PCCP

PAPER

View Article Online
View Journal | View Issue

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

Cite this: *Phys. Chem. Chem. Phys.*, 2015, **17**, 5887

Wobble pairs of the HDV ribozyme play specific roles in stabilization of active site dynamics

Kamali N. Sripathi,[ab] Pavel Banáš,[c] Kamila Réblová,[e] Jiří Šponer,[de] Michal Otyepka[c] and Nils G. Walter*[b]

Received 3rd November 2014,
Accepted 20th January 2015

DOI: 10.1039/c4cp05083e

www.rsc.org/pccp

The hepatitis delta virus (HDV) is the only known human pathogen whose genome contains a catalytic RNA motif (ribozyme). The overall architecture of the HDV ribozyme is that of a double-nested pseudoknot, with two GU pairs flanking the active site. Although extensive studies have shown that mutation of either wobble results in decreased catalytic activity, little work has focused on linking these mutations to specific structural effects on catalytic fitness. Here we use molecular dynamics simulations based on an activated structure to probe the active site dynamics as a result of wobble pair mutations. In both wild-type and mutant ribozymes, the in-line fitness of the active site (as a measure of catalytic proficiency) strongly depends on the presence of a C75(N3H3+)N1(O5′) hydrogen bond, which positions C75 as the general acid for the reaction. Our mutational analyses show that each GU wobble supports catalytically fit conformations in distinct ways; the reverse G25U20 wobble promotes high in-line fitness, high occupancy of the C75(N3H3+)G1(O5′) general-acid hydrogen bond and stabilization of the G1U37 wobble, while the G1U37 wobble acts more locally by stabilizing high in-line fitness and the C75(N3H3+)G1(O5′) hydrogen bond. We also find that stable type I A-minor and P1.1 hydrogen bonding above and below the active site, respectively, prevent local structural disorder from spreading and disrupting global conformation. Taken together, our results define specific, often redundant architectural roles for several structural motifs of the HDV ribozyme active site, expanding the known roles of these motifs within all HDV-like ribozymes and other structured RNAs.

1 Introduction

The hepatitis delta virus (HDV) ribozyme is a member of the family of small catalytic RNAs. In addition to the HDV ribozyme, this family comprises several well-characterized members such as the hairpin and hammerhead ribozymes.[1–4] Despite adopting distinct architectures, all small ribozymes catalyse the same reaction, site-specific cleavage of their phosphodiester backbone. This reaction proceeds through S_N2-type in-line attack, in which a nucleophilic 2′ OH directly upstream of the scissile phosphate attacks the electrophilic phosphorus. Ribozyme backbone cleavage results in 3′ and 5′ cleavage products with 5′ OH and 2′–3′ cyclic phosphate groups, respectively.[1–3] The HDV ribozyme is of special importance because it is the only known member of this class with versions found in both a human pathogen and the human genome.[1–6] Its reaction is believed to proceed through general acid–base catalysis, with cytosine 75 (C75) long identified as a catalytic nucleotide.[7,8]

The HDV ribozyme in its minimal, 85 nt form adopts a complex tertiary structure, consisting of five helical regions numbered P1–P4 and P1.1.[9] These helices are arranged in a double-nested pseudoknot, resulting in the formation of two coaxial stacks (Fig. 1a and b)[9,10] and a deep, narrow cleft in which the active site residues interact (Fig. 1c). The active site is bounded above by a pair of type I and type II A-minor interactions (involving the two green A residues in Fig. 1a).[11,12] Its lower bound is two-fold, consisting first of the 2-base-pair helix P1.1, and below it a unique set of quadruple interactions involving a protonated C41 residue.[10,13,14] These latter two features have been shown to be critical hallmarks of the native state, as both the presence of the P1.1 helix and correct formation of the C41 quadruple are crucial for ribozyme activity[10,14,15] and architecture.[11]

Several crystal structures have been solved to first ascertain the structure of the active site, and later to elucidate the role of the catalytic C75.[10,13,16] The initial structure, a *cis*-acting reaction

[a] *Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1065, USA*
[b] *Department of Chemistry, Single Molecule Analysis Group, University of Michigan, Ann Arbor, Michigan 48109-1055, USA. E-mail: nwalter@umich.edu*
[c] *Regional Centre of Advance Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University, tr. 17 listopadu 12, 771 46 Olomouc, Czech Republic*
[d] *Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic*
[e] *CEITEC – Central European Institute of Technology, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic*

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

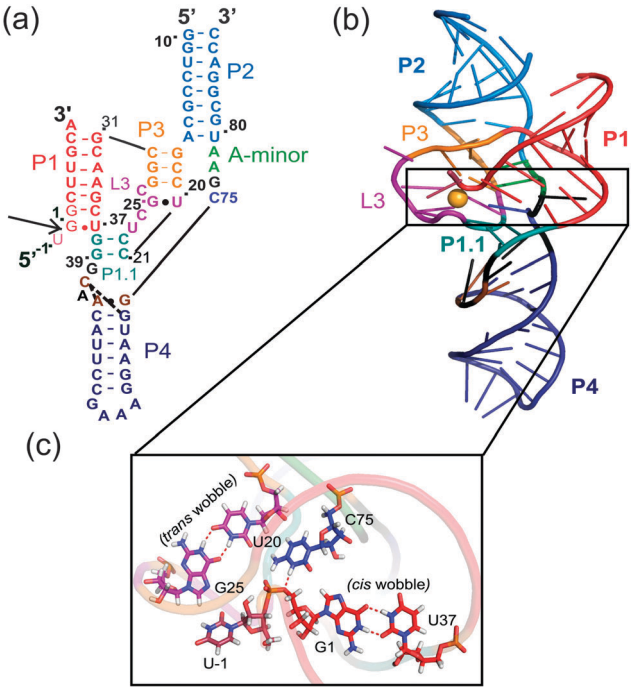

Fig. 1 Secondary and tertiary structure depiction of the HDV ribozyme. Please note that the numbering we use in this manuscript matches previous, well-established conventions for numbering HDV ribozyme residues,[10] and thus does not follow the residue numbering in PDB file 3NKB. Noteworthy secondary (a) and tertiary (b) interactions between motifs are coloured as indicated. The active site is outlined in a black box in the tertiary structure depiction (right), and the magnesium ion from structure 3NKB is shown in its active site binding pocket in gold. (c) Active site residues of the HDV ribozyme, with *cis* and *trans* GU wobbles indicated. Please note that panels b and c are not in the same orientation; rather, each has been optimized to provide the best views of the residues involved. Specifically, the field of view in panel c is angled ~30° upward compared to that in panel b.

product, revealed the mini-helix P1.1 and the wealth of functional groups within close proximity in the HDV ribozyme active site.[10] Later, several *cis*-acting reaction precursor structures showed an *anti* orientation of the G25 base, in contrast to its *syn* position in the product structure.[13] In the most recent *trans*-acting structure, the G25 and its partner U20 occurred in a rare "reverse wobble",[16] in which G25 is in a syn conformation and the pair's glycosidic bonds are in a *trans* orientation.[17] (Please note that the terms "*cis*" and "*trans*" are used in two different contexts in this work, as is common. The terms "*cis*-acting" and "*trans*-acting" refer to ribozymes for which the cleavage site and active site are on the same ("*cis*") or different ("*trans*") strands. By contrast, the terms "*cis* orientation" and "*trans* orientation" refer to the orientation of the glycosidic bonds (between sugar and base) to the base pair axis: glycosidic bonds pointing in the same direction are in a *cis* orientation, whereas bonds pointing in opposite directions are in a *trans* orientation. Ref. 18 offers an effective visual of the latter usage.) Additionally, the position of C75 was resolved between U-1(O2′) and G1(O5′), indicating that it could act as either a general base or acid, respectively. Global comparison showed a collapse of the P1 and P3 helices towards the active site and an overall lengthening along the longitudinal axis from the precursor to the product structure.[10,13,18] All of these observations indicated significant conformational changes upon cleavage at both the whole-molecule and active-site levels. By contrast, a more recent *trans*-acting crystal structure exhibits more product-like global and local conformations.[16] To prevent cleavage, the authors used an external substrate analogue containing three DNA residues at the −1, 1 and 2 positions (see Fig. 1 for positions). The authors used the product structure as the most appropriate for molecular replacement.[16] G25 is again in the *syn* conformation, and C75 has the potential to hydrogen bond with G1(O5′). This structure was solved at pH 5 to ensure protonation of C75 as expected for it to act as general acid; however, such protonation is less favoured at the physiological pH at which the ribozyme is active, and at which typical *in vitro* experiments are conducted. Accordingly, this catalytically crucial conformation was suggested to represent a functionally important population only partially occupied under standard experimental conditions,[19] consistent with the significant conformational change found to accompany cleavage in solution in this as in all other *trans*-acting ribozymes characterized so far.[20–26]

In addition to the reverse G25U20 wobble, these crystal structures also provided atomic-level resolution of a more common wobble in G1U37, which flanks the ribozyme active site and is in a *cis* orientation.[10,13,16] In general, GU wobble pairs play a special role in RNA biology due to their function in rendering the genetic code promiscuous, and as underscored by their high abundance and frequent conservation in functionally important sites of RNAs.[27–29] This ubiquity of GU wobbles takes advantage of their thermodynamic and geometric near-equivalence to standard Watson–Crick base pairs, as well as their more unique properties as discussed in the following.

Extensive research has shown that both the more common G1U37 wobble and the rarer reverse G25U20 wobble have distinct structural and chemical characteristics.[10,13,16,30–34] Initial biochemical characterization of the HDV ribozyme indicated that mutation of either of its two GU wobbles results in significant decreases in catalytic activity;[33–36] however, more recent evaluation of the G1U37 wobble directly adjacent to the scissile phosphate appeared to suggest more tolerance for alternate base pairing at this position.[30] Recent molecular dynamics (MD) simulations indicate that each GU pair plays a distinct role in metal binding; the G1U37 wobble binds ions more transiently, while the reverse G25U20 wobble exhibits a better-defined binding site to attract longer-residence ions.[37,38] Subsequent studies on a G25AU20C mutant showed disruption of this metal binding, further highlighting the crucial function of these wobbles.[31] However, no systematic tests have been performed to explore the atomic-level structural contributions of these structural motifs to HDV ribozyme active site architecture and stability, providing us with a unique opportunity to probe the functional importance of two distinct GU wobbles within a well-defined structural context.

To this end, we performed extensive comparative MD simulations (for a total of >1 μs simulation time, Table 1) of wild-type, G1U37 and G25U20 mutants using the most recent *trans*-acting crystal structure as a model of a conformation especially poised for catalysis.[16,19] Our results indicate that, on the timescales

 This journal is © the Owner Societies 2015

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

Table 1 List of simulations

Type of simulation	Number of simulations	Duration of simulations (ns)	Mg^{2+} present?
Control simulations			
Wildtype_Na	6	40, 4 × 20, 83	No
Wildtype_Mg	4	120, 78, 2 × 20	Yes
A25Y20 mutants			
A25C20	4	4 × 20	No
A25U20	4	4 × 20	No
R1Y37 mutants			
A1 + C37	4	4 × 20	No
A1°C37	4	4 × 20	No
A1U37	4	4 × 20	No
G1C37	4	4 × 20	No
Y1R37 mutants			
C1G37	4	4 × 20	No
U1G37	4	4 × 20	No

studied, both types of GU wobbles are stable in the presence of either sodium or magnesium at the active site, but that the identity of the active site ion leads to varying in-line fitness of the scissile phosphate. We also find that in-line fitness is further fine-tuned by interactions between the G28 and U-1 hydroxyl groups in simulations containing only sodium. Our mutations of the reverse G25U20 (Table 1) wobble exhibit disruptions of in-line fitness, C75(N3H3+)G1(O5′) general-acid hydrogen bonding, and A25U20 and G1U37 hydrogen bonding. Conversely, alternative hydrogen bonds with nearby functional groups of R1Y37 mutants (Table 1) destabilize active site architecture due to low in-line fitness and disruption of the general-acid hydrogen bond. Further investigation of this site by means of Y1R37 mutants (Table 1) shows comparable, and in some cases increased, base pair and active site stability compared to R1Y37 mutants, indicating the active site's tolerance of pyrimidine–purine pairs at this position. We also find that any active site disorder present in either wild-type or mutant ribozymes is prevented from spreading to the global conformation on the timescale of our simulations due to the type I A-minor and P1.1 motifs above and below the active site, respectively. Taken together, our results predict specific structural rationales for the functional conservation of the two types of GU wobbles in maintaining a catalytically proficient active site. We further suggest that the complex tertiary structure of the HDV ribozyme in part serves to provide redundant, shell-like layers of global structural stabilization. These insights pave the path for understanding structure–function relationships within a plethora of newly discovered HDV-like ribozymes.[5,39,40]

2 Computational details

2.1 Structure preparation

All simulations (Table 1) were based on the most recent, *trans*-acting HDV ribozyme precursor, PDB ID 3NKB.[16] In this common starting structure, the deoxy U-1, G1 and G2 residues were replaced with their ribonucleotide equivalents using the hammerhead ribozyme cleavage site (PDB ID 2OEU), as previously described.[16] Wildtype_Mg structures (Table 1) contained a single magnesium ion near the active site, placed based on structure 3NKB,[16] and were consequently neutralized with sodium. Thus, the equilibration results (see Section 2.2 for simulation procedure) for this structure reflect the behaviour of the magnesium ion when interacting with the modelled-in U-1(O2′).

All mutants were generated from equilibrated structures with an in-line fitness value of approximately 0.5 or greater, before being subjected to a specialized equilibration protocol (see Section 2.2 for details).[41] A25N20 and R1Y37 (Table 1) mutations were made using an equilibrated Wildtype_Na structure with the C75(N3)G1(O5′) "general-acid" heavy-atom distance of 3.1 Å and a good in-line fitness[41] (see below) of 0.66. In all of these cases, (1) all exocyclic atoms and hydrogens were deleted from the bases to be mutated in the text version of the PDB file; (2) the residue and atom names were changed appropriately; and (3) the LEaP module of AMBER10/11 was used to add in any missing heteroatoms and hydrogens. Y1R37 (Table 1) mutations were made in an ~10 ns snapshot from a G1C37 simulation containing C75(N3)G1(O5′) distance of 3.0 Å and an in-line fitness[41] value of 0.46. The Y1R37 mutants required an extra step compared to the A25N20 and R1Y37 protocol described above: a deletion of all atoms of the mutated bases except for the one directly involved in the glycosidic bond (*i.e.*, N1 for pyrimidines and C9 for purines), and the two atoms flanking it (*i.e.*, C2 and C6 for pyrimidines, and C8 and C6 for purines). Steps 1–3 were then followed as for the A25N20 and R1Y37 mutants above. All structures were inspected in XLEaP to ensure base pairing; in cases where pairing did not occur, the bases were adjusted manually to the best possible positions.

Wildtype and all mutant structures (Table 1) were net-neutralized with sodium ions placed by the LEaP module at the points of greatest electrostatic favourability, resulting in an ion concentration of ~0.13−0.16 M. Despite the difference from typical experimental conditions (typically ~11–12 mM Mg^{2+}, *e.g.*, ref. 19, 21 and 33), we used neutralizing monovalents as we have in most of our previous work (*e.g.*, ref. 11, 12, 19 and 42–44). Our rationale for this protocol has been previously summarized;[43,45] most prominently, force field parameters for sodium are very well-defined, unlike those for Mg^{2+}, especially given the latter's strong tendency to polarize the phosphate group, which is ill-described in current force fields. In addition, counter ions naturally condense around the RNA solute so that a bulk experimental ion concentration will not likely provide accurate results.[43,45] The structures were solvated with the TIP3P water box model[46] in a rectangular box reaching 10 Å away from the outermost solute atoms. It took between 18 000–24 000 water molecules to solvate our starting various structures. These structures were then submitted for equilibration (see Section 2.2 for details).

2.2 Simulation procedure

All equilibrations and production runs used the AMBER10/AMBER11 suites of programs[47,48] and the ff99bsc0$_{\chi OL3}$ force field (with modified α/γ[49] and χ[50] torsion angles, essential for preventing large-scale RNA perturbations in longer simulations).

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

Structures were equilibrated in one of two ways. In the first method, Wildtype_Na replicate 6 and Wildtype_Mg replicates 1 and 2 were equilibrated such that all hydrogens were relaxed in the first minimization phase. Next, the solute, counter ions and water were minimized. The water and counter ions were then allowed to mix for 3 ps. Next, harmonic restraints are applied to the backbone atoms of the solute with decreasing force constants of 1000 kcal mol^{-1}, 500 kcal mol^{-1}, 125 kcal mol^{-1} and then 25 kcal mol^{-1} while the rest of the system was allowed to minimize recursively. The system was then heated to 298.16 K over 100 ps, and this resulting structure was used for production runs. The second method was used for all other structures. Wildtype_Na replicates 1–5 were equilibrated according to the previously defined ABC protocol,[51,52] while all mutants (Table 1) were equilibrated with a modified version of the protocol. Briefly, the unaltered ABC protocol began a first minimization of the system using harmonic restraints applied to the solute at 25 kcal mol^{-1}. The system was slowly heated to 300 K over 100 ps, then subjected to minimization/equilibration cycles with decreasing harmonic restraints applied at 5, 4, 3, 2, 1, and 0.5 kcal mol^{-1}. A last 50 ps of unrestrained equilibration was performed, and the final coordinates from this last step were the starting points for production MD. Because we saw no distinct behaviour in Wildtype_Na replicate 6 or Wildtype_Mg replicates 1 and 2 (in-house protocols) compared the rest of our Wildtype_Na and Wildtype_Mg simulations (ABC protocol), we deemed both equilibration methods equally able to relax our system.

We amended the ABC protocol in two ways for all of our mutants. (1) Prior to the initial minimization with 25 kcal mol^{-1} restraints, we applied a minimization/equilibration cycle in which only the mutated base pairs were allowed to relax. (2) We added harmonic restraints (using Sander's ability to implement NMR restraints) between the hydrogen-bonding heavy atoms of the mutated base pairs, as well as the heavy-atom general-acid distance (C75(N3)N1(O5′)). The restraints were implemented as a square parabola to restrain the distances between 2.7 Å and 3.2 Å. The left side of the parabola had a force constant of 50 kcal mol^{-1}, while the right side had a force constant of 100 kcal mol^{-1}. These restraints were implemented during all phases of equilibration, to ensure that our mutants and active site would start each production run with productive geometries. For our A1°C37 mutants, which have only one stabilizing hydrogen bond (Fig. 5b), we implemented the A1 + C37 heavy atom restraints throughout the whole equilibration protocol, as well as for 5 ns into the production runs to compare dynamics before and after restraints were released.

The PMEMD-Sander module of the AMBER10 and AMBER 11[47,48] suite of programs was used for equilibration and production runs. We employed the particle mesh Ewald (PME) method using a 1.0 Å grid spacing, cubic-spline interpolation, and a heuristic pair list update incorporating a 9.0 Å Lennard-Jones cutoff. Our production runs were performed at 300 K (or at 298.16 K for Wildtype_Na replicate 6 and Wildtype_Mg replicates 1 and 2) and constant pressure boundary conditions, using the Berendsen temperature coupling algorithm[53] with a 1.0 ps time constant. We restrained bonds involving hydrogens using the SHAKE protocol.[54] Hydrogen bonds were analysed using the Simulaid suite of programs.[55] and in-line fitness components were calculated using the ptraj module in AMBER10/11.[47,48] Following the rationale in our previous work (*e.g.*, ref. 11, 12, 19 and 42–44), we chose to conduct 4 replicates of each 20 ns for each of our mutants, instead of longer simulations, for two reasons: (1) in longer simulations, the later time points are dependent on the evolution of earlier time points, whereas in our smaller simulations, we are able to investigate greater phase space with multiple fresh simulations; and (2) longer simulations are more vulnerable to force field artefacts.

In-line fitness (F) was calculated using the following equation as described:[41,56]

$$F = \frac{(\tau - 45)}{180 - 45} \times \frac{3^3}{d_{O2',P}{}^3}$$

where τ is the angle between the 2′ oxygen nucleophile (O2′), the phosphorus electrophile (P) and the 5′ oxygen of the leaving group (O5′) and d is the distance between the 2′ oxygen nucleophile and the phosphorus electrophile Briefly, this parameter combines the nucleophile's in-line attack angle and the distance between nucleophile and electrophile into a single parameter that, normalized to a range of 0–1 with 1 being highest, describes the propensity of that phosphodiester bond to be cleaved. Ideal ranges for τ and d are 160–180° and 3.0–3.5 Å, respectively.[41]

3 Results and discussion

3.1 The wildtype active site architecture is better stabilized by Mg^{2+} than Na^+

MD simulations are well suited to capture fast, microscopic biomolecular dynamics at atomic resolution.[45,57,58] To interrogate the HDV ribozyme, we used the most recent *trans*-acting (two-stranded) precursor structure 3NKB[16] as a representation of a conformationally "activated" structure.[19] To relate the geometry of the scissile phosphate to potential effects on catalysis, we employed the previously described in-line fitness parameter.[41,56] The use of this parameter provided us with a straightforward metric to correlate with the retention of both GU pairs and other active site architectural parameters.

The *trans*-acting crystal structure[16] features a magnesium ion in the active site, allowing us to ask how it may contribute to maintaining the local architecture. Experimentally, the HDV ribozyme has shown distinct behaviours in monovalent and divalent metal ion environments.[7,59] Although cleavage is optimal in the presence of Mg^{2+}, it also occurs in high (molar) concentrations of Na^+.[7,59] pH-rate studies indicate a potential change in mechanism from monovalent to divalent metal ion conditions.[7] The magnesium ion in our Wildtype_Mg simulations (Table 1) was placed based on the Mg^{2+} ion near the crystal structure's L3 loop.[16] This specific location was proposed to support a potential role of Mg^{2+} as a general and/or Lewis acid during catalysis.[16] However, as discussed in Section 2 and in ref. 16, the electron densities of the scissile phosphate and U-1 nucleotide are missing from structure 3NKB. We modelled these missing moieties based

 This journal is © the Owner Societies 2015

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

on previous protocols,[16] and the structures used for our and all previous[37,38] production runs consequently are a result of the equilibration between the active-site magnesium ion and modelled scissile phosphate. To ask whether the active-site Mg^{2+} ion helps rationalize available experimental observations, we conducted comparative simulations (Table 1) termed Wildtype_Mg (containing the active-site magnesium and net-neutralizing sodium ions) and Wildtype_Na (neutralized with sodium ions alone).

Both Wildtype_Mg and Wildtype_Na simulations showed almost complete retention of the G1U37 and G25U20 wobbles throughout the simulation timescale (Fig. 2a). Despite exhibiting such stable lateral boundaries, however, we found that the scissile phosphate was able to sample a wide range of conformations dependent on metal ion identity (Fig. 2b). As previously observed,[56] Wildtype_Na simulations in particular did not exhibit high in-line fitness across the board, and generally revealed quite static behaviour; that is, a simulation typically started with a favourable or unfavourable value and maintained that preference throughout a single simulation. By comparison, the Wildtype_Mg simulations experienced more dynamics in in-line fitness, with a large percentage of values ≥ 0.7. We thus conclude that the presence of an active-site magnesium ion appears to induce the scissile phosphate to more often sample conformations optimal for catalysis. Our simulations thus provide a potential structural role for this magnesium ion that so far only has been strongly implicated in playing a more direct catalytic role.[16,31,37,38]

Next, we studied the effects of varying the active-site ion on the C75(N3H3+)G1(O5′) general-acid hydrogen bond (Fig. 2c), an additional important marker of active site stability. In our high in-line fitness Wildtype_Na simulations (replicates 4–6 in Fig. 2b), we found a C75(N3H3+)G1(O5′) occupancy of $\geq 80\%$ (Fig. 2c and d). By contrast, trajectories with low in-line fitness (replicates 1–3 in Fig. 2b) are only in rare cases able to retain the general-acid hydrogen bond (Fig. 2c). Such an exception is replicate 3, which exhibited unfavourable in-line fitness and a largely retained C75(N3H3+)G1(O5′) hydrogen bond. Conversely, in our Wildtype_Mg simulations, the catalytically important C75(N3H3+)G1(O5′) interaction was always retained (Fig. 2c and d). These results invoke a second structural role for the catalytically implicated magnesium ion.[16,31,37,38]

During investigation of additional, external influences on in-line fitness, we observed distinct interactions between the 2′ hydroxyl groups of U-1 and G28 in Wildtype_Na *versus* Wildtype_Mg simulations (Fig. 3). In most Wildtype_Na replicates of high in-line fitness, the U-1 2′ hydroxyl typically pointed towards the scissile phosphate (Fig. 3b), enabling the formation of a hydrogen bond between the G28(HO2′) hydrogen and U-1(O2′) (Fig. 3a). The exception to this observation was replicate 5, in which high in-line fitness was maintained, but the G28(HO2′)U-1(O2′) interaction was disrupted due to transient ionic interactions with U-1(O2′). Despite this exception, we generally observed this interaction to have a stabilizing effect on in-line fitness, partially compensating for the reduced structural stability due to the missing magnesium ion in this vicinity.

A different G28–U-1 interaction occurred in Wildtype_Mg simulations. With the active site Mg^{2+} ion closely coordinated to U-1(O2′) and G1(O2P) (Fig. 2d and 3b), the U-1(HO2′) typically pointed to the outer reach of the L3 loop. In this conformation, the U-1 2′ hydroxyl became the hydrogen bond donor to G28(O2′). This bond was often disrupted, however (Fig. 3a, left). Thus, this hydrogen bond appears to be more of a consequence of the prevalent Wildtype_Mg U-1(O2′) conformation rather than to be buttressing high in-line fitness as it does in Wildtype_Na simulations.

In summary, despite the lack of U-1 and scissile phosphate electron density in the 3NKB crystal structure, we find that the active-site magnesium ion appears to result in stabilization of architecturally important cleavage site interactions during the course of our trajectories. In our MD simulations, the active site magnesium leads to greater dynamics of the in-line fitness, supports the C75(N3H3+)G1(O5′) hydrogen bond, and renders interactions of the G28 and U-1 hydroxyls less critical. The distinct features of sodium-only simulations suggest a rationale for the proposed, ion-dependent multi-channel reaction mechanism of the HDV ribozyme.[7,59]

3.2 A25N20 mutants disrupt active site stability

Having ascertained that the HDV ribozyme's GU wobbles are stable regardless of ion identity (Fig. 2a), we performed a thorough mutational probing of each wobble. We began with the *trans* G25U20 wobble, taking advantage of recent characterizations of its role in metal binding.[31,37,38] In particular, this wobble has been suggested to bind a magnesium ion implicated as a Lewis-type acid, a Bronsted-type base, or both, during catalysis.[16,31] An A25C20 mutation inhibits the ability of the L3 loop to bind this catalytic Mg^{2+} ion.[31] Earlier mutational studies also showed that this wobble is crucial for effective catalysis.[33] However, no recent investigation has been conducted on the effects of the *trans* G25U20 wobble on active site architecture. We used a set of A25N20 simulations (Table 1) to address this question. An A25C20 mutant, isosteric to the native G25U20 wobble, exhibited a stable mutant base pair during our entire trajectory timescale, similar to previous results.[31] However, this mutant often showed low in-line fitness and loss of other active site hydrogen bonds (Fig. 4). Among our four replicates, two trajectories (replicates 2 and 4) exhibited high in-line fitness, and two trajectories displayed in-line fitness predominantly <0.5. Only replicate 2 had high C75(N3H3+)G1(O5′) hydrogen bond occupancy (Fig. 4a). In low in-line fitness replicates 1 and 3, the active site geometry was disrupted likely due to increased L3 dynamics as a result of the loss of the putative catalytic ion, a result that has been previously predicted.[31] The low in-line fitness may thus be partially explained by the previously proposed decrease in negative surface charge potential.[31] However, while an ion was also not present in replicate 2, in this trajectory the general acid C75(N3H3+)G1(O5′) hydrogen bond was retained.

Here, structural dynamics enabled alternative hydrogen bonds to form between C75 and the scissile phosphate's non-bridging oxygens, stabilizing new conformations (Fig. 4a). These alternative C75-scissile phosphate hydrogen bonds sometimes, but not always, affected in-line fitness; for example, while replicate 2 showed both high in-line fitness and stable C75(N3H3+)G1(O5′)

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

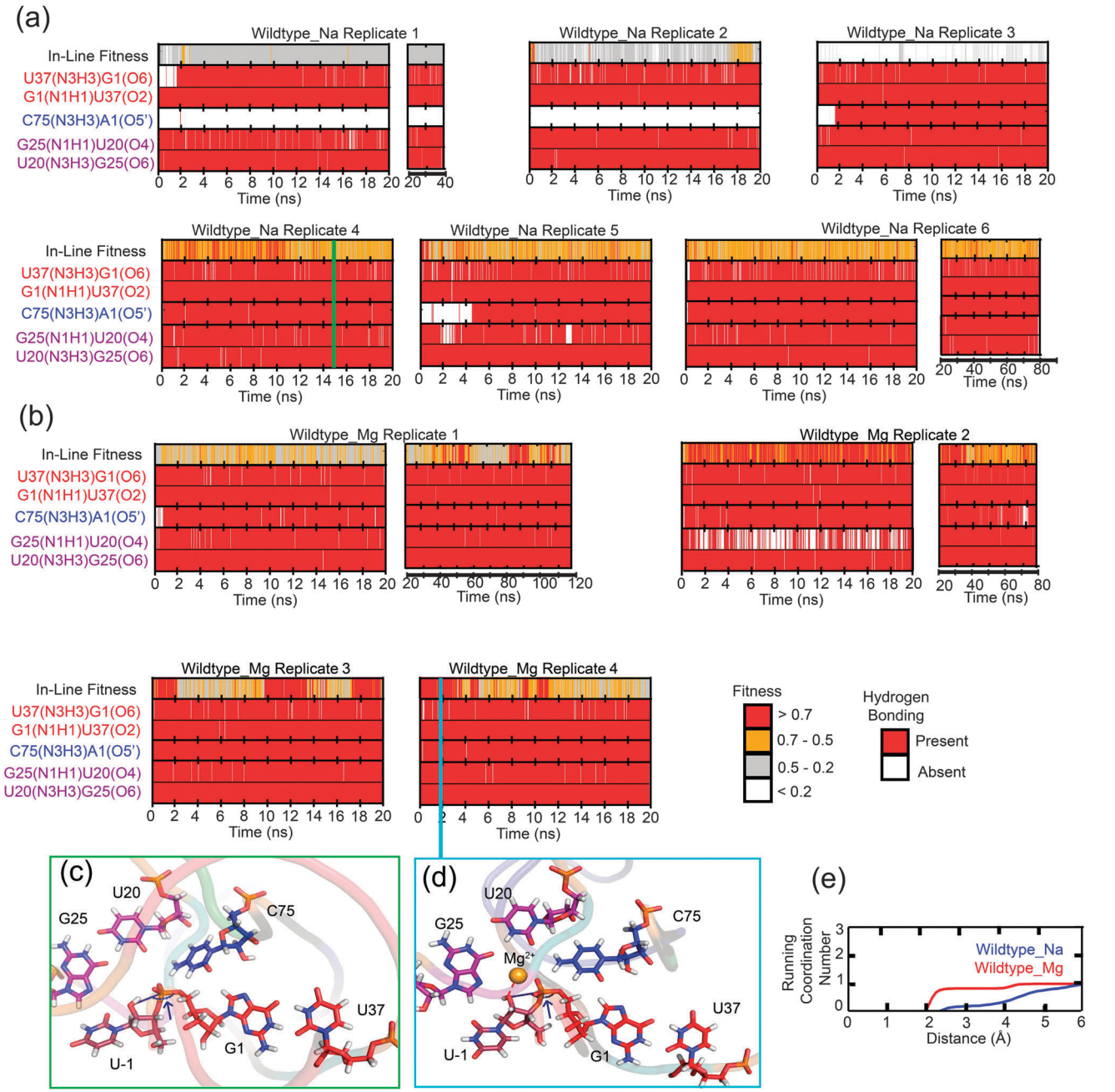

Fig. 2 Active site dynamics of the HDV ribozyme in the presence of Mg^{2+} *versus* Na^{+} in the active site. (a) In-line fitness, hydrogen bonding patterns of the G1U37 wobble, C75(N3H3)G1(O5′) and G25U20 wobble of the Wildtype_Na (a) and Wildtype_Mg (b) trajectories. Note that trajectories are of different lengths; the first 20 ns of each trajectory are given the same scale for comparison purposes, while the remainder of any longer trajectories are condensed in subplots to the right of their respective trajectories. These hydrogen bonds were deemed ''present'' if (i) the donor–hydrogen–acceptor angle was $\geq 120^{\circ}$ and (ii) the hydrogen–acceptor distance was ≤ 3 Å. Representative conformations of active site residues in the absence (c) and presence (d) of the long-residence active-site Mg^{2+} ion. Residues are coloured as in Fig. 1, with oxygens in red, nitrogens in blue, hydrogens in white and phosphorus in orange for clarity. Images were made from snapshots averaged over 100 ps. The bright green and cyan lines in panel (b) identify from where the snapshots were taken. Dark blue-and-red dashed lines (distance) and dark blue arcs and arrows (in-line attack angle) at the scissile phosphates indicate the parameters needed to calculate in-line fitness. Dashed bright red lines in the right panel indicate coordination of the U-1(O2′) to the Mg^{2+} ion. (e) Comparison of ion retention at the U-1(O2′) between Wildtype_Na and Wildtype_Mg simulations. Running coordination numbers indicate the number of ions at a given distance from the U-1(O2′). Curves are a weighted (by trajectory length) average of all six Wildtype_Na replicates and all four Wildtype_Mg replicates, respectively.

hydrogen bonding, replicate 4 showed 10 ns of high-in-line fitness that appeared to be disrupted at least partially by a shift in C75(N3H3+) hydrogen bonding from G1(O1P) to G1(O2P) (Fig. 4a). Despite the differences in in-line fitness and hydrogen bond occupancies among replicates, the G1U37 wobble stayed mostly intact except for a 4 ns interruption in the U37(N3H3)G1(O6)

 This journal is © the Owner Societies 2015

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

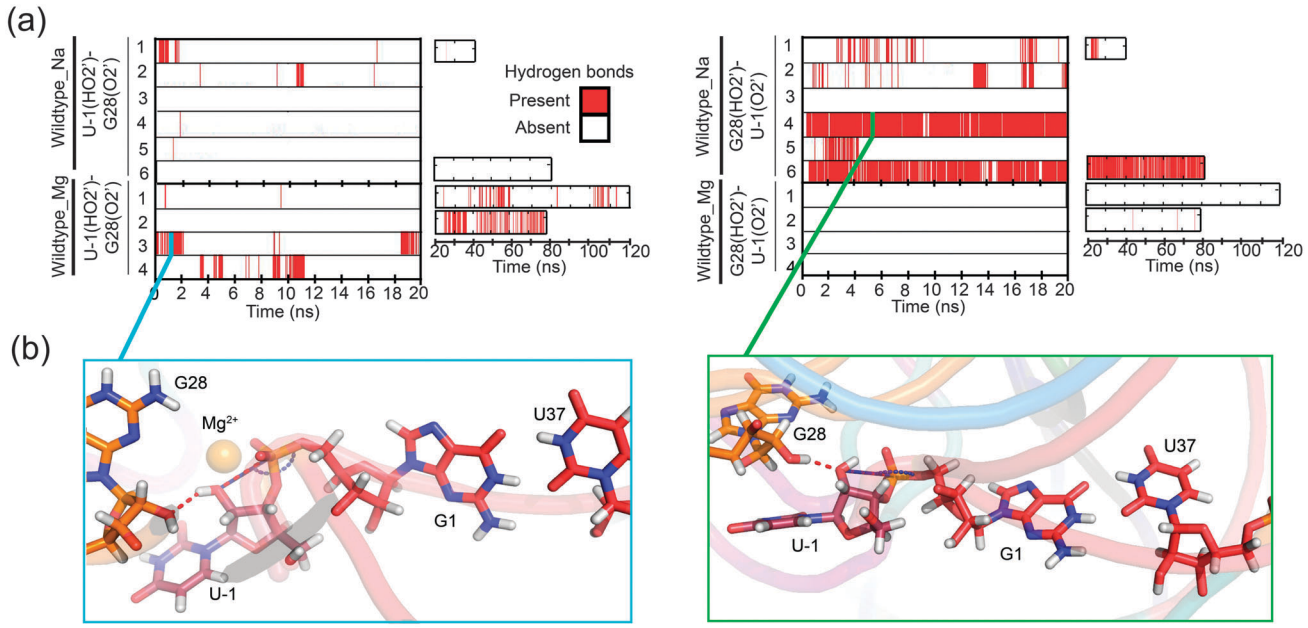

Fig. 3 Effects of G28 on scissile phosphate conformation. (a) Occurrence of the two possible hydrogen-bonding interactions between the 2′hydroxyls of G28 and U-1. Note that the G28(2OH′)U-1(O2′) hydrogen bonding patterns of the Wildtype simulations often match periods of high in-line fitness as observed in Fig. 2(b). Cyan and bright green lines indicate from which replicate and at which time the snapshots in panel b were taken. (b) Snapshots of G28–U-1 2′ hydroxyl interactions in Wildtype_Mg (left) and Wildtype_Na (right) simulations. Snapshots are from structures averaged over 101 ps or 102 ps of the indicated trajectory. Red dashed lines indicate hydrogen bonds formed in the snapshot. Dark blue-and-red dashed lines and dark blue arcs at the scissile phosphate indicate measured parameters used to calculate in-line fitness.

distance in replicate 1. This general retention may be attributable to the isostericity of the A25C20 Watson–Crick pair with the *trans* G25U20 wobble.

Next, we tested an A25U20 mutant and found it to sample high in-line fitness more often than the A25C20 mutant, yet significant active site disruption still occurred (Fig. 4b). In particular, while all four replicates of this mutant showed at least some high in-line fitness, retention of the C75(N3H3+)G1(O5′) general acid bond was comparably low to that observed for the A25C20 mutant (only one of four replicates), and similar alternative hydrogen bonds between C75 and G1's non-bridging oxygens formed in long residence (Fig. 4b). Interestingly, we observed greater dynamics in either the G1U37 wobble (replicates 2 and 3) or the A25U20 base pair (replicate 4) compared to the A25C20 mutants (Fig. 4a and b). Replicate 4 again exhibited the redundancy inherent in the HDV ribozyme active site; although the A25U20 base pair was disrupted, alternative C75-non-bridging oxygen and G1U37 hydrogen bonds formed stably, providing an environment favourable for high in-line fitness. The overall increased active site instability in this mutant compared to A25C20 may be due to the non-isostericity of A25U20 with the native G25U20 (Fig. 4c, right), which may contribute to the observed greater L3 dynamics that transfer to the G1U37 wobble.

In conclusion, we found that both isosteric and, to a greater extent, non-isosteric canonical mutations to the *trans* G25U20 wobble disrupt active site dynamics during the course of our simulations. In-line fitness is poorer compared to both types of Wildtype simulations, and the canonical C75 and G1U37 hydrogen bonds are disrupted in both A25U20 mutants. This relative active site instability in A25N20 mutants provides a structural rationale that, together with previously described ion binding deficiencies of the A25C20 mutant,[31] explains experimentally measured decreases in catalytic rate of these mutants.[33]

3.3 R1Y37 mutants disrupt local architecture

To determine whether mutations at the G1U37 wobble affect G25U20 stability, analogous to the A25U20 disruptions of G1U37 hydrogen bonding described in the last section, we simulated all possible R1Y37 mutants (Table 1). Over the timescales of our trajectories, all of these mutants showed a decrease in active site stability compared to both types of Wildtype simulations, in addition to a less stable mutant pair than observed for the A25C20 mutants (Fig. 5 and 6). The lack of stability in the mutant base pair was due to the formation of alternative hydrogen bonding interactions. In particular, most A1Y37 simulations exhibited interactions between the exocyclic amine of A1 with the non-bridging oxygens of nearby G76 (Fig. 5a, 6a and c), while the dynamics in the G1C37 base pair were due to shifting Watson–Crick interactions (Fig. 6b and d). No R1Y37 mutants disrupted the G25U20 wobble, however, to the same extent as the A25U20 mutants disrupted the G1U37 wobble (compare Fig. 4–6).

We also simulated protonated and unprotonated versions of the A1C37 wobble to model the effects of this mutant at low and high pH, respectively (Fig. 5). Although the A1 + C37 base pair represents an isosteric replacement to the G1U37 wobble (Fig. 5c, left), this mutant showed decreased stability of active site interactions compared to Wildtype (Fig. 5a and c). Two of the four replicates (replicates 1 and 4) showed stable mutated

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

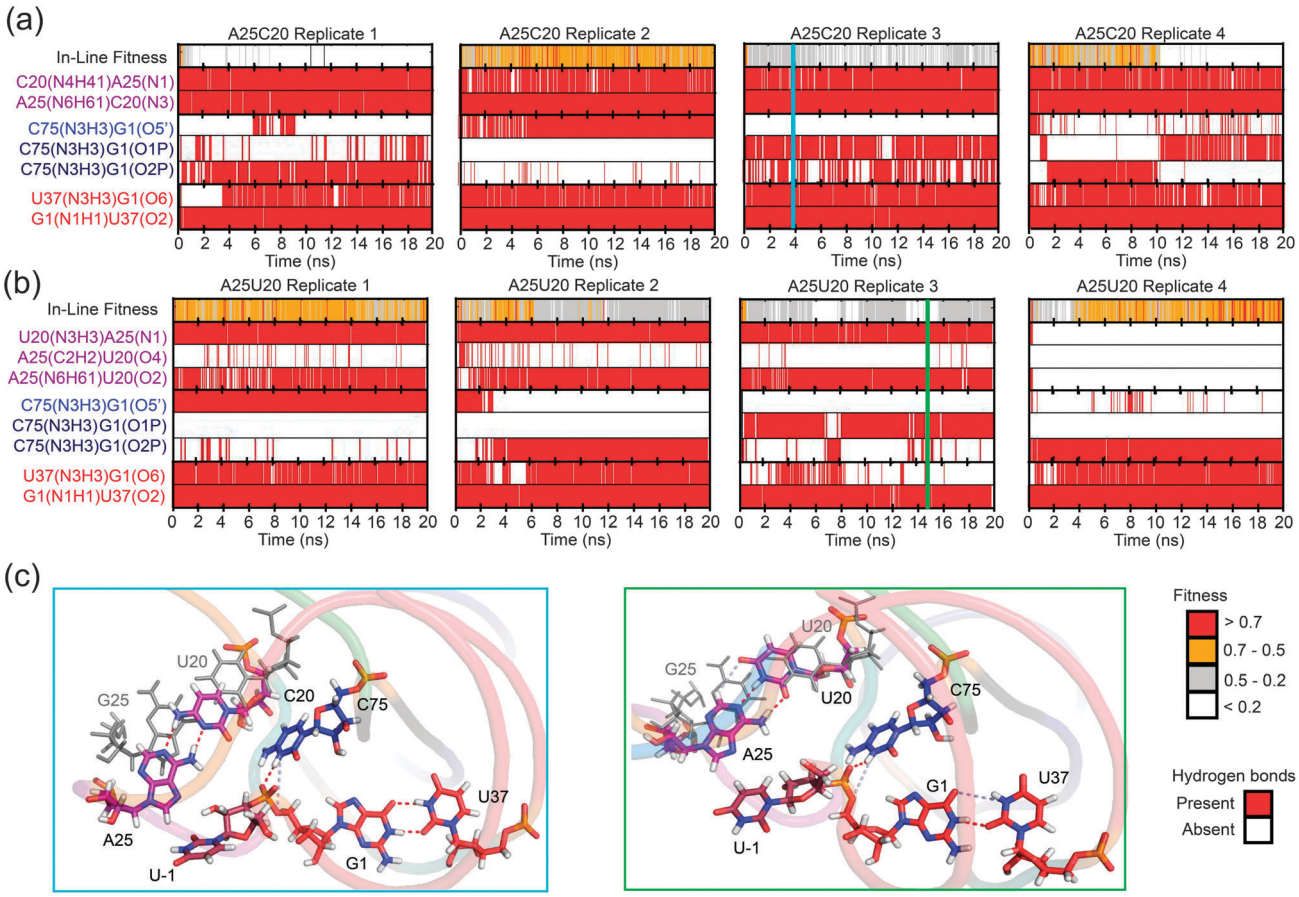

Fig. 4 Active site dynamics of A25N20 mutants. In-line fitness and active site hydrogen bonds of the (a) A25C20 and (b) A25U20 mutants. Text colours are consistent with Fig. 1 where purple indicates the A25N20 wobble; blue indicates canonical, and dark blue indicates non-canonical, hydrogen bonds made by C75; red indicates G1U37 hydrogen bonding interactions. Cyan and bright green lines indicate from which replicate and at which time the snapshots in panel C were taken. (c) Snapshots of representative active site dynamics in the A25C20 (left) and A25U20 (right) mutants. Red dashed lines indicate hydrogen bonds stable during the time of snapshot, while light blue dashed lines indicate hydrogen bonds that are destabilized during the length of the trajectory. Colours of nucleotides are consistent with Fig. 1, and oxygens are red, nitrogens are blue, hydrogens are white, and phosphorus is orange for clarity. In both snapshots, the native G25U20 wobble is shown in dark grey lines for reference.

A1 + C37 hydrogen bonding and high in-line fitness, similar to the A25C20 simulations (Fig. 5a); however, both replicates showed only partial retention of the C75(N3H3+)A1(O5′) hydrogen bond, as well as non-canonical C75(N3H3+)A1(O2P)(O1P) hydrogen bonds (Fig. 5c). The A1 + C37 mutation also exhibited dynamic hydrogen bonding in the other two replicates (replicates 2 and 3), in contrast to the G1U37 wobble in both types of Wildtype simulations (Fig. 2). Thus, we found that even isosteric mutations are not as structurally stable as the native G1U37 wobble, although the structural destabilization is not observed to the same extent as with the A25U20 mutants (see previous section).

We also simulated the unprotonated A1°C37 wobble, to emulate the structural effects of pH values >6. Interestingly, this wobble showed enhanced high in-line fitness and active site stability compared to the protonated version. Although the single A1(N6H61)C37(N3) hydrogen bond was not retained after restraints were released (see Materials and Methods), alternative Hoogsteen–Sugar Edge interactions between the A1 exocyclic amine and the C37 2′ hydroxyl oxygen appeared to aid in stabilization of the rearranged conformation (Fig. 5b and c, right). Our findings thus underscore the structural robustness that the G1U37 lends to the HDV ribozyme active site. When this wobble is not present, the active site is able to retain comparable-to-wildtype geometry, in which alternate hydrogen bonding between different faces of the N1N37 bases are able to support a catalytically fit geometry.

Our two Watson–Crick R1Y37 mutants, A1U37 and G1C37, showed greater dynamics in various active site structural motifs than observed in the A1C37 mutants (Fig. 6). For the A1U37 mutants, alternative A1(N6)G76(O1/O2P) interactions that formed neither stabilized nor destabilized the base pair, but appeared simply to be a result of stochastic base pair dynamics (Fig. 6a and c). These trajectories exhibited the lowest percentage of favourable in-line fitness of all mutants, with not even one entire trajectory exhibiting high values. C75(N3H3+)A1(O5′) interactions were also present for only a small percentage of one trajectory, with the alternative C75–A1 interactions stabilizing non-wildtype conformations (Fig. 6a and c).

G1C37 mutants showed similar behaviour to the A1 + C37 mutants in terms of active site stability (Fig. 6b and c). For these

 This journal is © the Owner Societies 2015

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

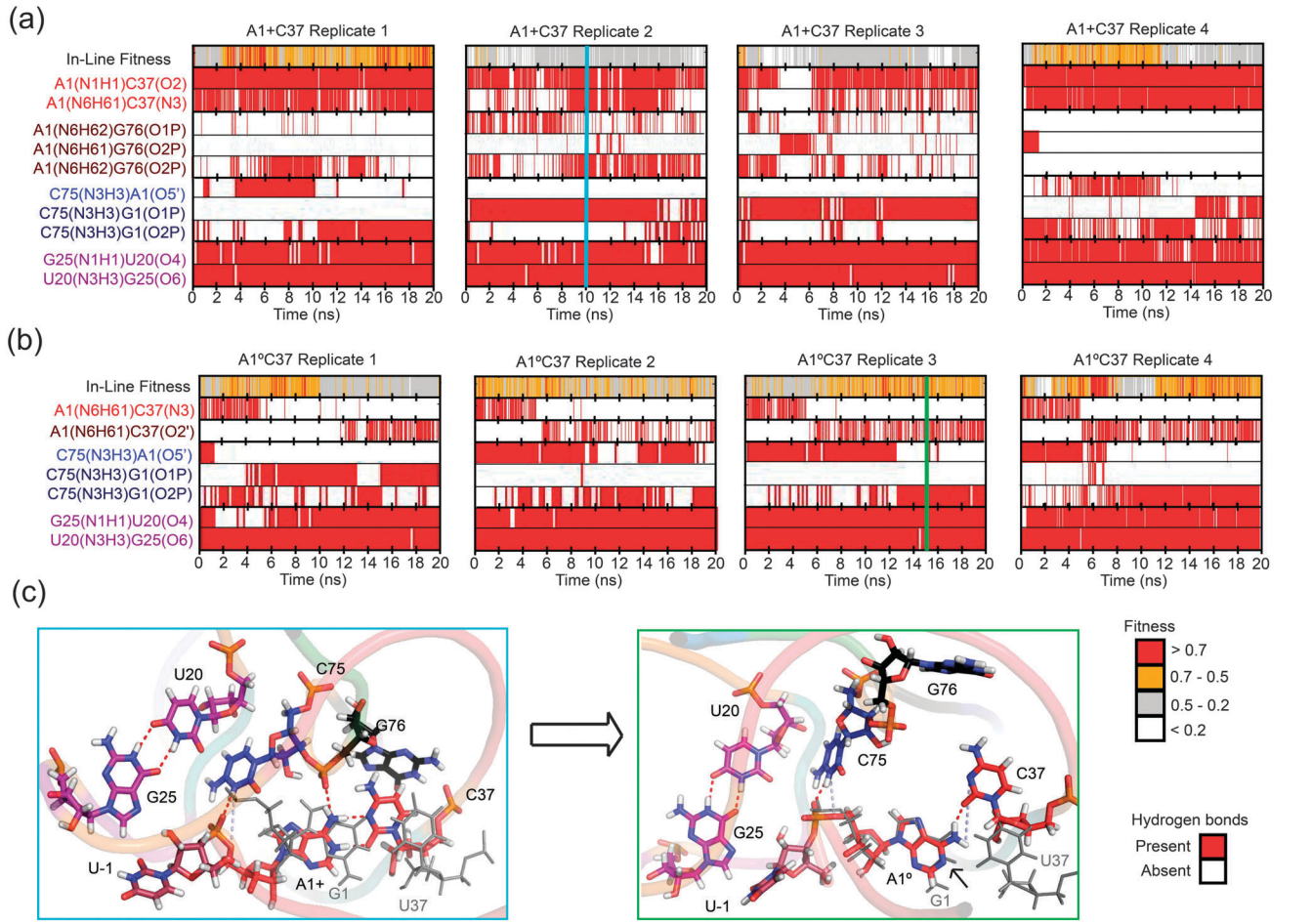

Fig. 5 Active site dynamics of the A1C37 wobble. In-line fitness and active site hydrogen bonds of the (a) A1 + C37 and (b) A1°C37 mutants. Colours correspond to Fig. 1; canonical and non-canonical A1C37 hydrogen bonds are labelled in red and dark red respectively; canonical and non-canonical C75 interactions are in blue and dark blue respectively; and G25U20 interactions are in purple. (c) Conformational differences between protonated (left) and unprotonated (right) A1C37 wobbles. The arrow indicates a change in pH. Colours of the nucleotides are retained from Fig. 1, while important heavy atoms are coloured as in previous figures. Similar to Fig. 4, the native G1U37 base pair is shown in dark grey lines.

mutants, altered dynamics were observed in the form of a shifting of the Watson–Crick interactions. In particular, we observed dynamic exchanges between the canonical (C37(N4H41)G1(O6), G1(N1H1)C37(N3)) and non-canonical (G1(N1H1)C37(O2)) hydrogen bonds, often within a single trajectory (Fig. 6b and d). The former canonical hydrogen bonds appeared to occur when the G1C37 base pair occupied a similar location to the native G1U37 pair (Fig. 6d, right). This shifting resulted in loss of high in-line fitness and lack of C75(N3H3+)G1(O5′) interactions in replicates 1 and 2 (Fig. 6b). Interestingly, neither high in-line fitness nor general acid C75(N3H3+)G1(O5′) hydrogen bonding appeared to be dependent solely on G1C37 base pair stability; these two features are relatively stable in both replicates 3 and 4, despite varying dynamics of the G1C37 base pair (Fig. 6b). In the G1C37 mutant simulations, we observed for the first time an extended disruption in G25U20 hydrogen bonding (Fig. 6b, replicate 4), occurring in tandem with a relatively stable G1C37 base pair. This behaviour is similar to that seen in our A25U20 mutations; for both of these mutants, it thus appears that hydrogen bonds can only be stable in either the 1 : 37 or the 25 : 20 base pair of the HDV ribozyme active site, but not in both simultaneously. This behaviour contrasts with that of both native structures (Fig. 2), in which both the GU pairs are stable.

Decreased in-line fitness in the G1C37 mutants is especially interesting in the context of the human HDV-like (CPEB3) ribozyme,[5] found in an intron of the cytoplasmic polyadenylation element-binding protein gene. The CPEB3 gene sequence exhibits a single nucleotide polymorphism (SNP) of G1C37,[5] which has been shown to increase the CPEB3 ribozyme's cleavage rate, contrary to our observations here. We note, however, that the structural differences between the HDV and CPEB3 ribozymes likely translate into different roles for their 1–37 base pairs.[5] Specifically, the P1.1 helix of the CPEB3 ribozyme consists of only one base pair, which is known to strongly reduce catalytic activity.[15] A G1C37 polymorphism may thus add needed stability in the CPEB3 ribozyme that in the HDV ribozyme is supplied instead by its second P1.1 base pair.

Our findings thus indicate that a canonical Watson–Crick base pair in either of the two GU wobble positions may introduce differential conformational alignment that has a destabilizing

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

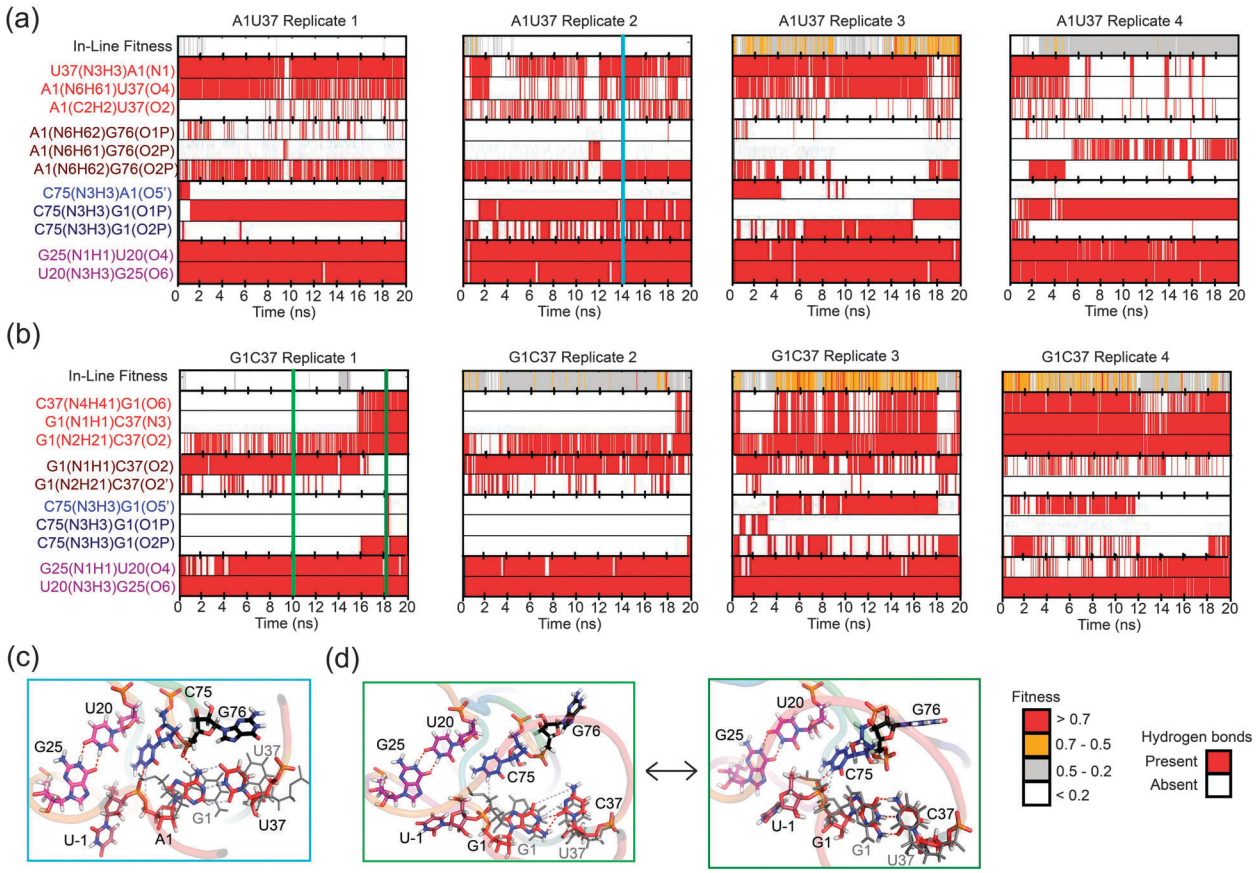

Fig. 6 In-line fitness and active site hydrogen bonds of (a) A1U37 and (b) G1C37 mutants. Colouring of labels and plots is in keeping with previous figures. An additional coloured line is added in panel (b) to indicate the replicate and time from which the second snapshot of the G1C37 mutant in panel (d) is taken. Snapshots of (c) A1U37 and (d) G1C37 representative structures. As in previous figures, strongly retained hydrogen bonds are shown in red dashed lines, while canonical, dynamic hydrogen bonds are shown in blue. In panel (d), the shifted (left) and canonical (right) Watson–Crick interactions of the G1C37 mutant are both shown, to indicate their interconverting nature. Similar to previous figures, the native G1U37 base pair is shown in dark grey lines.

effect on the other position over the time course of our simulations. Taken together, our results thus suggest that no type of canonical purine–pyrimidine Watson–Crick base pair is tolerated at the active site, due to a decrease in in-line fitness and the destabilization of both canonical active site interactions as well as of the base pairs themselves. These mutants are, however, able to form alternative hydrogen bonding interactions that may aid in preventing active site disorder from spreading throughout the entire molecule.

A wealth of experimental data exists for a variety of *cis*- and *trans*-acting genomic and antigenomic ribozyme mutants of the G1U37 wobble pair.[31,33–35,37,38] Early studies on A1°C37 *trans*-acting mutants already showed a decreased cleavage rate, but an almost wild-type extent of cleavage.[35] Although these studies contained magnesium, unlike our simulations, our A1°C37 simulations exhibit similar trends: high in-line fitness, yet disrupted C75(N3H3+)A1(O5′) general acid bonds. A1 + C37, and all other R1Y37, rate constants were measured in recent kinetic studies.[30] This study showed similar activity for all R1Y37 mutants in the pH range within which A1+ remained protonated. Beyond the protonation range of A1+, the work reported low activity and an alternate, magnesium-dependent mechanism.[30] Our MD simulations exhibit alternative trends, which indicate that active site architecture is less stable with the protonated A1 + C37 base pair than with the isosteric G1U37 wildtype, which may be due to the lack of magnesium in our simulations. Our ~50% catalytically fit G1C37 MD simulation trajectories are more consistent with trends observed for a *cis*-acting ribozyme mutant,[34] despite the variety of ribozyme constructs used in these early activity studies.[35,60] The still low percentage of catalytically fit trajectories that we observe may again be due to lack of magnesium in our simulations.

In conclusion, all of our R1Y37 mutants display low in-line fitness, decreased occupancy of the general acid C75(N3H3+)N1(O5′) hydrogen bond and dynamic Watson–Crick mutated base pairing (during the course of our trajectories), explaining the high conservation of the wildtype G1U37 wobble in clinical isolates.[61] The highest probability of favourable in-line fitness for these mutants is found in A1C37 wobbles, but is still not comparable in stability to that seen in both Wildtype structures. These results indicate that the G1U37 wobble is unsurpassed in stabilizing favourable in-line fitness and active site architecture in the HDV ribozyme. We also found that alternative hydrogen bonding interactions are prevalent in this

 This journal is © the Owner Societies 2015

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

group of mutants, which may help contain active site disorder. These alternative interactions occur to a greater extent in most of our R1Y37 mutants than in our A25U20 mutants, indicating that the G1U37 wobble has a greater effect on local conformational retention, while the G25U20 wobble effectively stabilizes elements both proximal and distal to its position in the L3 loop (Fig. 1).

3.4 Y1R37 mutants show greater active site stability compared to R1Y37 mutants

An interesting feature of GU wobbles is that they are not isosteric to the inverse UG wobbles.[17,27,62,63] We thus investigated U1G37 mutant "inverse" wobbles, to study their effects on in-line fitness and active-site geometry. Interestingly, we found that all of our U1G37 replicates exhibited nearly ideal active site stability, greater even than our Wildtype_Na and Wildtype_Mg simulations, despite their non-isostericity (Fig. 7a and c). The mutant base pair is stable, and both the G25U20 and, in all but one replicate, C75(N3H3+)U1(O5′) hydrogen bonds are retained (Fig. 7a). The active site is indeed so rigid that even in the absence of the general acid bond in replicate 4 (Fig. 7a) high fitness and stable hydrogen bonding are maintained. Based on previous work,[41] our results indicate that experimental constructs of our U1G37 mutants would be more active than constructs of our wildtype structures. The high occurrence of favourable in-line fitness in our U1G37 simulations is thus in direct contrast to early experimental results, which predict a loss in activity for these mutants in both *cis*- and *trans*-acting, genomic and antigenomic HDV ribozymes.[34,35,64] This discrepancy may be due to changes in Mg^{2+} binding properties upon mutation, or due to the fact that these early experimental studies reported significant miscleavage of this mutant at the +1 position[35] and contained a significantly longer 5′ leader sequence[34,35,64] than does our simulated, experimentally optimized ribozyme.[16]

We next tested the active site conformational stability of a C1G37 mutant and found replicates to show some occurrence of high in-line fitness (≥ 0.5) over the simulation timescale, although to a lesser extent than our U1G37 mutant (Fig. 7b). Similar to the G1C37 mutant (Fig. 6d), the C1G37 base pair seems to result in high in-line fitness when it occupies a position similar to the native G1U37 pair (Fig. 7c, left). The observed favourable fitness values are not always dependent on retained C75(N3H3+)C1(O5′) hydrogen bonding, but in contrast to the R1Y37 mutants, the mutated base pair is always stable (Fig. 7b). Thus, the decreased activity observed in mutational studies is likely more dependent on disruption of the general acid bond than on unfavourable in-line fitness, similar to replicate 4 of the U1G37 mutant (Fig. 7a). In further contrast to the R1Y37 mutants, the alternate C1(N4H42)G76(O2P) interaction is so high in occupancy (Fig. 7b) that it may act as a stabilizing feature for the mutated C1G37 base pair. The observed fluctuations in in-line fitness, coupled with stable active site geometry, explain the high cleavage extent, yet low catalytic rate, observed in experimental

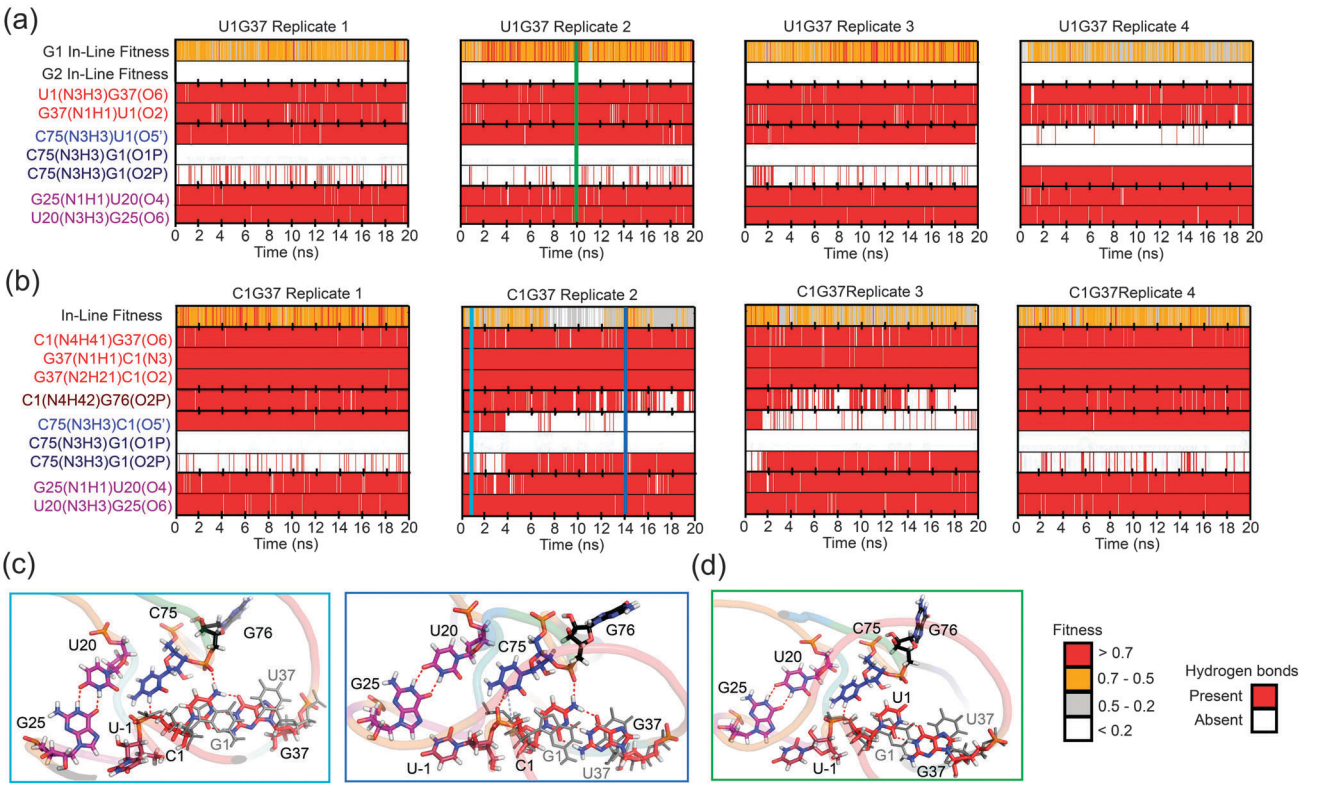

Fig. 7 In-line fitness, Y1R37, C75 interactions and G25U20 hydrogen bonds of (a) U1G37 and (b) C1G37 mutants. Snapshots of relevant conformations of (c) C1G37 and (d) U1G37 mutants. Colouring is similar to previous figures. An extra dark blue line in the C1G37 trajectories defines a snapshot with alternative hydrogen bonding interactions. Similar to previous figures, the native G1U37 base pair is shown in dark grey lines.

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

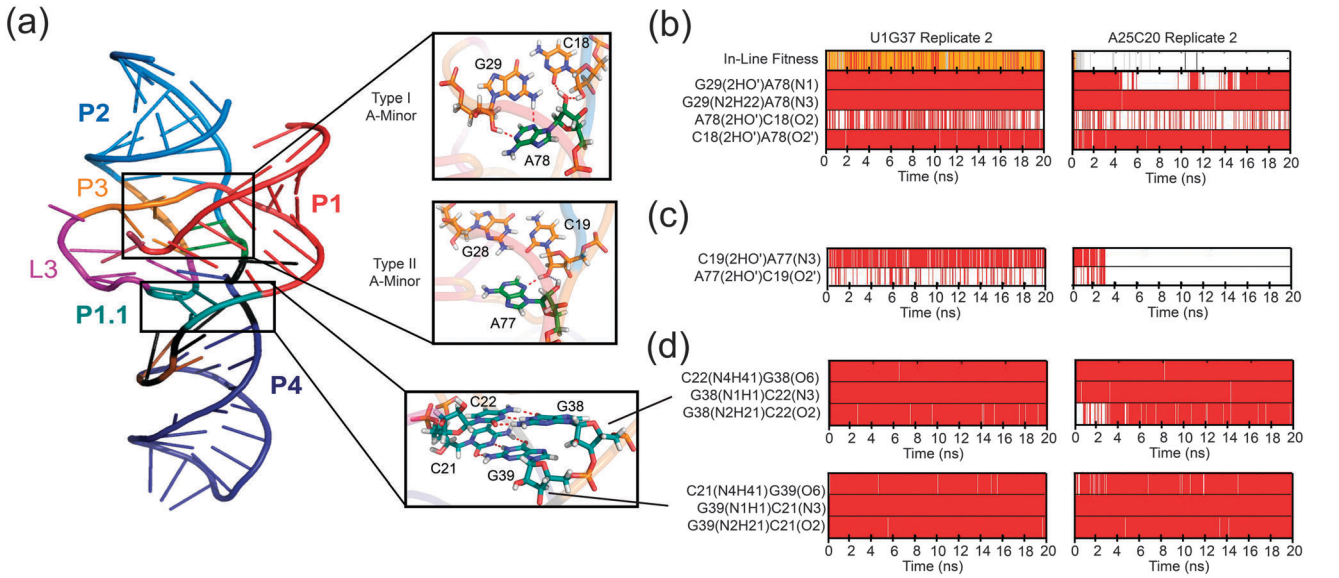

Fig. 8 (a) Global and close-up views of type I and type II A-minor interactions and the P1.1 helix. Representative time traces of type I (b) and type II (c) A-minor interactions as well as P1.1 helix (d) hydrogen bonds for high in-line fitness (U1G37, left) and low in-line fitness (A25C20, right) trajectories.

studies of both genomic and antigenomic C1G37 ribozymes.[35,64] Taken together, our results indicate that Y1R37 mutations are tolerated at the G1U37 wobble better than are R1Y37 mutants.

3.5 Active site disorder is contained above and below by structurally inert motifs

The HDV ribozyme's complex tertiary structure is held together by several specific motifs: type I and type II A-minor interactions with P3 stabilize coaxial stacking of P2 onto P3 and form the top boundary of the active site (Fig. 1 and 8a).[10,13,65] Additionally, the 2-base-pair P1.1 minihelix buttresses the bottom of the active site (Fig. 8a).[10,13] Because we observed active site disorder in all but our Wildtype simulations, we investigated how each of these motifs is affected by that disorder (Fig. 8). Disruption of either or both of these motifs and consequent global structural disruption could provide an additional explanation for experimentally observed loss in activity for some of these mutants.[31,33–35,64]

A-minor interactions represent the most frequent tertiary interactions in folded RNAs.[65] Upon their initial discovery, their characterization diversified into four categories based on the number of hydrogen bonds they employ. The A77 and A78 nucleotides form consecutive A-minor type II and type I interactions, which are the most common types of A-minor interactions due to their consistency with the native helicity of RNA. During our investigation of the active site, we observed that a significant number of replicates in almost every type of simulation showed partial or complete disruption of the type II A-minor interactions (Fig. 8c, middle panel). However, in almost all cases, the type I A-minor interactions remained intact over the course of our simulations, and reformed if ever disrupted (Fig. 8b, top panel). A typical reason for the rare disruption of the type I A-minor interactions was the presence of dynamic ions within close proximity in a highly perturbed active site. The overall greater stability of the type I *versus* type II interactions is consistent with the previously reported comparatively weaker nature of the latter.[65] Additionally, and also consistent with previous observations,[11] A77 is only able to form interactions with C19 due to the open nature of the active site (Fig. 8a, middle). We noticed that this attribute of the HDV ribozyme's type II A-minor interactions resulted in disruption of these hydrogen bonds whenever the active site was even minimally disturbed. By contrast, because A78 is more fully connected to both partners of the C18:G29 hydrogen bond, its type I interactions are more stable (Fig. 8a, top). They thus provide an upper boundary that encases most active site disorder and prevents it from spreading to other portions of the ribozyme during our accessible MD simulation times.

The P1.1 minihelix was not accounted for in initial pseudoknot predictions, and was first discovered upon resolution of the product HDV ribozyme crystal structure.[10] Mutational experiments proved that it is essential for optimal catalysis of the HDV ribozyme.[15] Because the top base pair of the P1.1 helix, C22G38, forms the junction of the coaxial stack with P1, we tested if disruptions at the scissile phosphate and N1N37 base pair resulted in further destabilization of the P1.1 helix. In all cases, the P1.1 helix was fully stable during our simulations regardless of G1U37 or G25U20 mutations (Fig. 8b, bottom). We thus conclude that, similar to the type 1 A-minor interactions above the active site, the P1.1 helix forms a lower bound that prevents active site disorder from spreading to the P4 helix below.

4 Conclusions

GU wobble pairs generate a distinct structural and chemical environment not observed with canonical Watson–Crick base pairs.[27,63] Their unique combination of ion- and water-binding functional groups, propagation of up- and downstream backbone distortion, and lack of reversible isostericity have made

 This journal is © the Owner Societies 2015

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

them key recurring players in functional RNAs across all kingdoms of life.[27,63] Two GU wobbles of the HDV ribozyme active site, one *cis* (standard) and one *trans* (reverse) wobble, have been shown to be crucial for optimal catalysis. Mutational studies of both wobbles have outlined their structural[31,33,66] and chemical[34,37,38,67] importance. Although MD simulations have proven effective in elucidating the crucial nature of the two GU wobbles for ion binding, no in-depth studies have been conducted to study the conformational dependence of active site dynamics on each of the bounding GU wobbles, providing for an opportunity to probe the structural context of their functional conservation within the tightly folded HDV ribozyme. Here we employ a total of >1 μs of MD simulations to study the sub-microsecond, small-scale conformational effects of each GU wobble in an activated HDV ribozyme active site. The resulting simulations indicate that, within such an activated conformation, each wobble plays a distinct set of structural roles to promote optimal catalysis. Our analyses show that, in addition to roles in ion binding, the *trans* G25U20 wobble acts to stabilize conformational dynamics both directly adjacent to and somewhat distal from the L3 loop. Our further mutations of the G1U37 wobble reveal that this motif acts more locally to stabilize, specifically, the general-acid hydrogen bond and in-line fitness. These effects are in contrast to the more distal influence of the G25U20 wobble; indeed, the *trans* wobble's propensity for chelated ion binding may be the mechanism by which it is able to effect comparatively longer-range conformational changes. Taken together, our results provide specific conformational functions for each type GU wobble in the HDV ribozyme active site. More broadly, our findings rationalize a number of experimental observations on the HDV ribozyme, provide a platform to understand the roles of non-canonical GU base pairs in the ever-expanding family of HDV-like ribozymes,[5,39,68] as well as in a plethora of highly structured RNAs where such wobbles are abundant and often functionally conserved.[27,28]

Acknowledgements

The authors would like to thank Dr Wendy Tay for use of her Wildtype_Na simulations and Simulaid protocols, as well as Mr Pavel Polakovič for use of two of his Wildtype_Mg simulations. The authors also gratefully acknowledge Dr Nad'a Špačková for protocols and simulation assistance. This work was supported by CEITEC – Central European Institute of Technology (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund (J.S. and K.R.). We also acknowledge Operational Program Research and Development for Innovations–European Regional Development Fund (Project CZ.1.05/2.1.00/03.0058), the Operational Program Education for Competitiveness–European Social Fund (CZ.1.07/2.3.00/20.0058) of the Ministry of Education, Youth and Sports of the Czech Republic (including travel support for N.G.W.), and grant P208/12/1878 from the Czech Science Foundation (J.S., M.O. and P.B.), as well as NIH grant R01 GM062357 (N.G.W.).

References

1 J. A. Doudna and J. R. Lorsch, *Nat. Struct. Mol. Biol.*, 2005, **12**, 395–402.
2 M. J. Fedor and J. R. Williamson, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 399–412.
3 M. A. Ditzler, E. A. Aleman, D. Rueda and N. G. Walter, *Biopolymers*, 2007, **87**, 302–316.
4 N. G. Walter and S. Perumal, *Springer Ser. Biophys.*, 2009, **13**, 103–127.
5 K. Salehi-Ashtiani, A. Luptak, A. Litovchick and J. W. Szostak, *Science*, 2006, **313**, 1788–1792.
6 C. H. T. Webb, N. J. Riccitelli, D. J. Ruminski and A. Luptak, *Science*, 2009, **326**, 953.
7 S. Nakano, D. M. Chadalavada and P. C. Bevilacqua, *Science*, 2000, **287**, 1493–1497.
8 A. T. Perrotta, I. Shih and M. D. Been, *Science*, 1999, **286**, 123–126.
9 M. D. Been, *Curr. Top. Microbiol. Immunol.*, 2006, **307**, 47–65.
10 A. R. Ferre-D'Amare, K. Zhou and J. A. Doudna, *Nature*, 1998, **395**, 567–574.
11 M. V. Krasovska, J. Sefcikova, N. Spackova, J. Sponer and N. G. Walter, *J. Mol. Biol.*, 2005, **351**, 731–748.
12 J. Sefcikova, M. V. Krasovska, J. Sponer and N. G. Walter, *Nucleic Acids Res.*, 2007, **35**, 1933–1946.
13 A. Ke, K. Zhou, F. Ding, J. H. Cate and J. A. Doudna, *Nature*, 2004, **429**, 201–205.
14 T. S. Wadkins, I. Shih, A. T. Perrotta and M. D. Been, *J. Mol. Biol.*, 2001, **305**, 1045–1055.
15 T. S. Wadkins, A. T. Perrotta, A. R. Ferre-D'Amare, J. A. Doudna and M. D. Been, *RNA*, 1999, **5**, 720–727.
16 J. H. Chen, R. Yajima, D. M. Chadalavada, E. Chase, P. C. Bevilacqua and B. L. Golden, *Biochemistry*, 2010, **49**, 6508–6518.
17 N. B. Leontis, J. Stombaugh and E. Westhof, *Nucleic Acids Res.*, 2002, **30**, 3497–3531.
18 D. A. Harris, R. A. Tinsley and N. G. Walter, *J. Mol. Biol.*, 2004, **341**, 389–403.
19 K. N. Sripathi, W. W. Tay, P. Banas, M. Otyepka, J. Sponer and N. G. Walter, *RNA*, 2014, **20**, 1112–1128.
20 D. A. Harris, D. Rueda and N. G. Walter, *Biochemistry*, 2002, **41**, 12051–12061.
21 M. J. Pereira, D. A. Harris, D. Rueda and N. G. Walter, *Biochemistry*, 2002, **41**, 730–740.
22 S. Jeong, J. Sefcikova, R. A. Tinsley, D. Rueda and N. G. Walter, *Biochemistry*, 2003, **42**, 7727–7740.
23 R. A. Tinsley, D. A. Harris and N. G. Walter, *J. Am. Chem. Soc.*, 2003, **125**, 13972–13973.
24 R. A. Tinsley, D. A. Harris and N. G. Walter, *Biochemistry*, 2004, **43**, 8935–8945.
25 M. E. Gondert, R. A. Tinsley, D. Rueda and N. G. Walter, *Biochemistry*, 2006, **45**, 7563–7573.
26 Y. Tanaka, M. Tagaya, T. Hori, T. Sakamoto, Y. Kurihara, M. Katahira and S. Uesugi, *Genes Cells*, 2002, **7**, 567–579.
27 G. Varani and W. H. McClain, *EMBO Rep.*, 2000, **1**, 18–23.
28 A. Mondragon, *Annu. Rev. Biophys.*, 2013, **42**, 537–557.

Published on 21 January 2015. Downloaded by University of Michigan Library on 10/02/2015 16:28:59.

29 A. Mokdad, M. V. Krasovska, J. Sponer and N. B. Leontis, *Nucleic Acids Res.*, 2006, **34**, 1326–1341.
30 A. L. Cerrone-Szakal, D. M. Chadalavada, B. L. Golden and P. C. Bevilacqua, *RNA*, 2008, **14**, 1746–1760.
31 J. Chen, A. Ganguly, Z. Miswan, S. Hammes-Schiffer, P. C. Bevilacqua and B. L. Golden, *Biochemistry*, 2013, **52**, 557–567.
32 J. H. Chen, B. Gong, P. C. Bevilacqua, P. R. Carey and B. L. Golden, *Biochemistry*, 2009, **48**, 1498–1507.
33 N. K. Tanner, S. Schaff, G. Thill, E. Petit-Koskas, A. M. Crain-Denoyelle and E. Westhof, *Curr. Biol.*, 1994, **4**, 488–498.
34 H. N. Wu, J. Y. Lee, H. W. Huang, Y. S. Huang and T. G. Hsueh, *Nucleic Acids Res.*, 1993, **21**, 4193–4199.
35 F. Nishikawa, H. Fauzi and S. Nishikawa, *Nucleic Acids Res.*, 1997, **25**, 1605–1610.
36 H. N. Wu, Y. J. Wang, C. F. Hung, H. J. Lee and M. M. Lai, *J. Mol. Biol.*, 1992, **223**, 233–245.
37 N. Veeraraghavan, A. Ganguly, J. H. Chen, P. C. Bevilacqua, S. Hammes-Schiffer and B. L. Golden, *Biochemistry*, 2011, **50**, 2672–2682.
38 N. Veeraraghavan, A. Ganguly, B. L. Golden, P. C. Bevilacqua and S. Hammes-Schiffer, *J. Phys. Chem. B*, 2011, **115**, 8346–8357.
39 C. H. Webb, N. J. Riccitelli, D. J. Ruminski and A. Luptak, *Science*, 2009, **326**, 953.
40 C. Guerriertakada, K. Gardiner, T. Marsh, N. Pace and S. Altman, *Cell*, 1983, **35**, 849–857.
41 G. A. Soukup and R. R. Breaker, *RNA*, 1999, **5**, 1308–1325.
42 M. A. Ditzler, J. Sponer and N. G. Walter, *RNA*, 2009, **15**, 560–575.
43 M. V. Krasovska, J. Sefcikova, K. Reblova, B. Schneider, N. G. Walter and J. Sponer, *Biophys. J.*, 2006, **91**, 626–638.
44 J. Sefcikova, M. V. Krasovska, N. Spackova, J. Sponer and N. G. Walter, *Biopolymers*, 2007, **85**, 392–406.
45 M. A. Ditzler, M. Otyepka, J. Sponer and N. G. Walter, *Acc. Chem. Res.*, 2010, **43**, 40–47.
46 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
47 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin and P. A. Kollman, *AMBER 10*, University of California, San Francisco, CA, 2008.
48 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 11*, University of California, San Franscisco, CA, 2010.
49 A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, 3rd, C. A. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.
50 M. Zgarbova, M. Otyepka, J. Sponer, A. Mladek, P. Banas, T. E. Cheatham, 3rd and P. Jurecka, *J. Chem. Theory Comput.*, 2011, **7**, 2886–2902.
51 R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. Cheatham, 3rd, S. Dixit, B. Jayaram, F. Lankas, C. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova and J. Sponer, *Nucleic Acids Res.*, 2010, **38**, 299–313.
52 D. L. Beveridge, T. E. Cheatham, 3rd and M. Mezei, *J. Biosci.*, 2012, **37**, 379–397.
53 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
54 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
55 M. Mezei, *J. Comput. Chem.*, 2010, **31**, 2658–2668.
56 A. Hernandez-Cid, S. Aguirre-Sampieri, A. Diaz-Vilchis and A. Torres-Larios, *IUBMB Life*, 2012, **64**, 521–528.
57 S. E. McDowell, N. Spackova, J. Sponer and N. G. Walter, *Biopolymers*, 2007, **85**, 169–184.
58 J. Sponer, P. Banas, P. Jurecka, M. Zgarbova, P. Kuhrova, M. Havrila, M. Krepl, P. Stadlbauer and M. Otyepka, *J. Phys. Chem. Lett.*, 2014, **5**, 1771–1782.
59 S. Nakano, A. L. Cerrone and P. C. Bevilacqua, *Biochemistry*, 2003, **42**, 2982–2994.
60 F. Nishikawa, J. Kawakami, A. Chiba, M. Shirai, P. K. R. Kumar and S. Nishikawa, *Eur. J. Biochem.*, 1996, **237**, 712–718.
61 M. D. Been and G. S. Wickham, *Eur. J. Biochem.*, 1997, **247**, 741–753.
62 J. Stombaugh, C. L. Zirbel, E. Westhof and N. B. Leontis, *Nucleic Acids Res.*, 2009, **37**, 2294–2312.
63 B. Masquida and E. Westhof, *RNA*, 2000, **6**, 9–15.
64 A. T. Perrotta and M. D. Been, *Nucleic Acids Res.*, 1996, **24**, 1314–1321.
65 P. Nissen, J. A. Ippolito, N. Ban, P. B. Moore and T. A. Steitz, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 4899–4903.
66 J. Kawakami, P. K. R. Kumar, Y. A. Suh, F. Nishikawa, K. Kawakami, K. Taira, E. Ohtsuka and S. Nishikawa, *Eur. J. Biochem.*, 1993, **217**, 29–36.
67 S. Nakano and P. C. Bevilacqua, *J. Am. Chem. Soc.*, 2001, **123**, 11333–11334.
68 C. H. Webb and A. Luptak, *RNA Biol.*, 2011, **8**, 719–727.

 This journal is © the Owner Societies 2015