Anticodon-like loop-mediated dimerization in the crystal structures of HdV-like CPEB3 ribozymes

Anna Ilaria Przytula-Mally [a]; Sylvain Engilberge [b]; Silke Johannsen [a]; Vincent Olieric [b]; Benoît Masquida [c*]; Roland K.O. Sigel [a*]

[a] Department of Chemistry, University of Zurich, CH-8057 Zurich, Switzerland

[b] Swiss Light Source, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland;

[c] UMR 7156, CNRS - Université de Strasbourg, IPCB, 21 rue René Descartes, Strasbourg 67084, France

*To whom correspondance should be addressed

roland.sigel@chem.uzh.ch

b.masquida@unistra.fr

Keywords: RNA, HDV ribozyme, CPEB3 ribozyme, X-ray crystallography, structure, dimerization

ABSTRACT:

Cytoplasmic polyadenylation element-binding (CPEB) proteins are involved in many different cellular processes including cell division, synaptic plasticity, learning, and memory. A highly conserved, short mammalian ribozyme has been found within the second intron of the CPEB3 gene. Based on its cleavage mechanism and structural features, this ribozyme belongs to the hepatitis delta virus (HDV)-like ribozyme family. Here, we present the first crystallographic structures of human and chimpanzee CPEB3 ribozymes both confirming the general topology of the HDV ribozyme with two parallel co-axial helical stacks. However, the residues involved in the formation of the characteristic double-nested pseudoknot P1.1 instead participate in a seven nucleotides loop with a conformation identical to the one found in the anticodon (AC) loop of tRNAs. The conformation of the loop supports the formation of a 4 base pair helix by interacting with an AC-like loop from a symmetry-related ribozyme leading to ribozyme dimer formation. Comparing the present crystal structures to the one from the genomic HDV ribozyme shows how nucleotides with different identities adopt distinct behaviours. These new snapshots suggest unforeseen relationships between the genomic, the antigenomic versions of the HDV ribozyme and the CPEB3 ribozyme.

INTRODUCTION

Since its discovery, the genomic HDV ribozyme has been one of the most studied model small self-cleaving ribozymes, contributing to better understand the physicochemical basis of the self-scission mechanism, and illustrating the principles of metal ion/nucleic acid interaction and more generally RNA structural features. The first crystal structure of the post-cleavage state of the genomic HDV ribozyme revealed that it comprises five double stranded A-helical regions (P1, P1.1, P2, P3 and P4) joined by three single stranded regions (J1/2, J1.1/4 and J4/2) organized into two parallel helical stacks where P1/P1.1/P4 constitute one of them and P2/P3 the other (1). The structure also revealed the prominent role of the P1.1 minihelix and of the P3-P1 cross-over in the organization of the catalytic core (**Figure 1A**). The highly complex P1.1 nested pseudoknot fold weaves the two helical stacks together and builds up a platform on top of which sits the catalytic site, notably the catalytically important C75 (1-3).

In the following years, the structures of the uncleaved (4), mutant-inhibited (5) and ion-complexed forms (6) were published, confirming that both pre- and post-cleavage states of the genomic HDV ribozyme, adopt nearly identical folds, including the double nested pseudoknot (simultaneous formation of P1-P2 and P3-P1.1). The catalytic core of the HDV ribozyme is flanked by two highly conserved G•U wobble base pairs (bp) playing different structural functions. The first one corresponds to the ribozyme first bp of P1, located right at the cleavage site. This *cis* G1•U37 wobble has a local stabilizing effect on the C75-G1 hydrogen bonds and has been shown to bind Mg^{2+} (7) together with the second wobble bp located in P3, the *trans* G25oU20 (4). This bp adopts different conformations during self-cleavage, and

has a stabilizing action on both proximal and distal elements located within the L3 loop, which forms the P1.1 pseudoknot (8).

In 2006, using a human transcriptomic library designed for searching ribozymes, the Szostak group discovered a new self-cleaving ribozyme located in the second intron in the pre-mRNA encoding the cytoplasmic polyadenylation element binding 3 protein (CPEB3) (9) (**Figure 1B**). Based on biochemical and secondary structure characterization, this ribozyme was classified as the first representative of the HDV-like ribozyme family (10) (**Figure 1C**). The *cpeb3* gene encodes the prion-like CPEB3 protein, a protein involved in synaptic plasticity and long-term memory (11-13). The ribozyme sequence is highly conserved among mammals and the self-cleavage activity has persisted over > 170 millions years of evolution (14). However, despite its apparent similarity to the HDV ribozyme, very little is known about the ribozyme mechanism and function. It has been suggested that coupling between splicing and ribozyme self-cleavage may regulate the CPEB3 protein level and thus have an impact on synaptic plasticity (15).

Here, we focus on the human and chimpanzee homologues of the CPEB3 ribozyme (humCPEB3r and chimpCPEB3r). The sequences of these ribozymes differ only at one position located at the tip of P1, and this seems to be sufficient to make the chimpanzee ribozyme cleave faster than the human one (**Figure 1C**) (14,16). Comparative sequence analysis suggests that P1 from humCPEB3r is closed by a wobble C•A pair (also found in sloth, armadillo and dolphin), while other mammals, including the chimpanzee, present a *cis* Watson-Crick C-G bp (10).

In an effort to investigate the three-dimensional structure of HDV-like family members, we report here the crystal structures of the cleavage products of the human and chimpanzee CPEB3 ribozymes. The present structures confirm the general topology observed in the HDV

ribozyme with two coaxial helical stacks. However, we observed an unforeseen open conformation. The two bp pseudoknot P1.1 is replaced by a four bp stem formed between the L3 loops of two (pseudo)symmetry-related CPEB3 ribozymes (**Figure 1D**). Strikingly, the L3 loops fold like the 7 nucleotides (nt) tRNA anticodon loops in which absence of chemical modifications leaves the four palindromic nt downstream from the U-turn free to participate in a four bp P3.1 dimerization stem. Interestingly, RNA dimerization has been considered a crystallization artefact in the case of tetraloop-containing oligonucleotides in the 90's (17-20). Nevertheless, it is worth to note that ribozyme dimerization can also be biologically relevant, as in the case of the Varkud Satellite ribozyme (21), the hammerhead ribozymes from Penelope-like transposable elements (22), as well as the Hatchet ribozyme (23). Nucleic acids oligomerization and nucleic acid/protein cluster formation is nowadays an area of intensive investigations showing a functional relevance of those structures (24). A precise structural analysis of the HDV ribozyme representatives indicates that the antigenomic HDV ribozyme presents conserved features from the CPEB3 ribozyme that are not found in the genomic HDV ribozyme. This work gives additional support to the hypothesis made by Szostak that the HDV ribozymes have evolved from the CPEB3 ribozyme.

MATERIAL AND METHODS

RNA and protein purification: The DNA templates were designed to encode the human or chimpanzee CPEB3 ribozyme sequence with a U1ABD within P4 and a 9 nt 5' leader sequence to allow self-scission of the ribozyme during transcription. Transcription was performed according to Gallo et al. (25) using 0.85 µM of DNA template (Microsynth), 0.85 µM complementary strand (Microsynth), 5 mM of each NTP (GE Healthcare), 0.01% Triton X-100, 35 mM $MgCl_2$, 200 mM Tris-HCl pH 7.5, 200 mM DTT, 10 mM Spermidine, and T7 polymerase

(the amount was adapted according to the activity of each enzyme batch) and were incubated for 4 hours at 38°C. The RNA was precipitated overnight at -20°C upon addition of 3 volumes of ethanol and 1/20 volume of 5 M NaCl. The RNA was purified on an 8 M urea, TBE, 12% polyacrylamide gel, visualized by UV shadowing, excised from the gel and recovered by electroelution. For desalting, Vivaspin® Concentrators with 5000 Da MWCO (Viva products) cutoff were used. By ultrafiltration, the RNA samples were washed 3 times with 1 M KCl, pH 8.0. Afterwards, the samples were washed with water. The final pH was adjusted to 7 with 3.2% HCl. After lyophilization, the RNA was dissolved in 100 mM KCl, 10 mM HEPES-KOH pH 7.5, 10 µM EDTA buffer to reach the final concentration of 0.4 mM.

Double-Mutant U1A protein plasmid was a kind gift from Adrian Ferré-d'Amaré. The protein was produced in *Escherichia coli* BL-21 Star™ (DE3) pLysS, induced with 0.5 mM IPTG. The protein purification was described by Ferré-d'Amaré (26), except that the last purification step by hydroxyapatite (CHT-I, BioRad) chromatography was omitted. The final product was confirmed by SDS-PAGE and Mass Spectrometry.

Complex formation: The RNA purified from the transcription reactions was refolded by 1 minutes incubation at 90°C followed by 4 minutes on ice. The refolded RNA was annealed to the Double-Mutant U1A protein 1:1 molar equivalent ratio to yield a complex concentration of 0.2 mM (in 50 mM KCl, 6.7 mM HEPES-KOH pH 7.5, 5 nM EDTA, 5 mM $MgCl_2$, and 1 mM spermine).

Crystallization: For crystallization by the vapor-diffusion technique, the sitting drops of 400 nL of complex solution were mixed with 200 nL of a reservoir solution. The initial crystals of

both complexes were obtained at the Protein Crystallization Center (PCC, UZH) using a Gryphon LCP robot (ART - Art Robbins Instruments) and the Index HT sparse matrix screen (Hampton Research) for the human CPEB3 complex and PEG/Ion screen (Hampton Research) for the chimpanzee complex. The human CPEB3 complex was crystallized in a solution containing 20-24% PEG 3350, 0.2 M sodium citrate tribasic dihydrate and 30-200 mM HEPES-KOH pH 7.5. For the chimpanzee CPEB3 complex, the reservoir solution contained 19-24% PEG 3350, 0.2 M sodium formate. The crystallization setups were incubated at 4°C. The human and chimpanzee CPEB3 complex crystals appeared within 5 days, and grew over 25 days, and reached maximum dimensions of 250 µm x 50^2 μm^2, space group $C222_1$. For cryogenic conditions, the crystals of the human CPEB3 complex were transferred to cryoprotecting solution containing 25% PEG3350, 10 mM HEPES-KOH pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 2 mM spermine, 0.2 M sodium citrate tribasic dihydrate, and 5% or 10 % (v/v) glycerol. The crystals of the chimpanzee CPEB3 complex were transferred to a solution containing 22% PEG3350, 10 mM HEPES-KOH, pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 2 mM spermine, 0.2 M sodium formate, 10% or 18% (v/v) glycerol. Crystals were mounted into CryoCap loops (Molecular Dimensions) and flash-cooled in liquid nitrogen.

Data collection: X-ray structure determination and refinement: Native datasets were collected at the beamline PXIII (X06DA) at the Paul Scherrer Institute (PSI), using a Pilatus 2M-F detector and the multi-axis PRIGo goniometer at 1.00004 Å wavelength with an exposure of 0.2 s per 0.2° for 360° data set.

The autoPROC pipeline was used to process the data. Diffraction frames were integrated using XDS and scaled with Aimless (27). Staraniso (28) was used to correct the anisotropy. The human complex structure was solved by the molecular replacement method in Phaser (29),

using 1drz as a search model. Remaining nucleotides were built by hand. The chimpanzee complex structure was solved using the human complex as a template. In addition, for the chimpanzee CPEB3 crystals, anomalous data were collected with X-ray energy at 5.975 keV to help the localization of phosphorus and sulfur atoms. Iterative model building and refinement was performed using the programs PHENIX (30), COOT (31), and Buster 2.10.3 (32). More than 370 crystals were screened to solve the humCPEB3 structure, owing to the temperature sensitivity of the crystals and a high solvent content of 74%. The highest resolution was 2.83 Å. Only 20 crystals were screened to determine the chimpCPEB3 structure. They had a solvent content of 62% and diffracted to a resolution of 2.18 Å.

The humCPBE3-U1Ar is a crystallographic dimer while the chimpanzee ribozyme is pseudo-symmetric, meaning that two different but very close conformations of the chimpCPEB3 ribozyme are observed. The refined models were validated with MolProbity (33). Data collection and refinement statistics are shown in **Table 1**. All figures were prepared using PyMOL (34).

Pair Fit Alignment: Root Mean Square Deviation determination was performed by Pair Fit Alignment in PyMol (34). For alignment with the genomic HDV ribozyme structure (PDB ID: 4pr6), Phosphorus atoms from nucleotides' stretches ranging from 101 to 122, from 163 to 172 (HDV structure) and from 1 to 22, 59-68 for both human and chimpanzee CPEB3 ribozymes were used. These parameters were chosen to clearly identify structural differences between HDV and CPEB3 ribozymes without focusing on the P4 and U1A-RNP domains. Alignment of all CPEB3 ribozyme structures was based on their full RNA sequence (nucleotides from 1 to 69).

Sequences: provided in DNA

Short template (forward strand):

5'-GAAATTAATACGACTCACTATAGG-3'

Human-CPEB3-U1A ribozyme (reverse strand):

3'-

CTTTAATTATGCTGAGTGATATCCCTATTGTCCCCCGGTGTCGTCTTCGCAAGTGCAGCGTCGGGGAC

AGTCGGTAACGTGAGGCCGACGCTTAAGACGA-5'

Chimpanzee CPEB3-U1A ribozyme (reverse strand):

3'-

CTTTAATTATGCTGAGTGATATCCCTATTGTCCCCCGGTGTCGTCTTCGCAAGTGCAGCGCCGGGGAC

AGTCGGTAACGTGAGGCCGACGCTTAAGACGA-5'

RESULTS

Design of the CPEB3 ribozyme cleavage product

Based on pioneering work from Ferré-D'Amaré and colleagues (1), we engineered the CPEB3 ribozyme to overcome structural heterogeneities by harvesting the RNA product of the self-scission reaction. A nine nt leader sequence 5'-GGGAUAUCA-3' was added upstream from the CPEB3 ribozyme cleavage site to let the catalytic reaction proceed in the course of *in vitro* transcription (IVT) reactions. The leader sequence allows distinguishing catalytically active ribozymes from RNAs trapped in non-active conformations by PAGE fractionation. Additionally, to promote crystal packing interactions, non-essential regions of the RNA secondary structure can be replaced by known RNA motifs. For example, a GNRA tetraloop or

the human U1A RNA binding protein are frequently used to engineer the RNAs to crystallize (26,35-40).

Previous studies of the HDV ribozymes showed that the length of P4 can be changed without noticeable influence on the catalytic activity. Moreover, in some HDV-like ribozymes this region does not exist (10). Accordingly, both human and chimpanzee CPEB3 ribozymes were designed with the U1ABD substitution in the catalytically dispensable P4 helix (**Figure 1C**). The effectiveness of the cleavage reaction in the course of the IVT reaction (shown by the presence of both cleaved and uncleaved populations in denaturing PAGE), indicates that the insertion of the U1A binding site has no inhibitory effect on the ribozyme (data not shown). Prior to crystallization, the formation of the RNA/U1A protein complex was assessed by electrophoretic mobility shift assay (EMSA) (Data not shown). The present crystal structures unravel for the first time the human and the chimpanzee CPEB3 ribozymes in their product state, captured as dimers.

A L3 sequence compatible with the formation of an anticodon-like conformation favours CPEB3 ribozyme dimerization over folding of a catalytically competent conformation

At first glance, the 2D structures of the HDV and CPEB3 ribozymes look very similar. However, the sequence determinants and ribozyme length differ (**Figure 1A and C**). The minimum HDV ribozyme (PDB ID: 4pr6) is about 85 nt (2) versus 67 nt for the CPEB3 RNA. Interestingly, the most important differences are located around the catalytic core: (i) the CPEB3 ribozyme putative P1.1 minihelix has one canonical G-C bp and one U•U mismatch while in the HDV ribozyme this segment consists of two canonical G-C bp; (ii) the 3 nt joining region J1.1/4 does not exist in the CPEB3 ribozymes; and (iii) the 5'-$A_{23}CGU_{26}$-3' stretch of L3 presents 4 nt versus 5 nt (including only one purine) in the genomic HDV ribozyme; (iv) P4 in the CPEB3 ribozyme

is shorter than in the genomic HDV ribozyme. The dimerization of the CPEB3 ribozymes observed in this study seem to result from these sequence and secondary structure changes, which affect ribozyme folding, 3D structure, and cleavage rate (**Figure 2A**).

In the observed conformation of the 7 nt L3 loop, residues U_{21} and C_{22} are splayed apart by the formation of a U-turn, preventing their pairings to their respective targets U_{38} and G_{37} to form the P1.1 pseudoknot (**Figure 2)**. This feature is accompanied by the formation of an new 4 bp helix (P3.1) taking place between the L3 palindromic stretches (5'-$A_{23}CGU_{26}$-3') of two (pseudo)symmetry-related ribozymes, partially corresponding to the loop residues forming the codon-anticodon interaction during ribosome translation. In this unforeseen ribozyme dimer, a competent catalytic core cannot be formed. The dimerization could arise from the competition between a weak intramolecular two bp P1.1 mini-helix (formed by one G-C bp and one weaker U•U mismatch) and a stronger intermolecular four bp helix formed by nucleotides 5'-$A_{23}CGU_{26}$-3' from L3. The P3.1 interaction propagates through all L3 residues adopting a helical conformation, including C_{22}, which forms a loose triple interaction with U_{20} and U_{26} from a symmetry-related molecule. Some contacts are also observed between two U1A proteins, which may contribute to further stabilize the dimer. However, RNA crystal structures obtained with the help of the U1A binding domain usually present U1A-mediated crystal packing contacts that do not interfere with the RNA global fold (37). Consequently, the G_{37} and U_{38} residues from J1/4 supposed to form P1.1 interact instead with those from J4/2, resulting in extending P4 and in J1/4 sequestration. These two non-canonical additional base pairs mediate stacking continuity with P1 (**Figure 2B**), and also prevent the assembly of the catalytic site. Despite CPEB3 ribozyme dimerization, the type I A-minor interaction (41,42) involving A_{62} and the P3 penultimate bp (C_{18}-G_{28}) is observed (corresponding to $A_{77}sC_{18}wcG_{29}$

in HDV (pdb: 1drz)) whereas the type II that would be expected from the catalytic site assembly observed in the HDV ribozyme structure ($A_{76}sC_{19}wcG_{28}$) is dissociated.

The HDV ribozyme catalytic core is flanked by the *cis* G1•U37 and the *trans* G25oU20 wobble pairs (HDV numbering, Figure 1A), which are crucial for ribozyme activity since their mutations significantly reduce cleavage rate. In the crystal structures of the CPEB3 ribozyme the G25oU20 is disrupted by ribozyme dimerization (Figure 2). Another notable difference between all HDV ribozymes is the presence of the three nt J1.1/4 stretch only in the genomic ribozyme (**Figure 1A**) where P4 starts by a non-canonical A_{43}•G_{74} bp, and J1.1/4 adopts a specific turn, which resembles an AA-platform (the sugar and Hoogsteen edges of G_{40} and A_{42}, respectively form hydrogen bonds) (43) with a bulging C inserted. Moreover, the WC edge from G_{40} interacts with the Hoogsteen edge of G_{74}. The area mediated by this triple interaction provides a stacking surface for both P1.1 on one side and P4 on the other. The stacking continuity between the P1, P1.1 and P4 elements thus mainly results from the presence of J1.¼. The absence of J1.¼ raises the question of how P1.1 can be stabilized in the antigenomic HDV and CPEB3 ribozymes beyond the identities of the individual involved nucleotides. In the CPEB3 ribozymes, the stacking continuity between P1 and P4 results from involvement of residues (G_{37}, U_{38}) that should be embedded in P1.1 in a catalytically competent conformation. How these residues could be untrapped to reach a catalytically competent conformation remains to be studied.

A single nucleotide change between the human and chimpanzee CPEB3 ribozymes structures supports their different catalytic properties

A unique sequence difference between the human (A_{30}) and the chimpanzee (G_{30}) homologues induces a local structural adaptation (**Figure 3 and 4**). The putative C•A mismatch

predicted to close P1 in the humCPEB3r does not form. Instead, C_7 is pushed away towards the solvent and its presumed partner A_{30} is in close distance to A_8(O2') (**Figure 4**). Setting up this conformation led to a significant decrease of R_{free} from 0.2870 to 0.2602 and better interpretable electron density maps. The proximity of A_{30} to O2' and possible stacking on G_{31} obviously helps stabilizing J1/2. As C_7 is pushed away, C_6, A_8 and C_9 are stacking one on each other positioning A_{10} to form a standard WC bp with U_{69} (**Figure 4b**). Per contra, the chimpCPEB3-U1Ar conserves the C7-G30 canonical bp, and the backbone makes a smooth curve placing all nucleobases towards P2 (**Figure 4a**). J1/2 appears shorter in the human version of the ribozyme, due to the A_{30}-A_8 base pair, which relegates C_7 to a backward position. Consequently the proximal ends of P1 and P2 are closer in the human than in the chimpanzee ribozyme (d_{chimp}(O3'(A7)-P(A10)=12.2 Å; d_{hum}(O3'(A8)-P(A10)=7.4 Å). The helical continuity of the human P1 is broken by a non-canonical A-A pair, potentially affecting stability. Apart from the difference in P1 helix closure, all 3 ribozyme conformations are very similar, conserving the two parallel coaxial helical stacks described earlier, and the same dimer orientation.

In contrast, the differences observed between the two conformers of the chimpCPEB3r, mainly consisting in the position of P1 relative to the ribozyme, may result from crystal packing constraints.

DISCUSSION

The genomic HDV ribozyme has been the main focus of structural studies over years, and the first crystal structure of the HDV ribozyme was published over two decades ago (1). Here, we present the crystal structures of both the self-cleavage products of the HDV-like human and

the chimpanzee CPEB3 ribozymes. As in the case of the historic HDV crystal structure, the P4 region of both ribozymes were modified to harbour the U1A binding domain, a well-known effective crystallization module (37). Initial secondary structure modeling based on the HDV crystal structure integrated tertiary contacts leading to a similar conformation. Yet, the observed conformations significantly depart from the HDV archetype, which significantly expands the structural data on this ribozyme family.

It has been suggested that the chimpanzee ribozyme is more prone to adopt its native state due to a more stable P1 helix than its human counterpart (16). Indeed, the closure of the chimpCPEB3 P1 stem by a canonical C-G bp may stabilize it better than the human ribozyme where a bulge (C_7) followed by a WC-sugar edge A-A pair takes place (**Figure 1D and 4B**). Consequently, P1 stability and the backbone conformation between C_6 and A_{10} differ between the two CPEB3 homologues. These disparities of the critical J1/2 region may influence ribozyme compaction and further the catalytic activities and rationalize their different catalytic activities.

Although the CPEB3 ribozyme is highly conserved among mammals, and its localization within an intronic sequence remains unchanged since 170 millions years (14), its precise mechanism of action remains unclear. Both *Aplysia* and *Drosophila* CPEB3 homologues (CPEB and ORB2, respectively) have been shown to play a crucial role in neuronal plasticity and their function strongly correlates with their ability to form prion-like oligomers. In humans, the only CPEB protein able to aggregate is CPEB3. Work from Kandel's laboratory has shown that the humCPEB3 protein is necessary for synaptic plasticity by stimulating cytoplasmic polyadenylation-induced translation of plasticity related proteins (PRPs) like AMPA receptor subunits GluA1 and GluA2 (12,13). CPEB3 exists in both soluble and aggregated states depending on SUMOylation, having a regulatory effect on the translation of several mRNAs

(44). However the link between the CPEB3 ribozyme activity, the protein expression level, and the hippocampal-dependent memory remains to be fully elucidated. A recent study from the Luptak laboratory has shown that antisense oligonucleotides (ASO) competing with P1 formation, prevent CPEB3 ribozyme folding and catalysis, leading to increased levels of both the CPEB3 protein and its mRNA (15). The authors conclude that the ribozyme activity influences the expression of CPEB3, which confirmed the original hypothesis from Salehi-Ashtiani et al. (9) that the low background activity is linked to the splicing completion.

The present crystal structures present inhibitory conformations, which gives the opportunity to directly compare the structural roles of the nucleotides according to their identity and position in the CPEB3 and HDV ribozymes. Interestingly, in the present conformers, the folding of the L3 loop in an anticodon-like conformation would be sufficient to repress catalytic activity, with the concomitant sequestration of the J1/4 residues by those from J4/2, including the catalytic C_{57}, which prevents the formation of the P1.1 pseudoknot. Additionally, the anticodon-like conformation is poised to dimerize through formation of a 4 bp double-stranded helix (**Figure 3**). Even though we cannot infer from the structures whether dimerization is biologically relevant, we can assume that the potential to do so would significantly prevent formation of a catalytically active conformation. Two sequence elements seem to drive dimerization, the palindrome (5'-ACGU-3') corresponding to the tRNA anticodon nucleotides 35 to 38 (tRNA numbering). C_{22} upstream from the palindrome (34 in tRNA numbering) contributes a loose triple interaction with the first and the ultimate residues of L3, U_{20} and U_{26} of a symmetry-related RNA. The second important element is U_{21} corresponding to nucleotide 33 (tRNA numbering), which mediates the U-turn typical of tRNAs anticodon loop. This residue is also expected to form the weak UoU pair involved in

the P1.1 pseudoknot. The 4 nt palindrome together with the dual role of U_{21} may favour an inactive conformation of the CPEB3 ribozyme, and further its dimerization.

It has been proposed by the Szostak group that the HDV ribozymes have originated recently from the CPEB3 ribozyme encoded by the human transcriptome (Salehi-Ashtiani, 2006). The HDV ribozymes may thus have evolved faster cleavage properties by drifting from the original CPEB3 ribozyme sequence. To achieve this, the HDV ribozymes may have gradually lost the propensity to fold L3 as an anticodon-like loop prone to dimerize in order to directly favor a catalytically competent conformation without the risk of being kinetically trapped. These drifting events included (i) mutation of the U-turn resulting in strengthening the P1.1 pseudoknot required for building up the catalytic pocket, (ii) mutations breaking down the palindrome. Beyond these changes, the genomic HDV ribozyme intermediates may have acquired (iii) an additional nucleotide in L3 and (iv) a J1.1/4 stretch that should also contribute improved catalytic properties. In between the two extreme points represented by the genomic HDV and the CPEB3 ribozymes seems to stand the antigenomic HDV ribozyme, which still harbours a Y-rich 7 nt L3 loop not expected to dimerize, but has not yet gained the J1.1/4 junction (**Figure 5**).

The present structures indicate that the structural repertoire of the HDV ribozymes is greater than thought earlier. P1.1 can adopt sequence patterns confering poor catalytic properties to the CPEB3 ribozymes. This observation has to be put in the context of the dual roles of the P1.1 residues embedded in L3, which competitively adopts an anticodon-like loop conformation poised to dimerize by means of a palindromic four nt sequence. However, the present results do not rule out whether the dimers observed in the human and chimpanzee CPEB3 ribozymes have any biological relevance. Interesting data from the Been laboratory have shown that replacement of the second P1.1 G-C base pair by a C•C mismatch leads to a

3900-fold activity drop of the genomic HDV ribozyme (45). Since the U•U mismatch conserved in all mammalian CPEB3 ribozymes is theoretically isosteric to a C•C mismatch (46), it has been suggested that both Y-Y base pairs may provoke a dramatic catalytic decrease, indicating a lesser stability of the Y-Y containing P1.1 (45). On the other hand, multimerization of the humCPEB3r has already been reported in two instances by the Bevilacqua laboratory. Strulson et al. (47) have reported detection of humCPEB3r aggregation by PAGE under native conditions in the presence of high molecular weight crowding agents and 10 mM [Mg^{2+}]. Indeed, looking carefully at figure S7 of the study from Strulson (47) suggests that the band interpreted as aggregated CPEB3 ribozymes could correspond to dimers. Chadalavada et al. (16) have also observed the precursor of the CPEB3 by PAGE native conditions and have noted two distinct species they called R1 and R2. These species are separated one from each other on native gels so that they could correspond to dimer and monomer in the light of the present crystal structures (See figure 2 of (16)).

The question we will need to address in the future is whether the present dimeric structures results from a crystallization artefact. A number of studies point to ribozyme dimers as functional entities. In 2017, numerous (±150 000) small hammerhead ribozymes have been discovered characterized by a short (1 bp) stem III and relatively poor cleavage rate (22). These ribozymes become reactive upon forming dimers presenting a recapitulated stem III. The VS ribozyme has also been shown to be active as a dimer in its natural context (21). The Hatchet ribozyme has also been crystallized as a dimer of cleavage products (23). In this case, the 3’ tail of one ribozyme entangles the core of a pseudo-symmetric ribozyme and vice-versa. Strikingly, the sequence, which permits this inter-strand exchange contains the same palindromic sequence 5’-ACGU-3’ than the CPEB3 ribozymes herein (48). All these data point towards the necessity to investigate the biological significance of the inactive CPEB3

ribozymes conformations prone to dimerize to better understand the relationship between CPEB3 ribozyme activity and the expression of the CPEB3 protein.

ACCESSION NUMBERS

Crystal structure models and structure factors were deposited at the wwPDB. PDB IDs are 7qr3 for the chimpanzee and 7qr4 for the human CPEB3 ribozymes.

FUNDING

This work was supported by the fund of Forschungskredit (FK-18-097) of the University of Zürich to A.I.P.M; the MNF to R.K.O.S. B.M. is supported by the Interdisciplinary Thematic Institute IMCBio, as part of the ITI 2021-2028 program at the University of Strasbourg, CNRS and Inserm, by IdEx Unistra (ANR-10-IDEX-0002), and EUR (IMCBio ANR-17-EUR-0023), under the framework of the French Investments Program for the Future. We also thank Dr. Adrian Ferré-D'Amaré for the kind gift of the double mutant-U1A protein plasmid.

CONFLICT OF INTEREST

None

REFERENCES

1. Ferré-D'Amaré, A.R., Zhou, K.H. and Doudna, J.A. (1998) Crystal structure of a hepatitis delta virus ribozyme. *Nature*, **395**, 567-574.
2. Perrotta, A.T. and Been, M.D. (1991) A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature*, **350**, 434-436.
3. Perrotta, A.T., Shih, I. and Been, M.D. (1999) Imidazole rescue of a cytosine mutation in a self-cleaving ribozyme. *Science*, **286**, 123-126.
4. Chen, J.H., Yajima, R., Chadalavada, D.M., Chase, E., Bevilacqua, P.C. and Golden, B.L. (2010) A 1.9 A crystal structure of the HDV ribozyme precleavage suggests both

Lewis acid and general acid mechanisms contribute to phosphodiester cleavage. *Biochemistry*, **49**, 6508-6518.
5. Kapral, G.J., Jain, S., Noeske, J., Doudna, J.A., Richardson, D.C. and Richardson, J.S. (2014) New tools provide a second look at HDV ribozyme structure, dynamics and cleavage. *Nucleic Acids Res*, **42**, 12833-12846.
6. Ke, A., Ding, F., Batchelor, J.D. and Doudna, J.A. (2007) Structural Roles of Monovalent Cations in the HDV Ribozyme. *Structure*, **15**, 281-287.
7. Skilandat, M., Rowinska-Zyrek, M. and Sigel, R.K. (2016) Secondary structure confirmation and localization of Mg2+ ions in the mammalian CPEB3 ribozyme. *Rna*, **22**, 750-763.
8. Sripathi, K.N., Banas, P., Reblova, K., Sponer, J., Otyepka, M. and Walter, N.G. (2015) Wobble pairs of the HDV ribozyme play specific roles in stabilization of active site dynamics. *Phys Chem Chem Phys*, **17**, 5887-5900.
9. Salehi-Ashtiani, K., Luptak, A., Litovchick, A. and Szostak, J.W. (2006) A Genomewide Search for Ribozymes Reveals an HDV-Like Sequence in the Human CPEB3 Gene. *Science*, **313**, 1788-1792.
10. Webb, C.H. and Luptak, A. (2011) HDV-like self-cleaving ribozymes. *RNA biology*, **8**, 719-727.
11. Stephan, J.S., Fioriti, L., Lamba, N., Colnaghi, L., Karl, K., Derkatch, I.L. and Kandel, E.R. (2015) The CPEB3 Protein Is a Functional Prion that Interacts with the Actin Cytoskeleton. *Cell reports*, **11**, 1772-1785.
12. Fioriti, L., Myers, C., Huang, Y.Y., Li, X., Stephan, J.S., Trifilieff, P., Colnaghi, L., Kosmidis, S., Drisaldi, B., Pavlopoulos, E. *et al.* (2015) The Persistence of Hippocampal-Based Memory Requires Protein Synthesis Mediated by the Prion-like Protein CPEB3. *Neuron*, **86**, 1433-1448.
13. Ford, L., Ling, E., Kandel, E.R. and Fioriti, L. (2019) CPEB3 inhibits translation of mRNA targets by localizing them to P bodies. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 18078-18087.
14. Bendixsen, D.P., Pollock, T.B., Peri, G. and Hayden, E.J. (2021) Experimental Resurrection of Ancestral Mammalian CPEB3 Ribozymes Reveals Deep Functional Conservation. *Mol Biol Evol*, **38**, 2843-2853.
15. Chen, C.C., Han, J., Chinn, C.A., Li, X., Nikan, M., Myszka, M., Tong, L., Bredy, T.W., Wood, M.A. and Lupták, A. (2021) The CPEB3 ribozyme modulates hippocampal-dependent memory. *bioRxiv*, 2021.2001.2023.426448.
16. Chadalavada, D.M., Gratton, E.A. and Bevilacqua, P.C. (2010) The human HDV-like CPEB3 ribozyme is intrinsically fast-reacting. *Biochemistry*, **49**, 5321-5330.
17. Holbrook, S.R., Cheong, C., Tinoco, I. and Kim, S.H. (1991) Crystal structure of an RNA double-helix incorporating a track of non-Watson-Crick base pairs. *Nature.*, **353**, 579-581.
18. Baeyens, K.J., De Bondt, H.L. and Holbrook, S.R. (1995) Structure of an RNA double helix including uracil-uracil base pairs in an internal loop. *Nature Struct. Biology*, **2**, 56-62.
19. Baeyens, K.J., Debondt, H.L., Pardi, A. and Holbrook, S.R. (1996) A curved RNA helix incorporating an internal loop with G . A and A . A non-Watson-Crick base pairing. *Proc. Natl. Acad. Sci. USA.*, **93**, 12851-12855.
20. Lietzke, S., Barnes, C.L., Berglund, J.A. and Kundrot, C.E. (1996) The structure of an RNA dodecamer shows how tandem U-U base-pairs increase the range of stable RNA structures and the diversity of recognition sites. *Structure.*, **4**, 917-930.
21. Suslov, N.B., DasGupta, S., Huang, H., Fuller, J.R., Lilley, D.M., Rice, P.A. and Piccirilli, J.A. (2015) Crystal structure of the Varkud satellite ribozyme. *Nat Chem Biol*, **11**, 840-846.
22. Lunse, C.E., Weinberg, Z. and Breaker, R.R. (2017) Numerous small hammerhead ribozyme variants associated with Penelope-like retrotransposons cleave RNA as dimers. *RNA biology*, **14**, 1499-1507.

23. Zheng, L., Falschlunger, C., Huang, K., Mairhofer, E., Yuan, S., Wang, J., Patel, D.J., Micura, R. and Ren, A. (2019) Hatchet ribozyme structure and implications for cleavage mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 10783-10791.
24. Bakthavachalu, B., Huelsmeier, J., Sudhakaran, I.P., Hillebrand, J., Singh, A., Petrauskas, A., Thiagarajan, D., Sankaranarayanan, M., Mizoue, L., Anderson, E.N. *et al.* (2018) RNP-Granule Assembly via Ataxin-2 Disordered Domains Is Required for Long-Term Memory and Neurodegeneration. *Neuron*, **98**, 754-766 e754.
25. Gallo, S., Furler, M. and Sigel, R. (2005) In vitro Transcription and Purification of RNAs of Different Size. *CHIMIA International Journal for Chemistry*, **59**, 812-816.
26. Ferré-D'Amaré, A.R., Zhou, K.H. and Doudna, J.A. (1998) A general module for RNA crystallization. *Journal of molecular biology*, **279**, 621-631.
27. Evans, P.R. and Murshudov, G.N. (2013) How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography*, **69**, 1204-1214.
28. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C. *et al.* (2011). Global Phasing Ltd, United Kingdom.
29. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J Appl Crystallogr*, **40**, 658-674.
30. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, **66**, 213-221.
31. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, **66**, 486-501.
32. Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S.M. and Bricogne, G. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta crystallographica. Section D, Biological crystallography*, **60**, 2210-2221.
33. Davis, I.W., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2004) MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res*, **32**, W615-619.
34. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
35. Reiter, N.J., Osterman, A., Torres-Larios, A., Swinger, K.K., Pan, T. and Mondragon, A. (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*, **468**, 784-789.
36. Costa, M., Walbott, H., Monachello, D., Westhof, E. and Michel, F. (2016) Crystal structures of a group II intron lariat primed for reverse splicing. *Science*, **354**, aaf9258.
37. Ferre-D'Amare, A.R. (2010) Use of the spliceosomal protein U1A to facilitate crystallization and structure determination of complex RNAs. *Methods*, **52**, 159-167.
38. Cochrane, J.C., Lipchock, S.V. and Strobel, S.A. (2007) Structural investigation of the GlmS ribozyme bound to Its catalytic cofactor. *Chemistry & biology*, **14**, 97-105.
39. Kulshina, N., Edwards, T.E. and Ferre-D'Amare, A.R. (2010) Thermodynamic analysis of ligand binding and ligand binding-induced tertiary structure formation by the thiamine pyrophosphate riboswitch. *Rna*, **16**, 186-196.
40. Rupert, P.B. and Ferre-D'Amare, A.R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*, **410**, 780-786.
41. Doherty, E.A., Batey, R.T., Masquida, B. and Doudna, J.A. (2001) A universal mode of helix packing in RNA. *Nature structural biology*, **8**, 339-343.
42. Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B. and Steitz, T.A. (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 4899-4903.

43. Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Szewczak, A.A., Kundrot, C.E., Cech, T.R. and Doudna, J.A. (1996) RNA tertiary structure mediation by adenosine platforms. *Science.*, **273**, 1696-1699.
44. Drisaldi, B., Colnaghi, L., Fioriti, L., Rao, N., Myers, C., Snyder, A.M., Metzger, D.J., Tarasoff, J., Konstantinov, E., Fraser, P.E. *et al.* (2015) SUMOylation Is an Inhibitory Constraint that Regulates the Prion-like Aggregation and Activity of CPEB3. *Cell reports*, **11**, 1694-1702.
45. Wadkins, T.S., Perrotta, A.T., Ferre-D'Amare, A.R., Doudna, J.A. and Been, M.D. (1999) A nested double pseudoknot is required for self-cleavage activity of both the genomic and antigenomic hepatitis delta virus ribozymes. *Rna*, **5**, 720-727.
46. Stombaugh, J., Zirbel, C.L., Westhof, E. and Leontis, N.B. (2009) Frequency and isostericity of RNA base pairs. *Nucleic Acids Res*, **37**, 2294-2312.
47. Strulson, C.A., Yennawar, N.H., Rambo, R.P. and Bevilacqua, P.C. (2013) Molecular crowding favors reactivity of a human ribozyme under physiological ionic conditions. *Biochemistry*, **52**, 8187-8197.
48. Bou-Nader, C. and Zhang, J. (2020) Structural Insights into RNA Dimerization: Motifs, Interfaces and Functions. *Molecules*, **25**.
49. Leontis, N.B. and Westhof, E. (2001) Geometric nomenclature and classification of RNA base pairs. *Rna*, **7**, 499-512.
50. Shi, H. and Moore, P.B. (2000) The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: a classic structure revisited. *Rna*, **6**, 1091-1105.

TABLE AND FIGURE LEGENDS

Table 1

Data collection	**Human CPEB3 ribozyme**	**Chimpanzee CPEB3 ribozyme**
Space group	$C222_1$	$C222_1$
Unit cell parameters *(a; b; c (Å))*	79.451; 131.879; 90.456	120.254; 135.976; 83.254
Resolution range *(Å)*	68.055-2.825 (3.124-2.825)[a]	67.958-2.184 (2.355-2.184)[a]
Total no. of reflections	87407 (4303)[a]	374221 (19001)[a]
Unique reflections	6828 (342)[a]	28178 (1410)[a]
Completeness *(%) (ellipsoidal)*	92.6 (73.7)[a]	93.7 (57.0)[a]
R_{merge} *(all I^+ and I^-)*	0.081 (1.998)[a]	0.064 (1.904)[a]
R_{meas} *(all I^+ and I^-)*	0.085 (2.083)[a]	0.067 (1.985)[a]
Multiplicity	12.8	13.3
MeanI/σI	16.7	24.9
$CC_{1/2}$	0.998	1.000
Models		
PDB number	7qr4	7qr3
R/R_{free}	0.220/0.268	0.241/0.277

No. Reflections used to R_{free}	348	1431
Wilson *B* factor	93.31	57.16
Number of atoms RNA Protein	 1461 (chain B) 728 (chain A)	 1462/1462 (chains C/D) 738/742 (chains A/B)
Rmsd from ideal values Bond lengths (Å) Bond angles (°) Dihedral angles (°)	 0.003 0.615 17.250	 0.002 0.633 13.880

Table 1. Data collection and model refinement statistics. [a] Highest resolution shell

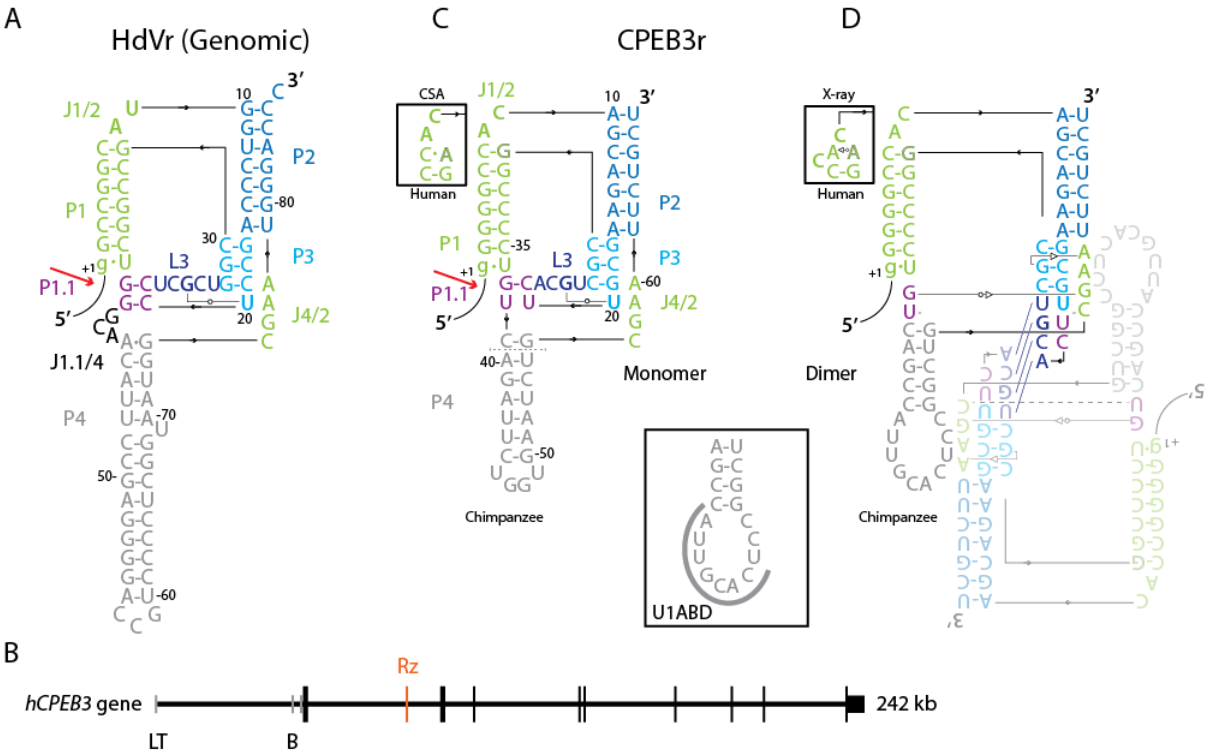

Figure 1. Secondary structure of (**A**) the genomic HDV ribozyme with individually colored secondary structure elements. The red arrow indicates the cleavage site. (**B**) Mapping of the human CPEB3 gene. Horizontal lines (black) correspond to introns. Tissue specific exons are shown by gray vertical lines (L stands for liver, T for testis and B for brain). Translated exons are shown by black vertical lines (black). The CPEB3 ribozyme (Rz) location within the second intron is marked by an orange vertical line (adapted from Salehi-Ashtiani (9)). **C**) The secondary structure model of the CPEB3 wild-type ribozyme as deduced from the HDV ribozyme crystal structure suggests the presence of a weak P1.1 pseudoknot containing a G-C pair and a putative U-U base pair deduced by comparison to the genomic HDV ribozyme crystal structure. The single nucleotide difference between human (A_{30}) and chimpanzee (G_{30}) CPEB3 ribozymes is shown in the upper inset, which shows the secondary structure of the human J1/2 region deduced from comparative sequence analysis (CSA). Non-canonical base pairs are highlighted using symbols from the Leontis-Westhof nomenclature (49). The replacement of P4/L4 by the U1A-RDB used for crystallographic purposes is shown in the lower inset and takes place right below the dashed line shown on the WT model. The secondary structure model of the humanCPEB3 ribozyme J1/2 as deduced from the crystal structure is shown in the upper left corner inset. The connectivity between nucleotides is drawn using black arrows. The leader sequence preceding the cleavage site is

represented by a black curve. P stands for base-Paired regions; L for loops and J for single-stranded joining regions. The catalytically important C residue is the first from J4/2 (C_{57}). **D**) Schematic representation of the CPEB3 ribozyme dimer as interpreted from the crystal structures. The upper inset indicates the actual secondary structure taking place in the human J1/2 region.

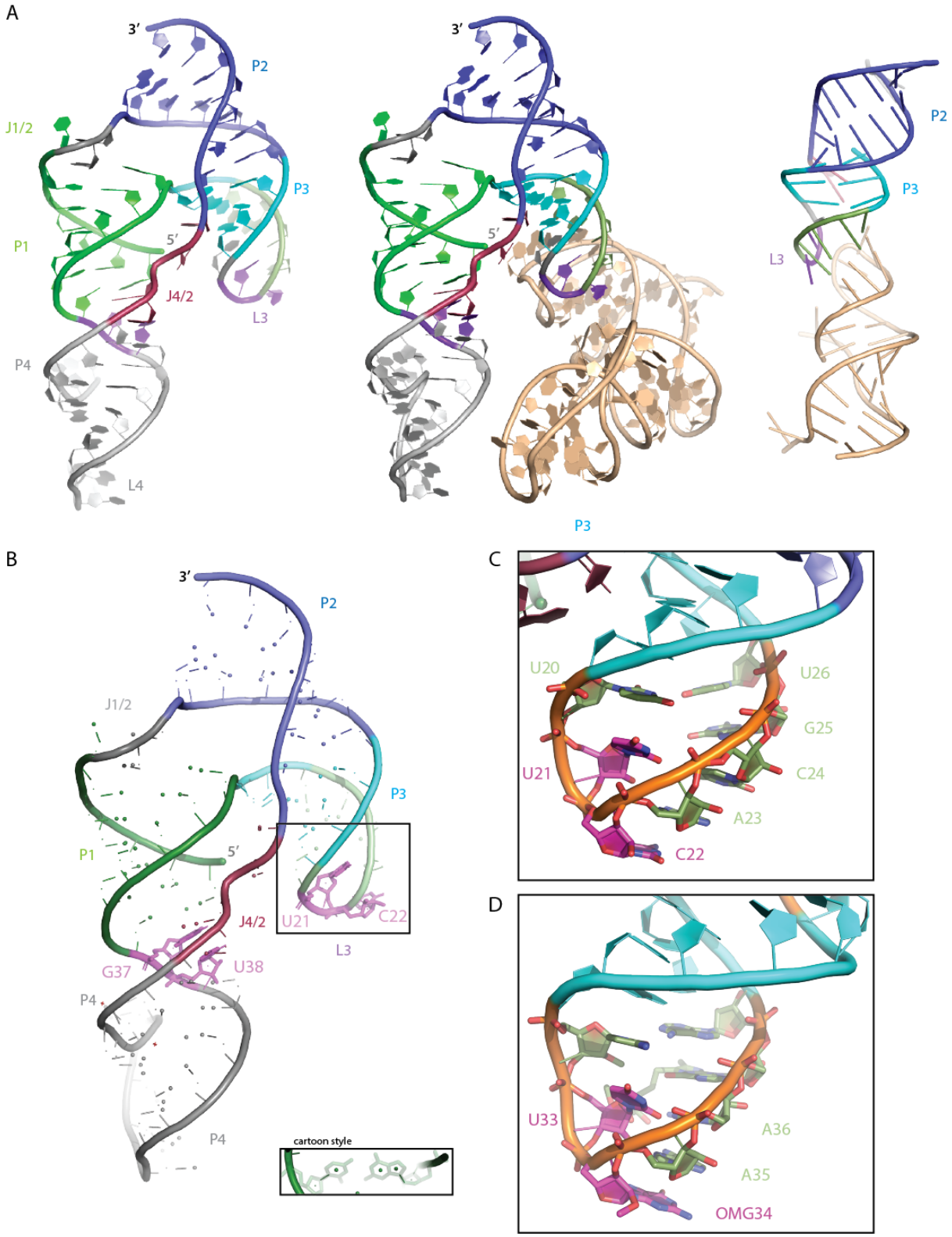

Figure 2. Crystal structure of the human (left) and the chimpanzee (right) CPEB3-U1A RBD ribozymes. (A) The humCPEB3-U1A ribozyme (left panel) is a crystallographic dimer (central panel) while the chimpCPEB3-U1A is pseudo-symmetric. The dimerization site (right panel) comprises nucleotides A23, C24, G25 and U26 from the L3 loop. The U1A protein binding L4 is omitted for clarity. (B) The residues

predicted to form the nested pseudoknot P1.1 (purple) are highlighted in sticks on the cartoon 3D backbone spline. The mode of representation is illustrated in the inset entitled “Cartoon style”. (C)shows a close-up of L3 highlighting residues supposed to form the nested pseudoknot (U21, C22, purple) and the five residues involved in dimerization (C22, A23, C24, G25 and U26, green). The nucleobases of U20 and U26 are coplanar but could only form water-bridging hydrogen bonds deduced from the distance in between the two proximal carbonyl groups. A23 through U26 adopt a helical conformation. U21 adopts a classical U-turn conformation and its N3 contacts the O2P oxygen of C24. (D) The anticodon loop of tRNAPhe (pdb: 1ehz (50)) is depicted for a direct visual comparison, which indicates that each nucleotide conformation is similar to the one observed in L3 from the CPEB3 ribozyme.

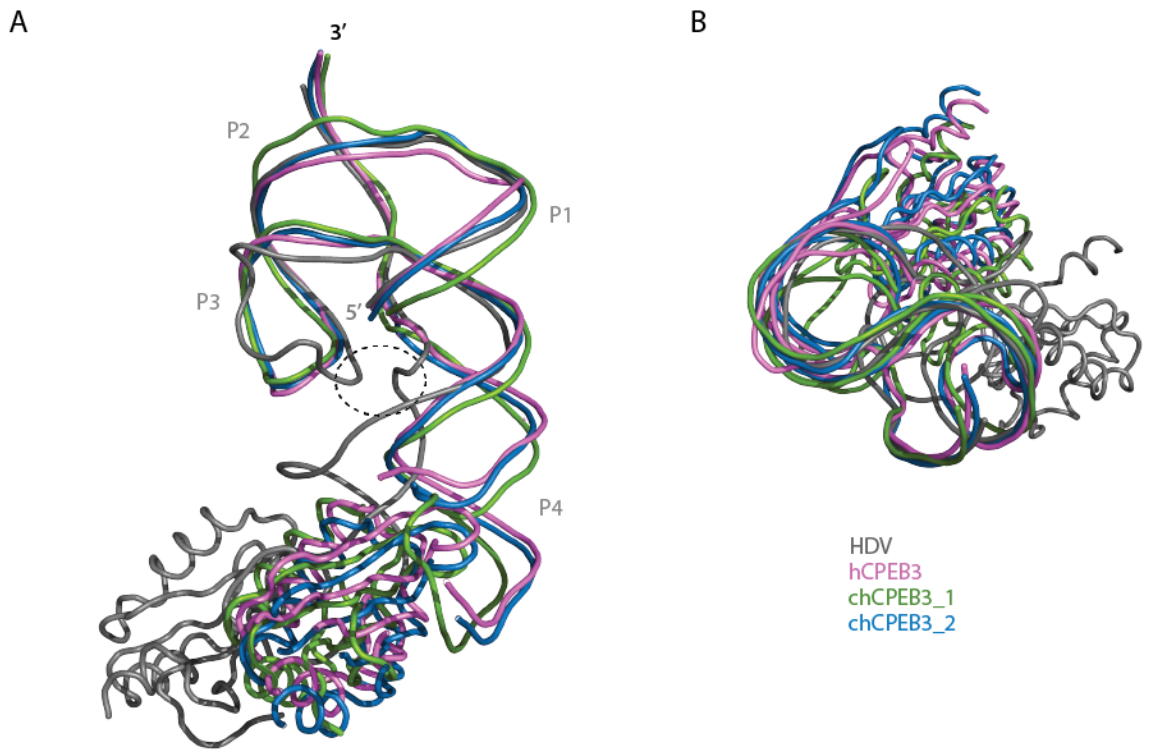

Figure 3. Structural similarity between Reorganization of the CPEB3 ribozymes and HDV ribozyme structures. Alignment of the genomic HDV (PDB ID: 4PR6), of the humCPEB3-U1A ribozyme (PDB ID: 7qr4) and of the chimpCPEB3-U1A (PDB ID: 7qr3) ribozymes. Both conformations of the chimpCPEB3-U1A ribozymes are represented. The HDV ribozyme is shown in gray, the humCPEB3-U1Ar in violet and both chimpCPEB3-U1Ar conformations in green and blue. The dashed ellipsoid delineates the shorter distance taking place in the genomic HDV ribozyme structure due to the formation of P1.1.

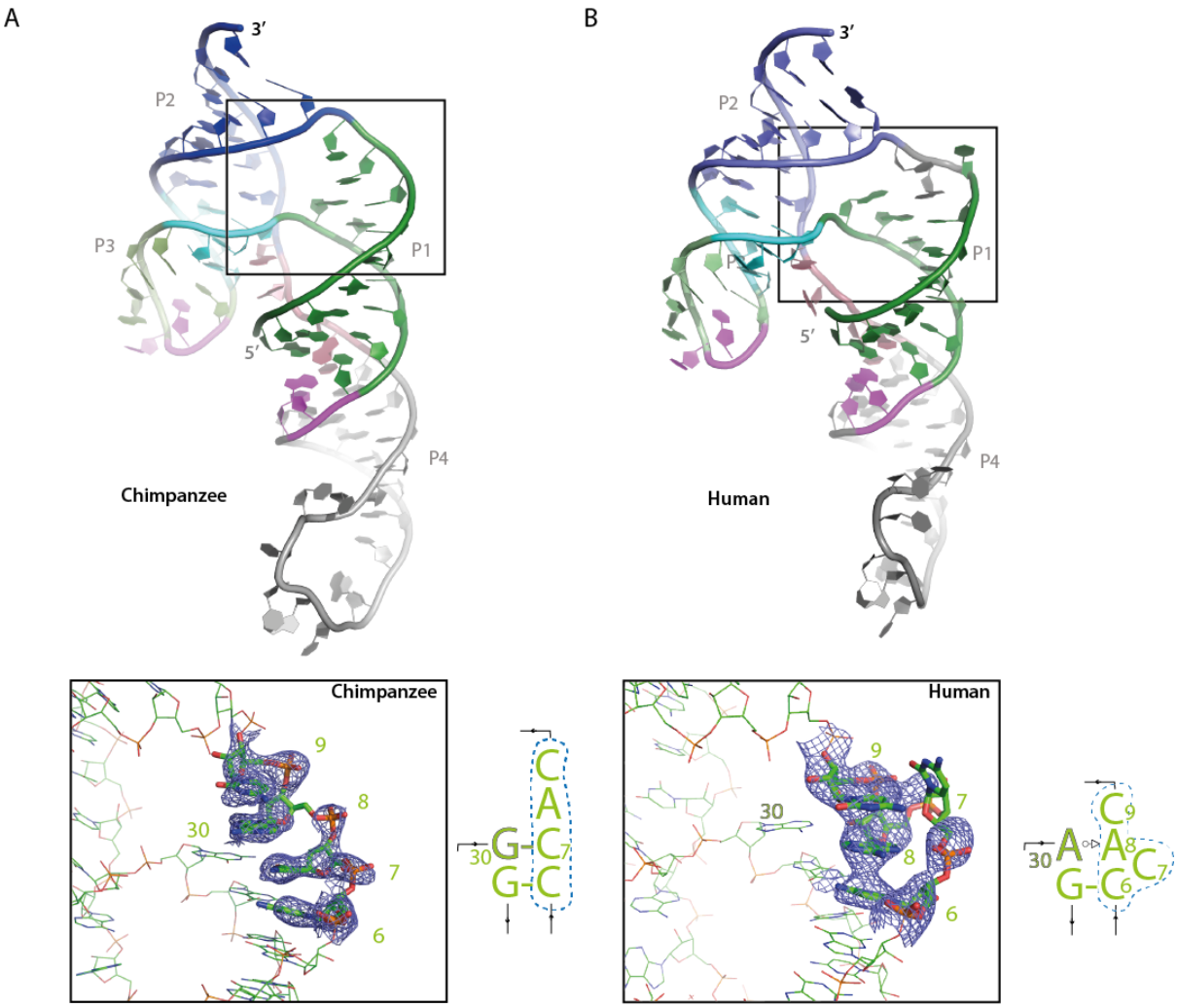

Figure 4. Structural differences between human and chimpanzee CPEB3 ribozymes around J1/2, at the P1/ P2 strand cross-over. Both chimpanzee (A) and human (B) CPEB3 ribozymes present the same J1/2 sequences, although they are structured differently due to a sequence difference at position 30 (outlined nucleobase). A30w_osA base pair instead of a G30-C base pair. Consequently, residues 6 to 9 form a stacked column in the chimpanzee ribozyme whereas in the human ribozyme, only 3 residues can stack with C7 bulging out. The distance between the proximal extremities of P1 and P2 is thus reduced to less than 10 Å in the human ribozyme. The weighted 2Fobs-Fcalc electron density maps are drawn contoured at 1.0 sigma. The corresponding region is highlighted by a blue dashed frame on the secondary structure diagrams corresponding to each situation.

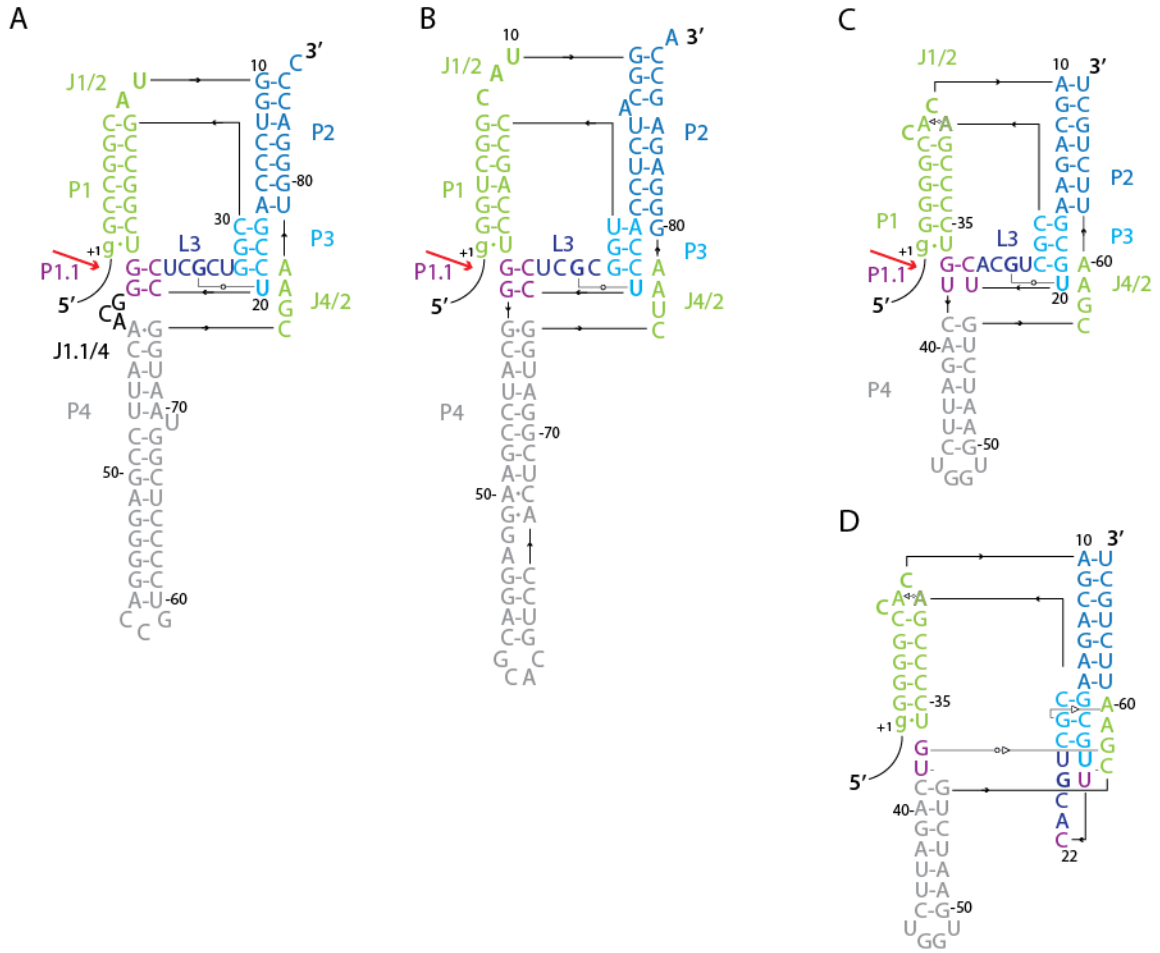

Figure 5. Secondary structures of the HDV and CPEB3 ribozymes. The differences between these ribozymes mostly reside in L3 (deep blue), P1.1 (purple) and J1.1/4 (black). (A) The genomic HDV ribozyme presents an 8-membered loop not expected to adopt an anticodon conformation. P1.1 is composed of two GC pairs with a C_{21} residue not supporting the appearance of a U-turn. The J1.1/4 junction supports formation of P1.1. (B) The antigenomic HDV ribozyme has no palindrome and is Y-rich. L3 has still 7 residues. Yet the presence of C_{21} is not expected to support the U-turn conformation of L3. (C and D) The CPB3 ribozyme presents a seven membered L3 loop containing a four nt palindrome 5'-ACGU-3'. U21 supports the formation of a U-turn inducing competition between the formation of an antidcodon-like conformation for L3 (D) and closure of P1.1 to drive the ribozyme towards catalysis (C).