Crystal structure of a hepatitis delta virus ribozyme

Adrian R. Ferré-D'Amaré*, Kaihong Zhou† & Jennifer A. Doudna*†

** Department of Molecular Biophysics and Biochemistry, and † Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520-8114, USA*

The self-cleaving ribozyme of the hepatitis delta virus (HDV) is the only catalytic RNA known to be required for the viability of a human pathogen. We obtained crystals of a 72-nucleotide, self-cleaved form of the genomic HDV ribozyme that diffract X-rays to 2.3 Å resolution by engineering the RNA to bind a small, basic protein without affecting ribozyme activity. The co-crystal structure shows that the compact catalytic core comprises five helical segments connected as an intricate nested double pseudoknot. The 5′-hydroxyl leaving group resulting from the self-scission reaction is buried deep within an active-site cleft produced by juxtaposition of the helices and five strand-crossovers, and is surrounded by biochemically important backbone and base functional groups in a manner reminiscent of protein enzymes.

HDV is an RNA satellite virus of hepatitis B virus (HBV). Infection of humans by both HBV and HDV is generally associated with more severe hepatitis than that caused by HBV alone[1]. HDV has a circular, single-stranded RNA genome of 1,700 nucleotides, with ~70% self-complementarity. This genome is thought to be replicated by the host RNA polymerase II through a double-rolling-circle mechanism[1,2]. Replication produces linear multimers of both genomic and antigenomic RNAs, both of which possess a self-cleaving activity that processes the RNAs to unit length[3–5]. The activity has been mapped to a catalytic-RNA, or ribozyme, domain whose function is required for HDV replication *in vivo*[1]. The HDV ribozyme, which is active *in vitro* in the absence of any proteins, is the only known example of a catalytic RNA associated with an animal virus. The circular single-stranded RNA genome, its mode of replication by a host RNA polymerase, and the self-cleaving-RNA activity of HDV have close parallels in the plant viroids. Apart from the genomic and antigenomic ribozymes, however, there are no known homologues of HDV ribozymes, and sequence variation of the HDV ribozymes in clinical isolates is minimal[2].

The self-cleavage reaction catalysed by the HDV ribozyme is a transesterification reaction, which yields products with 2′,3′-cyclic phosphate and 5′-hydroxyl termini. Eighty-five contiguous nucleotides are required for activity of both genomic and antigenomic sequences[6]. Remarkably, the presence of a single nucleotide located immediately 5′ to the cleavage site is sufficient for cleavage, and the identity of this −1 nucleotide has only small effects on the reaction rate[2]. On the basis of sequence comparison and compensatory mutagenesis studies, Perrotta and Been[7] proposed closely related secondary-structure models for the genomic and antigenomic HDV ribozymes that have a pseudoknot as their distinguishing feature (Fig. 1).

The HDV ribozyme is the fastest known naturally occurring self-cleaving RNA, and also stands out among ribozymes because of its stability to denaturants and its lack of requirement for specific metal ions. The first-order rate constant for a genomic ribozyme has been estimated to be 52 reactions per min at 37 °C (ref. 8). As the optimal temperature *in vitro* is about 65 °C (ref. 9), this ribozyme can cleave itself at a rate of more than 1 per second. For comparison, the hammerhead ribozyme cleaves itself at rates of about 1 per min (refs 10, 11), and ribonuclease A cleaves UpG with a first-order rate constant of 69 per second (ref. 12). Remarkably, the HDV ribozyme is active in 5 M urea or 18 M formamide[13,14]. The ribozyme seems to have a nonspecific requirement for divalent cations for activity; cleavage is observed even in very low (<0.1 mM) concentrations of Ca^{2+}, Mg^{2+}, Mn^{2+} or Sr^{2+}, and at a reduced rate in the presence of several other cations[15].

We have now determined the structure of the self-cleaved form of a genomic HDV ribozyme by X-ray crystallography and refined the atomic model against diffraction data extending to 2.3 Å resolution. The HDV ribozyme adopts an intricate fold that buries the active site deep within a catalytic cleft. The structure is in agreement with

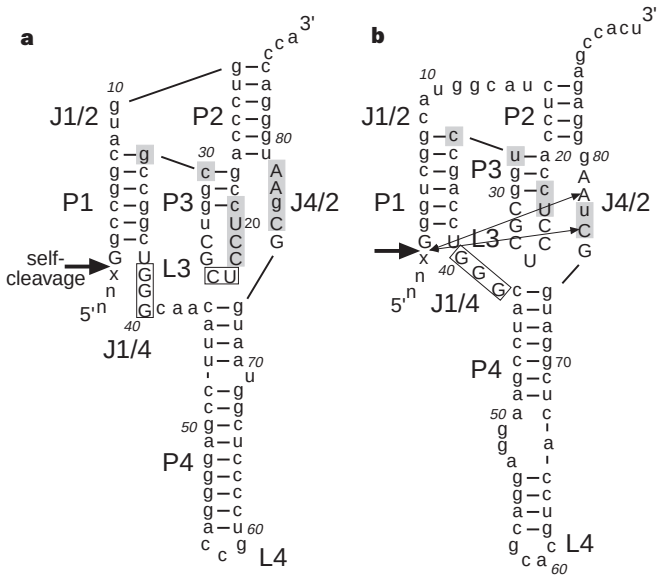

Figure 1 Pseudoknotted secondary-structure models of **a**, genomic and **b**, antigenomic HDV ribozymes proposed by Perrotta and Been[7]. A pseudoknot is a nucleic acid structure characterized by base-pairing between nucleotides in the loop of a conventional hairpin duplex with complementary residues outside the hairpin[48]. RNAs are numbered starting at the cleavage site[1] (large arrows). Base-paired regions are denoted P1 to P4 in a 5′ to 3′ direction; L3 and L4 are loops capping the corresponding paired regions. Single-stranded segments joining helices are indicated by J, followed by the helices they connect, in a 5′ to 3′ direction, as J1/2, J1/4 and J4/2. X denotes the nucleotide preceding the cleavage site, which can have any base. Upper-case letters denote nucleotides whose mutation results in marked reduction in ribozyme activity, and could not be rescued by compensatory mutations elsewhere. Lower-case letters mark nucleotides whose identity seems to be unimportant for ribozyme function[8,9,30]. Boxed nucleotides correspond to those found to be strongly (grey) or weakly (white) protected from hydroxy-radical cleavage[25]. Two-headed arrows mark positions of photocrosslinks formed between an azidophenacyl group placed 5′ to G1 and nucleotides in J4/2 (ref. 25).

existing biochemical data, and provides a three-dimensional scaffold for understanding the mode of action of this unique RNA catalyst.

Crystallization and structure determination

The P4 stem (Fig. 1) of the HDV ribozyme can be drastically shortened or mutagenized without detriment to the ribozyme activity, indicating that it stabilizes the active structure without participating directly in catalysis[2]. We reported previously that placing a crystallization module based on an RNA tertiary interaction into the P4 stem of an HDV ribozyme yielded active RNAs that crystallized readily[16]. To obtain well-ordered crystals, we developed a new type of crystallization module. The solvent-exposed P4 stem was engineered to contain a high-affinity binding site for the small, basic RNA-binding domain of the U1A spliceosomal protein (U1A-RBD, Fig. 2a). U1A-RBD has been extensively characterized biochemically, and the structure of its complex with a cognate 21-nucleotide RNA stem loop has been determined previously at 1.92 Å resolution[17]. The alteration of P4 did not adversely affect the self-cleavage activity of the HDV ribozyme, nor did the binding of U1A-RBD (Fig. 2b). We reasoned that co-crystallization might result in better ordered crystals, because the protein–RNA complex would have a greater variety of surface functional groups than would a 'naked' RNA. This approach enabled crystallization of two unrelated types of ribozyme (A.R.F. and J.A.D., unpublished observations) and might be a powerful general technique for the crystallization of RNA.

As a significant proportion of HDV ribozymes transcribed *in vitro* is misfolded (ref. 2; A.R.F. and J.A.D., unpublished observations), we used self-cleaved HDV ribozymes, which have gone through the transition state, for crystallization. Biochemical[2] and spectroscopic[14] data indicate that the structures of product and precursor ribozymes

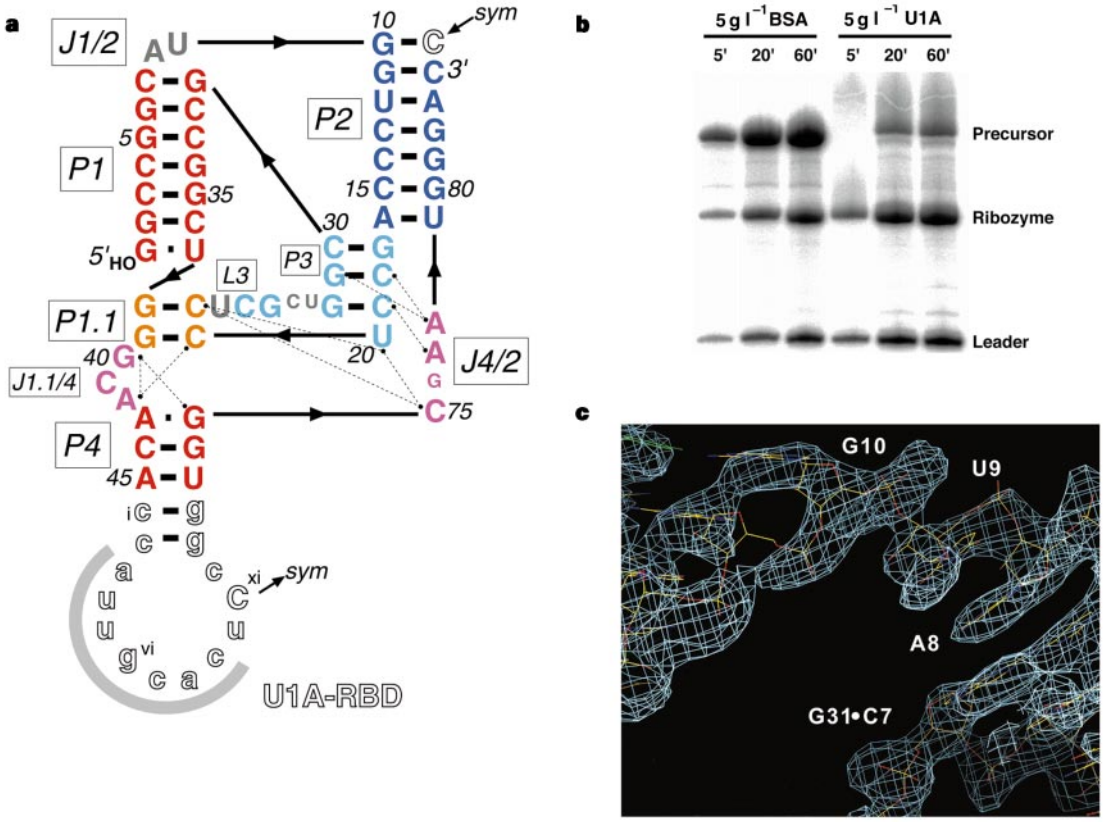

Figure 2 Complex of the HDV ribozyme and U1A-RBD. **a**, The genomic HDV ribozyme RNA used for structure determination, rearranged to reflect aspects of the three-dimensional structure. Upper-case letters correspond to nucleotides in the wild-type HDV ribozyme, coloured to match the three-dimensional structure representations in subsequent figures. Small upper-case letters denote nucleotides that are extruded from helical stacks. Nucleotides of the U1A-binding site engineered into helix P4 (compare with Fig. 1a), and which do not participate in the ribozyme structure, are shown in outlined lower-case letters; those making close contacts with the U1A-RBD face a light grey semicircle. The numbering is as in Fig. 1a, except for the U1A-binding loop, for which the numbering is in lower-case Roman numerals. Thick short lines represent Watson–Crick base pairs; small squares indicate a wobble G·U and an A·G base pair; lines with embedded arrowheads represent covalent connections; and thin dotted lines indicate some long-range contacts found in the structure. The nucleotide Cxi forms base pairs with a symmetry-related molecule completing the 3′ terminus of P2. The complementarity to J1/2 ends at C85 (Fig. 1a), defining the length of P2 of the genomic ribozyme as being seven base pairs. **b**, self-cleavage activity of the crystallized HDV ribozyme construct during transcription (see Methods) in the presence of either 5 g l^{-1} bovine serum albumin (BSA) or U1A-RBD (U1A), and traces of [α-^{32}P]GTP. Aliquots were removed at the indicated times, and reactions stopped by mixing with formamide and EDTA and boiling. The RNAs were resolved on an 8 M urea/12% polyacrylamide gel, and visualized by autoradiography. The build-up of product is similar in both reactions; the precursor band is smeared in the presence of U1A-RBD under these conditions. Gel mobility-shift experiments showed that both precursor and ribozyme are bound by U1A-RBD under conditions suitable for transcription (not shown). **c**, Experimental electron-density map of J1/2 contoured at one s.d. from the mean density level, superimposed on the refined atomic model. The map[44] was calculated using all reflections from the crystal II data set and four-wavelength MAD phases (Table 1) extended to 2.7 Å with SOLOMON[43]. The model shown comprises the uppermost base pairs of helices P1 and P2, and nucleotides A8 and U9 which form the J1/2 'bridge'. The backbone torsion angles are of A-form up to U9, and then the chain makes a sharp bend around the phosphate of G10 which, with Cxi, forms the top base pair of P2. The slight misalignment of the model and of the electron-density map reflects differences in unit-cell dimensions of the crystals used for MAD phasing and for refinement (Table 1 and Methods).

are similar, as would be expected if no nucleotides located 5′ of the cleavage site participated in forming the ribozyme architecture.

Our best crystals were of the complex of a product RNA (Fig. 2a) and the selenomethionyl version of a 98-residue double-mutant U1A-RBD construct[17]. The structure of the complex was determined by multiwavelength anomalous diffraction (MAD) at 2.7 Å resolution, and subsequently refined at 2.3 Å to a free *R* factor of 28.4% (Fig. 2c and Table 1; see Methods). Experimental details, exhaustive refinement, and analysis of crystal contacts will be described elsewhere (A.R.F. and J.A.D., manuscript in preparation).

Overview of the structure

The HDV ribozyme folds into a compact structure comprising five helical segments connected as a nested, double pseudoknot. In addition to the four biochemically identified base-paired regions, P1 to P4, the crystal structure reveals the presence of the two-base-pair helix P1.1 (Figs 2a and 3; see Figs 1 and 2a for numbering and segment name conventions). The five helical segments form two parallel stacks: P1, P1.1 and P4 stack nearly co-axially, whereas P2 and P3 form a second co-axial stack (Fig. 3). The two helical stacks are joined side-by-side by five strand-crossovers, located at the joining regions J1/2 and J4/2, the connection between P3 and P1, and at either end of P1.1. It is the unexpected formation of P1.1 by nucleotides of J1/4 and loop L3, believed until now to be single-stranded (Fig. 1), that introduces a second pseudoknot into the structure. The interactions that produce P1.1 are the vital interactions that constrain the ribozyme into its unique three-dimensional fold by bracing together the P1 substrate helix, the P2–P3 stack, the P4 helix, and the functionally important (see below) J4/2 region. No tightly bound metal ions were found in the HDV ribozyme; the structure appears to be stabilized entirely by base-pairing, stacking, and non-canonical base–backbone and backbone–backbone interactions. The compact, convoluted fold buries the 5′-hydroxyl leaving group of the self-cleavage reaction deep within an active-site cleft (Figs 3 and 4a) that is lined by nucleotides that are known to be important for catalysis.

The helices P1.1 and P3, which constitute the second pseudoknot, stack with the helices (P1 and P2) that form the pseudoknot predicted by Perrotta and Been[7]. This stacking nests the P1.1–P3 pseudoknot within the P1–P2 pseudoknot. This order of stacking is different from that of classical pseudoknots, in which two consecutive stems stack co-axially and in which the two strand-crossovers interact with the major and minor grooves[18]. The three-dimensional fold of the nested double pseudoknot probably occurs in other RNAs. For example, such a structure was proposed for the translational repressor of the α-operon of *Escherichia coli*[19]. The thermal unfolding of the two inner helices of the α-repressor (these inner helices correspond to P2 and P3 in the HDV ribozyme) is coupled[20], indicating that they might be stacked like P2 and P3 are in the HDV ribozyme. Another double-pseudoknot fold has been proposed for a very active RNA ligase identified by *in vitro* selection[21].

A computer graphics model of the HDV ribozyme[8] did not predict the presence of P1.1 or, therefore, the unique, doubly pseudoknotted fold of the ribozyme. Because that model lacked the structural constraints imposed by P1.1, P4 was rotated roughly 180° away along its helical axis from the correct confirmation. As a result, the active site of this model bears no resemblance to the active site seen here in the crystal structure. Two NMR structure determinations of RNA sequences corresponding to isolated P2–P3–L3 segments have been reported[22,23]. As the base-pairing interactions of L3 with J1/4 are absent in these constructs, the two NMR structures exhibit different base-pairing schemes from those seen in the crystal structure.

If an RNA fold is defined as a three-dimensional arrangement of at least two separate helical stacks, the HDV ribozyme represents the fourth RNA fold whose structure is known; the other three are transfer RNA, the hammerhead ribozyme, and the P4–P6 domain of the *Tetrahymena* group I intron. In tRNA, the two helical stacks are perpendicular, whereas in the hammerhead ribozyme they join in one region to form a Y-shaped molecule. The HDV ribozyme superficially resembles the P4–P6 domain in that two helical stacks are arranged side by side (Fig. 3c). However, the 160-nucleotide P4–P6 domain is not pseudoknotted, and its two helical stacks are brought together by two sets of loop-to-helix-groove contacts involving non-canonical interactions (the A-rich-bulge–P4 and tetraloop–receptor contacts). The smaller HDV ribozyme comprises

Table 1 Crystallographic statistics

Diffraction data

Data set	Wavelength (Å)	Resolution (Å)	Number of reflections* observed/unique	Data coverage (%) overall/last shell	$\langle I\rangle/\langle\sigma(I)\rangle$ overall/last shell	R_{sym}† (%) overall/last shell
Crystal I	0.9761 (λ_1)	41.3–2.9	55,889/17,817	95.7/97.6	21.7/5.2	6.5/31.8
	0.9794 (λ_2)	42.5–2.9	85,008/18,459	99.1/99.9	26.2/6.5	7.8/39.3
	0.9792 (λ_3)	35.3–2.9	56,668/17,962	96.4/98.0	21.0/3.6	6.8/41.1
Crystal II	1.1390 (λ_4)	41.5–2.7	68,210/21,826	93.5/93.9	17.7/4.4	9.0/34.7
Crystal III	0.9209	28.6–2.3	72,648/18,914	95.5/96.5	25.9/4.1	6.6/41.4

MAD analysis (crystals I and II)

Phasing power‡ (acentric reflections)

Resolution range (Å)	λ_1 ano	λ_2 iso/ano	λ_3 iso/ano	λ_4 iso/ano
42.5–2.9	2.83	2.39/2.64	2.22/3.10	0.76/0.58
3.2–2.9	1.41	1.65/1.39	1.37/1.44	1.04/0.21

R_{kraut}§ (acentric reflections, %)

Resolution range (Å)	λ_1 iso/ano	λ_2 iso/ano	λ_3 iso/ano	λ_4 iso/ano
42.5–2.9	1.7/4.3	2.5/3.8	2.9/4.6	9.1/4.0
3.2–2.9	6.1/13.1	6.2/12.0	8.7/16.0	5.9/11.7

Mean figure of merit‖ (acentric reflections)

Resolution range (Å)		
42.5–2.9	0.635	
3.1–2.9	0.425	

* Anomalous reflections counted separately for crystals I and II.
† $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the observed intensity and $\langle I\rangle$ is the statistically weighted average intensity of multiple measurements of symmetry-related reflections.
‡ Phasing power $= \langle|F_H|\rangle/\langle||F_{PH}| - |F_P + F_H||\rangle$, where $\langle||F_{PH}| - |F_P + F_H||\rangle$ is the residual lack-of-closure error; 'iso' and 'ano' refer to contributions from isomorphous and anomalous differences.
§ $R_{kraut} = \Sigma||F_{PH}| - |F_P + F_H||/\Sigma|F_{PH}|$.
‖ Mean figure of merit $= \langle\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)\rangle$, where α is the phase and $P(\alpha)$ is the phase-probability distribution.

mainly standard Watson–Crick base pairs (Fig. 2a), and achieves its compact structure by adopting its intricate double-pseudoknot fold.

The convoluted fold of the HDV ribozyme appears to contribute significantly to its stability, as disruption of either of its two pseudoknots results in a marked loss of activity. Cleavage of the ribozyme at J1/2, which deletes the first pseudoknot, produces molecules that are active in *trans*. However, those molecules have cleavage rates that are about two orders of magnitude slower than the rates of their parent *cis*-ribozymes, and, unlike the intact ribozymes, are inactivated by denaturants[2]. Mutagenesis has been used to demonstrate the importance of the four conserved nucleotides of P1.1 for ribozyme function. Their deletion, or their mutation into nucleotides that would disrupt the P1.1 helix and hence delete the second pseudoknot, reduced ribozyme activity by three to five orders of magnitude[8,24].

Structural components of the active site

Hydroxy-radical footprinting of genomic and antigenomic ribozymes (Fig. 1) showed strong protection of the RNA backbone on the 5′ side of P3 and L3, of the nucleotides at the junction of P3 and P1, of J4/2, and weaker protection of guanosines 38–40 (genomic numbering)[25]. In good agreement with this footprinting pattern, the crystal structure shows that these segments line a buried cleft around the 5′-hydroxyl leaving group, resulting from the self-scission reaction (Figs 3 and 4a). The position of this functional group identifies the active site of the HDV ribozyme. Three sides of this active-site cleft are formed by the substrate-bearing P1 helix, J4/2, and a niche formed by the minor groove of P3 and L3. The cleft is enclosed on top by the arching RNA backbone in the junction of P3 and P1, and supported from below by the P1.1 helix.

Positioning of the substrate-bearing helix. The substrate-bearing P1 helix is positioned in the active site of the HDV ribozyme both by a pair of helical crossovers at its top and by the stacking of a G·U wobble (G1·U37) on P1.1 at its bottom (Figs 2a, 3 and 4b). Functional groups in the grooves or the phosphate backbone of the helix do not make lateral contacts with the rest of the ribozyme. Site-directed mutagenesis showed that, provided that base-pairing and total helix length were maintained, the composition of P1 could be varied freely, except for the G1·U37 wobble pair[2]. The structure shows that because of the co-axial stacking of P1 on P1.1, changing the length of P1 would place the crossovers out of register and destabilize the ribozyme. The mutagenesis studies indicated that both stacking energy and placement of the 5′-hydroxyl group (which bears the scissile phosphate in the precursor) resulting from the wobble pair are probably essential for activity, as A·C or G·C, but not any other standard or non-canonical pairs, could partially replace the G·U wobble[26].

The structure shows that, unlike in the group I intron, in which the cleavage site is recognized in part through the N2 amino group of guanosine exposed as a result of wobble base pairing[27], the HDV ribozyme exploits the backbone distortion inherent in the G·U wobble. The 5′-hydroxyl group of G1 is pushed deeper towards the centre of the ribozyme because of the twist of G1 into the minor groove and of U37 into the major groove (Figs 3c and 4b). In addition, the helix twist, together with a displacement of the helix axes of P1 and P1.1, results in G1 stacking onto G38 at the top of the opposite strand of P1.1 (Figs 3a and 4b).

A strand-crossover overhangs the core. The P3–P1 strand-crossover arches over the active site (Figs 3a and 4c). The 3′ side of P3 exits the P2–P3 stack at C30. Without any spacer nucleotides, the backbone makes a sharp change in direction about the phosphate of G31, which then participates in the uppermost base pair of P1. The P2–P3 stack is continued by the incursion of the strand coming from J4/2, which forms the 3′ side of P3. The strong hydroxy-radical protection of C30 and G31 (ref. 25) (Fig. 1) might result from the riboses of these two nucleotides facing an enclosed cavity formed by P1, P2 and the J1/2 bridge (Figs 2c and 3).

A triple-base interaction stabilizes this P3–P1 strand-crossover (Fig. 4c): A78, the last nucleotide of J4/2, forms four hydrogen

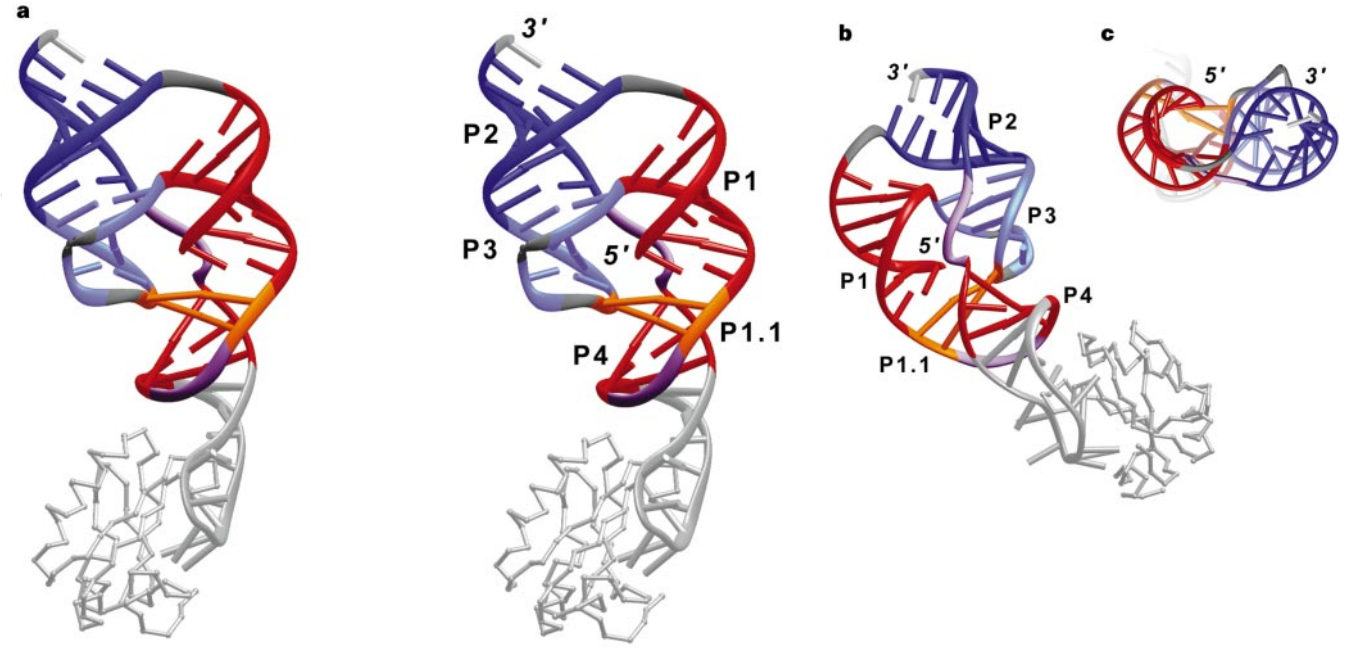

Figure 3 Ribbon and stick representations of the crystal structure of the HDV ribozyme, colour-coded as in Fig. 2a. The path of the RNA backbone is shown as a ribbon, and base-paired nucleotides as sticks. The unpaired nucleotides from J1/2 (grey), J1.1/4 and J4/2 (pink) and L3 (grey) are omitted for clarity. The U1A-RBD protein is represented as an α-carbon ball-and-stick model. Paired regions and RNA termini are labelled. Cxi at the 3′ terminus is in white. **a**, Stereoview of the front of the structure. **b**, Back view of the molecule, tilted to emphasize the snug fit of the J4/2 region between the two helical stacks, and the intricate pseudoknot crossovers. **c**, Top view of the structure, showing that the two helical stacks are nearly parallel to each other.

bonds with the minor groove face of the penultimate pair of P3 (C18·G29). This base triplet has been observed before, both at the P5abc junction of the P4–P6 domain[28], and as part of a base-quadruplet interaction between GAAA tetraloops and either an intramolecular tetraloop receptor[28] or an RNA stem in a crystal contact[29]. This interaction explains, in part, the strict requirement for A78 (refs 8, 30). The exocyclic N6 of A78 is not involved in hydrogen bonding and points towards the active site, where it would be available for coordination, and may participate in catalysis.

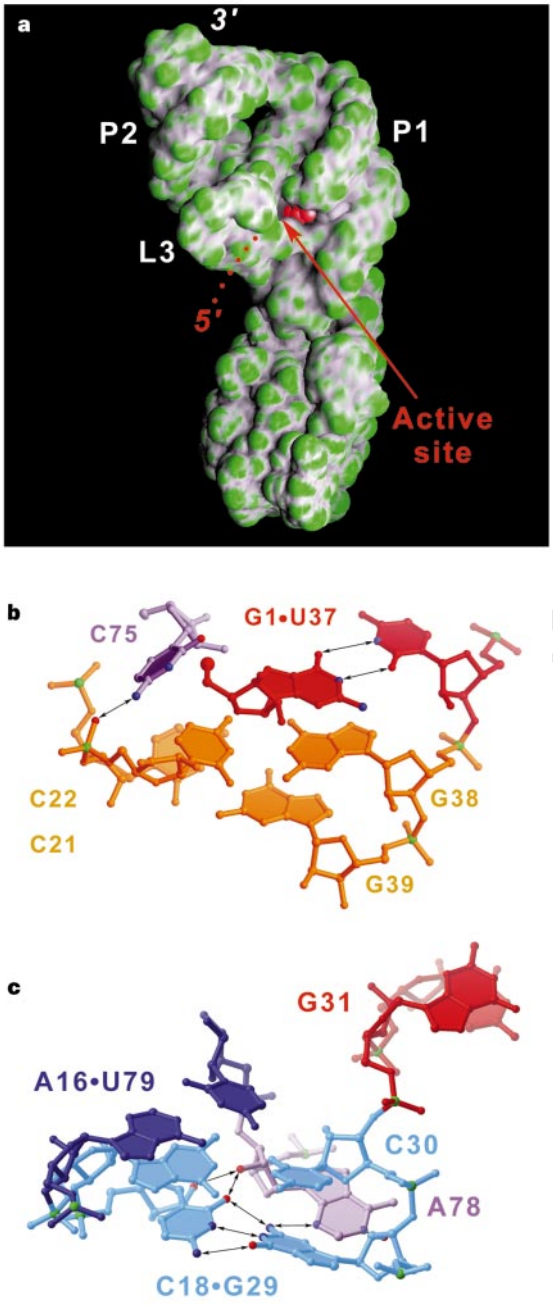

Figure 4 Structural components of the active site. **a**, Solvent-accessible surface of the ribozyme–U1A-RBD complex, orientated as in Fig. 3a, colour-coded by its curvature (green, convex; grey, concave[49]). The atoms of the ribose ring of G1 are shown as red spheres; the 5′-hydroxyl group points away from the reader into the ribozyme. The trajectory proposed to be followed by nucleotides located 5′ of the cleavage site in the precursor is indicated by the row of small red dots. **b**, The substrate-bearing nucleotide G1 forms a wobble pair with U37 which stacks on P1.1. The 5′-hydroxyl group (shown as a larger sphere) is within hydrogen-bonding distance of the Watson–Crick face of C75. Nucleotides are coloured as in Figs 2 and 3. Phosphorus atoms are coloured green. Oxygen and nitrogen atoms that participate in noteworthy hydrogen bonding (denoted by two-headed arrows) are coloured red and blue, respectively. **c**, Triple-strand junction between P3 and P1, orientated as in Fig. 3a. A hydrogen bond between the N1 group of A78 and the O2′ group of G29 is obscured in this view.

P3 and L3 form a niche. P3 is composed of three base pairs as predicted[7] (Figs 3 and 4c). Consistent with mutagenesis studies[2,9,31], P3 makes no sequence-specific contacts with the rest of the ribozyme. The minor groove face of P3 and the unpaired nucleotides U20, C24 and G25 form a niche that flanks the active site (Figs 3, 4a and 5a). The helical stack of P3, but not the base pairing, is continued by U20 and G25 below the C19·G28 pair. U20 and G25 are too far from each other to form any hydrogen bonds between their Watson–Crick faces. Mutagenesis of these two nucleotides to sequences that would force pairing (such as G25A) resulted in ribozymes that were 3,000 times less active[8]. The P2–P3 helical stack ends with C24 stacking underneath G25, where it makes a hydrogen bond between its N4 and the pro-S_p phosphate of C22 (Fig. 5a). Of the remaining nucleotides in L3, C21 and C22 are part of P1.1, U23 lies underneath C24, and C26 and U27 extrude from the helical stack, in good agreement with chemical-modification experiments[32]. The pyrimidine bases of C26, U27 and, to a lesser extent, U23 (grey in Fig. 3) are not well-ordered in our crystals (see Methods). Results of biochemical experiments indicate that these three mobile nucleotides act mainly to connect catalytically important segments of P3 and L3 and that their bases have only modest structural and energetic roles[9,32–34].

P1.1 and a pedestal support the core. Two strand-crossovers corresponding to the second pseudoknot occur at either end of P1.1, and form part of a pedestal that rests upon P4 and supports both the P2–P3–L3 stack and the cleavage-site G·U wobble (Fig. 5b). The non-canonical A43·G74 base pair at the top of P4 (Fig. 5d) occurs at the beginning of the P4 helix, and results in an opening of the minor groove. The pedestal is completed by novel interactions around J1.1/4. The C21·G39 pair at the bottom of P1.1 stacks on the 3′ side of the A43·G74 pair only, because of the large (100°) twist angle between these pairs. The J1.1/4 segment spanning G39–A43 (Figs 3 and 5b) adopts a unique turn conformation that is different from the well-characterized U- or π-turns. G40, the first of the turn nucleotides, stacks under G39, and also forms hydrogen bonds with A42 and G74. The backbone then buckles, pushing the unpaired C41 into a narrowed major groove and unstacking its bottom face. The backbone turns again, now pushing the base of A42 into the minor groove, where it stacks on A43 and buttresses C21 by hydrogen bonding (Fig. 5b). Beyond this, the chain enters P4 where it adopts an A-conformation. The turn breaks the helical stack between P1.1 and P4 on one strand but not on the other, connecting P1 and P4 while also reinforcing the minor groove of P1.1.

A turn and ribose zipper buttress the core. The single-stranded region J4/2 closes the rear side of the active site by connecting P4 with P2, and incorporates a 'trefoil' turn (Fig. 5c, d) that positions important nucleotides[2,8] in the active site, arranging G74, C75 and G76 as the three leaflets of a clover and fully extruding the base of G76 into solvent. C75 projects deep into the core of the ribozyme, and is held in place by a hydrogen bond between its N4 and the pro-R_p oxygen of C22, as well as by stacking interactions with A77 and backbone interactions. The top two nucleotides of J4/2, namely A77 and A78, are held in place by the base-triplet interaction in the P3–P1 junction (Fig. 4c), and are also involved in a two-tiered ribose-zipper[28] interaction with C19 and C18, respectively (Fig. 5d). Consistent with their marked protection from hydroxy-radical attack[25] (Fig. 1), chemical modification[32] and enzymatic digestion[35] the riboses of C75–A78 are buried by the conformation of J4/2 and the dense network of backbone hydrogen bonds.

As expected from its position within the active site next to the 5′-hydroxyl leaving group produced by the self-scission reaction (Figs 4b and 5a), C75 cannot be mutated to any other nucleotide

without completely abolishing ribozyme activity[8]. However, the extruded G76 can be mutated to uridine with only modest loss of activity[8]; furthermore, results of a modification-interference study showed that opening the purine ring of this guanosine resulted in an increase in ribozyme activity[33]. Both A77 and A78 are also sensitive to changes[8].

Active site and catalysis

Together with existing biochemical information, the crystal structure of the HDV ribozyme offers some hints to the chemical mechanism of this RNA catalyst. A solvent-accessible surface representation of the crystal structure (Fig. 4a) shows that the 5′-hydroxyl group, which marks the location of the active site, lies buried at the bottom of a deep cleft, surrounded by the structural elements described above. Mutations in all of these elements are deleterious to the function of the ribozyme, indicating that the product of the self-scission reaction has a structure similar to that of the precursor and the transition state. Consistent with a transesterification mechanism in which the 2′-hydroxyl group of the −1 nucleotide performs a nucleophilic attack on the phosphorus group, substitution of the −1 ribose with a 2′-deoxyribose abolished the reaction[2]. Thus, structural elements present in the product RNA whose structure we have determined must position the −1 ribose for nucleophilic attack, provide a base with which to activate its 2′-hydroxyl group, and perhaps stabilize the transition state by neutralizing the developing negative charge in the leaving group.

Our structure of the HDV ribozyme and the available biochemical data do not reveal how metal ions might participate in catalysis. The rate of cleavage of the ribozyme does not show a marked pH dependence[5,36]. Phosphorothioate substitution at the cleavage site showed that the pro-R_p oxygen cannot be substituted by a sulphur without loss of activity[37]; cleavage was not recovered by the addition of Mn^{2+}. The only other phosphate oxygen sensitive to sulphur substitution was the pro-R_p oxygen of C22. This substitution was also insensitive to Mn^{2+}. This pro-R_p oxygen forms a hydrogen bond with the *trans* face of the N4 amino group of C75 (Figs 4b and 5a, d), so the phosphorothioate substitution probably results in steric clash. The lack of requirement for a specific metal ion indicates that the ion does not participate directly in activating the nucleophile, but instead acts as a charge sink.

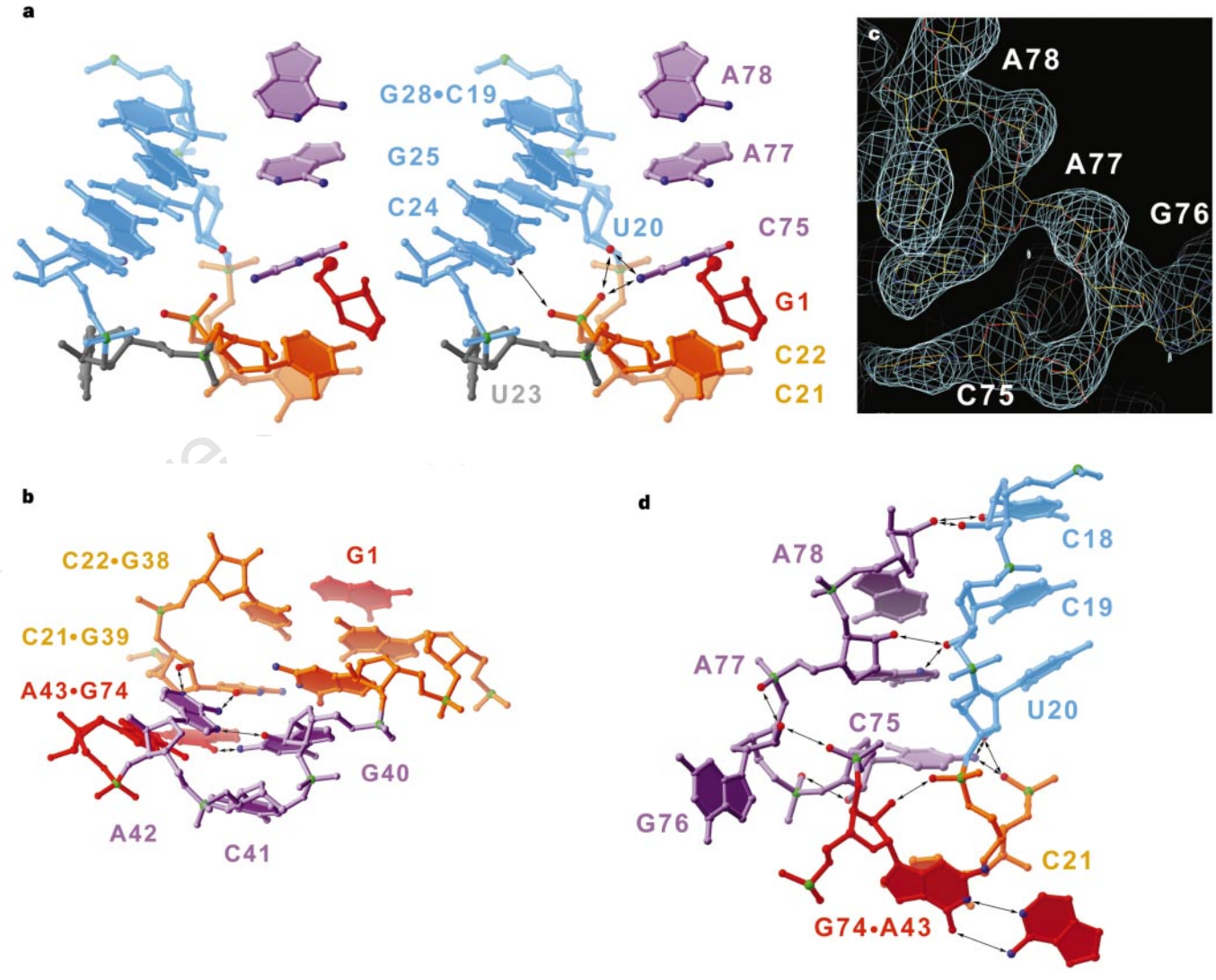

Figure 5 Structural features of the active site. **a**, The P3–L3 niche cradles the active site. The stereoview shows how the phosphate ribose backbone crosses over from the P3–L3 stack to P1.1 and back, passing through the unstacked nucleotide U23. The ribose of G1 and the bases of the stacked nucleotides from J4/2 are shown. **b**, The P1.1 helix and the pedestal. **c**, A σA-weighted $2|F_o| - |F_c|$ map[44,45], with data from 20 to 2.3 Å, contoured at 1 s.d., showing the nucleotides of J4/2 as seen from the major groove of P1. **d**, The trefoil turn (orientated roughly as in Fig. 3b) is the sharp twist in the RNA chain between G74 and A77. This is stabilized by a complex network of interactions between O2′ groups and phosphate oxygens, a two-tiered ribose zipper, and base–backbone hydrogen bonds. Figures 3, 4b, c and 5a, b, d were prepared with RIBBONS[50].

In a *trans* ribozyme, an azidophenacyl group placed at the cleavage site formed crosslinks with residues at the 3′ side of L3 (equivalent to G25 and C26)[25]. Photocrosslinking of a cleavable *trans* ribozyme substituted with deoxythiouridine at position −2 resulted in crosslinks from this deoxythiouridine to the nucleotide corresponding to C75. Deoxythiouridines at positions −2 and −1 in non-cleavable ribozymes formed crosslinks with nucleotides corresponding to C21 and G25 in the L3 niche[38]. We propose that the RNA chain leading to the scissile bond enters the core of the HDV ribozyme through the niche formed by these mutation-sensitive nucleotides (Figs 4a and 5a). If so, the phosphodiester backbone must make a sharp bend at or near the scissile bond, so that it can adopt the helical conformation it has in P1. Qualitatively, it appears that the parts of bases in the niche that form minor grooves would be well placed to position the ribose for attack. The extrusion of U23 from the base stack creates a pocket at the bottom of L3 in which the base of nucleotide −1 could be accommodated without making specific interactions (Fig. 5a).

From results of crosslinking and mutagenesis studies, and because of its proximity to the 5′-hydroxyl group in our structure, it seems that C75 is probably important in catalysis. As metal ions do not appear to participate directly in catalysis, we propose that C75 acts as the general base that activates the 2′-hydroxyl group of nucleotide −1 for nucleophilic attack. The *trans* face of the N4 amino group of C75 participates in an interesting network of hydrogen-bond interactions (Fig. 5a, d). It is positioned so that it could donate a hydrogen bond to either the O2′ group of U20, or the pro-R_p phosphate oxygen of C22. The O2′ of U20 probably donates a hydrogen bond to the same phosphate oxygen. There are, therefore, three polar ligands arranged at the vertices of a triangle with sides of length ~2.8 Å and only two hydrogens (Fig. 5a, d). This, and the negative electrostatic potential in the region of C75 (resulting from the trefoil turn), could perturb the pK_a values of C75, making it basic.

In summary, the crystal structure of a genomic HDV ribozyme reveals a unique, nested double-pseudoknot fold that allows this catalytic RNA to form a deep, solvent-inaccessible active-site cleft, much like that of protein enzymes. The only other ribozyme for which a structure is available, the hammerhead ribozyme, places its scissile bond in a solvent-exposed portion of a typical A-form helix[10,11]. Our structure of the HDV ribozyme explains the available biochemical data, and will act as the starting point for detailed studies of the chemical mode of action of this very active catalytic RNA. □

Methods

RNA and protein preparation. Plasmid pDU9 is a derivative of pUC19 that encodes the HDV ribozyme we studied, preceded by a 41-nucleotide leader sequence and a T7 RNA polymerase promoter. The insert was prepared by the polymerase chain reaction from overlapping synthetic DNA oligonucleotides, and is followed by a *Bsa*I restriction site, such that run-off transcription of the linear plasmid followed by ribozyme self-cleavage produces the 72-nucleotide HDV ribozyme RNA (Fig. 2a). The RNA has homogeneous 5′, and heterogeneous 3′ termini, because of addition of a few random nucleotides by the polymerase. Attempts at producing a homogeneous 3′ end failed. Transcription and purification were done as described[16].

Selenomethionyl U1A-RBD was produced in the methionine auxotroph *Escherichia coli* strain B834 (Novagen), transformed with the (A1-98 Y31H, Q36R) T7 expression plasmid[17] (provided by K. Nagai) and grown in minimal medium supplemented with selenomethionine. Cell lysate was clarified by centrifugation, and the supernatant fractionated with polyethyleneimine and ammonium sulphate. The protein was further purified by cation-exchange (SP-Sepharose FF, Pharmacia), size-exclusion (Superdex-75 PG, Pharmacia), and hydroxyapatite (CHT-I, BioRad) chromatography.

Crystallization and data collection. Annealed RNA was mixed with one molar equivalent of selenomethionyl U1A-RBD protein to yield a complex concentration of 0.5 mM (in 1.25 mM $MgCl_2$). For crystallization by the sitting-drop vapour-diffusion technique, 4 μl complex solution was mixed with 2 μl of a reservoir solution consisting of 12.5–14% (w/v) polethyleneglycol monomethyl ether of average molecular weight 2,000 (PEG-MME 2000), 100 mM Tris-HCl, pH 7.0, 200–250 mM Li_2SO_4 and 4 mM spermine-HCl. In addition, reservoir solutions contained 0.2 mM cobalt(III) hexammine chloride (CoHex, crystals I and III); 0.4 mM osmium(III) hexammine chloride triflate (OsHex, crystal II); and 5 mM $MgCl_2$ (crystal III). Trigonal trapezohedra grew at 25 °C over several weeks, and attained maximum dimensions of $0.8 \times 0.6 \times 0.6\,mm^3$ (space group $R32$, $\alpha = 90°$, $\gamma = 120°$). Cell dimensions were $a = 108.7$ Å, 109.15 Å and 109.35 Å, and $c = 190.38$ Å, 190.93 Å and 190.68 Å, for crystals I, II and III, respectively. For cryogenic data collection, crystals were transferred to solutions containing 25% (*w/v*) PEG-MME 2000, 100 mM TRIS-HCl pH 7.0, 250 mM Li_2SO_4, 6 mM (crystals I and II) or 25 mM (crystal III) spermine-HCl, 0.4 mM or 2.0 mM CoHex (crystals I and III, respectively), or 0.5 mM OsHex (crystal II), 2.5 mM, 2.0 mM or 10 mM (crystals I, II and III, respectively) $MgCl_2$, 5% (w/v) PEG 6000, and 3–4% (v/v) (2*R*,3*R*)-(−)-2,3-butanediol. Crystals were mounted and flash-cooled as described[16].

MAD data from crystals I and II were collected at beamline X4A of the National Synchrotron Light Source using an Raxis-IV imaging plate area detector (Rigaku) and the inverse-beam mode. Data from crystal III was obtained at beamline F-1 of the Cornell High Energy Synchrotron Source on a Quantum-4 CCD detector (Area Detectors Systems) (Table 1). Oscillation photographs were processed using the programs DENZO and SCALEPACK[39], and the CCP4 suite[40].

Phase determination and structure refinement. Four selenium sites were located in crystal I using the package SOLVE[41] (www.solve.lanl.gov). Density modification[40] of three-wavelength SOLVE phases produced a 2.9 Å electron-density map with clear protein and RNA features. Fourier maps calculated using these phases and anomalous differences from data from crystal II did not contain significant peaks except for the four selenium sites, indicating that the OsHex was not bound specifically. Selenium parameters were further refined and phase distributions and centroids were calculated with SHARP[42] using data collected at three wavelengths from crystal I (high-energy remote, rising inflection point, and peak, respectively), and one from crystal II (low-energy remote, Table 1). Three of the four seleniums are well-ordered; the site corresponding to Met 97 is not. Density modification and phase extension to 2.7 Å (SOLOMON[43]) produced an electron-density map into which most of the protein and nucleic acid residues could be built unambiguously, using program O (ref. 44). Correct registration of both macromolecule chains was confirmed by superimposing the methionine sites of the U1A-RBD–RNA 21-mer complex coordinates[17] on the selenium sites. The U1A–RNA 21-mer structure[17] was not used as a source of phase information at any stage.

Rounds of manual rebuilding, interspersed with positional, torsion-angle-simulated annealing, and tightly restrained individual *B*-factor refinement with CNS[45], produced a model with an *R*-free value of 38.9% to 2.7 Å. This model was further refined against all crystal III amplitude data between 20 and 2.3 Å (17,042 and 1,861 reflections in the work and free sets, respectively), and against experimental phase-probability distributions from SHARP, using the phased maximum-likelihood target, purely repulsive van der Waals parameters, and a bulk solvent mask[45]. Use of the phased maximum-likelihood target[45] resulted in a *R*-free factor typical of a well refined model at this resolution[46], with negligible overfitting ($R_{free} = 28.4\%$, $R_{work} = 28.0\%$). Despite the high average *B*-factors (44.2 and 91.7 $Å^2$, respectively, for protein and RNA atoms), electron density for the RNA is of good quality (Figs 2c and 5c). In a σA-weighted $2|F_o| - |F_c|$ map[45] contoured at 0.9 s.d., there is a total of seven minor breaks in the RNA backbone, in P1, P2 and P4, and two larger breaks in the backbone of L3, in which electron density for the extruded bases of residues U23, C26 and U27 is also weak. Electron density for the protein moiety is good throughout, with continuous backbone density from residues 7 to 96 in the same map contoured at 2.5 s.d. All nucleotides are in *anti* conformation. Ribose puckers are all C3′-endo except for nucleotides 22, 41, 74, 75, 76, 79, v, vii, and x, which were restrained to C2′-endo. The present model consists of all RNA residues, U1A–RBD residues 4–98 (H31, R36 and S71 are all in two conformations), two sulphate and three magnesium ions, and nineteen water molecules (2,366 non-hydrogen atoms), and has root-mean-square (r.m.s.) deviations from the ideal of 0.009 Å and 1.43° for bond lengths and angles,

respectively. The r.m.s. deviation of the *B*-factor of covalently bonded atoms is 0.47 Å^2. PROCHECK[47] analysis shows no outliers in the Ramachandran plot, and an overall *G*-factor better (2.7 bandwidths from the mean) than expected at this resolution.

Received 17 July; accepted 7 September 1998.

1. Lai, M. M. The molecular biology of hepatitis delta virus. *Annu. Rev. Biochem.* **64,** 259–286 (1995).
2. Been, M. D. & Wickham, G. S. Self-cleaving ribozymes of hepatitis delta virus RNA. *Eur. J. Biochem.* **247,** 741–753 (1997).
3. Kuo, M. Y.-P., Sharmeen, L., Dinter-Gottlieb, G. & Taylor, J. Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus. *J. Virol.* **62,** 4439–444 (1988).
4. Sharmeen, L., Kuo, M. Y.-P., Dinter-Gottlieb, G. & Taylor, J. Antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. *J. Virol.* **62,** 2674–2679 (1988).
5. Wu, H.-N. *et al.* Human hepatitis delta virus RNA subfragments contain an autocleavage activity. *Proc. Natl Acad. Sci. USA* **86,** 1831–1835 (1989).
6. Perrotta, A. T. & Been, M. D. The self-cleaving domain from the genomic RNA of hepatitis delta virus: sequence requirements and the effects of denaturant. *Nucleic Acids Res.* **18,** 6821–6827 (1990).
7. Perrotta, A. T. & Been, M. D. A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature* **350,** 434–436 (1991).
8. Tanner, N. K. *et al.* A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses. *Curr. Biol.* **4,** 488–498 (1994).
9. Thill, G., Vasseur, M. & Tanner, N. K. Structural and sequence elements required for the self-cleaving activity of the hepatitis delta virus ribozyme. *Biochemistry* **32,** 4254–4262 (1993).
10. Pley, H. W., Flaherty, K. M. & McKay, D. B. Three-dimensional structure of a hammerhead ribozyme. *Nature* **372,** 68–74 (1994).
11. Scott, W. G., Finch, J. T. & Klug, A. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **81,** 991–1002 (1995).
12. Richards, F. M. & Wyckoff, H. W. in *The Enzymes* (ed. Boyer, P. D.) 647–806 (Academic, New York, 1971).
13. Rosenstein, S. P. & Been, M. D. Self-cleavage of hepatitis delta virus genomic strand RNA is enhanced under partially denaturing conditions. *Biochemistry* **29,** 8011–8016 (1990).
14. Duhamel, J. *et al.* Secondary structure content of the HDV ribozyme in 95% formamide. *Nucleic Acids Res.* **24,** 3911–3917 (1996).
15. Suh, Y.-A., Kumar, P. K. R., Taira, K. & Nishikawa, S. Self-cleavage activity of the genomic HDV ribozyme in the presence of various divalent metal ions. *Nucleic Acids Res.* **21,** 3277–3280 (1993).
16. Ferré-D'Amaré, A. R., Zhou, K. & Doudna, J. A. A general module for RNA crystallization. *J. Mol. Biol.* **279,** 621–631 (1998).
17. Oubridge, C., Ito, N. Evans, P. R., Teo, C.-H. & Nagai, K. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372,** 432–438 (1994).
18. Kolk, M. H. *et al.* NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. *Science* **280,** 434–438 (1998).
19. Tang, C. K. & Draper, D. E. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. *Cell* **57,** 531–536 (1989).
20. Gluick, T. C. & Draper, D. E. Thermodynamics of folding a pseudoknotted mRNA fragment. *J. Mol. Biol.* **241,** 246–262 (1994).
21. Ekland, E. H., Szostak, J. W. & Bartel, D. P. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* **269,** 364–370 (1995).
22. Kolk, M. H., Heus, H. A. & Hilbers, C. W. The structure of the isolated, central hairpin of the HDV antigenomic ribozyme: novel structural features and similarity of the loop in the ribozyme and free in solution. *EMBO J.* **16,** 3685–3692 (1997).
23. Lynch, S. R. & Tinoco, I. Jr The structure of the L3 loop from the hepatitis delta virus ribozyme: a *syn* cytidine. *Nucleic Acids Res.* **26,** 980–987 (1998).
24. Perrotta, A. T. & Been, M. D. Core sequences and a cleavage site wobble pair required for HDV antigenomic ribozyme self-cleavage. *Nucleic Acids Res.* **24,** 1314–1321 (1996).
25. Rosenstein, S. P. & Been, M. D. Hepatitis delta virus ribozymes fold to generate a solvent-inaccessible core with essential nucleotides near the cleavage site phosphate. *Biochemistry* **35,** 11403–11413 (1996).
26. Nishikawa, F., Fauzi, H. & Nishikawa, S. Detailed analysis of base preferences at the cleavage site of a trans-acting HDV ribozyme: a mutation that changes cleavage site specificity. *Nucleic Acids Res.* **25,** 1605–1610 (1997).
27. Strobel, S. A. & Cech, T. R. Minor groove recognition of the conserved G-U pair at the *Tetrahymena* ribozyme reaction site. *Science* **267,** 675–679 (1995).
28. Cate, J. H. *et al.* Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273,** 1678–1685 (1996).
29. Pley, H. W., Flaherty, K. M. & McKay, D. B. Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature* **372,** 111–113 (1994).
30. Kumar, P. K. R. *et al.* Random mutations to evaluate the role of bases at two important single-stranded regions of the genomic HDV ribozyme. *Nucleic Acids Res.* **20,** 3919–3924 (1992).
31. Kumar, P. K. R., Suh, Y.-A., Taira, K. & Nishikawa, S. Point and compensation mutations to evaluate essential stem structures of genomic HDV ribozyme. *FASEB J.* **7,** 124–129 (1993).
32. Kumar, P. K. R., Taira, K. & Nishikawa, S. Chemical probing studies of variants of the genomic hepatitis delta virus ribozyme by primer extension analysis. *Biochemistry* **33,** 583–592 (1994).
33. Belinsky, M. G., Britton, E. & Dinter-Gottlieb, G. Modification interference analysis of a self-cleaving RNA from hepatitis delta virus. *FASEB J.* **7,** 130–136 (1993).
34. Wadkins, T. S. & Been, M. D. Core-associated non-duplex sequences distinguishing the genomic and antigenomic self-cleaving RNAs of hepatitis delta virus. *Nucleic Acids Res.* **25,** 4085–4092 (1997).
35. Rosenstein, S. P. & Been, M. D. Evidence that genomic and antigenomic RNA self-cleaving elements from hepatitis delta virus have similar secondary structures. *Nucleic Acids Res.* **19,** 5409–5416 (1991).
36. Huang, Z.-S., Ping, Y.-H. & Wu, H.-N. An AU at the first base pair of helix 3 elevates the catalytic activity of hepatitis delta virus ribozymes. *FEBS lett.* **413,** 299–303 (1997).
37. Jeoung, Y.-H., Kumar, P. K. R., Suh, Y.-A., Taira, K. & Nishikawa, S. Identification of phosphate oxygens that are important for self-cleavage activity of the HDV ribozyme by phosphorothioate substitution interference analysis. *Nucleic Acids Res.* **22,** 3722–3727 (1994).
38. Bravo, C., Lescure, F., Laugâa, P., Fourrey, J.-L. & Favre, A. Folding of the HDV antigenomic ribozyme pseudoknot structure deduced from long-range photocrosslinks. *Nucleic Acids Res.* **24,** 1351–1359 (1996).
39. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
40. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–776 (1994).
41. Terwilliger, T. C. & Kim, S.-H. Generalized method of determining heavy-atom positions using the difference Patterson function. *Acta Crystallogr. A* **43,** 1–5 (1987).
42. de la Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276,** 472–494 (1997).
43. Abrahams, J. P. & Leslie, A. G. W. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* **54,** 905–921 (1998).
44. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
45. Brünger, A. T. *et al.* Crystallography and NMR system: a new software system for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
46. Brünger, A. T. Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355,** 472–475 (1992).
47. Laskowski, R. J., Macarthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–290 (1993).
48. Pleij, C. W. A., Rietveld, K. & Bosch, L. A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res.* **13,** 1717–1731 (1985).
49. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insight from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11,** 281–296 (1991).
50. Carson, M. Ribbons 2.0. *J. Appl. Crystallogr.* **24,** 958–961 (1991).

Acknowledgements. We thank K. Nagai for U1A-RBD expression plasmids; C. Ogata and D. Cook for help at beamline X4A of the National Synchrotron Light Source (Brookhaven National Laboratory); the staff of the Cornell High Energy Synchrotron Source for support at beamlines A-1 and F-1; R. Batey, D. Battle, J. Kieft, A. Luptak and R. Rambo for help with synchrotron data collection and useful comments; P. Adams, M. Been, S. Bellon, C. Correll, R. Gaudet, D. Herschlag, J. Ippolito, P. Moore, L. Silvian, T. Steitz, S. Strobel, O. Uhlenbeck and D. Wilson for helpful discussions; and the staff of the Yale Center for Structural Biology for crystallographic and computational support. A.R.F. was a Fellow of the Jane Coffin Childs Memorial Fund for Medical Research. This work was supported in part by grants from the Jane Coffin Childs Memorial Fund for Medical Research, the Searle Scholars program, the NIH, and a Beckman Young Investigator award. J.A.D. is an assistant investigator of the Howard Hughes Medical Institute, a Searle scholar, and a Fellow of the David and Lucile Packard Foundation.

Correspondence and requests for materials should be addressed to J.A.D. (e-mail: doudna@csb.yale.edu).
Atomic coordinates have been deposited at the Protein Data Bank under accession number 1drz, and are also available directly from the authors (www.csb.yale.edu/people/doudna).