research papers

A posteriori design of crystal contacts to improve the X-ray diffraction properties of a small RNA enzyme

Celeste MacElrevey,[a] Robert C. Spitale,[b] Jolanta Krucinska[a] and Joseph E. Wedekind[a,b]*

[a]Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642, USA, and [b]Department of Chemistry, Biological Chemistry Cluster, River Campus Box 270216, University of Rochester, Rochester, New York 14627-0216, USA

Correspondence e-mail: joseph.wedekind@rochester.edu

The hairpin ribozyme is a small catalytic RNA comprising two helix–loop–helix domains linked by a four-way helical junction (4WJ). In its most basic form, each domain can be formed independently and reconstituted without a 4WJ to yield an active enzyme. The production of such minimal junctionless hairpin ribozymes is achievable by chemical synthesis, which has allowed structures to be determined for numerous nucleotide variants. However, abasic and other destabilizing core modifications hinder crystallization. This investigation describes the use of a dangling 5′-U to form an intermolecular U·U mismatch, as well as the use of synthetic linkers to tether the loop A and B domains, including (i) a three-carbon propyl linker (C3L) and (ii) a nine-atom triethylene glycol linker (S9L). Both linker constructs demonstrated similar enzymatic activity, but S9L constructs yielded crystals that diffracted to 2.65 Å resolution or better. In contrast, C3L variants diffracted to 3.35 Å and exhibited a 15 Å expansion of the c axis. Crystal packing of the C3L construct showed a paucity of 6_1 contacts, which comprise numerous backbone to 2′-OH hydrogen bonds in junctionless and S9L complexes. Significantly, the crystal packing in minimal structures mimics stabilizing features observed in the 4WJ hairpin ribozyme structure. The results demonstrate how knowledge-based design can be used to improve diffraction and overcome otherwise destabilizing defects.

1. Introduction

The hairpin ribozyme is a small self-cleaving RNA derived from the negative strand of the 359 nt tobacco ringspot virus satellite RNA. *In vivo*, this ribozyme generates unit-length circular transcripts from concatenated replication intermediates produced by rolling-circle genome replication (Hampel & Tritz, 1989). As with other naturally occurring small ribozymes such as the hammerhead, hepatitis δ-virus, Varkud satellite and metabolite-sensing varieties, hairpin-ribozyme cleavage generates a free 5′-hydroxyl and a cyclic 2′,3′-phosphate as cleavage products. The reaction occurs without the need for a metal hydroxide and leads to site-specific phosphodiester-bond cleavage of the order of 0.1– 0.3 min^{-1} using minimal hinged constructs that lack a natural four-way helical (4WJ) junction (reviewed in Walter & Burke, 1998). The active site of the hairpin ribozyme forms at the confluence of two helix–loop–helix domains, called A and B, that dock together through interactions between the respective internal loops. Docking is stabilized by a cross-strand Watson–Crick base pair between G+1 and C25, as well as a 'ribose-zipper' motif (Rupert & Ferré-D'Amaré, 2001; Chowrira et al., 1993). In the wild-type sequence, the A and B domains are joined distally through a single adenine residue

and this connection is bolstered further by a 4WJ (reviewed in Hohng *et al.*, 2004). Although dispensable for the catalytic activity, the 4WJ enhances folding relative to minimally hinged constructs (Hampel & Tritz, 1989; Tan *et al.*, 2003).

The global fold and active-site architecture of the hairpin ribozyme were initially revealed as a 2.4 Å resolution crystal structure (Fig. 1*a*; Rupert & Ferré-D'Amaré, 2001). With the exception of the substrate strand, this 113-nucleotide (nt) construct was generated by *in vitro* transcription and included both the 4WJ motif that flanks helices 2 and 3, as well as a stem-loop sequence located at the end of helix 4. The latter stem-loop was enlarged and adapted to bind a well characterized RNA-recognition motif (RRM) derived from the U1A splicing factor (Oubridge *et al.*, 1994). This protein oligomerizes while bound to specific RNA sequences, thus creating a self-assembling crystallization platform (Ferré-D'Amaré *et al.*, 1998). In an alternative minimalist approach, an all-RNA construct of the hairpin ribozyme was produced by solid-phase chemical synthesis. The ease of manipulation of such minimal constructs offered two advantages. Firstly, this approach enabled well controlled high-throughput structural investigations designed to elucidate the roles of specific functional groups located in the enzyme active site (Alam *et al.*, 2005; Salter *et al.*, 2006). Secondly, *a posteriori* engineering of constructs enabled incremental improvements in X-ray diffraction resolution (Alam *et al.*, 2005; Salter *et al.*, 2006). Initial efforts were based on the observation that minimal

constructs, which included the A14 hinge residue but lacked the surrounding 4WJ motif, had a tendency to misfold *via* coaxial stacking of helix 3 on helix 2 (Esteban *et al.*, 1998). This problem could be overcome by reconstituting the fold from independent loop A and B domains *in trans* (*i.e.* a junctionless hairpin ribozyme; Butcher *et al.*, 1995) and led to the crystallization of a 64-mer construct that diffracted to 3.17 Å resolution (Grum-Tokars *et al.*, 2003). Empirical observations based on the 64-mer lattice led to the design of a 61 nt junctionless construct that diffracted to 2.05 Å resolution (Alam *et al.*, 2005; Salter *et al.*, 2006). This variant eliminated the A14 residue at the interdomain junction to yield a 5′-dangling U on the opposite strand. This residue was designed to engage in an intermolecular U·U mismatch and thereby promote pseudo-continuous helical stacking (Alam *et al.*, 2005; Salter *et al.*, 2006).

Junctionless ribozymes demonstrate a 10^4-fold increase in the apparent K_m compared with hinged variants that include A14. The reduced association of these domains has been attributed to an increase in the entropic penalty of docking as well as an enhanced dependence of domain association on the rate of diffusion (Butcher *et al.*, 1995). These qualities represent potential caveats in the crystallization of junctionless ribozymes that incorporate transition-state mimics or abasic residues, since these modifications have the potential to further reduce interdomain docking and substrate affinity because they require either (i) the introduction of breaks into
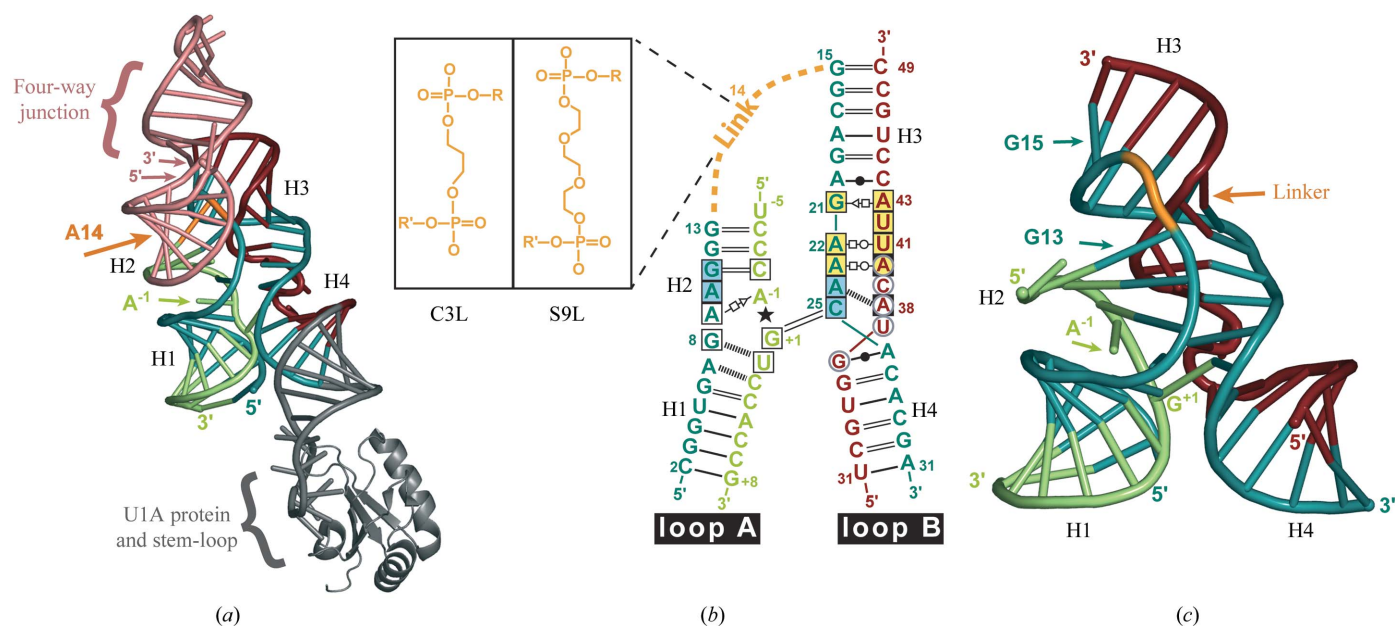

Figure 1
Schematic depictions of the global hairpin-ribozyme fold and the RNA sequences used in this investigation. (*a*) Ribbon diagram of the four-way helical junction (4WJ) hairpin ribozyme (PDB code 1m5k). The junction region, which is not present in minimal constructs, is shown in salmon. RNA and protein residues included for the purposes of forming the U1A crystallization platform are colored gray. Other strands are colored as described for (*b*). (*b*) Secondary structure of the minimal hinged hairpin ribozyme adapted from *RNAview* (Yang *et al.*, 2003). The substrate strand is depicted in green, the S-turn strand in red and the 29-mer strand in blue. The linker position is colored orange. The helix–loop–helix domains are labeled loop A and loop B; helices are labeled H1–H4. The inset displays two alternate synthetic linkages: C3L and S9L. The site of enzymatic cleavage is highlighted by a star. The A−1 residue is a 2′-deoxy A in the C3L structure and a 2′-deoxy-2′-amino U in the S9L structure. Conserved residues are boxed; blue backgrounds indicate residues of the ribose zipper, yellow boxes indicate E-loop residues and gray circled residues belong to the S-turn. Hydrogen-bond pairings: open-square, Hoogsteen; open triangle, *trans*-sugar; open circle, Watson–Crick face; closed circle, wobble pair. Double and single lines indicate Watson–Crick pairs; black dashed lines indicate single hydrogen bonds. (*c*) Ribbon diagram of the minimal hinged hairpin ribozyme solved in this investigation.

research papers

Table 1
Intensity and refinement statistics.

Values in parentheses are for the highest resolution shell.

Construct	61-mer junctionless (JL)†	61-mer position 14 C3 linker‡	61-mer position 14 S9 linker§
Intensity statistics			
Resolution range (Å)	2.05–30.0 (2.05–2.12)	3.35–34.7 (3.35–3.47)	2.65–38.9 (2.65–2.74)
Space group	$P6_122$	$P6_122$	$P6_122$
Unit-cell parameters (Å)	$a = 93.3$, $c = 131.3$	$a = 91.1$, $c = 145.0$	$a = 94.3$, $c = 131.1$
Solvent content (%)	80	81	80
No. of measurements	178418	61410	158609
No. of unique reflections	21669 (2142)	5453 (522)	10389 (1006)
Average redundancy	8.2 (4.6)	11.1 (12.2)	15.0 (16.1)
Completeness (%)	99.3 (99.1)	98.6 (99.1)	98.6 (99.6)
R_{sym}¶ (%)	3.0 (43.2)	10.6 (51.0)	5.9 (45.7)
$I/\sigma(I)$	24.0 (3.3)	16.2 (4.7)	25.5 (6.0)
Refinement statistics			
No. of RNA atoms	1295	1304	1308
No. of water atoms	87	3	14
No. of ions	2 $Co(NH_3)_6^{3+}$, 1 SO_4^{2-}	1 $Co(NH_3)_6^{3+}$	1 $Co(NH_3)_6^{3+}$
$R_{cryst}/R_{work}/R_{free}$†† (%)	24.8/25.2/26.8	26.1/26.6/27.3	19.9/19.5/22.2
R.m.s. deviations from ideality			
Bond lengths (Å)	0.006	0.008	0.005
Bond angles (°)	1.4	1.4	1.3
Temperature factors			
B factor, RNA (Å^2)	69.5	95.6	78.5
B factor, water (Å^2)	74.0	78.4	77.0

† From PDB entry 2oue. ‡ From PDB entry 2npz. § From PDB entry 2npy. ¶ $R_{sym} = [\sum |I(h)_j - \langle I(h)\rangle|/\sum I(h)_j] \times 100$, where $I(h)_j$ is the observed intensity of the jth measurement of reflection h and $\langle I(h)\rangle$ is the mean intensity of reflection h. †† $R = [\sum ||F_o| - k|F_c||/\sum |F_o|] \times 100$, where F_o and F_c are the observed and calculated structure-factor amplitudes and k is a scale factor. R_{cryst} is calculated using the residual target in *CNS* with all reflections in the resolution range of refinement. R_{work} is calculated using all reflections except those randomly removed for the test set (5–7%); R_{free} is calculated using the test set of reflections.

the RNA backbone (Torelli *et al.*, 2007; Rupert *et al.*, 2002) or (ii) the potential to form cavities within the ribozyme core (Kuzmin *et al.*, 2005). In this study, we describe strategies to improve the interdomain-docking properties of minimal all-RNA hairpin-ribozyme crystallization constructs. Our goal is to produce variants that are more resilient to the destabilizing modifications that are central to our structure–function studies. Strategies include the use of 5′-U·U overhanging mismatches and incorporation of flexible interdomain linkers (Fig. 1b), the effects of which on crystal packing cannot be readily established with solution-activity assays. To assess the outcome of this approach, four hairpin-ribozyme crystal structures were compared, including the 4WJ structure (PDB code 1m5k), the refined junctionless hairpin ribozyme (PDB code 2oue) and two new structures harbouring a linker at position 14 that included an S9 linker (Fig. 1c; PDB code 2npy) or a shorter C3 linker (PDB code 2npz). The results provide a set of knowledge-based principles that can be applied generally to RNA constructs to optimize X-ray diffraction for structure–function studies.

2. Materials and methods

2.1. Minimal junctioned all-RNA hairpin-ribozyme construct

The 61 nt minimal junctionless (JL) all-RNA hairpin-ribozyme construct has been described previously (Alam *et al.*, 2005). The sequence of the ribozyme used in the current study

(Fig. 1b) included the gain-of-function mutation U39C to prevent conformational heterogeneity in the S-turn. Position A14 was substituted with either of two commercially available synthetic linkers (Fig. 1b) that permitted the preservation of the standard phosphodiester bond. In the C3L residue, a three-carbon propyl linker was included as a flexible mimic of the ribose backbone. The S9L residue contained three ethylene glycol subunits (nine atoms), which added five more atoms to the backbone compared with a standard nucleotide. Cleavage activity during crystallization was inhibited by modification of the 2′-OH nucleophile of A−1 in the substrate strand: a 2′-deoxy A residue was utilized in the C3L structure, whereas a 2′-deoxy-2′-amino U (U2N) group was incorporated into the S9L structure. The stem-loop sequence present at the end of helix 4 in the natural sequence was removed to promote blunt-ended crystal packing as in other minimal hairpin-ribozyme constructs. Strand sequences outside the conserved core (Fig. 1b) were optimized to prevent self-complementarity and the GC content of helical ends was enriched to minimize helical end fraying. All strands were synthesized by Dharmacon Inc. (Lafayette, CO, USA) with subsequent deprotection and HPLC purification performed at home as described by Wedekind & McKay (2000). The three strands comprising the substrate and ribozyme were mixed together in 10 m*M* sodium cacodylate pH 6.0; these included a 13-mer substrate, a 19-mer strand comprising the S-turn and a 29-mer harboring either the C3L or S9L linker. The final ribozyme concentration was 0.3 m*M*. Docking was promoted by addition of $Co(NH_3)_6Cl_3$ from a 100 m*M* stock solution, which was slowly added to the RNA with mixing until reaching a final metal concentration of 1 m*M*.

2.2. Crystallization and X-ray diffraction experiments

Crystals grew as hexagonal rods in hanging-drop vapor-diffusion experiments conducted at 293 K containing the following components for the S9L and C3L constructs, respectively: 20%(*w/w*) PEG 2K MME, 0.1 *M* sodium cacodylate pH 6.4 or 6.2, 0.25 *M* Li$_2$SO$_4$, 2 m*M* spermidine–HCl and 1 m*M* $Co(NH_3)_6Cl_3$. All crystals were cryoprotected by 3 min serial transfers through four synthetic mother liquors containing 5–18%(*v/v*) glycerol. Crystals were captured in thin nylon loops (Hampton Research) and flash-cooled in a 100 K stream of nitrogen gas (X-stream, Rigaku/MSC). Diffraction data for the C3L structure were recorded at home on an R-AXIS IV image-plate system equipped with confocal optics

(Osmic). Cu *Kα* X-rays were generated using a Rigaku RUH2R rotating-anode generator operated at 4.5 kW and equipped with a 0.3 mm focal cup. 220 images were collected at a crystal-to-detector distance of 12.5 cm with an exposure time of 25 min per 0.5° oscillation. Diffraction data were collected from the S9L crystal at 100 K on a Quantum210 CCD (ADSC) at the A1 station of the Cornell High Energy Synchrotron Source (CHESS) using X-rays of wavelength 0.977 Å. A total of 300 images were collected at a crystal-to-detector-distance of 24 cm using an exposure of 30 s per 0.5°. All data were reduced and scaled using *CrystalClear* (Pflugrath, 1999). Intensity and data-reduction statistics are provided in Table 1.

2.3. Structure determination and refinement

Each hairpin-ribozyme construct in this study crystallized in space group $P6_122$ with a single 61-mer per asymmetric unit. Phases were derived through difference Fourier analysis starting from the highest resolution JL structure available, *i.e.* PDB entry 1zfr. Refinement employed conventional methods for RNA as implemented in *CNS* and *O* (Brünger *et al.*, 1998; Jones *et al.*, 1991). Rigid-body refinement was first performed on the entire structure and subsequently on the individual loop A and B domains, followed by simulated annealing, positional minimization and individual *B*-factor refinement (Alam *et al.*, 2005). *B* factors were set to refine between values of 1 and 400 Å^2, as described for other refined RNA structures

(Salter *et al.*, 2006). The anisotropic bulk-solvent correction was applied during both refinements; however, in the C3L structure solvent density (k) and *B* factors (B) were manually adjusted as necessary. Model building employed reduced-bias σ_A coefficients throughout the refinement process (Pannu & Read, 1996). Parameter and topology files for the C3 linker were adapted from existing ribonucleotide files available in *CNS* v.1.0. The S9L files were developed with *XPLO2D* (Kleywegt, 1995) based on a compilation of atomic resolution phosphate and ethylene glycol molecules obtained from the Cambridge Structural Database (Kleywegt, 1995; Kraut, 1961; Britton & Chantooni, 2001). Waters were assigned to structures based on the following criteria: >1σ electron density in $2mF_o - DF_c$ maps, reasonable geometry, hydrogen-bond distances between donor and acceptor groups of 2.6–3.7 Å and refined temperature factors comparable to surrounding RNA atoms. One Co(NH$_3$)$_6$ molecule, located near G21 and A40 of the S-turn, has been consistently observed in all minimal hairpin-ribozyme crystal structures; a second previously characterized site within the major groove of H2 has been modeled here as a water in the C3L and S9L structures owing to low occupancy. Sugar-pucker restraints were derived from the 2.05 Å resolution 1zfr structure where appropriate. Refinement statistics are provided in Table 1. To verify the orientation of U−5 and linker residues, simulated-annealing omit maps with coefficients ($mF_o - DF_c$) were generated in *CNS* (Figs. 2c, 2d, 4a and 4b). Superpositions were generated from conserved residues using *LSQKAB* as implemented in
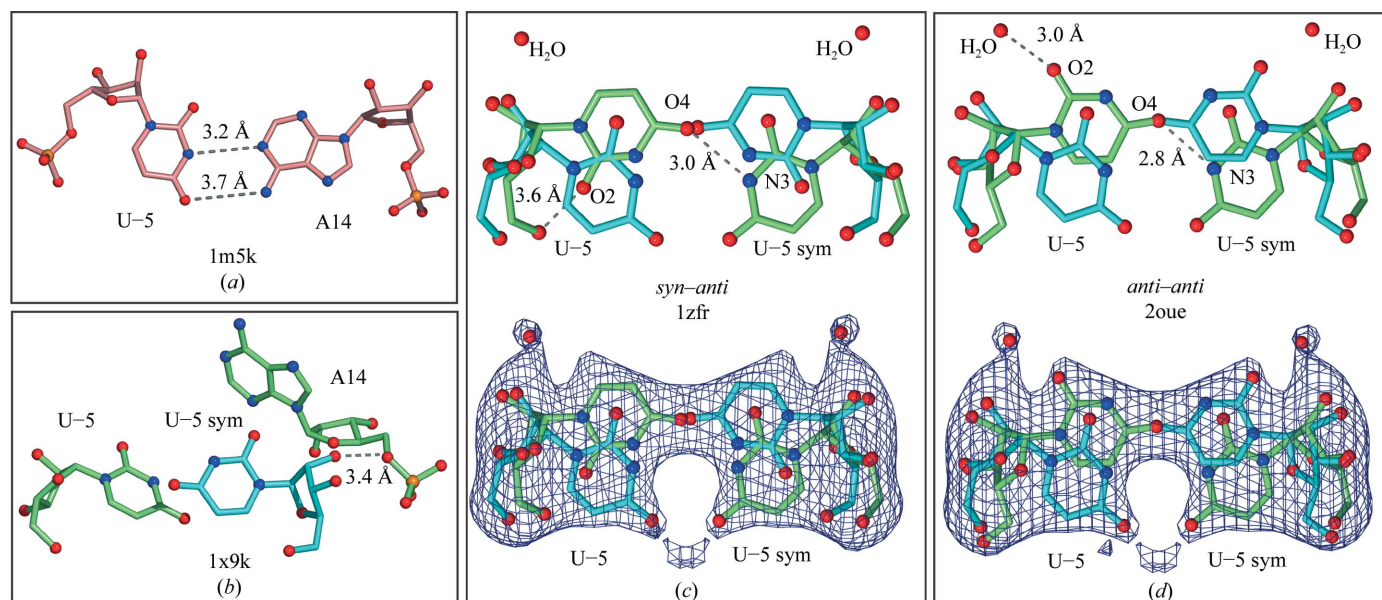

Figure 2
Ball-and-stick and electron-density maps of the engineered U·U mismatch at the top of helix 2 (H2). (*a*) The U−5·A14 base pair as observed in one of two molecules in the asymmetric unit of the 4WJ structure (PDB code 1m5k). O atoms are colored red and N atoms are blue. Dashed gray lines indicate putative hydrogen bonds with corresponding distances. (*b*) The minimal junctionless ribozyme 64-mer structure (PDB code 1x9k) with an orientation equivalent to that in (*a*). Two symmetry-related U−5 residues are separated by a 3 Å translation and are not base-paired. (*c*) The original *syn–anti* base-pairing conformation of U−5 modeled for the 2.05 Å resolution junctionless ribozyme (PDB code 1zfr). Alternate conformations were modeled with equal occupancy and are colored green or cyan. A single set of hydrogen bonds is indicated; the symmetry-related hydrogen bonds and labels were omitted for clarity. A σ_A-weighted simulated-annealing omit electron-density map calculated for the U−5 residue and nearby solvent molecules is shown contoured at 3.5σ with coefficients $mF_o - DF_c$. (*d*) The revised 2.05 Å resolution minimal hairpin-ribozyme structure of this study depicting the new *anti–anti* or 'Calcutta' model for U−5 fitted into an omit map as described in (*c*).

research papers

CCP4 (Kabsch, 1976; Collaborative Computational Project, Number 4, 1994). Solvent-accessible surface areas were calculated using GRASP (Nicholls et al., 1991). Helical parameters were analyzed using CURVES and are based on the best curvilinear fit of the helical axis (Lavery & Sklenar, 1989). All figures were generated in PyMOL, supplemented by Nuccyl in Fig. 1 (DeLano, 2004; Jovine, 2003; Yang et al., 2003). Coordinates were deposited in the PBD under accession codes 2npy (S9L), 2npz (C3L) and 2oue (the redeposited 2.05 Å 1zfr JL structure).

2.4. Activity assays

Cleavage reactions for the hairpin ribozyme were conducted at 298 K under single-turnover conditions in order to compare the relative activities of the various minimal crystallization constructs of this study. The respective loop A and B strands of the JL ribozyme or the three strands of the hinged S9L or C3L constructs (Fig. 1b) were combined in a reaction buffer comprising 0.10 M Tris–HCl pH 7.5 and 0.20 M NaCl. To assist with pre-reaction folding, the respective loop A and B domains of the JL construct were heated to 343 K for 3 min and cooled to 298 K over 15 min. The JL hairpin-ribozyme reaction was initiated upon addition of 50 mM MgCl$_2$ (Butcher et al., 1995); the hinged-ribozyme reactions were initiated by addition of 12 mM MgCl$_2$ (Hampel & Tritz, 1989; Chowrira et al., 1993). The starting concentration of ribozyme strand for each assay was 200 µM for all constructs, whereas the concentration of substrate in each reaction was 2.0 µM. 10 µl aliquots were removed from each reaction at time points spanning 4 h and were quenched by the addition of two volume equivalents of denaturing sample buffer comprising 7 M urea at pH 8.0 without tracking dye. Time points were measured in duplicate from independent assays and the variance between duplicates was <10%. The cleavage products were heated to 370 K for 1 min and then cooled rapidly on ice before separation on 15% polyacrylamide gels containing 7 M urea (Sambrook et al., 1989). The 13-mer substrate strand was pre-labeled with Cy5 at the 5′-end and handled under light-restricted conditions. Substrate and product (8-mer) were detected directly within the gel by use of a Storm 860 imaging system (GE Healthcare Inc.) operated in red fluorescent mode (662 nm). The fraction of substrate cleaved relative to input substrate was quantified using ImageQuant software (Molecular Dynamics). Experimental data from time-dependent cleavage assays were fitted to a double-exponential equation (described by Esteban et al., 1997, 1998; Rueda et al., 2004),

$$F(t) = A_0 + A_1[1 - \exp(-k_1 t)] + A_2[1 - \exp(-k_2 t)],$$

where A_1 and A_2 represent amplitudes of the biphasic time course, k_1 and k_2 are the corresponding first-order rate constants of the fast and slow phases, respectively, and t is time. A_0 represents the initial amount cleaved, which was about 3%. Amplitudes and rate constants were estimated by the Marquardt–Levenberg nonlinear least-squares regression routine (SigmaPlot 9.0). The standard error for fitted parameters was <15%.

3. Results and discussion

3.1. Significance of and modeling a U·U mismatch

A common strategy to improve the X-ray diffraction properties of an RNA crystal is to alter the construct itself, rather than subject it to numerous screening conditions (Anderson et al., 1996). The initial minimal JL hairpin-ribozyme 64-mer was designed without the benefit of a known structure and included residue A14 at the end of helix H2 (Fig. 1b). This seemed sensible even after the 4WJ hairpin-ribozyme structure had been solved since an A14·U−5 Watson–Crick pair was apparent in this structure (Fig. 2a). Nonetheless, the diffraction of 64-mer crystals was limited to a nominal 3.17 Å resolution (Alam et al., 2005) and, upon solving the structure (PDB code 1x9k), it was apparent that the lattice packing environment necessitated the extrusion of A14 from helix H2 in deference to a symmetry-related U−5 (Fig. 2b). It was therefore hypothesized that the two U−5 residues, vertically separated by 3 Å along the c axis, might base-pair across a strict crystallographic dyad if A14 were removed, thereby promoting pseudo-continuous helical packing.

A thermodynamic consideration of the U·U pairing possibility directed us to the work of Turner and colleagues. Their results revealed that while a dangling 3′-U in the context of a helix improved stability by ∼4 kJ mol^{-1}, the 5′-variant that we used (Fig. 1b) would contribute a negligible 0.4 kJ mol^{-1} (Sugimoto et al., 1987). This result suggests that any observed advantages in crystallization would not be the result of increased stability of the folded ribozyme in solution. In contrast, the 5′-overhang optimizes the free energy gained through formation of a pseudo-continuous helix. The base-stacking interaction at the helix–helix interface of a sticky-ended coaxial stack has been reported to contribute ∼4 kJ mol^{-1} more favorability to the free energy than if all nucleotides were connected by a phosphodiester linkage (Walter et al., 1994; Walter & Turner, 1994). The rationale for this observation is that a missing linkage in an overhanging base stack allows the nucleotides more freedom to maximize intermolecular interactions and reduce charge repulsion. This provides one explanation why pseudo-continuous helices predominate in the crystal packing of nucleic acids, particularly since such interfaces involve two missing backbone linkages. Lastly, the sequence context of the 5′-U·U mismatch (i.e. flanked by 5′-G and 3′-C; Fig. 1b) offered the greatest favorable free energy of internal loop formation among those sequences characterized by Kierzek et al. (1999). The latter study also suggested that the U·U mismatch is more favorable in regions where there are fewer helical constraints on its shape, as demonstrated by observed increases in stability when placed near a helix end (Kierzek et al., 1999).

Pairings of U·U mismatches have been reported in several conformations that must be considered for modeling into electron-density maps. U·U mismatches are most commonly positioned such that the Watson–Crick faces orient towards each other to engage in a single imino-to-keto hydrogen bond, according to a survey of structures reported in the Non-

canonical Base-Pair Database (Nagaswamy *et al.*, 2002). This configuration could readily be fitted into our electron-density maps with two stipulations. Firstly, the requirement for strict twofold symmetry for this interaction assumed static conformational disorder, as supported by the broadened electron-density features (Fig. 2c). Secondly, an alternate *syn* conformation modeled for U−5 was necessary to achieve opposing Watson–Crick faces and maximization of hydrogen bonding, although this configuration is rarely observed (Burkard *et al.*, 1999). Nonetheless, a similar *syn–anti* orientation was reported in the 30S *Thermus thermophilus* ribosomal subunit (Selmer *et al.*, 2006). Care was taken to avoid an O2 to O5′ clash when modelling the *syn* conformation and the resulting model exhibited both reasonable stereochemistry and a good fit to the electron density (Fig. 2c).

The *syn–anti* conformer model for U−5 required that uridine simultaneously occupy both *syn* and *anti* conformations with equal occupancy ($q = 0.5$) owing to its location near a crystallographic twofold axis of symmetry, which complicated the interpretation of electron-density maps (Fig. 2c). However, NMR studies and molecular-dynamics simulations each suggested that uridine seldom assumes the *syn* orientation (Neumann *et al.*, 1980; Foloppe & Nilsson, 2005), leading us to re-evaluate our original published model (PDB code 1zfr). As such, we explored an alternative model in which both bases adopt the *anti* orientation, as proposed in two structures of superior resolution with PDB codes 1osu and 413d (Wahl *et al.*, 1996; Tanaka *et al.*, 1999). This conformation, termed the 'Calcutta' base-pair by Sundaralingam, exhibited one hydrogen bond between N3 and O4 and one between O4 and C5 that presumably involved the hydrogen of the latter. This model provided a slightly improved fit to the electron density of the current study, although a 4.2 Å distance between O4 and C5 ruled out a second hydrogen bond in the JL structure

(Fig. 2d). Parallel refinements of both possible U·U pairs were conducted (Figs. 2c and 2d) in which R values were compared after identical rounds of positional and individual B-factor minimization (holding weighting factors constant in *CNS*). The Calcutta conformation yielded R_{cryst} and R_{free} values that were 0.13% and 0.08% lower than the *syn–anti* combination, respectively. This observation and the paucity of known *syn–anti* conformers suggested that the Calcutta conformation is likely to be a more accurate representation of the data. As such, new coordinates of the JL hairpin ribozyme at 2.05 Å resolution (Table 1) were redeposited in the PDB with code 2oue and the hinged structures of this study were also refined with the Calcutta conformation. This result emphasizes that crystallographic contacts and alternative conformations must be modeled carefully and provides an important precedent for the use of databases to assist in the assignment of unusual base-pairing conformations during RNA model building.

3.2. Design and crystallization of synthetically hinged hairpin ribozymes

The limitations of our JL hairpin-ribozyme construct became apparent during attempts to crystallize the hairpin ribozyme with transition-state analogues or abasic substitutions in the active site. In attempts to capture vanadium oxide as a transition-state mimic between A−1 and G+1 (Fig. 1b), it was necessary to introduce a break into the cleavage site of the substrate strand (Torelli *et al.*, 2007). As such, the four-stranded JL construct became five strands plus a vanadium ligand. This six-component complex failed to crystallize, presumably owing to improper folding and interdomain docking. Thus, reconnecting the loop A and B domains was a logical step towards reducing the number of RNA strands. This effort proceeded with the stipulation that constructs incorporating the interdomain connection should remain small enough for efficient solid-phase chemical synthesis and be amenable to crystallization (and high-quality X-ray diffraction) in the lattice obtained previously.

The native A14 hinge residue found in the 4WJ structure is base-paired to U−5 and further stabilized by the surrounding 4WJ (Rupert & Ferré-D'Amaré, 2001). However, the intermolecular U·U mismatch formed in the JL crystal-packing scheme precludes re-introducing a base at the hinge position. Additionally, natural nucleotide linkers adjoining H3 and H2 were observed to produce coaxial helical stacks in other minimal constructs (Esteban *et al.*, 1998). Taken together with the observation that lattice formation by RNA favors pseudo-continuous helical packing (Brünger *et al.*, 1998), the presence of a natural nucleotide linker during crystallization was deemed likely to exacerbate improper docking between the loop A and loop B domains. It was therefore hypothesized that tethering the loop A and B domains with flexible synthetic linkers (or spacers) would introduce sufficient conformational freedom at the helical interface to dissuade unfavorable H2-to-H3 end-to-end stacking with minimal disruption to the existing crystal lattice. Similar modifications using non-nucleotide linkers were employed in single-
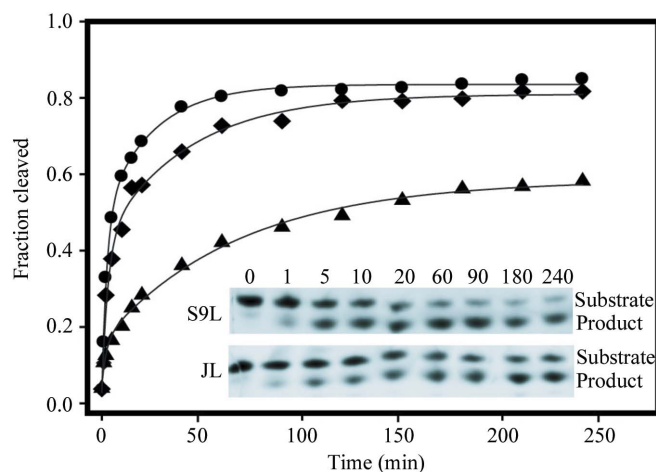

Figure 3
Kinetic analysis of hairpin-ribozyme constructs used in this study. Substrate-cleavage reactions were carried out under single-turnover conditions. For each construct (circles, S9-linked; diamonds, C3-linked; triangles, junctionless, JL) time courses were fitted to a double-exponential kinetic equation (see §2.4). Inset: representative gel images of the cleavage time course for the S9 and JL constructs; similar results were observed for the C3 construct (data not shown). Bands represent 5′-Cy5-labeled 13-mer substrate and 8-mer product strands.

molecule fluorescence resonance energy transfer (FRET) studies. Specifically, when a C3L spacer (Fig. 1b) was incorporated between A14 and A15 of a minimal hinged hairpin ribozyme, a 35-fold improvement in the interdomain docking rate constant was observed relative to the wild-type sequence (Rueda et al., 2004). The apparent favorability of this substitution was tempered by increased docking heterogeneity and a 25-fold slower bond-breaking step (i.e. k_{cleave}). However, this construct still exhibited overall cleavage rates that were nearly double that of the A14 linkage alone. In light of these data, we chose to substitute A14 with the S9L spacer (Fig. 1b), yielding a ribozyme with the same number of backbone atoms as that used in the FRET investigation but with no adenine base present to interfere with the U·U mismatch. Similarly, we mimicked the native linkage with the C3L spacer, a substitution that preserved the native backbone in atom number and type but offered increased flexibility since it lacked the furanose ring.

Cleavage assays were conducted on the respective crystallization constructs to compare relative activities for our S9L, C3L and JL constructs prior to crystallization. The hinged constructs exhibited similar cleavage profiles (Fig. 3) that fitted best to a biphasic double-exponential equation (Esteban et al., 1997, 1998; Rueda et al., 2004). S9- and C3-linked constructs exhibited total amplitudes ($A_1 + A_2$) of 81% (46.1 + 34.6%) and 78% (39.3 + 38.7%) for cleavage of a 13-mer substrate and showed similar first-order rate constants for the fast phase, where $k_1 = 0.390 \pm 0.055$ min^{-1} and $k_1 = 0.325 \pm 0.080$ min^{-1}, respectively. The slow phase, described by k_2, displayed rates of 0.041 ± 0.007 and 0.024 ± 0.005 min^{-1}, respectively. For the purposes of this study, we believe these values are indiscernible. Likewise, the JL construct showed a fast-phase first-order rate constant, k_1, of 0.429 ± 0.150 min^{-1} and a slow phase with $k_2 = 0.014 \pm 0.001$ min^{-1}. The $t_{1/2}$ for the fast phase of each construct was ~2 min, whereas the slow phases displayed $t_{1/2}$ values of 17 min (S9L), 29 min (C3L) and 50 min (JL), respectively. Furthermore, the JL construct cleaved only 54% (A_1, 11.4% + A_2, 42.3%) of the input substrate. The results suggested that each ribozyme construct could adopt a catalytically competent fold, as demonstrated by the comparable rate constants in the fast phase. However, greater populations of hinged hairpin ribozymes docked productively relative to the JL construct based on the total fraction that was cleaved ($A_1 + A_2$). Significantly, the conformation needed for solution activity appeared to be equally accessible by both the C3L and S9L variants, which has implications for their use in the development of hinged constructs for crystallization.

Despite the kinetic similarities between the hinged ribozymes, constructs harboring the S9L spacer exhibited superior crystal growth relative to C3L constructs. The latter crystals displayed multiple morphologies, with the best samples exhibiting a hexagonal habit that reached dimensions of 0.2 × 0.2 × 0.15 mm. These crystals were harvested after three weeks and diffracted X-rays to a maximum resolution of 3.35 Å. Their diffraction was not significantly improved by synchrotron radiation. S9L crystals grew much more consistently in a hexagonal habit that reached dimensions of up to 0.4 × 0.3 × 0.6 mm within 2–3 weeks. The initial S9L crystal used for this investigation diffracted to 2.65 Å resolution, although comparable examples have since been refined to resolutions as high as 2.05 Å (Torelli et al., 2007). Both S9L and C3L constructs crystallized in space group $P6_122$, although the C3L unit cell was longer by 15 Å along the c axis, which translates into a difference of 2.5 Å per asymmetric unit. Crystals of the minimal JL 61-mer hairpin ribozyme have been described elsewhere (Alam et al., 2005; Salter et al., 2006). In general, these crystals require higher concentrations of RNA for growth and reach maximum dimensions of 0.25 × 0.25 × 0.35 mm in 1–3 months.

3.3. Structural comparison of C3L and S9L hinged hairpin ribozymes

3.3.1. Overall model quality. The quality of the models is indicated by the observation that both linker structures fit well to electron-density maps and refined with reasonable geometric parameters (Table 1; Figs. 4a and 4b). The electron density of each structure was continuous throughout the molecule and there were no breaks in the RNA backbone. However, the C3L structure was inferior in several respects. Most notably, the resolution of X-ray diffraction and the agreement of reflections measured in multiplicity were poorer than for the JL and S9L structures (Table 1). The 3.35 Å resolution C3L structure (PDB code 2npz) refined to an R_{cryst} of 26% and an R_{free} of 27%, compared with the 2.65 Å resolution S9L structure (PDB code 2npy), which refined to an R_{cryst} of 20% and an R_{free} of 22%, or the newly refined 2.05 Å resolution JL structure (PDB code 2oue), which produced an R_{cryst} of 24.8% and an R_{free} of 26.8%. Coordinate errors estimated from cross-validated σ_A-weighting were 0.67 and 0.49 Å for the C3L and S9L structures, respectively. Electron-density maps were contiguous for all regions of each structure, with the exception of the C3L residue itself in which the central C atom could not be observed above the 3σ contour level in an $mF_o - DF_c$ omit electron-density map (Figs. 4a and 4b). Taking this into account, the conformation of the C3 linker was based largely on the positions of the flanking P atoms and bond-angle geometries. The average B factor for the C3L structure (95.6 Å^2) was higher than that for the S9L structure (78.5 Å^2), although both of these values were lower than those observed for the initial 3.19 Å resolution 1x9k structure (161 Å^2). Tertiarily folded RNA structures commonly exhibit higher average B factors compared with comparably sized proteins. As an example, the average value reported for the 2.4 Å resolution 4WJ structure, PDB code 1m5k, was 87 Å^2; however, the average B factor for the RNA atoms in this structure was 99 Å^2, whereas the average for the protein atoms was only 50 Å^2.

Helix H1 was the region of each hinged structure with the poorest quality electron density. This trend has also been documented in all crystal structures of the minimal JL hairpin ribozyme (Alam et al., 2005; Salter et al., 2006) and thus is not a consequence of the linkers. Increasing the GC content in this
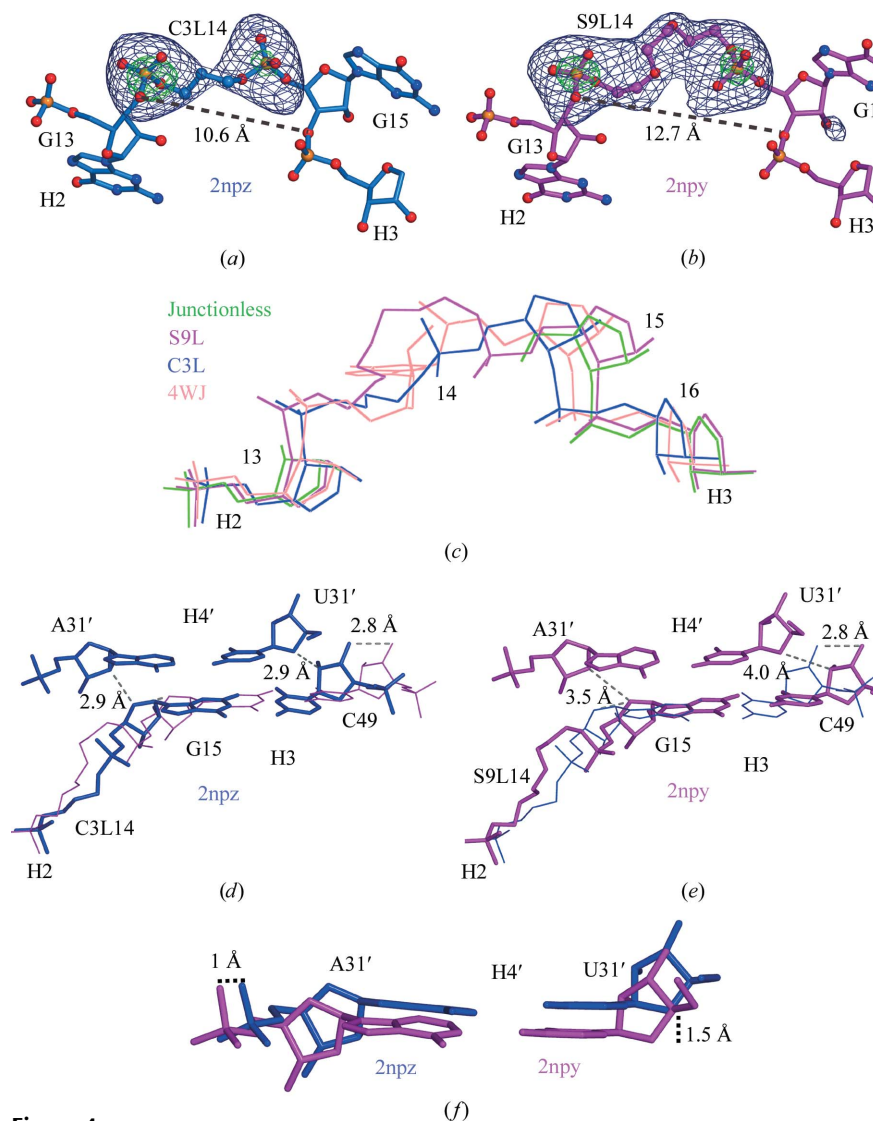
Figure 4
Comparison of the local structural effects of the synthetic C3L and S9L residues observed in the hinge region between the loop A and B domains. (*a*) σ_A-Weighted simulated-annealing $mF_o - DF_c$ omit electron-density map calculated for the C3L residue and flanking G15 phosphate. Blue density is contoured at 3σ and green density at 9σ. The distance between atom O3′ of residues 13 and 15 is shown as a dashed line. (*b*) Omit map for the S9L residue as described in (*a*); blue density is contoured at 3.5σ and green density at 15σ. (*c*) Linker and neighboring ribose positions as observed in the superposition of four alternatively hinged hairpin-ribozyme structures. Note the trend in the ribose positions of residue 13 compared with 16. The C3L structure (blue) most closely mimics the natural A14 linkage of the 4WJ structure (salmon), whereas the S9L structure (magenta) most closely mimics the junctionless structure (green). Bases are omitted for clarity, with the exception of A14. (*d*) Structural influence of the C3L structure on the terminus of H3. The C3L structure is depicted by blue sticks and the S9L structure by magenta lines. The lateral shift of the H3 terminal base pair is represented by the 2.8 Å distance between the O3′ atoms of residue C49 in the respective C3L and S9L structures. The 2_1 pseudo-helical base-stacking interaction with H4′ is shown to demonstrate the more flush base-stacking interaction exhibited by the C3L structure compared with S9L. The 2.9 Å distance between the O4′ and O3′ atoms of ribose moieties engaged in crystal packing in the C3L structure is compared with the equivalent distances of 3.5 and 4.0 Å displayed for the S9L structure in (*e*). (*e*) The equivalent region in the S9L structure, as described in (*d*). S9L residues are shown as magenta sticks and C3L residues as blue lines. Equivalent O4′–O3′ distances between the H3 and H4′ helices are longer in the S9L structure, which is representative of a more staggered base-stacking interaction. (*f*) A superposition of the terminal H4 base pair of the C3L and S9L structures demonstrates the influence of the shortened linker on pseudo-helical base stacking. The more flush base stack in the C3L structure is necessarily accompanied by a 1.5 Å upward shift of U31′ to avoid a steric clash with the symmetry-related C49 residue. Note that even after this movement U31′ and C49 are still closer in the C3L structure (*d* and *e*).

region minimized the helical fraying observed in the initial 1x9k 64-mer structure, but the temperature factors for this region remained high. The most likely reason for this is the scarcity of crystal-packing contacts in this region. This is evident by an examination of helix H1 of the 4WJ structure, which does not exhibit the same difficulties. Although it is engaged in a similar end-to-end base-stacking crystal contact as the minimal construct, helix H1 of the 4WJ structure is stabilized further by a modest number of packing interactions along the exposed length of the helix (Rupert *et al.*, 2002).

An all-atom superposition of the C3L and S9L hinged structures revealed an overall r.m.s.d. of 0.77 Å. The A and B domains exhibited similar molecular dimensions and their active-site residues overlaid well. Both structures displayed nearly the same 'pre-catalytic' active-site conformation described previously for the minimal JL 61-mer (Salter *et al.*, 2006). Deviations between the two structures were more pronounced at the helical ends, which are areas where the minimal hairpin ribozyme engages in few tertiary contacts, thus rendering them more susceptible to crystal-packing forces. The H2 stem of the S9L structure forms an intermolecular U·U mismatch (Figs. 1*b* and 2*d*) that is stabilized by a 2.7 Å hydrogen bond between N3 and O2 and a 3.3 Å hydrogen bond between O4 and C5 as reported in similar U·U pairs (Wahl *et al.*, 1996; Tanaka *et al.*, 1999). The equivalent distances for the C3L model are longer (3.3 and 3.7 Å), but are subject to greater coordinate error (Table 1). In light of the 4.2 Å distance exhibited by the JL structure for the second hydrogen bond (discussed above), it seems plausible that steric effects attributable to the linker favorably influence the spatial proximity of the symmetry-related U−5, leading to the improved hydrogen-bonding distances observed here.

3.3.2. Contrasting features of two different interdomain hinges. A comparison of the distance between the O3′ atoms of G13 and G15, which flank the interdomain junction, reveals significant differences between the two constructs. As anticipated, this distance was longer for the S9 linker (12.7 Å) relative to the C3 linker (10.6 Å; Figs. 4*a* and 4*b*) owing to the lengths of the respective linkers. An all-atom superposition of these hinged ribo-

zymes suggested that structural differences arising from variable linker lengths were not distributed equally along the lengths of the newly tethered helices (H2 and H3; Fig. 4c, note the ribose superposition of residue 13 versus residue 16). Rather, structural disparities localized mostly to the top of H3, as observed by the 2.8 Å displacement between the G15–C49 base pairs of each hinged structure (Figs. 4d and 4e). The basis of this localized change can be explained by the observation that the terminal base pairs of H2 and H3 are oriented nearly perpendicular to each other (Fig. 1c). G13 is buried further into the core of the loop B domain, whereas G15 is more solvent-exposed. Thus, G15 is better suited to adapt to the structural influences of the linker. The 2.8 Å movement of the terminal H3 base pair in the C3L structure toward the center of mass of the hairpin ribozyme increased the local twist of H3 to 38° compared with 32° for the S9L structure, effectively narrowing the end of the C3L variant. This twist was alleviated gradually as H3 progressed toward the internal loop of the B domain.

Contraction of the C3L H3 helical end affects H4′ of the symmetry-related molecule owing to pseudo-continuous helical crystal-packing interactions (Fig. 4d). The terminal
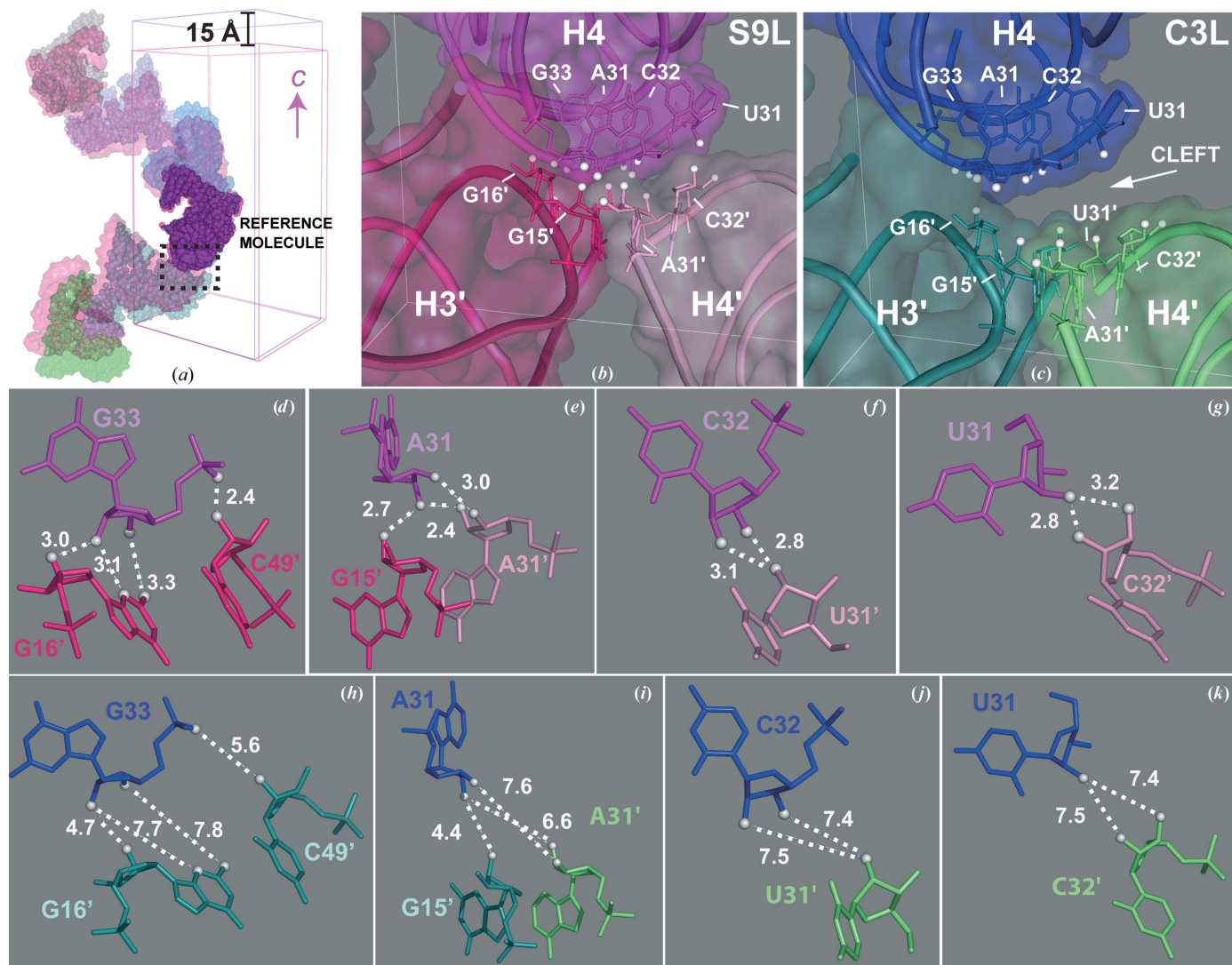

Figure 5
Schematic surface and ball-and-stick diagrams illustrating the 6_1 packing interactions of hinged hairpin-ribozyme constructs. (a) Comparison of C3L (blue) and S9L (magenta) unit cells. The perspective represents the interaction of molecules about the 6_1 screw axis. (For twofold and 2_1 operations, refer to Fig. 6a.) Superposition of C3L onto S9L initiates at the third molecule from the bottom (i.e. the 'reference' molecule), represented in bold magenta and blue overlay. Symmetry operations used to generate the remaining molecules demonstrate the degeneration of the superposition as a result of the 15 Å elongated unit cell of the C3L structure. C3L molecules are shown in blue and green, with S9L structures shown in pink. The dashed box denotes the inset for (b) and (c). (b) Expanded view of the 6_1 packing scheme for the S9L structure. Atoms engaged in potential hydrogen bonds are depicted as white spheres. Helices are labeled as in Fig. 1(c) and primes (′) denote symmetry-related molecules. An additional H3″ helix that is packed in a blunt-ended base stack with H4 has been omitted for clarity. The asymmetric unit (H4) is colored magenta, with symmetry molecules in red (H3′) or light pink (H4′). (c) The C3L 6_1 packing scheme as described in (b), but the asymmetric unit is colored blue and symmetry mates are coloured teal (H3′) or green (H4′). White spheres identify equivalent atoms in the C3L structure that are engaged in hydrogen bonding in the S9L structure. (d)–(g) Detailed view of hydrogen-bond interactions with each of four H4 residues. The A31 and U31 residues are base-paired; A31 ends the linker strand, while U31 begins the S-turn strand. (h)–(k) Equivalent distances in the C3L structure demonstrate the loss of 6_1-fold packing interactions.

base pair of H4′ (A31′–U31′) experiences a modest lateral movement similar to, but less than, that observed for H3. This is illustrated in a comparison of the distances between the C3′ of C49 and the O4′ of U31′ (4 Å in S9L versus 2.9 Å in C3L; Figs. 4d and 4e). Because the lateral shift exhibited by the

A31′–U31′ base pair in the C3L structure is not fully commensurate with that of H3, it is necessarily accompanied by a 1.5 Å upward shift of U31′ (Fig. 4f). This upward movement circumvents an otherwise inescapable steric clash between the ribose moieties of U31′ and C49. The net result is
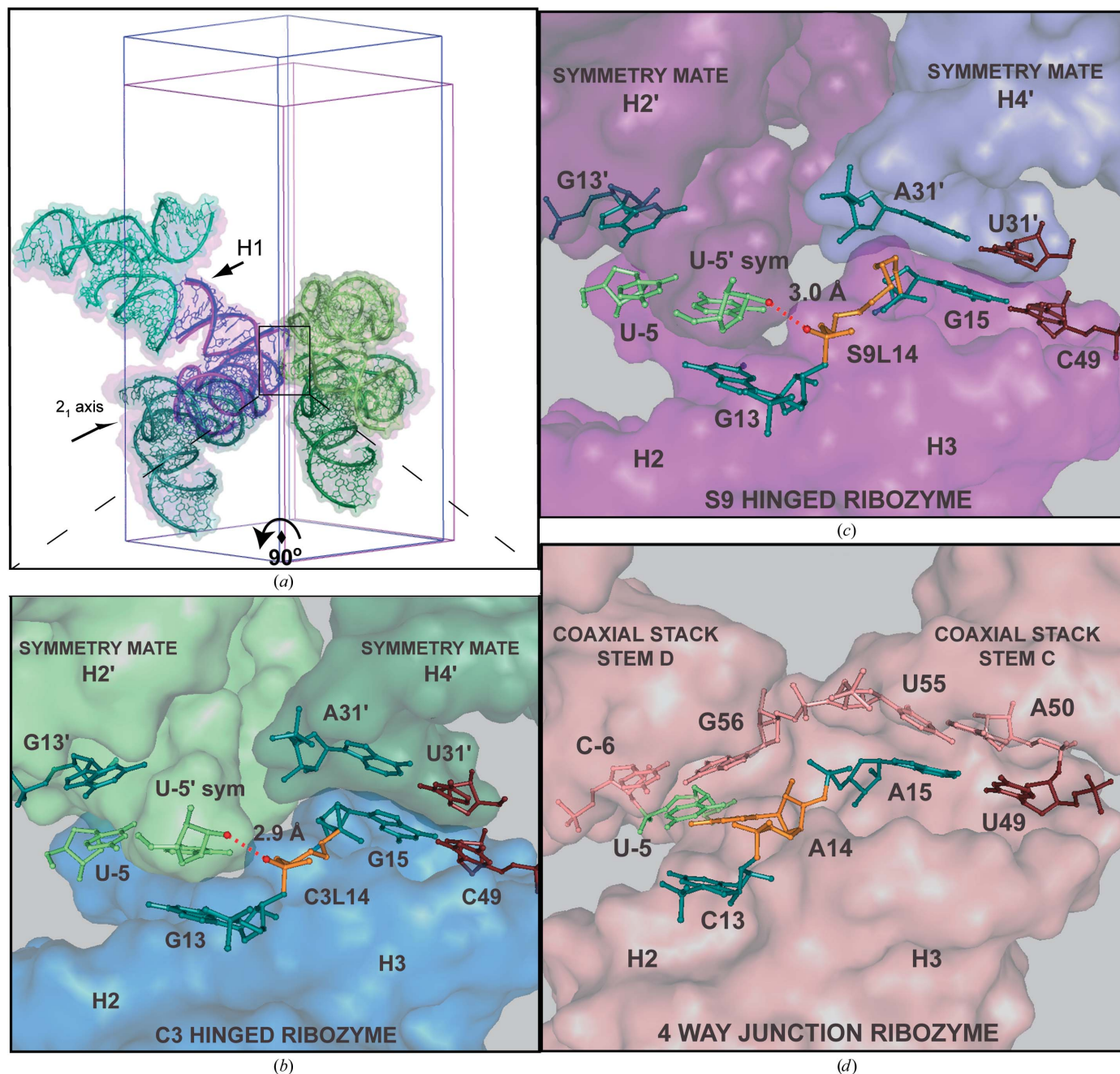

Figure 6
Schematic surface and ball-and-stick diagrams illustrating a pseudo-four-way helical junction generated by crystal-packing interactions of minimal hairpin ribozymes. (a) Unit-cell perspective of the twofold and 2_1-fold symmetry contacts within the asymmetric units of both C3L and S9L structures. Superposed C3L (blue) and S9L (magenta) molecules are shown overlaid in the center of the diagram (i.e. reference molecule). The minimal difference in the lengths of the a and b axes of the C3L and S9L unit cells (3.2 Å) allows greater conservation in the superposition as it is extended from the reference molecule's asymmetric unit; therefore, symmetry-related S9L molecules (pink surface representations) are largely hidden behind the C3L molecules (blue and green cartoon and stick depiction). Helix H1 is labeled to demonstrate its high degree of solvent exposure. The 2_1 screw axis depicts the relationship between H3 and H4′. The boxed region is expanded in (b)–(d). (b)–(d) Individual panels of the junction region of the C3L, S9L and 4WJ structures. Residues are colored according to strand conventions in Fig. 1(b) and numbered according to hairpin-ribozyme conventions (Chowrira & Burke, 1991). Surface representations for the C3L molecules are coloured as observed in (a) and rotated 90° anticlockwise from that perspective. O atoms engaged in putative hydrogen bonds are designated as red spheres connected by dashed lines.

a more flush base stack between H3 and H4′ in the C3L structure compared with the more staggered interaction observed in the S9L (Figs. 4d and 4e) and JL structures (not shown). From an engineering and design perspective, there are two opposing factors that appear to influence this end-to-end stacking interaction. The first is the local influence of the shortened C3 linker, which draws the G15–C49 base pair closer to the center of mass of the RNA, effectively improving the crystallographic base-stacking interactions. The second factor originates from crystal contacts observed in both the S9L and JL hairpin-ribozyme structures. In these two structures the terminal residues of both H3′ and H4 are engaged in backbone and minor-groove hydrogen bonds that confer stability along the 6_1-fold screw axis (for perspective, see Figs. 5a and 5b). These interactions support the more staggered helical packing, but are notably absent from the C3L structure (Fig. 5c, cleft), which is relevant to understanding the principles of RNA interaction that influence high-resolution X-ray diffraction.

3.3.3. Variations in crystal-packing interactions within the minimal constructs. All termini of the three minimal hairpin-ribozyme structures discussed here (Table 1) engage in end-to-end helical base stacking. These interactions sustain the twofold and 2_1 axes present in space group $P6_122$. The base stacks observed for H1, H3 and H4 are blunt-ended, whereas the U·U mismatch of H2 is sticky. Each pseudo-helical packing interaction buries approximately 400 Å^2 of hydrophobic surface. The terminus of H1 and H2 each stacks self-to-self on dyad axes; the self-to-self (i.e. H1–H1′) nature of this interaction offers little support to the mostly solvent-exposed length of H1 and may contribute to the disorder observed in this helix. In contrast, H3 and H4′ stack onto each other in support of the 2_1 symmetry axis. In this manner, linker influences originating in H3 are propagated into H4′ of the symmetry-related molecule.

Minor-groove and backbone interactions contribute to the formation of the 6_1 screw axis (Fig. 5). The backbone of H4 of a 'reference' molecule (Fig. 5a, dark molecule) bridges the blunt stack formed by helices H3′ and H4′, which are related by a crystallographic 2_1 symmetry axis (Figs. 5b and 5c). According to the symmetry operators of space group $P6_122$, H4 of the reference molecule forms a blunt-end stack to symmetry-related molecule H3″, which is comparable to the H3′ interaction with H4′ (Figs. 5b or 5c). If all helices were included at this junction, the rudimentary shape would be a Greek cross or '+'. (Note: H3″ was omitted for clarity, thus resulting in the letter 'T' in Figs. 5b and 5c.) The cross is formed by two base-stacked smaller helices that form a single pseudo-continuous helix. These longer helices contribute the respective horizontal or vertical cross components. The intersection of these components comprises minor-groove and backbone contributions of varied extent that depend

upon the choice of linker (Figs. 5b and 5c). At the intersection of these helices, H4 buries 1200 Å^2 of surface through interactions with three of the six ribozymes that contact the asymmetric unit. The minor-groove/backbone component of this interaction within the S9L structure buries 800 Å^2, within which there are 11 hydrogen bonds (Figs. 5d–5g). In contrast, the comparable interaction of H4 in the C3L structure buried a total of only 700 Å^2. While the 2_1 base stacks were equivalent in the two structures (400 Å^2), the minor-groove contacts were largely absent in the C3L structure, giving rise to a large cleft (Fig. 5c). Significantly, no stabilizing hydrogen bonds were identified within the 300 Å^2 of buried surface that constituted the minor-groove symmetry contact in the C3L structure (Figs. 5h–5k). The bridged lattice-packing interaction observed in the S9L structure was also lacking in the C3L ribozyme, such that H4 made contact with only H3′, rather than with both H3′ and H4′ as in the S9L structure (Figs. 5b and 5c). The packing deficits in the C3L structure undoubtedly contribute to its poor diffraction properties relative to the S9L variant.

The paucity of crystallographic packing interactions along the 6_1 axis of the C3L structure appears directly attributable to the choice of interdomain linker. The shift of H3 toward the RNA center of mass of the asymmetric unit moves it away from the minor-groove interaction with symmetry-related helices H3′ and H4′ (Fig. 5b versus 5b). The ensuing loss of hydrogen-bond interactions and changes in shape complementarity between intermolecular surfaces appear to destabilize the 6_1-fold packing scheme (Figs. 5h–5k), thus contributing to the elongated c axis (Fig. 5a). These observations suggest that base-stacking interactions promote formation of the crystal lattice, but the 2′-OH-mediated intermolecular hydrogen-bond contacts dictate lattice stability. Thus, the C3L lattice exhibits sufficient flexibility to maintain pseudo-continuous helical packing at all four termini, but the 6_1-fold minor-groove interactions are necessarily sacrificed. Although a number of small-molecule compounds from the crystallization medium, such as spermidine or glycerol, could theoretically span the gap between H4 and H4′ of C3L, no such ligands from the
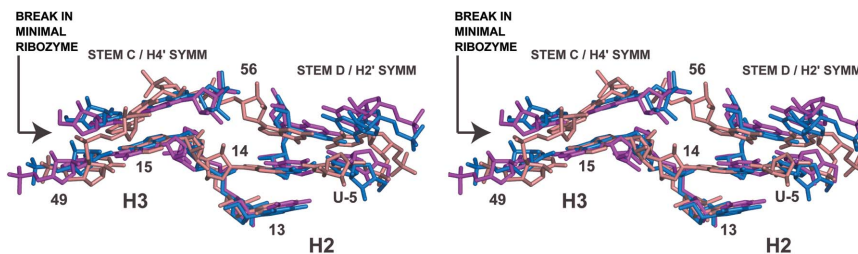

Figure 7
A stereoview stick diagram of a superposition of the 4WJ, C3L and S9L hairpin-ribozyme structures. The orientation of the stereoview is rotated 180° around the vertical axis compared with Fig. 6 and was chosen to emphasize the agreement between residues of the intramolecular four-way junction with those of the intermolecular packing environment observed in the C3L and S9L structures. The 4WJ structure is shown in salmon, C3L in blue and S9L in magenta. Residues are numbered according to hairpin-ribozyme conventions and are labeled only where the sequence is conserved in all three structures except position 56 of the natural 4WJ..

crystallization/cryoprotection medium were observed in $2F_o - F_c$ or $F_o - F_c$ electron-density maps, which has implications for the use of such additives to stabilize an RNA lattice.

3.4. Comparison of four hairpin-ribozyme structures with different interdomain linkages

3.4.1. Local comparison. Four alternatively hinged ribozyme structures are now available, offering the first opportunity to examine the impact of the hinge region on the overall RNA fold. The S9L and C3L structures of this investigation are the first to link the A and B domains without the addition of the remaining two helices of the four-way helical junction. As such, they provide a means of understanding the structural influence of the hairpin-ribozyme interdomain linkage (hinge) outside the context of the four-way helical junction. This comparison also has implications for the use of synthetic linkers to produce crystal contacts in RNA constructs, as well as the inclusion of crystal contacts in searches aimed at identifying potentially biologically relevant motifs that promote tertiary or quaternary RNA folding.

A superposition of all four hairpin-ribozyme variants demonstrated excellent agreement between residues at the active site. The pairwise r.m.s.d. values between the 4WJ structure and the C3L and S9L structures were 1.5 and 1.6 Å, respectively. A closer inspection of the local differences indicated that the C3 linker most closely resembled the 4WJ A14 linkage (Fig. 4c). This result seems reasonable since each of these linkages possesses three C atoms between the phosphate groups at positions 14 and 15. The G13 O3′ to G15 O3′ distances clustered accordingly (Figs. 4a and 4b and described above); this distance was measured as 11.1 Å in the 4WJ structure (10.6 Å for C3L) and the JL distance was 12.7 Å (also 12.7 Å for S9L). Local twist values for G15 at the top of H3 showed the same trend: i.e. the C3L and 4WJ structures exhibited increased twist (38° and 48°), whereas the S9L and JL structures were more relaxed (31.9° and 31.6°); for reference, standard A-form RNA exhibits a twist of 32.7° (Saenger, 1984). While H3 of the C3L and 4WJ structures is likely to be overwound in response to the shortened linker, the mildly underwound twist values observed for the S9L and JL structures may reflect crystal packing. Twist exhibits a linear dependence on minor-groove width, as calculated by *CURVES* (Boutonnet *et al.*, 1993). The 6_1-fold packing scheme of the S9L and JL structures is mediated largely by backbone and ribose contacts along the outside of the minor grooves of H3 and H4, with two genuine minor-groove hydrogen bonds to the guanine base of residue 16 (Fig. 5d). These contacts may contribute to the widened minor-groove values calculated for the penultimate base pair using *CURVES*, reported as 10.2 Å (JL) and 10.4 Å (S9L) compared with 9.4 Å (C3L) and 9.3 Å (4WJ); notably, the 15–49 base pair interaction was not amenable to measurement by *CURVES* owing to its terminal location. Standard A-form RNA exhibits a minor-groove width of 11 Å (Saenger, 1984) and the discrepancy from this value may be accounted for by the higher than normal local

twist, where these deviations appear to be confined. In contrast, variation in the minor-groove width for H4, which contacts H3 in the crystal lattice, exhibited a much smaller range of values (±0.2 Å), with a mean of 9.5 Å. The twist values for this helix also exhibited a smaller range (±0.4°), with an average of 33°. Most importantly, the distortion here is mitigated as one approaches the active site of the hairpin ribozyme. This observation has evolutionary implications for RNA enzymes and suggests that the folds of globular RNAs may be somewhat self-correcting and tolerant of flanking structural perturbations, even though they lack the true hydrophobic core that confers proteins with the substantial plasticity and stability required to accommodate insertions (Vetter *et al.*, 1996).

3.4.2. Implications for molecular mimicry through engineered RNA crystal packing. A novel feature of the minimal ribozyme lattice was observed during the comparison of minimally hinged and 4WJ structures. The prevalence of pseudo-continuous helical packing within the lattice of minimal constructs (Fig. 6a) enabled the formation of a rudimentary 4WJ motif. H4′ of one symmetry mate stacked bluntly onto H3, while H2′ of another symmetry mate stacked onto H2 in a staggered fashion through the U·U interaction (*e.g.* Fig. 6b); additionally, the 2′-OH of a symmetry-related U−5 molecule was positioned appropriately to engage in an ∼3 Å hydrogen bond with a nonbridging phosphoryl O atom of either linker (Figs. 6b and 6c). Overall, the stacking interactions at this helical intersection bury 490 and 540 Å^2 of surface area within the C3L and S9L structures, respectively, and exhibit an uncanny resemblance to the 4WJ motif (Fig. 6d). A similar series of interactions were described above as stabilizing forces for the interdomain linker residue, A14, which is present in the 4WJ structure. In the latter molecule, the base pair between U−5 and A14 at the top of H2 was flanked by a coaxial interaction to form an energetically favorable flush stack; the end of H3 was supported similarly (Fig. 6d; Rupert *et al.*, 2002). The strand of RNA in the 4WJ structure that comprised these two coaxial stacks was single-stranded as it crossed over A14, further contributing to the ∼710 Å^2 of surface area buried in the interface of this biologically relevant motif. Although the natural 4WJ is an intramolecular interaction, in contrast to the intermolecular contacts of the minimal constructs, a superposition of these three structures further emphasizes the excellent agreement between the modes of stacking at this interface (Fig. 7).

As demonstrated here, base-stacking of pseudo-continuous helices is a powerful packing restraint commonly observed in natural RNA structures and should be considered as a key driving force when engineering blunt or overhanging duplex sequences for crystallization. The inclusion of synthetic linkers between helices and the generation of junctions through symmetry contacts represent rational approaches to tailor RNA crystallization constructs to the needs of specific structural studies. This approach requires *a posteriori* structural knowledge, but the outcome demonstrates that significant advances in crystallization and diffraction can be attained by

modifying a crystallization construct in a manner that uses variations of the naturally occurring RNA architecture.

4. Summary and conclusions

This investigation represents the first crystallographic characterization of the hairpin ribozyme incorporating synthetic linkages between the loop A and B domains. As such, it is relevant on two fronts. From a crystallographic point of view, it represents a novel method for connecting RNA strands. Synthetic linkers appear to be less prone to nonproductive coaxial stacking compared with natural nucleotide linkages, which have limited flexibility as well as greater hydrophobicity and steric bulk. The structures of this study also highlight the potential for deliberately promoting 4WJ mimicry in RNA crystal-packing interactions, with an obvious rationale for screening multiple linkers. Moreover, the use of crystallographic screening and X-ray diffraction analyses on multiple constructs is still worthwhile, despite the apparent similarities of constructs in solution enzymatic assays.

From a functional perspective, comparison of the four structures discussed here (JL, C3L, S9L and 4WJ) provides the first opportunity to analyze the influences of the hinge region on the global fold of the ribozyme. Significantly, this investigation was conducted in the context of multiple lattice packing schemes. As such, the comparisons have facilitated a dissection of characteristics relevant in a biological setting from those generated solely from crystal lattice contacts, which can be especially problematic to parse in nucleic acid structures (Yajima *et al.*, 2007; Wedekind & McKay, 2003). Overall, these results suggest a rationale for the production of minimal RNA constructs based on natural 4WJ motifs that have led to well diffracting hinged hairpin-ribozyme constructs amenable to incorporation of synthetic abasic residues and transition-state analogues for use in structure–function studies.

References

Alam, S., Grum-Tokars, V., Krucinska, J., Kundracik, M. L. & Wedekind, J. E. (2005). *Biochemistry*, **44**, 14396–14408.

Anderson, A. C., Earp, B. E. & Frederick, C. A. (1996). *J. Mol. Biol.* **259**, 696–703.

Boutonnet, N., Hui, X. & Zakrzewska, K. (1993). *Biopolymers*, **33**, 479–490.

Britton, D. & Chantooni, M. K. Jr (2001). *J. Chem. Crystallogr.* **31**, 5–16.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Burkard, M. E., Turner, D. H. & Tinoco, I. (1999). *The RNA World*, edited by R. F. Gesteland & J. F. Atkins, pp. 675–680. Plainview, NY, USA: Cold Spring Harbor Laboratory.

Butcher, S. E., Heckman, J. E. & Burke, J. M. (1995). *J. Biol. Chem.* **270**, 29648–29651.

Chowrira, B. M., Berzal-Herranz, A., Keller, C. F. & Burke, J. M. (1993). *J. Biol. Chem.* **268**, 19458–19462.

Chowrira, B. M. & Burke, J. M. (1991). *Biochemistry*, **30**, 8518–8522.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

DeLano, W. L. (2004). *The PyMOL Molecular Graphics System*, v. 0.97. http://www.pymol.org.

Esteban, J. A., Banerjee, A. R. & Burke, J. M. (1997). *J. Biol. Chem.* **272**, 13629–13639.

Esteban, J. A., Walter, N. G., Kotzorek, G., Heckman, J. E. & Burke, J. M. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 6091–6096.

Ferré-D'Amaré, A. R., Zhou, K. & Doudna, J. A. (1998). *J. Mol. Biol.* **279**, 621–631.

Foloppe, N. & Nilsson, L. (2005). *J. Phys. Chem. B*, **109**, 9119–9131.

Grum-Tokars, V., Milovanovic, M. & Wedekind, J. E. (2003). *Acta Cryst.* D**59**, 142–145.

Hampel, A. & Tritz, R. (1989). *Biochemistry*, **28**, 4929–4933.

Hohng, S., Wilson, T. J., Tan, E., Clegg, R. M., Lilley, D. M. & Ha, T. (2004). *J. Mol. Biol.* **336**, 69–79.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Jovine, L. (2003). *Nuccyl.* http://www.biosci.ki.se/groups/ljo/software/nuccyl.html.

Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.

Kierzek, R., Burkard, M. E. & Turner, D. H. (1999). *Biochemistry*, **38**, 14214–14223.

Kleywegt, G. J. (1995). *Dictionaries for Heteros.* http://xray.bmc.uu.se/usf/factory_5.html.

Kraut, J. (1961). *Acta Cryst.* **14**, 1146–1152.

Kuzmin, Y. I., Da Costa, C. P., Cottrell, J. W. & Fedor, M. J. (2005). *J. Mol. Biol.* **349**, 989–1010.

Lavery, R. & Sklenar, H. (1989). *J. Biomol. Struct. Dynam.* **6**, 655–667.

Nagaswamy, U., Larios-Sanz, M., Hury, J., Collins, S., Zhang, Z., Zhao, Q. & Fox, G. E. (2002). *Nucleic Acids Res.* **30**, 395–397.

Neumann, J. M., Bernassau, J. M., Gueron, M. & Tran-Dinh, S. (1980). *Eur. J. Biochem.* **108**, 457–463.

Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.

Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. (1994). *Nature (London)*, **372**, 432–438.

Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.

Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.

Rueda, D., Bokinsky, G., Rhodes, M. M., Rust, M. J., Zhuang, X. & Walter, N. G. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 10066–10071.

Rupert, P. B. & Ferré-D'Amaré, A. R. (2001). *Nature (London)*, **410**, 780–786.

Rupert, P. B., Massey, A. P., Sigurdsson, S. T. & Ferré-D'Amaré, A. R. (2002). *Science*, **298**, 1421–1424.

Saenger, W. (1984). *Principles of Nucleic Acid Structure.* New York: Springer-Verlag.

Salter, J., Krucinska, J., Alam, S., Grum-Tokars, V. & Wedekind, J. E. (2006). *Biochemistry*, **45**, 686–700.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. Plainview, NY, USA: Cold Spring Harbor Laboratory Press.

Selmer, M., Dunham, C. M., Murphy, F. V. IV, Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R. & Ramakrishnan, V. (2006). *Science*, **313**, 1935–1942.

Sugimoto, N., Kierzek, R. & Turner, D. H. (1987). *Biochemistry*, **26**, 4554–4558.

Tan, E., Wilson, T. J., Nahas, M. K., Clegg, R. M., Lilley, D. M. & Ha, T. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 9308–9313.

Tanaka, Y., Fujii, S., Hiroaki, H., Sakata, T., Tanaka, T., Uesugi, S., Tomita, K. & Kyogoku, Y. (1999). *Nucleic Acids Res.* **27**, 949–955.

Torelli, A. T., Krucinska, J. & Wedekind, J. E. (2007). In the press.

Vetter, I. R., Baase, W. A., Heinz, D. W., Xiong, J. P., Snow, S. & Matthews, B. W. (1996). *Protein Sci.* **5**, 2399–2415.

Wahl, M. C., Rao, S. T. & Sundaralingam, M. (1996). *Nature Struct. Biol.* **3**, 24–31.

Walter, A. E. & Turner, D. H. (1994). *Biochemistry*, **33**, 12715–12719.

Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Muller, P., Mathews, D. H. & Zuker, M. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 9218–9222.

Walter, N. G. & Burke, J. M. (1998). *Curr. Opin. Chem. Biol.* **2**, 303.

Wedekind, J. E. & McKay, D. B. (2000). *Methods Enzymol.* **317**, 149–168.

Wedekind, J. E. & McKay, D. B. (2003). *Biochemistry*, **42**, 9554–9563.

Yajima, R., Proctor, D. J., Kierzek, R., Kierzek, E. & Bevilacqua, P. C. (2007). *Chem. Biol.* **14**, 23–30.

Yang, H., Jossinet, F., Leontis, N., Chen, L., Westbrook, J., Berman, H. & Westhof, E. (2003). *Nucleic Acids Res.* **31**, 3450–3460.