Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity

Melissa M Matthews1,3, Justin M Thomas1,3, Yuxuan Zheng1, Kiet Tran1, Kelly J Phelps1, Anna I Scott2, Jocelyn Havel1, Andrew J Fisher1,2 & Peter A Beal1

Adenosine deaminases acting on RNA (ADARs) are editing enzymes that convert adenosine to inosine in duplex RNA, a modification reaction with wide-ranging consequences in RNA function. Understanding of the ADAR reaction mechanism, the origin of editing-site selectivity, and the effect of mutations is limited by the lack of high-resolution structural data for complexes of ADARs bound to substrate RNAs. Here we describe four crystal structures of the human ADAR2 deaminase domain bound to RNA duplexes bearing a mimic of the deamination reaction intermediate. These structures, together with structure-guided mutagenesis and RNA-modification experiments, explain the basis of the ADAR deaminase domain’s dsRNA specificity, its base-flipping mechanism, and its nearest-neighbor preferences. In addition, we identified an ADAR2-specific RNA-binding loop near the enzyme active site, thus rationalizing differences in selectivity observed between different ADARs. Finally, our results provide a structural framework for understanding the effects of ADAR mutations associated with human disease.

RNA-editing reactions alter a transcript’s genomically encoded sequence by inserting, deleting, or modifying nucleotides1. Deamination of A, the most common form of RNA editing in humans, generates inosine (I) at the corresponding nucleotide position. Because I base-pairs with C, it functions similarly to G in gene expression processes such as splicing, translation, and reverse transcription2-3. A-to-I editing has wide-ranging consequences in RNA function, including altering microRNA-recognition sites, redirecting splicing, and changing the meaning of specific codons4-6. Two different enzymes carry out A-to-I editing in humans: ADAR1 and ADAR2 (ref. 7). ADAR activity is required for nervous-system function, and altered editing has been linked to neurological disorders such as epilepsy and Prader–Willi syndrome8-10. In addition, mutations in the gene encoding ADAR1 cause the autoimmune disease Aicardi–Goutières syndrome (AGS) and the skin disorder dyschromatosis symmetrica hereditaria (DSH)11-13. Hyperediting has been observed at certain sites in cancer cells, such as in the AZIN1 mRNA (for antizyme inhibitor 1)14-15. However, hypoediting also occurs in cancer-derived cell lines, as exemplified by the reduced editing observed in GLI1 mRNA (for glioma-associated oncogene 1)16.

The ADAR proteins have a modular structure with double-stranded RNA (dsRNA)-binding domains (dsRBDs) and a C-terminal deaminase domain17 (human [h] ADAR2 domains in Fig. 1a). ADARs efficiently deaminate specific adenosines in duplex RNA while leaving most adenosines unmodified18. The mechanism of adenosine deamination requires ADAR to flip the reactive base out of an RNA double helix to access its active site17. How an enzyme can accomplish this task with a duplex RNA substrate is not known. Furthermore, how the ADAR deaminase domain contributes to editing-site selectivity is also unknown; to our knowledge, no structures of ADAR deaminase domain–RNA complexes have been reported. To address these knowledge gaps, we set out to trap the human ADAR2 deaminase domain (amino acids [aa] 299–701, hADAR2d) bound to different duplex RNAs and to solve structures of the resulting complexes with X-ray crystallography. We then evaluated the importance of protein–RNA contacts by using structure-guided mutagenesis and RNA-modification experiments coupled with adenosine-deamination kinetics.

RESULTS

Trapping the flipped conformation

The ADAR reaction involves the formation of a hydrated intermediate that loses ammonia and thereby generates the inosine-containing product RNA17 (reaction scheme in Fig. 1b). The covalent hydrate of the nucleoside analog 8-azanebularine mimics the proposed high-energy intermediate19 (reaction scheme in Fig. 1b). For trapping hADAR2d bound to RNA for crystallography, we incorporated 8-azanebularine into the edited site of duplex RNAs that have recently been shown to be excellent substrates for deamination by hADAR2d20 (duplex sequence in Fig. 1c); characterization of protein–RNA complex in Supplementary Fig. 1). In addition, for one of these duplexes (Bdf2), we positioned the 8-azanebularine opposite either U or C to mimic an A-U pair or A-C mismatch at the editing site, creating a total of three different RNA substrates for structural studies (Fig. 1c).

The hADAR2d protein (without RNA bound) has previously been crystallized and structurally characterized, thus revealing features of the active site, including the presence of zinc21. In addition, an inositol

1Department of Chemistry, University of California, Davis, Davis, California, USA. 2Department of Molecular and Cellular Biology, University of California, Davis, Davis, California, USA. 3These authors contributed equally to this work. Correspondence should be addressed to P.A.B. (pabeal@ucdavis.edu) or A.J.F. (ajfisher@ucdavis.edu).

Received 1 February; accepted 11 March; published online 11 April 2016; doi:10.1038/nsmb.3203
hexakisphosphate (IHP) molecule has been found to be buried in the core of the protein hydrogen-bonded to numerous conserved polar residues. For crystallization of hADAR2d–RNA complexes, we used both the wild type (WT) deaminase domain and a mutant (E488Q) with enhanced catalytic activity20,22 (descriptions of the crystallization conditions, X-ray diffraction data collection and solution of the structures in Online Methods).

Four protein–RNA combinations generated diffracting crystals that resulted in high-resolution structures (hADAR2d WT–Bdf2-U, hADAR2d WT–Bdf2-C, ADAR2-D E488Q–Bdf2-C, and hADAR2d E488Q–GLI1) (Table 1). In each of these complexes, the protein binds the RNA on one face of the duplex, over ~20 bp, by using a positively charged surface near the zinc-containing active site (Fig. 2 and Supplementary Fig. 2a). The large binding site (1,493 Å² RNA surface area and 1,277 Å² protein surface area buried) observed for hADAR2d is consistent with the results from recent footprinting studies20. Both strands of the RNA contact the protein, and the majority of these interactions are mediated through the phosphodiester-ribose backbone near the editing site (Fig. 2c and Supplementary Fig. 2b–d).

The structures show a large deviation from an A-form RNA conformation at the editing site (Figs. 2 and 3 and Supplementary Video 1). The 8-azanebularine is flipped out of the helix and bound in the active site as its covalent hydrate, where it interacts with several amino acids including V351, T375, K376, E396, and R455 (Fig. 3a and Supplementary Fig. 3a). The side chain of E396 hydrogen-bonds to purine N1 and O6. This interaction was expected, given the proposed role of E396 in mediating proton transfer to and from N1 of the substrate adenosine17. The Z'-hydroxyl of 8-azanebularine hydrogen-bonds to the backbone carbonyl of T375 while the T375 side chain contacts its Z'-phosphodiester. R455 and K376 help position...

Table 1 Data collection and refinement statistics

| ADAR2-D E488Q–BDF2-GLI1 | ADAR2-D E488Q–BDF2-U | ADAR2-D WT–BDF2-C | ADAR2-D WT–BDF2-U |
|---------------------------|----------------------|-------------------|-------------------|
| ADAR2-D E488Q–BDF2-GLI1 | ADAR2-D E488Q–BDF2-U | ADAR2-D WT–BDF2-C | ADAR2-D WT–BDF2-U |
| P2₁2₁2₁                  | P2₁2₁2₁              | P2₁2₁2₁          | P2₁2₁2₁          |
| Space group               | Space group          | Space group      | Space group      |
| P2₁2₁2₁                  | P2₁2₁2₁              | P2₁2₁2₁          | P2₁2₁2₁          |
| Cell dimensions           | Cell dimensions      | Cell dimensions  | Cell dimensions  |
| a, b, c (Å)               | a, b, c (Å)          | a, b, c (Å)      | a, b, c (Å)      |
| Resolution (Å)            | Resolution (Å)       | Resolution (Å)   | Resolution (Å)   |
| 100–2.75 (2.82–2.75)      | 100–2.75 (2.82–2.75) | 100–2.98 (3.06–2.98) | 100–3.09 (3.17–3.09) |
| Rmerge (%)                | Rmerge (%)           | Rmerge (%)       | Rmerge (%)       |
| 7.0 (68.5)                | 9.6 (135.1)          | 14.4 (86.3)      | 11.6 (68.7)      |
| CC₁/₂                     | CC₁/₂                | CC₁/₂            | CC₁/₂            |
| 99.6 (66.3)               | 99.7 (47.6)          | 99.1 (75.3)      | 99.3 (77.1)      |
| Completeness (%)          | Completeness (%)     | Completeness (%) | Completeness (%) |
| 96.5 (98.8)               | 98.1 (98.9)          | 97.3 (90.0)      | 96.8 (89.1)      |
| Redundancy                | Redundancy           | Redundancy       | Redundancy       |
| 2.93 (3.00)               | 5.19 (5.13)          | 4.79 (4.56)      | 3.31 (2.84)      |
| Refinement                | Refinement           | Refinement       | Refinement       |
| Resolution (Å)            | Resolution (Å)       | Resolution (Å)   | Resolution (Å)   |
| 2.75                      | 2.95                 | 2.98             | 3.09             |
| No. of reflections (F > 0) | 27,153               | 35,727           | 21,376           |
| Rmerge / Rfree            | Rmerge / Rfree       | Rmerge / Rfree   | Rmerge / Rfree   |
| 16.27 / 22.34             | 18.79 / 20.75        | 16.67 / 24.67    | 16.29 / 23.79    |
| No. of atoms              | No. of atoms         | No. of atoms     | No. of atoms     |
| Protein                   | Protein              | Protein          | Protein          |
| 6,197                     | 6,038                | 6,168            | 6,157            |
| RNA                       | RNA                  | RNA              | RNA              |
| 973                       | 1,950                | 973              | 973              |
| Inositol Hexakisphosphate | Inositol Hexakisphosphate | Inositol Hexakisphosphate | Inositol Hexakisphosphate |
| 72                        | 72                   | 72               | 72               |
| Zn                        | Zn                   | Zn               | Zn               |
| 2                         | 2                    | 2                | 2                |
| Water                     | Water                | Water            | Water            |
| 33                        | –                    | 1                | 1                |
| 8 factors                 | 8 factors            | 8 factors        | 8 factors        |
| Protein                   | Protein              | Protein          | Protein          |
| 68.46                     | 90.65                | 63.63            | 67.92            |
| RNA                       | RNA                  | RNA              | RNA              |
| 88.24                     | 108.8                | 69.70            | 77.49            |
| Inositol Hexakisphosphate | Inositol Hexakisphosphate | Inositol Hexakisphosphate | Inositol Hexakisphosphate |
| 47.10                     | 65.57                | 44.23            | 43.77            |
| Zn                        | Zn                   | Zn               | Zn               |
| 48.47                     | 64.38                | 38.25            | 49.04            |
| Water                     | Water                | Water            | Water            |
| 48.11                     | –                    | 43.66            | 50.20            |
| R.m.s. deviations         | R.m.s. deviations    | R.m.s. deviations | R.m.s. deviations |
| Bond lengths (Å)          | Bond lengths (Å)     | Bond lengths (Å) | Bond lengths (Å) |
| 0.010                     | 0.007                | 0.009            | 0.010            |
| Bond angles (°)           | Bond angles (°)      | Bond angles (°)  | Bond angles (°)  |
| 1.342                     | 0.885                | 1.359            | 1.365            |
the flipped nucleotide in the active site by fastening the phosphate backbone flanking the editing site. The R455 side chain ion-pairs with the 5'-phosphodiester of 8-azanebularine while the K376 side chain contacts its 3'-phosphodiester. Last, the side chain of V351 provides a hydrophobic surface for interaction with the nucleobase of the edited nucleotide. RNA binding does not alter IHP binding or the hydrogen-bonding network linking IHP to the active site 21.

ADARs use a unique mechanism to modify duplex RNA

The ADAR2 base-flipping loop, bearing residue 488, approaches the RNA duplex from the minor-groove side at the editing site. The side chain of this amino acid penetrates the helix, where it occupies the space vacated by the flipped-out base and hydrogen-bonds to the complementary-strand orphaned base and to the 2'-hydroxyl of the nucleotide immediately 5' to the editing site (Fig. 3b,c). In the four structures reported here, we observed three different combinations of helix-penetrating residue and orphan base (i.e., E488 + U, E488 + C and Q488 + C), and all three combinations show the same side chain and base positions (Fig. 3b,c and overlay of all three in Supplementary Fig. 4a). For instance, in the complex with hADAR2d E488Q and the Bdf2-C duplex, the protein recognizes an orphaned C by donating hydrogen bonds from Nε2 to cytosine N3 and from its backbone NH to cytosine O2 (Fig. 3b). In the complex of hADAR2d WT and the Bdf2-U duplex, we observed a similar interaction, in which the E488 backbone NH hydrogen-bonds to the uracil O2, and the E488 side chain hydrogen-bonds to the uracil N3H (Fig. 3c).

Interestingly, the E488Q mutant was initially discovered in a screen for highly active ADAR2 mutants, and this residue has been suggested to be involved in base-flipping, given its effect on editing substrates with a fluorescent nucleobase at the editing site 22. ADARs react preferentially with adenosines in A•C mismatches and A-U pairs over A•A and A•G mismatches 23. A purine at the orphan base position (in its anti conformation) would clash with residue 488, thus explaining the preference for pyrimidines observed here.

The interaction of residue 488 with the orphaned base is reminiscent of an interaction observed for the HhaI DNA methyltransferase (MTase), a duplex-DNA-modifying enzyme that also uses a base-flipping mechanism to access 2'-dC for methylation 24,25. For that enzyme, Q237 hydrogen-bonds to an orphaned dG while it fills the void left by the flipped-out dC 25 (Supplementary Fig. 4b). In addition, two glycine residues flank Q237, thus allowing the loop to adopt the conformation necessary for penetration into the helix 24. The flipping loop in ADAR2 (i.e., aa 487-489) also has the helix-penetrating residue flanked by glycines. However, unlike the case of the DNA MTase, which approaches the DNA from the major groove, the ADAR2 loop approaches the duplex from the minor-groove side. Such an approach requires deeper penetration of the intercalating residue to access the hydrogen-bonding sites on the orphaned base, thus necessitating an additional conformational change in the RNA duplex. This change includes shifting of the base pairs immediately
5′ to the editing site toward the helical axis and a widening of the major groove opposite the editing site (Fig. 4a,b and Supplementary Video 1). In the case of the hADAR2d WT–Bdf2–U RNA, this shift is accompanied by a shearing of the U11-A13′ base pair with U11 shifted further in the direction of the major groove, thus creating an unusual U-A ‘wobble’ interaction with adenine N6 and N1, within hydrogen-bonding distance to uracil N3H and O2, respectively (Fig. 4c and Supplementary Fig. 3b). This type of wobble pair has been observed before and requires either the imino tautomeric of adenine or the enol tautomeric of uracil29. The ADAR-induced distortion in RNA conformation results in a kink in the RNA strand opposite to the editing site (Fig. 4b). This kink is stabilized by interactions of the side chains of R510 and S495 with phosphodiester s in the RNA backbone (Supplementary Fig. 5a). Interestingly, ADAR2’s flipping-loop approach to gain access to the modification site from the minor groove, thereby allowing for penetration of intercalating residues and damage recognition30. Whereas hADAR2d clearly alters the duplex conformation in the minor groove spanning the three base pairs that include the nearest-neighboring nucleotides flanking the edited base (Supplementary Fig. 3b,c). As described above, the base pair including the 5′-nearest-neighbor U (U11-A13′ in the Bdf2 duplex) is shifted from the position that it would occupy in a typical A-form helix to accommodate the loop (Fig. 4a). In addition, the minor-groove edge of this pair is juxtaposed with the protein backbone at G489. Modeling a G-C or C-G pair at this position (i.e., 5′-G or 5′-C) suggests that a 2-amino group in the minor groove would clash with the protein at G489 (Fig. 5a and Supplementary Fig. 7c). Indeed, replacing the U-A pair adjacent to the editing site with a C-G pair in the GLI1 duplex substrate resulted in an 80% reduction in the rate of hADAR2d-catalyzed deamination (Fig. 5b,c). To determine whether this effect arises from an increase in local duplex stability from the substitution of C-G for U-A or from the presence of the 2-amino group, we replaced the U-A pair with a U-2-aminopurine (2AP) pair. 2AP is an adenosine analog that forms

**Figure 3** ADAR recognition of the flipped-out and orphaned nucleotides. (a) Contacts to the editing-site nucleotide (N) in the active site. Colors correspond to those in Figures 1 and 2. (b) Orphan-nucleotide recognition in the hADAR2d E488Q–Bdf2–C complex. (c) Orphan-nucleotide recognition in the hADAR2d WT–Bdf2–U complex.

**Figure 4** Other ADAR-induced changes in RNA conformation. (a) hADAR2d-induced shift in the position of the U11-A13′ base pair from ideal A-form RNA helix (yellow). (b) Overlay of Bdf2 duplex RNA and an idealized A-form duplex of same sequence (yellow), illustrating a kink in the strand and the widening of the major groove opposite the editing site induced by hADAR2d. (c) Unusual wobble A13′-U11 interaction in the hADAR2d WT–Bdf2–U complex, shown in sticks with hydrogen bonds indicated with yellow dashes and distances shown in Å. The position of this base pair in the hADAR2d E488Q–Bdf2–C duplex is shown in wire with hydrogen bonds shown with gray dashes.
that deamination in the substrate with a 3′-I, inosine; d2AP, 2-aminopurine). In each of the hADAR2d–RNA structures reported here, the backbone carbonyl oxygen at S486 forms a hydrogen bond with the 3′-G 2-aminogroup (Fig. 5e). G is the only common nucleobase that presents a hydrogen-bond-donor capability in the RNA minor groove at this location for sequences with 5′-nearest-neighbor G or C. However, the observed clash is not severe, and the backbone at G489 that results from the presence of an amino group in the minor groove at this location for sequences with 5′-nearest-neighbor G or C. This residue is conserved in ADAR2s and ADAR1s but is glutamine in the editing-inactive ADAR3s (Supplementary Table 1). Mutation of hADAR2d at this site to glutamine (G593Q) or alanine (G593A) reduced the measured deamination rate constant by approximately an order of magnitude (Fig. 6c). In addition, the contact point near the 5′ end of the unedited strand involved G593, K594, and R348, residues completely conserved in the family of ADAR2s (Fig. 2c, Supplementary Table 1). Mutation of any of these residues to alanine (G593A, K594A, or R348A) substantially reduced editing activity (Fig. 6c). In addition, mutation of G593 to glutamate (G593E) resulted in a rate reduction of nearly two orders of magnitude, a result consistent with the proximity of this residue to the negatively charged phosphodiester backbone of the RNA (Fig. 6c).

RNA-binding loops of the ADAR catalytic domain

The structures reported here identify RNA-binding loops of the ADAR catalytic domain and suggest roles for several amino acids not previously known to be important for editing, either substrate binding or catalysis (Fig. 6). The side chain of R510 ion-pairs with the 3′-phosphodiester of the orphaned nucleotide (Fig. 3a,c). This residue is conserved in ADAR2s and ADAR1s but is glutamine in the editing-inactive ADAR3s (Supplementary Table 1). Mutation of hADAR2d at this site to glutamine (R510Q) or alanine (R510A) reduced the measured deamination rate constant by approximately an order of magnitude (Fig. 6c). In addition, the contact point near the 5′ end of the unedited strand involved G593, K594, and R348, residues completely conserved in the family of ADAR2s (Fig. 2c, Supplementary Table 1). Mutation of any of these residues to alanine (G593A, K594A, or R348A) substantially reduced editing activity (Fig. 6c). In addition, mutation of G593 to glutamate (G593E) resulted in a rate reduction of nearly two orders of magnitude, a result consistent with the proximity of this residue to the negatively charged phosphodiester backbone of the RNA (Fig. 6c).

RNA binding leads to an ordering of the 454–477 loop, which was disordered in the RNA-free hADAR2d structure (Fig. 2a,b and Supplementary Video 2). This loop binds the RNA duplex contacting the minor groove near the editing site and inserting into the adjacent major groove (Fig. 6e). This loop sequence is conserved in ADAR2s but is different in the family of ADAR1s (Fig. 6d). The substantial difference in sequence between the ADARs in this RNA-binding loop suggests that differences in editing-site selectivity between the two ADARs arise at least partially from differences in how this loop binds RNA substrates.

DISCUSSION

Base-flipping is a well-characterized mechanism by which nucleic acid–modifying enzymes gain access to sites of reaction that are otherwise buried in base-paired structures. DNA methylases, DNA-repair glycosylases, and RNA-loop-modifying enzymes that flip a nucleotide out of a base pair are known. However, none of the structurally characterized base-flipping enzymes access their reactive sites from a B-form DNA conformation.

Figure 5 Interactions with editing-site nearest-neighbor nucleotides. (a) Space-filling model of 5′-nearest-neighbor base pair and G489. The minor-groove edge of the U11-A13′ base pair from the Bdf2 duplex approaches G489; model with a C-G pair at this position suggests a clash with the G2-aminogroup. (b) RNA duplex substrates prepared with different 5′-nearest-neighbor nucleotides adjacent to the editing site, indicated in red (2AP, 2-aminopurine). (c) Comparison of deamination rate constants by hADAR2d at the editing-site adenosine (red) for duplexes bearing different 5′-nearest neighbors; $k_{\text{rel}} = k_{\text{obs}}/k_{\text{obs}}$ for unmodified RNA). Error bars, s.d. (n = 3 technical replicates). (d) hADAR2 S486 backbone hydrogen bond with the 3′-G 2-aminogroup. (e) RNA duplex substrates prepared with different 3′-nearest-neighbor nucleotides adjacent to the editing site, indicated in red (I, inosine; N2MeG, N2-methylguanosine; 2AP, 2-aminopurine). (f) Comparison of deamination rate constants by hADAR2d at the editing-site adenosine (red) for duplexes bearing different 3′-nearest neighbors. $k_{\text{rel}} = k_{\text{obs}}/k_{\text{obs}}$ for unmodified RNA). Error bars, s.d. (n = 3 technical replicates). Asterisk indicates no reaction product observed.
sites from within a normal base-paired RNA duplex. We are aware of one other protein-induced nucleotide-flipping mechanism from an RNA duplex region\(^3\). Bacterial initiation factor 1 (IF1) binds the 30S ribosomal subunit at helix 44 of 16S RNA, with A1492 and A1493 flipped out of the helix and bound in protein pockets (Supplementary Fig. 5b). However, these nucleotides are located in a highly distorted and dynamic duplex region containing several mismatches and are predisposed to undergo this conformational change\(^3\). Thus, this system is not illustrative of base-flipping from a normal duplex and does not involve an enzyme that must carry out a chemical reaction on the flipped-out nucleotide. Other RNA-modification enzymes are known that flip nucleotides out of loops, even from base pairs in loop regions (e.g., pseudouU synthetase\(^3\), RNA transglycosylase\(^4\), and restricticin bound to the sarcin-ricin loop of 28S rRNA\(^3\)) (Supplementary Fig. 5b). Because the modification sites are not flanked on both sides by normal duplex, these enzymes do not experience the same limits in approach to the substrate that ADARs experience. The requirement that ADARs must induce flipping from a normal duplex has implications on its preference for adenosines flanked by certain base pairs, a phenomenon that was not previously well understood.

In our structures, the flipped-out 8-azanebularine is hydrated, mimicking the tetrahedral intermediate predicted for deamination of adenosine (Figs. 1b and 3a and Supplementary Fig. 3a). Our use of 8-azanebularine, with its high propensity to form a covalent hydrate\(^4\), allowed us to capture a true mimic of the tetrahedral intermediate and to reveal the interactions between the deaminase active site and the reactive nucleotide. In addition, 8-azanebularine adopted a 2′ endo sugar pucker with its 2′ hydroxyl hydrogen-bonded to the protein backbone carbonyl at T375. The 2′ endo conformation appears to facilitate access of the nucleobase to the zinc-bound water for nucleophilic attack at C6.

Several other base-flipping enzymes stabilize the altered nucleic acid conformation by intercalation of an amino acid side chain into the space vacated by the flipped-out base\(^3\). For hADAR2, E488 serves this role. In the two structures with wild-type hADAR2, the E488 residue and orphan base are in nearly identical positions (overlay in Supplementary Fig. 4a). Thus, the E488 side chain directly contacts each orphan base, probably by accepting a hydrogen bond from uracil N3H or by donating a hydrogen bond to cytidine N3. The latter interaction requires E488 to be protonated. The pK\(_a\) of E488 in the ADAR–RNA complex has not been measured, but proximity to hydrogen-bond acceptors, such as the N3 of C, and insertion between stacked nucleobases, would undoubtedly elevate this value and might lead to a substantial fraction in the protonated state at physiologically relevant pH. Because the glutamine side chain is fully protonated under physiologically relevant conditions, a rate enhancement for the E488Q mutant would be expected if the reaction requires E488 protonation.

The interactions of hADAR2d with base pairs adjacent to the editing-site adenosine explain the known 5′ and 3′-nearest-neighbor preferences (Fig. 5). Although these studies indicate the ADAR2 catalytic domain makes an important contact to the 3′-nearest-neighbor G, Stefl et al. have suggested the 3′-G preference arises from dsRBD binding selectivity for ADAR2 (ref. 42). These authors have reported a model for ADAR2’s dsRBDs bound to an editing substrate, based on NMR data from the isolated dsRBDs (lacking the deaminase domain) and short RNA fragments derived from the Gluk-B (official symbol Gria2) R/G-site RNA\(^4\). They have described an interaction wherein the 3′-G 2-amino group hydrogen-bonds to the backbone carbonyl of S258 found in the β1–β2 loop of ADAR2’s dsRBD II. It is not possible for the S486–3′-G interaction that we describe here and the S258–3′-G interaction reported by Stefl et al. to exist in the same complex, because both involve protein loops bound in the RNA minor groove at the same location. Because our structures capture the edited nucleotide in the conformation required to access the active site, the interactions observed here are highly likely to occur during the deamination reaction at the editing site. However, because dsRBDs are known to bind promiscuously with duplex RNA, it is possible that the S258–3′-G interaction found in a complex lacking the deaminase domain may not be relevant to catalysis at the editing site\(^4\). It is also possible that the ADAR dsRBD and catalytic domain binding are sequential, such that release of the dsRBD from the RNA takes place before catalytic-domain engagement and base-flipping.

AGS and DSH are human diseases caused by mutations in hADAR1, and several disease-associated mutations have been found in the deaminase domain\(^1\). Given the conservation in RNA-binding surface and active site residues, we expect the hADAR1 catalytic domain to bind RNA with an orientation of the helix similar to that found in our hADAR2d–RNA structures. When the locations of the AGS-associated mutations are mapped onto the hADAR2d–RNA complex, two mutations involve residues in proximity to the RNA (<4 Å).
Methods and any associated references are available in the online Accession codes.

The authors declare no competing financial interests.

**ACKNOWLEDGMENTS**

The authors acknowledge funding from the US National Institutes of Health (NIH) grant R01GM061115 (P.A.B.). A.I.S. was supported by NIH training grant T32 GM007377. C. Palumbo is acknowledged for technical assistance.

**AUTHOR CONTRIBUTIONS**

J.M.T., M.M.M., A.I.S., and Y.Z. purified protein. K.J.P. and J.M.T. designed and purified RNA for crystallography and characterized protein-RNA binding. M.M.M. and A.I.S. conducted crystallization trials. M.M.M. and A.J.F. collected diffraction data and solved and refined the crystal structures. M.J.T., Y.Z., and J.H. measured enzyme reaction rates. K.T. synthesized 8-azanebularane phosphoramidite. J.M.T. and A.I.S. conducted mutagenesis. J.M.T., M.M.M., P.A.B., and A.J.F. analyzed the structures. P.A.B. wrote the initial manuscript draft. J.M.T., M.M.M., P.A.B., and A.J.F. edited the manuscript.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.
37. Yang, X., Gérlæzi, T., Glover, L.T. & Correll, C.C. Crystal structures of restrictocin-inhibitor complexes with implications for RNA recognition and base flipping. Nat. Struct. Biol. 8, 968–973 (2001).

38. Carter, A.P. et al. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. Science 291, 498–501 (2001).

39. Réblová, K. et al. Structure, dynamics, and elasticity of free 16s rRNA helix 44 studied by molecular dynamics simulations. Biopolymers 82, 504–520 (2006).

40. Xie, W., Liu, X. & Huang, R.H. Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. Nat. Struct. Biol. 10, 781–788 (2003).

41. Erion, M.D. & Reddy, M.R. Calculation of relative hydration free energy differences for heteroaromatic compounds: use in the design of adenosine deaminase and cytidine deaminase inhibitors. J. Am. Chem. Soc. 120, 3295–3304 (1998).

42. Stefl, R. et al. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. Cell 143, 225–237 (2010).

43. Fierro-Monti, I. & Mathews, M.B. Proteins binding to duplexed RNA: one motif, multiple functions. Trends Biochem. Sci. 25, 241–246 (2000).

44. Mannion, N.M. et al. The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. Cell Rep. 9, 1482-1494 (2014).
ONLINE METHODS

Unless otherwise stated, reagents were purchased from Fisher Scientific, Sigma-Aldrich, or Life Technologies. T4 polynucleotide kinase, T4 DNA ligase, molecular-biology-grade bovine serum albumin (BSA), and RNase inhibitor were purchased from New England BioLabs. [γ-32P]ATP was purchased from PerkinElmer Life Sciences. The avian myeloblastosis virus (AMV) reverse transcriptase, dNTP mix and RQI RNase-free DNase were purchased from Promega. Pfu Ultra II was purchased from Stratagene. Dpn I was purchased from Invitrogen. The QuickChange XL II mutagenesis kit was purchased from Agilent Technologies. RNA oligonucleotides were synthesized at the University of Utah DNA/Peptide Core Facility or purchased from GE Healthcare Dharmacan or Sigma Aldrich. DNA oligonucleotides were purchased from Integrated DNA Technologies. Storage phosphorimaging plates from Molecular Dynamics were imaged with a Molecular Dynamics 9400 Typhoon phosphorimager. Data were analyzed with Molecular Dynamics ImageQuant 5.2 software. Electrospray Ionization (ESI) mass spectrometry of oligonucleotide samples was performed at the Campus Mass Spectrometry Facilities, UC Davis. Oligonucleotide masses were determined with Mongo Oligo Mass Calculator v2.06.

Expression and purification of hADAR2 deaminase domain (hADAR2d) for crystallography. Protein expression and purification were carried out by modifying a previously reported protocol45. In brief, S. cerevisiae BCY123 cells were transformed with a pSc-ADAR construct encoding either hADAR2d WT or hADAR2d E488Q (corresponding to the deaminase domain; residues 299–701). Cells were streaked on yeast minimal medium minus uracil (CM–ura) plates. A single colony was used to inoculate a 15-ml CM–ura starter culture. After cultures were shaken at 300 rpm and 30 °C overnight, 10 ml of starter culture was used to inoculate each liter of yeast growth medium. After 24 h, cells were induced with the addition of 110 ml of sterile 30% galactose per liter, and protein was expressed for 5 h. Cells were collected by centrifugation and stored at −80 °C. Cells were lysed in buffer A (20 mM Tris-HCl, pH 8.0, 5% glycerol, 25 mM imidazole, 1 mM NaCl, and 0.01% Triton X-100) with 750 mM NaCl with a microfluidizer, and cell lysate was clarified by centrifugation (39,000 × g, 30 min). Lysate was passed over a 5-ml Ni–NTA column, which was then washed in three steps with 20–50 ml of lysis buffer, wash I buffer (buffer A + 300 mM NaCl), and wash II buffer (buffer A + 100 mM NaCl). Protein was eluted with a 35–300 mM imidazole gradient in wash II buffer over 80 min at a flow rate of 1 ml/min. Fractions containing the target protein were pooled and further purified on a 2-ml GE Healthcare Lifesciences Hi-Trap Heparin HP column in the absence of BME. The HiTrap fusion protein was cleaved with an optimized ratio of 1 mg of TEV protease per 1 mg of protein. Cleavage was carried out for 1–2 h with 1 mL for gel filtration on a GE Healthcare HiLoad 16/600 Superdex 200 PG column. Purification of the target protein was pooled and further purified on a 2-ml GE Healthcare Lifesciences Hi-Trap Heparin HP column in the absence of BME. The HiTrap fusion protein was cleaved with an optimized ratio of 1 mg of TEV protease per 1 mg of protein. Cleavage was carried out for 1–2 h with 1 mL for gel filtration on a GE Healthcare HiLoad 16/600 Superdex 200 PG column. Fractions containing purified protein were pooled and concentrated to 5–7 mg/ml for crystallography trials.

Purification of RNAs for crystallography. The 8-azanebularine (N) phosphoramide was synthesized as previously described39, and RNAs were synthesized as previously described46. Single-stranded RNAs (sequences in Supplementary Table 2) were purified by denaturing PAGE and visualized with UV shadowing. Bands were excised from the gel, crushed, and soaked overnight at 4 °C in 0.5 M NH4OAc, 0.1% sodium dodecyl sulfate (SDS), and 0.1 mM EDTA. Polyacrylamide fragments were removed with a 0.2-µm filter, and this was followed by desalting on a C18 Sep-Pak column. The RNA solutions were lyophilized to dryness, resuspended in nuclease-free water, quantified by absorbance at 260 nm, and stored at −70 °C. Oligonucleotide mass was confirmed by electrospray ionization mass spectrometry. Unmodified RNA stands were purchased from GE Healthcare Dharmacan and purified as described above. Duplex RNA was hybridized in a 1:1 ratio by heating to 95 °C for 5 min and slow cooling to 30 °C.

Crystallization of the hADAR2d–RNA complex. Crystals of the hADAR2d E488Q–Bdf2-C RNA complex were grown at room temperature by the sitting-drop vapor-diffusion method. A solution of 0.5-µl volume containing 4.5 mg/ml protein and 70 µl of Bdf2-C 23-mer RNA (1:0.7 ADAR2/RNA molar ratio) was mixed with 0.5 µl of 0.1 M MES-NaOH, pH 6.5, 9% (w/v) PEG 3350, 13% glycerol, and 0.015 M NAD, which was added to improve crystal growth. Crystals took several weeks to grow. A single cube-shaped crystal of approximately 120 µm in size was soaked briefly in a solution of mother liquor plus 30% glycerol before flash cooling in liquid nitrogen. Data were collected via fine-phi slicing with 0.2° oscillations on beamline 24-ID-C at the Advanced Photon Source at the Argonne National Laboratories in Chicago. To obtain crystals of the hADAR2d WT–Bdf2-C RNA, an identical procedure was used as above; however, the crystallization conditions had slightly different concentrations of reagents (10% PEG 3350, 15% glycerol, 0.1 M MES-NaOH, pH 6.5, and no NAD). For the hADAR2d WT–Bdf2-U construct, hanging-drop vapor diffusion with 200 nl of a mixture containing 4.5 mg/ml protein and 70 µl of Bdf2-U (1:0.7 molar ratio) and 200 ml of mother liquor (0.1 M ammonium acetate, 0.1 M Bis-Tris, pH 5.5, and 17% PEG 10,000) yielded several crystals with a morphology similar to that described above. All wild-type crystals were soaked briefly in a solution of mother liquor plus 30% glycerol before flash cooling in liquid nitrogen. Data were collected via fine-phi slicing with 0.2° oscillations on beamline 12-2 at the Stanford Synchrotron Radiation Lightsource. Crystals of the hADAR2d E488Q–GLI1 RNA complex were grown through hanging-drop vapor diffusion. A solution of volume 200 ml containing 4.5 mg/ml protein and 100 µm of GLI1 23-mer RNA (1:1 ADAR2/RNA molar ratio) was mixed with 200 ml of 0.1 M MES-NaOH, pH 6.5, and 12% PEG 20,000. At room temperature, a single diamond-shaped crystal approximately 150 µm long and 50 µm wide was observed approximately a week later. This crystal was soaked briefly in a solution of mother liquor plus 30% glycerol before flash cooling in liquid nitrogen. Data were collected on beamline 12-2 at the Stanford Synchrotron Radiation Lightsource with the fine-phi slicing described above.

Processing and refinement of crystallographic data. Data for the E488Q Bdf2-C–bound and GLI1–bound structures were processed with XDS47 and scaled with Aimless (CCP4). Diffraction data for hADAR2d wild-type structures were processed with XDS and scaled with XSCALE47. The RNA-free hADAR2d crystal structure (PDB 1ZY7)21 was used as a model for molecular replacement with PHENIX38. The structures were refined with PHENIX including TLS parameters and zinc coordination restraints. Ideal zinc–ligand distances were determined with average distances found for similar coordination models in the PDB database. Table 1 shows the statistics in data processing and model refinement. The asymmetric unit for GLI1–bound hADAR2d E488Q included two complexes of protein–RNA. In each of these complexes, the first 17 residues of the deaminase domain (residues 299–316) as well as a C-terminal proline (P701) were disordered and were therefore not included in the model. However, although the RNA-free structure (PDB 1ZY7) lacked electron density for residues 457–475, we observed density for the backbone atoms of these residues. These residues were initially modeled as polyalanine. After several rounds of refinement, electron density revealed the location of some side chains. Residues whose main side chains interact with the RNA backbone were clearly defined in the final density map. Although some non-RNA-binding side chains showed only weak density, the backbone density was strong. As observed in the original hADAR2d RNA-free structure, IHP was buried in the enzyme core31. The asymmetric units for Bdf2-bound ADARs contained one ADAR2d–RNA complex (protein chain A) and one RNA-free ADAR2d monomer (chain D). The N terminus of the Bdf2-bound structures included more residues than did the GLI1-bound structure, beginning at P305 in chain A and T304 for chain D in the mutant structure, and beginning at R307 in chain A and T304 or P305 in chain D in the wild-type structures. The first few residues (in structures in which the specified residues were modeled) had weak side chain density, including residues 305 and 307 in chain A, and residues 304–307 in chain D, and were modeled in the structure as alanine. The last residue of E488Q–Bdf2-C, P701, had very weak electron density for both protein subunits in the asymmetric unit. Unlike the E488Q–GLI1 structure, electron density was defined better in the originally disordered loop (residues 457–475) for most residues in the Bdf2–bound structures. With the exception of E466 in the wild-type structures, we were able to model–build in main chain and side chain atoms for all residues of this loop in the ADAR subunit in complex with the Bdf2 RNA duplexes. In the RNA-free subunit (chain D) of E488Q–Bdf2-C, a crystal contact stabilized this flexible loop so that we were able to model in the backbone for residues 457–475, but residues 465–475 were modeled as alanine because of poorly defined side chain density. An identical crystal contact was observed in the wild-type structures. In the WT–Bdf2-C complex, the density
Expression and purification of hADAR2d for in vitro deamination kinetics. Histidine-tagged human ADAR2 deaminase domain (hADAR2d) and hADAR2d mutant proteins were expressed in *Saccharomyces cerevisiae* strain BCY123 and purified as described above with the following modifications. Cell lysate was purified with a 0.45-µm filter after centrifugation and loaded three times through 5-ml Ni-NTA Superflow (Qiagen) at 3 ml/min. Washes of 50 ml with buffers 1, 2 and 3 at 4 ml/min were followed by elution with a 35-ml gradient from buffer 3 to elution buffer. Selected elution fractions from the Ni-NTA column were pooled and loaded at 0.5 ml/min on a 1-ml HiTrap Heparin HP column from GE. The column was washed with 10 ml of heparin 1 buffer at 0.5 ml/min and eluted with a 12 ml gradient from heparin 1 to heparin 2 buffer. Selected elution fractions from the heparin column were pooled and concentrated to <300 µL in a 10,000 MWCO Amicon Ultra 4 centrifugal filter at 6,500 RCF and 4 °C. TEF protein cleavage and gel-filtration steps were omitted. Buffer exchange was accomplished via three rounds of concentration to <300 µL and subsequent addition of 3 ml of storage buffer. After final concentration, protein concentrations were determined with BSA standards, as visualized by SYPRO Orange staining on SDS–PAGE gels, and the purified proteins were stored at −70 °C.

Site-directed mutagenesis. Mutagenesis of the hADAR2 catalytic domain was carried out via PCR site-directed mutagenesis with the primers listed in Supplementary Table 2. All primers were purchased from IDT and were PAGE purified as described above but were desalted by phenol-chloroform extraction, ethanol precipitation and 70% ethanol wash instead of C18 Sep-Pack. Sequences of RNAs used to prepare internally 32P-labeled substrates are shown in Supplementary Table 2. For comparison of hADAR2-D mutants, deamination kinetics was determined as described above with the following modifications. The final reaction conditions were 300 nM hADAR2d, 10 nM RNA, 16 mM Tris HCl, pH 7.4, 3.6% glycerol, 1.6 mM EDTA, 0.003% NP-40, 60 mM KCl, 8.6 mM NaCl, 0.5 mM DTT, 160 units/ml RNasin, and 1 µg/ml yeast tRNA.

**Deamination kinetics of transfected hGGL1** RNAs. Deamination kinetics of transfected RNAs was determined as previously described, but with the following modifications. The final reaction volume was 20 µl, the final enzyme concentration was 10 nM, and the final RNA concentration was 2 nM. The final reaction conditions were 17 mM Tris HCl, pH 7.4, 5.0% glycerol, 1.6 mM EDTA, 0.003% NP-40, 60 mM KCl, 15.6 mM NaCl, 0.5 mM DTT, 160 units/ml RNasin, and 1 µg/ml yeast tRNA. Reactions were quenched by addition of 10 µl 0.5% SDS at 95 °C and incubation at 95 °C for 5 min. cDNA was generated from RNA via RT–PCR, Sanger sequenced and quantified with SeqScanner 2 software from Applied Biosystems. The $k_{cat}$ (min$^{-1}$) of each assay was calculated as described previously.

45. Macbeth, M.R. & Bass, B.L. Large-scale overexpression and purification of ADARs from *Saccharomyces cerevisiae* for biophysical and biochemical studies. Methods Enzymol. 424, 319–331 (2007).
46. Pokharel, S. et al. Matching active site and substrate structures for an RNA editing reaction. *J. Am. Chem. Soc.* 131, 11882–11891 (2009).
47. Kabsch, W. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 (2010).
48. McCoy, A.J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674 (2007).
49. Alfonine, P.V. et al. Towards automated crystallographic structure refinement with phaserefin.* Acta Crystallogr. D Biol. Crystallogr. 68, 352–367 (2012).
50. Phelps, K.J., Ibarra-Soza, J.M., Tran, K.; Fisher, A.J. & Beal, P.A. Click modification of RNA at adenosine: structure and reactivity of 7-ethynyl- and 7-triazolyl-8-aza-7-deazaadenosine in RNA. *ACS Chem. Biol.* 9, 1780–1787 (2014).