Title

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

Journal

Nature Structural & Molecular Biology, 23(5)

ISSN

1545-9993

Authors

Matthews, Melissa M

Thomas, Justin M

Zheng, Yuxuan

et al.

Publication Date

2016-05-01

DOI

10.1038/nsmb.3203

Peer reviewed

Published in final edited form as:

Nat Struct Mol Biol. 2016 May ; 23(5): 426–433. doi:10.1038/nsmb.3203.

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

Melissa M. Matthews[1,†], **Justin M. Thomas**[1,†], **Yuxuan Zheng**[1], **Kiet Tran**[1], **Kelly J. Phelps**[1], **Anna I. Scott**[2], **Jocelyn Havel**[1], **Andrew J. Fisher**[1,2], and **Peter A. Beal**[1]

[1]Department of Chemistry, University of California, Davis, CA, USA

[2]Department of Molecular and Cellular Biology, University of California, Davis, CA, USA

Abstract

ADARs (adenosine deaminases acting on RNA) are editing enzymes that convert adenosine (A) to inosine (I) in duplex RNA, a modification reaction with wide-ranging consequences on RNA function. Our understanding of the ADAR reaction mechanism, origin of editing site selectivity and 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 deaminase domain of human ADAR2 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 for ADAR deaminase domain's dsRNA specificity, its base-flipping mechanism, and nearest neighbor preferences. In addition, an ADAR2-specific RNA-binding loop was identified near the enzyme active site 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.

Introduction

RNA editing reactions alter a transcript's genomically encoded sequence by inserting, deleting or modifying nucleotides[1]. Deamination of adenosine (A), the most common form of RNA editing in humans, generates inosine (I) at the corresponding nucleotide position. Since I base pairs with cytidine (C), it functions like guanosine (G) in cellular processes such as splicing, translation and reverse transcription[2,3]. A to I editing has wide-ranging

Correspondence to: pabeal@ucdavis.edu, ajfisher@ucdavis.edu.
†These authors contributed equally to this work.

Accession codes. Coordinates for the hADAR2d-RNA complexes have been deposited in the Protein Data Bank (PDB) with PDB IDs of 5ED1 and 5ED2 for the hADAR2d-E488Q–Bdf2-C RNA and hADAR2d-E488Q–Gli1 RNA complexes, respectively, and 5HP2 and 5HP3 for the hADAR2d-WT–Bdf2-U RNA and hADAR2d-WT–Bdf2-C RNA complexes, respectively.

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/refined crystal structure. J.M.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.

consequences on RNA function including altering miRNA recognition sites, redirecting splicing and changing the meaning of specific codons[4–6]. Two different enzymes carry out A to I editing in humans; ADAR1 and ADAR2[7]. ADAR activity is required for nervous system function and altered editing has been linked to neurological disorders such as epilepsy and Prader Willi Syndrome[8–10]. In addition, mutations in the *ADAR1* gene are known to cause the autoimmune disease Aicardi-Goutieres Syndrome (AGS) and the skin disorder Dyschromatosis Symmetrica Hereditaria (DSH)[11–13]. Hyper editing has been observed at certain sites in cancer cells, such as in the mRNA for *AZIN1* (antizyme inhibitor 1)[14,15]. However, hypo editing also occurs in cancer-derived cell lines exemplified by reduced editing observed in the message for glioma-associated oncogene 1 (*Gli1*)[16].

The ADAR proteins have a modular structure with double stranded RNA binding domains (dsRBDs) and a C-terminal deaminase domain (see Fig. 1a for hADAR2 domains)[17]. ADARs efficiently deaminate specific adenosines in duplex RNA while leaving most adenosines unmodified[18]. The mechanism of adenosine deamination requires ADAR to flip the reactive base out of an RNA double helix to access its active site[17]. How an enzyme could accomplish this task with a duplex RNA substrate is not known. Furthermore, how an ADAR deaminase domain contributes to editing site selectivity is also unknown, since 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 (aa299–701, hADAR2d) bound to different duplex RNAs and solve structures for the resulting complexes using x-ray crystallography. We then evaluated the importance of protein-RNA contacts 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 to generate the inosine-containing product RNA (for reaction scheme see Fig. 1b)[17]. The covalent hydrate of the nucleoside analog 8-azanebularine (N) mimics the proposed high-energy intermediate (for reaction scheme see Fig. 1b)[19]. For trapping hADAR2d bound to RNA for crystallography, we incorporated 8-azanebularine into duplex RNAs shown recently to be excellent substrates for deamination by hADAR2d (for duplex sequence see Fig. 1c) (for characterization of protein–RNA complex see Supplementary Fig. 1)[20]. In addition, for one of these duplexes (Bdf2), we positioned the 8-azanebularine opposite either uridine or cytidine 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 been previously crystallized and structurally characterized revealing features of the active site including the presence of zinc[21]. In addition, an inositol hexakisphosphate (IHP) molecule was found 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) that has enhanced catalytic activity[20,22]. A description of the crystallization conditions, X-ray diffraction data collection and solution of the structures can be found in Online Methods.

Four protein-RNA combinations generated diffracting crystals that resulted in high-resolution structures (hADAR2d WT–Bdf2-U, hADAR2d WT–Bdf2-C, hADAR2d E488Q–Bdf2-C, hADAR2d E488Q–Gli1) (Table 1). In each of these complexes, the protein binds the RNA on one face of the duplex over ~ 20 bp using a positively charged surface near the zinc-containing active site (Fig. 2, Supplementary Fig. 2a). The large binding site (1493 $Å^2$ RNA surface area and 1277 $Å^2$ protein surface area buried) observed for hADAR2d is consistent with recent footprinting studies[20]. Both strands of the RNA contact the protein with the majority of these interactions mediated through the phosphodiester-ribose backbone near the editing site (Fig. 2c, Supplementary Fig. 2 b–d).

The structures show a large deviation from A-form RNA conformation at the editing site (Fig. 2, Fig. 3, Supplementary Video 1). The 8-azanebularine is flipped out of the helix and bound into the active site as its covalent hydrate where it interacts with several amino acids including V351, T375, K376, E396 and R455 (Fig. 3a, Supplementary Fig. 3a). The side chain of E396 H-bonds to purine N1 and O6. This interaction was expected given the proposed role of E396 in mediating proton transfers to and from N1 of the substrate adenosine[17]. The 2'-hydroxyl of 8-azanebularine H-bonds to the backbone carbonyl of T375 while the T375 side chain contacts its 3'-phosphodiester. R455 and K376 help position 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. Lastly, 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 H-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 H-bonds to the complementary strand orphaned base and to the 2' hydroxyl of the nucleotide immediately 5' to the editing site (Figs. 3b, 3c). In the four structures reported here, three different combinations of helix-penetrating residue and orphan base are observed (i.e. E488 + U, E488 + C and Q488 + C) and all three combinations show the same side chain and base positions (Figs. 3b, 3c, Supplementary Fig. 4a for overlay of all three). For instance, in the complex with hADAR2d E488Q and the Bdf2-C duplex, the protein recognizes an orphaned C by donating H-bonds from Nε2 to cytosine N3 and from its backbone NH to cytosine O2 (Fig. 3b). In the complex with hADAR2d WT and the Bdf2-U duplex, a similar interaction is observed with the E488 backbone NH hydrogen bonded to the uracil O2 and the E488 side chain H-bonded to the uracil N3H (Fig. 3c). Interestingly, the E488Q mutant was discovered in a screen for highly active ADAR2 mutants and this residue was 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[24]. A purine at the orphan base position (in its anti conformation) would clash with the 488 residue explaining the preference for pyrimidines here.

The interaction of the 488 residue with the orphaned base is reminiscent of an interaction observed for Hha I DNA methyltransfersase (MTase), a duplex DNA modifying enzyme that also uses a base flipping mechanism to access 2'-deoxycytidine (dC) for methylation[25,26]. For that enzyme, Q237 H-bonds to an orphaned dG while it fills the void left by the flipped out dC (Supplementary Fig. 4b)[26]. In addition, two glycine residues flank Q237 allowing the loop to adopt the conformation necessary for penetration into the helix[25]. The flipping loop in ADAR2 (i.e. aa487–489) also has the helix-penetrating residue flanked by glycines. However, unlike the case of the DNA MTase that 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 H-bonding sites on the orphaned base, 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 (Figs. 4a, 4b, 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 creating an unusual U-A "wobble" interaction with adenine N6 and N1 within H-bonding distance to uracil N3H and O2, respectively (Fig. 4c, Supplementary Fig. 3b). This type of wobble pair has been observed before and requires either the imino tautomer of adenine or the enol tautomer of uracil[27]. The ADAR-induced distortion in RNA conformation results in a kink in the RNA strand opposite the editing site (Fig. 4b). This kink is stabilized by interactions of the side chains of R510 and S495 with phosphodiesters in the RNA backbone of the unedited strand (Fig. 4a). Interestingly, ADAR2's flipping loop approach from the minor groove side is like that seen with certain DNA repair glycosylases (e.g. UDG[28], HOGG1[29], and AAG[30]) that also project intercalating residues from loops bound in the minor groove (Supplementary Fig. 5a). However, these enzymes typically bend the DNA duplex at the site of modification to allow for penetration of intercalating residues and damage recognition[31]. While hADAR2d clearly alters the duplex conformation to gain access to the modification site from the minor groove, it does not bend the RNA duplex (Figs. 2a, 2b, 4b). Furthermore, ADARs do not modify duplex DNA. The DNA B-form helix has groove widths and depths that would prevent productive interactions with ADAR. For instance, ADAR can readily penetrate an A-form helix from the minor groove side and place the helix-penetrating residue in the space occupied by the editing site base (Supplementary Fig. 6). However, this residue cannot penetrate the minor groove enough to occupy the base position in a B-form helix (Supplementary Fig. 6). Furthermore, DNA lacks the 2' hydroxyls that are used by ADAR for substrate recognition (Fig. 2c). Indeed, in each of the four complexes reported here, the protein contacts at least five ribose 2' hydroxyl groups (Fig. 2c, Supplementary Fig. 2 b–d). Thus, hADAR2d uses a substrate recognition and base flipping mechanism with similarities to other known nucleic acid-modifying enzymes but uniquely suited for reaction with adenosine in the context of duplex RNA.

Structures explain nearest neighbor preferences

ADARs have a preference for editing adenosines with 5' nearest neighbor U (or A) and 3' nearest neighbor G[18,23,32]. The ADAR2 flipping loop occupies the minor groove spanning the three base pairs that include the nearest neighbor nucleotides flanking the edited base

(Figs. 3b, 3c). As described above, the base pair including the 5' nearest neighbor U (U11-A13' in the Bdf2 duplex) is shifted from the position it would occupy in a typical A-form helix to accommodate the loop (Fig. 4a). Also, the minor groove edge of this pair is juxtaposed to the protein backbone at G489. Modeling a G-C or C-G pair at this position (i.e. 5' G or 5' C) suggests a 2-amino group in the minor groove would clash with the protein at G489 (Fig. 5a, 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 (Figs. 5b, 5c). To determine whether this effect arises from an increase in local duplex stability from the C-G for U-A substitution 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 a base pair with uridine of similar stability to a U-A pair[33], but places an amino group in the minor groove (Fig. 5b). Importantly, this substitution also resulted in an 80% reduction in rate, illustrating the detrimental effect of the amino group in the minor groove at this location. These observations suggest that hADAR2's 5' nearest neighbor preference for U (or A) is due to a destabilizing clash with the protein 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. However, the observed clash is not severe and the enzyme would be able to accommodate G or C 5' nearest neighbors by slight structural perturbations, explaining why this sequence preference is not an absolute requirement.

In each of the hADAR2d-RNA structures reported here, the backbone carbonyl oxygen at S486 accepts an H-bond from the 2-amino group of the G on the 3' side of the edited nucleotide (Fig. 5d). Guanine is the only common nucleobase that presents an H-bond donor in the RNA minor groove suggesting that other nucleotides in this position would reduce editing efficiency. Indeed, mutating this base to A, C or U, while maintaining base pairing at this position, reduced the rate of deamination by hADAR2d in *Gli1* mRNA model substrates (Supplementary Fig. 7 a–b). To test the importance of the amino group on the 3' G in the hADAR2d reaction, we prepared RNA duplex substrates with purine analogs on the 3' side of the edited A (Fig. 5e). We tested a G analog that lacks the 2-amino group (inosine, I) and one that blocks access to this amino group (N^2-methylguanosine (N^2MeG). In addition, we compared a 3' A to a 3' 2AP since 2AP could form the H-bonding interaction observed with S486. We found the substrate with a 3' N^2MeG to be unreactive to hADAR2d-catalyzed deamination confirming the importance of the observed close approach by the protein to the 3' G 2-amino group (Fig. 5f). In addition, the substrate with a 3' I displayed a reduced deamination rate compared to the substrate with a 3' G suggesting the observed H-bond to the 2-amino group contributes to the 3' nearest neighbor selectivity (Fig. 5f). This conclusion is further supported by the observation that deamination in the substrate with a 3' 2AP is faster than in the substrate with a 3' A (Fig. 5f).

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 for R510 ion-pairs with the 3' phosphodiester of the orphaned nucleotide (Figs. 3a, 3c). 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 either glutamine (R510Q) or to 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 involves 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, R348A) substantially reduces editing activity (Fig. 6c). In addition, mutation of G593 to glutamic acid (G593E) resulted in a nearly two orders of magnitude reduction in rate, consistent with 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. 1d, green) (Supplementary Video 2)[21]. 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 different in the family of ADAR1s (Fig. 6d). The substantial difference in sequence between the ADARs in this RNA-binding loop suggests differences in editing site selectivity between the two ADARs arise, at least in part, 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[34]. DNA methylases, DNA repair glycosylases and RNA loop modifying enzymes are known that flip a nucleotide out of a base pair[25,35–37]. However, none of the structurally characterized base-flipping enzymes access their reactive sites from within a normal base-paired RNA duplex. We are aware of one other protein-induced nucleotide flipping from an RNA duplex region[38]. Bacterial initiation factor 1 (IF1) binds to the 30S ribosomal subunit at helix 44 of 16S RNA with A1492 and A1493 flipped out of the helix and bound into 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[39]. Thus, this system is not illustrative of base flipping from a normal duplex and does not involve an enzyme that must carryout 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 (pseudoU synthetase[35], tRNA transglycosylase[40], and restrictocin bound to sarcin/ricin loop of 28S rRNA[37]) (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 do. The fact 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 well understood prior to this work.

In our structures, the flipped out 8-azanebularine is hydrated, mimicking the tetrahedral intermediate predicted for deamination of adenosine (Figs. 1b, 3a, Supplementary Fig. 3 a–b). Our use of 8-azanebularine, with its high propensity to form a covalent hydrate[41], allowed us to capture a true mimic of the tetrahedral intermediate and reveal the interactions

between the deaminase active site and the reactive nucleotide. In addition, 8-azanebularine was found to adopt a 2'-endo sugar pucker with its 2'-hydroxyl H-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[28,31,35]. 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 (see Supplementary Fig. 4a for overlay). Thus, the E488 side chain directly contacts each orphan base, likely by accepting an H-bond from uracil N3H or by donating an H-bond to cytidine N3. The latter interaction requires E488 to be protonated. The pKa of E488 in the ADAR-RNA complex has not been measured, but proximity to H-bond acceptors, such as cytidine N3, and insertion between stacked nucleobases, would undoubtedly elevate this value and could lead to a substantial fraction in the protonated state at physiologically relevant pH. Since 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). While these studies indicate the ADAR2 catalytic domain makes an important contact to the 3' nearest neighbor G, Stefl et al. suggested the 3' G preference arises from dsRBD binding selectivity for ADAR2[42]. These authors 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 *GluR-B* R/G site RNA[42]. They describe an interaction wherein the 3' G 2-amino group H-bonds to the backbone carbonyl of S258 found in the β1-β2 loop of ADAR2's dsRBDII. It is not possible for the S486-3'G interaction we describe here and the S258-3'G interaction reported by Stefl et al. to exist in the same complex since both involve protein loops bound in the RNA minor groove at the same location. Because our structures have captured 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, since 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 is not relevant to catalysis at the editing site[43]. It is also possible that ADAR dsRBD and catalytic domain binding are sequential, with release of the dsRBD from the RNA taking place prior to catalytic domain engagement and base flipping.

Aicardi-Goutieres Syndrome (AGS) and Dyschromatosis Symmetrica Hereditaria (DSH) are human diseases caused by mutations in the human *ADAR1* gene and several of the disease-associated mutations are found in the deaminase domain[11,12]. Given the conservation in RNA binding surface and active site residues, we expect the hADAR1 catalytic domain to bind RNA with a similar orientation of the helix found in our hADAR2d-RNA structures. When one maps the locations of the AGS-associated mutations onto the hADAR2d-RNA complex, two mutations involve residues in close proximity to the RNA (< 4 Å) (Supplementary Fig. 8a). G487 of hADAR2 is found on the flipping loop near the RNA (Fig.

3b). Sequence in this loop is highly conserved among ADARs and corresponds to G1007 in hADAR1 (Supplementary Table 2). An arginine at this position would preclude close approach of the flipping loop to the RNA, preventing E1008 insertion and base flipping into the active site (Supplementary Fig. 8b). This is consistent with the observation that the G1007R mutation in hADAR1 inhibits RNA editing activity[11,44]. Also, K376 forms salt bridges with both the 5' and 3' phosphodiesters of the guanosine on the 3' side of the editing site (Fig. 2). The corresponding residue in hADAR1 (R892) could form similar contacts and the R892H mutation would likely alter this interaction.

In summary, the structures described here establish human ADAR2 as a base-flipping enzyme that uses a unique mechanism well suited for modifying duplex RNA. In addition, this work provides a basis for understanding the role of the ADAR catalytic domain in determining editing site selectivity and additional structural context to evaluate the impact of ADAR mutations associated with human disease.

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-[^{32}P] ATP was purchased from Perkin-Elmer Life Sciences. The Avian Myeloblastosis Virus (AMV) reverse transcriptase, deoxynucleotide triphosphate (dNTP) mix and RQ1 RNase free DNase were purchased from Promega. Pfu Ultra II was purchased from Stratagene. Dpn 1 was purchased from Invitrogen. 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 Dharmacon, Inc. or Sigma Aldrich. DNA oligonucleotides were purchased from Integrated DNA Technologies. Storage phosphor imaging plates from Molecular Dynamics were imaged using Molecular Dynamics 9400 Typhoon phosphorimager. Data were analyzed using Molecular Dynamics ImageQuant 5.2 software. Electrospray Ionization (ESI) mass spectrometry of oligonucleotide samples was carried out at the Campus Mass Spectrometry Facilities, UC Davis. Oligonucleotide masses were determined using 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 protocol[45]. In brief, 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 media minus uracil (CM-ura) plates. A single colony was used to inoculate a 15 mL CM-ura starter culture. After shaking at 300 rpm and 30 °C overnight, 10 mL of starter culture was used to inoculate each liter of yeast growth media. 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, 35 mM imidazole, 1mM BME, 0.01% Triton × 100) with 750 mM NaCl using a microfluidizer and cell lysate clarified by centrifugation (39,000 × *g* for 25 min). Lysate was passed over a 5 mL Ni-NTA column; washed in three steps with 20–50mL of Lysis Buffer, Wash I buffer (Buffer A + 300 mM NaCl), and Wash II buffer (Buffer A + 100 mM NaCl); and protein eluted by a 35–300 mM imidazole gradient in Wash II 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 10xHis fusion protein was cleaved with an optimized ratio of 1 mg of TEV protease for each 1 mg of protein. Cleavage was carried out for 1–2 h before passing the product over another Ni-NTA column at 0.5 mL/min. The flow-through and wash were collected; dialyzed against 20 mM Tris pH 8.0, 200 mM NaCl, 5% glycerol, and 1 mM BME; and concentrated to just under 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) phosphoramidite was synthesized as previously described[19] and RNAs were synthesized as previously described[46]. Single-stranded RNAs (**See** Supplementary Table 2 **for sequences**) were purified by denaturing polyacrylamide gel electrophoresis and visualized using UV shadowing. Bands were excised from the gel, crushed and soaked overnight at 4 °C in 0.5 M NH_4OAc, 0.1% sodium dodecyl sulfate (SDS) and 0.1 mM EDTA. Polyacrylamide fragments were removed using a 0.2 μm filter followed by desalting on C18 Sep-Pak column. The RNA solutions were lyophilized to dryness, re-suspended 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 Dharmacon-GE Life Sciences and purified as described above. Duplex RNA was hybridized in a 1:1 ratio by heating to 95 °C for 5 min and slowly cooling to 30 °C.

hADAR2d-RNA complex crystallization

Crystals of 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 μM of Bdf2-C 23mer RNA (1:0.7 ADAR2:RNA molar ratio) were mixed with 0.5 μL of 0.1 M MES:NaOH pH 6.5, 9% (w/v) PEG 3350, 13% glycerol, and 0.015M NAD, which was added to improve crystal growth. Crystals took several weeks to grow. A single, cube-shaped crystal about 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 using 0.2° oscillations on beamline 24-ID-C at the Advanced Photon Source at 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, no NAD). For the hADAR2d+WT:Bdf2-U construct, hanging drop vapor diffusion using 200 nL of a mixture containing 4.5 mg/mL protein and 70 μM of Bdf2-U (1:0.7 molar ratio) and 200 nL of a mother liquor (0.1 M ammonium acetate, 0.1M Bis-tris pH 5.5, 17% PEG 10,000) yielded several crystals with a morphology similar to the

one 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 using 0.2° oscillations on beamline 12-2 at the Stanford Synchrotron Radiation Lightsource. Crystals of the hADAR2d E488Q+Gli1 RNA complex were grown using hanging drop vapor diffusion. A solution of volume 200 nL containing 4.5 mg/mL protein and 100 μM of Gli1 23mer RNA (1:1 ADAR2:RNA molar ratio) were mixed with 200 nL of 0.1 M MES:NaOH pH 6.5 and 12% PEG 20,000. At room temperature, a single diamond-shaped crystal about 150 μm long and 50 μm wide was observed about 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 using the fine-phi splicing described above.

Processing and refinement of crystallographic data

Data for the E488Q Bdf2-C-bound and Gli1-bound structures were processed using XDS[47] and scaled with Aimless (CCP4 1994). Diffraction data for hADAR2d wild type structures were processed with XDS and scaled with SCALA (Kabsch, 2010). The RNA-free hADAR2d crystal structure (PDB ID: 1ZY7)[21] was used as a model for molecular replacement using PHENIX[48]. The structures were refined using PHENIX[49] including TLS parameters and Zn coordination restraints. Ideal Zn-ligand distances were determined using average distances found for similar coordination models in the PDB database. Table 1 gives the statistics in data processing and model refinement. The asymmetric unit for Gli1-bound hADAR2d E488Q includes 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 (Pro701) are disordered and were therefore not included in the model. However, although the RNA-free ADAR2 structure (PDB ID: 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 a several rounds of refinement, electron density revealed the location of some side chains. Residues whose basic side chains interact with the RNA backbone are clearly defined in the final density map. Although some non-RNA-binding side chains show only weak density, backbone density is strong. As observed in the original hADAR2d RNA-free structure, inositol hexakisphosphate (IHP) was buried in the enzyme core[21]. The asymmetric units for Bdf2-bound ADARs contain one ADAR2d:RNA complex (protein chain A) and one RNA-free ADAR2d monomer (chain D). The N-terminus of the Bdf2-bound structures include more residues than Gli1-bound, beginning at Pro305 in chain A and Thr304 for chain D in the mutant structure, and beginning at Arg307 in chain A and Thr304 or Pro305 in chain D in the wild type structures. The first few residues (in structures in which the specified residues are modeled) had weak side chain density, including residues 305 and 307 in chain A, and residues 304–307 in chain D, and are modeled in the structure as alanine. The last residue of E488Q+Bdf2-C, Pro701, 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 Glu466 in the wildtype structures, we were able to model-build in main chain and side chain atoms for all residues of this loop in the ADAR subunit complexed to 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 wildtype structures. In the wt+Bdf2-C complex, density for residues 467–470 was strong enough for side chains to be included in these structures; however, side chain density was not strong for residues 465, 466, 471, 473–475, and 477. Therefore, these side chains were not included in the model. In wt+Bdf2-U, density for side chains 465, 466, 470, 471, and 473–475 was too weak to model. IHP was observed in all ADAR deaminase domains in the asymmetric unit. To model the hydrated 8-azanebularine nucleotide in all RNAs, a CCP4 dictionary file for adenosine (A) was modified to replace the 6-amino group with hydrogen, to change atom 8 to nitrogen and to include an additional hydroxyl group off carbon 6. Additionally, an energy minimization calculated idealized structure was used to determine ideal bond angles and distances for the modified base of the hydrated 8-azanebularine (unpublished data, Professor Dean Tantillo, University of California- Davis). The refinement restraint dictionary file was edited to match these parameters.

Expression and purification of hADAR2d for *in vitro* deamination kinetics

Histidine-tagged human ADAR2 deaminase domain (hADAR2d) and hADAR2d mutant proteins were expressed in *S. cerevisiae* strain BCY123 and purified as described above with the following modifications. Cell lysate was 0.45 μm filtered after centrifugation and loaded 3 times through 5 mL Ni-NTA Superflow (Qiagen) at 3 mL/min. Washes of 50 ml with buffer 1, 2 and 3 at 4 mL/min followed by elution with 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 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 10,000 MWCO Amicon Ultra 4 centrifugal filter at 6500 RCF and 4 °C. TEV protease cleavage and gel filtration steps were omitted. Buffer exchange was accomplished via 3 rounds of concentration to <300 μL followed by addition of 3 mL of Storage buffer. After final concentration, protein concentrations were determined using BSA standards visualized by SYPRO Orange staining on SDS-polyacrylamide gels and the purified proteins were stored at −70 °C.

Site-directed mutagenesis

Mutagenesis of hADAR2 catalytic domain was carried out via PCR site directed mutagenesis using the primers listed in Supplementary Table 2. All primers were purchased from IDT and PAGE purified as described above but were desalted by phenol chloroform extraction, ethanol precipitation and 70% ethanol wash instead of C18 Sep-Pack. Sequences for mutant plasmids were confirmed by Sanger sequencing.

Preparation of hGLI1 splint-ligated RNA

Oligonucleotides were purified as described above but were desalted by phenol chloroform extraction, ethanol precipitation and 70% ethanol wash. The 3’ GLI1 top strand 12 mer RNAs were radiolabeled with [γ-^{32}P] at the 5’ end with T4 PNK as described previously[50]. Labeled 3’ GLI1 top strand 12 mer RNAs were ligated as previously described to give an

internally labeled RNA. The splint ligation products were PAGE purified as described above. Labeled RNAs were hybridized with the complementary GLI1 bottom strand 24 mer RNA (Y is chosen based on the identity of X, see Fig. 3b) in 10 mM Tris-HCl, 0.1 mM EDTA pH 7.5 and 100 mM NaCl. See Supplementary Table 2 for RNA sequences.

In vitro deamination kinetics with internally ^{32}P labeled substrates

Deamination kinetics of analog containing RNAs were carried out as previously described[45] but with the following modifications. The final reaction volume was 10 μL. Final enzyme concentration was 300 nM. Final RNA concentration was 10 nM. Final reaction conditions were: 16 mM Tris HCl pH 7.4, 3.3% glycerol, 1.6 mM EDTA, 0.003% Nonidet NP-40, 60 mM KCl, 7.1 mM NaCl, 0.5 mM DTT, 160 units/mL Rnasin, 1 μg/mL yeast tRNA. Reactions were quenched by adding 190 μL 95 °C nuclease-free water followed by incubation at 95 °C for 5 min or by 10 μL 0.5% SDS at 95 °C followed by incubation at 95 °C for 5 min. Each experiment was carried out in triplicate, and the rate constants reported in the text are average values ± standard deviations. Sequences of RNAs used to prepare internally labeled substrates are shown in Supplementary Table 2. For comparison of hADAR2-D mutants, deamination kinetics were carried out as described above with the following modifications. 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% Nonidet NP-40, 60 mM KCl, 8.6 mM NaCl, 0.5 mM DTT, 160 units/mL Rnasin, 1 μg/mL yeast tRNA.

EMSA analysis of radioactively labelled RNA

Duplex RNAs containing 8-azanebularine and ^{32}P labeled were prepared as previously described[20]. Samples containing 0.25 nM RNA and different concentrations of hADAR2d E488Q (128, 64, 32, 16, 8, 4, 2, 1, 0.5, 0.25 and 0 nM) were equilibrated in 20 mM Tris-HCl, pH 7, 6% glycerol, 0.5 mM DTT, 60 mM KCl, 20 mM NaCl, 0.1 mM BME, 1.5 mM EDTA, 0.003% NP-40, 160 units/ml RNasin, 100 μg/ml BSA and 1.0 μg/ml yeast tRNA for 30 min at 30 °C. Assay and data analysis were carried out as previously described[20]. See Supplementary Table 2 for RNA sequences.

In vitro transcription of RNA

A truncation of *hGLI1* mRNA incorporating 81 nucleotides upstream and 65 nucleotides downstream of the edited site was transcribed and purified as previously described[23]. 3' Nearest neighbor mutants of *hGLI1* RNA were generated by site directed mutagenesis to generate G to A, G to C and G to U nearest neighbor mutants. A second site −32 bases from the edit site was mutated to maintain the original secondary structure of the RNA. See Supplementary Table 2 for primers used for mutagenesis.

Deamination kinetics of transcribed *hGLI1* RNAs

Deamination kinetics of transcribed RNAs were carried out as previously described[23] but with the following modifications. Final reaction volume was 20 μL. Final enzyme concentrations was 10 nM. Final RNA concentration was 2 nM. Final reaction conditions were: 17 mM Tris HCl pH 7.4, 5.0% glycerol, 1.6 mM EDTA, 0.003% Nonidet NP-40, 60 mM KCl, 15.6 mM NaCl, 0.5 mM DTT, 160 units/mL RNasin, 1 μg/mL yeast tRNA.

Reactions were quenched by adding 10 μL 0.5% SDS at 95 °C followed by incubation at 95 °C for 5 min. cDNA was generated from RNA via RT-PCR, Sanger sequenced and quantified using SeqScanner 2 software from Applied Biosystems. The k_{obs} (min^{-1}) of each assay was calculated as described previously[23].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge funding from the US National Institutes of Health (NIH) grand # R01GM061115 (P.A.B). A.I.S. was supported by NIH training grant T32 GM007377. C. Palumbo is acknowledged for technical assistance. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the US Department of Energy Office of Biological and Environmental Research, and by the NIH, US National Institute of General Medical Sciences (including P41GM103393). Part of this work is also based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the NIH (P41 GM103403). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

References

1. Grosjean, H. Fine-Tuning of RNA Functions by Modification and Editing. Springer; 2005.
2. Bass BL. RNA editing by adenosine deaminases that act on RNA. Ann. Rev. Biochem. 2002; 71:817–846. [PubMed: 12045112]
3. Nishikura K. Functions and regulation of RNA editing by ADAR deaminases. Ann. Rev. Biochem. 2010; 79:321–349. [PubMed: 20192758]
4. Wang Q, et al. ADAR1 regulates ARHGAP26 gene expression through RNA editing by disrupting miR-30b-3p and miR-573 binding. RNA. 2013; 19:1525–1536. [PubMed: 24067935]
5. Rueter SM, Dawson TR, Emeson RB. Regulation of alternative splicing by RNA editing. Nature. 1999; 399:75–80. [PubMed: 10331393]
6. Yeo J, Goodman RA, Schirle NT, David SS, Beal PA. RNA editing changes the lesion specificity for the DNA repair enzyme NEIL1. Proc. Natl. Acad. Sci. USA. 2010; 107:20715–20719. [PubMed: 21068368]
7. Bass BL, et al. A standardized nomenclature for adenosine deaminases that act on RNA. RNA. 1997; 3:947–949. [PubMed: 9292492]
8. Maas S, Kawahara Y, Tamburro KM, Nishikura K. A to I RNA Editing and Human Disease. RNA Biology. 2006; 3:1–9. [PubMed: 17114938]
9. Slotkin W, Nishikura K. Adenosine-to-inosine RNA editing and human disease. Genome Med. 2013; 5:105. [PubMed: 24289319]
10. Morabito MV, et al. Mice with altered serotonin 2C receptor RNA editing display characteristics of Prader-Willi syndrome. Neurobiol. Dis. 2010; 39:169–180. [PubMed: 20394819]
11. Rice GI, et al. Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature. Nat. Genetics. 2012; 44:1243–1248. [PubMed: 23001123]
12. Miyamura Y, et al. Mutations of the RNA-specific adenosine deaminase gene (DSRAD) are involved in dyschromatosis symmetrica hereditaria. Am. J. Hum. Genet. 2003; 73:693–699. [PubMed: 12916015]
13. Zhang XJ, et al. Seven novel mutations of the ADAR gene in Chinese families and sporadic patients with disychromatosis symmetrica hereditaria (DSH). Hum. Mutat. 2004; 23:629–630. [PubMed: 15146470]

14. Chen L, et al. Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. Nat. Med. 2013; 19:209–216. [PubMed: 23291631]
15. Gallo A. RNA editing enters the limelight in cancer. Nat. Med. 2013; 19:130–131. [PubMed: 23389604]
16. Shimokawa T, et al. RNA editing of the GLI1 transcription factor modulates the output of Hedgehog signaling. RNA Biology. 2013; 10:321–333. [PubMed: 23324600]
17. Goodman RA, Macbeth MR, Beal PA. ADAR proteins: Structure and Catalytic Mechanism. Curr. Top. Microbiol. Immunol. 2012; 352:1–33. [PubMed: 21769729]
18. Li JB, et al. Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. Science. 2009; 324:1210–1213. [PubMed: 19478186]
19. Haudenschild BL, et al. A transition state analogue for an RNA-editing reaction. J. Am. Chem. Soc. 2004; 126:11213–11219. [PubMed: 15355102]
20. Phelps KJ, et al. Recognition of Duplex RNA by the Deaminase Domain of the RNA editing Enyzme ADAR2. Nucleic Acids Res. 2015; 43:1123–1132. [PubMed: 25564529]
21. Macbeth MR, et al. Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA editing. Science. 2005; 309:1534–1539. [PubMed: 16141067]
22. Kuttan A, Bass BL. Mechanistic Insights into edtiing-site specificity of ADARs. Proc. Natl. Acad. Sci. USA. 2012; 109:E3295–E3304. [PubMed: 23129636]
23. Eifler T, Pokharel S, Beal PA. RNA-Seq Analysis Identifies A Novel Set of Editing Substrates for Human ADAR2 Present in Saccharomyces cerevisiae. Biochemistry. 2013; 52:7857–7869. [PubMed: 24124932]
24. Wong SK, Sato S, Lazinski DW. Substrate recognition by ADAR1 and ADAR2. RNA. 2001; 7:846–858. [PubMed: 11421361]
25. Klimasauskas S, Kumar P, Roberts RJ, Cheng X. HhaI methyltransferase flips its target base out of the DNA helix. Cell. 1994; 76:357–369. [PubMed: 8293469]
26. Daujotyte D, et al. HhaI DNA Methyltransferase Uses the Protruding Gln237 for Active Flipping of Its Target Cytosine. Structure. 2004; 12:1047–1055. [PubMed: 15274924]
27. Thiyagarajan S, Rajan SS, Gautham N. Cobalt hexamine induced tautomeric shift in Z-DNA: the structure of d(CGCGCA).d(TGCGCG) in two crystal forms. Nucleic Acids Res. 2004; 32:5945–5953. [PubMed: 15534365]
28. Slupphaug G, et al. A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase boudn to DNA. Nature. 1996; 384:87–92. [PubMed: 8900285]
29. Bruner SD, Norman DP, Verdine GL. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. Nature. 2000; 403:859–866. [PubMed: 10706276]
30. Lau AY, Scharer OD, Samson L, Verdine GL, Ellenberger T. Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. Cell. 1998; 95:249–258. [PubMed: 9790531]
31. Brooks SC, Adhikary S, Rubinson EH, Eichman BF. Recent advances in the structural mechanisms of DNA glycosylases. Biochim. Biophys. Acta. 2013; 1834:247–271. [PubMed: 23076011]
32. Eggington JM, Greene T, Bass BL. Predicting sites of ADAR editing in double stranded RNA. Nat. Commun. 2011; 2:1–9.
33. Peacock H, Maydanovych O, Beal PA. N2-Modified 2-aminopurine ribonucleosides as minor-groove-modulating adenosine replacements in duplex RNA. Org. Lett. 2010; 12:1044–1047. [PubMed: 20108910]
34. Roberts RJ, Cheng X. Base flipping. Ann. Rev. Biochem. 1998; 67:181–198. [PubMed: 9759487]
35. Hoang C, Ferre-D'Amare AR. Cocrystal Structure of a tRNA psi55 Pseuduridine Synthase: Nucleotide Flipping by an RNA-modifying Enyzme. Cell. 2001; 107:929–939. [PubMed: 11779468]
36. Parikh SS, et al. Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. EMBO J. 1998; 17:5214–5226. [PubMed: 9724657]
37. Yang X, Gerczei T, Glover L, Correll CC. Crystal structures of restrictocin-inhibitor complexes with implications for RNA recognition and base flipping. Nat. Struct. Biol. 2001; 8:968–973. [PubMed: 11685244]

38. Carter AP, et al. Crystal Structure of an Initiation Factor Bound to the 30S Ribosomal Subunit. Science. 2001; 291:498–501. [PubMed: 11228145]

39. Reblova K, et al. Structure, Dynamics, and Elasticity of Free 16S rRNA Helix 44 Studied by Molecular Dynamics Simulations. Biopolymers. 2006; 82:504–520. [PubMed: 16538608]

40. Xie W, Liu X, Huang RH. Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. Nat. Struct. Biol. 2003; 10:781–788. [PubMed: 12949492]

41. Erion MD, Reddy MR. 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. 1998; 120:3295–3304.

42. Stefl R, et al. The Solution Structure of the ADAR2 dsRBM-RNA Complex Reveals a Sequence-Specific Readout of the Minor Groove. Cell. 2010; 143:225–237. [PubMed: 20946981]

43. Fierro-Monti I, Mathews MB. Proteins binding to duplexed RNA: one motif, multiple functions. Trends Biochem. Sci. 2000; 25:241–246. [PubMed: 10782096]

44. Mannion NM, et al. The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. Cell Rep. 2014; 9:1482–1494. [PubMed: 25456137]

Online methods references

45. Macbeth MR, Bass BL. Large-Scale Overexpression and Purification of ADARs from Saccharomyces cerevisiae for Biophysical and Biochemical Studies. Methods Enzymol. 2007; 424:319–331. [PubMed: 17662848]

46. Pokharel S, et al. Matching Active Site Structure to Substrate Analog for an RNA Editing Reaction. J. Am. Chem. Soc. 2009; 131:11882–11891. [PubMed: 19642681]

47. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. [PubMed: 20124692]

48. McCoy AJ, et al. Phaser crystallographic software. J. Appl. Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

49. Afonine PV, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012; 68:352–367. [PubMed: 22505256]

50. Phelps K, Ibarra-Sosa J, Tran K, Fisher AJ, Beal PA. Click Modification of RNA at Adenosine: Structure and Reactivity of 7-Ethynyl- and 7-Triazolyl-8-aza-7-deazaadenosine in RNA. ACS Chem. Biol. 2014; 9:1780–1787. [PubMed: 24896732]

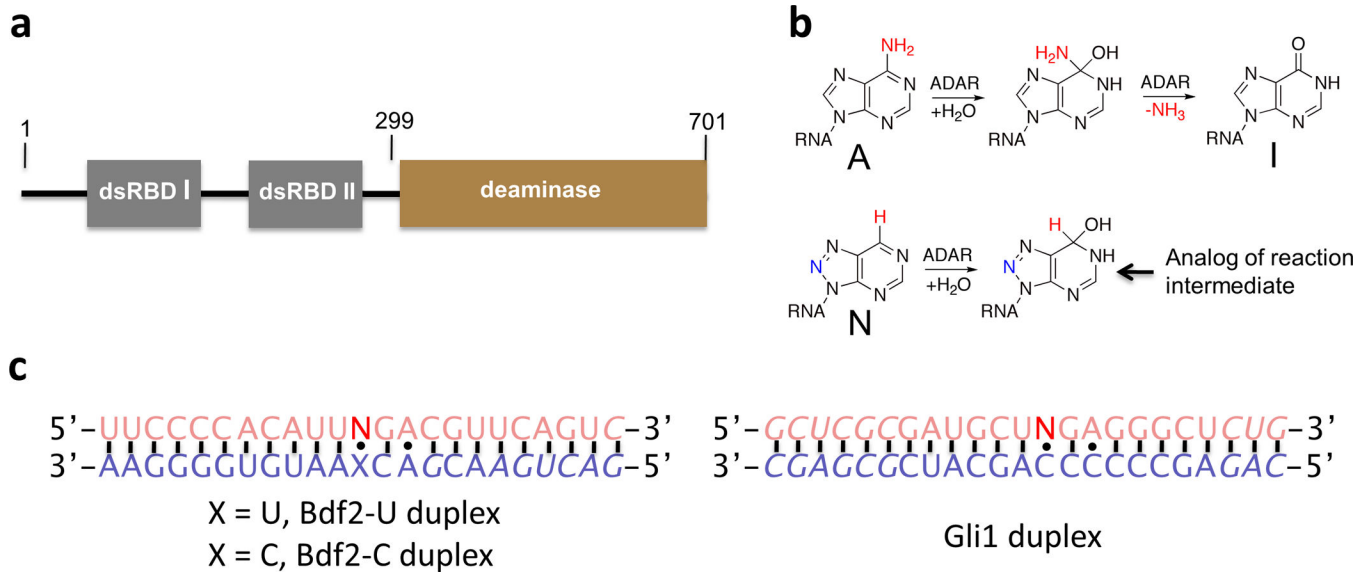

Figure 1. Human ADAR2 and modified RNAs for crystallography

a, Domain map for human ADAR2 **b**, ADAR reaction showing intermediate and 8-azanebularine (N) hydrate that mimics this structure **c**, Duplex RNAs used for crystallization. Bdf2 duplex sequence is derived from an editing site found in *S. cerevisiae* Bdf2 mRNA[23] and Gli1 duplex has sequence surrounding the human *Gli1* mRNA editing site[16]. Italics indicate nucleotides added for duplex stability.

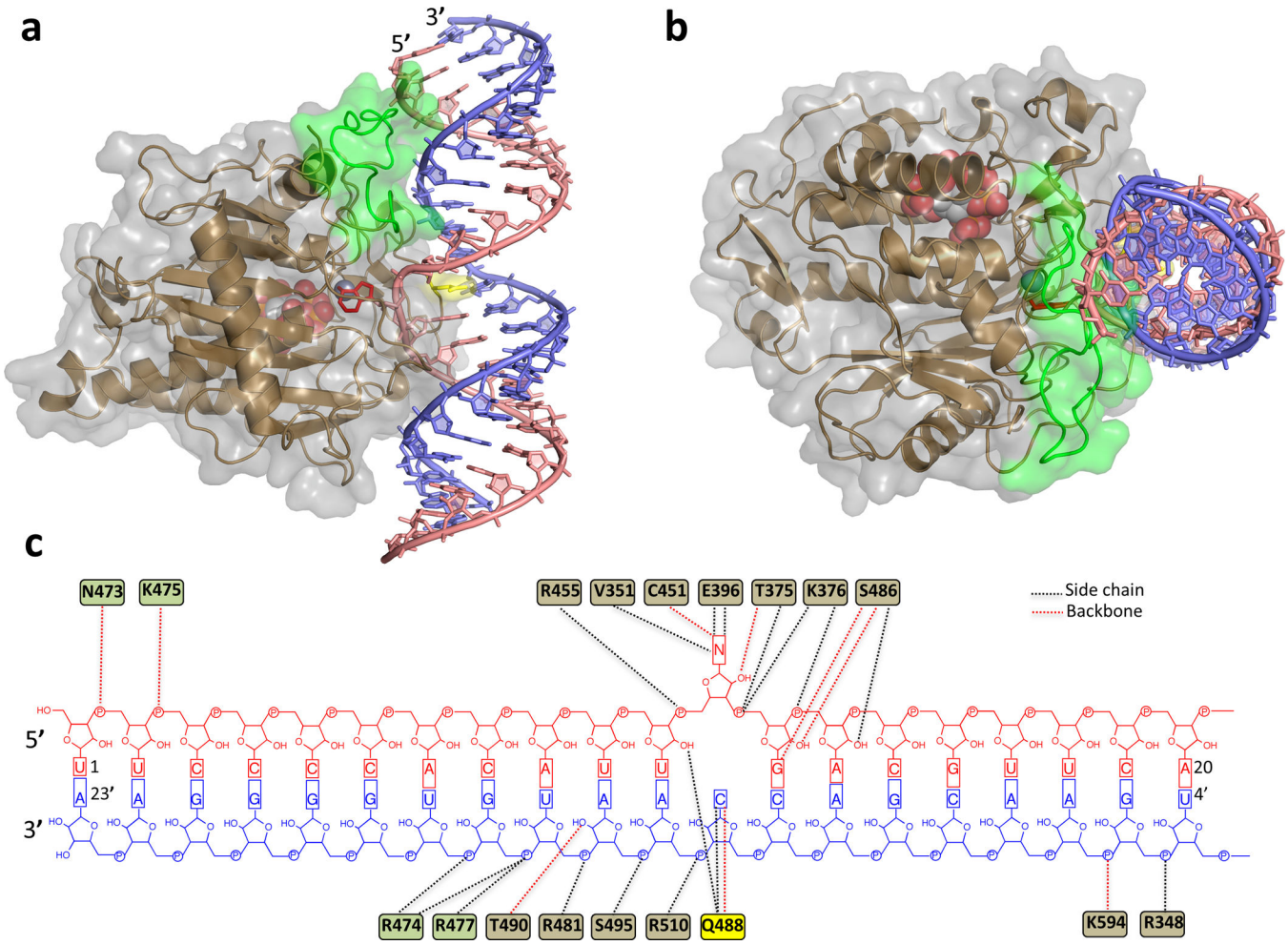

Figure 2. Structure of hADAR2d E488Q bound to the Bdf2-C RNA duplex at 2.75 Å resolution

a, View of structure perpendicular to the dsRNA helical axis. Colors correspond to those in Figs. 1a and 1c; flipped out base N is highlighted red, zinc in grey space-filling sphere, Q488 in yellow, previously disordered aa454–477 loop in green and inositol hexakisphosphate (IHP) in space filling. A transparent surface is shown for the hADAR2d protein. **b**, View of structure along the dsRNA helical axis. **c**, Summary of the contacts between hADAR2d E488Q and the Bdf2-C RNA duplex.

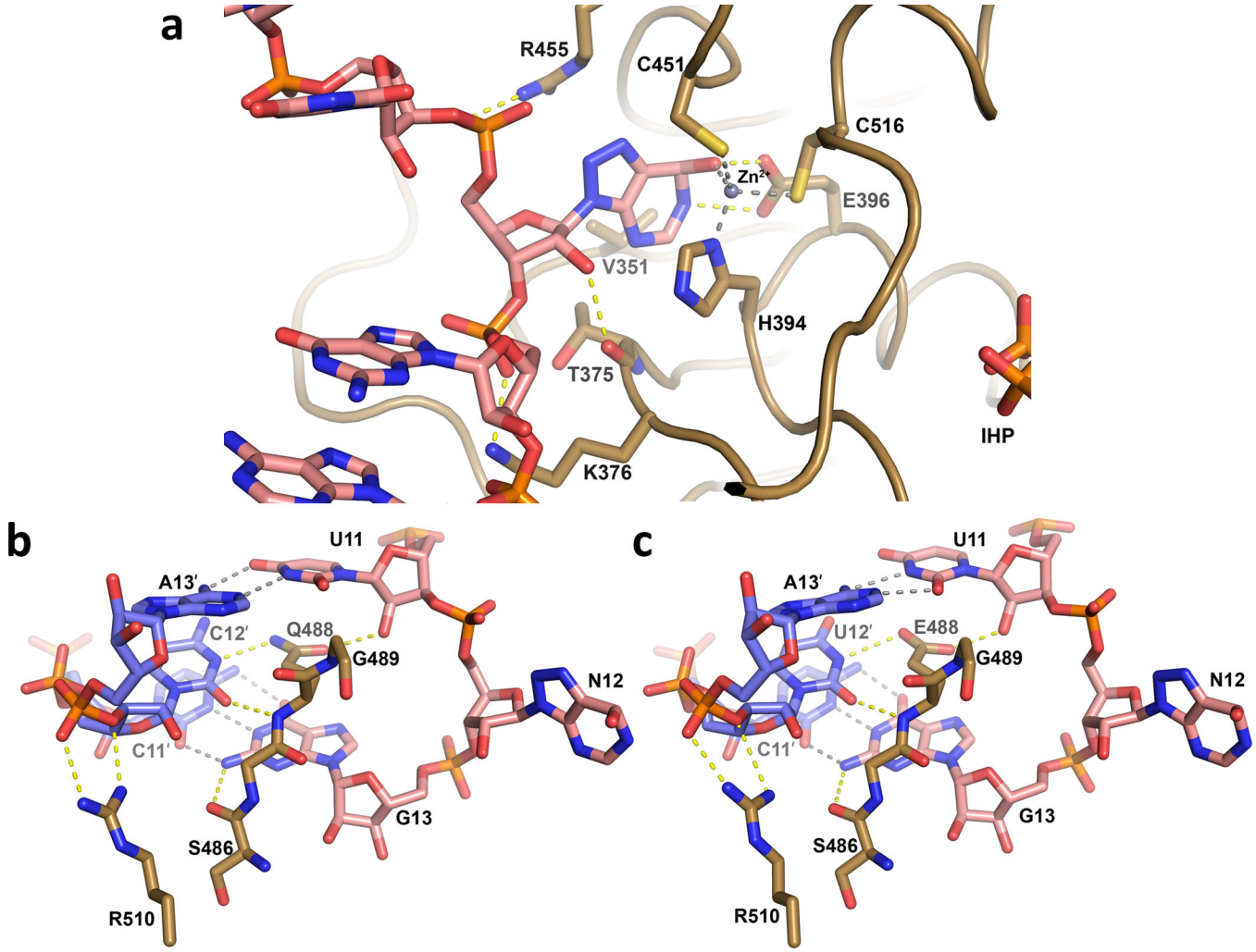

Figure 3. ADAR recognition of the flipped out and orphaned nucleotides

a, Contacts to the editing site nucleotide (N) in the active site. Colors correspond to those in Figs. 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.

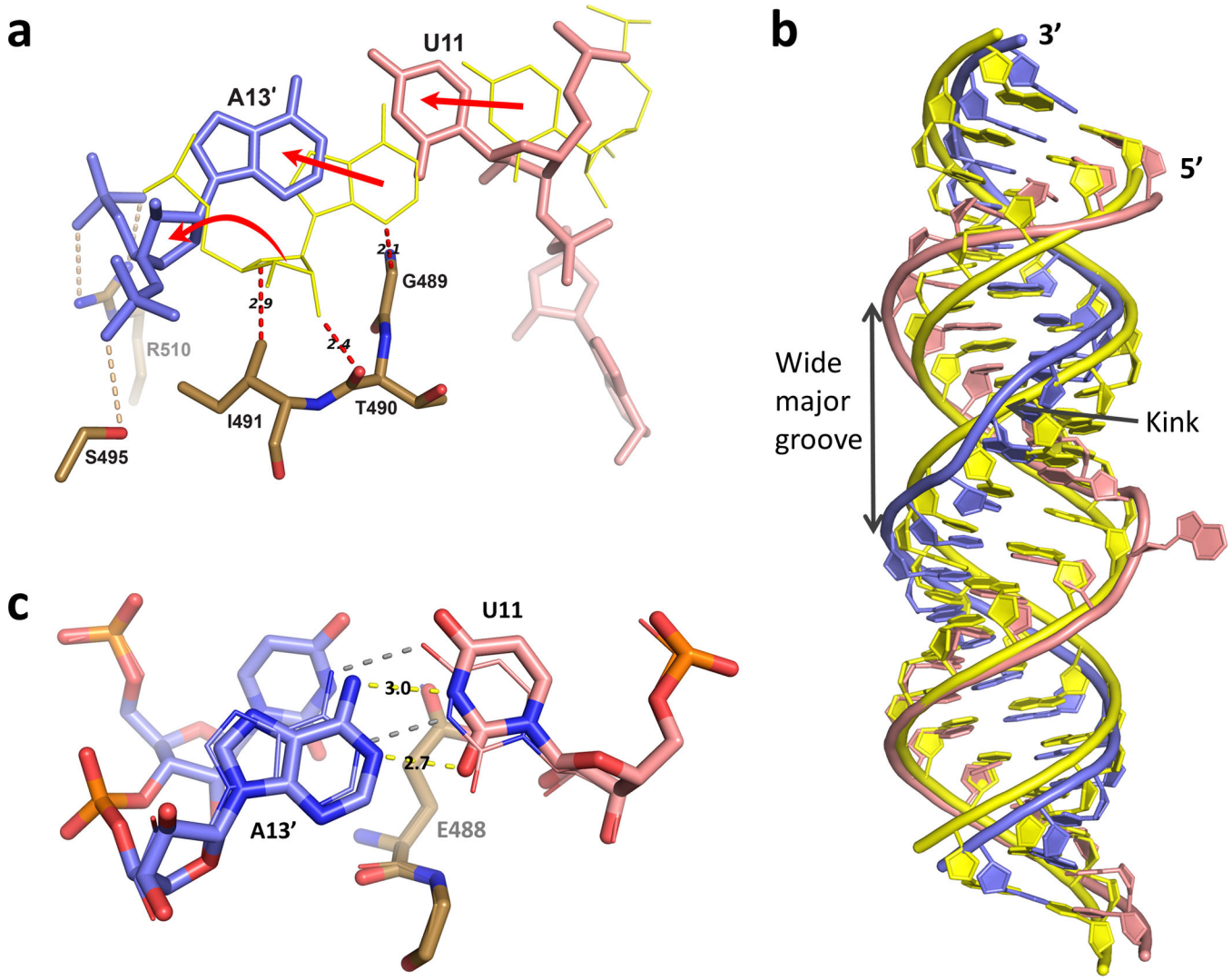

Figure 4. Other ADAR-induced changes in RNA conformation

a, hADAR2d shifts the position of U11-A13' base pair from ideal A-form RNA helix (yellow). **b**, Overlay of Bdf2 duplex RNA and idealized A form duplex of same sequence (yellow) illustrating kink in strand and widening of major groove opposite editing site induced by hADAR2d. **c**, Unusual "wobble" A13'-U11 interaction in the hADAR2d WT–Bdf2-U complex shown in stick with H-bond 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 H-bonds shown with gray dashes.

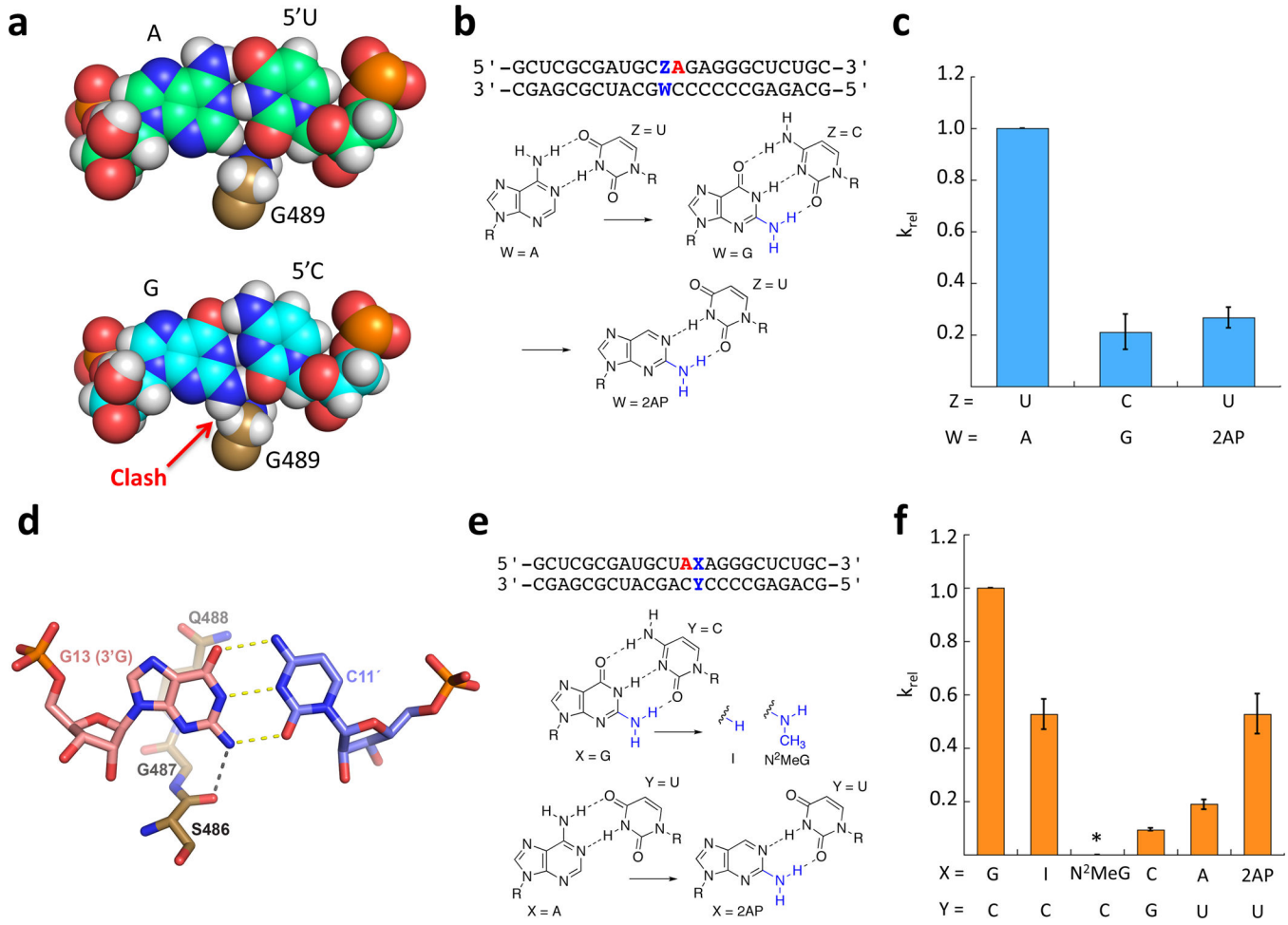

Figure 5. Interactions with editing site nearest neighbor nucleotides

a, 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 G 2-amino group **b**, RNA duplex substrates prepared with different 5' nearest neighbor nucleotides adjacent to 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; $k_{rel} = k_{obs}/(k_{obs}$ for unmodified RNA). Error bars, s.d (n=3 technical replicates). **d**, hADAR2 S486 backbone H-bond with 3' G 2-amino group; **e**, RNA duplex substrates prepared with different 3' nearest neighbor nucleotides adjacent to editing site indicated in red (I = inosine, N^2MeG = N^2-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. $k_{rel} = k_{obs}/(k_{obs}$ for unmodified RNA). Error bars, s.d (n=3 technical replicates). * indicates no reaction product observed.

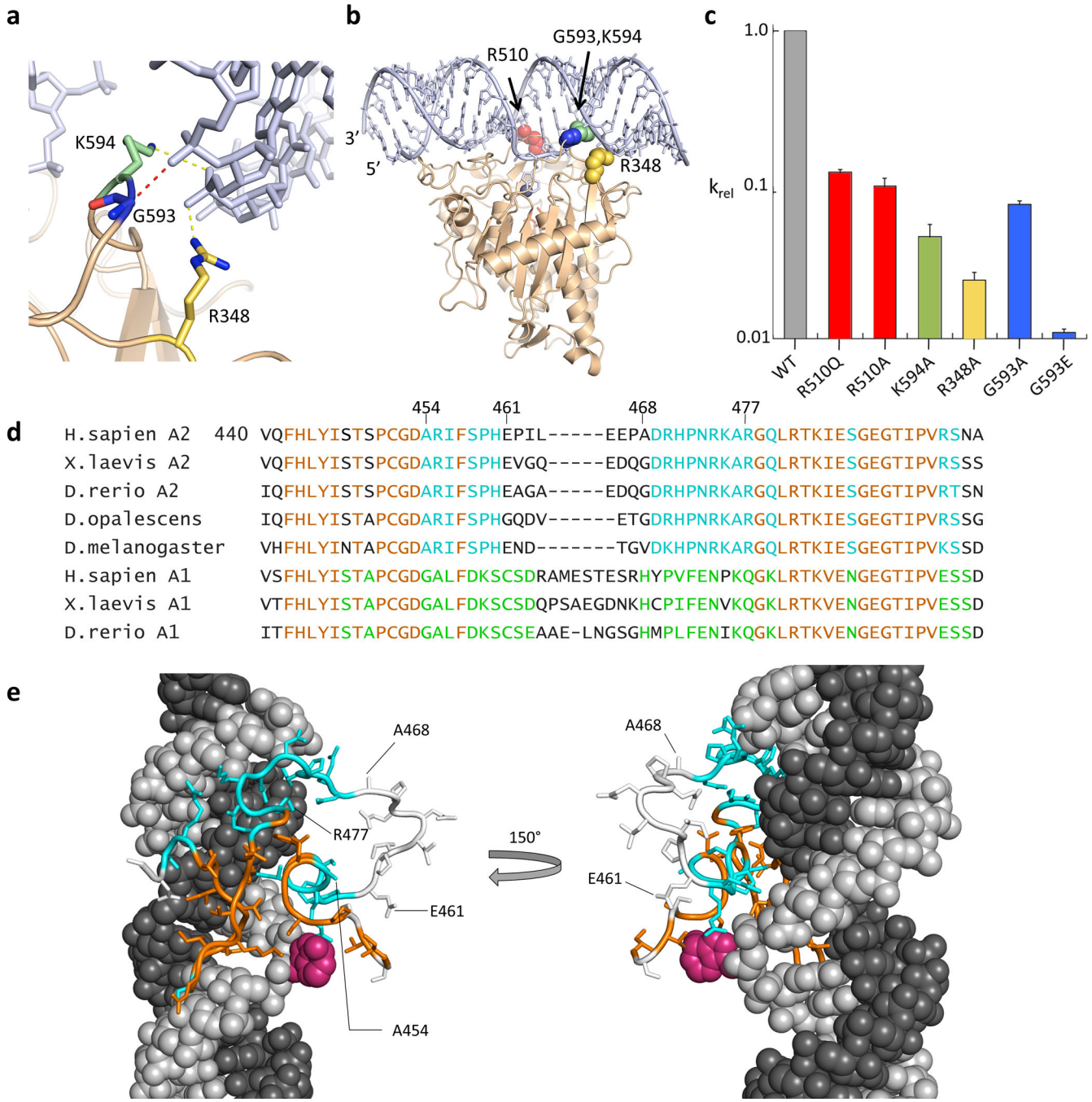

Figure 6. RNA-binding loops in the ADAR catalytic domain

a, hADAR2 residues that contact phosphodiester backbone near 5' end of unedited strand. **b**, Location of mutations introduced at protein-RNA interface. **c**, Comparison of deamination rate constants of the different hADAR2d mutants (Log scale). $k_{rel} = k_{obs}$ for mutant/k_{obs} for WT. Error bars, s.d (n=3 technical replicates). **d**, Sequence alignment of ADAR2s (A2) and ADAR1s (A1) from different organisms with different levels of conservation colored (Yellow: conserved in all ADAR1s and ADAR2s, red: conserved in ADAR2s, blue: conserved in ADAR1s. **e**, Interaction of the ADAR-specific RNA-binding loop near the 5'

end of the edited strand. Colors as in **d**, white: not conserved, flipped out base is shown in pink.

Table 1

Data Processing and Refinement Statistics.

Complex	ADAR2-D_E488Q:BDF2-C 23mer	ADAR2-D_E488Q:GLI1 23mer	ADAR2-D_wt:BDF2-U 23mer	ADAR2-D_wt:BDF2-C 23mer
Data Collection				
Synchrotron (Beamline)	APS (24-ID-C)	SSRL (12-2)	SSRL (12-2)	SSRL (12-2)
Wavelength (Å)	0.9792	0.9795	0.9795	0.9795
Space Group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions				
a, *b*, *c*, (Å)	82.36, 107.50, 121.10	79.13, 81.61, 256.62	81.32, 106.68, 120.49	81.51, 107.21, 120.62
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	100 – 2.75 (2.82 – 2.75)	50.0 – 2.95 (3.03 – 2.95)	100 – 2.98 (3.06 – 2.98)	100 – 3.09 (3.17 – 3.09)
R_{merge} (%)	7.0 (68.5)	9.6 (135.1)	14.4 (86.3)	11.6 (68.7)
$CC_{1/2}$	99.6 (66.3)	99.7 (47.6)	99.1 (75.3)	99.3 (77.1)
I/σ (I)	11.57 (1.52)	12.27 (1.16)	10.00 (1.88)	10.56 (1.75)
Completeness (%)	96.5 (98.8)	98.1 (98.9)	97.3 (90.0)	96.8 (89.1)
Redundancy	2.93 (3.00)	5.19 (5.13)	4.79 (4.56)	3.31 (2.84)
Refinement				
Resolution (Å)	2.75	2.95	2.98	3.09
No of reflections (F>0)	27,153	35,727	21,376	19,325
R_{work}/ R_{free}	16.27/22.34	18.79/20.75	16.67 / 24.67	16.29 / 23.79
No. of atoms				
Protein	6197	6038	6168	6157
RNA	973	1950	973	973
Inositol Hexakisphosphate (IHP)	72	72	72	72
Zn	2	2	2	2
Waters	33	0	1	1
***B* factors**				
Protein	68.46	90.65	63.63	67.92
RNA	88.24	108.8	69.70	77.49
Inositol Hexakisphosphate (IHP)	47.10	65.57	44.23	43.77
Zn	48.47	64.38	38.25	49.04
Waters	48.11	N/A	43.66	50.20
r.m.s deviations				
Bond lengths (Å)	0.010	0.007	0.009	0.010
Bond angles (°)	1.342	0.885	1.359	1.365