ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Methods 28 (2002) 307–322

METHODS

www.academicpress.com

Analysis of substrate recognition by the ribonucleoprotein endonuclease RNase P

Eric L. Christian, Nathan H. Zahler, Nicholas M. Kaye, and Michael E. Harris*

Center for RNA Molecular Biology, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

Accepted 30 July 2002

Abstract

Ribonuclease P (RNase P), is a ribonucleoprotein complex that catalyzes the site-specific cleavage of pre-tRNA and a wide variety of other substrates. Although RNase P RNA is the catalytic subunit of the holoenzyme, the protein subunit plays a critical role in substrate binding. Thus, RNase P is an excellent model system for studying ribonucleoprotein function. In this review we describe methods applied to the in vitro study of substrate recognition by bacterial RNase P, covering general considerations of reaction conditions, quantitative measurement of substrate binding equilibria, enzymatic and chemical protection, cross-linking, modification interference, and analysis of site-specific substitutions. We describe application of these methods to substrate binding by RNase P RNA alone and experimental considerations for examining the holoenzyme. The combined use of these approaches has shown that the RNA and protein subunits cooperate to bind different portions of the substrate structure, with the RNA subunit predominantly interacting with the mature domain of tRNA and the protein interacting with the 5′ leader sequence. However, important questions concerning the interface between the two subunits and the coordination of RNA and protein subunits in binding and catalysis remain.

Keywords: Bacterial RNase P RNA; Holoenzyme; Substrate recognition

1. Introduction

Ribonuclease P (RNase P; EC 3.1.26.5) is distinct among cellular ribonucleases in that its catalytic subunit is composed of RNA, while its substrate binding interface is composed of both RNA and protein which cooperate to provide specificity (reviewed in [1,2]). The key biological role of RNase P is the formation of the mature 5′ end of the full complement of tRNA precursors within the cell. In addition, *Escherichia coli* RNase P processes the 5′ end of tRNA-like pseudoknotted structures in viral RNAs [3], 4.5S RNA [4], tmRNA [5], C4 antisense RNA from bacteriophages P1 and P7 [6], and a polycistronic pre-mRNA [7]. Together, the cooperation of RNA and protein in substrate binding, the breadth of RNase P substrates, and the catalytic nature of the bacterial RNA subunit make this an intriguing system for investigating how a single enzyme can process multiple substrates. Accordingly, a significant degree of effort has been directed at elucidating the structural basis for recognition of cognate substrates by RNase P. Such studies continue to be important, not only for revealing the role of RNase P in metabolism but also because the enzyme is an excellent paradigm for examining the coordination of RNA and protein function in ribonucleoprotein enzymes.

Although there is as yet no high-resolution structure for the RNase P RNA or holoenzyme, phylogenetic comparative sequence analysis, cross-linking, and mutational studies of the bacterial form of the enzyme have led to a detailed understanding of RNase P RNA secondary structure and low-resolution three-dimensional models of the ribozyme–substrate complex (Figs. 1 and 2). Bacterial RNase P RNAs are large (ca. 400 nucleotides) and highly structured with approximately 18 double-stranded regions [8]. The overall structure contains two distinct, independently folding domains: a substrate binding or "S-domain" containing helices 7–14 and a catalytic or "C-domain" containing helices 1–

* Corresponding author. Fax: 1-216-368-3055.
E-mail address: meh2@pop.cwru.edu (M.E. Harris).

1046-2023/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S1046-2023(02)00238-4

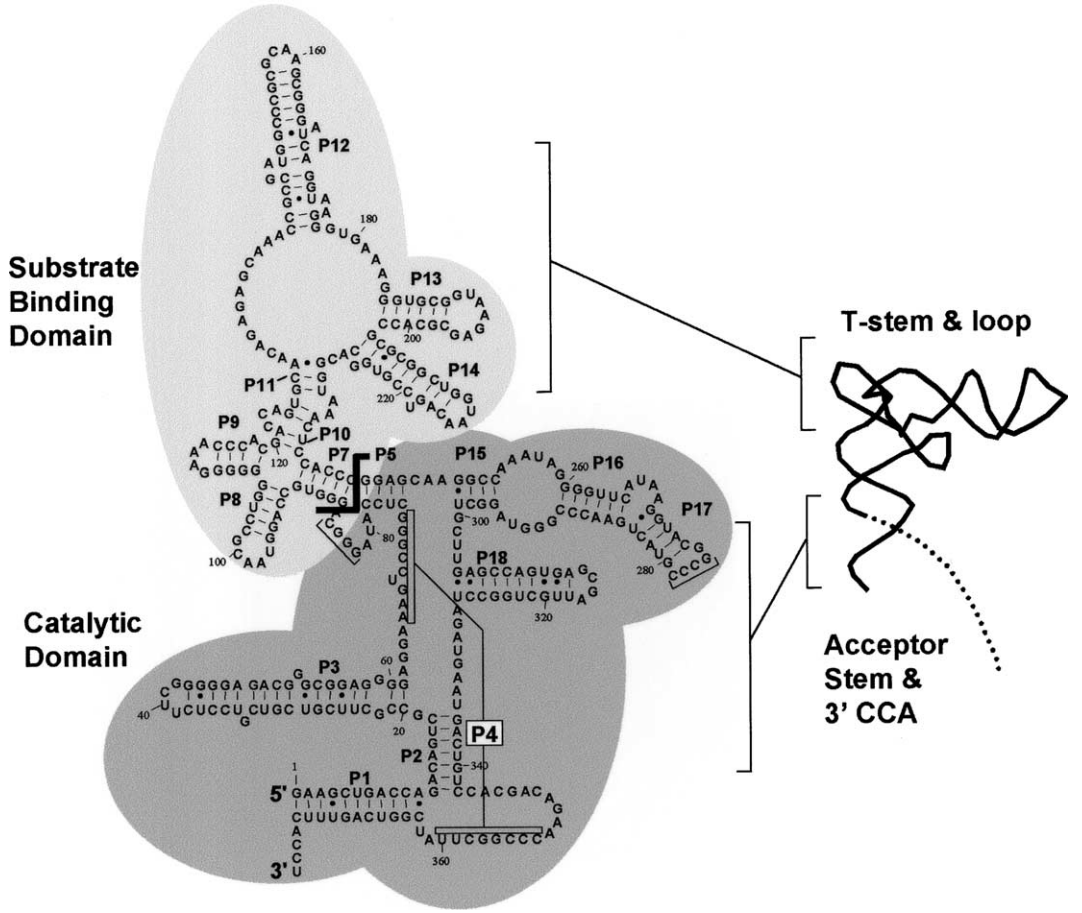

Fig. 1. Secondary structure and domain organization of *E. coli* RNase P RNA. The secondary structure derived from phylogenetic comparative sequence analysis is shown [8]. Helices are designated P, for paired; sequences joining helices are referred to in the text as J, for joining (e.g., J3/4 connects helices P3 and P4). The regions of the secondary structure that make up the substrate binding domain and catalytic domain are highlighted in light gray and dark gray, respectively. A line between P7 and P5 indicates the junction between the two domains. As described in the text, the substrate binding domain contacts the T-stem and loop of the pre-tRNA substrate, while the catalytic domain interacts with the acceptor stem and cleavage site. This general feature of substrate recognition is indicated by brackets connecting tRNA and the appropriate domain of RNase P RNA. The mature tRNA sequences are depicted as a solid line tracing the path of the phosphodiester backbone. The pre-tRNA leader sequences are depicted as a dotted line.

6, and 15–18 [9,10]. With the exception of the ribosome, RNase P is unlike other ribozymes characterized to date in that it recognizes its substrates in trans and binds them with a significant dependence on RNA tertiary structure. Using the approaches outlined in this review, the S-domain has been shown to interact with the T-stem and loop of the pre-tRNA substrate, while the C-domain has been associated with recognition of the acceptor stem, the cleavage site, and the highly conserved 3′ CCA sequence (Fig. 1) [11–17]. These and other findings have led to the current understanding that the major determinants of substrate recognition by RNase P RNA are contained in, or immediately adjacent to, the coaxially stacked acceptor stem, T-stem, and loop of the pre-tRNA substrate [18–20].

In bacteria, the protein component of RNase P contributes only one tenth the mass of the holoenzyme and is essential for function in vivo [21] (Fig. 2A). The bacterial RNase P protein subunit is approximately 120 residues, including an 18-residue conserved consensus sequence or "RNR" motif [8,22,23]. High-resolution structures have recently been solved for the protein subunit from *Bacillus subtilis* (X-ray) and *Staphylococcus aureus* (NMR) [24,25]. These studies reveal a globular structure of approximately $40 \times 35 \times 30\,\text{Å}$ that adopts the fold of an α-β sandwich and an overall topology of αββββαβα (Fig. 2B). In addition, these structures reveal three potential RNA binding motifs, including the RNR motif, a metal-binding loop, and a conserved cleft formed by an alpha helix and the four-stranded β-sheet, which appear to interact with the pre-tRNA substrate and RNase P RNA [26–28].

Numerous biochemical studies have demonstrated that the protein has profound effects on reactivity of the RNA subunit in vitro by increasing the affinity of the complex for substrate [29–31]. Although the protein component of RNase P appears to interact directly with the C-domain near nucleotide regions known to be

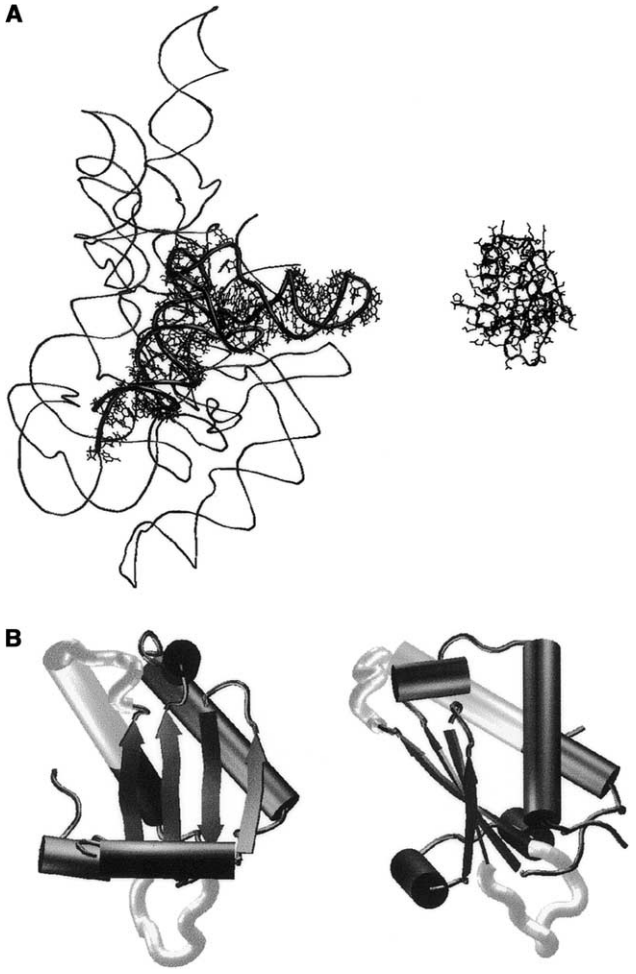

Fig. 2. Structural features of *B. subtilis* RNase P protein. (A) A proposed model of the tertiary structure of the *E. coli* P RNA–tRNA complex [77] and the crystal structure of *B. subtilis* RNase P protein [24] drawn to scale. (B) Ribbon diagrams of the *B. subtilis* RNase P protein. Two potential RNA interaction domains, the highly conserved RNR domain at the top of the structure as shown and the metal binding loop at the bottom, are highlighted in light gray. The figure on the right is rotated to show the central cleft which interacts with the substrate leader sequence [26]. Graphical representation of structures were generated using VMD [118].

important for catalytic function [28,32–36], it does not appear to play a direct role in catalysis [31]. Rather, the protein subunit contributes to substrate binding affinity by directly contacting the 5′ leader sequence [26,31,37]. Binding of the 5′ leader enhances the ability of the holoenzyme to discriminate between substrate and product and is thought to prevent product inhibition by mature tRNA, whose intracellular concentration is high compared to that of pre-tRNA [31].

In this review, we describe methods applied to the study of substrate recognition in RNase P that led to the model for substrate binding described above, focusing on in vitro studies of bacterial ribozymes, in which the majority of the work on substrate recognition has been done. Although many of the techniques discussed can also be applied to substrate recognition in eukaryotic model systems, comprehensive reviews have been published elsewhere [38,39]. The current review is divided into five sections covering general considerations of reaction conditions, quantitative measurement of substrate binding, enzymatic and chemical protection, cross-linking, modification interference and analysis of site-specific substitutions. Each section describes application of the specific method to substrate binding by RNase P RNA alone and to the holoenzyme. While many of the methods discussed also apply to the analysis of catalytic function in RNase P, the experimental considerations for this type of analysis are, to a large extent, similar to those used for other ribozymes. A thorough description of kinetic approaches for analyzing RNase P catalytic function can be found in the studies by Fierke and co-workers [31,37 and references therein, 40].

2. General considerations of reaction conditions

RNase P requires divalent metal ions, optimally magnesium (Mg^{2+}), for folding of the RNA, for binding of protein and substrate, and for catalytic activity [41 and references therein]. In the absence of protein and under near physiological conditions (e.g., 10 mM Mg^{2+}, 100 mM NaCl), the RNA has weak affinity for the substrate. This affinity can be enhanced by higher concentrations of monovalent or divalent ions, which presumably act to screen electrostatic repulsion between enzyme and substrate. For catalytic reactions a typical buffer used in our laboratory contains 1 M NaCl, 25 mM $MgCl_2$, 50 mM Pipes, pH 6.0, and 0.01% Nonidet P-40. For binding assays, calcium (Ca^{2+}), which suppresses catalysis but supports RNA folding and substrate binding, is substituted for Mg^{2+}. Many different ratios and species of monovalent and divalent ions, in addition to other components, have been used in different laboratories to achieve optimal activity under different experimental conditions. For examples of the range of experimental conditions see [39,41,42]. Otherwise, because of the general nature of this review, the reader should refer to the individual methods cited below.

Under all experimental conditions both ribozyme and substrate must be properly folded. Although renaturation conditions vary, nearly uniformly folded populations (>95%) of native ribozyme or pre-tRNA substrate can be formed by heating to 95 °C for 3 min in reaction buffer lacking divalent metal, followed by cooling to 37 °C and incubating in the presence of divalent metal for at least 15 min. The fraction of active enzyme can, under some conditions, be determined by measuring the amplitude of the pre-steady-state burst in product formation in a time course of substrate cleavage by the ribozyme ([40]; see below).

RNase P cleavage reactions are typically initiated by the mixing of prefolded enzyme and substrate and terminated by the addition of EDTA in a twofold molar excess over the divalent metal ion concentration, although efficient quenching can also be achieved by adding 2.5 volumes of ethanol or an equal volume of 10% trichloroacetic acid [40]. The RNase P RNA cleavage reaction has a log-linear dependence on pH [42], with the rate of native ribozyme becoming too rapid for accurate manual measurement above ~pH 7. Reactions occurring over short time periods are measured mechanically, by rapid quench, which can accurately measure incubation times between 5 ms and 10 s [40]. For a more comprehensive review of transient kinetic approaches see [43].

3. Quantitative analysis of substrate affinity

Quantitative measurement of substrate binding has been achieved by examining individual rate constants for association and dissociation and by several methods for the measurement of apparent binding equilibria. These studies have involved the application of pre-steady-state and pulse-chase kinetic approaches and the thermodynamic methods of gel shift, gel filtration, cross-linking, and fluorescence, which are described below. The importance of these approaches cannot be understated, as they have provided both the means and the framework upon which mechanistic comparisons have been made and from which structure-probing studies by footprinting, cross-linking, and modification interference are properly designed and interpreted.

Kinetic and thermodynamic studies have led to the description of the reaction mechanism for cleavage of pre-tRNA by RNase P which includes (i) rapid association and slow dissociation of pre-tRNA, (ii) irreversible cleavage, (iii) rapid dissociation of the 5′ leader sequence, and (iv) slow dissociation of the tRNA product [40]. Importantly, substrate dissociation from RNase P (k_{-1}) is not significantly faster than catalysis (k_2) and thus K_m does not accurately reflect the dissociation constant (K_D) for substrate binding. This finding and the observation that multiple turnover reactions are rate-limited by product dissociation have led to the predominant use of pre-steady-state or single-turnover reactions ([E] ≫ [S]), which simplifies the interpretation of kinetic data since only the chemical and preceding steps such as binding are considered.

Measurement of the rate constant for pre-tRNA binding to the ribozyme (k_1; Scheme 1) requires the examination of the reaction under specific conditions [40]. If k_2 (cleavage) is fast relative to k_{-1} (dissociation), which is the case at elevated pH, then k_1 can be determined from the slope of a plot of the observed rate versus enzyme concentration. This approach is limited to low concentrations of RNase P RNA (<1.4 μM for *Bacillus subtilis* RNase P RNA), where the association of substrate is slower than cleavage ($k_1 \ll k_2$). At high concentrations of RNase P RNA (>19 μM for *B. subtilis* RNase P RNA), the cleavage rate becomes independent of substrate binding and approximates k_2 [40].

$$\mathrm{E} + \text{pre-tRNA} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{E} \cdot \text{pre-tRNA} \xrightarrow{k_2} \mathrm{E} \cdot 5'\text{Leader} \cdot \text{tRNA}$$

Scheme 1.

The rate constants for substrate association and dissociation can also be measured using a pulse-chase or "tRNA trap" approach, which directly tests the assumption that $k_{-1} \ll k_2$ [44]. Here, single-turnover reactions are allowed to proceed for various times and are then chased by a large (ca. 1000-fold) excess of unlabeled pre-tRNA prior to the addition of EDTA to terminate the reaction. This chase prevents further binding of labeled pre-tRNA but allows labeled pre-tRNA enzyme–substrate complexes to partition between dissociation and cleavage. Accumulation of products after the chase is described by a single first-order exponential, and the second-order rate of association is calculated from the slope of a plot of the observed rate constant versus enzyme concentration. The extent of product formation after the quench is also a direct reflection of the relative rate constants for cleavage (k_2) and substrate dissociation (k_{-1}). The rate constant for pre-tRNA dissociation from the ribozyme–substrate complex can be estimated from Eq. (1),

$$\frac{[P]_{\mathrm{obs}}}{[P]_{\infty}} = \frac{k_2}{k_2 + k_{-1}}, \tag{1}$$

where $[P]_{\mathrm{obs}}$ is the fraction of pre-tRNA that partitions to form products under conditions of sufficient excess enzyme to maximize the formation of enzyme–substrate complex and $[P]_{\infty}$ is the end point of the reaction, reflecting the fraction of substrate capable of being cleaved [44]. Significant levels of product formation after the chase indicate that $k_{-1} < k_2$, while a lack or near absence of product formation indicates that $k_{-1} > k_2$. The measurement of similar values of k_1 by both pre-tRNA quench and the single-turnover method described above validates the assumption of the single-turnover approach that $k_{-1} < k_2$.

Kinetic and thermodynamic methods have also been employed to directly measure the equilibrium dissociation constant (K_D). These experiments require that the rate of cleavage be significantly slower than dissociation of pre-tRNA from the ribozyme ($k_{-1}/k_2 \geqslant 10$). This decrease in reaction rate is typically achieved by inhibiting the cleavage reaction (approximately 400-fold) through the introduction of a 2′-deoxy substitution at the pre-tRNA cleavage site or through the substitution of Ca^{2+} for Mg^{2+} (see above; [42,45,46]). It is important

to note that the 2′-deoxy substitution modifies a functional group that is important for both substrate recognition and catalysis [42,47]. Similarly, the use of Ca^{2+} in place of Mg^{2+} is limited by the fact that although Ca^{2+} supports both RNA folding and formation of a ribozyme–substrate complex, whether Ca^{2+} binds to the enzyme–substrate complex in an identical fashion is not known. With these modifications a kinetic determination of K_D can be obtained directly from a plot of the single-turnover rate versus enzyme concentration. In contrast, the measurement of substrate affinity with equilibrium binding methods has been achieved principally through the use of gel shift, size-exclusion column chromatography, and cross-linking, which are generally applicable to both the RNA alone and the holoenzyme systems.

In the gel shift approach, a fixed amount of radioactively labeled substrate is combined with varying concentrations of enzyme and incubated together for sufficient time for binding to reach equilibrium (for examples see [40,48,49]). Free and bound substrate are separated into distinct complexes on nondenaturing polyacrylamide gels. Nondenaturing gel conditions are generally similar to those used during the initial incubation with respect to temperature and monovalent and divalent salt (e.g., 1 M ammonium acetate, 25 mM $CaCl_2$, 10 mM Pipes, pH 6, 37 °C). Gels are dried and the amount of bound substrate is quantified using a phosphorimager and associated software. The observed K_D is determined by fitting a plot of the fraction of substrate bound vs [E] to a single binding isotherm (Eq. (2)),

$$\frac{[\text{pre-tRNA}]_{\text{bound}}}{[\text{pre-tRNA}]_{\text{total}}} = \frac{[\text{E}]}{[\text{E}] + K_D}, \tag{2}$$

where $[\text{pre-tRNA}]_{\text{bound}}/[\text{pre-tRNA}]_{\text{total}}$ is the fraction of bound substrate at a given concentration of enzyme and [E] is the free enzyme concentration in solution, which is approximated by the total enzyme concentration ($[\text{E}]_{\text{total}}$) when $[\text{E}]_{\text{total}} \gg [\text{pre-tRNA}]_{\text{total}}$. This approach offers the advantage of examining multiple samples in parallel with relatively little scatter in the experimental data (typically 10–20%). However, while it has been suggested that the gel matrix effectively traps the ribozyme–substrate complex by a ''caging effect'' which prevents further association or irreversible dissociation during electrophoresis [50], other observations indicate that gel retardation can discriminate against more labile complexes, leading to an overestimation of losses in binding free energy for complexes with low affinities ($K_{D_{app}} > 1\,\mu M$) [51].

Gel filtration through spin columns [52,53] has been used as an alternative method for separating bound and free populations of substrate and product [37]. Typically, small volumes of a binding reaction (~20 μL) are loaded into centrifuge columns containing ~600 μL packed Sephadex G-75 (or G-100 for larger substrates) which has been preequilibrated in binding buffer and centrifuged briefly. The radioactively labeled pre-tRNA that passes through the column into the eluate is then measured by scintillation or Cerenkov counting. Sephadex G-75 retains the majority (approximately 75%) of unbound labeled pre-tRNA while allowing the much larger ribozyme–substrate complex to pass through the column. Nevertheless, background levels of pre-tRNA are found in the eluate in the absence of enzyme. In addition, the resin retains some portion of the bound substrate even at saturating enzyme concentrations. This can be reduced by the inclusion of 0.01% Nonidet P-40 in the binding reaction and column wash buffers. Both the background pre-tRNA in the eluate and the enzyme–substrate complex retention can be accounted for during curve fitting using Eq. (3),

$$[\text{pre-tRNA}]_{\text{eluate}} = \frac{[\text{pre-tRNA}]_{\max}[\text{E}]}{[\text{E}] + K_D} + [\text{pre-tRNA}]_{\text{background}}, \tag{3}$$

where $[\text{pre-tRNA}]_{\text{eluate}}$ is the concentration of substrate in the eluate, $[\text{pre-tRNA}]_{\max}$ is the maximum amount of substrate in the eluate, $[\text{pre-tRNA}]_{\text{background}}$ is the concentration of substrate in the eluate in the absence of enzyme, and [E] is the enzyme concentration as described for Eq. (2). Importantly, pre-tRNA concentrations are proportional to, and can be substituted with, radioactivity levels determined from scintillation or Cerenkov counting.

Although spin columns tend to have a greater level of experimental error than gel shift (20–30% vs 10–20%, respectively) the short time required for its execution and the ability to examine many different solution conditions simultaneously make it an attractive method for screening effects on substrate binding in the context of both the RNA alone and the holoenzyme. It is also important to note that the use of gel filtration or gel shift methods are not limited to equilibrium studies, but can be combined with other methods such as pulse-chase or ''tRNA trap'' experiments (e.g., [40,49]) and chemical protection and modification interference studies [33,54–58].

Cross-linking provides a third method of measuring the extent of complex formation in both the RNA alone and the holoenzyme reaction by trapping the bound substrate in a stable covalent linkage [59,60]. Although the efficiency of cross-link formation varies depending on the type and position of the cross-link, the extent of cross-linking is proportional to the concentration of the enzyme–substrate complex. Cross-linked radioactively labeled substrate can easily be separated from uncross-linked material on low-percentage (4%) denaturing polyacrylamide gels. A plot of observed radioactivity in the cross-linked species versus enzyme concentration

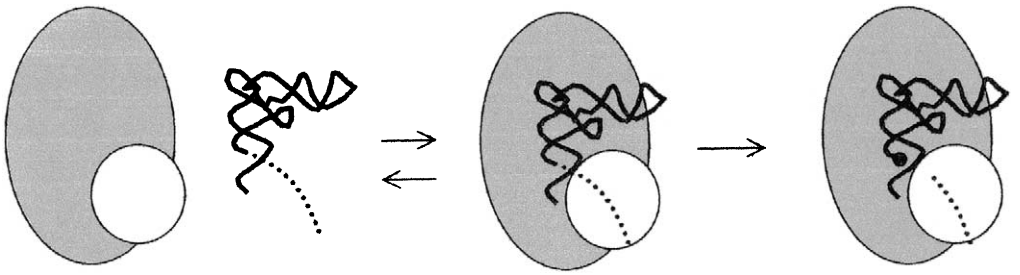

Fig. 3. Cooperation between RNA and protein during substrate recognition and cleavage as demonstrated by kinetic studies [31,37]. A diagram of the RNase P holoenzyme is shown in which the RNA and protein subunits are depicted by ovals proportional to their relative molecular weights. The RNA is shown in gray and the protein in white. The pre-tRNA substrate is shown as a backbone diagram as in Fig. 1. The filled circle at the end of the mature tRNA indicates a newly formed 5′ end resulting from cleavage. As described in the text, RNase P RNA contacts the mature tRNA portion of the substrate, while the protein interacts with 5′ leader sequences proximal to the cleavage site. These observations place the substrate cleavage site very near the interface between the RNA and the protein subunits of RNase P.

can be fit to a single binding isotherm (Eq. (2)) to determine K_D.

Equilibrium binding has also recently been examined by fluorescence [51]. Reduction in fluorescence emission has been observed on binding of a 3′-fluorescein-labeled tRNA to RNase P RNA. A plot of the change in fluorescence versus enzyme concentration yields values similar to those obtained by spin column and gel shift assays. A potential limitation of this approach is the requirement for large amounts of enzyme in the case of enzyme variants with significantly reduced affinity for substrate. Nevertheless, given the significant application of fluorescence in other RNA systems, the use of this approach offers a frontier for future work in RNase P [61]. Application of spectroscopic approaches, for example, may allow the examination of potential conformational changes that have been proposed to occur after initial binding to RNase P RNA [11,62–67].

An excellent example of the power of these quantitative approaches is illustrated in the elucidation of the function of the RNase P protein (Fig. 3) [31,37]. Here, the kinetic and thermodynamic comparison of *B. subtilis* RNase P RNA and holoenzyme were able to pinpoint and quantitatively assess the mechanistic role of the protein in increasing affinity for the precursor over product, to exclude other potential roles such as a direct enhancement of catalysis, and to identify the apparent position of protein contacts on the 5′ leader sequence. These observations not only positioned the protein near the active site in RNase P RNA and explained how the holoenzyme can catalyze pre-tRNA cleavage in the large intracellular excess of tRNA products but also provided the framework for more directed chemical protection and cross-linking studies as described below.

4. Enzymatic and chemical protection

Enzymatic and chemical protection, using a wide variety of reagents, has been used to identify regions of contact between substrate and RNase P RNA and with the holoenzyme (Fig. 4; for a general review of reagents and protocols see [68]). Monitoring changes in enzymatic or chemical protection can also provide a qualitative measure of the extent to which individual modifications or mutations influence structure or contribute to substrate binding (e.g., [12,62,65,69]). Moreover, given that the holoenzyme interacts with a broader array of substrates than RNase P RNA alone [4,70], differences in the protection pattern of substrate binding to RNase P RNA and holoenzyme could be used to begin to address the structural basis for this important aspect of substrate selectivity (e.g., [36]).

In the study of RNase P, enzymatic and chemical protection has focused almost exclusively on the protection of RNA components because of the relative ease of this type of analysis. The overall approach is to subject substrate or RNase P RNA (alone or in the context of holoenzyme) to limited RNA hydrolysis or chemical modification using specific chemical or enzymatic probes. An average of less than one cleavage or modification per molecule is essential to avoid potential effects arising from multiply modified molecules. Two general methods of nucleic acid sequencing can then be used to resolve the position of RNA cleavage or modification, depending both on the length of RNA being studied and on the type of structural probe. The first method employs radioactively end-labeled RNA targets and structural probes that break the RNA chain either by enzymatic cleavage or by subsequent chemical treatment followed by separation on denaturing gels. This method is well suited to the analysis of pre-tRNA substrates; however, the analysis of RNase P RNA is more difficult because of the limited resolution of larger RNA fragments on sequencing gels. In addition, analysis from the 3′ terminus can be problematic due to end heterogeneity associated with runoff transcription [71]. The second method uses primer extension and the termination of synthesis at cleaved or modified nucleotides [72]. This approach not only provides a clearer analysis

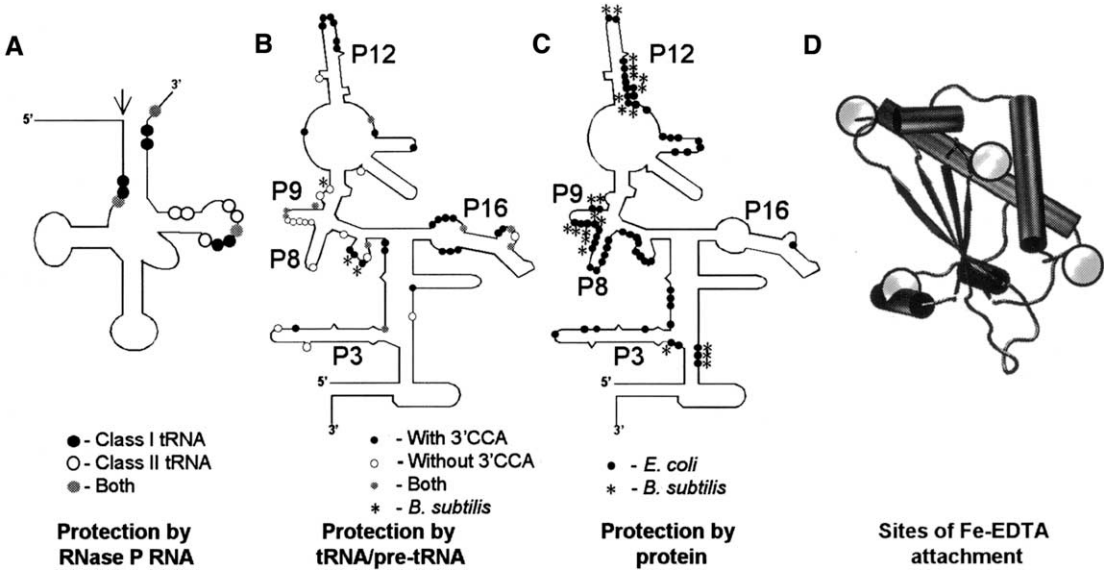

Fig. 4. Summary of studies using enzymatic and chemical protection. (A) tRNA secondary structure showing sites of phosphorothioate incorporation at which iodine cleavage is either enhanced or suppressed on binding of pre-tRNA to RNase P RNA [65]. Black circles indicate positions of enhancement or suppression by tRNAs with a short variable arm (Class I); white circles are similar sites by tRNAs with a long variable arm (Class II); gray circles were detected with both classes of tRNA. The cleavage site is indicated by an arrow. (B) Protection of RNase P RNA from chemical modification on pre-tRNA or tRNA binding. Positions marked with circles were observed in *E. coli* RNase P RNA and asterisks mark homologous positions observed in the *B. subtilis* ribozyme. Black circles indicate protections by full-length tRNA [12,62,69]; white circles indicate positions protected by a tRNA lacking a 3′ terminal CCA [12,62]; gray circles indicate protections by tRNAs with and without a 3′ terminal CCA. (C) Enzymatic and chemical protection of RNase P RNA by RNase P protein [32,33,36]. Circles and asterisks as in (B). (D) Ribbon diagram of the *B. subtilis* RNase P protein crystal structure [24] showing locations of Fe–EDTA attachment to *E. coli* C5 protein [28]. Sites of modification were mapped to homologous locations in the crystal structure using a protein sequence alignment [25]. Note that sites of induced cleavage on the RNA subunit are shown in Fig. 5C.

of nucleotides in the middle of large RNAs but is the only practical method to resolve numerous chemical modifications that cannot be revealed by chemical scission (see below).

The reagents used in protection can be roughly divided into three general groups according to the level of structural detail that they can provide. Nuclease protection has been used as a rapid but low-resolution means of establishing areas of potential contact with a target RNA (e.g., [32]). Care must be taken in the interpretation of such data because of the potential steric limitation of enzymatic cleavage and the inherent level of variation in cleavage signal in these types of experiments. RNases differ both in their requirements for reaction conditions, which can influence the structure of the RNA, and in their primary and secondary structure specificity, which may limit the ability to probe specific regions of potential contact. Nevertheless, commercially available nucleases offer the ability to examine a range of primary and secondary structural features [68].

A more general and less sterically restrictive probe for the solvent-accessible regions of the substrate and ribozyme is hydroxyl-radical footprinting. The Fe(II)–EDTA reagent has typically been used to generate hydroxide radicals which attack the phosphate backbone, causing strand scission [73,74]. Ribose moieties that are less accessible to solvent, due to folding or binding of other components, are protected from hydroxyl-radical attack [73,74]. This approach has been used to identify protected regions arising from substrate (Fig. 4B) or protein binding (Fig. 4C) with RNase P RNA (for conditions see [34,36,73]). However, the resolution of this approach is still somewhat limited in that it cannot distinguish between direct contacts and indirect structural changes resulting from binding. A more focused application of this approach has been obtained by tethering the Fe–EDTA reagent to a specific position in the protein subunit (Fig. 4D) [28] (see below) and could be applied to the RNA subunit also.

Chemical protection of individual functional groups can provide greater structural detail in binding studies. Base and backbone functional groups involved in substrate recognition have been examined by a standard set of chemical reagents used to probe RNA structure [12,62,69]. These studies have applied the more commonly used reagents, including dimethyl sulfate, which methylates guanine N7, adenosine N1, and cytosine N3; kethoxal, which reacts with guanine N1 and N2; and diethyl pyrocarbonate, which reacts with adenosine N7, although other probes for base and backbone functional groups are useful also (for a list of probes and assay conditions see [68]). Importantly, the reactivity of individual functional groups is dependent on their local structural and electrostatic environments [75]. For

example, functional groups tend to be resistant to modification if they are involved in hydrogen bonding interactions such as base pairing. In this respect, sensitivity to chemical modification is generally limited to a subset of functional groups not otherwise involved in secondary or tertiary interactions. In addition, as with many of the other methods that have been discussed, there are numerous indirect sources of experimental signal, such as conformational heterogeneity of the folded RNA and conformational changes in the RNA subunit induced by protein or substrate binding. Footprinting by enzymatic or chemical protection, therefore, cannot be interpreted beyond the proximity of general regions and should be performed in parallel with other methods such as cross-linking and modification interference.

An example of effective application of footprinting is provided by Pace and co-workers [12] in work examining protection by wild-type and truncated forms of pre-tRNA and tRNA. In addition to identifying regions of RNase P RNA subsequently shown to contact the T-stem and loop of the pre-tRNA substrate (see below), these experiments identified specific sets of nucleotides associated with the 3′ RCCA motif and the 5′ leader sequence adjacent to the cleavage site, in agreement with concurrent and subsequent cross-linking and mutational studies [11,13,60,76]. Serial deletion of the 3′ RCCA also provided a means of orienting this conserved tRNA motif with respect to its recognition element in RNase P RNA. Notably, these findings were shown to be consistent in three homologous but structurally distinct RNase P RNAs, an approach widely used to distinguish idiosyncratic features from those that are conserved and thus more likely to be central to RNase P function.

5. Cross-linking

Cross-linking has been used extensively in the study of RNase P structure and substrate recognition to more precisely determine the orientation of specific elements of the RNA and protein components (Figs. 5A–C). Cross-linking has also been instrumental in the determination of physical constraints that, together with constraints derived from sequence comparisons, have led to the development of three-dimensional models of the ribozyme–substrate complex [77,78]. Cross-linking has been achieved using ultraviolet (UV) light to induce cross-linking between unmodified RNAs [79,80] and by random and site-specific incorporation of photoaffinity reagents (reviewed in [81]). UV cross-linking of unmodified RNAs offers the advantages of producing a short-distance constraint (2–3 Å) and requiring no prior modification of substrate or enzyme. Experimentally, UV cross-linking is also particularly straightforward, requiring a short (~3-min) exposure of the enzyme–substrate complex to 254-nm light, the purification of slowly migrating cross-linked species from polyacrylamide gels, and the identification of the site of cross-linking by reverse transcription as described above. The disadvantage of this approach is that the strong structural dependence of UV cross-linking produces relatively few cross-links to the pre-tRNA substrate, leaving much of the interface between substrate and enzyme relatively undefined.

A much larger number of structural constraints has been obtained through the use of site-specific incorporation of long-range and short-range photoaffinity reagents. Long-range cross-linking using arylazides has produced cross-links from numerous positions in the tRNA substrate, in RNase P RNA, and in the protein subunit (Fig. 5). In the study of RNase P, the most commonly used cross-linking reagents are azidophenacyl (APA) derivatives which contain an azide moiety approximately 9 Å from their point of attachment to RNA or protein [81]. APA reagents can be used over a wide range of experimental conditions and remain inert unless exposed to 302-nm light, allowing RNAs to be in fully folded complexes prior to photoagent activation. Although these reagents have provided a large number of distance constraints, they do so with considerable uncertainty (±9 Å), often producing a cluster of cross-links in specific regions of the RNA secondary structure.

In RNA, APA derivatives are often incorporated at the 5′ or 3′ termini because of the speed, yield, and simplicity of the coupling reaction. APA is usually linked to a 5′ terminal guanosine monophosphorothioate (GMPαS) that has been introduced by priming transcription with a large excess of this modified nucleotide [81]. The 5′ terminal phosphorothioate reacts quantitatively with azidophenacyl bromide and the resulting product can be purified by standard phenol extraction and ethanol precipitation. Attachment of APA to the 3′ end of RNA has been accomplished by modification of the 3′ terminal ribose with sodium periodate and an alkyldiamine to obtain a primary amine, which is subsequently coupled to the azide moiety through an *N*-hydroxysuccinimidyl group [13].

The APA cross-linking agent can be placed at different positions by using circularly permuted forms of the substrate or RNase P RNA, which join the native 5′ and 3′ ends by a short oligonucleotide sequence and introduce novel termini elsewhere in the structure [82,83]. Circularly permuted constructs for in vitro transcription are easily generated by polymerase chain reaction from tandem genes [17,45,82]. However, when moving the 5′ and 3′ termini through circular permutation, it is critical to ensure that the permutation does not significantly alter substrate binding and catalytic activity. Fortuitously, both tRNA and RNase P RNA can tolerate movement of the termini to many different positions, including sites proximal to the pre-tRNA cleavage site, the 3′ RCCA sequence, and the T-stem

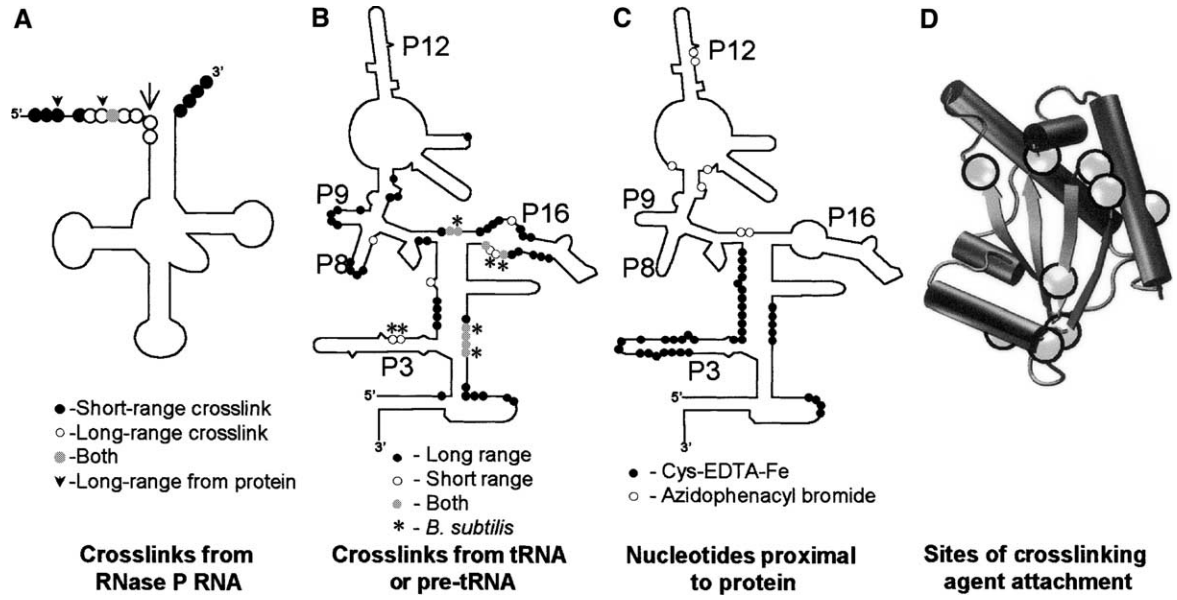

Fig. 5. Summary of studies using cross-linking. (A) Sites of cross-linking to the substrate. Black circles indicate sites of short-range cross-links from the ribozyme [79,91]; white circles indicate long-range cross-links [82]; light gray circles indicate sites of both long-range and short-range cross-links. Small arrows indicate positions of long-range cross-links from the *B. subtilis* protein [26,27]. The large arrow indicates the site of pre-tRNA cleavage. (B) Cross-linking between the tRNA or pre-tRNA and the ribozyme. Positions marked with circles were observed in the *E. coli* ribozyme and asterisks mark homologous positions observed in *B. subtilis*. Long-range cross-links are shown as black circles [13,16,17,78,82]; short-range cross-links are shown as white circles [60,76,79,80,91]; sites of both long-range and short-range cross-links are shown as gray circles. (C) Regions of the ribozyme proximal to the protein. Black circles are sites determined by hydroxyl-radical cleavage [28]; white circles are sites of protein cross-linking to the ribozyme [27]. (D) Ribbon diagram of the *B. subtilis* RNase P protein crystal structure [24] showing locations of attachment for photo-cross-linking reagents which cross-link to the 5′ leader sequence of pre-tRNA [26].

and loop. Thus, the use of long-range photoaffinity agents has helped to define clusters of highly conserved nucleotides and structural motifs that are likely to mediate recognition (Fig. 5) [13,16,17,78,83].

Incorporation of long-range cross-linking reagents into the protein subunit has been accomplished by the introduction of unique cysteine residues at specific positions within the protein and the subsequent modification of the cysteine's thiol moiety with an APA reagent (Fig. 5D) [26,27]. The positioning of unique cysteines has also been applied in two other methods useful in studying substrate recognition in the holoenzyme. The sulfhydryl-specific reagent (1-oxyl-2,2,5,5-tetramethyl-Δ^3-pyrroline-3-methyl)methane thiosulfonate, has been used to introduce a nitroxide spin label, allowing the assessment of the proximity of specific protein positions to RNase P RNA by examining changes in the electron paramagnetic resonance spectra [84]. Unique cysteine mutants have also been modified with EDTA-2-aminoethyl-2-pyridyl disulfide, effectively converting the protein into a site-specific chemical (Fe–EDTA) nuclease that cleaves the RNA backbone within approximately 20 Å from the site of attachment [28]. The Fe–EDTA reagent can also cleave proteins, including its host molecule, and may be useful in mapping the relative orientation of protein in different holoenzyme–substrate complexes [85–87] (for review of the use and methods of attaching these hydroxyl-radical reagents to proteins, see [88]). Combined, these techniques have significantly shaped our understanding of the structure of the holoenzyme. In particular, these studies have provided evidence for the direct interaction of the 5′ leader sequence with the central cleft of the RNase P protein, the orientation of the protein relative to the substrate, and the positioning of the protein near the active site in RNase P RNA.

Short-range photo-cross-linking reagents have provided significant refinement of our understanding of substrate binding by both decreasing distance constraints to approximately 2–3 Å and reducing the number of regions and nucleotides involved in cross-linking in the target molecule [81]. Short-range cross-linking in RNase P has almost exclusively focused on the use of the thionucleotide analogs 4-thiouridine (s^4U) and 6-thioguanosine (s^6G), although other analogs (e.g., s^2C) should also be useful in this type of analysis. For review of thionucleotides see [91] and for examples of conditions and specific applications of s^4U and s^6G see [61,65,77,92,93]. In contrast to long-range cross-linking reagents, the introduction of a single sulfur atom in these nucleotide analogues offers the important advantage of minimizing the perturbation of RNA structure.

Thionucleotide cross-linking agents have been used as structure probes by both random and site-specific incorporation. Random incorporation of these analogs during in vitro transcription provides a rapid means of

surveying potential contacts over an entire RNA transcript. In practice, the large number of cross-links that can be generated by this approach can be difficult to separate on polyacrylamide gels and to characterize by primer extension sequencing. However, this ambiguity can be reduced or eliminated by selective deletion or mutation of nucleotide positions contributing to the cross-linking signal [76]. In contrast, site-specific incorporation eliminates the ambiguity of the source of cross-linking and generally increases the experimental signal. The primary disadvantages of site-specific incorporation are that each position must be examined independently and that each position generally requires a distinct construct. Site-specific modification can be easily achieved by priming transcription of wild-type and circularly permuted constructs with the monophosphate form of s^6G (s^6GMP) (see [60]). Although the monophosphate form is required to render s^6G sufficiently soluble for in vitro transcription, it is not commercially available and is generally synthesized by chemical phosphorylation of s^6G [81]. Priming transcription with nucleotides derived from U is much less efficient, although it may be incorporated as part of dinucleotide primers (e.g., s^4UpG) [89].

Internal site-specific modification with cross-linking reagents can be achieved by incorporating thionucleotides into short oligonucleotide fragments of the pre-tRNA substrate or RNase P RNA, which are subsequently joined to the remainder of the molecule using olignucleotide-directed RNA ligation [90]. This approach offers the advantage of being able to introduce a broader range of cross-linking reagents and controls for the potential structural complications of cross-linking from a 5′ terminal nucleotide, whose structure may be perturbed by circular permutation. Oligonucleotide-directed ligation, however, can be inefficient and thus it may be difficult to obtain sufficient material to map weaker cross-links.

Although individual studies have varied somewhat in their methods for validation and interpretation of individual cross-links, there are a number of common approaches that have been used to insure functional relevance of the structural information. First and foremost, the conformational state of the interacting molecules is examined by determining whether an individual cross-linked species retains catalytic activity. Second, the presence of a particular distance constraint is confirmed by the demonstration of reciprocal cross-links. Third, the generality of a cross-linking interaction can be tested by comparing homologous, but structurally distinct, RNA structures (e.g., *E. coli* RNase P RNA versus *B. subtilis* RNase P RNA). Finally, it should be noted that in all cross-linking studies the strong geometrical and chemical requirements for bond formation dictate that the absence of a cross-link cannot be interpreted as the absence of proximity and that the presence of a cross-link does not provide evidence of a native or direct interaction but rather provides only a maximal distance constraint.

Using the approaches above, site-specific positioning of s^6G and s^4U at the pre-tRNA cleavage site and along the 5′ leader sequence has allowed the identification of distinct single-stranded regions in RNase P RNA associated with the 3′ CCA sequence, the nucleotides 5′ and 3′ of the scissile phosphate, and regions of the ribozyme which appear to be associated with different positions of the 5′ leader sequence (Fig. 5) [60,76,91]. Taken together, this series of cross-links both orients the pre-tRNA cleavage site within the conserved core of RNase P RNA and defines the position of the 5′ leader on the surface of the ribozyme. Interestingly, the regions of RNase P RNA identified by cross-linking from the substrate generally overlap with regions implicated in protein binding by the methods described above (Figs. 5 and 7) [28,33,34,36].

6. Modification interference and analysis of site-specific substitutions

The analysis of specific functional groups involved in substrate recognition has come from a combination of two approaches: an open-ended survey of functional groups by modification interference and the analysis of site-specific substitutions. Modification interference has been the method of choice for initial identification and comparative screening of candidate functional groups due to its ability to test nearly all nucleotide or backbone functional groups in a molecule simultaneously for functional significance (Fig. 6). However, a complete understanding of biological function requires the significantly more laborious kinetic and thermodynamic analysis of particular functional groups, in isolation or in combination with potential interacting partners. Due to the significantly greater relative complexity and effort required for characterization of individual site-specific substitutions, such studies have lagged well behind the data collected by modification interference.

Briefly, modification interference analyzes the ability of a pool of molecules with a limited number of randomly distributed modifications to perform a specific function (reviewed in [92]). In the current discussion, the ability to form an enzyme–substrate complex is used as a means of separating active variants from those with impaired function due to modification at a specific position. The positions of modification in active and inactive populations are then identified by chemical cleavage in a manner analogous to that described for enzymatic or chemical protection. Positions that are reduced (or enhanced) in the bound fraction relative to that in the unselected population indicate chemical groups important for substrate recognition.

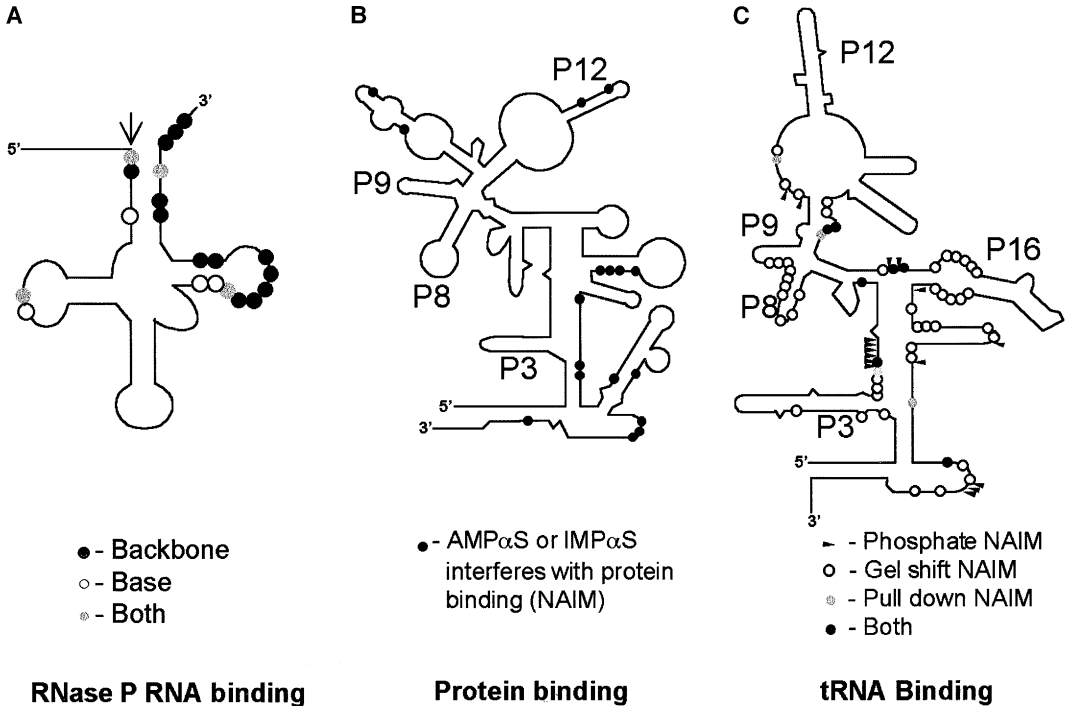

Fig. 6. Summary of studies using modification interference. (A) Sites on pre-tRNA molecules identified as important for substrate binding by NAIM and not likely to be due to structural perturbations [57 and references therein]. Black circles indicate positions of interference due to backbone modifications; white circles indicate base functional group interferences; light gray circles indicate both. The cleavage site is indicated by an arrow. (B) Functional groups of the *B. subtilis* ribozyme important for protein binding as determined by NAIM using gel shift [95]. (C) Sites on the ribozyme which are important for substrate binding as determined by NAIM. Arrows designate phosphate groups important for binding determined using gel shift [55]; white circles represent nucleotides with functional groups important for binding determined by gel shift [54,57,58]; gray circles represent base and backbone positions determined by biotin pulldown [45,94]; black circles are sites identified by both gel shift and biotin pulldown.

Limited modification (optimally an average of no more than one per molecule) can be achieved by three general methods: by the addition of nucleoside analog triphosphates during in vitro transcription [54,55,57,58, 93–95], by random chemical synthesis of short oligonucleotide fragments which can then be ligated [90] or reconstituted with the remaining portions of the substrate or ribozyme [96–98], or by posttranscriptional modification by enzymatic and chemical treatment [14,19,20,45,68]. Of these, the first method, known as nucleotide–analog interference mapping (NAIM) has been the most widely applied method in the context of RNase P RNA and holoenzyme systems. NAIM utilizes a wide range of modified nucleotides that delete or modify individual functional groups on the base and the sugar phosphate backbone and offers the advantage of introducing modifications uniformly throughout large RNA transcripts. This approach, however, is limited to the subset of potential analogs that can serve as substrates for in vitro transcription, although mutant forms of phage T7 RNA polymerase have helped to increase the permissiveness of analogue incorporation [92,99].

Separation of functional and nonfunctional RNA pools in substrate binding has been achieved through two main strategies and requires the inhibition of catalysis as described in the general discussion of reaction conditions. The first strategy utilizes gel shift in a manner analogous to that described for the quantitative analysis of substrate binding [33,46,58 and references therein]. This method allows the separation of unbound substrate and authentic enzyme–substrate complexes from nonspecific aggregates which are normally trapped in the loading well. This method also allows the identification and separation of multiple or anomalous bands that can arise from complex formation, which would otherwise complicate the interpretation of the interference signal. The second approach involves the use of biotinylated pre-tRNA attached to streptavidin–agarose or streptavidin-coated magnetic particles to affinity select enzyme–substrate complexes (see [45,94]). This approach offers the advantage of simple and rapid separation of bound and unbound RNAs by centrifugation; however, care must be taken to determine that biotinylation or immobilization do not perturb substrate binding [45]. Studies with the holoenzyme have applied an analogous method of selection based on the utilization of recombinant RNase P protein containing an N-terminal histidine tag and nickel–nitrilotriacetic acid resin, which allows separation of bound and unbound modified RNAs by centrifugation [95].

As noted above, the identification of a particular chemical group important to substrate recognition is obtained by comparing the relative sequencing band intensities of bound and unbound RNA fractions at a particular position. Changes in band intensity, however, can be quite subtle and can be difficult to distinguish when examining sequences far from the 5′ or 3′ end. Although a qualitative distinction can often be made by visual inspection, the significance of a change in band intensity can be assessed quantitatively by measurement of band intensity (using a phosphorimager) and correction for both the background signal and the differences in loading of individual lanes [92,94,100]. Significant changes are generally regarded as two standard deviations from the average background at an individual position to assure a 95% probability that the observed signal does not arise by random variation.

In comparing findings by modification interference, it is important to note that different experimental conditions or selection methods can often produce overlapping, but not identical sets of data (Fig. 6). For example, the folding of RNase P RNA is particularly sensitive to the concentrations of monovalent and divalent ions and consequently different ionic conditions can significantly alter both pre-tRNA affinity and functional groups identified by modification interference (see [101]). In addition, selection methods such as gel shift appear to allow greater sensitivity to general perturbations of tertiary structure than pulldown methods [58]. Nevertheless, NAIM analysis from a set of different laboratories using a variety of analogues and conditions has identified a surprisingly consistent set of functional groups that perturb substrate recognition, thus focusing the number of positions to be examined by further investigation [54,56–58,93,94].

To more fully explore the largely qualitative nature of findings by modification interference or other structure-probing methods discussed above, the analysis of individual functional groups by site-specific modification has utilized the quantitative and interpretive power of traditional enzyme kinetics (see [31,36 and reference therein, 37]). As noted above, site-specific functional group modifications have been introduced by mutation using standard protocols (see [102–105]) or by incorporation of specific analogues into short oligonucleotides by transcription or chemical synthesis and subsequent joining to the remainder of the target molecule by ligation (see [15,90,106]). In addition to quantifying the contribution of an individual functional group to binding (see above), the identification of tertiary contacts has been largely examined by "rescue" of the original structural perturbation. This is accomplished by altering the base or functional group of interest (e.g., [14]), by compensatory mutation of the interacting nucleotide (e.g., [11,14]), or by the use of thiophilic metal ions (e.g., [55]).

An example of how modification interference can be combined with the analysis of site-specific modifications to identify specific interactions is illustrated in the work of Pan and co-workers [14]. The identification of 2′-hydroxyl (2′ OH) groups involved in substrate binding was achieved by exposing a circularized tRNA to limited alkaline hydrolysis, converting 2′ OH groups to 2′,3′-cyclic phosphates at random positions, and identifying sites important to binding. While 2′ OH moieties are both hydrogen bond acceptors and donors, 2′,3′-cyclic phosphates are capable only of accepting hydrogen bonds. A second round of modification interference analysis was then performed in which T4 polynucleotide kinase was used to resolve 2′,3′-cyclic phosphates into 2′ or 3′ terminal phosphates. Comparison of results from the two rounds of modification interference identified positions at which 2′ OH groups act as functionally important hydrogen bond donors. Hydrogen bonding partners were subsequently identified by partial rescue of 2′,3′-cyclic phosphate-containing substrates through site-specific mutations of RNase P RNA.

Additional interactions have been identified through the analysis of site-specific mutations and modifications. For example, the 3′ terminal RCCA motif of pre-tRNA substrates has been shown to be critical to substrate binding and to be adjacent to the J15/16 internal bulge by cross-linking and chemical protection [12,13]. However, direct demonstration of the specific interaction has come most convincingly from the analysis of specific combinations of compensatory mutations in pre-tRNA and RNase P RNA [11]. In addition, nonbridging phosphate oxygens, which were identified as important for substrate binding by modification interference, have been shown to interact with magnesium ions through metal ion specificity switch experiments [55]. In all these examples, rescue is consistent with identification of a direct contact; however, care must be taken to insure that the rescue is not due to indirect effects, such as RNA folding or changes in metal ion concentration and identity [107].

7. Overview

The combined use of multiple low- and higher-resolution biochemical methods described above has led to a significant increase in our understanding of substrate recognition by bacterial RNase P. As illustrated in Fig. 7, cross-linking, chemical protection, and modification interference studies have identified a series of conserved nucleotides located in P11, J5/15, J18/2, and P15–P16 internal bulge in RNase P RNA that participate in substrate binding. In the substrate binding domain, combined cross-linking, kinetic, and interference data show that P11 contacts 2′OH groups in the T-stem [14]. Nucleotides adjacent to P11 in P9 also cross-link to the

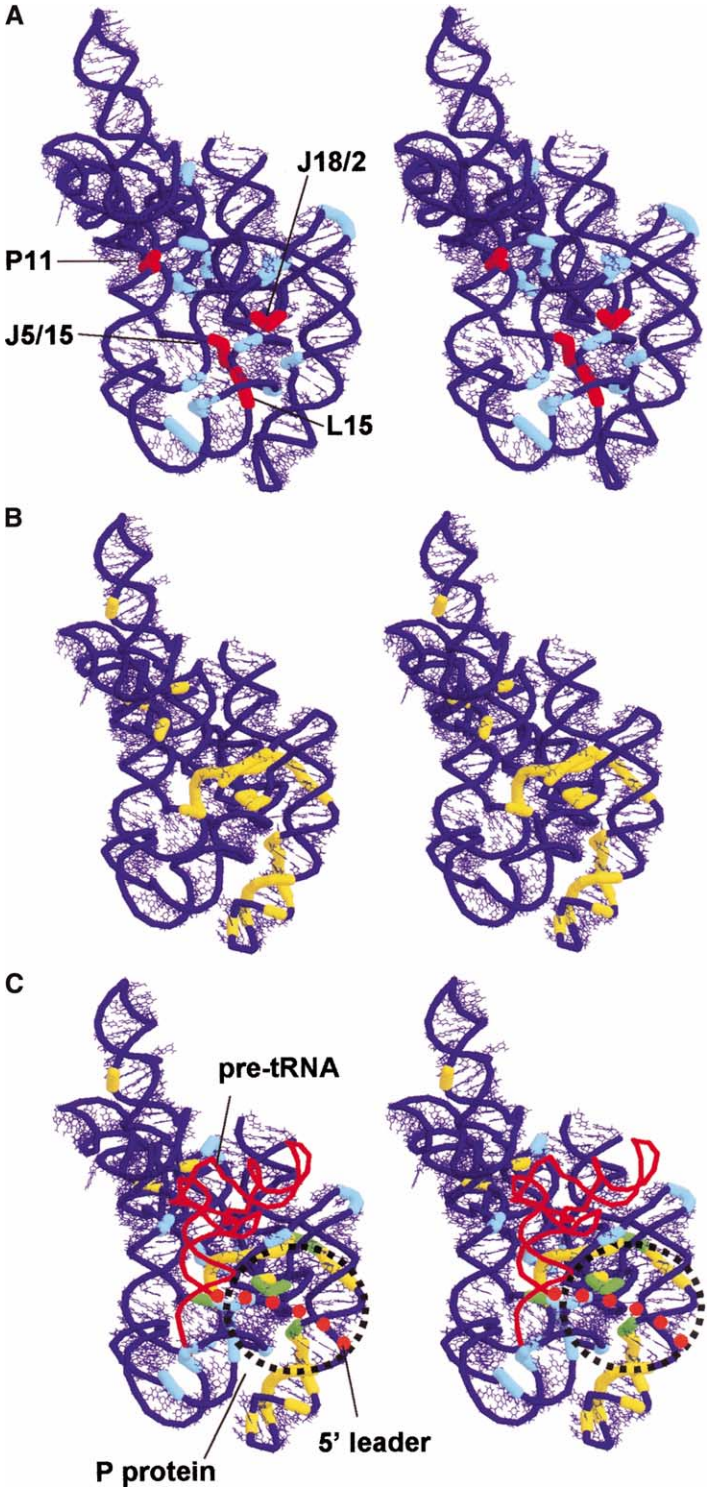

Fig. 7. Sequences and structures involved in substrate interactions and protein binding as indicated by combined cross-linking, chemical protection, and interference data. (A) Data from intermolecular cross-linking, chemical protection, and selected modification interference studies (light blue) are shown in the context of a three-dimensional model of the enzyme–substrate complex (dark blue) [77]. The subset of regions of RNase P structure where contacts between the enzyme and the substrate have been defined (P11, J5/15, J18/2, and P15–P16 internal bulge) are shown in red. The pre-tRNA is omitted from (A) and (B) for clarity. (B) Data from analysis of the P protein binding site using site-specific cross-linking and chemical probing are shown in yellow. (C) Position of the pre-tRNA substrate (red) based on cross-linking and chemical protection data. The biochemical data suggesting sites of substrate (blue) and protein (yellow) contact are indicated as in (A) and (B), with overlapping information shown in green. Based on these data, a hypothetical position for the P protein is indicated by a dashed circle that is proportional to the relative size of the protein. A probable path for the 5′ leader sequence based on the protein position and intermolecular cross-links obtained from the RNA-alone reaction is indicated by a red dotted line.

T-stem and numerous positions in or adjacent to P9, P10, and P11 have been shown to be linked to substrate binding by NAIM or chemical protection [12,17,54,55,57,58,62, 69,94]. Although multiple lines of evidence also point to a direct role in substrate binding for J11/12 and J12/13, the structure and function of this region is not well defined and indeed is an important area for further study.

A more detailed perspective has been gained for the catalytic domain and interactions at or adjacent to the pre-tRNA cleavage site. Here the ribozyme makes contact with the 3′ terminal CCA sequence and functional groups at the cleavage site, while the protein binds to the 5′ leader sequence. Initial kinetic studies of substrate deletion mutants showed that the 3′ CCA contributed significantly to binding affinity [108–110]. Subsequent cross-linking and chemical protection studies, performed using native and mutant tRNAs, showed that the P15–P16 internal bulge region of the ribozyme is involved in contacting these sequences [12,13]. The structural basis for these interactions was defined by analysis of compensatory effects of mutations in the ribozyme and CCA sequence [11]. Additional modification interference and mutagenesis experiments and NMR studies have provided a detailed structural model of these interactions [36,51,57,58,105,111–113]. In addition to the 3′ CCA interactions, other functional groups, including the 2′OH at the cleavage site and the G(1)–C(72) pre-tRNA basepair, appear to act as specificity determinants [42,47,114]. Short-range cross-linking studies consistently detect J5/15 and J18/2 as proximal to the nucleotides flanking the cleavage site [60,76]. Additionally, both regions contain functional groups important for substrate binding and catalysis [45,54,55,57,58,94,115,116]. Recently we observed that mutations in J5/15 can suppress miscleavage supporting a direct role for this element in substrate contacts (N.H. Zahler and M.E. Harris, unpublished). J18/2 is protected from chemical modification by the substrate, but not the tRNA product, providing additional evidence that this region is near the 5′ leader sequence. However, even though mutations in J18/2 disrupt binding, a direct role in contacting the leader sequences has not been established. Several studies indicate that leader nucleotides proximal to the cleavage site influence recognition by the ribozyme [12,37,47,93,114,117]. Thus, a focus of future studies will be to further define the matrix of contacts between the C-domain and the substrate cleavage site.

Because the protein subunit has been shown to contact the leader sequence [26,27], the functions of J5/15 and J18/2 or other elements of RNase P RNA which contact substrates near the cleavage site may be coordinated with, or modulated by, the binding of the protein. Recent studies of the protein binding site using site-specific Fe–EDTA hydrolysis demonstrate that the majority of proposed contacts are in the catalytic domain and include regions important for binding and catalysis in the RNA-alone reaction. In several instances, the identified regions overlap with regions of RNA structure that have been shown to be associated with the substrate cleavage site (Figs. 5C and 7B) [28]. Potential RNase P protein contacts in the substrate binding domain were also detected in cross-linking and protection studies [27,32,49,95]. Although these latter results seem difficult to rationalize with respect to current structure models of the ribozyme–substrate complex, their consistent observation again suggests that models of this region of RNase P RNA are still in need of significant refinement.

Some clues with respect to the coordination of RNA and protein function can be gained when the data for protein binding and substrate binding are mapped together onto the current three-dimensional model of the ribozyme–substrate complex (Fig. 7C). The observation that regions of the ribozyme that are proximal to the cleavage site are also contacted by the protein permits the general positioning of the protein binding site in the enzyme substrate complex. In particular, the available biochemical and structure-probing data suggest that the substrate cleavage site is recognized at the interface between the RNA and the protein subunits of the RNase P holoenzyme. Taken together, the data point to a distinct strategy for cooperation of RNA and protein structure in ribonucleoprotein function. Rather than utilizing protein to facilitate RNA folding and stabilization of catalytic RNA structure as in other ribozymes, the protein contributes directly to substrate binding and specificity by contacts to the cleavage site and possibly to the active site itself.

The methods discussed in this review have clearly defined numerous regions and functional groups, in both the substrate and the enzyme, that are important for substrate binding. However, only a small number of specific interactions have been defined. Particularly conspicuous in their absence are interactions involving substrate functional groups immediately adjacent to the pre-tRNA scissile phosphate. These interactions will be important for the formation of the active site common to all substrates, particularly in the context of the holoenzyme where the importance of contacts to the T-stem and loop are diminished [36]. Also, despite the growing perspective on the contacts between substrate and enzyme, we know little about the order in which intermolecular contacts are established and whether they are associated with conformational changes. Moreover, the principles defined for RNase P recognition of pre-tRNA have not been generally extended to encompass the recognition of the wide variety of other substrates with which RNase P is known to interact, and thus much remains to be done. In sum, the methods presented here provide a powerful tool chest for elucidating the structural basis for recognition of multiple substrates by RNase P and the recognition of RNA substrates in general.

Acknowledgments

We gratefully acknowledge Dr. Jonathan Gott for helpful comments and careful reading of the manuscript. The work was supported by NIH Grant GM56740 to M.E.H. N.H.Z is supported by NIH Training Grant GM008056.

References

[1] S. Altman, L. Kirsebom, Ribonuclease P, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999.
[2] J.C. Kurz, C.A. Fierke, Curr. Opin. Chem. Biol. 4 (2000) 553–558.
[3] R.M. Mans, C. Guerrier-Takada, S. Altman, C.W. Pleij, Nucleic Acids Res. 18 (1990) 3479–3487.
[4] K.A. Peck-Miller, S. Altman, J. Mol. Biol. 221 (1991) 1–5.
[5] Y. Komine, M. Kitabatake, T. Yokogawa, K. Nishikawa, H. Inokuchi, Proc. Natl. Acad. Sci. USA 91 (1994) 9223–9227.
[6] R.K. Hartmann, J. Heinrich, J. Schlegl, H. Schuster, Proc. Natl. Acad. Sci. USA 92 (1995) 5822–5826.
[7] P. Alifano, F. Rivellini, C. Piscitelli, C.M. Arraiano, C.B. Bruni, M.S. Carlomagno, Genes Dev. 8 (1994) 3021–3031.
[8] E.S. Haas, J.W. Brown, Nucleic Acids Res. 26 (1998) 4093–4099.
[9] T. Pan, M. Jakacka, EMBO J. 15 (1996) 2249–2255.
[10] A. Loria, T. Pan, RNA 2 (1996) 551–563.
[11] L.A. Kirsebom, S.G. Svard, EMBO J. 13 (1994) 4870–4876.
[12] T.E. LaGrandeur, A. Huttenhofer, H.F. Noller, N.R. Pace, EMBO J. 13 (1994) 3945–3952.
[13] B.K. Oh, N.R. Pace, Nucleic Acids Res. 22 (1994) 4087–4094.
[14] T. Pan, A. Loria, K. Zhong, Proc. Natl. Acad. Sci. USA 92 (1995) 12510–12514.
[15] A. Loria, T. Pan, Biochemistry 36 (1997) 6317–6325.
[16] A.B. Burgin, N.R. Pace, EMBO J. 9 (1990) 4111–4118.
[17] J.M. Nolan, D.H. Burke, N.R. Pace, Science 261 (1993) 762–765.
[18] W.H. McClain, C. Guerrier-Takada, S. Altman, Science 238 (1987) 527–530.
[19] D. Kahle, U. Wehmeyer, G. Krupp, EMBO J. 9 (1990) 1929–1937.
[20] D.L. Thurlow, D. Shilowski, T.L. Marsh, Nucleic Acids Res. 19 (1991) 885–891.
[21] R. Kole, M.F. Baer, B.C. Stark, S. Altman, Cell 19 (1980) 881–887.
[22] S. Altman, Adv. Enzymol. Relat. Areas Mol. Biol. 62 (1989) 1–36.
[23] N.R. Pace, J.W. Brown, J. Bacteriol. 177 (1995) 1919–1928.
[24] T. Stams, S. Niranjanakumari, C.A. Fierke, D.W. Christianson, Science 280 (1998) 752–755.
[25] C. Spitzfaden, N. Nicholson, J.J. Jones, S. Guth, R. Lehr, C.D. Prescott, L.A. Hegg, D.S. Eggleston, J. Mol. Biol. 295 (2000) 105–115.
[26] S. Niranjanakumari, T. Stams, S.M. Crary, D.W. Christianson, C.A. Fierke, Proc. Natl. Acad Sci. USA 95 (1998) 15212–15217.
[27] S.M. Sharkady, J.M. Nolan, Nucleic Acids Res. 29 (2001) 3848–3856.
[28] R. Biswas, D.W. Ledman, R.O. Fox, S. Altman, V. Gopalan, J. Mol. Biol. 296 (2000) 19–31.
[29] C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, Cell 35 (1983) 849–857.
[30] C. Reich, G.J. Olsen, B. Pace, N.R. Pace, Science 239 (1988) 178–181.
[31] J.C. Kurz, S. Niranjanakumari, C.A. Fierke, Biochemistry 37 (1998) 2393–2400.
[32] A. Vioque, J. Arnez, S. Altman, J. Mol. Biol. 202 (1988) 835–848.
[33] S.J. Talbot, S. Altman, Biochemistry 33 (1994) 1399–1405.
[34] E. Westhof, D. Wesolowski, S. Altman, J. Mol. Biol. 258 (1996) 600–613.
[35] A. Loria, T. Pan, Nucleic Acids Res. 29 (2001) 1892–1897.
[36] A. Loria, S. Niranjanakumari, C.A. Fierke, T. Pan, Biochemistry 37 (1998) 15466–15473.
[37] S.M. Crary, S. Niranjanakumari, C.A. Fierke, Biochemistry 37 (1998) 9409–9416.
[38] F. Houser-Scott, W.A. Ziehler, D.R. Engelke, Methods Enzymol. 342 (2001) 101–117.
[39] N. Jarrous, S. Altman, Methods Enzymol. 342 (2001) 93–100.
[40] J.A. Beebe, C.A. Fierke, Biochemistry 33 (1994) 10294–10304.
[41] J.A. Beebe, J.C. Kurz, C.A. Fierke, Biochemistry 35 (1996) 10493–10505.
[42] D. Smith, N.R. Pace, Biochemistry 32 (1993) 5273–5281.
[43] C.A. Fierke, G.G. Hammes, Methods Enzymol. 249 (1995) 3–37.
[44] I.A. Rose, Methods Enzymol. 64 (1980) 47–59.
[45] T. Pan, K. Zhong, Biochemistry 33 (1994) 14207–14212.
[46] W.D. Hardt, J. Schlegl, V.A. Erdmann, R.K. Hartmann, Nucleic Acids Res. 21 (1993) 3521–3527.
[47] A. Loria, T. Pan, Biochemistry 37 (1998) 10126–10133.
[48] W.D. Hardt, J. Schlegl, V.A. Erdmann, R.K. Hartmann, J. Mol. Biol. 247 (1995) 161–172.
[49] S.J. Talbot, S. Altman, Biochemistry 33 (1994) 1406–1411.
[50] D. Lane, P. Prentki, M. Chandler, Microbiol. Rev. 56 (1992) 509–528.
[51] S. Busch, L.A. Kirsebom, H. Notbohm, R.K. Hartmann, J. Mol. Biol. 299 (2000) 941–951.
[52] H.S. Penefsky, Methods Enzymol. 56 (1979) 527–530.
[53] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Plainview, NY, 1989.
[54] W.D. Hardt, V.A. Erdmann, R.K. Hartmann, RNA 2 (1996) 1189–1198.
[55] W.D. Hardt, J.M. Warnecke, V.A. Erdmann, R.K. Hartmann, EMBO J. 14 (1995) 2935–2944.
[56] C. Heide, S. Busch, R. Feltens, R.K. Hartmann, RNA 7 (2001) 553–564.
[57] C. Heide, R. Feltens, R.K. Hartmann, RNA 7 (2001) 958–968.
[58] C. Heide, T. Pfeiffer, J.M. Nolan, R.K. Hartmann, RNA 5 (1999) 102–116.
[59] D. Smith, A.B. Burgin, E.S. Haas, N.R. Pace, J. Biol. Chem. 267 (1992) 2429–2436.
[60] E.L. Christian, D.S. McPheeters, M.E. Harris, Biochemistry 37 (1998) 17618–17628.
[61] D.M. Lilley, T.J. Wilson, Curr. Opin. Chem. Biol. 4 (2000) 507–517.
[62] A.K. Knap, D. Wesolowski, S. Altman, Biochimie 72 (1990) 779–790.
[63] J. Ciesiolka, W.D. Hardt, J. Schlegl, V.A. Erdmann, R.K. Hartmann, Eur. J. Biochem. 219 (1994) 49–56.
[64] J. Kufel, L.A. Kirsebom, J. Mol. Biol. 263 (1996) 685–698.
[65] R.K. Gaur, A. Hanne, F. Conrad, D. Kahle, G. Krupp, RNA 2 (1996) 674–681.
[66] A. Loria, T. Pan, Biochemistry 38 (1999) 8612–8620.
[67] D.A. Pomeranz Krummel, O. Kent, A.M. MacMillan, S. Altman, J. Mol. Biol. 295 (2000) 1113–1118.
[68] C. Ehresmann, F. Baudin, M. Mougel, P. Romby, J.P. Ebel, B. Ehresmann, Nucleic Acids Res. 15 (1987) 9109–9128.
[69] L. Odell, V. Huang, M. Jakacka, T. Pan, Nucleic Acids Res. 26 (1998) 3717–3723.
[70] F. Liu, S. Altman, Cell 77 (1994) 1093–1100.
[71] J.F. Milligan, D.R. Groebe, G.W. Witherell, O.C. Uhlenbeck, Nucleic Acids Res. 15 (1987) 8783–8798.
[72] D. Moazed, S. Stern, H.F. Noller, J. Mol. Biol. 187 (1986) 399–416.
[73] D.W. Celander, T.R. Cech, Science 251 (1991) 401–407.
[74] J.A. Latham, T.R. Cech, Science 245 (1989) 276–282.

[75] R. Lavesry, A. Pullman, Biophys. Chem. 19 (1984) 171–181.
[76] E.L. Christian, M.E. Harris, Biochemistry 38 (1999) 12629–12638.
[77] C. Massire, L. Jaeger, E. Westhof, J. Mol. Biol. 279 (1998) 773–793.
[78] J.L. Chen, J.M. Nolan, M.E. Harris, N.R. Pace, EMBO J. 17 (1998) 1515–1525.
[79] C. Guerrier-Takada, N. Lumelsky, S. Altman, Science 246 (1989) 1578–1584.
[80] J. Kufel, L.A. Kirsebom, J. Mol. Biol 244 (1994) 511–521.
[81] M.E. Harris, E.L. Christian, Methods 18 (1999) 51–59.
[82] M.E. Harris, J.M. Nolan, A. Malhotra, J.W. Brown, S.C. Harvey, N.R. Pace, EMBO J. 13 (1994) 3953–3963.
[83] M.E. Harris, A.V. Kazantsev, J.L. Chen, N.R. Pace, RNA 3 (1997) 561–576.
[84] V. Gopalan, H. Kuhne, R. Biswas, H. Li, G.W. Brudvig, S. Altman, Biochemistry 38 (1999) 1705–1714.
[85] M.R. Ermacora, D.W. Ledman, H.W. Hellinga, G.W. Hsu, R.O. Fox, Biochemistry 33 (1994) 13625–13641.
[86] T.M. Rana, C.F. Meares, Proc. Natl. Acad. Sci. USA 88 (1991) 10578–10582.
[87] M.R. Ermacora, D.W. Ledman, R.O. Fox, Nat. Struct. Biol. 3 (1996) 59–66.
[88] K.B. Hall, R.O. Fox, Methods 18 (1999) 78–84.
[89] J.F. Milligan, O.C. Uhlenbeck, Methods Enzymol. 180 (1989) 51–62.
[90] M.J. Moore, P.A. Sharp, Science 256 (1992) 992–997.
[91] D.A. Pomeranz Krummel, S. Altman, RNA 5 (1999) 1021–1033.
[92] S.P. Ryder, S.A. Strobel, Methods 18 (1999) 38–50.
[93] R.K. Gaur, G. Krupp, Nucleic Acids Res. 21 (1993) 21–26.
[94] D. Siew, N.H. Zahler, A.G. Cassano, S.A. Strobel, M.E. Harris, Biochemistry 38 (1999) 1873–1883.
[95] C. Rox, R. Feltens, T. Pfeiffer, R.K. Hartmann, J. Mol. Biol. 315 (2002) 551–560.
[96] C. Guerrier-Takada, S. Altman, Proc. Natl. Acad. Sci. USA 89 (1992) 1266–1270.
[97] D.S. Waugh, N.R. Pace, FASEB J. 7 (1993) 188–195.
[98] J. Schlegl, W.D. Hardt, V.A. Erdmann, R.K. Hartmann, EMBO J. 13 (1994) 4863–4869.
[99] R. Sousa, R. Padilla, EMBO J. 14 (1995) 4609–4621.
[100] E.L. Christian, M. Yarus, J. Mol. Biol. 228 (1992) 743–758.
[101] W.D. Hardt, J.M. Warnecke, R.K. Hartmann, Mol. Biol. Rep. 22 (1995) 161–169.
[102] S.G. Svard, L.A. Kirsebom, Nucleic Acids Res. 21 (1993) 427–434.
[103] W.D. Hardt, R.K. Hartmann, J. Mol. Biol. 259 (1996) 422–433.
[104] V. Gopalan, A.D. Baxevanis, D. Landsman, S. Altman, J. Mol. Biol. 267 (1997) 818–829.
[105] A. Tallsjo, J. Kufel, L.A. Kirsebom, RNA 2 (1996) 299–307.
[106] E.L. Christian, N.M. Kaye, M.E. Harris, RNA 6 (2000) 511–519.
[107] S. Basu, S.A. Strobel, RNA 5 (1999) 1399–1407.
[108] C. Guerrier-Takada, W.H. McClain, S. Altman, Cell 38 (1984) 219–224.
[109] C.J. Green, B.S. Vold, J. Biol. Chem. 263 (1988) 652–657.
[110] C.K. Surratt, B.J. Carter, R.C. Payne, S.M. Hecht, J. Biol. Chem. 265 (1990) 22513–22519.
[111] S.G. Svard, U. Kagardt, L.A. Kirsebom, RNA 2 (1996) 463–472.
[112] E.V. Puglisi, J.D. Puglisi, J.R. Williamson, U.L. RajBhandary, Proc. Natl. Acad. Sci. USA 91 (1994) 11467–11471.
[113] C. Glemarec, J. Kufel, A. Foldesi, T. Maltseva, A. Sandstrom, L.A. Kirsebom, J. Chattopadhyaya, Nucleic Acids Res. 24 (1996) 2022–2035.
[114] S.G. Svard, L.A. Kirsebom, J. Mol. Biol. 227 (1992) 1019–1031.
[115] A.V. Kazantsev, N.R. Pace, RNA 4 (1998) 937–947.
[116] N.M. Kaye, E.L. Christian, M.E. Harris, Biochemistry 41 (2002) 4533–4545.
[117] A. Hansen, T. Pfeiffer, T. Zuleeg, S. Limmer, J. Ciesiolka, R. Feltens, R.K. Hartmann, Mol. Microbiol. 41 (2001) 131–143.
[118] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. 14 (1996) 33–38, 27–28.