Published online 3 July 2009

Nucleic Acids Research, 2009, Vol. 37, No. 17 ***5793–5802***
doi:10.1093/nar/gkp537

Crystal structure and assembly of the functional *Nanoarchaeum equitans* tRNA splicing endonuclease

Michelle Mitchell[1], Song Xue[1], Rachel Erdman[1], Lennart Randau[2], Dieter Söll[2,3] and Hong Li[1,*]

[1]Department of Chemistry and Biochemistry, Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, [2]Department of Molecular Biophysics and Biochemistry and [3]Department of Chemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA

Received May 13, 2009; Revised June 5, 2009; Accepted June 8, 2009

ABSTRACT

The RNA splicing and processing endonuclease from *Nanoarchaeum equitans* (NEQ) belongs to the recently identified $(\alpha\beta)_2$ family of splicing endonucleases that require two different subunits for splicing activity. *N. equitans* splicing endonuclease comprises the catalytic subunit (NEQ205) and the structural subunit (NEQ261). Here, we report the crystal structure of the functional NEQ enzyme at 2.1 Å containing both subunits, as well as that of the NEQ261 subunit alone at 2.2 Å. The functional enzyme resembles previously known α_2 and α_4 endonucleases but forms a heterotetramer: a dimer of two heterodimers of the catalytic subunit (NEQ205) and the structural subunit (NEQ261). Surprisingly, NEQ261 alone forms a homodimer, similar to the previously known homodimer of the catalytic subunit. The homodimers of isolated subunits are inhibitory to heterodimerization as illustrated by a covalently linked catalytic homodimer that had no RNA cleavage activity upon mixing with the structural subunit. Detailed structural comparison reveals a more favorable hetero- than homodimerization interface, thereby suggesting a possible regulation mechanism of enzyme assembly through available subunits. Finally, the uniquely flexible active site of the NEQ endonuclease provides a possible explanation for its broader substrate specificity.

INTRODUCTION

In all domains of life, some tRNAs contain introns that must be removed to reveal the mature and functional structure. In bacteria, the intron removal mechanism is autocatalytic, carried out by the self-splicing Group I introns (1,2). In archaeal RNA and eukaryal nuclear tRNA genes, intron removal is carried out by the stepwise action of the splicing endonuclease, ligase and, in some cases, the 2′-phosphotransferase (3,4). Approximately 5% of tRNAs contain introns that must be removed (5–12). Recently, a related function of the tRNA splicing endonuclease was found in two archaea, one of which is the Nanoarchaeota *Nanoarchaeum equitans*, in which it processes precursors of split tRNA genes leading to the formation of functional tRNA (13–15). RNA processing is a fundamental process in a variety of cellular events. Importantly, missense mutations in three of the four human splicing endonuclease subunits (Sen2, Sen34 and Sen54) were recently identified to be linked to two variants of the human genetic disease Pontocerebellar hypoplasia (16). The mutations in the catalytic subunits are thought to interfere with the subunit and domain interaction and those in the structural subunit result in a reduced level of the functional subunit.

The RNA substrates recognized by the tRNA splicing endonuclease vary widely in structure and locations. Introns are found predominantly in the anticodon stem of the precursor tRNA molecule (17–19), invariably between positions 37 and 38 in eukaryotes, but at various positions in archaea (10). In split tRNA genes, two or three separate tRNA fragments are joined by base pairing to form sites similarly located in the anticodon stem that are recognized and processed by the tRNA splicing endonuclease (13–15). In eukaryotic tRNA, the location of the intron relative to the mature domain is conserved and in fact is essential for the recognition of the splice sites by the splicing endonuclease, in a process called the ruler mechanism (20–22). The archaeal splicing endonuclease differs from its eukaryotic counterpart in that it does not require the mature domain for splice site recognition, rather it uses a specific RNA secondary structure, the

*To whom correspondence should be addressed. Tel: +1 850 644 6785; Fax: +1 850 644 7244; Email: hong.li@fsu.edu

bulge–helix–bulge (BHB) motif, the canonical form of which is made up of two three-nucleotide bulges separated by a 4-bp helix (the 3–4–3 BHB) (Figure 1B) (10,23). The simple and portable feature of the BHB motif helps to explain the prevalence of tRNA splicing endonuclease-facilitated RNA processing events in archaea.

Recognition of the RNA substrates by the splicing endonuclease depends on its subunit composition. Splicing endonucleases are classified into four families according to the number and arrangement of subunits. The earliest studies of splicing endonuclease were carried out in yeast, and it was determined that the yeast splicing endonuclease comprises four distinct subunits in an $\alpha\beta\gamma\delta$ arrangement, with two subunits serving catalytic roles and two subunits serving structural roles (24). The work done with archaeal splicing endonucleases focused on two families found in Euryarchaeota, the α_4, which includes *Methanocaldococcus jannaschii* (MJA), and the α_2, which includes *Archaeoglobus fulgidus* (AFU), that assemble as a homotetramer or a homodimer, respectively (25,26). More recently, a fourth family has been characterized, the $(\alpha\beta)_2$ family, which contains members from Crenarchaeota and *N. equitans*. The $(\alpha\beta)_2$ family of endonucleases differs from the other two archaeal families in that it requires a second subunit of weak homology (13,27,28). The $(\alpha\beta)_2$ is thought to have arisen from gene duplication as an example of subfunctionalization, resulting in a second gene coding for the structural subunit, one with significantly less sequence homology but retaining conserved sequences required for its role in the assembly of the enzyme (27). Even though there are differences in the individual subunit makeup between families, the catalytic subunit is highly conserved and a 'four-subunit' architecture, as well as the means to position the catalytic subunits for recognition and cleavage of the splice sites, is believed to be universally conserved.

Intriguingly, a clear difference exists between the $(\alpha\beta)_2$ and the other archaeal splicing endonucleases in their ability to accommodate deviations from the canonical BHB motif. The natural substrates found in Crenarchaeota including organisms such as *Sulfolobus solfataricus* (SSO) and *S. tokodaii* (STO), as well as *N. equitans* (NEQ), the only sequenced member of the Nanoarchaeota, contain a significant number of non-canonical BHB motifs with slightly varied bulge structures (Figure 1) (10,29). Deviations from the canonical 3–4–3 BHB are believed to be linked with unique architectural features of the $(\alpha\beta)_2$ family of splicing endonucleases. *In vitro* splicing experiments with reconstituted recombinant splicing endonucleases support this hypothesis (13,27,28). While the homotetrameric endonucleases are only able to cleave those substrates containing canonical BHB motifs, the heterotetrameric, and to a lesser extent the homodimeric variant are able to cleave more relaxed versions of the BHB motif (Figure 1) (13,27,28). This would follow the supposition that the tRNA substrates coevolved with the splicing endonucleases (29).

To understand the process of the tRNA splicing endonuclease assembly, we determined the crystal structure of the functional splicing endonuclease and that of the structural subunit from *N. equitans*. The structural and complementary biochemical analyses confirm the hypothesized architecture of the $(\alpha\beta)_2$ endonuclease and further suggest a regulation mechanism of enzyme assembly through available subunits.

MATERIALS AND METHODS

Protein expression and purification

The NEQ205 and NEQ261 genes were cloned into the pET-Duet vector that allows the production of proteins with a hexahistidine tag. Residues L135 and L137 of NEQ205 were mutated to methionine by site-directed mutagenesis in preparation for structure determination by a selenomethionine anomalous method and is referred throughout as NEQ205. This modification did not affect RNA cleavage activity (Figure 1A). BL21 (DE3) *Escherichia coli* cells were transformed with the pET-Duet vector containing genes encoding NEQ205 and NEQ261 for the coexpression of both subunits. The cells were grown at 37°C until 0.6 OD_{595} before induction with 750 μM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Induced cell culture was incubated for 16 h at 20°C and was harvested by centrifugation. The cells were lysed in a buffer containing 0.03 M Tris pH 7.5, 1.0 M NaCl, 5% glycerol and 5 mM β-mercaptoethanol (βME), and the NEQ205–NEQ261 complex was purified using a Nickel–nitrilotriacetic acid (Ni-NTA) affinity column followed by a gel filtration procedure. Eluted fractions were pooled and concentrated. RNA splicing activity was used to confirm the function of the complex (30).

The gene encoding NEQ261 was subcloned into pET28a with a fused N-terminal hexahistidine tag. NEQ261 was purified using a similar procedure as that for the NEQ205–NEQ261 complex.

To produce the covalently linked SSO catalytic dimer, primers were designed for the stepwise PCR amplification of the tandem *SSO043* gene duplication linked by a sequence expressing a 10 amino acid linker. This gene construct (Linker10) was then cloned into a pET21a vector with an N-terminal hexahistidine tag for overexpression in BL21-Codon + (DE3) *E. coli* cells. Additionally, primers were designed to reduce the length of the linker peptide by three residues (Linker7) using a Quikchange procedure. Both linker mutant proteins were purified using a similar protocol to that for the NEQ205–NEQ261 complex. The production and purification of the covalently linked subunits were confirmed by SDS–PAGE.

NEQ205–NEQ261 cleavage assay

N. equitans precursor $tRNA^{Glu}$ transcript was produced and 3′-radiolabeled as described (28). A 200 nM concentration of this transcript was incubated with 1 μM NEQ205–NEQ261 splicing endonuclease at 65°C for 20 min. In some reactions, the mixture was cooled down for 3 min on ice followed by incubation with 10 units of T4 polynucleotide kinase and 20 units of T4 RNA ligase (New England Biolabs). Reaction products were analyzed by electrophoresis on denaturing 12% polyacrylamide gel. The mature tRNA product was eluted from the gel in 100 mM sodium acetate of pH 6.4 and precipitated and

Table 1. Data collection and refinement statistics (values in parentheses refer to those of the highest resolution shell)

Diffraction data	NEQ205–NEQ261	NEQ261
Space group	$I4_1$	$P3_221$
Unit-cell parameters (Å)		
a	124.10	95.85
b	124.10	95.85
c	69.82	47.85
Resolution range (Å)	50–2.1 (2.18–2.10)	100–2.2 (2.28–2.20)
No. of unique reflections	25 031 (1602)	13 062 (1311)
Redundancy	5.5 (4.2)	19.0 (11.5)
Completeness (%)	81.6 (48.3)	100 (99.9)
$I/\sigma(I)$	44.2 (3.86)	60.1 (4.9)
R_{sym} (%)	11.0 (50.1)	9.7 (62.4)
Refinement statistics		
Resolution range (Å)	43.9–2.1 (2.16–2.11)	100–2.2 (2.28–2.20)
R_{work} (%)	23.2 (30.5)	25.4 (35.8)
R_{free} (%)	25.9 (33.9)	27.2 (40.7)
Model information		
No. of protein monomers	1	1
No. of amino-acid residues	288	149
No. of protein atoms	2437	1270
R.m.s. deviations of the model		
Bond length (Å)	0.011	0.012
Bond angle (°)	1.204	1.401
Average B-factors		
B, bonded main chain (Å^2)	57.6	46.8
B, bonded side chain (Å^2)	60.2	49.7
Ramachandran plot		
Residues in the most favored region	233 (91.7%)	122 (92.4%)
Residues in additionally allowed region	19 (7.5%)	10 (7.6%)
Residues in the generously allowed region	1 (0.4%)	0 (0.0%)
Residues in the disallowed region	1 (0.4%)	0 (0.0%)

resuspended in water. RT-PCR and sequencing were performed as described (14) using the following oligonucleotides: Forward: 5′-GCCCCCGTGGTGTAGCC aGGTctAGCATAC-3′ and Reverse: 5′-TGCCCCCGCC GGGATTTGAACC-3′.

Crystallization and structure determination of the functional NEQ endonuclease

NEQ functional protein was crystallized at 60 mg/ml concentration at 30°C using the hanging drop vapor diffusion method. Drops were set up with 1.5 µl of a protein solution and equal amount of a well solution that contained 44% Tacsimate. Crystals were obtained after 3 days and could be directly frozen without additional cryoprotectant. Data collection was done at Argonne National Laboratory APS SER-CAT Beamline 22ID (Argonne, IL). Data statistics are reported in Table 1. The crystal structure was solved by molecular replacement (31) using a polyalanine model of the MJA monomer as the search probe. The chain was built in COOT (32) and was refined using REFMAC5 (33) and CNS (34).

Crystallization and structure determination of the NEQ261 subunit

NEQ261 was crystallized by the hanging drop vapor diffusion method at a concentration of 30 mg/ml at 30°C. Drops were set up with 1.5 µl of a protein solution and equal amount of well solution that contained 0.1 M sodium acetate (pH 4.6), 0.30 M ammonium acetate and 35% PEG 4000. Crystals were obtained after 5 days. The crystals were transferred to a cryoprotectant solution containing the mother liquor plus 10% glycerol before being flash frozen in a liquid nitrogen stream. Data collection was done at Argonne National Laboratory APS SER-CAT Beamline 22BM (Argonne, IL). Data statistics are reported in Table 1. The structure was solved by molecular replacement using the NEQ261 from the functional structure as the search model and was refined to 2.2 Å using REFMAC5 (33) and CNS (34). The crystal structure coordinates have been deposited with RCSB Protein Data Bank under IDs 3IEY and 3IF0.

Linker mutants cleavage assay

In vitro RNA cleavage assay was carried out as previously described (28). Briefly, a synthetic 29-mer or 21-mer oligo was [α-^{32}P]-labeled by T4 kinase (Stratagene) in the presence of trace amount of [α-^{32}P]-ATP and was annealed with a cold complementary oligo to form the canonical BHB motif. The cleavage reaction was carried out by combining 2 nM annealed RNA and 10 µM proteins in a cleavage buffer that contains 0.04 M Tris (pH 7.4), 0.0025 M $MgCl_2$, 0.001 M EDTA and RNase inhibitor (New England Biolabs). Reactions were incubated at 60°C

for 30 min and products were separated on a 20% denaturing polyacrylamide gel. Bands were visualized by phosphorimaging with Molecular Dynamics Storm 860 Scanner and Image Quant analysis software (GE Healthcare).

RESULTS AND DISCUSSION

Overall structure

The functional *N. equitans* splicing endonuclease and the NEQ261 subunit were purified following a previous strategy with some modifications (13). Functionality of the NEQ205–NEQ261 splicing endonuclease was verified by the cleavage assay using the previously described *N. equitans* precursor $tRNA^{Glu}$ (Figure 1A) (28). This tRNA contains a relaxed BHB motif containing one 3-nt and one 4-nt bulge (Figure 1B). The site of cleavage was further verified by sequencing the ligated splicing products. Briefly, a product was generated that matched the size of the mature $tRNA^{Glu}$ (78 nt). The corresponding band was excised from the gel, amplified and sequenced. Our sequencing results showed that the standard, mature tRNA was produced, thereby confirming the specificity of the splicing endonuclease with relaxed BHB motifs (Figure 1C).

Crystals of both NEQ205–NEQ261 complex and NEQ261 alone were obtained by the standard vapor diffusion method. The NEQ205–NEQ261 complex was crystallized using 44% Tacsimate as precipitant and NEQ261 was crystallized using 35% PEG 4000 as precipitant (see 'Materials and Methods' section for details). The NEQ205–NEQ261 complex crystals were in space group $I4_1$ with cell dimensions $a = 124.102$, $b = 124.102$ and $c = 69.820$, with one complex in the asymmetric unit. The NEQ261 crystals were in space group $P3_221$ with unit cell dimensions of $a = 95.434$, $b = 95.434$ and $c = 47.791$ and one NEQ261 in the asymmetric unit. The NEQ205–NEQ261 complex structure was determined first by molecular replacement methods using the MJA monomer structure as the search probe. The NEQ261 structure was then determined by molecular replacement using its coordinate from the NEQ205–NEQ261 complex structure. The structures were refined to a resolution of 2.1 Å (NEQ205-NEQ261) and 2.2 Å (NEQ261), respectively, and are shown in ribbon diagrams in Figures 2 and 3.

The overall structure of the NEQ endonuclease, representative of the $(\alpha\beta)_2$ endonuclease family, was that of a heterotetramer, retaining the four-subunit arrangement observed in both the α_4 and α_2 families (Figure 2A). The two catalytic subunits were arranged at a diagonal with the two structural subunits acting to stabilize this orientation. This arrangement positions the active sites of the catalytic subunits for interaction with the RNA bulges containing the splice sites, as demonstrated in the crystal structure of the full-length AFU enzyme bound to an RNA substrate (35). The superimposition of the MJA and the AFU structures onto the NEQ enzyme (Figure 2B) revealed significant similarity in the overall structure (MJA RMSD = 1.271 Å for 130 C^{α} atoms, AFU RMSD = 1.427 Å for 108 C^{α} atoms). It has been shown structurally as well as biochemically that the four-subunit architecture is the functional assembly that interacts and cleaves the RNA substrates (3,35,36). The four-unit assembly of NEQ endonuclease provides the link between the $(\alpha\beta)_2$ family of endonucleases and the α_2 and α_4 families of endonucleases and, therefore, underscores a common origin of all splicing endonucleases including the eukaryotic $\alpha\beta\gamma\delta$ endonucleases.

Although a single NEQ261 subunit is found in the asymmetric unit, a homodimer with extensive buried

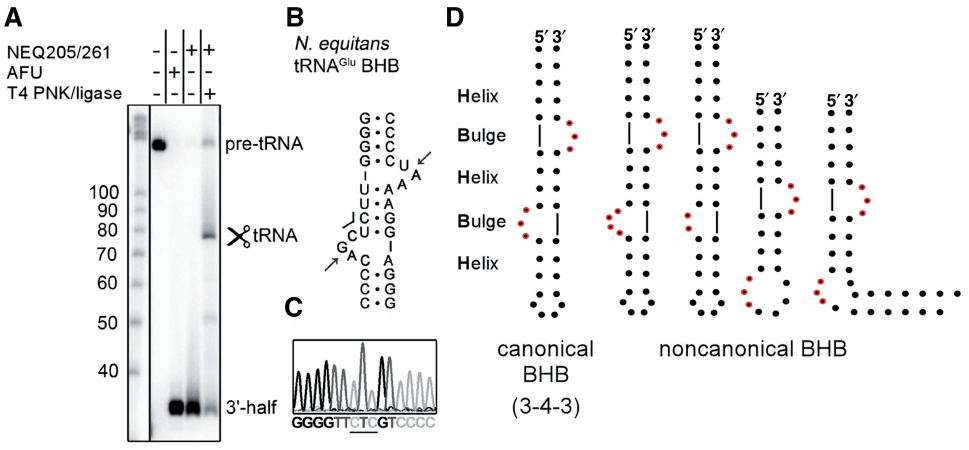

Figure 1. Splicing activity and specificity of the NEQ RNA splicing endonuclease. (**A**) A 3′-radiolabeled RNA transcript of the *N. equitans* $tRNA^{Glu}$ precursor was incubated without enzyme (–) or with either 1 μM NEQ205-NEQ261 splicing endonuclease at 65°C for 20 min alone or followed by incubation with T4 polynucleotide kinase (PNK) and T4 RNA ligase. AFU splicing endonuclease was incubated with a substrate as a positive control. (**B**) Secondary structure of the relaxed BHB motif of *N. equitans* $tRNA^{Glu}$ precursor. The predicted cleavage sites are indicated by arrows and the CUC anticodon is indicated by a line. (**C**) The mature tRNA product was excised, amplified by RT-PCR and sequenced. The anticodon loop was correctly assembled and the anticodon is underlined. (**D**) Examples of the RNA substrates cleaved by the tRNA splicing endonuclease that have been confirmed biochemically, i.e. canonical bulge–helix–bulge (BHB) RNA substrate (left panel) (13,27,28) and non-canonical BHB substrates (right panel). For non-canonical substrates, from the left: a synthetic 4–3–3 and 2–3–3 BHB (28), a bulge–helix–loop (BHL) (29) and a *trans*-spliced BHL formed by two split half tRNA genes (13,29).

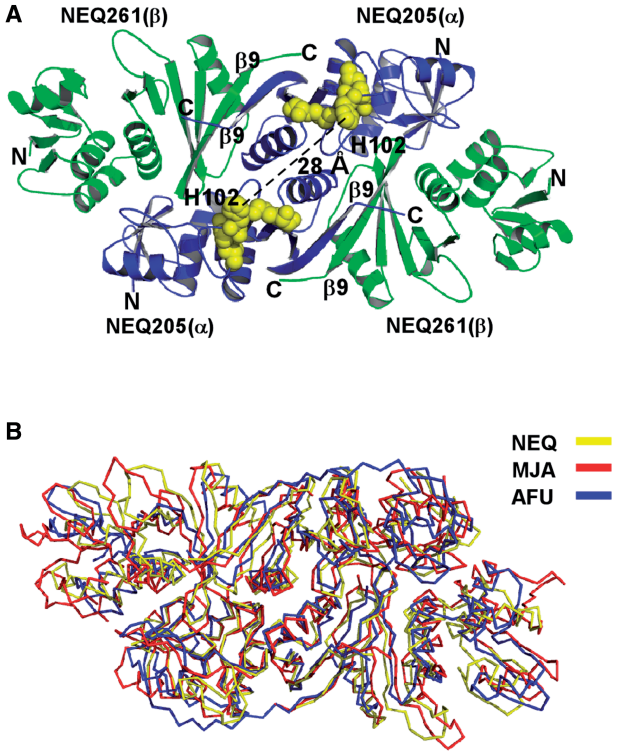

Figure 2. Structure overview of the NEQ RNA splicing endonuclease. (**A**) Structure of the functional $(\alpha\beta)_2$ NEQ205–NEQ261 splicing endonuclease. Catalytic subunits are colored in blue and structural subunits in green. Putative catalytic triad residues are shown as yellow spheres, with catalytic histidine (H102) denoted. (**B**) Ribbon figure showing the overlay of the NEQ205–NEQ261 (in yellow), MJA (PDBID 1A79) (in red) and AFU (PDBID 1RLV) (in blue) full-length enzymes.

solvent accessible surface (880 Å^2) can be formed through the crystallographic 2-fold symmetry operation (Figure 3A) and retains the same orientation as the heterodimer (Figure 3B). The formation of the inactive homodimers has also been observed for the isolated catalytic subunit from SSO (28) and STO (PDBID:2CV8), suggesting that the dimer is the most stable form of isolated subunits (Figure 3B).

Competing functional and non-functional subunit assemblies

The functional NEQ structure allowed us to analyze features responsible for enzyme assembly among all currently known endonuclease or subunit structures. We calculated buried solvent accessible surface areas among the three interfaces using the protein–protein interaction interface server (PISA) (37). Unique to the $(\alpha\beta)_2$ family of endonuclease is the fact that the heterotetramer is formed through dimerization of two isologous heterodimers (as opposed to a homodimer or a monomer of two repeats). The heterodimers comprises catalytic and structural subunits coming together at their C-terminal ends (α–β assembly), through the antiparallel β9–β9′ strands that confer hydrophobic (55%) and hydrogen bonding (14 hydrogen bonds) interactions (β9–β9′ interaction) (Figure 4A). Similar compositions are seen in the β9–β9′ interactions of the α_4 (59% hydrophobic and 16 hydrogen bonds) and α_2 (50% hydrophobic and 12 hydrogen bonds) endonucleases. Furthermore, the total buried solvent accessible area by the β9–β9′ interactions is 1330 Å^2 for the $(\alpha\beta)_2$, 1130 Å^2 for the α_2 and 1240 Å^2 for the α_4 endonucleases, respectively, suggesting a similar mechanism of assembly at this interface.

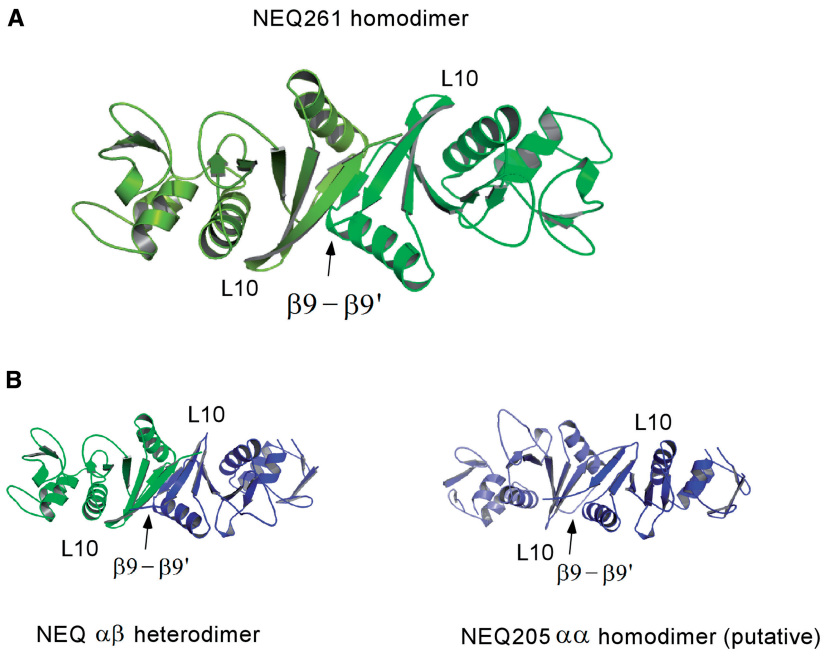

Figure 3. Structure of a NEQ261 subunit homodimer. (**A**) NEQ261 homodimer generated by crystallographic 2-fold symmetry. Conserved loop L10 and β9–β9′ features are labeled. (**B**) Comparison of the NEQ αβ heterodimer from the NEQ205–NEQ261 complex structure (left panel) and NEQ205 α–α homodimer generated by superimposition with the SSO0439 homodimer structure (28).

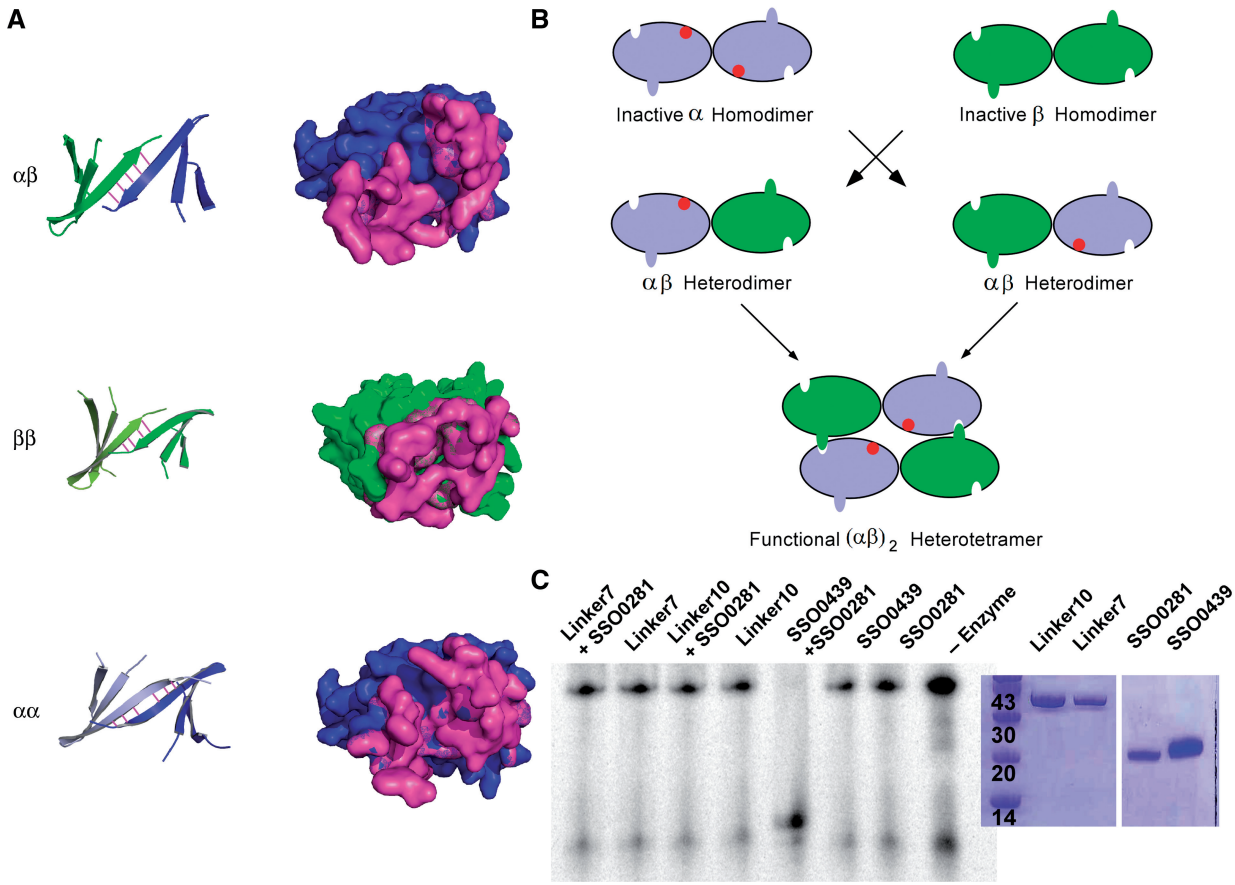

Figure 4. Endonuclease dimer interfaces and the model of enzyme assembly. (**A**) (Left panel) Illustration of the β9–β9′ interaction and its interface (right panel) at the C-terminal interface between α–β, α–α and β–β dimers. Lines indicate hydrogen bonds at the β9–β9′ interface. Cutaway view of the space-filling model of the interface (magenta) with back monomer in green or blue and front monomer not shown. (**B**) A proposed model of assembly for the $(\alpha\beta)_2$ tRNA splicing and processing endonuclease. Catalytic and structural subunits form inactive homodimers through homo-associated β9–β9′ interactions, and reassemble to form heterodimers through heteroassociated β9–β9′ interactions. Functional tetramers are formed through the L10–pocket interaction. (**C**) RNA cleavage activity of the wild-type SSO endonuclease and two linker mutants. Individual subunits purified separately are shown in a SDS–PAGE gel on right side.

Analysis of the NEQ261 homodimer structure and the previously determined structure of the SSO catalytic subunit (SSO0439) (28) revealed a mechanism of competitive subunit assembly (Figure 4B). In both non-functional homodimers of the isolated catalytic and structural subunits, the same β9–β9′ interface required for the α–β heterodimer formation was used to form the α–α and β–β homodimers except that the β9 strands were from two identical subunits in the homodimers (Figure 4A). Thus, subunits engaged in homodimer assembly are unable to form the functional four-unit functional assembly (Figure 4C).

Comparison of the buried solvent accessible surface areas at all β9–β9′ interfaces suggests a favorable formation of the heterodimer. The α–β β9–β9′ interface in the NEQ205–NEQ261 complex had slightly greater buried solvent accessible surface area (1323 Å^2) than that of the α–α β9–β9′ interface in the modeled NEQ205 homodimer (1270 Å^2), although both values were similar to those of the β9–β9′ interface in the α_2 and α_4 homoassociated endonucleases (3,36). This comparison alone suggests that the non-functional homodimer of the catalytic subunit is similar in interaction strength to the functional catalytic–structural heterodimer. Strikingly, however, the β–β β9–β9′ interface seen in the NEQ261 homodimer has a significantly less solvent accessible surface area (880 Å^2) than that involving the catalytic subunit. This result implicates a weaker interaction between two structural subunits than that between an α and a β subunit, suggesting a favorable formation of the α–β heterodimer over the β–β homodimer when both subunits are present.

Further interaction of these heterodimers to form the heterotetramer is facilitated by an electrostatic interaction between a negatively charged loop L10 of the structural subunit and a positively charged pocket of the opposing catalytic subunit (Figure 5A). While this interaction involves similar buried solvent accessible area [900 Å^2 for $(\alpha\beta)_2$, 1100 Å^2 for α_2 and 1080 Å^2 for α_4], the interaction appears to be less extensive than the β9–β9′ interaction, as evidenced by the large proportions of the polar interactions [58% for $(\alpha\beta)_2$, 66% for α_2 and 55% for α_4]. Nonetheless, sequence analysis indicates that the L10–pocket interaction is well conserved (Figure 5B). Indeed, it was recently shown that the ability

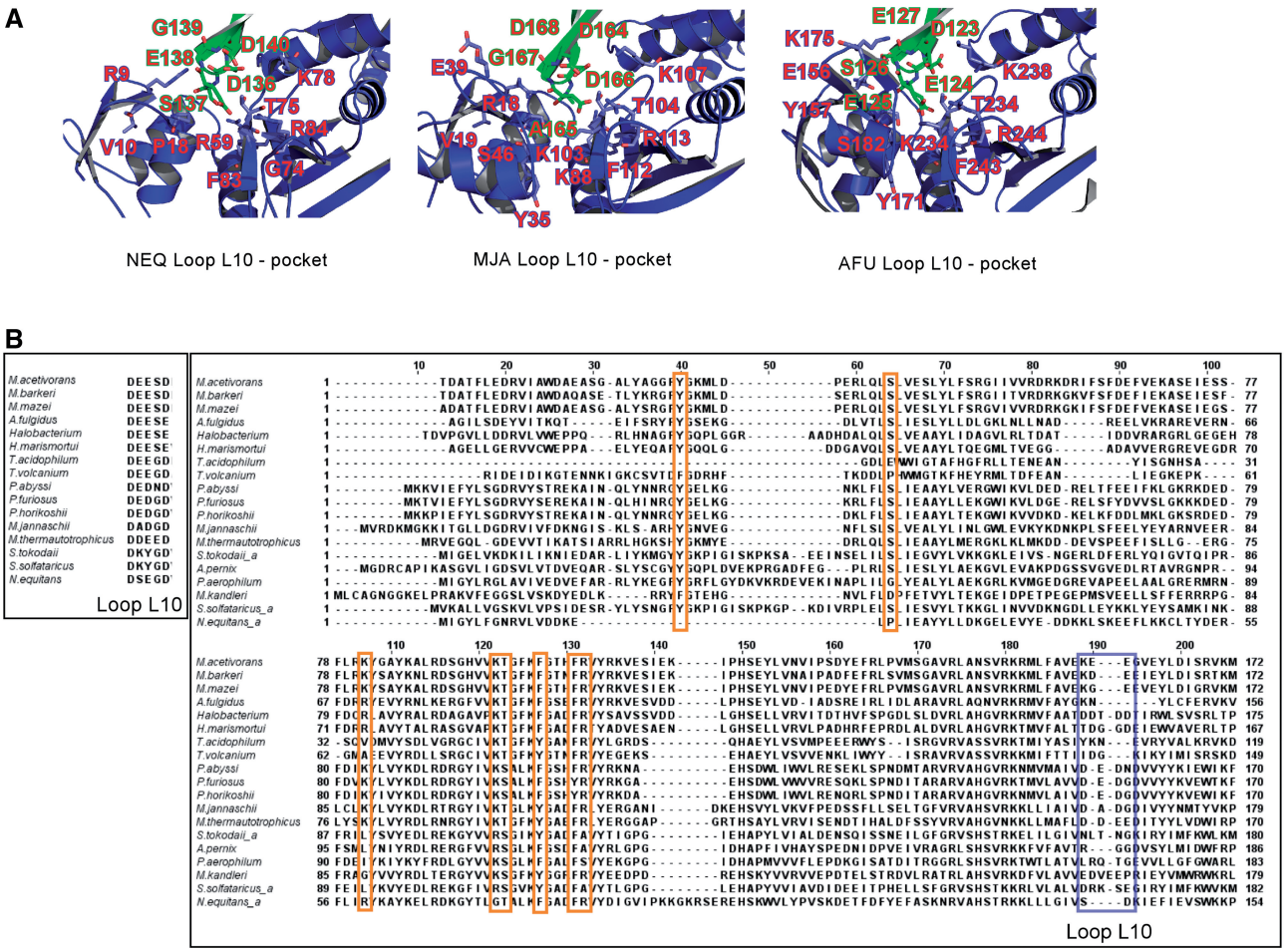

Figure 5. Conserved L10–pocket interactions in $(\alpha\beta)_2$ endonuclease. (**A**) Comparison of the L10–pocket interface structures of NEQ (left panel), MJA (middle panel) and AFU (right panel) endonucleases. (**B**) Sequence comparison of the conserved L10 region of the structural subunits (left panel) and of the pocket residues from catalytic subunits (orange, right panel). Loop L10 sequences of catalytic subunits, with little sequence similarity, are denoted by a blue box, and conserved pocket residues are denoted by orange boxes.

of the subunits to form the heterotetramer resided in the loop L10 sequence, and that 'swapping' the L10 sequence of the $(\alpha\beta)_2$ catalytic subunit with that of the structural subunit was sufficient to allow the assembly into a functional heterotetrameric enzyme (30).

Based on the structural findings, we propose that the heteroassociated $(\alpha\beta)_2$ endonuclease is assembled from homodimers of isolated subunits by competition for the same interface (Figure 4B). This mechanism of assembly allows regulation of the available functional endonuclease through available subunits. At low concentrations, the functional heterodimer is the preferred form of assembly whereas at high concentrations, homodimers of an individual subunit are formed that may act as a homeostasis devise in maintaining a constant level of the active enzyme.

An assembly deficient mutant confirms the assembly model

The NEQ structure indicates that the formation of the four-subunit arrangement as seen with the MJA and AFU structures does indeed occur and is the functional assembly of this enzyme. However, the structures of the NEQ structural subunit as well as of the SSO and STO catalytic subunits consistently show the formation of homodimers that compete for the formation of the functional tetramer. These structural findings suggest that a reassembly step is required for the formation of the active enzyme.

In order to explore this reassembly step in the formation of the heterotetramer, we constructed plasmids that express two linked SSO catalytic subunits (SSO0439) by a peptide linker of either ten (Linker10) or seven (Linker7) amino-acid residues. The peptide linkers were designed based on both the natural linker of the AFU endonuclease and the three-dimensional structure of NEQ205–NEQ261 so that they do not disrupt the structural integrity of the catalytic subunit. We hypothesized that the catalytic subunit linker mutants, regardless the length of the linker, would be incapable of dissociating to reassemble with the β-subunit and therefore not exhibit any activity toward the RNA substrate either alone or when combined with the structural subunit, SSO0281. Indeed, neither linker mutants cleaved the canonical BHB RNA substrate

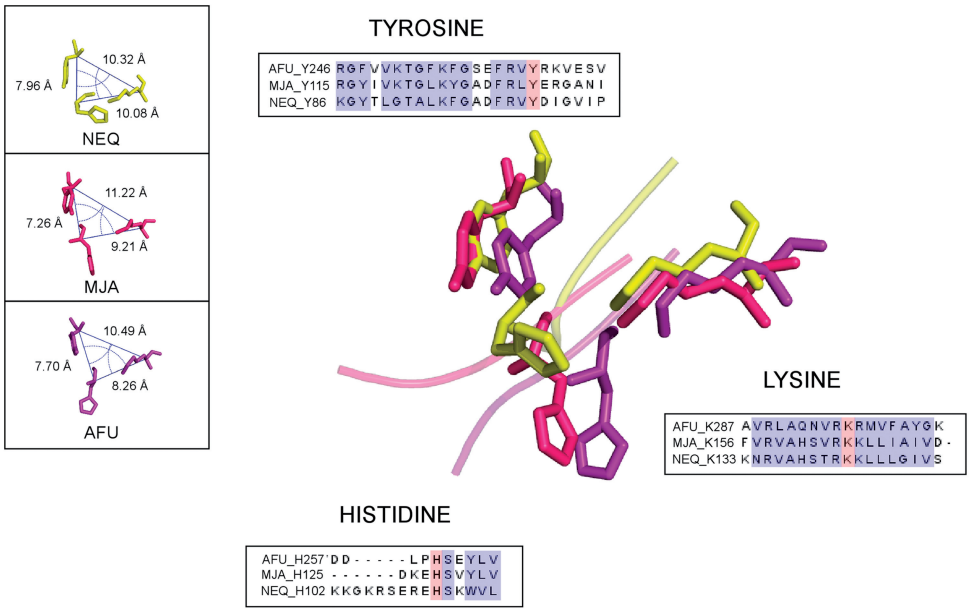

Figure 6. Comparison of (putative) RNA-binding pockets. Superimposition of the active site catalytic residues of MJA (in hot pink) and AFU (in purple) onto NEQ (in yellow). NEQ residues, His102, Tyr86 and Lys133, MJA residues, His125, Tyr115 and Lys156, and AFU residues, His257, Tyr246 and Lys287, are shown. Sequence alignments around each of the three catalytic triad residues are shown with catalytic residues highlighted in pink and conserved residues in blue. (Inset) Comparison of the geometry of the RNA-binding pocket represented by the catalytic triad residues.

when incubated alone or in the presence of the SSO0281 structural subunit, while the wild-type SSO0439 catalytic subunit did (Figure 4C). These results indicate that preventing the reassembly of the homodimers into the heterodimer does prohibit further reassembly into the functional heterotetramer and in turn inhibits cleavage, thus supporting the necessary reassembly step in the overall assembly process of the $(\alpha\beta)_2$ splicing endonuclease (Figure 4B).

Active site structure and the implication of broad RNA substrate specificity

An earlier protein engineering study showed that a functional splicing endonuclease comprised entirely of the catalytic subunit of the SSO $(\alpha\beta)_2$ endonuclease retained the broad substrate specificity characteristic of the $(\alpha\beta)_2$ family of endonucleases (30). This indicates that the basis of the enzymes' substrate specificity is inherent to the catalytic subunit, most likely the active site itself.

To explore the unique RNA substrate specificity of the $(\alpha\beta)_2$ endonuclease, the active site structure of NEQ is compared to that of the MJA and AFU endonucleases. In the previously determined AFU endonuclease–RNA complex structure (35), three conserved residues, a histidine, a tyrosine, and a lysine, are identified to be possible catalytic residues. In the proposed catalytic mechanism for the splicing endonuclease, the tyrosine facilitates deprotonation of the 2′-nucleophilic oxygen, and histidine acts as the general acid donating a proton to the 5′ leaving group (35). In addition to the putative general acid role, the conserved histidine is also believed to be responsible for orientating the scissile phosphate group nearly in-line for the phosphotransfer reaction by stacking its imidazole ring on the third bulge nucleobase (35,38). Superimposed active site structures of the three endonucleases show a general agreement in the location of the catalytic residues (3 C_α atoms RMSD 1.4 Å with AFU and 1.9 Å with MJA) (Figure 6), which can be expected since all three families are able to cleave the canonical RNA substrate containing the canonical BHB motif. However, notable differences are observed in the orientation of NEQ Tyr86 and His102 in comparison to the AFU and MJA equivalent residues, with the more significant deviation in the location of His102. An examination of the NEQ catalytic residues with respect to a bound BHB RNA (Figure 7A) model based on the AFU endonuclease–RNA complex structure (35) showed that while Lys133 is in a similar position to AFU Lys287, both His102 and Tyr86 differ in their current conformations to those in the AFU counterparts (Figure 7B). In addition, NEQ also shares less similar residues flanking Tyr86 and His102 than those in MJA and AFU endonucleases (Figure 6) and has slightly larger catalytic pocket based on distances of the catalytic triad residues (Figure 6, inset). The unique environment of the catalytic pocket may further enhance the enzyme's ability to act on all non-canonical RNA substrates.

CONCLUSIONS

The crystal structure of the NEQ RNA splicing endonuclease, the representative member of the $(\alpha\beta)_2$ family of endonucleases, has been reported. This structure has confirmed that the $(\alpha\beta)_2$ endonuclease adopts the same four-subunit arrangement that has been determined for the α_4 and α_2 endonucleases. Additionally, the crystal structure

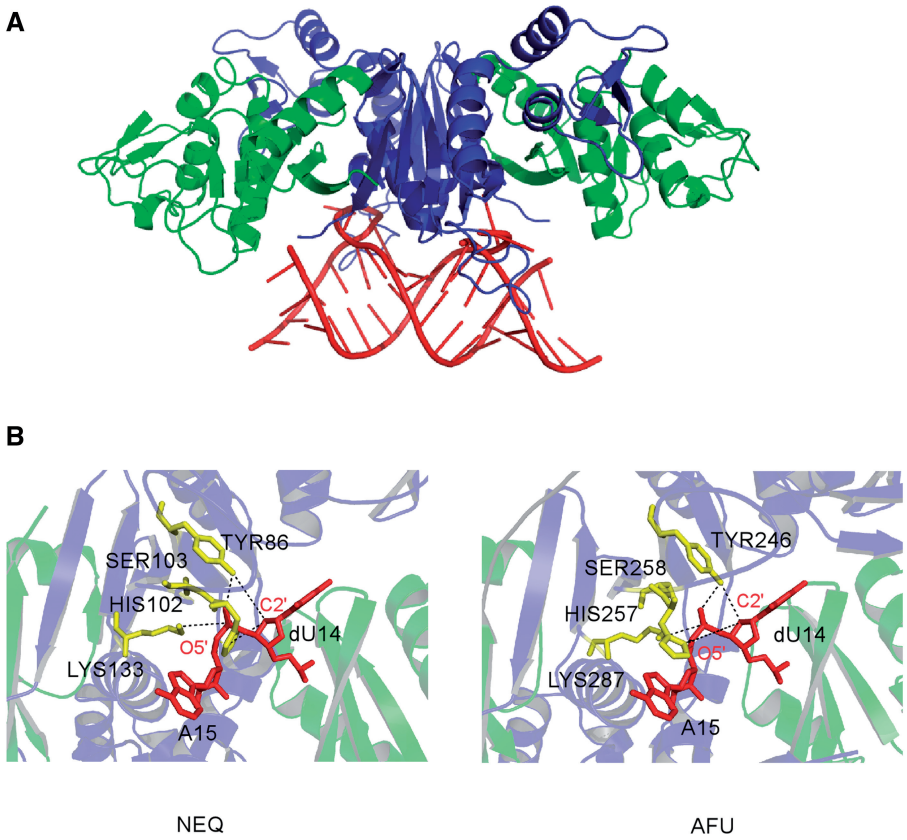

Figure 7. Structure model of NEQ splicing endonuclease bound to a BHB RNA substrate. (**A**) NEQ205–NEQ261 heterotetramer bound to the BHB RNA substrate obtained by superimposition of the NEQ endonuclease with AFU endonuclease that is bound to the BHB RNA (PDBID 2GJW). The BHB RNA largely fits the NEQ endonuclease. (**B**) Detailed model of the NEQ endonuclease–RNA complex (left) at the active site and comparison to that of the AFU endonuclease–RNA complex (right). Two of the three bulge nucleotides are shown in red sticks with C2′ of the 5′ nucleotide and O5′ of the 3′ nucleotide labeled. NEQ catalytic triad residues Tyr86, His102 and Lys133 as well as AFU catalytic residues are shown in yellow.

of the NEQ261 homodimer indicates that the structural subunit forms a homodimer in the absence of its NEQ205 catalytic subunit counterpart. This structure, along with the previously determined structures of another $(\alpha\beta)_2$ catalytic subunit, which also form homodimers when in isolation, indicates that a reassembly step is needed in the overall assembly process in order to form the functional heterotetramer. Mutational studies using a covalently linked catalytic subunit indicate that it fails to reassemble and to cleave substrate RNA. We suggest that the equilibrium established among homodimers of subunits and heterotetramers of the functional enzyme provides a means to regulate the amount of functional enzyme.

Data previously reported on maintenance of broad specificity by the catalytic subunit of an $(\alpha\beta)_2$ endonuclease alone indicate that assembly is not a factor in specificity. Structural comparison of the NEQ active site and those in MJA and AFU endonucleases indicates an inherent flexibility in the NEQ active site that may act to accommodate non-canonical RNA bulge structures abundant in this organism. This result and the earlier study suggest that the active site pocket is likely the primary structural unit responsible for the broad substrate specificity. Further structural and biochemical studies including a cocrystal structure of NEQ and an RNA substrate are needed to establish the origin of broad RNA specificity in the NEQ splicing endonuclease.

ACKNOWLEDGEMENTS

X-ray diffraction data were collected from the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions for APS beamlines may be found at http://necat.chem.cornell.edu/ and http://www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

FUNDING

National Institute of General Medical Sciences Grants (to D.S.); the National Science Foundation (to H.L.); and the American Heart Association Florida/Puerto Rico Affiliate (to S.X.).

Conflict of interest statement. None declared.

REFERENCES

1. Tourasse,N.J. and Kolsto,A.B. (2008) Survey of group I and II introns in 29 sequenced genomes of the *Bacillus cereus* group: insights into their spread and evolution. *Nucleic Acids Res.*, **36**, 4529–4548.
2. Biniszkiewicz,D., Cesnaviciene,E. and Shub,D.A. (1994) Self-splicing group I intron in cyanobacterial initiator methionine tRNA: evidence for lateral transfer of introns in bacteria. *EMBO J.*, **13**, 4629–4635.
3. Li,H., Trotta,C.R. and Abelson,J. (1998) Crystal structure and evolution of a transfer RNA splicing enzyme. *Science*, **280**, 279–284.
4. Abelson,J., Trotta,C.R. and Li,H. (1998) tRNA splicing. *J. Biol. Chem.*, **273**, 12685–12688.
5. Abelson,J. (1979) RNA processing and the intervening sequence problem. *Annu. Rev. Biochem.*, **48**, 1035–1069.
6. Culbertson,M.R. and Winey,M. (1989) Split tRNA genes and their products: a paradigm for the study of cell function and evolution. *Yeast*, **5**, 405–427.
7. Knapp,G., Beckmann,J.S., Johnson,P.F., Fuhrman,S.A. and Abelson,J. (1978) Transcription and processing of intervening sequences in yeast tRNA genes. *Cell*, **14**, 221–236.
8. Valenzuela,P., Venegas,A., Weinberg,F., Bishop,R. and Rutter,W.J. (1978) Structure of yeast phenylalanine-tRNA genes: an intervening DNA segment within the region coding for the tRNA. *Proc. Natl Acad. Sci. USA*, **75**, 190–194.
9. Marck,C. and Grosjean,H. (2002) tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA*, **8**, 1189–1232.
10. Marck,C. and Grosjean,H. (2003) Identification of BHB splicing motifs in intron-containing tRNAs from 18 archaea: evolutionary implications. *RNA*, **9**, 1516–1531.
11. Grosjean,H., Gaspin,C., Marck,C., Decatur,W.A. and de Crecy-Lagard,V. (2008) RNomics and Modomics in the halophilic archaea *Haloferax volcanii*: identification of RNA modification genes. *BMC Genomics*, **9**, 470.
12. Chan,P.P. and Lowe,T.M. (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.*, **37**, D93–D97.
13. Randau,L., Calvin,K., Hall,M., Yuan,J., Podar,M., Li,H. and Soll,D. (2005) The heteromeric *Nanoarchaeum equitans* splicing endonuclease cleaves noncanonical bulge-helix-bulge motifs of joined tRNA halves. *Proc. Natl Acad. Sci. USA*, **102**, 17934–17939.
14. Randau,L., Munch,R., Hohn,M.J., Jahn,D. and Soll,D. (2005) *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their 5′- and 3′- halves. *Nature*, **433**, 537–541.
15. Fujishima,K., Sugahara,J., Kikuta,K., Hirano,R., Sato,A., Tomita,M. and Kanai,A. (2009) Tri-split tRNA is a transfer RNA made from 3 transcripts that provides insight into the evolution of fragmented tRNAs in archaea. *Proc. Natl Acad. Sci. USA*, **106**, 2683–2687.
16. Budde,B.S., Namavar,Y., Barth,P.G., Poll-The,B.T., Nurnberg,G., Becker,C., van Ruissen,F., Weterman,M.A., Fluiter,K., te Beek,E.T. *et al.* (2008) tRNA splicing endonuclease mutations cause pontocerebellar hypoplasia. *Nat. Genet.*, **40**, 1113–1118.
17. Lee,M.C. and Knapp,G. (1985) Transfer RNA splicing in *Saccharomyces cerevisiae*: secondary and tertiary structures of the substrates. *J. Biol. Chem.*, **260**, 3108–3115.
18. Swerdlow,H. and Guthrie,C. (1984) Structure of intron-containing tRNA precursors: analysis of solution conformation using chemical and enzymatic probes. *J. Biol. Chem.*, **259**, 5197–5207.
19. Phizicky,E.M., Consaul,S.A., Nehrke,K.W. and Abelson,J. (1992) Yeast tRNA ligase mutants are nonviable and accumulate tRNA splicing intermediates. *J. Biol. Chem.*, **267**, 4577–4582.
20. Baldi,M.I., Mattoccia,E., Bufardeci,E., Fabbri,S. and Tocchini-Valentini,G.P. (1992) Participation of the intron in the reaction catalyzed by the *Xenopus* tRNA splicing endonuclease. *Science*, **255**, 1404–1408.
21. Greer,C.L., Soll,D. and Willis,I. (1987) Substrate recognition and identification of splice sites by the tRNA-splicing endonuclease and ligase from *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **7**, 76–84.
22. Reyes,V.M. and Abelson,J. (1988) Substrate recognition and splice site determination in yeast tRNA splicing. *Cell*, **55**, 719–730.
23. Thompson,L.D. and Daniels,C.J. (1990) Recognition of exon-intron boundaries by the *Halobacterium volcanii* tRNA intron endonuclease. *J. Biol. Chem.*, **265**, 18104–18111.
24. Trotta,C.R., Miao,F., Arn,E.A., Stevens,S.W., Ho,C.K., Rauhut,R. and Abelson,J.N. (1997) The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonuclease. *Cell*, **89**, 849–858.
25. Kleman-Leyer,K., Armbruster,D.W. and Daniels,C.J. (1997) Properties of *H. volcanii* tRNA intron endonuclease reveal a relationship between the archaeal and eucaryal tRNA intron processing systems. *Cell*, **89**, 839–847.
26. Lykke-Andersen,J. and Garrett,R.A. (1997) RNA-protein interactions of an archaeal homotetrameric splicing endoribonuclease with an exceptional evolutionary history. *EMBO J.*, **16**, 6290–6300.
27. Tocchini-Valentini,G.D., Fruscoloni,P. and Tocchini-Valentini,G.P. (2005) Structure, function, and evolution of the tRNA endonucleases of Archaea: an example of subfunctionalization. *Proc. Natl Acad. Sci. USA*, **102**, 8933–8938.
28. Calvin,K., Hall,M.D., Xu,F., Xue,S. and Li,H. (2005) Structural characterization of the catalytic subunit of a novel RNA splicing endonuclease. *J. Mol. Biol.*, **353**, 952–960.
29. Tocchini-Valentini,G.D., Fruscoloni,P. and Tocchini-Valentini,G.P. (2005) Coevolution of tRNA intron motifs and tRNA endonuclease architecture in Archaea. *Proc. Natl Acad. Sci. USA*, **102**, 15418–15422.
30. Calvin,K. and Li,H. (2007) Achieving specific RNA cleavage activity by an inactive splicing endonuclease subunit through engineered oligomerization. *J. Mol. Biol.*, **366**, 642–649.
31. Vagin,A. and Teplyakov,A. (1997) MOLREP: an automated program for molecular replacement. *J. App. Crystallogr.*, **30**, 1022–1025.
32. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular replacement graphics. *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, **60**, 2126–2132.
33. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, **53**, 240–255.
34. Brunger,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. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, **54**, 905–921.
35. Xue,S., Calvin,K. and Li,H. (2006) RNA recognition and cleavage by a splicing endonuclease. *Science*, **312**, 906–910.
36. Li,H. and Abelson,J. (2000) Crystal structure of a dimeric archaeal splicing endonuclease. *J. Mol. Biol.*, **302**, 639–648.
37. Krissinel,E. and Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **372**, 774–797.
38. Min,D., Xue,S., Li,H. and Yang,W. (2007) 'In-line attack' conformational effect plays a modest role in an enzyme-catalyzed RNA cleavage: a free energy simulation study. *Nucleic Acids Res.*, **35**, 4001–4006.