina Hiseq2000 with a single-end 50-bp read length. The adaptors were trimmed by using FASTX-Toolkit[31]. The deep sequencing data were mapped to Human genome version hg19 by Tophat version 2.07 (ref. 32) without any gaps and allowed for at most two mismatches. PAR-CLIP data were analyzed by PARalyzerv1.1 (ref. 33) with default settings. The motif of PAR-CLIP peaks was generated by HOMER motif analysis. The motif with the lowest *P* value is shown in **Figure 1b**. The distances between GGAC and T to C mutation sites were calculated by setting the A of GGAC as position zero (**Supplementary Fig. 3e**).

24. Kabsch, W. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
25. Evans, P.R. & Murshudov, G.N. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
26. McCoy, A.J. *et al. J. Appl. Crystallogr.* **40**, 658–674 (2007).
27. Lebedev, A.A. *et al. Acta Crystallogr. D Biol. Crystallogr.* **68**, 431–440 (2012).
28. Murshudov, G.N. *et al. Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
29. Chen, V.B. *et al. Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
30. Hafner, M. *et al. J. Vis. Exp.* **2** doi:10.3791/2034 (2010).
31. Pearson, W.R., Wood, T., Zhang, Z. & Miller, W. *Genomics* **46**, 24–36 (1997).
32. Trapnell, C., Pachter, L. & Salzberg, S.L. *Bioinformatics* **25**, 1105–1111 (2009).
33. Corcoran, D.L. *et al. Genome Biol.* **12**, R79 (2011).

doi:10.1038/nchembio.1654

CORRIGENDUM

Structural basis for selective binding of m^{6}A RNA by the YTHDC1 YTH domain

Chao Xu, Xiao Wang, Ke Liu, Ian A Roundtree, Wolfram Tempel, Yanjun Li, Zhike Lu, Chuan He & Jinrong Min

Nat. Chem. Biol. **10**, 927–929 (2014); published online 21 September 2014; corrected after print 19 August 2015

In the version of this Brief Communication initially published, the email address for corresponding author, Chao Xu, was incorrect. It should be listed as chaor.xu@utoronto.edu. This error has been corrected in the PDF and HTML versions of the article.

ERRATUM

Structural basis for selective binding of m^{6}A RNA by the YTHDC1 YTH domain

Chao Xu, Xiao Wang, Ke Liu, Ian A Roundtree, Wolfram Tempel, Yanjun Li, Zhike Lu, Chuan He & Jinrong Min

Nat. Chem. Biol. **10**, 927–929 (2014); published online 21 September 2014; corrected after print 19 August 2015 and 11 November 2015

In the version of this Brief Communication initially published, as well as in a corrected version of 19 August 2015, incorrect e-mail addresses were given for corresponding author Chao Xu. The correct address is chaor.xu@utoronto.ca. This error has been corrected in the PDF and HTML versions of the article.