Chemistry & Biology, Vol. 10, 799–806, September, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S1074-5521(03)00171-6

Substrate 2′-Hydroxyl Groups Required for Ribozyme-Catalyzed Polymerization

Ulrich F. Müller and David P. Bartel*
Whitehead Institute
9 Cambridge Center
Cambridge, Massachusetts 02142

Summary

A polymerase ribozyme has been generated that uses nucleoside triphosphates to elongate an RNA primer by the successive addition of nucleotides complementary to an RNA template. Its polymerization is accurate, with an average error rate less than 3%, and it is general in terms of the sequence and the length of the primer and template RNAs. To begin to understand how the substrate contacts contribute to this accurate and general activity, we investigated which primer and template 2′-hydroxyl groups are involved in substrate recognition. We identified eight positions where 2′-deoxy substitutions can influence polymerization kinetics. All eight are within five nucleotides of the primer 3′ terminus. Some, but not all, of the 2′-deoxy effects appear to be sequence dependent. These results begin to build a picture of how the polymerase ribozyme recognizes its substrates.

Introduction

During a hypothetical era in the early evolution of life, known as the RNA world, RNA is thought to have served both as carrier for genetic information and as biological catalyst [1–4]. Catalytic RNAs (ribozymes) exist in nature, supporting this idea [5, 6], and artificial ribozymes can be obtained by in vitro selection methods from pools of random sequences to expand the known catalytic repertoire of RNAs [7, 8]. To generate a set of ribozymes that is capable of self-replication and evolution, the choice of ribozymes would depend in part on the available precursor molecules. However, for its replication, every such system needs an activity that polymerizes mono- or oligonucleotides [9, 10]. If created, such a system would support the RNA world hypothesis and could be used to investigate fundamental questions in biology, such as which network topologies are suited for self-replication and evolution of molecular networks [11, 12].

An RNA-dependent RNA polymerase ribozyme was recently developed, starting with a ribozyme selected from random sequences and improving this activity by rational design and further randomization and selection [13]. This ribozyme can extend the 3′ terminus of an RNA primer by the successive and templated addition of up to 14 nucleotides. Its polymerization activity is general in terms of the sequence and the length of the primer and template RNAs, provided that the 3′ terminus of the primer pairs with the template. Its polymerization is quite accurate, with error rates between 0.0004 and 0.079, depending on the templating nucleotide. Its ability to use different sequences requires a set of sequence-independent contacts to the substrate; its high fidelity requires precise substrate positioning via those contacts.

The 2′-hydroxyl groups of RNA have important roles in ribozyme substrate positioning and more generally in RNA structure. The RNase P RNA and the group I self-splicing intron largely depend on contacts to substrate 2′-hydroxyl groups. The RNase P ribozyme specifically recognizes the tertiary structure of pre-tRNAs with the help of several 2′-hydroxyl contacts, but requires the pre-tRNA CCA-3′ terminus for optimal activity [14, 15]. The group I intron recognizes the helix that presents the 5′ splice site using four discrete 2′-hydroxyl contacts [16–18] and additionally requires sequence-specific contacts to a G:U base pair [17, 19]. Two of those four 2′-hydroxyl contacts appear to be part of a ribose zipper motif [20]. The ribose zipper motif is defined by the interaction of four 2′-hydroxyl groups, in which two consecutive 2′-hydroxyl groups on one strand interact with two consecutive 2′-hydroxyl groups of an antiparallel, interacting strand [21, 22]. The ribose zipper motif often coincides with another tertiary structural motif that uses 2′-hydroxyl contacts, the A-minor motif. This motif involves the insertion of adenines in the minor grove of neighboring helices, where they form hydrogen bonds with one or both of the helix 2′-hydroxyl groups [23]. A-minor motifs tend to cluster, often in the form of ribose zipper motifs. The importance of these two motifs for RNA tertiary structures can be seen from their abundance and conservation in ribosomal RNA, which has dozens of ribose zipper [22] and A-minor motifs [23].

Other RNA binding macromolecules in addition to ribozymes employ 2′-hydroxyl groups for sequence-independent RNA recognition. Many proteins that bind double-stranded RNA (dsRNA) contain a sequence called the dsRNA binding motif, dsRBM [24, 25], which relies mostly on 2′-hydroxyl contacts [26]. Proteinaceous RNA-dependent RNA polymerases (RdRps) also use 2′-hydroxyl groups for substrate recognition. Those enzymes can discriminate between RNA and DNA templates. However, only very few substrate 2′-hydroxyl groups contribute to substrate recognition [27, 28], and 2′-hydroxyl groups are not always required for polymerization [27, 29].

To investigate the involvement of 2′-hydroxyl groups in the substrate recognition of the polymerase ribozyme, we used substrates with single 2′-deoxy modifications and quantified the effects on polymerization kinetics. The results showed that the ribozyme uses defined 2′-hydroxyl groups to recognize its substrate. All of those contacts to 2′-hydroxyl groups in the single-stranded region are sequence dependent. In contrast, only a part of the contacts in the double-stranded region appear to be sequence dependent.

Results

To identify substrate 2′-hydroxyl groups involved in ribozyme-substrate contacts, we used substrates with sin-

*Correspondence: dbartel@wi.mit.edu

Table 1. Primers and Templates Used in This Study

Sequence 1
All-Ribo Primers
(8 nt) 5′-CUGCCAAC-3′
(9 nt) 5′-CUGCCAACC-3′
(10 nt) 5′-CUGCCAACCG-3′
(11 nt) 5′-CUGCCAACCGU-3′
(12 nt) 5′-CUGCCAACCGUG-3′
(13 nt) 5′-CUGCCAACCGUGC-3′
(14 nt) 5′-CUGCCAACCGUGCG-3′
Primers with 2′-Deoxy Substitutions
5′-CUGCCAACCG-3′
5′-CUGCCAACCG-3′
5′-CUGCCAACCG-3′
5′-CUGCCAACCG-3′
5′-CUGCCAACCG-3′
All-Ribo Template
3′-GACGGUUGGCACGCUUCGCAG-5′
Templates with 2′-Deoxy Substitutions
3′-GACGGUUGGCACGCUUCGCAG-5′
3′-GACGGUUGGCACGCUUCGCAG-5′
3′-GACGGUUGGCACGCUUCGCAG-5′
3′-GACGGUUGGCACGCUUCGCAG-5′
3′-GACGGUUGGCACGCUUCGCAG-5′
Sequence 2
5′-CUGCCAACUGUG-3′
3′-GACGGUUGACACGCUUCGCAG-5′
3′-GACGGUUGACACGCUUCGCAG-5′

Positions of 2′-deoxy substitutions are underlined. Using different combinations of primers and templates, the substitution on the template strand can be set to all positions from −10 nt to +10 nt relative to the 3′ end of the primer. For sequence 2, the primer is shown on top of the template sequences. Sequence 2 is identical to sequence 1, but carries a C9U mutation in the primer and a corresponding G*13*A mutation in the template.

gle 2′-deoxy substitutions and studied their effects on polymerization kinetics. The effect of 2′-deoxy modifications on ribozyme kinetics has been used before to study ribozyme-substrate interactions [17]. It is especially suitable for the characterization of a polymerase because a single modification can be used to study several consecutive contacts during the progression of polymerization.

Six modified templates were employed in this study (Table 1). By the use of primers with different lengths, the 3′ end of the primer was set to every position relative to a 2′-deoxy substitution in the template strand, between 10 nt upstream and 10 nt downstream. Additionally, five modified primers were used (Table 1). The position of 2′-deoxy substitutions was numbered relative to the 3′ terminus of the primer (Figure 1, bold numbers). The 3′-terminal nucleotide of the primer was labeled as "0," as was the template nucleotide that pairs to it. The single-stranded part of the substrate was assigned positive numbers and the double-stranded part negative numbers. During polymerization, the 3′ end of the primer elongates, and the 2′-deoxy substitutions, fixed in the substrate sequence, shift relative to the new 3′ terminus of the primer strand. Because the ribozyme had to act on each new 3′ terminus, the modifications successively probed possible contacts to the ribozyme, if we assume that the active site is recognizing a uniform substrate duplex with analogous successive contacts at each register. The kinetic analysis of each intermediate monitored the strength of those possible contacts because the substrate affinity is limiting with respect to polymerization efficiency [30].

The progress of polymerization was followed by the quantitation of polymerization intermediates. The accumulation patterns differed between the all-ribose template and a template that contained a 2′-deoxy substitution. For example, with a 2′-deoxy substitution at the initial template position −1, the second intermediate accumulated more than with an all-ribose substrate (Figure 1A). In the second intermediate, the 2′-deoxy modification had shifted to position −3. The accumulation of this specific intermediate showed that the template position −3 required a 2′-hydroxyl group for most efficient polymerization (Figure 1A).

To quantitate the kinetic effects of specific 2′-deoxy substitutions, we measured the band intensities and fitted different kinetic models to the data. A kinetic model with only one variable for each polymerization step did not fit the data well (not shown). A kinetic model with an additional variable that accounted for the partial processivity of the reaction [30] fit the data well, but gave indeterminate solutions for products with more than two nucleotides added (data not shown). A third method was chosen to process the data. First, the fraction of a certain polymerization intermediate that reacted between two time points was calculated ("reacting fraction"; Figure 1B). Second, the reacting fraction values were averaged over the timecourse (Figure 1C). This average was used to describe the reactivity for each intermediate. Third, each value for an all-ribose intermediate was divided by the corresponding value of the 2′-modified intermediate. Each quotient showed how many-fold the polymerization rate was reduced by the 2′-deoxy substitution at the respective position (Figures 2 and 3).

The largest kinetic influence of a 2′-hydroxyl group (100-fold) was at the 3′ end of the primer (position 0). This can be explained by the inductive/resonance effects of cis-1,2-diols. The presence of a 2′-hydroxyl group decreases the pK_A of the adjacent 3′-hydroxyl group by about 3 pH units [31, 32], thereby influencing the nucleophilicity of the reacting 3′-oxygen, which could account for the 100-fold decrease in polymerization kinetics. Alternatively, the terminal 2′-hydroxyl group could be involved in binding a metal ion at the catalytic site [33].

The influences of all other substrate 2′-hydroxyl groups were 8-fold or less. These influences can be explained by the formation of hydrogen bonds because one hydrogen bond in a hydrophilic microenvironment can shift equilibrium constants between 2- and 15-fold [34–36]. We considered 2′-hydroxyl groups as relevant if their deoxy modification resulted in at least 2-fold slowing of the polymerization [37]. This applied for the positions 0, −1, and −2 on the primer (Figure 2) and −3, −2, +3, +4, and +5 on the template (Figure 3).

The use of primers with different lengths and templates with 2′-deoxy substitutions at different positions enabled contacts to be examined in different sequence contexts. Several 2′-deoxy effects were strongly se-

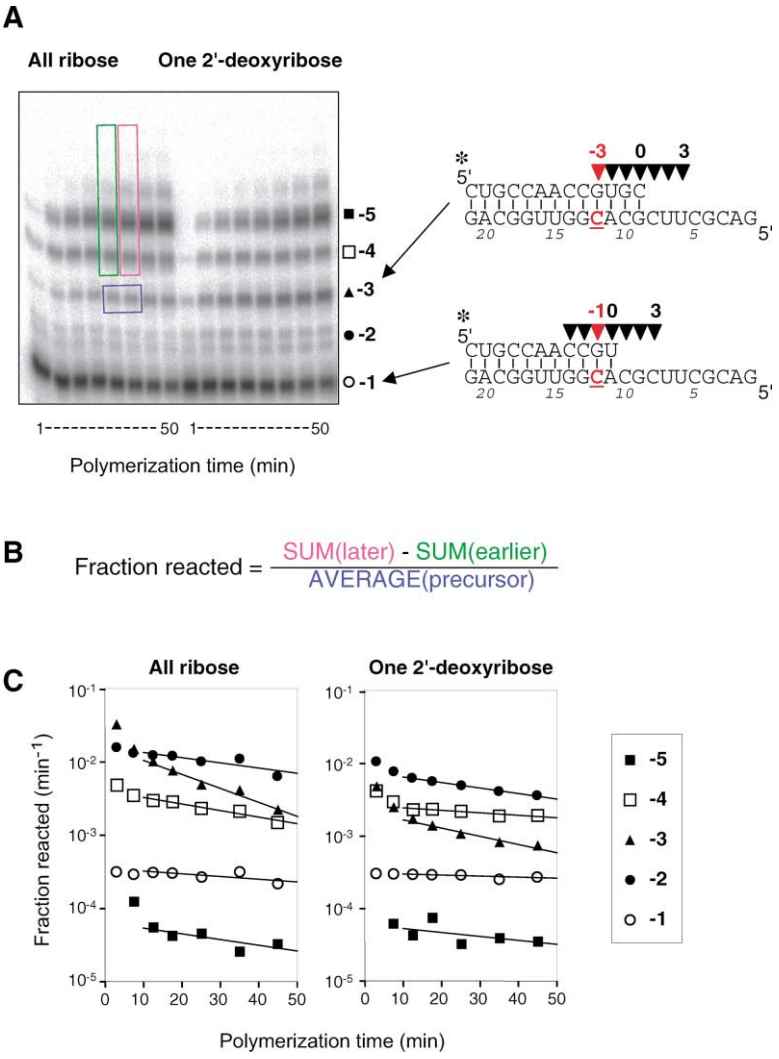

Figure 1. Polymerization Timecourse with an Unmodified Substrate and a Substrate that Carries a 2′-Deoxy Modification in the Template

The primer was 5′ [^{32}P]-labeled and the polymerization products were separated by denaturing PAGE.

(A) PhosphorImage of a typical timecourse (1–50 min). The product increase differed between the time course of the unmodified substrate and 2′-deoxy modified substrate. The first nucleotide addition resulted in a double band, which was caused by a misincorporated nucleotide (see Discussion). The sequences of the starting primer template and its polymerization intermediate extended by two nt are shown. The bold numbers denote the nucleotide positions relative to the 3′ terminus of the primer, illustrating the shifting position of the 2′-deoxy modification (red) during polymerization. The italicized numbers denote the identity of the template nucleotides (referred to in the legend of Figure 3).

(B) The method that was used to quantify the reaction efficiency. The colored rectangles illustrate the bands for the analysis of intermediate +2 between the fifth and sixth time point. The "reacting fraction" is the proportion of a polymerization intermediate that was extended by one or more nucleotides between two time points.

(C) Quantifying the kinetic effects of 2′-deoxy substitutions. The reacting fraction is plotted as a function of polymerization time for each intermediate. By exponential fitting (lines in the semi-logarithmic plot), the geometric means of the data were calculated, correcting for the uneven distribution of the time points (see Experimental Procedures). The data of different intermediates are labeled by the position of their 2′-deoxy substitutions (−1, open circle; −2, filled circle; −3, filled triangle; −4, open square; and −5, filled square). The drop in average reactivity can be seen when the 2′-deoxy group is in position −2 (2.4-fold) and −3 (5-fold).

quence dependent. The most obvious sequence dependence was in the single-stranded part of the substrate. The involvement of 2′-hydroxyl groups +3, +4, and +5 was present in one single-stranded region of the template (Figure 3, upper panel) and absent in a different single-stranded region of the template (Figure 3, middle panel). The two different template segments probably favored different conformations and established different hydrogen bonds to the ribozyme. In contrast, the 2′-deoxy effect at template position −3 was present in all three contexts studied (Figure 3). The sequence-independent geometry of double-stranded RNA could readily explain analogous contacts of the ribozyme to different primer/template sequences. However, at another position in the double-stranded part of the template, the 2′-deoxy effect was sequence dependent. When a 2′-deoxy C*12* was at position −2, the effect was 1.33-fold (+0.08/−0.07; Figure 3), whereas when a 2′-deoxy C*10* was at position −2, the effect was 3.7-fold (+0.7/−0.6; data not shown). The nucleotide that carried the 2′-deoxy substitution was a C in both cases and thus could not have caused the difference. The 3′-proximal nucleotide, which was an A for the strong effect and a G for the weak effect, seemed like a plausible candidate to be responsible for the different 2′-deoxy effects. Perhaps the exocylic amino group of the 3′-proximal guanine substituted for the missing 2′-hydroxyl group. Because adenine does not carry this exocyclic amino group, it would not have substituted. To test this hypothesis, we changed the relevant guanosine in sequence 1 to an adenosine, expecting to restore the effect of the 2′-deoxy modification. However, the modification showed no effect (Figure 3, sequence 2). Therefore, the sequence dependence of the 2′-deoxy effect at position −2 requires a more complex explanation than the hypothesized local interactions.

The effects of 2′ substitutions can be monitored at successive nucleotide additions in a single polymerization reaction. This allowed the same 2′-deoxy effect to be examined in two different scenarios. In one, the added nucleotide was the first to be added to the input primer (Figure 3, red columns). In the other, the added

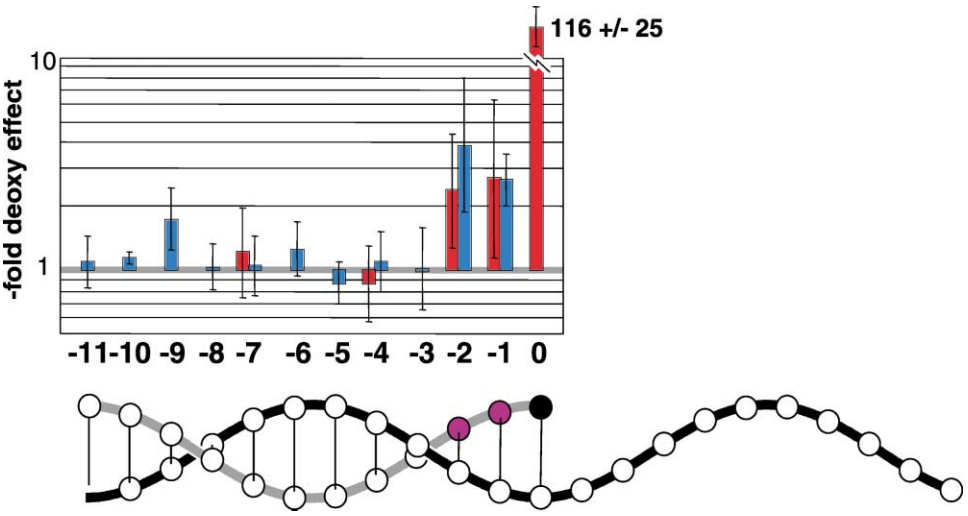

Figure 2. Influences of 2′-Deoxy Substitutions in the Primer on Polymerization Kinetics

Significant effects are in the positions 0, −1, and −2. The red columns show effects on the addition of the first nucleotide (nonprocessive); the blue columns show effects on later additions (partially processive). Unlike in Figure 3, the blue and red columns in one position describe the effects in different sequence contexts. Error bars are standard deviations of 3 to 10 experiments.

nucleotide was the second or later to be added (Figure 3, blue columns). In general, both scenarios examined the effect in the identical sequence contexts, but the second scenario monitored the extension after the input primer had already been extended by the polymerase.

The 2′-deoxy effects generally were the same for the first nucleotide added to the primer and later nucleotides added (Figure 3, red and blue columns). However, a significant difference was observed in the template strand at position +3 (sequence 1a). Here, a 2′-deoxy substitution did not slow the first nucleotide addition (1.18 +0.16/−0.14), but it slowed later nucleotide additions (3.15 +1.5/−0.8). The differential 2′-deoxy effect at position +3 can be explained by the polymerization processivity. The processivity in a polymerization reaction describes what fraction of substrate stays bound to the ribozyme between two catalytic events. The first nucleotide addition (red columns) is nonprocessive because there is no previous step and the fraction of prebound substrate is negligible [30]. The later nucleotide additions are partially processive, with a processivity that is specific for each sequence context [30]. If a nucleotide is added processively, the substrate stays bound between two catalytic events and the 2′-deoxy effect on the preceding event could substantively influence the amount of bound primer template for the next nucleotide addition, which in turn would influence the accumulation of the next polymerization intermediate. This can explain the difference in 2′-deoxy effects at template position +3 of sequence 1a (Figure 3): The 2′-deoxy effect in the partially processive addition (blue) describes mainly the preceding nucleotide addition (2′-deoxy at position +4) while the 2′-deoxy effect in the first addition (red) characterizes the current addition (2′-deoxy at position +3). The processivity in the observed nucleotide addition is 0.92 at our conditions and NTP concentrations [30]. That is, in 92% of the first nucleotide additions, a second nucleotide gets added without release of the substrate. Therefore, over the analyzed timecourse, the large majority of the later additions is to a primer template that remained bound to the ribozyme, and the apparent 2′-deoxy effect at position +3 in the later addition can reflect the 2′-deoxy effect at position +4.

In contrast to the evaluation of 2′-deoxy effects in the template strand (Figure 3), the evaluation of 2′-deoxy effects in the primer strand (Figure 2) used only primers with a length of 10 nt. Therefore, the first nucleotide addition (red column) and later nucleotide additions (blue column) in the same position are not based on the same templating nucleotide. In other words, the blue column and the red column describe a different sequence context. The geometric mean for the blue and the red columns in position −1 are identical (Figure 2). The same is true at primer position −2. This appears to suggest that at both of these positions, the 2′-deoxy effects are not sequence dependent. However, the experimental variation at these positions makes such a conclusion tenuous (Figure 2).

Discussion

The polymerase ribozyme does not require particular substrate sequences [13]. Therefore, it is interesting to look at how the polymerase ribozyme recognizes its substrates. In contrast to protein enzymes that can use positively charged residues to establish sequence-nonspecific ionic interactions to the phosphate backbone, a ribozyme has to rely on hydrogen bonds and metal-mediated contacts. The hydrogen bonds can involve ribose hydroxyl groups, phosphate oxygens, or functional groups of bases. The present study shows that defined substrate 2′-hydroxyl groups are used to mediate substrate recognition. The results suggest that the polymerase ribozyme forms four contacts to 2′-hydroxyl groups of the double strand, at least one of which is present in multiple sequence contexts, and three con-

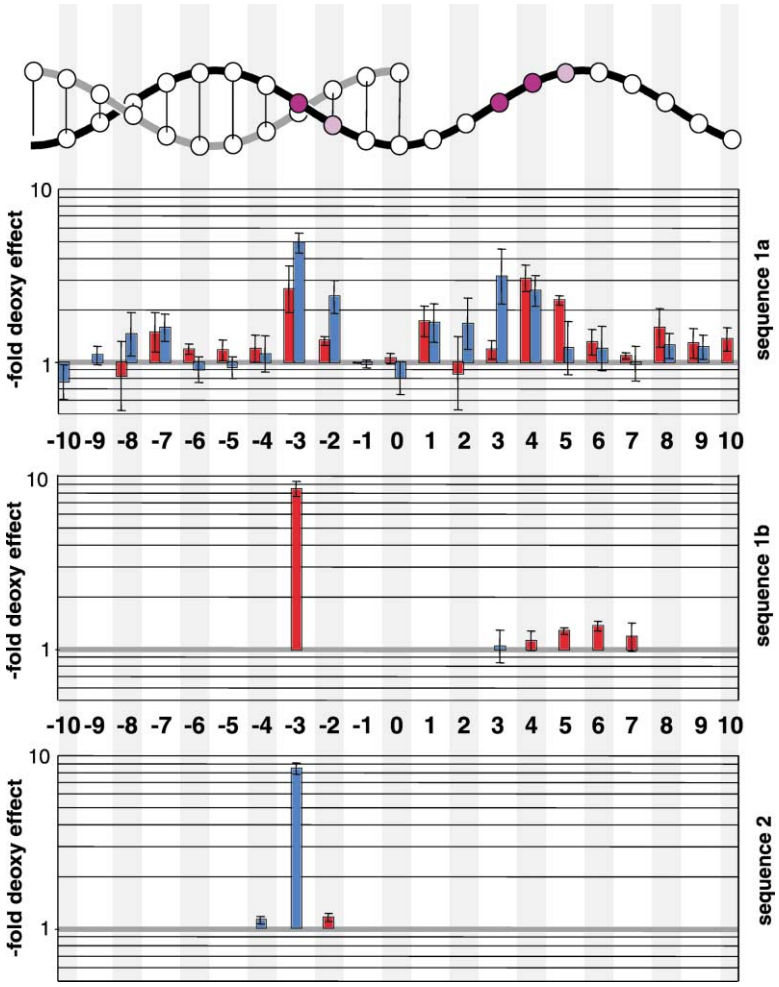

Figure 3. Influences of 2′-Deoxy Substitutions in the Template on Polymerization Kinetics

Significant effects are in the positions 5, 4, 3, −2, and −3. The sequence 1a refers to the template 3′-GACGGUUGGCACGCUUCGCAG-5′, with the underlined nucleotides as templating nucleotides. Sequence 1b is 3′-GACGGUUGGCACGCUUCGCAG-5′ and sequence 2 is 3′-GACGGUUGACACGCUUCGCAG-5′. The reasons for the separate analyses of those template segments are their different templating abilities. In sequence 1a, the polymerization rates are 1 hr^{-1} or more; in sequence 1b, the polymerization rates are 0.1 hr^{-1} or less [28]. The following list corresponds to the italicized numbers in Figure 1A. It provides the templating nucleotides of sequence 1a and is in the format ($x/nt_Y/nt_Z$) where x is the position of the 2′-deoxy substitution, nt_Y describes the nucleotide for the first addition, and nt_Z describes the nucleotide for later additions, counting from the 5′ to the 3′ terminus of the template strand: (−10/-/U_6), (−9/-/U_7), (−8/C_9/C_8), (−7/G_9/G_9), (−6/C_{10}/C_{10}), (−5/A_{11}/U_6), (−4/U_7/U_7), (−3/C_9/C_8), (−2/G_9/G_9), (−1/C_{10}/C_{10}), (0/A_{11}/U_6), (1/U_7/U_7), (2/C_8/C_8), (3/G_9/G_9), (4/C_{10}/C_{10}), (5/A_{11}/U_6), (6/U_7/U_7), (7/C_9/C_8), (8/G_9/G_9), (9/C_{10}/C_{10}), (10/A_{11}/-). For sequences 1b and 2, the templating nucleotides were (−3/G_{13}/-), (3/-/C_{12}), (4/G_{13}/-), (5/G_{14}/-), (6/C_{12}/-), and (7/G_{13}/-). The red columns show effects on the addition of the first nucleotide (nonprocessive); the blue columns show effects on later additions (partially processive). In positions +3 and +5, the kinetics of the first addition differed from the kinetics of the later addition. Note that the difference at position +5 could be due to sequence dependence because two different templating nucleotides are used in this case. Error bars are standard deviations of 3 to 17 experiments. Additional experiments repeated less than three times are consistent with the findings of conserved effects in the double-stranded part of the template and diverse effects in the single strand (data not shown).

tacts to the single-stranded part of the template, all of which are dependent on the template sequence.

The energetic contribution of 2′-hydroxyl groups to the substrate binding energy is −17 kJ/mol (−4.1 kcal/mol). This is based on the following assumption and calculation: one 2′-hydroxyl group (at the primer 3′ terminus) is involved in the chemical step; all others are probably involved in substrate binding. The product of all 2′-deoxy effects that appear to be involved in substrate binding is 1000 (sequence 1a). A 1000-fold shift of the reaction equilibrium corresponds to a $\Delta\Delta G$ of −17 kJ/mol (with $\Delta\Delta G = -RT \ln(K_{D2}/K_{D1})$ and R = the gas constant of 8.31 J/mol*K and T = the reaction temperature, 295 K). In the group I ribozyme, the undocked P1 helix is bound into the docked state via tertiary interactions with the same free energy of −17 kJ/mol [38]. However, the undocked P1 helix is linked to the group I intron, which is why the docking energy of −17 kJ/mol is sufficient for efficient splicing. In contrast, the polymerase ribozyme is not attached to its substrate helix and binds the substrate with a K_D in the range of 3 mM [30]. The energetic comparison is confounded by two factors. First, the measurements of the two ribozymes are made at different salt concentrations and temperatures, 10 mM $MgCl_2$ and 50°C for the group I intron in contrast to 200 mM $MgCl_2$ and 22°C for the polymerase ribozyme [30, 36]. Second, the way in which the single energetic contributions add up to the total binding energy depends on the distance between their binding sites [39], which differs between the two ribozymes [20 and this paper]. Nevertheless, the calculation shows that 2′-hydroxyl groups make an important energetic contribution to substrate binding by the polymerase ribozyme.

Proteinaceous RNA-dependent RNA polymerases (RdRps) also use 2′-hydroxyl groups for substrate recognition. However, most RdRps require less 2′-hydroxyl groups than are important for the polymerase ribozyme. Some RdRps accept DNA as primer strand [29] or use DNA/RNA chimeric template strands [40]. For example, the brome mosaic virus RdRp can use a DNA template with an elongation efficiency comparable to that with an RNA template, if a single ribonucleotide is incorporated at position −11 relative to the initiation site [27]. The brome mosaic virus RdRp and the cucumber mosaic virus RdRp both require a 2′-hydroxyl group at position

−3 on the template strand relative to the template 5′ terminus for the synthesis of full-length RNA [40]. The small number of 2′-hydroxyl groups required by RdRps suggests that substrate 2′-hydroxyl groups are necessary for RdRps to discriminate between RNA and DNA templates, but not for tight substrate binding.

The replacement of a ribose sugar by 2′-deoxyribose in an RNA backbone not only removes a hydrogen bond donating group in that sugar moiety but also perturbs the structure of the RNA backbone [41]. Therefore, the influence of a 2′-deoxy substitution on polymerization kinetics might be mediated by structural perturbations in the substrate. This is not likely to be the case in the double-stranded part of the substrate. Ribonucleotides adopt almost completely RNA conformations in double-stranded RNA/DNA hybrids [42]. Additionally, they force neighboring and pairing deoxynucleotides into conformations that resemble RNA more closely than DNA [42]. Because only single 2′-deoxy substitutions are used, the substrate should not undergo important structural changes in the double-stranded part. In contrast to double strands, the structural effect of a 2′-deoxy substitution is more pronounced in single strands [41]. Such a structure perturbation of the template could reduce its affinity to the ribozyme. This interpretation would fit for the 2′-deoxy effects at the positions +3, +4, and +5 (Figure 3). However, a simpler explanation for these effects is hydrogen bonding of the substituted 2′-hydroxyl groups to the ribozyme, increasing the affinity between ribozyme and substrate.

We observed a prominent double band at one sequence position (Figure 1A, 1st intermediate). This double band was caused by both the properly incorporated G and a misincorporated C, across from a C template residue. The ratio between the incorporation rates of both monomers indicate an error rate of 0.10 in this position (data not shown), much higher than the error rate of 0.0002 that is observed for the incorporation of a G across from a templating C in another sequence context [13]. We do not know why this position is prone to errors in this context and we are puzzled by the insinuated formation of a C:C base pair because this base pair is of low thermodynamic stability and has a different geometry than that of Watson-Crick base pairs [43]. This misincorporation illustrates that the explanations for high fidelity are more complex than a combination of base pairing energy and A-form helix geometry. Some of the mismatched products were elongated, but at a slower rate as judged by the observation that the proportion of the second band in the next intermediate was much lower (Figure 1A). A slower rate of extension of the mismatched products would explain why the double band contained two bands of similar intensity, when the matched base (G) is incorporated nine times more efficiently than the C mismatch. An accumulation of intermediates with misincorporations could also explain why the reacting fraction of several polymerization intermediates decreased over the polymerization timecourse (Figure 1C).

Extension fidelity is not the only context-dependent phenomenon for the polymerase ribozyme. The efficiency of the polymerase ribozyme is also sequence dependent. The rates of individual nucleotide additions vary by more than two orders of magnitude, even for extension by the same NTP within a single template sequence [30]. Furthermore, the contacts to substrate 2′-hydroxyl groups vary between different substrate sequences, especially in the single-stranded region of the substrate (Figure 3). These two observations are correlated. Those substrate intermediates that extend slowly lack the 2′-hydroxyl contacts of the single-stranded region, while those substrate intermediates that extend rapidly have the 2′-hydroxyl contacts at template nucleotides +3, +4, and +5 (Figure 3). This correlation is expected because substrate affinity is limiting for polymerization efficiency under standard conditions [30]. Every factor that stabilizes the ribozyme/substrate complex increases polymerization efficiency. This view puts the finding that the ribozyme is not limited to particular template sequences [13] in a more quantitative light. The polymerization efficiency with a certain template increases with the number and strength of hydrogen bonds that this template established to the ribozyme. Nonetheless, despite these sequence-dependent preferences, the ribozyme is able to utilize to some degree all the primer-template RNAs that have been tried. Perhaps, proteinaceous RNA polymerases also have substantial preferences for different templates, but their preferences are less obvious because those polymerases have much higher efficiencies than does the polymerase ribozyme [30].

Significance

The RNA world hypothesis assumes the previous existence of a self-replicating and evolving RNA system. A polymerase ribozyme would have been essential in such a system. Efforts to generate such a ribozyme have yielded a polymerase with some of the necessary attributes. Its fidelity is close to that required, but its polymerization efficiency, which is limited by a low substrate affinity, would have to undergo major improvements to enable self-replication [30]. As a first step in exploring the molecular basis of the polymerization fidelity and substrate affinity, we studied the importance of 2′-hydroxyl groups for efficient substrate utilization. The 2′-deoxy substitutions at eight substrate positions decrease polymerization efficiency, suggesting that 2′-hydroxyl groups are more important for the ribozyme substrate recognition than for the characterized proteinaceous RNA-dependent RNA polymerases. Seven of these 2′-deoxy effects are of a magnitude consistent with the loss of a hydrogen-bond contact to the ribozyme. All are within five nucleotides of the 3′ terminus of the primer, providing the first clues about the ribozyme-substrate interface and a starting point for investigating the ribozyme side of the interface through experiments such as nucleotide analog interference suppression [20]. Knowledge of the ribozyme-substrate interface will also guide efforts to improve substrate affinity using both combinatorial and design approaches.

Experimental Procedures

Ribozyme and Substrates

The ribozyme, known as round 18 polymerase ribozyme, was prepared by runoff transcription from a PCR template as described

[13]. All oligoribonucleotides but the polymerase ribozyme were purchased from Dharmacon. All oligoribonucleotides in the ribozyme reactions were gel purified. Primers were radiolabeled using T4 polynucleotide kinase and $\gamma[^{32}P]$ATP. All chemicals were Molecular Biology grade or higher (99%+).

Polymerization Reactions

All reactions were in final volumes of 35 μl and included final concentrations of 2 μM ribozyme, 2.5 μM stem oligonucleotide (GGCACCA), 0.5 μM unlabeled primer, 1 μM template, and less than 50 nM (approximately 0.3 μCi) 5′-labeled primer. The RNAs were mixed in H_2O, heat-denatured for 2 min at 80°C, and cooled to room temperature. The reactions were started by addition of Mg^{2+} and NTPs final concentrations of 200 mM $MgCl_2$, 50 mM Tris/HCl (pH 8.6), and 4 mM of each sodium-NTP (pH 7.5). During incubation at 22°C, samples were taken at appropriate time points and diluted with 1.5 volumes of stop solution containing 4.5 M urea, 200 mM sodium-EDTA (pH 8.6), and an oligonucleotide that is complementary and in 10-fold excess to the template [13]. The diluted samples were heat-denatured and separated on denaturing 20% polyacrylamide gels.

Data Analysis

The radioactive signals were recorded using a phosphorimager (Fuji BAS 2000), quantified using ImageGauge version 3.3, and processed using Excel 98 (see below). The reactivities, expressed as fraction of polymerization intermediates that reacted per second, were plotted as function of time for each polymerization intermediate (Figure 1C). The values were averaged because the reacting fraction changed over time for several intermediates. The averages of the reacting fractions were derived after mono-exponential least squares fitting to the reacting fraction as a function of reaction time. The data points that were used for fitting were from a time range in which the signals could be quantified reliably (indicated by the length of the fitted lines in Figure 1C). To account for the uneven distribution of the time points and for changes in the reacting fraction over time, the midpoint of the fitted line was used as the mean reactivity that served to compare different substrates. All statistical analysis used geometric means.

Acknowledgments

We thank Wendy Johnston for conducting pilot experiments and Mike Lawrence for helpful discussions, and Scott Baskerville, Mike Lawrence, and Erik Schultes for critical reading of the manuscript. This work was supported by the DFG grant MU-1708 1/1 and the NIH grant R01 GM61835.

Received: May 22, 2003
Revised: June 27, 2003
Accepted: June 30, 2003
Published: September 19, 2003

References

1. Woese, C.R. (1967). The fundamental nature of the genetic code: prebiotic interactions between polynucleotides and polyamino acids or their derivatives. Proc. Natl. Acad. Sci. USA *59*, 110–117.
2. Crick, F.H.C. (1968). The origin of the genetic code. J. Mol. Biol. *38*, 367–379.
3. Orgel, L.E. (1968). Evolution of the genetic apparatus. J. Mol. Biol. *38*, 381–393.
4. Joyce, G.F., and Orgel, L.E. (1999). Prospects for understanding the origin of the RNA world. In The RNA World, Second Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 49–77.
5. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell *31*, 147–152.
6. DeRose, V.J. (2002). Two decades of RNA catalysis. Chem. Biol. *9*, 961–969.
7. Bartel, P.D., and Szostak, J.W. (1993). Isolation of new ribozymes from a large pool of random sequences. Science *261*, 1411–1418.
8. Jäschke, A. (2001). Artificial ribozymes and deoxyribozymes. Curr. Opin. Struct. Biol. *11*, 321–326.
9. Ekland, E.H., and Bartel, D.P. (1996). RNA-catalysed RNA polymerization using nucleoside triphosphates. Nature *382*, 373–376.
10. Bartel, D.P.D. (1999). Re-creating an RNA replicase. In The RNA World, Second Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 143–162.
11. Farkas, I., Jeong, H., Vicsek, T., Barabasi, A.-L., and Oltvai, Z.N. (2003). The topology of the transcription regulatory network in the yeast, *Saccharomyces cerevisiae*. Physica A *318*, 601–612.
12. Szostak, J.W., Bartel, P.D., and Luisi, P.L. (2001). Synthesizing life. Nature *409*, 387–390.
13. Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E., and Bartel, D.P. (2001). RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science *292*, 1319–1325.
14. Kirseboom, L.A., and Svard, S.G. (1992). The kinetics and specificity of cleavage by RNase P is mainly dependent on the structure of the amino acid acceptor stem. Nucleic Acids Res. *20*, 425–432.
15. Frank, D.N., and Pace, N.R. (1998). Ribonuclease P: unity and diversity in a tRNA processing ribozyme. Annu. Rev. Biochem. *67*, 153–180.
16. Bevilacqua, P.C., and Turner, D.H. (1991). Comparison of binding of mixed ribose-deoxyribose analogues of CUCU to a ribozyme and to GGAGAA by equilibrium dialysis: evidence for ribozyme specific interactions with 2′ OH groups. Biochemistry *30*, 10632–10640.
17. Pyle, A.M., and Cech, T.R. (1991). Ribozyme recognition of RNA by tertiary interactions with specific ribose 2′-OH groups. Nature *350*, 628–631.
18. Strobel, S.A., and Cech, T.R. (1993). Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the Tetrahymena ribozyme. Biochemistry *32*, 13593–13604.
19. Doudna, J.A., Cormack, B.P., and Szostak, J.W. (1989). RNA structure, not sequence, determines the 5′ splice-site specificity of a group I intron. Proc. Natl. Acad. Sci. USA *86*, 7402–7406.
20. Szewczak, A.A., Ortoleva-Donnelly, L., Ryder, S.P., Moncoeur, E., and Strobel, S.A. (1998). A minor groove RNA triple helix within the catalytic core of a group I intron. Nat. Struc. Biol. *5*, 1037–1042.
21. Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. Science *273*, 1678–1685.
22. Tamura, M., and Holbrook, S.R. (2002). Sequence and structural conservation in RNA ribose zippers. J. Mol. Biol. *320*, 455–474.
23. 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. Proc. Natl. Acad. Sci. USA *98*, 4899–4903.
24. St. Johnston, D., Brown, N.H., Gall, J.G., and Jantsch, M. (1992). A conserved double-stranded RNA-binding domain. Proc. Natl. Acad. Sci. USA *89*, 10979–10983.
25. Fierro-Monti, I., and Mathews, M.B. (2000). Proteins binding to duplexed RNA: one motif, multiple functions. Trends Biochem. Sci. *5*, 241–246.
26. Bevilacqua, P.C., and Cech, T.R. (1996). Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. Biochemistry *35*, 9983–9994.
27. Siegel, R.W., Bellon, L., Beigelman, L., and Kao, C.C. (1999). Use of DNA, RNA, and chimeric templates by a viral RNA-dependent RNA polymerase: Evolutionary implications for the transition from the RNA to the DNA world. J. Virol. *73*, 6424–6429.
28. Kim, M.-J., Zhong, W., Hong, Z., and Kao, C.C. (2000). Template nucleotide moieties required for *de novo* initiation of RNA synthesis by a recombinant viral RNA-dependent RNA polymerase. J. Virol. *74*, 10312–10322.
29. Arnold, J.J., Ghosh, S.K.B., and Cameron, C.E. (1999). Poliovirus

RNA-dependent RNA polymerase ($3D^{pol}$). J. Biol. Chem. *274*, 37060–37069.

30. Lawrence, M.S., and Bartel, D.P. (2003). Processivity of a polymerase ribozyme. Biochemistry *42*, in press.
31. Izatt, R.M., Rytting, J.H., Hansen, L.D., and Christensen, J.J. (1966). Thermodynamics of proton dissociation in dilute aqueous solution. V. An entropy titration study of adenosine, pentoses, hexoses, and related compounds. J. Am. Chem. Soc. *88*, 2641–2645.
32. Stewart, R. (1985). The proton: applications to organic chemistry. In Organic Chemistry, H. H. Wasserman, ed. (London: Academic press Inc.), pp. 42–44.
33. Shan, S.-O., Yoshida, A., Sun, S., Piccirilli, J.A., and Herschlag, D. (1999). Three metal ions at the active site of the *Tetrahymena* group I ribozyme. Proc. Natl. Acad. Sci. USA *96*, 12299–12304.
34. Silverman, S.K., and Cech, T.R. (1999). Energetics and cooperativity of tertiary hydrogen bonds in RNA structure. Biochemistry *38*, 8691–8702.
35. Williams, D.J., Boots, J.L., and Hall, K.B. (2001). Thermodynamics of 2′-ribose substitutions in UUCG tetraloops. RNA *7*, 44–53.
36. Herschlag, D., Eckstein, F., and Cech, T.R. (1993). Contributions of 2′-hydroxyl groups of the RNA substrate to binding and catalysis by the Tetrahymena ribozyme. An energetic picture of an active site composed of RNA. Biochemistry *32*, 8299–8311.
37. Sood, V.D., Yekta, S., and Collins, R.A. (2002). The contribution of 2′-hydroxyls to the cleavage activity of the Neurospora VS ribozyme. Nucleic Acids Res. *30*, 1132–1138.
38. Narlikar, G.J., Khosla, M., Usman, N., and Herschlag, D. (1997). Quantitating binding energies of 2′ OH groups on the P1 duplex of the Tetrahymena ribozyme: Intrinsic binding energy in an RNA enzyme. Biochemistry *36*, 2465–2477.
39. Shamoo, Y., Abdul-Manan, N., and Williams, K.R. (1995). Multiple RNA binding domains (RBDs) just don't add up. Nucleic Acids Res. *23*, 725–728.
40. Tayon, R., Jr., Kim, M.J., and Kao, C.C. (2001). Completion of RNA synthesis by viral RNA replicases. Nucleic Acids Res. *17*, 3576–3582.
41. Lindqvist, M., Sarkar, M., Winqvist, A., Rozners, E., Strömberg, R., and Gräslund, A. (2000). Optical spectroscopic study of the effects of a single deoxyribose substitution in a ribose backbone: Implications in RNA-RNA interaction. Biochemistry *39*, 1693–1701.
42. Mellema, J.R., Haasnoot, C.A., van der Marel, G.A., van Boeckel, C.A., van Boom, J.H., and Altona, C. (1983). Proton NMR studies on the covalently linked RNA-DNA hybrid r(GCG)d(TATACAC). Assignment of proton resonances by application of the nuclear Overhauser effect. Nucleic Acids Res. *11*, 5717–5738.
43. Saenger, W. (1984). Principles of Nucleic Acid Structure (New York: Springer-Verlag), pp. 120 and 131.