Enhancing the Prebiotic Relevance of a Set of Covalently Self-Assembling, Autorecombining RNAs Through In Vitro Selection

Aaron S. Burton · Niles Lehman

Abstract An in vitro form of the self-splicing group-I intron interrupting the *Azoarcus* tRNA^{Ile} was shortened by ~10% with the removal of helix P6a. This deletion reduced the reverse-splicing activity of the ribozyme about 10-fold. Through in vitro selection, this activity was restored in several low-error mutants. A number of mutations were found that improved reverse-splicing activity through both increased k_{obs} and better folding. The deletion mutant could be fragmented into as many as three discrete pieces, which, when incubated together, were capable of covalent self-assembly through energy-neutral transesterification reactions, a process called autorecombination. A subset of the mutations identified through in vitro selection for reverse-splicing were exaptations in that they were also shown to augment the autorecombination reactions, leading to higher yields of covalently self-assembled products, making this the smallest such system yet discovered.

Keywords In vitro selection · Autocatalysis · Recombination · Group-I intron · Self-replication · Molecular evolution · *Azoarcus*

Introduction

The RNA World hypothesis describes a time during the origins of life, prior to the advent of DNA and proteins, when RNA served as both catalyst and informational molecule. A great deal of evidence has been obtained in support of this hypothesis (Rich 1962; Kuhn 1972; Orgel 1986; Joyce 1991,

A. S. Burton · N. Lehman (✉)
Department of Chemistry, Portland State University,
P.O. Box 751, Portland, OR 97207, USA
e-mail: niles@pdx.edu

1998, 2002; Szostak and Ellington 1993; Gesteland et al. 2006; Boussau et al. 2008). Current research is focused on finding self-replicating RNA systems. A number of advances have been made toward an RNA replicase molecule based on RNA-dependent RNA polymerization (Johnston et al. 2001; McGinness and Joyce 2003; Zaher and Unrau 2007). Unfortunately, the polymerase ribozymes discovered to date are able to copy less than 1/4 of their total length. In addition, the molecules created by these polymerases are complementary copies of their targets, meaning that two complete copies have to be made for each replication event. These limitations have led to the exploration of other ribozyme chemistries for self-replicating RNA systems, including simple *trans*-esterification, which can be used to build up longer RNA molecules through recombination (Gilbert 1986; Burke and Willis 1998; Lehman 2003; Striggles et al. 2006).

Group-I introns are a class of structurally conserved RNAs that catalyze their own excision from nascent transcripts. These self-splicing ribozymes ligate the exons they interrupt via consecutive *trans*-esterification reactions, the first employing exogenous guanosine as a nucleophile to free the 5′-exon, followed by attack of the 3′-hydroxyl of the now-free 5′-exon on the 3′-splice site. The two *trans*-esterification steps have been extensively studied for several group-I introns. Because the *trans*-esterification reactions swap one 3′,5′-phosphoester bond for another, they are essentially energy-neutral, and therefore, readily reversible. The reverse of the second step of splicing reaction, ligation of the 3′-exon to the 3′-end of the ribozyme, has been exploited for in vitro selection of ribozymes with altered catalytic properties, such as the ability to use Ca^{2+} ions instead of Mg^{2+} ions for catalysis (Lehman and Joyce 1993; Burton and Lehman 2006).

The reversibility of the second step of splicing has also been used to achieve the recombination of exogenous

substrates. To date, the *Tetrahymena*, *Azoarcus*, *Candida*, and *Pneumocystis* ribozymes have been shown to catalyze recombination in vitro (Bell et al. 2002; Riley and Lehman 2003; Dotson et al. 2008). The *Azoarcus* ribozyme has demonstrated the ability to make products ranging from essentially non-structured, short RNAs to a ~200-nucleotide group-I introns, from as many as four smaller substrate pieces (Riley and Lehman 2003; Hayden et al. 2005; Hayden and Lehman 2006). Furthermore, the ribozyme itself can be fragmented into as many as four pieces ranging from 39 to 63 nucleotides that are able to generate full-length, covalent ribozymes by performing recombination reactions as a non-covalent *trans* assembly. The recombinant ribozymes are active in the milieu in which they are made, and generate additional full-length ribozymes by recombining the remaining non-covalently bonded pieces of themselves (autorecombination). Because these autorecombinases make identical copies of themselves, they actually perform autoreplication.

The aforementioned system provides an energy-neutral mechanism for the build-up of genetic information from smaller fragments and for molecular replication (Lehman 2008). Formation of the substrates for recombination, however, almost certainly required some sort of activating agent, perhaps as nucleotide triphosphates, imidazolates, or phophoramidates. Although it has been demonstrated that RNA polymers up to 50–60-nucleotides long can be generated by polymerization of activated monomers on clay minerals, the bulk of the products generated were 20–30 nucleotides in length (Ferris 2002). Clearly, then, smaller would have been better for any self-replicating system, both from an availability-of-materials standpoint and requirements for replication fidelity in light of Eigen's paradox (1971): no enzyme without a large genome, and no large genome without enzymes. To that effect, we have removed about 10% of the length of the autorecombinase form of the *Azoarcus* group-I intron (Fig. 1). This decreased the *trans*-splicing activity of the ribozyme by an order of magnitude, which we restored through in vitro selection. We were thus able to identify and characterize truncated *Azoarcus* ribozyme mutants capable of covalent self-assembly and autoreplication by autorecombination.
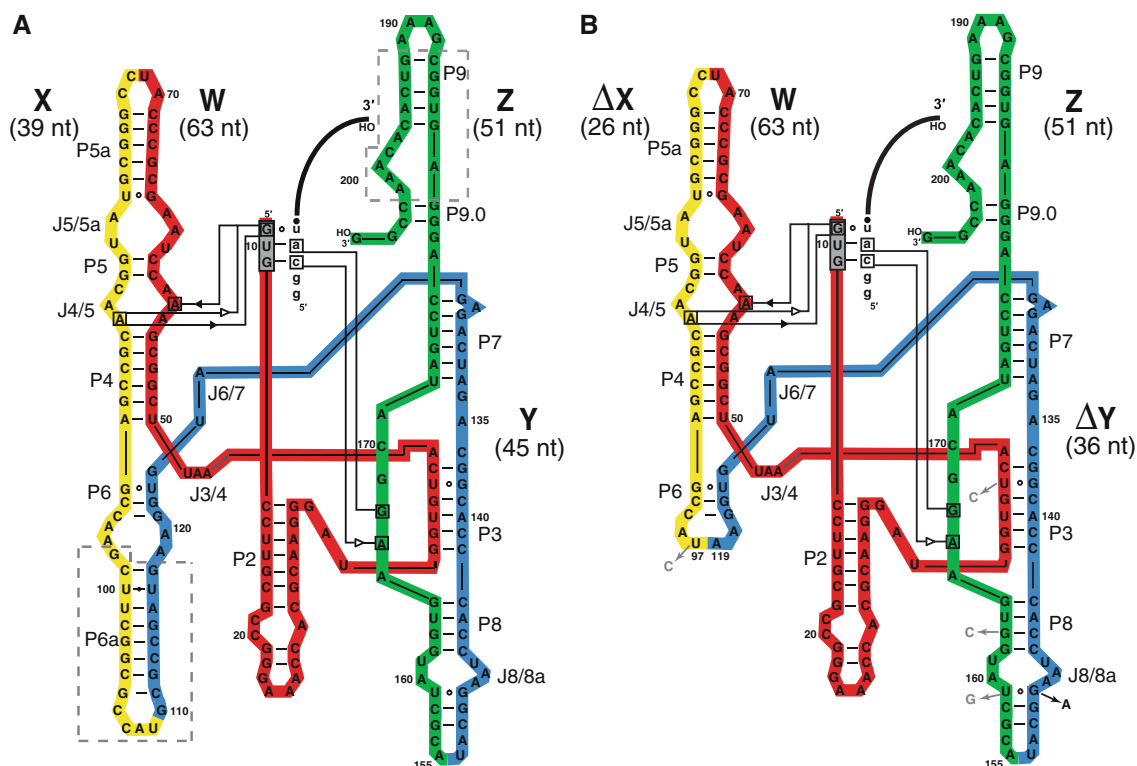

Fig. 1 Secondary structures of the *Azoarcus* ribozyme autorecombinase. Fragments are separated at the end of each *colored* portion. Names and sizes are listed next to each fragment. The internal guide sequence of the each ribozyme is boxed in *gray*; the interaction between substrate and ribozyme is indicated. **a** The 198-nucleotide ribozyme contains CAUs at L5, L6, and L8 that enable it to be autorecombined. Potential deletion sites in the *Azoarcus* ribozyme are outlined by *dotted lines*. The resulting covalently attached ribozymes were assayed for reverse-splicing ability. **b** The 176-nucleotide ΔP6aRC ribozyme. *Gray arrows* pointing from nucleotides 44, 97, 159, and 163 represent mutations present in the four-error mutant, 4EM. The *black arrow* shows the mutant G150A

Materials and Methods

Construction of G_0 Selection Population

DNA oligonucleotides were purchased from IDT. The DNA template for the ΔP6aRC form (Fig. 1b) of the ribozyme was generated by recursive gene synthesis as described (Engels and Uhlmann 1988; Rydzanicz et al. 2005; Hayden and Lehman 2006). The initial mutant population was generated by mutagenic PCR (Cadwell and Joyce 1992; Vartanian et al. 1996) to give a mutation level of 10% per position, as described previously (Burton and Lehman 2006). RNAs were transcribed by run-off transcription and gel-purified prior to selection.

In Vitro Selection

For selection generations 0–4, reactions contained 5-μM mutagenized *Azoarcus* ΔP6aRC RNA, 2.5-μM S9 substrate RNA (5′-AACAU•CCAAUCGCAGGCUCAGC-3′; selection is affected because the portion 3′ of the "•" is appended only to the ends of active ribozymes upon reaction), 100-mM $MgCl_2$, and 30-mM EPPS buffer (pH 7.5) in a total volume of 20 μl. In generations 5–10, ribozyme and substrate concentrations were halved to 2.5 and 1.25 μM, respectively. Solutions containing ribozyme and substrate RNAs were heat-denatured at 80°C for 5 min, after which the reaction buffer was added and reactions were incubated at 48°C for 15 min. Reactions were quenched by adding 4 μl 0.5-M EDTA and diluting sample volumes to 500 μl with water. Substrate molecules and salts were removed by concentrating the sample to <10 μl in a Nanosep 10K column. An additional 500 μl of water were added and the sample was concentrated again. Active ribozymes were selectively amplified in reverse transcription reactions that used all the desalted and concentrated RNA, T9 primer DNA (5′-GCTGAGCCTGCGATTGG-3′; complementary to the S9 tail and present in a 2:1 M ratio relative to total ribozymes), 10-mM $MgCl_2$, 5-mM dithiothreitol, 50-mM Tris–HCl (pH 7.5), 0.2-mM dNTPs, and 15-U AMV reverse transcriptase (USB) in 20 μl volumes at 37°C for 1 h. Double-stranded DNA templates containing the T7 promoter sequence were generated by standard PCR using T21a (5′-CTGCAGAATTCTAATACGACTCACTATAGTGCCTTGCGCCGGGAA-3′) and T20a (5′-CCGGTTTGTGTGACTTTGCC-3′).

Genotyping of Selection Population

PCR products of the G_{10} population were cloned into the vector pJET1.2 (Fermentas) and transformed into *E. coli* Top 10 competent cells (GeneJet). Individual colonies were picked as templates for PCR reactions using the primers pJET1.2-F and pJET1.2-R (Fermentas), which generate products of ∼325 bp (=the insert size of 205 bp plus 120 bp). Products of the correct size were genotyped using BigDye (v.3) cycle sequencing chemistry.

Assay of Selection Population and Mutants

Colony PCR products containing frequently occurring mutations were diluted and amplified by PCR using primers T21a and T20a to regenerate dsDNA templates for transcription. Following run-off transcription and gel purification, 2.5-μM ribozymes were incubated with 1.25-μM 3′-α^{32}P•dATP-radiolabeled S9a (Huang and Szostak 1996; Burton and Lehman 2006; Burton et al. 2009) in 100-mM $MgCl_2$, 30-mM EPPS (pH 7.5). Reactions were quenched by adding an equal volume of an 3:1 solution of denaturing gel loading buffer and 0.5-M EDTA, respectively. Products were separated on 15%/8-M urea or 20% polyacrylamide/8-M urea gels and visualized by phosphorimaging. Ribozyme activities were calculated using ImageQuant software (v.5.2, GE Health Sciences).

For the most active mutants, values for k_{obs} for reverse-splicing were obtained under single-turnover conditions. Ribozymes were heat-denatured at 80°C for 5 min, then allowed to fold for 5 min at 48°C in the reaction buffer (final concentrations of 100-nM ribozymes, 100-mM $MgCl_2$, and 30-mM EPPS; pH 7.5). Reactions were initiated by the addition of 3′-radiolabeled S9a (∼0.1 nM) and quenched with an equal volume of 3:1 :: 2x denaturing loading dye:0.5-M EDTA. Products were separated on 15% polyacrylamide/8-M urea gels. Values for k_{obs} were estimated by non-linear regression according to the equation $F = A\left(1 - e^{-kt}\right)$, where F is the fraction of substrate appended to the ribozymes, and A was set as 0.204, the regression-estimated limit of reverse splicing under these conditions for the fastest mutant, G150A. Values for k_{obs} varied by less than 50% on different days.

For this same set of mutants, the fraction of ribozymes capable of catalysis (i.e., correctly folded) was determined by allowing the reverse splicing reaction to approach equilibrium with 60-fold substrate excess. Reactions contained 0.5-μM ribozyme and 30-μM S9a substrate. RNAs were heated at 80°C for 5 min, and the reaction initiated by addition of the selection buffer. Aliquots of the reaction were quenched after 15 min (matching the selection conditions) and 60 min (near equilibrium) at 48°C. These aliquots of the reaction were quenched and separated on a denaturing 8% polyacrylamide/8-M urea gel. Products were visualized by SYBR green I staining using the fluorescence mode on the Typhoon scanner and quantified using ImageQuant software.

Covalent Self-Assembly Assays

RNA fragments as partitioned in Fig. 1b were either pur-
chased from IDT or synthesized via run-off transcription.
Fragments other than **W** contained a five-nucleotide 'head'
h of 5′-GGCAU-3′ with the CAU serving as the recombi-
nation tag. Reactions contained 2 μM of each fragment in
100-mM $MgCl_2$, 30-mM EPPS (pH 7.5), using 5′-radiola-
beled **W** molecules. RNAs were heated at 80°C for 5 min
prior to the addition of buffer, then immediately placed at
48°C. Aliquots were quenched with an equal volume of
3:1 :: denaturing loading buffer:0.5-M EDTA. Products
were separated on 8% polyacrylamide/8-M urea gels and
quantified by phosphorimaging.

Results and Discussion

Evaluation of Potential Deletion Sites for a Shorter Autorecombinase

We considered several deletion sites to reduce the size of
the autorecombining form of the *Azoarcus* ribozyme
(Fig. 1a). Two constructs contained deletions in P9, the
first removing nucleotides 181, 182, and 197–201 (ΔP9$_{-7}$);
the second removing nucleotides 181–187 and 192–201
(ΔP9$_{-17}$). Because the tetraloop in L9 is involved in a
tertiary interaction with P4, we expected both of these
deletions to reduce the reverse-splicing activity of the
ribozyme, which they did (Fig. 2). Unexpectedly, the
ΔP9$_{-17}$ form retained more activity than the ΔP9$_{-7}$ form
despite containing all the deletions in ΔP9$_{-7}$. It may be that
removing the bulged region of P9 still allows the L9–P4
tetraloop–receptor interaction to take place but alters the
overall conformation of the ribozyme, while the further
deletions in ΔP9$_{-17}$ spare the correct global fold at the
expense of the L9–L4 interaction. We also made a con-
struct with nucleotides 98–118 deleted (ΔP6aRC) that also
resulted in a ribozyme with reduced activity. Because the
ΔP9 forms were less active than the ΔP6aRC version and
formation of the *trans* recombination complex is likely to
be heavily dependent on the L9–P4 tetraloop interaction,
we chose the ΔP6aRC form as the starting point for the
smaller autorecombinase.

In Vitro Selection to Improve Catalytic Activity

The four-piece self-construction reactions demonstrated
previously (Hayden and Lehman 2006) require that the
RNA be separated into fragments that contain the com-
plement (5′-CAU-3′) to the internal guide sequence of the
Azoarcus ribozyme (5′-GUG-3′). When we placed the
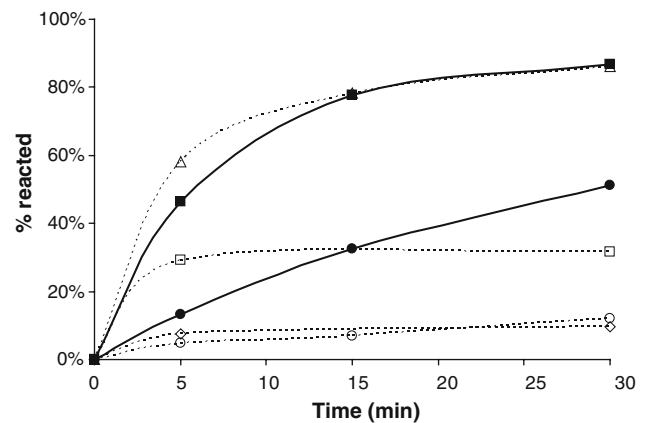CAUs in L5, L6, and L8 of the ΔP6aRC ribozyme to make

Fig. 2 Reverse-splicing activity of *Azoarcus* ribozyme variants.
Ribozymes (2.5 μM) were incubated with radiolabeled substrate
S9a (1.25 μM) ratio for 0–30 min at 48°C in 100-mM $MgCl_2$ and 30-
mM EPPS (pH 7.5). Products were separated on 15% polyacrylamide/
8-M urea gels. The fraction of substrate reverse-spliced was
calculated using ImageQuant software. *Dashed lines* represent pre-
selection ribozymes. *Open triangles*, *open rectangles*, *open diamonds*,
and *open circles* represent the wildtype *Azoarcus*, ΔP9$_{-17}$, ΔP9$_{-7}$,
and ΔP6aRC ribozymes, respectively. *Solid lines* represent the two
most active selection-derived mutants from the ΔP6aRC pre-
cursor. *Closed squares* and *circles* represent the 4EM (U44C/U97A/U159G/
G163C) and G150A ribozymes, respectively

the autorecombining form ΔP6aRC (Fig. 1b), the reverse-
splicing activity of the covalent ribozyme was reduced
roughly 7-fold (Fig. 2). As a means to restore this activity,
we generated a pool of ribozymes mutagenized at 10% per
position by PCR and selected for reverse splicing ability
under optimal autorecombination conditions (48°C in 100-
mM $MgCl_2$ and 30-mM EPPS; pH 7.5), keeping selection
pressure constant by allowing 15 min for folding and
reaction for every round. We started a more stringent
selection line, based on the generation 6 population that
was incubated at 48°C for just 2 min, but this line was
unable to accommodate the selection criteria.

Genotyping of Selection Population

After ten rounds of selection, we cloned and determined
the sequence of 23 individuals from the G_{10} population
(Fig. 3a). Six genotypes were identical to that of the initial
ΔP6aRC ribozyme, which was not unexpected given that it
possesses some catalytic activity. Yet two mutations were
observed with moderate frequency: U44C and G150A (five
times each, Fig. 3b). We transcribed and assayed for
reverse-splicing ability those genotypes that contained
either of these mutations (Fig. 3a). We chose the two most
active of these: a four-error mutant (U44C/U97A/U159G/
G163C, referred to as 4EM hereafter) and a single-error
mutant (G150A), for further characterization alongside the
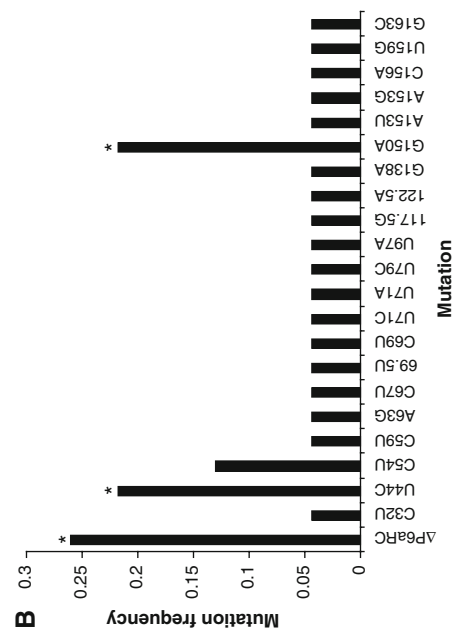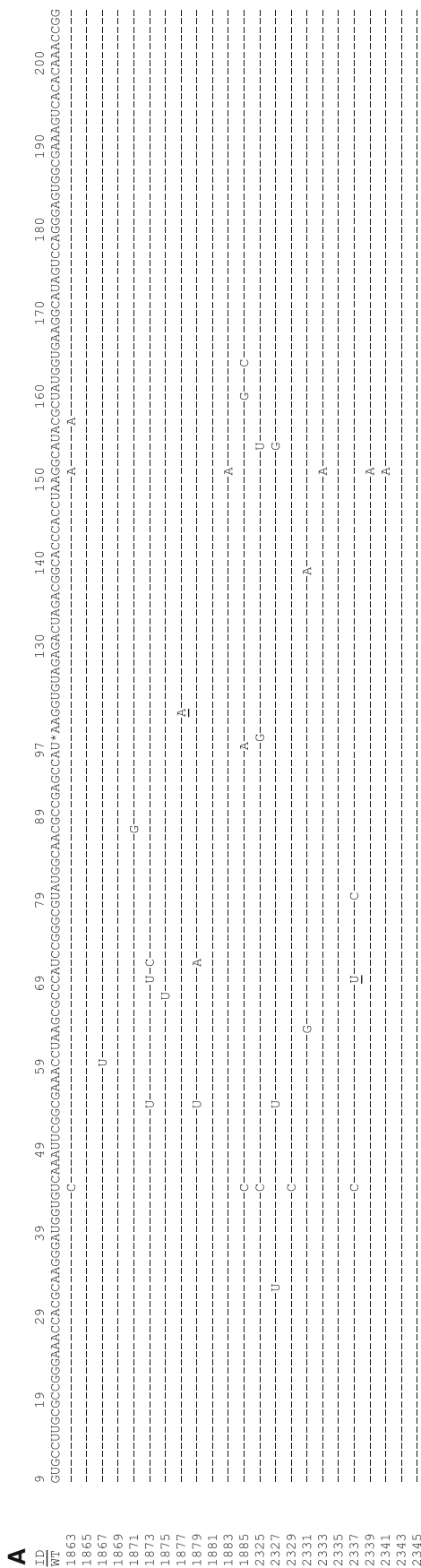full-length wildtype *Azoarcus* ribozyme, and two less

Fig. 3 Genotyping of the generation 10 selection population. **a** Sequence alignment of 23 *Azoarcus* ribozyme ΔP6aRC-truncated isolates obtained through in vitro selection. The *asterisks* (*) in the wildtype (wt) sequence indicates the point of truncation. The units digit of numbers above the sequence indicates that position in the sequence (i.e., the '0' of 160 is above nucleotide 160). *Underlined* nucleotides represent insertions in the sequence and occur after the position in the sequence they occupy. **b** Mutation frequency in the G_{10} population. Mutations with an *asterisk* (*) above them were characterized further. Those listed at decimal positions represent insertions into the ΔP6aRC sequence, between the two flanking integers (i.e., 69.5 U is an insertion of U between nucleotides 69 and 70). The nucleotide numbering in Fig. 1 has been retained for clarity

Table 1 Catalytic parameters of *Azoarcus* ribozyme variants

Ribozyme genotype	k_{obs} (min^{-1})	Fraction correctly folded (15/60 min)
Wildtype	0.079	0.89/0.93
ΔP6aRC	0.080	0.61/0.73
G150A	0.24	0.79/0.79
U44C/G150A	0.11	0.77/0.85
4EM	0.13	0.85/0.9

active variants, the ΔP6aRC starting molecule, and a two-error mutants (U44C/G150A).

General Characterization of Selection Mutants

To investigate how the most-active ribozymes adapted to the selection conditions, we determined k_{obs} values for each ribozyme through single-turnover kinetics, and the maximum fraction of ribozymes able to adopt a catalytically active fold by incubation in the presence of excess substrate (Table 1). The decrease in reverse-splicing activity between the wildtype ribozyme and ΔP6aRC is apparently not due to a difference in k_{obs}. Instead, the reduced catalytic ability of ΔP6aRC likely a consequence of a lessened propensity to fold into a catalytically active conformation. This phenomenon has been seen before in the behaviors of other pairs of related functional RNAs (Schmitt and Lehman 1999; Huang et al. 2009). Following incubation with excess substrate in 100-mM MgCl$_2$ and 30-mM EPPS (pH 7.5) for the time allowed in selection process, 15 min, only 61% of ΔP6aRC have reacted when compared to 89% of the wildtype ribozyme molecules. After 60 min, those percentages increase to 73 and 93 indicating that the ΔP6aRC population folds more slowly and less completely than the wildtype ribozyme, which has nearly reached full reactivity after just 15 min.

The selection mutants seem to exploit both folding and kinetic improvements to accommodate the selection criteria; G150A, U44C/G150A, and 4EM adapted to the selection through improved kinetics, increasing their k_{obs} values by 3-, 1.4-, and 1.6-fold, respectively, over that of the wildtype ribozyme. The mutants all fold better than the ΔP6aRC form, although none reaches the same fraction of correctly folded molecules as the wildtype ribozyme. These kinetic enhancements would be sufficiently large to drive a population towards fixation rapidly. Assuming, in the case of the G150A mutant, that a 3-fold kinetic enhancement gives a 3-fold selection advantage to that mutant, then we can approximate the number of generations it would take for that mutation to become the dominant sequence according to basic population-genetic theory of haploids:

$$\ln\left(\frac{p_t}{q_t}\right) = \ln\left(\frac{1-\mu}{\mu}\right) + t \ln w$$

where p_t and q_t are the frequencies of wildtype and mutants at generation t, respectively, μ is the mutation rate to the mutant allele, and w is the relative fitness disadvantage that the wildtype has to the mutant. If we assume a 10% mutation rate per position per generation, not unreasonable in prebiotic conditions, then the G150A mutation would occur in ∼3.3% of progeny molecules. In this scenario, with minimal epistatic effects, over 99% of the species present would contain the G150A mutation after only eight generations of direct competition.

Analysis of the Selection-Derived Mutations

Perhaps the most interesting mutant is the G150A molecule because it is the kinetically fastest ribozyme identified through the selection and folds the most quickly, reaching its reaction equilibrium by 15 min. This mutation is in the P8 tetraloop receptor and G150 has been identified as a monovalent metal ion binding site (Basu et al. 1998). Considering that we performed the selection in much higher Mg^{2+} concentrations than occur in vivo and in the absence of any monovalent ions, this mutation likely stabilizes formation of the L2–P8 tetraloop/receptor interaction, aiding faster and more accurate folding. Extra stabilization here may also help counter the detrimental effects of losing a portion of the P4–P6 fast-folding domain. Less intuitive, however, is how this mutation provides such a dramatic improvement to k_{obs}; although, presumably the new L2–P8 interaction may have long-range effects that stabilize the active site in the presence of elevated Mg^{2+}.

The most active overall mutant, 4EM, contains a suite of mutations (U44C/U97A/U159G/G163C). One, U159G, may have a similar effect as G150A, as it lies directly opposite G150 and may allow access to conformations more applicable to Mg^{2+}-binding or stacking with A148 and A149. A second mutation, U97A, reverses the A97U mutation we introduced in the ΔP6aRC variant to make it autorecombination-compatible. Perhaps, the resulting AAAA L6 is more stable than the original AUAA loop. Alternatively, the mutation could be essentially neutral. The remaining mutations, U44C and G163C, appear to work together to prevent misfolding. Because nucleotides 139–142 (P3) and 143–146 (P8) are both 5′-CACC-3′, their corresponding base-pairing partners (40–43 and 162–165, consisting of 5′-GGUG-3′), have the potential to incorrectly base-pair. In general, formation of the P3–P7 pseudoknot is a limiting step in group-I intron folding (Zarrinkar and Williamson 1994; Pan and Woodson 1998; Sclavi et al. 1998; Pan et al. 2000; Treiber and Williamson

2001; Zhang et al. 2005). Native P3 formation is bolstered by the U44C mutation, which trades the U•G wobble for a C–G basepair. Both the non-native "P3" (162–165 base-pairing with 139–142) and native P8 are destabilized by G163C. Both the effects combine to allow the critical P3–P7 pseudoknot to form more frequently, helping to offset the loss of part of the P4–P6 fast-folding domain and result in a better folding ribozyme population.

Covalent Self-Assembly is Improved by Incorporating Selection-Derived Mutations

The mutations identified through in vitro selection improved the reverse-splicing ability of the ΔP6aRC ribozyme. However, reverse-splicing ability is only one of several properties required for covalent self-construction and autorecombination as observed in the *Azoarcus* ribozyme. The necessary fragments must also be able to fold correctly, both individually and in conjunction with their cognate pieces to form the *trans* autorecombinase, and then perform both forward- and reverse-splicing reactions. We thus turned a rational eye towards which of the selection mutations would actually benefit the self-construction process. Within 4EM, three of the four mutations were deemed to be deleterious to covalent self-assembly: U97A because it removes the L6 CAU recombination tag and both U159G and G163C because they almost certainly would destabilize the necessary *trans* complex formation. Instead, we chose to focus on the two mutations that were the most common during the selection: G150A, which had the largest effect on reverse-splicing kinetics, and U44C, which would likely enhance *trans* complex formation.

We first attempted simple two-piece self-construction reactions with **W** + *h*•**ΔXΔYZ** and **W$_{U44C}$** + *h*•**ΔXΔYZ** to verify that the shortened system was still active (Fig. 4). Surprisingly, the ΔP6aRC construct was as active as the U44C version in this context, despite its notably lower reverse-splicing ability. This parity was not maintained, however, in three-piece self-construction reactions. We tested two different connectivities, **W** + *h*•**ΔX** + *h*•**ΔYZ** and **W** + *h*•**ΔXΔY** + *h*•**Z**, with drastically different results. For the first connectivity set, the effects of the U44C and G150A mutations are roughly additive, giving a robust yield of ∼8.5% after 6 h, equal to or greater than the sum of either mutant alone, and about 10-fold greater than the ΔP6aRC. (The recombinant product from **W$_{U44C}$** + *h*•**ΔX** + *h*•**ΔYZ** was excised from the gel and genotyped to confirm it was the correct molecule.) The second set, however, was only capable of detectable self-construction with the G150A mutation, and was rendered essentially inactive with the ΔP6aRC, U44C, and U44C/G150A double-error mutant molecules. One factor for the difference in activities observed between these two
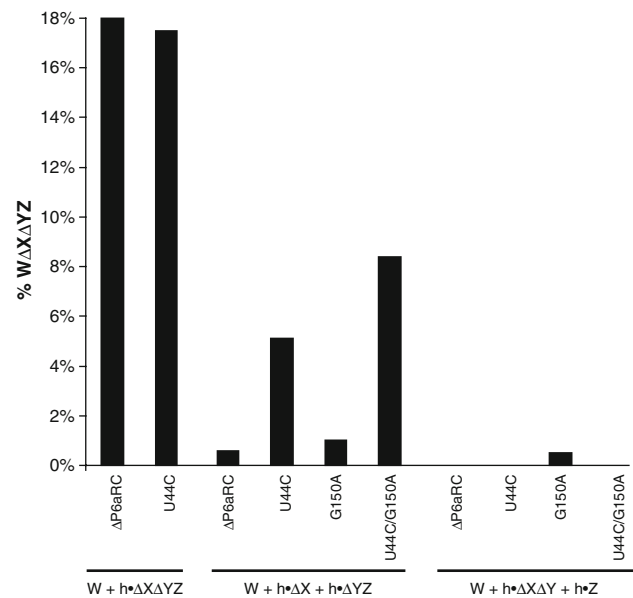

Fig. 4 Effects of selection-derived mutations on the self-construction activity of the *Azoarcus* ΔP6aRC ribozyme. Self-construction reactions contained 2 μM of each fragment and were reacted at 48°C in 100-mM MgCl₂ for 6 h. The number and breakpoints of the fragments are listed below the reactions. Products were separated on 8% polyacrylamide/8-M urea gels. **W**-fragments were 5′-radiolabeled, and the percent of **WΔXΔYZ** was calculated as the ratio between full-length molecules to the sum of all **W**-containing (radiolabeled) molecules

connectivity groups may be a lack of *trans* recombinase enzymes resulting from difficulty in forming the native non-covalent **ΔY**–**Z** interaction, something overcome to a minor extent by the G150A mutation. Additionally, it is possible that one or both parts of the **ΔY**–**Z** junction (the 3′-end of **ΔY** or the 5′-end of *h*•**Z**) are poor substrates for recombination, and the reaction is limited by the reactivity at this junction. Both factors probably play a role, because only ∼1.4% of **WΔXΔY$_{G150A}$** is obtained after 6 h, compared to **WΔX** yields of 7–9% in the **W** + *h*•**ΔX** + *h*•**ΔYZ** system (depending on which mutations are included). Four-piece self-construction reactions, **W** + *h*•**ΔX** + *h*•**ΔY** + *h***Z** were attempted, but no full-length ribozymes—and only trace amounts of **WΔXΔY** (in any of the ΔP6aRC or mutant systems)—were obtained (data not shown).

Conclusion

Using strictly Darwinian in vitro selection techniques, it is not obvious how one could directly improve autorecombinase ability—the majority of the full-length products from a randomized pool of RNA fragments that cooperate to perform covalent self-assembly would tend to be the best

recombination *substrates*, not the best *catalysts*. We thus performed in vitro selection to improve the reverse-splicing activity of the *Azoarcus* ΔP6aRC deletion mutant with the idea that some of the mutations enhancing catalytic ability would also augment self-construction activity. The selection-isolated mutants were able to regain some of the catalytic ability lost by the deletion of P6a by improving both their kinetic parameters and overall folding efficiency, to varying degrees. Some of the reverse-splicing mutations were in fact exaptations (Gould and Vrba 1982), helping to restore self-construction ability in three-piece reactions despite not being evolved for that specific purpose. Our attempt to improve the autorecombination system described above highlights the unique difficulty of evolution in an RNA World context in that mutations that improve or expand the catalytic activity of a ribozyme must not too severely hamper its ability to be acted upon as a substrate and vice versa. Our findings underscore the requisite for multifunctionality in primordial genetic molecules; the ability of RNA or an RNA-like molecule to acquire self-replication required selection to favor catalysis, structure, and cooperation simultaneously prior to the advent of purely selfish genomes.

References

Basu S, Rambo RP, Strauss-Soukup J, Cate JH, Ferré-D'Amaré AR, Strobel SA, Doudna JA (1998) A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. Nat Struct Biol 5:986–992

Bell MA, Johnson AK, Testa SM (2002) Ribozyme-catalyzed excision of targeted sequences from within RNAs. Biochemistry 41:15327–15333

Boussau B, Blanquart S, Necsulea A, Lartillot N, Gouy M (2008) Parallel adaptations to high temperatures in the Archaean eon. Nature 456:942–945

Burke DH, Willis JH (1998) Recombination, RNA evolution, and bifunctional RNA molecules isolated through chimeric SELEX. RNA 4:1165–1175

Burton AS, Lehman N (2006) Calcium(II)-dependent catalytic activity of the *Azoarcus* ribozyme: testing the limits of resolution for in vitro selection. Biochimie 88:819–825

Burton AS, Madix RA, Vaidya N, Riley CA, Hayden EJ, Chepetan A, Dìaz Arenas C, Larson BC, Lehman N (2009) Gel purification of radiolabeled nucleic acids via phosphorimaging: Dip-N-Dot. Anal Biochem 388:351–352

Cadwell RC, Joyce GF (1992) Randomization of genes by PCR mutagenesis. PCR Methods Appl 2:28–33

Dotson PP, Johnson AK, Testa SM (2008) *Tetrahymena thermophila* and *Candida albicans* group I intron-derived ribozymes can catalyze the trans-excision-splicing reaction. Nucleic Acids Res 36:5281–5289

Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465–523

Engels J, Uhlmann E (1988) Gene synthesis. Adv Biochem Eng Biotechnol 37:73–127

Ferris JP (2002) Montmorillonite catalysis of 30–50 mer oligonucleotides: laboratory demonstration of potential steps in the origin of the RNA world. Orig Life Evol Biosph 32:311–332

Gesteland R, Cech TR, Atkins J (eds) (2006) The RNA World, 3rd edn. Cold Spring Harbor Press, New York

Gilbert W (1986) The RNA world. Nature 319:618

Gould SJ, Vrba ES (1982) Exaptation-A missing term in the science of form. Paleobiology 8:4–15

Hayden EJ, Lehman N (2006) Self-assembly of a group I intron from inactive oligonucleotide fragments. Chem Biol 13:909–918

Hayden EJ, Riley CA, Burton AS, Lehman N (2005) RNA-directed construction of structurally complex and active ligase ribozymes through recombination. RNA 11:1678–1687

Huang Z, Szostak J (1996) A simple method for 3′-labeling of RNA. Nucl Acids Res 24:4360–4361

Huang Z, Pei W, Han Y, Jayaseelan S, Shekhtman A, Shi H, Niu L (2009) One RNA aptamer sequence, two structures: a collaborating pair that inhibits AMPA receptors. Nucl Acids Res 37:4022–4032

Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP (2001) RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science 292:1319–1325

Joyce GF (1991) The rise and fall of the RNA world. New Biol 3:399–407

Joyce GF (1998) Nucleic acid enzymes: playing with a fuller deck. Proc Natl Acad Sci USA 95:5845–5847

Joyce GF (2002) The antiquity of RNA-based evolution. Nature 418:214–221

Kuhn H (1972) Self-organization of molecular systems and evolution of the genetic apparatus. Angew Chem Int Ed Engl 11:798–820

Lehman N (2003) A case for the extreme antiquity of recombination. J Mol Evol 56:770–777

Lehman N (2008) A recombination-based model for the origin and early evolution of genetic information. Chem Biodivers 5:1707–1717

Lehman N, Joyce GF (1993) Evolution in vitro of an RNA enzyme with altered metal dependence. Nature 361:182–185

McGinness KE, Joyce GF (2003) In search of an RNA replicase ribozyme. Chem Biol 10:5–14

Orgel LE (1986) RNA catalysis and the origins of life. J Theor Biol 123:127–149

Pan J, Woodson SA (1998) Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. J Mol Biol 280:597–609

Pan J, Deras ML, Woodson SA (2000) Fast folding of a ribozyme by stabilizing core interactions: evidence for multiple folding pathways in RNA. J Mol Biol 296:133–144

Rich A (1962) On the problems of evolution and biochemical information transfer. In: Kasha M, Pullmann B (eds) Horizons in biochemistry. Academic Press, New York, pp 103–126

Riley CA, Lehman N (2003) Generalized RNA-directed recombination of RNA. Chem Biol 10:1233–1243

Rydzanicz R, Zhao XS, Johnson PE (2005) Assembly PCR oligo maker: a tool for designing oligodeoxynucleotides for constructing long DNA molecules for RNA production. Nucleic Acids Res 33:W521–W525

Schmitt T, Lehman N (1999) Non-unity molecular heritability demonstrated by continuous evolution in vitro. Chem Biol 6:857–869

Sclavi B, Sullivan M, Chance MR, Brenowitz M, Woodson SA (1998) RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. Science 279:1940–1943

Striggles JC, Martin MB, Schmidt FJ (2006) Frequency of RNA–
        RNA interaction in a model of the RNA World. RNA 12:353–
        359
Szostak J, Ellington A (1993) In vitro selection of functional RNA
        sequences. In: Gesteland R, Atkins J (eds) The RNA world, 1st
        edn. Cold Spring Harbor Press, New York, pp 511–533
Treiber DK, Williamson JR (2001) Beyond kinetic traps in RNA
        folding. Curr Opin Struct Biol 11:309–314
Vartanian JP, Henry M, Wain-Hobson S (1996) Hypermutagenic PCR
        involving all four transitions and a sizeable proportion of
        transversions. Nucleic Acids Res 24:2627–2631

Zaher HS, Unrau PJ (2007) Selection of an improved RNA
        polymerase ribozyme with superior extension and fidelity.
        RNA 13:1017–1026
Zarrinkar P, Williamson JR (1994) Kinetic intermediates in RNA
        folding. Science 265:918–924
Zhang L, Xiao M, Lu C, Zhang Y (2005) Fast formation of the P3–P7
        pseudoknot: a strategy for efficient folding of the catalytically
        active ribozyme. RNA 11:59–69