Supplemental Data
for

Crystal structure and RNA binding properties of the RRM/AlkB domains in human ABH8, an enzyme catalyzing tRNA hypermodification *

Chiara Pastore [1,2], Irini Topalidou [1], Farhad Forouhar [1,2], Amy C. Yan [3], Matthew Levy [3], and John F. Hunt [1,2].

[1] From the Department of Biological Sciences, 702A Fairchild Center, MC2434, Columbia University, New York, NY 10027, USA;
[2] Northeast Structural Genomics Consortium;
and
[3] Department of Biochemistry, Albert Einstein College of Medicine, Michael F. Price Center, 1301 Morris Park Avenue, Room 519, Bronx, NY 10461, USA.

*Running Title: *Structure and RNA binding of the RRM/AlkB domains in ABH8.*

To whom correspondence may be addressed: John F. Hunt, Department of Biological Sciences, 702A Fairchild Center, MC2434, Columbia University, New York, NY 10027, USA, Tel: (212)-854-5443; Fax (212)-865-8246; E-mail: jfhunt@ biology.columbia.edu.

Keywords: tRNA modification; DNA repair; AlkB; x-ray crystallography; SELEX; *C. elegans*.

Friday, October 28, 2011 @ 4:14 PM.

SUPPLEMENTAL METHODS

Molecular biology. Because a complete human cDNA was not available for ABH8, the gene as reported in ENSEMBL (http://useast.ensembl.org/index.html) was reconstructed with native codon usage using maxi-primer PCR (1). PCR products were obtained by amplification of the coding sequence with Hot-Start Pfu polymerase (Stratagene, La Jolla, CA), with the oligonucleotide primer sequences filling gaps in the commercially available cDNAs. The full-length gene bearing a C-terminal TEV-protease recognition sites (ENLYFQG) was inserted into the NdeI and XhoI sites of the pET26b vector (EMD Biosciences, Darmstadt, Germany), resulting in addition of a C-terminal hexahistidine tag. Constructs containing the MTase domain of ABH8 could only be obtained in soluble form when co-expressed with the Trm112 accessory protein (2,3). The ABH8-derived construct and full-length human Trm112 from I.M.A.G.E. clone 3163327 (http://image.hudsonalpha.org/) were cloned into the first and second multiple cloning sites of the pET-Duet vector (EMD Biosciences, Gibbstown, NJ) between the SalI/NotI and NdeI/XhoI sites, respectively. Site-directed mutagenesis was performed using the QuickChange kit (Qiagen, Valencia, CA).

Construction and analysis of C14B1.10- GFP fusions in C. elegans. For plasmid TU#1001 ($P_{C14B1.10}$*gfp*), a DNA fragment starting 533 bp upstream of the start codon of the C14B1.10 gene and extending to the end of its first intron was amplified by PCR (using the *C.e.* ABH8 5' and 3'-i1 primers in Table S1) and cloned into the PstI and BamHI sites in vector pPD95.75 (ftp://www.ciwemb.edu/pub/FireLabInfo/FireLabVectors). For plasmid TU#1002 (C14B1.10::GFP), a DNA fragment with the same start site but extending to the last codon of the C14B1.10 gene was amplified by PCR (using the *C.e.* ABH8 5' and 3'-cs primers in Table S1) and cloned into the same restriction sites in pPD95.75. Wild type (N2) animals were microinjected with 20 ng/µl of the relevant plasmid and 25 ng/µl of the dominant roller plasmid (pRF4; (11)). GFP fluorescence was observed using a Zeiss Axiophot II microscope.

SELEX. *In vitro* selection of a high affinity RNA ligand for the 1-354 construct was achieved via ten iterative rounds of binding, reverse transcription (RT), PCR amplification, and *in vitro* transcription. The initial template sequence (Table S1) contained a central region with 30 randomized nucleotides flanked by 26mer and 24mer constant regions at the 5' and 3' ends, respectively. For the initial round of selection, ~200 pmols of the template, comprising ~10^{14} unique sequences, was heated to 65 ºC for 5 minutes and then cooled slowly to room temperature in RNA Binding Buffer (150 mM NaCl, 1 mM $MgCl_2$, 20 mM HEPES, pH 7.5) prior to equilibration with 5 µg of protein for 30 minutes at room temperature. Protein-bound RNA was selected by passing this mixture through a 0.45 µm Hawp filter (Millipore Inc., Billerica, MA) previously equilibrated in Binding Buffer. The filter with bound RNA was washed with 1 ml of RNA Binding Buffer prior to heating it to 95 ºC for 2 minutes in 200 µl of elution buffer (8 M Urea, 300 mM NaOAc, 3 mM EDTA). The elutions from two such extractions were combined, extracted with phenol/chloroform, precipitated with ethanol, re-suspended in 20 µl of water, and then diluted to a final volume of 50 µl containing 20 µM SELEX reverse primer (Table S1), 4 mM dNTPs, 10 mM DTT, and 1X First Strand Buffer. Reverse transcription was conducted using MMLV RT (Invitrogen, Carlsbad, CA) at 37 ºC for 1 hour. The resulting single-stranded DNA was amplified by PCR using the appropriate primers (Table S1). Following amplification, the double-stranded DNA library was precipitated with ethanol and subsequently used as a template for *in vitro* transcription reactions conducted in a 20 µl volume containing 0.2–1.0 µg PCR product, T7 RNA polymerase produced in the Levy lab, 4 mM NTPs, 1 mM spermidine, 0.01 % Triton X-100, 5% (w/v) PEG8000, 5 mM DTT, 40 mM Tris, pH 8.3. Following incubation at 37 ºC for 2 hours, residual DNA was eliminated by treatment with 1 unit of DNase I (Epicentre Biotechnologies, Madison, WI) at 37 ºC for 15 minutes. The transcribed RNA was isolated on a denaturing 8% polyacrylamide gel containing 7 M urea, identified by UV shadowing, excised, and extracted overnight in 450 µl of 300 mM NaOAc. The RNA in this sample was recovered by ethanol precipitation and, following UV quantification, used for subsequent rounds of selection.

The selection stringency was progressively increased by decreasing the amount of protein in the

RNA-binding reaction (1 μg in cycle 2 and 0.2 μg in cycles 3-10) and increasing the volume of the washing steps (3, 5, 6, 10, and 10 ml in cycles 6-10). Starting in round 2, RNA sequences with affinity for the filter material were depleted from the pool with a negative selection step in which the transcribed RNA was passed through the protein-binding filter prior to addition of protein. Protein and RNA integrity were both checked after rounds 0 and 4 using polyacrylamide gel electrophoresis, and the affinity of the protein for the RNA pool was checked via filter-binding assays after rounds 4, 7, 9 and 10. Because no increase in affinity was observed in round 10, the reverse-transcribed DNA from round 9 was amplified with Taq polymerase and cloned into pCR4-TOPO (Invitrogen, Carlsbad, CA). This plasmid pool was transformed into NovaBlues cells (Novagen, Merck KGaA, Darmstadt, Germany), plated on LB agar with kanamycin, and plasmids from 20 of the resulting colonies were sequenced. RNA aptamers transcribed from each of these clones were evaluated using filter-binding assays, and one was selected for further characterization based on its high protein-binding affinity. Progressive truncations were used to demonstrate that sequences from the 3' constant region of this aptamer contribute to its protein-binding affinity, while its 5' constant region is dispensable. These experiments identified aptamer ABH8-2.2 (Table S2) as the minimal RNA species retaining full affinity.

Thermal denaturation assays. The thermal denaturation of proteins was monitored indirectly using a fluorescence-based thermal shift assay (10). Protein (10 μM) was mixed with SYPRO Orange dye (Invitrogen, Carlsbad, CA, 5x final concentration) and additives (75 μM, 1mM or 10 mM of $MnCl_2$, 10 mM 2OG, 5 mM EDTA) in a buffer containing 150 mM NaCl, 100 mM K-HEPES, pH 7.5. Fluorescence at 590 nm was monitored using 492 nm excitation in the Mx3005P Real-Time PCR system (Stratagene) as the temperature was raised 1 ˚C per min. The fluorescence level at each temperature is plotted relative to baselines extrapolated from the pre- and post-transition regions using linear least-squares regressions in the program EXCEL (Microsoft, Inc., Redmond, WA).

Curve-fitting. Binding isotherms were fit to a quadratic binding equation using the Marquardt and Levenberg algorithm as implemented in PRISM 5 (GraphPad, San Diego, CA):

$$Y = B_0 + 0.5\Delta B\left\{\left(1+\frac{(K_d + x)}{sR_t}\right) - \sqrt{\frac{(K_d + x)^2}{sR_t^{\,2}} + 2\frac{(K_d - x)}{sR_t} + 1}\right\}$$

In this equation, B_0 is the observed signal in the absence of protein, ΔB is the change in signal at saturation, sR_t is the concentration of the RNA ligand, and K_d is the dissociation constant. To improve the stability of curve-fitting, the value of sR_t was fixed at the measured RNA concentration used in the assays. The mean and sample variance from triplicate measurements were fit for the filter-binding data, with the fully saturated signal constrained to be greater than 0 and less than 1, while the mean and sample variance from ~60 seconds of observation were fit for the anisotropy data. For some anisotropy assays in which saturation was not observed clearly experimentally, the fully saturated signal ($r_{saturated}$) was fixed at a value estimated by reference to that in an assay showing clear saturation using a form of the Perrin equation:

$$r_{saturated} = r_0 * \theta/(\theta+\tau)$$

The value of r_0 was set to 0.4 and the fluorescence lifetime τ to 3.8 ns, and the apparent rotational correlation coefficient θ linearly scaled according to the predicted molecular weight of the 1:1 protein-RNA complex. This scaling generally agreed with the observed value of $r_{saturated}$ for assays showing full saturation.

Molecular graphics. Except for Fig. 6B in the main text which was prepared using GRASP2 (12), all molecular graphics figures in this paper were prepared using the program PYMOL (PyMOL Molecular Graphics System, Schrödinger, LLC).

SUPPLEMENTAL REFERENCES

1. Heckman, K. L., and Pease, L. R. (2007) *Nat Protoc* **2**, 924-932
2. Songe-Moller, L., van den Born, E., Leihne, V., Vagbo, C. B., Kristoffersen, T., Krokan, H. E., Kirpekar, F., Falnes, P. O., and Klungland, A. (2010) *Mol Cell Biol* **30**, 1814-1827
3. Fu, D., Brophy, J. A., Chan, C. T., Atmore, K. A., Begley, U., Paules, R. S., Dedon, P. C., Begley, T. J., and Samson, L. D. *Mol Cell Biol* **30**, 2449-2459
4. Otwinowski, Z. a. M., W. (1997) *Methods in Enzymology* **276**, 307-326
5. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J Appl Crystallogr* **40**, 658-674
6. Yu, B., Edstrom, W. C., Benach, J., Hamuro, Y., Weber, P. C., Gibney, B. R., and Hunt, J. F. (2006) *Nature* **439**, 879-884
7. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr* **54**, 905-921
8. McRee, D. E. (1999) *J Struct Biol* **125**, 156-165
9. Vagin, A. A., and Isupov, M. N. (2001) *Acta Crystallogr D Biol Crystallogr* **57**, 1451-1456
10. Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., and Salemme, F. R. (2001) *J Biomol Screen* **6**, 429-440
11. Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991) *EMBO J* **10**, 3959-3970
12. Petrey, D. a. H., B. (2003) *Methods in Enzymology* **374**, 492-509
13. Drenth, J. (1994) *Principles of Protein X-Ray Crystallography* Springer, New York
14. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993) *J. App. Cryst.* **26**, 283-291
15. Fu, Y., Dai, Q., Zhang, W., Ren, J., Pan, T., and He, C. (2010) *Angew Chem Int Ed Engl* **49**, 8885-8888
16. van den Born, E., Vagbo, C. B., Songe-Moller, L., Leihne, V., Lien, G. F., Leszczynska, G., Malkiewicz, A., Krokan, H. E., Kirpekar, F., Klungland, A., and Falnes, P. O. *Nat Commun* **2**, 172
17. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) *Bioinformatics* **23**, 2947-2948
18. Han, Z., Niu, T., Chang, J., Lei, X., Zhao, M., Wang, Q., Cheng, W., Wang, J., Feng, Y., and Chai, J. (2010) *Nature* **464**, 1205-1209
19. Sundheim, O., Vagbo, C. B., Bjoras, M., Sousa, M. M., Talstad, V., Aas, P. A., Drablos, F., Krokan, H. E., Tainer, J. A., and Slupphaug, G. (2006) *EMBO J* **25**, 3389-3397
20. Yang, C. G., Yi, C., Duguid, E. M., Sullivan, C. T., Jian, X., Rice, P. A., and He, C. (2008) *Nature* **452**, 961-965
21. Auweter, S. D., Fasan, R., Reymond, L., Underwood, J. G., Black, D. L., Pitsch, S., and Allain, F. H. (2006) *EMBO J* **25**, 163-173
22. Skrisovska, L., Bourgeois, C. F., Stefl, R., Grellscheid, S. N., Kister, L., Wenter, P., Elliott, D. J., Stevenin, J., and Allain, F. H. (2007) *EMBO Rep* **8**, 372-379
23. Hobor, F., Pergoli, R., Kubicek, K., Hrossova, D., Bacikova, V., Zimmermann, M., Pasulka, J., Hofr, C., Vanacova, S., and Stefl, R. (2011) *J Biol Chem* **286**, 3645-3657
24. Fribourg, S., Gatfield, D., Izaurralde, E., and Conti, E. (2003) *Nat Struct Biol* **10**, 433-439
25. Holm, L., and Rosenstrom, P. (2010) *Nucleic Acids Res* **38**, W545-549

Table S1. ***Buffers used for protein purification.*** [a]

Construct	Lysis	Equilibration	Wash 1	Wash 2	Elution	Dialysis	Gel filtration
1-133 **12-125** **25-125**	500 mM NaCl 10 mM IMDZ 0.2% Tween [d] 2 mM β-ME [e] 0.5 mg/ml LZ [f] A-Pr-T [h] 20 mM Tris-Cl, pH 7.0	500 mM NaCl 10 mM IMDZ 0.2% Tween [d] 2 mM β-ME [e] 20 mM Tris-Cl, pH 7.0	500 mM NaCl 10 mM IMDZ 2 mM β-ME [e] 20 mM Tris-Cl, pH 7.0	1 M NaCl 10 mM IMDZ 2 mM β-ME [e] 20 mM Tris-Cl, pH 7.0	500 mM NaCl 300 mM IMDZ 2 mM β-ME [e] 20 mM Tris-Cl, pH 7.0	150 mM NaCl 10 mM IMDZ 10 mM β-ME [e] 5% glycerol 20 mM Tris-Cl, pH 7.0	150 mM NaCl 5% glycerol Tris20, pH 7
132-354	10 mM Tris-Cl, pH 7.5	500 mM NaCl 5 mM IMDZ [b] Tris20, pH 7.5	500 mM NaCl 5 mM IMDZ Tris10, pH 7.5	1 M NaCl 5 mM IMDZ Tris20, pH 7.5	500 mM NaCl 100 mM IMDZ Tris20, pH 7.5	100 mM NaCl 0.5 mM EDTA Tris20, pH 7.5	100 mM NaCl Tris20, pH 8
1-354 **1-354** [j] **25-354** **25-354-$_6$his**	500 mM NaCl 10 mM IMDZ 20 mM AS [c] 100 μM $ZnSO_4$ 0.2% Tween [d] 2 mM β-ME [e] 0.5 mg/ml LZ [f] 2 mM $MnCl_2$ 1 mM 2OG [g] A-Pr-T [h] 20 mM Tris-Cl, pH 7.0	500 mM NaCl 10 mM IMDZ 20 mM AS 100 μM $ZnSO_4$ 0.2% Tween 2 mM β-ME 2 mM $MnCl_2$ 1 mM 2OG 20mM Tris-Cl, pH 7.0	500 mM NaCl 10 mM IMDZ 20 mM AS 100 μM $ZnSO_4$ 2 mM β-ME 2 mM $MnCl_2$ 1 mM 2OG Tris-Cl 20mH, pH 7.0	1 M NaCl 10 mM IMDZ 20 mM AS 100 μM $ZnSO_4$ 2 mM β-ME 2 mM $MnCl_2$ 1 mM 2OG Tris-Cl 20mH, pH 7.0	1 M NaCl 300 mM IMDZ 20 mM AS 100 μM $ZnSO_4$ 2 mM β-ME 2 mM $MnCl_2$ 1 mM 2OG Tris-Cl 20mH, pH 7.0	150 mM NaCl 20 mM IMDZ 20 mM AS 100 μM $ZnSO_4$ 10 mM β-ME Tris-Cl 20mH, pH 7.0	150 mM NaCl 5% glycerol Tris10, pH 7
352-663 [k] **126-663** [k] **1-663** [k]	150 mM NaCl 5 mM IMDZ 10 % glycerol 0.5 % NP-40 [i] 0.5 mg/ml LZ A-Pr-T 50 mM Tris-Cl, pH 7.5	150 mM NaCl 5 mM IMDZ 10 % glycerol 0.5 % NP-40 50 mM Tris-Cl, pH 7.5	500 mM NaCl 5 mM IMDZ 50 mM Tris-Cl, pH 7.5		150 mM NaCl 100 mM IMDZ 2 mM β-ME 50 mM Tris-Cl pH 7.5	150 mM NaCl 5% glycerol Tris20 pH 7.5	150 mM NaCl 5% glycerol Tris20 pH 7.5

[a] **Tris10, Tris20 and Tris50 buffers contained 1 mM DTT and respectively 10 mM, 20 mM ,or 50 mM Tris-Cl.**

[b] **IMDZ: abbreviation for imidazole.**

[c] **AS: abbreviation for ammonium sulfate -- $(NH_4)_2SO_4$.**

[d] **0.2% (w/v) Tween 20 non-ionic detergent.**

[e] **β-ME: abbreviation for β-mercaptoethanol.**

[f] **LZ: abbreviation for lysozyme.**

[g] **2OG: abbreviation for 2-oxoglutarate.**

[h] **A-Pr-T: abbreviation for Roche mini EDTA-free antiprotease tablets, 1 per 50 ml.**

[i] **0.5% (w/v) NP-40 non-ionic detergent.**

[j] **This construct contained the C341A/C349A double mutation in the C-terminal structural Zn-binding motif in the AlkB domain.**

[k] **These constructs were purified in complex with human Trm112.**

Table S2. ***Nucleotide sequences used in studies of ABH8.*** [a]

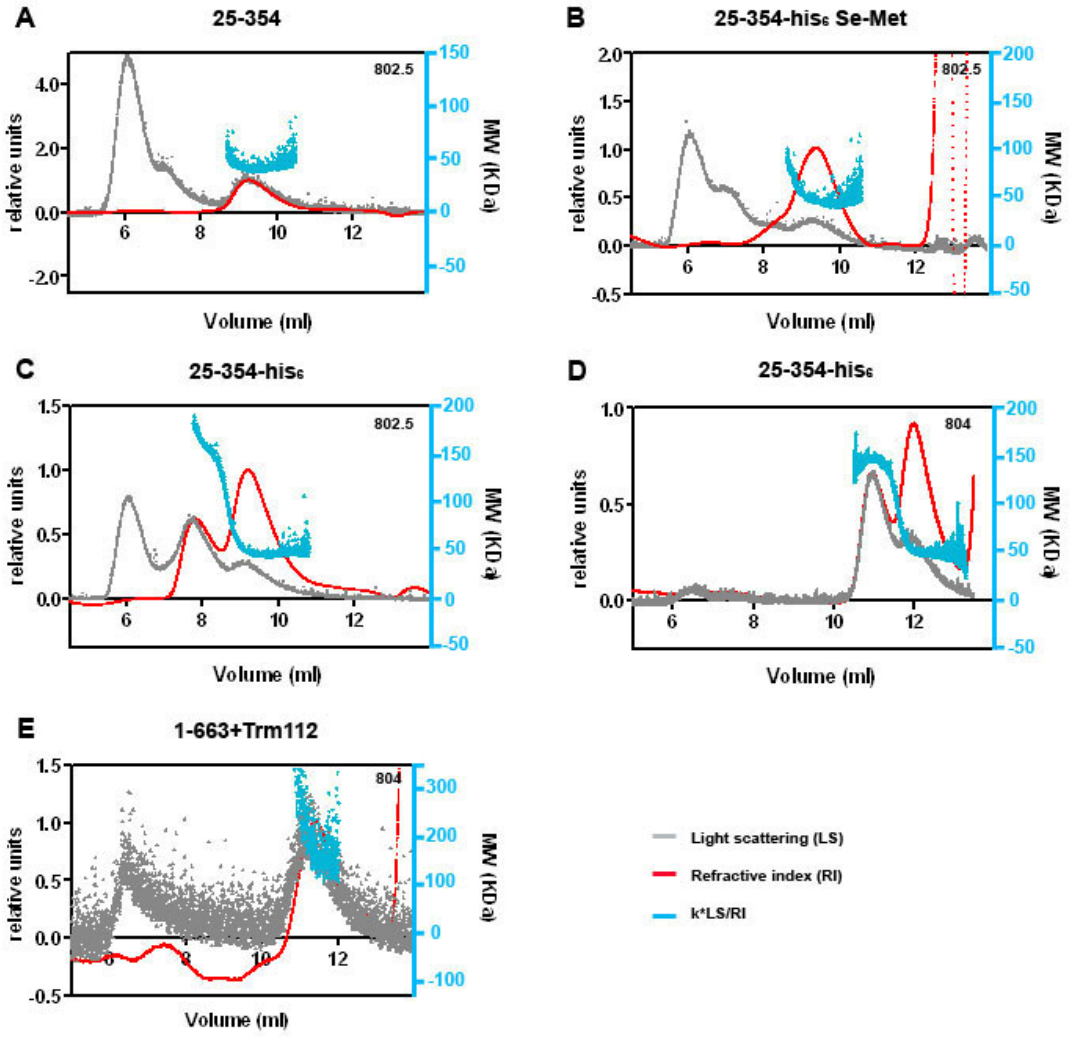
DNA primers [b]	
tRNA-Glu 5'	5'-**TAATACGACTCACTATA*GG***TCCCATATGGTCTAGCGGTTAGGA-3'
tRNA-Glu middle	5'-ATGGTCTAGCGGTTAGGATTCCTGGTTTTCACCCAGGTGGCCCGGGTT-3'
tRNA-Glu 3'	5'-TmTCCCATACCGGGAGTCGAACCCGGGCCACCTGGGT-3'
tRNA-Gly 5'	5'-**TAATACGACTCACTATA*GG***GCGTTGGTGGTATAGTGGTA-3'
tRNA-Gly middle	5'-GGTGGTATAGTG GTAAGCATAGCTGCCTTCCAAGCAGTTGACCCGGGTT-3'
tRNA-Gly 3'	5'-TmGCGTTGGCCGGGAATCGAACCCGGGTCAACTGCTT-3'
SELEX primer 5'	5'- GATAATACGACTCACTATAGGGAATGGATCCACATCTACGA-3'
SELEX primer 3'	5'-AAGCTTCGTCAAGTCTGCAGTGAA-3'
***C.e.* ABH8 3'**	5'- ACCCCTGCAGGGGTGCAAGTAGCGTGTCCTAG -3'
***C.e.* ABH8 5'-i1**	5'- CCAGGGATCCGCTGGAATAAAATTGTTTTCTTTAAA -3'
***C.e.* ABH8 5'-cs**	5'- AATTGGATCCCCAATTTTCTTCGCAATAATAATATAATTTC -3'
RNA species	
SELEX template	5'-GGGAAUGGAUCCACAUCUACGAAUUC$(N)_{30}$CACUGCAGACUUGACGAA GCUU-3'
aptamer ABH8-2.2	5'-CAAGCCUAAGACAAAUAAAAAACGGAUGAGUUCACUGCAGACUU-3'
17mer stem-loop	5'- GCUGCCUUCCAAGCAGU-3'
control 17mer	5'- ACCAUCACUUAAAAAAA-3'

[a] All sequences are given from 5' to 3' reading left to right.

[b] The T7 promoter sequence is shown in bold, sequences complementary to other primers are underlined. The italicized G nucleotides were added to increase T7 polymerase activity. The nucleotides "mT" and 'mG' represent 2′-*O*-methyl-deoxythymidine and 2′-*O*-methyl-deoxyguanosine, which are used to obtain transcripts with a uniform 3′-end.

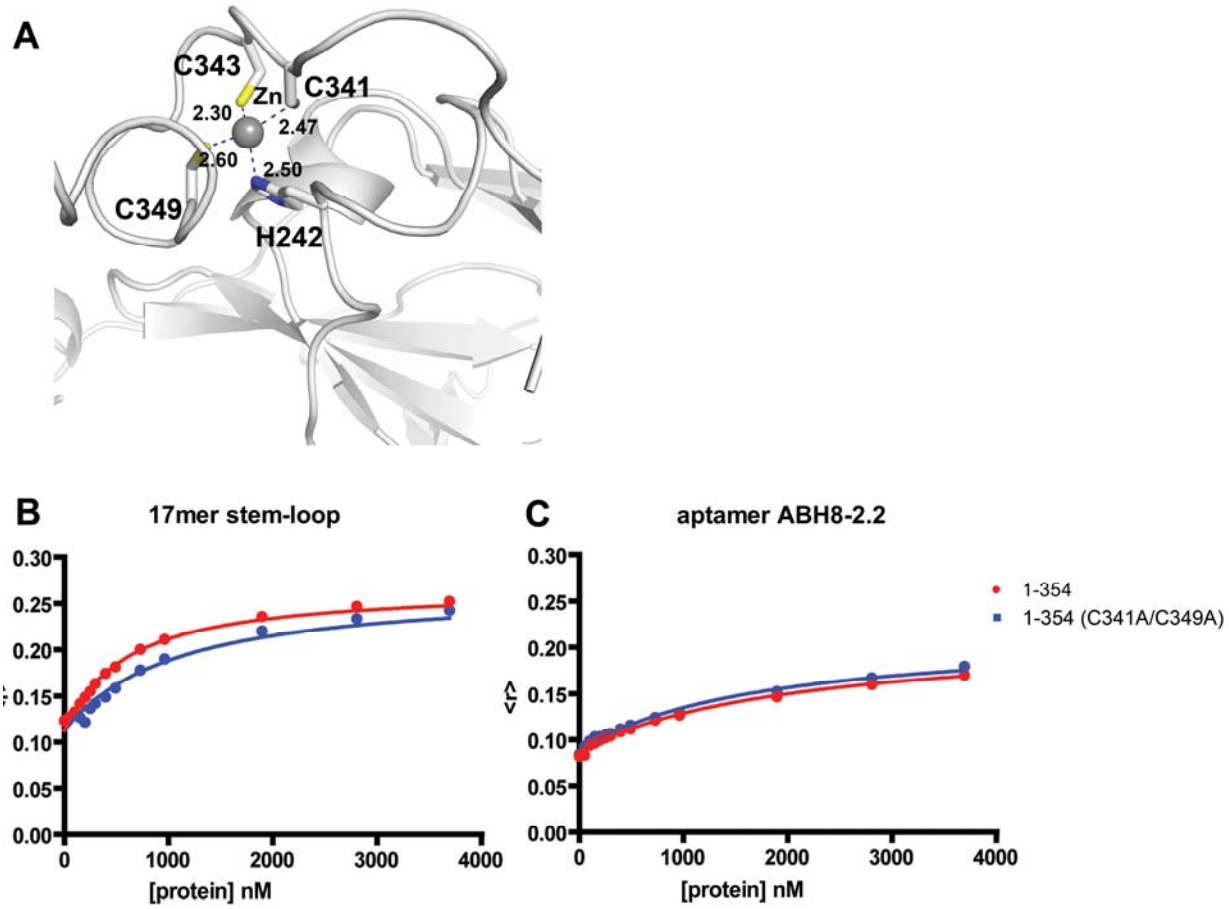
[c] The abbreviations *C.e.*, i1, and cs stand for "*C. elegans*", "intron 1" and "coding sequence", respectively.

Figure S1



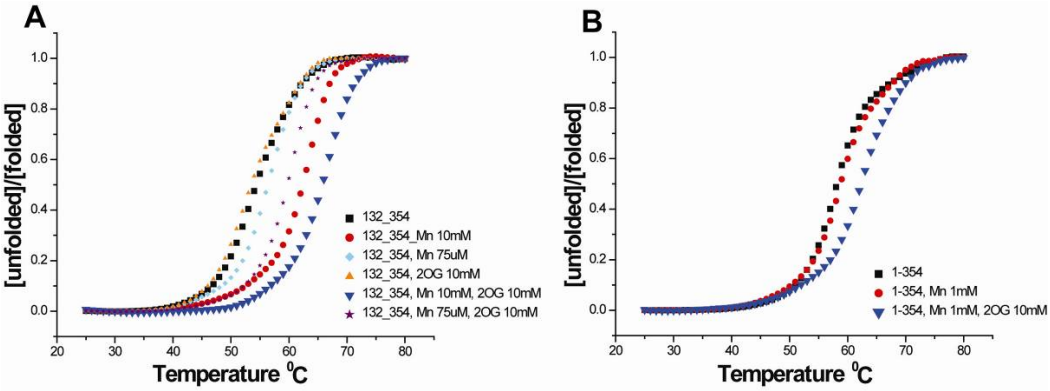
SUPPLEMENTAL FIGURE 1. **Hydrodynamic characterization of human ABH8 protein constructs using analytical gel-filtration chromatography monitored by static light-scattering.** The plots show refractive index (red) and 90˚ light-scattering (gray) signals from the column effluent on the vertical axis on the left and the inferred mass-averaged molecular weight (cyan) on the vertical axis on the right. The identity of the evaluated protein construct is indicated above each plot. The refractive index trace is normalized so that the value at the highest protein peak is ~1.0, while the light-scattering trace is normalized to match the height of the refractive index trace for one of the protein peaks. The mass-averaged molecular weight was calculated as the ratio of the 90˚ light-scattering signal to the refractive index signal at each point of the chromatogram, with a linear scale-factor (530.0) applied to match the results of Debye analyses at the top of the protein peaks. A 100 µl aliquot of each concentrated protein stock was injected onto a Shodex 802.5 or 804 analytic gel filtration column (as indicated at the top right in each plot) running at 4 ˚C and 0.5 ml/min in 150 mM NaCl, 1mM DTT, 5% glycerol and 20mM Tris-HCl. The buffers for the 802.5 and 804 columns were at pH 7.5 and pH 7.0, respectively. The mass of protein injected onto the column was 180 µg for 25-354, 230 µg for 25-354-his_6 Se-Met, 470 µg for 25-354-his_6, and 160 µg for 1-663+Trm112. Constructs not explicitly containing "his_6" in their names have had the tag cleaved off with TEV protease. The high noise in the refractive index and light-scattering traces for the 1-663+Trm112 construct is attributable to a poor mass recovery during chromatography.

Figure S2



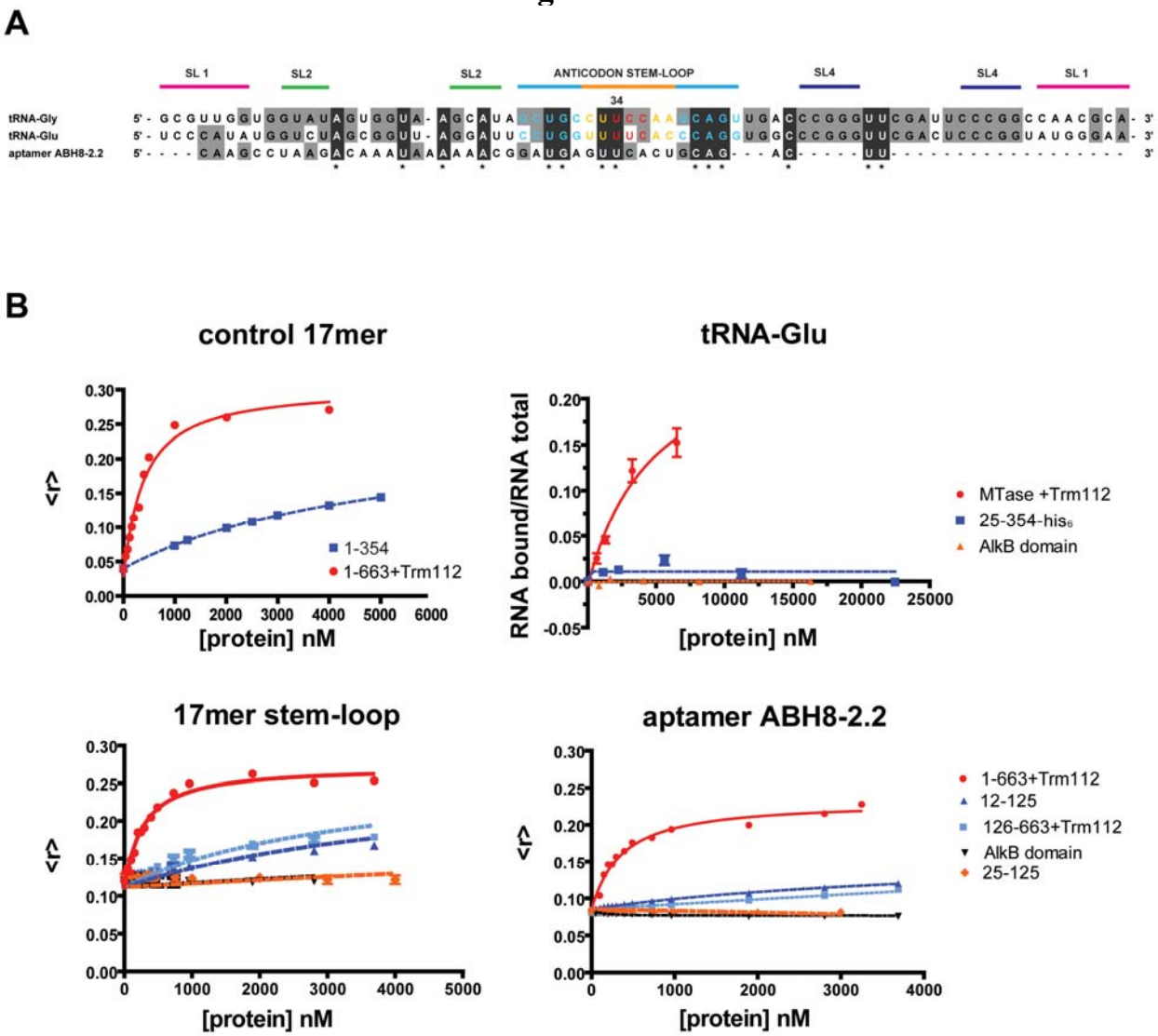
SUPPLEMENTAL FIGURE 2. **The C-terminal structural Zn-binding site does not participate in binding unmodified RNAs.** **(A)** Close-up view of the structural Zn-binding motif at the C-terminus of the AlkB domain in the crystal structure of the RRM/AlkB domains from ABH8 (Table 2 and Fig. 6 in the main text). The Zn(II) ion is chelated by three cysteine residues in the C-terminal extension of the AlkB domain and a histidine on the surface of the core of the domain. All of these residues are invariant in ABH8 orthologs (Fig. 2 in the main text). **(B,C)** Fluorescence anisotropy assays demonstrate that the structural Zn-binding motif does not contribute to binding fluorescently labeled synthetic RNAs *in vitro*. The wild-type 1-354 (red) and corresponding C341A/C349A double-mutant (blue) were titrated onto the 17mer stem-loop (panel **B**) or aptamer ABH8-2.2 (panel **C**) at 25 ºC in RNA Binding Buffer. Each point represents the average value of the anisotropy recorded over 60 seconds.

Figure S3



SUPPLEMENTAL FIGURE 3. **Mn(II)/2OG stabilize the folding of the isolated ABH8 AlkB domain.** Thermal denaturation assays starting at 25 ˚C were monitored using the fluorescent reporter SYPRO Orange were conducted in 150 mM NaCl, 50mM Hepes, pH 7.5. **(A)** Thermal denaturation of the 132-354 AlkB domain construct without any additions (black) or with the addition of 75 μM Mn(II) (light blue), 10 mM Mn(II) (red), 10 mM 2OG (orange), 75 μM Mn(II) plus 10 mM 2OG (purple stars), or 10 mM Mn(II) plus 10 mM 2OG (blue). **(B)** Thermal denaturation of the wild-type 1-354 RRM/AlkB double-domain construct without any additions (black) or with the addition of 1 mM Mn(II) (red) or 1 mM Mn(II) plus 10 mM 2OG (blue). The stabilization of the 132-354 AlkB domain construct by high concentrations of Mn(II) in the absence of 2OG could be attributable either to low-affinity metal-ion binding in the active site of the enzyme or to replacement of Zn(II) lost from the C-terminal structural Zn-binding site during purification of this construct.

Figure S4

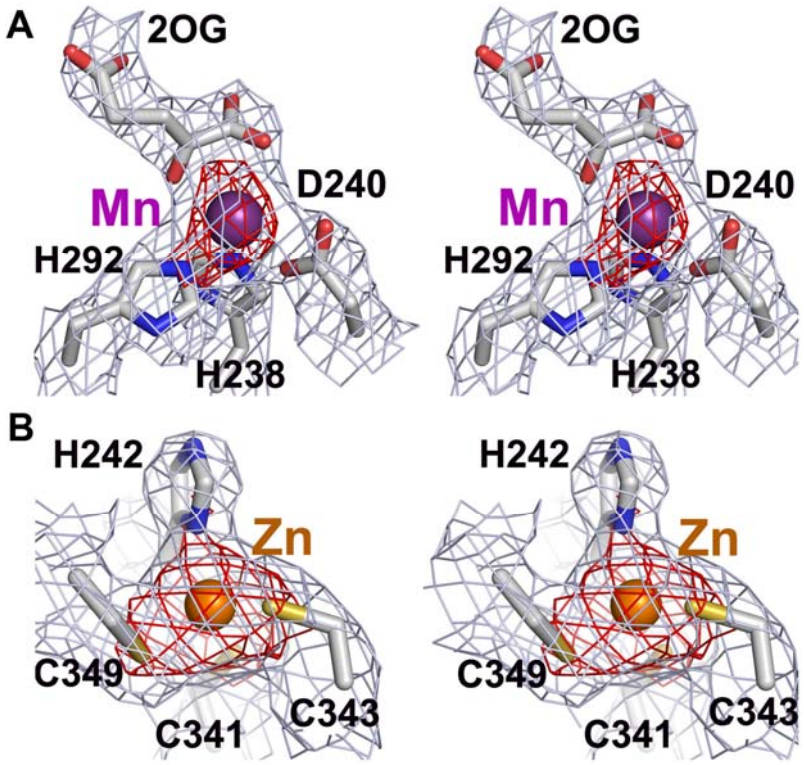


SUPPLEMENTAL FIGURE 4. **Sequence similarity of RNAs and their affinity for ABH8 protein constructs.** **(A)** Sequence alignment of aptamer ABH8-2.2 with human tRNAs. The aptamer shares 41% and 53% sequence identity, respectively, with tRNA-Glu and tRNA-Gly (both 72 bases long), including a perfect match in the anticodon loop of tRNA-Gly. The asterisks indicate the 14 bases in the 44 base aptamer that are identical to both tRNA-Glu and tRNA-Gly in this alignment. The bases in the anticodons are colored red, the other bases in the anticodon loop are colored orange, and the bases in the anticodon stem are colored blue. The base number of the wobble-codon uridine (34) is indicated above the sequence alignment. The other structural elements in the tRNA molecules are indicated above the alignment (with SL standing for stem-loop). The sequence alignment was generated using the program ClustalW (17) to align the aptamer to the 17 base anticodon stem-loop from tRNA-Gly (Table S2) (3) and the two tRNAs to one another. These two sub-alignments were manually combined and adjusted to maximize sequence identity in the regions flanking the anticodon stem-loop. **(B)** Additional RNA-binding experiments. Titrations in RNA-Binding Buffer were conducted and analyzed as described for the equivalent experiments in Fig. 5 in the main text. Fluorescent anisotropy assays (top left) were used to monitor the binding of a randomly chosen 5'-fluorescein-labeled control 17mer (Table S2) to either the

Figure S4 (cont.)

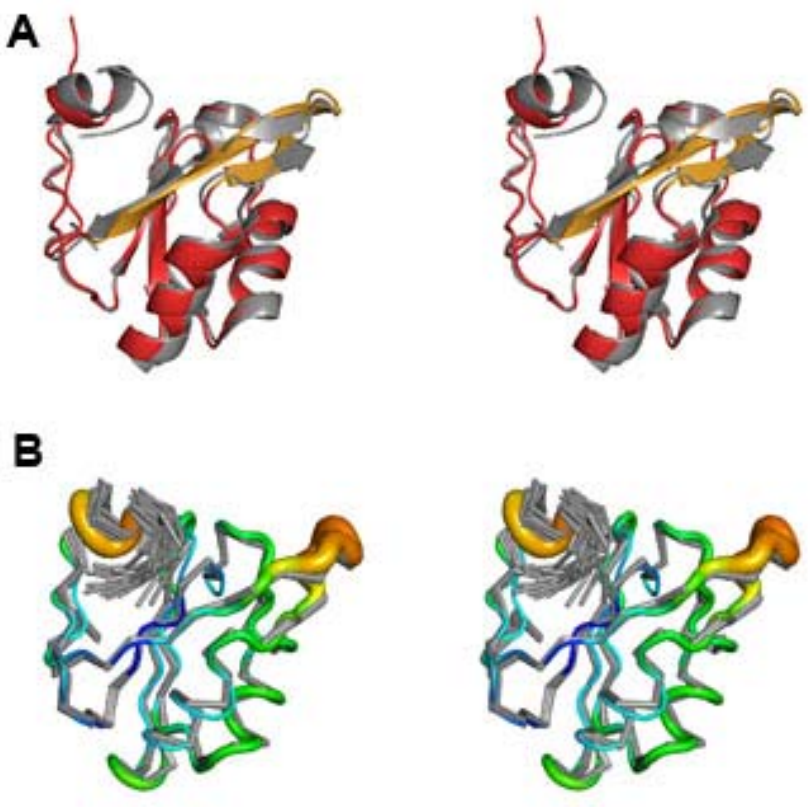
1-354 RRM/AlkB double-domain construct (blue) or the 1-633-his_6+Trm112 complex (red); the presence or absence of the C-terminal hexahistidine tag does not influence the observed interactions (data not shown). Filter-binding assays (top right) were used to assay the interaction of 5 nM radiolabeled tRNA-Glu with the 25-354 RRM/AlkB double-domain construct lacking the basic N-terminal α-helix (blue), the 132-354 AlkB single-domain construct (orange), or the 352-633+Trm112, MTase domain construct (red). Fluorescent anisotropy assays were used to monitor the binding of 5'-fluorescein-labeled 17mer stem-loop (bottom left panel) or aptamer ABH8-2.2 (bottom right panel) to N-terminally truncated RRM single-domain constructs (12-125 in dark blue and 25-125 in orange), the 132-354 AlkB single-domain construct (black), the 126-663-+Trm112 AlkB/MTase double-domain construct (light blue), or the entirety of ABH8 in complex with Trm112 (1-663-his_6+Trm112, red).

Figure S5



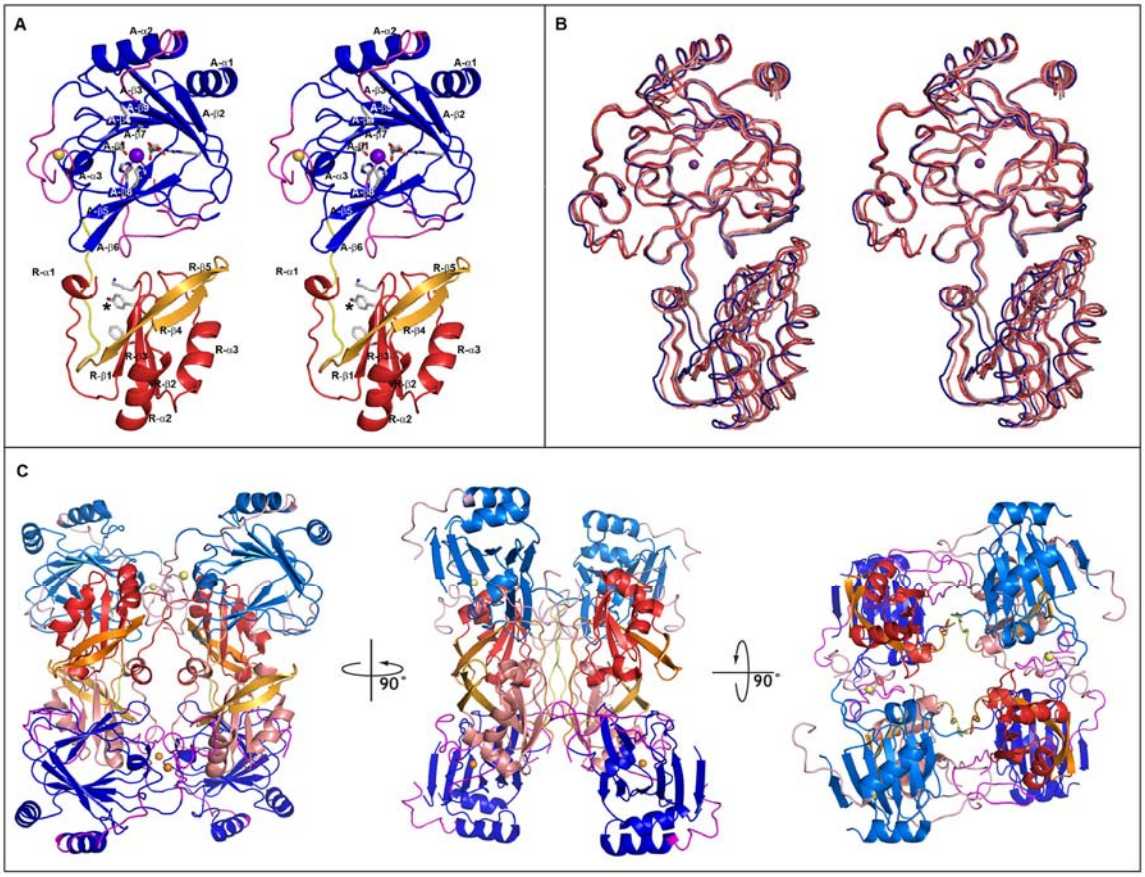
SUPPLEMENTAL FIGURE 5. **Simulated annealing electron density map of the active site and the C-terminal structural Zn-binding site in the ABH8 AlkB domain.** The stereopairs show regions of the Fo–Fc electron density map at 3 Å resolution after after 1000 K simulated annealing of the refined model omitting the metal ions and the 2OG co-substrate. Panel **A** shows the active site in the AlkB domain, while panel **B** shows the structural Zn-binding site at the C-terminus of this domain. The electron density map is shown contoured at both 1 σ (gray) and 5 σ (red). The bound Mn(II) and Zn(II) ions are represented by spheres, while the 2OG co-substrate and protein sidechains are represented by sticks (with carbon colored light gray, oxygen colored red, nitrogen colored blue, and sulfur colored yellow). Both panels show protomer B in the non-SeMet labeled crystal structure in spacegroup C2.

Figure S6



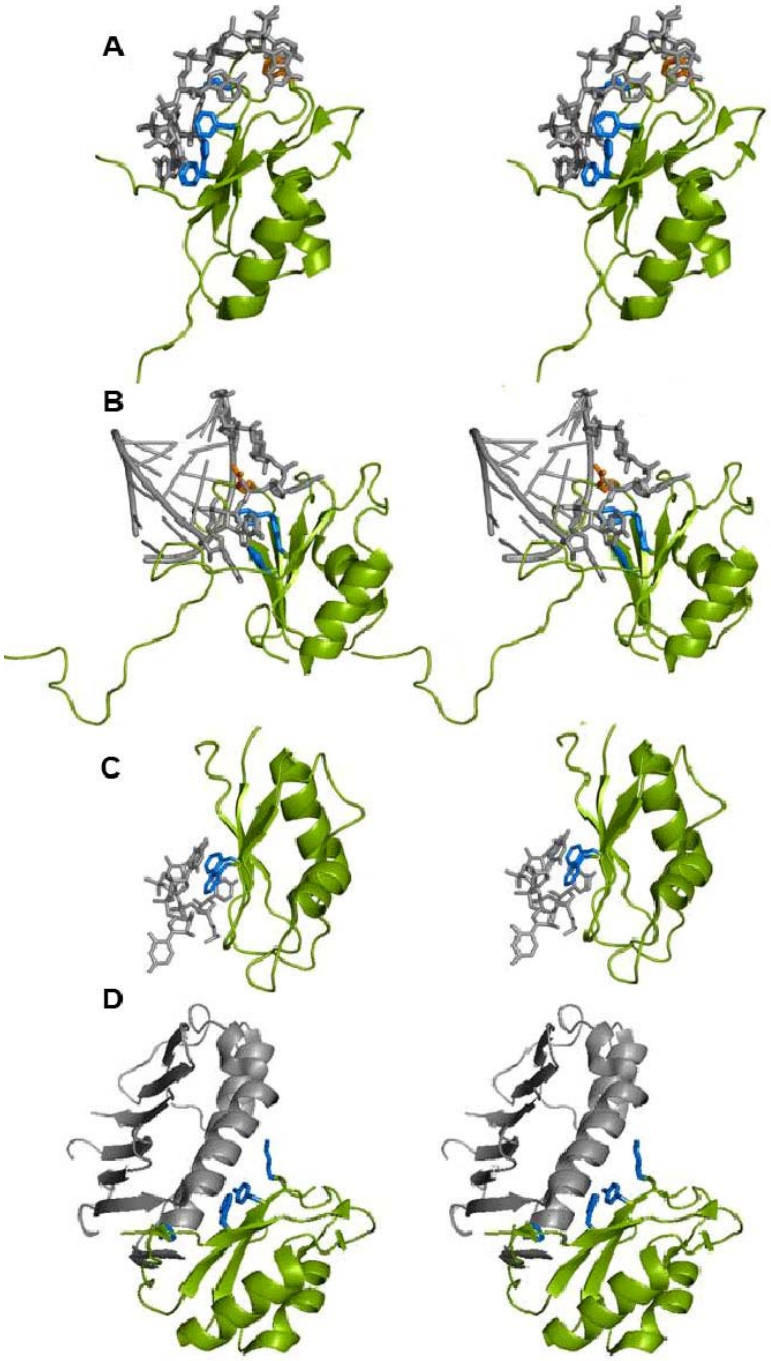
SUPLLEMENTAL FIGURE 6. **Comparison of the crystal structure of the RRM domain in ABH8 to its lowest energy NMR structure.** The crystal structure of the RRM domain of ABH8 (Table 2 and Fig. 6 in the main text) superimposes with its NMR structure with a 0.9 Å RMSD for Cα atoms. **(A)** Cartoon diagrams showing least-squares alignment of the crystal structure of the domain (colored red and orange as in Fig. 6A) with the lowest energy conformation from the NMR ensemble (gray, PDB id 2CQ2). **(B)** An equivalent view of the crystal structure superimposed with the backbone traces of the entire NMR ensemble (20 different conformers in gray). In this panel, the thickness and coloring of the backbone trace of the crystal structure encode the refined backbone B-factors (using the default encoding in Pymol). The x-ray and NMR structures diverge significantly only at the C-terminus of the domain, which contains non-native residues in the construct used for the NMR structure.

Figure S7



SUPPLEMENTAL FIGURE 7. **Detailed analyses of the crystal structures of the RRM/AlkB domains in human ABH8.** **(A)** This stereo ribbon diagram shows secondary structural elements labeled according to their connectivity within the RRM and AlkB domains, both of which retain the core folds characteristic of homologous proteins (6,18-20). Elements with labels starting with "R" belong to the RRM domain, while those starting with "A" belong the AlkB domain. The structure is colored and oriented identically to Fig. 6A in the main text. The blue asterisk highlights the position of the hydrophobic pocket formed in part by the RNP1 motif in the RRM domain. The C-terminal TEV-cleavable hexahistidine tag retained in the crystallized construct is omitted from display in this panel (but shown in panels B-C in this figure). **(B)** The crystal structures of the same ABH8 construct in different spacegroups (Table 2 in the main text) show some flexibility in the linkage between the RRM and AlkB domains. The four protomers in the asymmetric unit of the unlabeled crystal form (shades of red) are superimposed here on the single protomer in the SeMet crystal form (blue) based on least-squares alignment of the Cα atoms in their AlkB domains. These images show the C-terminal TEV-cleavable hexahistidine tag that makes the same inter-protomer packing contacts in all crystallographically observed subunits. **(C)** The same tetrameric assembly with *222* symmetry is observed in both crystal forms of the RRM/AlkB domains in ABH8. Two of the subunits are displayed in the same colors as in Fig. 6A in the main text (and panel **A** here), while the other two are displayed in lighter shades of the same colors. The C-terminal TEV-cleavable hexahistidine tag makes packing contacts that can be seen at the center of the leftmost image here (across 2-fold axes running vertically up the page). See the main text for further description of this oligomer and observations suggesting that it is not a physiological assembly.

Figure S8

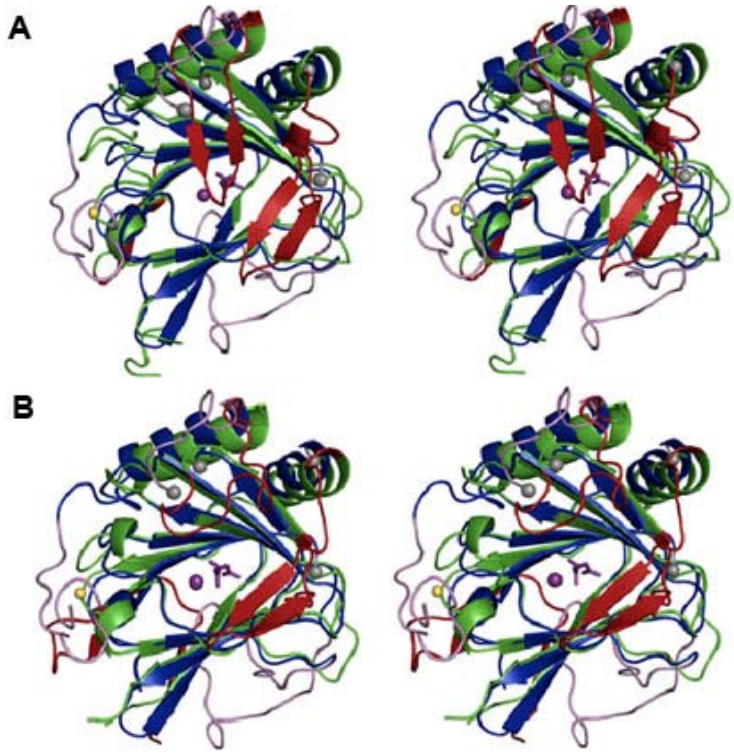


SUPPLEMENTAL FIGURE 8. **Additional structural comparisons of the RRM domain of ABH8.** These stereopairs show the RRM domains from four different proteins in which the domain participates in stereochemically divergent intermolecular interactions. The RRM domains are colored green, with the residues in their RNP1 motifs represented by blue sticks and additional residues involved in RNA binding represented by orange sticks. The macromolecules interacting with the homologous domains are colored gray. **(A)** The RRM domain of the human splicing factor Fox-1 (PDB id 2ERR) binds RNA (gray) using the aromatic residues in the RNP1 motif in strands β1 and β3 (blue), but additional contact are also made

Figure S8 (cont.)

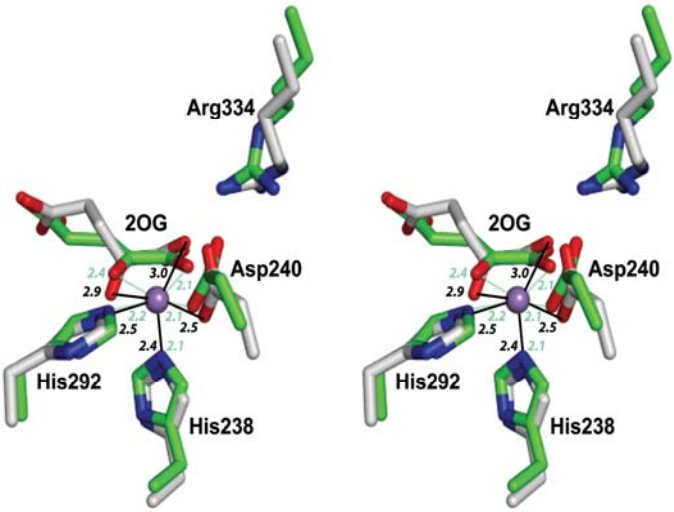
by an aromatic residue (orange) in the loop between strand β1 and helix α1 (21). **(B)** In the RRM domain of the human protein RMBY (PDB id 2FY1), the loop between strands β3 and β4 inserts into the RNA substrate (gray), while the RNP1 motif makes sequence-specific contacts to additional RNA bases (22). **(C)** The yeast protein Nab3 (PDB id 2L41) recognizes RNA (gray) via contacts from the canonical RNP1 and RNP2 motifs, including base-stacking interactions with aromatic residues from these motifs (blue) (23). **(D)** The RRM domain of the eukaryotic protein Y14 interacts with the protein Mago (gray) in the exon junction complex (PDB id 1HL6), illustrating the involvement of an RRM domain in mediating protein-protein rather than protein-RNA interactions (24).

Figure S9



SUPPLEMENTAL FIGURE 9. **Additional structural comparisons of the AlkB domain of ABH8.** In these stereo ribbon diagrams, the regions of the ABH8 domain conserved in other AlkB homologs are colored blue, while the non-conserved regions are colored magenta (like in Fig. 6A in the main text). The Fe/2OG-binding cores in the homologs are colored green while their nucleotide-binding lids are colored red. The Mn(II) cofactor and 2OG co-substrate bound to ABH8 are shown in space-filling and stick representations, respectively. The gray spheres indicate the boundaries of the disordered backbone segments in the ABH8 AlkB domain. **(A)** Superposition with human ABH2 (Z-score of 16 and a 2.4 Å RMSD for alignment of 157 out of 194 Cα atoms with 21% sequence id in PDB id 3BUC) (20,25). **(B)** Superposition with human ABH3 (Z-score of 16 and 2.4 Å RMSD for alignment of 153 of 194 Cα atoms with 19% sequence identity in PDB id 2IUW) (19,25).

Figure S10



SUPPLEMENTAL FIGURE 10. **Active-site geometry in the AlkB domain of human ABH8.** The catalytic Fe(II) center in the AlkB domain of ABH8 appears to be in an inactive conformation in the crystal structures reported in this paper (Table 2 in the main text). The stereopair shows a stick representation of a least-squares alignment of the catalytic center in our 3 Å crystal structure of ABH8 with that of *E. Coli* AlkB (PDB id 2FDH (6)). Carbon atoms are colored gray in ABH8 and green in *E. coli* AlkB, while oxygen atoms are colored red and nitrogen atoms are colored blue in both structures. The bound Mn(II) cofactors are shown as purple spheres. Inter-nuclear distances in Å are shown in black and green for the ABH8 and *E. coli* AlkB structures, respectively. Residue numbers refer to the ABH8 AlkB domain. The least-squares alignment was calculated using just the metal cofactor and its liganding atoms in the two structures, which were subsequently translated (without rotation) to exactly superimpose the metal cofactors. In both of the refined crystal structures of the RRM/AlkB domains of human ABH8 (Table 2 in the main text), the 1-carboxylate group and 2-oxo atom of 2OG are rotated ~70˚ away from the proper octahedral coordination geometry observed in *E. coli* AlkB (6) and most other Fe/2OG dioxygenases. See the main text and the legend to Fig. 6D therein for additional information.