

## CHAPTER 1

# *Ribozymes and RNA Catalysis: Introduction and Primer*

DAVID M.J. LILLEY<sup>a</sup> AND FRITZ ECKSTEIN<sup>b</sup>

<sup>a</sup> Cancer Research UK Nucleic Acid Structure Research Group, MSI/WTB Complex, The University of Dundee, Dow Street, Dundee DD1 5EH, UK;

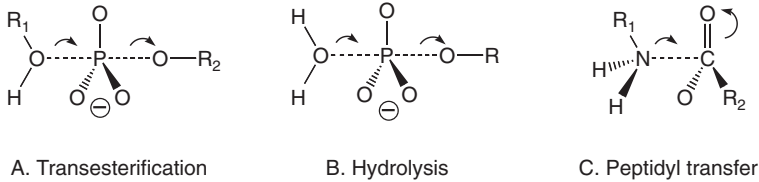
<sup>b</sup> Max-Planck-Institut Für Experimentelle Medizin, Hermann-Rein-Str. 3, Göttingen D-37075, Germany

## 1.1 What are Ribozymes?

Ribozymes are RNA molecules that act as chemical catalysts, a shortening of *ribonucleic acid enzymes*. In the contemporary biosphere, the known ribozymes carry out a relatively limited range of reactions (Figure 1.1), mostly involving phosphoryl transfer, notably transesterification (the large majority) and hydrolysis reactions. However, the discovery that peptidyl transferase is catalysed by the rRNA component of the large ribosomal subunit significantly extends the range of chemistry to include the condensation of an amine with an  $sp^2$ -hybridized carbonyl centre. A significantly greater range of chemical reactions may be catalysed by RNA species selected for the purpose, so that ribozymes catalyzing a broader set of reactions may have existed in the past.

## 1.2 What is the Role of Ribozymes in Cells?

Contemporary ribozymes (Table 1.1) are used for various biological purposes. The nucleolytic ribozymes bring about the site-specific cleavage (or the reverse ligation process) of RNA by attack of a 2'-hydroxyl group on the adjacent 3'-phosphorus (Figure 1.2A) (Chapters 2–8). This reaction is exploited for the processing of replication intermediates, and in the control of gene expression by metabolite-induced cleavage of mRNA. Ribonuclease P carries out the processing of tRNA in all kingdoms of life, using a hydrolytic reaction (Chapter 9). Several introns are spliced out autocatalytically by ribozyme action, initiated either by the attack of a 2'-hydroxyl group located remotely within the intron



**Figure 1.1** Reactions catalysed by the known natural ribozymes observed in biology. The biggest number carry out transesterification reaction (A), notably the nucleolytic ribozymes and the self-splicing introns. RNaseP carries out a hydrolytic reaction to process tRNA molecules (B). In the peptidyl transfer reaction (C) the amine of the A-site aminoacyl-tRNA attacks the carbonyl group of the peptidyl-tRNA (or initiator aminoacyl t-RNA) held in the P site of the large ribosomal subunit.

**Table 1.1** Classes of natural ribozymes.

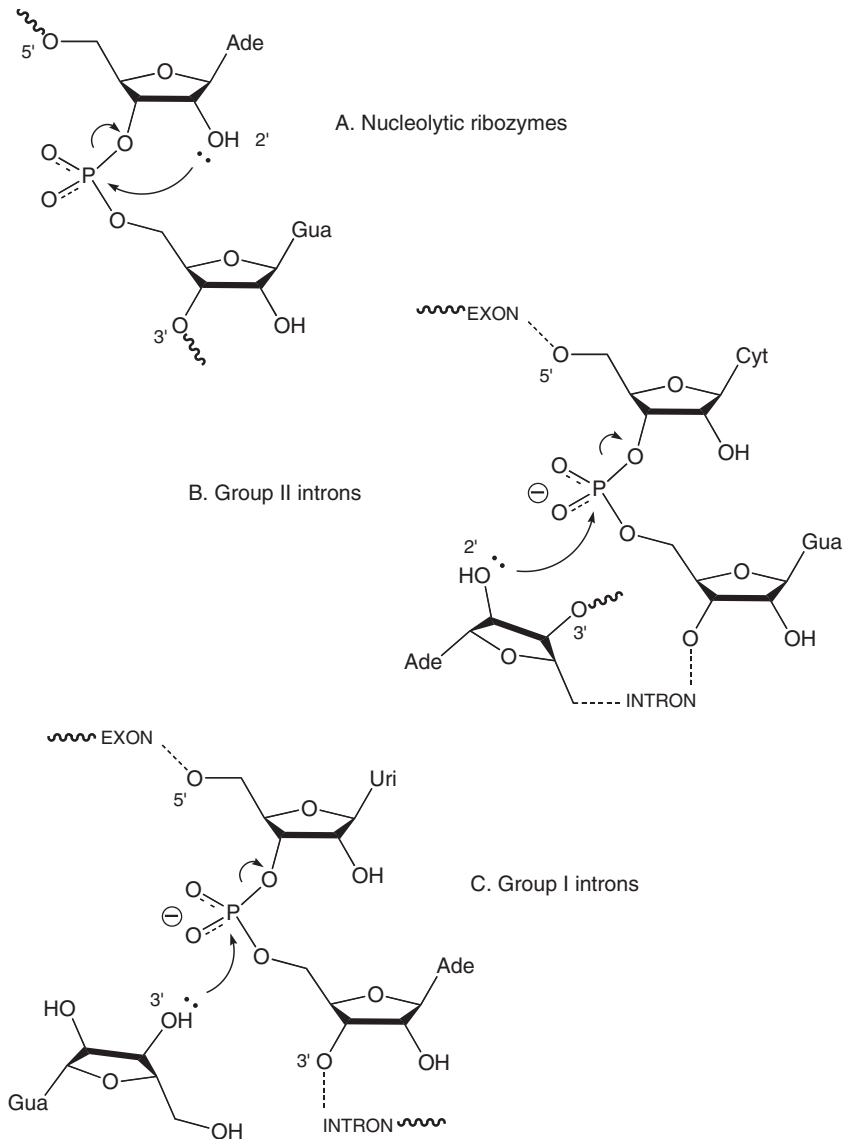
<i>Ribozyme</i>	<i>Mechanism</i>	<i>Reference</i>
Hammerhead	Transesterification 2'-O	Chapters 3, 4 <sup>a</sup>
Hairpin	Transesterification 2'-O	Chapter 5
Hepatitis Delta Virus	Transesterification 2'-O	Chapters 6, 7 <sup>b</sup>
Varkud Satellite	Transesterification 2'-O	Chapter 5
<i>glmS</i>	Transesterification 2'-O	Chapter 8 <sup>c</sup>
Group I Intron	Nucleotidyl transfer	Chapter 10
Group I-like	Nucleotidyl transfer	Chapter 12
Group II Intron	Nucleotidyl transfer	Chapter 11
RNase P	Hydrolysis	Chapter 9
Ribosome	Peptidyl transfer	Chapter 14

<sup>a</sup> This is a large group of ribozymes. Although many are found in plant pathogens, hammerhead sequences have been found in amphibia (Newt) and plants (*Arabidopsis thaliana*).

<sup>b</sup> A closely related sequence has been identified in the genomes of mammals, in gene CPEB3.

<sup>c</sup> The *glmS* ribozyme is the only nucleolytic ribozyme that uses a small molecule cofactor in its catalytic chemistry.

(group II introns, Figure 1.2B) (Chapter 11), or by an exogenous guanosine molecule (group I introns, Fig 1.2C) (Chapter 10). While “smoking gun” evidence has not yet been found, the similarity of the chemistry of mRNA splicing in the spliceosome to that of the group II introns makes it very likely that this too is RNA catalysed, with the snU4/U6 RNA as the ribozyme (Chapter 13). Lastly, the peptidyl transferase activity of the ribosome catalyses what is arguably the most important reaction of the cell, the condensation of amino acids into polypeptides (Chapter 14). Ribozymes are widespread in nature, from bacteria and their phages, archaea, yeasts and fungi and higher eukaryotes. They are also present in clinically-significant human pathogens such as the hepatitis D virus (Chapter 6). New ribozymes are still being found, both by biochemical approaches and by the bioinformatic analysis of genome sequencing data.



**Figure 1.2** Three types of transesterification reactions catalysed by natural ribozymes. The cleavage reaction of the nucleolytic ribozymes (*A*) involves attack of the 2'-O on the adjacent 3'-P, with a 5'-oxyanion leaving group. In the reverse ligation reaction the 5'-O attacks the P of the cyclic phosphate. A similar reaction occurs in the group II intron (*B*), except that the 2'-O nucleophile is provided by a nucleotide located elsewhere in the intron. The nucleophile of the first reaction of the group I intron (*C*) is the 3'-O of an exogenous guanosine.

### 1.3 Ribozymes Bring about Significant Rate Enhancements

Protein enzymes can achieve some extraordinary catalytic rate enhancements. Values of almost  $10^{18}$ -fold are possible, although many generate much smaller accelerations. RNA catalysts tend to produce more modest rate enhancements. For example, the nucleolytic ribozymes typically accelerate their transesterification reactions by around a million-fold relative to the uncatalysed reaction in a dinucleotide, with rates of around  $1 \text{ min}^{-1}$ . For those ribozymes this may be as fast as it needs to be, since a given site needs to be cut just once. However, while this rate was previously discussed as some kind of speed limit,<sup>1</sup> it appears that this is not an intrinsic limitation, and some redesign of some ribozymes has resulted in very respectable catalytic rates  $\geq 10 \text{ s}^{-1}$ .<sup>2,3</sup>

### 1.4 Why Study Ribozymes?

There are several reasons for studying ribozymes. First, they are active in contemporary living cells, carrying out reactions that are critical for cell viability in some cases; they are therefore legitimate subjects of interest in the complete description of cellular metabolism.

Second, they may have had a key role in the evolution of life on the planet.<sup>4,5</sup> There is clearly a rather severe “chicken and egg” problem involved in the origins of proteins and translation systems, both of which seem to require the prior existence of the other. Yet in principle a biosphere in which RNA was simultaneously the informational and catalytic macromolecule provides a temporary solution to that problem. Such an RNA world might have existed around 3.5 billion years ago, yet would have been relatively short lived in geological terms, being swiftly replaced by polypeptides that it would have produced. Some of the ribozymes that currently exist, most notably the ribosome perhaps, may be molecular fossils from that time, and therefore their study may offer a partial glimpse of that early metabolism. Although contemporary ribozymes carry out a very limited range of chemistries, selected ribozymes provide an indication of what is achievable by RNA catalysts,<sup>6–12</sup> and potentially offer a kind of proof-of-principle of an RNA world.

A third reason for studying ribozymes is that they are rather basic biocatalysts, providing a simplified and contrasting perspective on macromolecular catalytic mechanisms compared with enzymes. The last few years have seen significant advances in our understanding of the chemical origins of ribozyme catalysis, and this may cast light on protein-based catalysis in turn.

Lastly, there has been some effort to exploit the potential specificity of ribozymes as therapeutic agents. In principle, the great selectivity of ribozyme-induced cleavage of a chosen sequence could provide an opportunity to interfere with gene expression if targeted to a specific mRNA; this should ideally be the basis for their development into therapeutic drugs. However, this

requires that many more problems be overcome, including stability in serum, delivery to the required location of the chosen cell and correct folding into the active conformation in competition with the native folding of the target RNA *in vivo*. So far only two ribozymes have found their way into clinical trials. One is a chemically synthesized and modified hammerhead ribozyme targeting the vascular endothelial growth factor receptor-1 (VEGFR) mRNA.<sup>13</sup> In preclinical trials it has exhibited antitumor and antimetastatic activity by interfering with VEGF-dependent angiogenesis. Angiogenesis inhibition is important in patients with refractory solid tumours. The other example involved the use of a hammerhead ribozyme as part of a vector to combat HIV-1.<sup>14</sup> The ribozyme directs the cleavage of the transcript of the chemokine receptor CCR5 that is essential for HIV-1 infection. To optimize efficiency the vector contains in addition a TAR decoy and a short hairpin RNA targeting the *rev* and *tat* mRNA of HIV-1. Potent inhibition of HIV-1 replication was achieved with this construct in a human T cell line.

## 1.5 Folding RNA into the Active Conformation

Just as protein enzymes must be correctly folded into the conformation required for catalytic activity, so must RNA. Moreover, it is clear that the folding processes of ribozymes is intimately associated with their function in many cases. Marked differences between the chemical nature of RNA and proteins results in very different folding processes. In general the precise nature of Watson–Crick basepairing leads to the relatively easy formation of secondary structure, although a requirement to “un-do” unfavourable pairings can provide significant barriers to correct folding. But most of the attention in RNA folding is focussed on the formation of the tertiary structure. The polyelectrolyte character of RNA results in a strong electrostatic contribution to this process, and thus a dependence on the presence of metal ions. The resulting folded RNA structure can bind metal ions, either site-specifically or diffusely, and these bound ions can play a direct role in catalysis. Site-bound ions are inner-sphere complexes where one of more water molecules in the first coordination sphere have exchanged with ligands provided by the RNA; such ions exchange slowly with bulk solvent. Diffusely bound ions do not undergo ligand substitution, and consequently exchange with solvent much more rapidly. They can nevertheless exhibit high occupancy in sites of strong electrostatic potential.

The smaller ribozymes, notably the nucleolytic ribozymes, exhibit some common structural themes, and their architectures are based around either helical junctions (hammerhead, hairpin and VS) or pseudoknots (HDV, *glmS*); evidently these motifs are efficient ways to construct small, autonomously folding species. Furthermore, some of these ribozymes contain additional elements that are not strictly essential for catalytic activity, yet result in marked enhancement of folding, such as the interacting loops of the hammerhead<sup>2,15,16</sup> and the four-way junction of the hairpin ribozyme.<sup>17</sup>

Most studies of RNA folding *in vitro* have therefore focussed on ion-induced folding. The small nucleolytic ribozymes generally exhibit relatively simple folding, typically two or three-state processes. However, larger ribozymes like the group I introns undergo complex folding pathways, beset with kinetic traps (Chapter 15). In the cell, proteins may assist the folding processes.

## 1.6 The Catalytic Resources of RNA – Making a Lot of a Little

The chemical nature of proteins has evolved to provide a highly adaptable catalytic framework with a broad repertoire of functional groups. It is based on an electrically neutral backbone, with sidechains that introduce a wide variety of chemistries, including carboxylic acids, amines, hydroxyl and thiol groups as well as hydrophobic side chains that may be either aliphatic or aromatic. By contrast, RNA consists of just four nucleotide bases of rather similar chemical nature, connected by an electrically-charged ribose-phosphate backbone.

So what resources are available to RNA that can be used to build a catalyst? Firstly, there are the nucleobases. These have hydrogen bond donors and acceptors that can be used to bind the substrate, and potentially to stabilize a transition state. In principle they could also act as general acids and bases. However, first they must overcome the problem of their  $pK_a$  values, which are unsuitable for general acid–base catalysis at neutral pH. Adenine N1 and cytosine N3 have low  $pK_a$  values, while those of guanine N1 and uracil N3 are relatively high. For example, a cytosine with a  $pK_a$  of 4 is a relatively strong acid, but only one molecule in 1000 is protonated at neutral pH. Thus most ribozymes will be in the wrong form to carry out a protonation. The great majority of molecules are in the deprotonated form and able to act as a general base, but the conjugate base of a strong acid is weak, so it is rather unreactive. However, the situation can be improved because nearby anionic phosphate groups may raise the  $pK_a$  significantly, and values of 5.5–6.5 are quite possible,<sup>18,19</sup> making the nucleobase more available as an acid. Similarly, the  $pK_a$  of guanine might be reduced if it is located close to a bound metal ion, thereby making it basic at a lower pH.

The second potential players are metal ions, and their associated water molecules. The folding of a ribozyme may create specific ion binding sites, or pockets in which there is high occupancy of more weakly bound ions. Metal ions can act as Lewis acids, or provide electrostatic stabilization of negative charge such as a dianionic phosphorane transition state. Water molecules contained within the inner sphere of coordination may participate in general acid–base catalysis, as exemplified by the HDV ribozyme.<sup>20,21</sup>

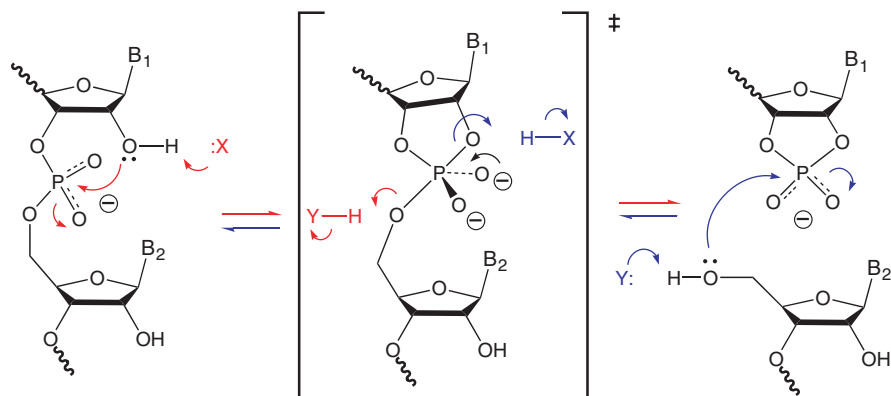
In addition to chemical participants, RNA can also potentially exploit its structure to contribute to catalysis. Substrate binding can result in acceleration of reaction velocity due to proximity and orientation, together with structural stabilization of the transition state.

## 1.7 Mechanisms and Catalytic Strategies of Ribozymes

Given the relative paucity of potential catalytic groups present in RNA molecules, ribozymes achieve some impressive rate accelerations. How they achieve this is a major topic of this volume, but we consider this briefly here. We will take the nucleolytic ribozymes as an example – most ribozymes carry out phosphoryl transfer reactions of various kinds, so that similar considerations will apply.

The chemical mechanism is shown in Figure 1.3. The cleavage reaction involves a nucleophilic attack of the 2'-oxygen on the adjacent 3'-phosphorus, with departure of the 5'-oxygen to create a cyclic 2',3'-phosphate product. The chirality of the phosphorus becomes inverted during the reaction,<sup>22–24</sup> indicating concerted bond formation and breaking to some degree, and passage through a phosphorane transition state (or possibly high energy intermediate). This requires an in-line attack by the nucleophile, and thus a degree of rate enhancement can arise from prealigning the reactants into this geometry. The phosphorane transition state might be stabilized relative to the ground state by specific hydrogen bonding, or electrostatically. The latter might be achieved by juxtaposition of a metal ion, or perhaps a protonated nucleobase.

A hydroxyl group is a relatively weak nucleophile. Removal of its proton by a base would create a much more reactive alkoxide ion.<sup>25</sup> The reaction would



**Figure 1.3** Mechanism of the nucleolytic ribozymes, together with a suggested role of general acid–base catalysis. Cleavage is rightward (red arrows) while ligation is leftward (blue arrows) in this scheme. The trigonal bipyramidal phosphorane is indicated as the transition state, although it could be a high energy intermediate flanked by a less symmetrical transition state. The cleavage reaction could be catalysed by a general base (X) to remove the proton from the 2'-O nucleophile and a general acid (YH) to protonate the 5'-oxyanion leaving group (both indicated red). By the principle of microscopic reversibility, the same groups will play converse roles in the ligation reaction, i.e. XH is the general acid that protonates the 2'-oxyanion leaving group, and Y is the base removing the proton from the 5'-O nucleophile (both shown red).

also be assisted by protonation of the 5'-oxyanion leaving group. Thus it would be expected that the reaction would be subject to general acid–base catalysis, and considerable evidence has been collected that this is generally the case in the nucleolytic ribozymes (Chapter 2). To date the active participants have included the nucleobases adenine, cytosine and guanine, and hydrated metal ions. Note that, the groups that act as acid and base in the cleavage reaction will reverse roles in the ligation reaction by the principle of microscopic reversibility.

## 1.8 Impact of New Methodologies to Study Ribozymes

While a lot of mechanistic insight into ribozyme action has come from the application of more or less conventional enzymological approaches, structural and biophysical methods have played key roles. Atomic resolution X-ray crystallographic structures have been determined for all the nucleolytic ribozymes except the VS ribozyme, and multiple forms have been determined in general. Crystal structures have also been solved for some of the larger ribozymes, including several examples of the group I ribozyme<sup>26–28</sup> (Chapter 10), RNaseP<sup>29–31</sup> (Chapter 9) and of course the peptidyl transferase centre within the 50S ribosomal subunit<sup>32</sup> (Chapter 14). All these studies have provided a wealth of structural data that then feeds back into mechanistic studies in an iterative process.

Single-molecule methods have also had a significant impact in the study of ribozyme folding<sup>17,33–35</sup> and activity.<sup>36</sup> These can provide a different perspective upon kinetic processes, free of the averaging that occurs with the ensemble and opening up the study of processes that cannot be synchronized. Both fluorescence and force spectroscopy have been applied to ribozymes. Other biophysical methods are also providing valuable information on RNA folding processes, such as small-angle X-ray scattering<sup>37–39</sup> and the combination of chemical footprinting and rapid reaction techniques.<sup>40,41</sup>

## 1.9 Finally

The field of RNA catalysis provides great challenges, and tremendous excitement. It has seen enormous development over the last 20 or so years, and it continues to spring surprises on a regular basis. This chapter provides an introduction for the reader who might not be directly involved in ribozyme research, Much more detail is provided in the following chapters. So please read on.

## References

1. R.R. Breaker, G.M. Emilsson, D. Lazarev, S. Nakamura, I.J. Puskarz, A. Roth and N. Sudarsan, *RNA*, 2003, **9**, 949–957.

2. M.D. Canny, F.M. Jucker, E. Kellogg, A. Khvorova, S.D. Jayasena and A. Pardi, *J. Am. Chem. Soc.*, 2004, **126**, 10848–10849.
3. R. Zamel, A. Poon, D. Jaikaran, A. Andersen, J. Olive, D. De Abreu and R.A. Collins, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 1467–1472.
4. C. Woese, in *The Genetic Code*, Harper & Row, New York, 1967, pp. 179–195.
5. F.H.C. Crick, *J. Mol. Biol.*, 1968, **38**, 367–379.
6. E.H. Eklund, J.W. Szostak and D.P. Bartel, *Science*, 1995, **269**, 364–370.
7. T.M. Tarasow, S.L. Tarasow and B.E. Eaton, *Nature*, 1997, **389**, 54–57.
8. H. Suga, P.A. Lohse and J.W. Szostak, *J. Am. Chem. Soc.*, 1998, **120**, 1151–1156.
9. P.J. Unrau and D.P. Bartel, *Nature*, 1998, **395**, 260–263.
10. W.K. Johnston, P.J. Unrau, M.S. Lawrence, M.E. Glasner and D.P. Bartel, *Science*, 2001, **292**, 1319–1325.
11. G. Sengle, R.A. Eisenfuh, P.S. Arora, J.S. Nowick and M. Famulok, *Chem. Biol.*, 2001, **8**, 459–473.
12. G.F. Joyce, *Nature*, 2002, **418**, 214–221.
13. D.E. Weng, P.A. Masci, S.F. Radka, T.E. Jackson, P.A. Weiss, R. Ganapathi, P.J. Elson, W.B. Capra, V.P. Parker, J.A. Lockridge, J.W. Cowens, N. Usman and E.C. Borden, *Mol. Cancer Therap.*, 2005, **4**, 948–955.
14. M.J. Li, J. Kim, S. Li, J. Zaia, J.K. Yee, J. Anderson, R. Akkina and J. Rossi, *J. Mol. Therap.*, 2005, **12**, 900–909.
15. A. Khvorova, A. Lescoute, E. Westhof and S.D. Jayasena, *Nat. Struct. Biol.*, 2003, **10**, 1–5.
16. J.C. Penedo, T.J. Wilson, S.D. Jayasena, A. Khvorova and D.M.J. Lilley, *RNA*, 2004, **10**, 880–888.
17. E. Tan, T.J. Wilson, M.K. Nahas, R.M. Clegg, D.M.J. Lilley and T. Ha, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 9308–9313.
18. P. Legault and A. Pardi, *J. Am. Chem. Soc.*, 1997, **119**, 6621–6628.
19. S. Ravindranathan, S.E. Butcher and J. Feigon, *Biochemistry*, 2000, **39**, 16026–16032.
20. S. Nakano, D.M. Chadalavada and P.C. Bevilacqua, *Science*, 2000, **287**, 1493–1497.
21. A. Ke, K. Zhou, F. Ding, J.H. Cate and J.A. Doudna, *Nature*, 2004, **429**, 201–205.
22. H. van Tol, J.M. Buzayan, P.A. Feldstein, F. Eckstein and G. Bruening, *Nucleic Acids Res.*, 1990, **18**, 1971–1975.
23. M. Koizumi and E. Ohtsuka, *Biochemistry*, 1991, **30**, 5145–5150.
24. G. Slim and M.J. Gait, *Nucleic Acids Res.*, 1991, **19**, 1183–1188.
25. M. Oivanen, S. Kuusela and H. Lonnberg, *Chem. Rev.*, 1998, **98**, 961–990.
26. P.L. Adams, M.R. Stahley, J. Wang and S.A. Strobel, *Nature*, 2004, **430**, 45–50.
27. F. Guo, A.R. Gooding and T.R. Cech, *Mol. Cell*, 2004, **16**, 351–362.
28. B.L. Golden, H.D. Kim and E. Chase, *Nat. Struct. Mol. Biol.*, 2005, **12**, 82–89.
29. A.S. Krasilnikov, X. Yang, T. Pan and A. Mondragon, *Nature*, 2003, **421**, 760–764.

30. A.S. Krasilnikov, Y. Xiao, T. Pan and A. Mondragon, *Science*, 2004, **306**, 104–107.
31. A.V. Kazantsev, A.A. Krivenko, D.J. Harrington, S.R. Holbrook, P.D. Adams and N.R. Pace, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 13392–13397.
32. P. Nissen, J. Hansen, N. Ban, P.B. Moore and T.A. Steitz, *Science*, 2000, **289**, 920–930.
33. X. Zhuang, L.E. Bartley, H.P. Babcock, R. Russel, T. Ha, D. Herschlag and S. Chu, *Science*, 2000, **288**, 2048–2051.
34. X.W. Zhuang, H.D. Kim, M.J.B. Pereira, H.P. Babcock, N.G. Walter and S. Chu, *Science*, 2002, **296**, 1473–1476.
35. G. Bokinsky, D. Rueda, V.K. Misra, M.M. Rhodes, A. Gordus, H.P. Babcock, N.G. Walter and X. Zhuang, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 9302–9307.
36. M.K. Nahas, T.J. Wilson, S. Hohng, K. Jarvie, D.M.J. Lilley and T. Ha, *Nat. Struct. Mol. Biol.*, 2004, **11**, 1107–1113.
37. R. Russell, I.S. Millett, S. Doniach and D. Herschlag, *Nat. Struct. Biol.*, 2000, **7**, 367–370.
38. R. Das, L.W. Kwok, I.S. Millett, Y. Bai, T.T. Mills, J. Jacob, G.S. Maskel, S. Seifert, S.G. Mochrie, P. Thiyagarajan, S. Doniach, L. Pollack and D. Herschlag, *J. Mol. Biol.*, 2003, **332**, 311–319.
39. G. Caliskan, C. Hyeon, U. Perez-Salas, R.M. Briber, S.A. Woodson and D. Thirumalai, *Phys. Rev. Lett.*, 2005, **95**, 268–303.
40. M.L. Deras, M. Brenowitz, C.Y. Ralston, M.R. Chance and S.A. Woodson, *Biochemistry*, 2000, **39**, 10975–10985.
41. T. Adilakshmi, P. Ramaswamy and S.A. Woodson, *J. Mol. Biol.*, 2005, **351**, 508–519.