Targeting RNA with Small Molecules

Jason R. Thomas, and Paul J. Hergenrother

Chem. Rev., **2008**, 108 (4), 1171-1224 • DOI: 10.1021/cr0681546

CHEMICAL REVIEWS

Volume 108, Number 4

Targeting RNA with Small Molecules

Jason R. Thomas and Paul J. Hergenrother*

Department of Chemistry, Roger Adams Laboratory, University of Illinois, Urbana, Illinois 61822

Received September 7, 2007

Contents

* To whom correspondence should be addressed. E-mail: hergenro@uiuc.edu.

10.1021/cr0681546 CCC: $71.00
Published on Web 03/25/2008

1. Introduction

The recently discovered functions of RNA dramatically expand the cellular roles for this macromolecule. RNA can no longer be viewed as a passive element, working simply as an intermediary between genomic information and the primary sequence of proteins. Rather, it is now recognized that RNA is essential for transcriptional regulation,[1] translational regulation,[2,3] protein function,[4,5] and catalysis,[6] responsibilities that have classically been reserved for proteins. This recent explosion in RNA biology underscores the importance of RNA in normal and aberrant cellular functions and highlights the potential of targeting RNA for the treatment of a multitude of disease states.

Given the diversity of RNA function, small molecules that selectively bind to RNA may provide novel points of therapeutic intervention. In fact, it is becoming increasingly clear that for certain pathways and diseases RNA targets may be the best and only option due the intractability of targeting certain segments of the proteome and the inherent difficulty of targeting DNA. The percentage of proteins that are considered "drugable" is a matter of some debate. Although the predicted value varies between investigations,[7−14] a recent study concluded that only 207 proteins encoded within the human genome are targeted by the current FDA-approved small molecule drugs.[14] These current protein targets are only a small subset of the 1620 proteins directly linked to genetic disease and a minute portion of the hundreds of thousands (including post-translational modifications and splice variants) of human proteins.[14] Furthermore, of the 207 protein targets >50% are class I GPCRs, nuclear receptors, or ion channels.[14] Compounding issues further, computational surveys suggest that as little as 15% of proteins within the proteome may have a suitable binding site for drug-like compounds,[9] and expanding drug discovery to the disruption of protein−protein interactions has proven to be a formidable challenge.[15] Thus, even though protein targets are (and will continue to be) the focus of drug discovery, alternative strategies to complement traditional protein-based targets are needed.

Dervan and co-workers developed an elegant paradigm of targeting DNA with small molecules that exhibit exquisite sequence specificity.[16] Drawing inspiration from DNA-binding small molecule natural products, a "modular code"[17] has been developed wherein the unique pairing of pyrrole, imidazole, and hydroxypyrrole rings are used to selectively recognize all combinations of Watson−Crick base pairs. By

Jason R. Thomas was born in Savannah, GA, on January 10, 1980. He attended Temple University, where he received his B.S. degree in Biophysics in 2002. He then began his graduate career in the laboratory of Professor Paul J. Hergenrother at the University of Illinois at Urbana−Champaign. In his graduate work, Jason developed a novel strategy to combat multidrug-resistant bacteria using small molecules to induce the elimination of plasmids that carry resistance-mediating genes. These compounds act by targeting a specific regulatory element in the mRNA of an essential protein for plasmid replication. He also developed small molecule ligands with specificity and affinity for RNA hairpin loops. After receiving his Ph.D. degree in July 2007 he joined the laboratory of Professor Benjamin F. Cravatt at the Scripps Research Institute in August 2007 and is currently an American Cancer Society Postdoctoral Fellow.

Paul J. Hergenrother was born in 1972 and raised in Akron, OH. He attended the University of Notre Dame, where he received his B.S. degree in Chemistry in 1994. From there he moved to the University of Texas at Austin to conduct graduate research under the direction of Professor Stephen F. Martin. While in the Martin laboratory Paul elucidated the catalytic mechanism of phospholipase C and completed the total synthesis of erythromycin B. He graduated with his Ph.D. degree in Chemistry in 1999 and moved on as an American Cancer Society postdoctoral fellow to Harvard University, where he worked in the laboratory of Professor Stuart L. Schreiber in the Department of Chemistry and Chemical Biology. While at Harvard he was involved in the development of small molecule microarrays as a platform for high-throughput compound screening. He established his own laboratory in the Department of Chemistry at the University of Illinois at Urbana−Champaign in 2001. During his Assistant Professor years he was the recipient of an NSF-CAREER Award, a Research Corporation Research Innovation Award, a Beckman Young Investigator Award, an Alfred P. Sloan Foundation Fellowship, the GlaxoSmithKline Chemistry Scholar Award, the ACS David Robertson Award for Excellence in Medicinal Chemistry, and the Camille Dreyfus Teacher-Scholar Award and named by *Technology Review* magazine as one of the top innovators under the age of 35. He was promoted to Associate Professor with tenure in 2006, and since then has been the recipient of the ACS Eli Lilly Award in Biological Chemistry. The Hergenrother laboratory seeks to use small molecules to identify and validate novel targets for the treatment of intractable diseases, including cancer, neurodegeneration, and multidrug-resistant bacteria. Through their work on novel antibacterials, the Hergenrother laboratory became interested in developing a general strategy for targeting RNA with small molecules.

targeting 6−16 DNA base pair segments (with binding affinities typically in the low- to subnanomolar range), these sequence-specific polyamides inhibit or induce transcription of specific genes in vitro and in cell culture.[17,18] As a general tool for chemical biology, the applicability of these sequence-specific polyamides in cell culture is case specific as their nuclear permeability is cell-type dependent.[17,19,20]

While drug−protein interactions are entrenched in the psyche of medicinal chemists and drug−DNA interactions are smaller in number but have multiple success stories, RNA has long been neglected as a primary drug target. From a targeting standpoint, RNA exhibits many attractive features similar to those possessed by DNA and/or proteins. Like DNA, the chemical building blocks of RNA are fewer and less complex than those of their protein counterparts. Both forms of nucleic acids adopt a regular helix featuring a major and minor groove. However, the regular A-form helix of RNA is often disrupted by regions of mismatched or unpaired bases that allow RNA to adopt more complex three-dimensional structures, analogous to those observed in proteins. Such structures give rise to defined pockets suitable for binding to other RNAs,[21,22] proteins,[23] and small molecule metabolites[24] in addition to allowing for intramolecular interactions within the transcript to form tertiary structure.[25] Further discussion regarding the structural properties shared by RNA with DNA and/or proteins is presented in section 3.

Some may contend, or even take for granted, that the drugability of RNA has already been demonstrated by various classes of antibiotics which bind to defined regions of the prokaryotic ribosome. Indeed, the aminoglycoside,[26,27] macrolide,[28] tetracycline,[29] and oxazolidinone[30] antibiotics exert their antibacterial effects by binding to ribosomal RNA (rRNA). While these drugs have conclusively demonstrated that rRNA is an excellent antibacterial target, there are caveats that temper the enthusiasm for the applicability of these results to general small molecule−RNA binding. The ribosome, unlike other forms of RNA, represents a catalytically competent and abundant form of cellular RNA. The "holy grail" in small molecule−RNA binding would be the selective targeting of *a single cellular RNA* (such as an mRNA transcript, a regulatory RNA, etc.), resulting in functional perturbation of a specific cellular process. Obviously, low copy, noncatalytically active transcripts (for example) present new targeting challenges relative to the ribosome. Thus, while the demonstrated ability of RNA-binding antibiotics to achieve their desired in vivo effect is certainly a motivation to attempt the targeting of novel RNAs, it does not guarantee success.

What then is the state of the small molecule−RNA binding field? While progress has been made in the last 10 years since a major *Chemical Reviews* article on this topic,[31] researchers in the field are very far from being able to design a ligand for an RNA simply based on knowledge of a RNA target sequence; we are nowhere near the development of an RNA-targeting paradigm akin to the Dervan polyamide−DNA targeting rules. Most small molecule ligands for RNA have only modest affinity and selectivity for their target. The aminoglycosides are still typically screened when one is searching for a lead ligand for a new target RNA, and relatively few compounds demonstrate any efficacy in cell culture, let alone in animal models. The goal of this review is to present a summary of the current state of targeting RNA with small molecules. First, an overview is given of RNA structure, RNA targets, and the methods for studying RNA−ligand interactions. Then, a comprehensive review of efforts to target the various classes of RNA secondary structures is presented. Finally, some challenges for the future are defined. As this review focuses on small molecules, the targeting of RNA with nucleotides or peptide−nucleic acid conjugates is not covered.[32−34]

2. Contribution of Secondary and Tertiary Structure to RNA Folding

RNA folding follows a hierarchical pathway analogous to that observed for proteins. The primary sequence dictates the type of secondary structure formed, which in turn leads to formation of possible tertiary structure via interaction of preformed secondary structures. Previously, RNA tertiary structure has been described as a form of "super-secondary structure", although this terminology is no longer used.[35,36] Formation of RNA secondary structure dominates the free energy of folding, as each base pair contributes 1−3 kcal/mol of free energy to the final fold.[36,37] For example, tRNAs have a uniquely evolved tertiary structure; the primary sequence of a tRNA dictates formation of a "clover leaf" secondary structure composed of three stem-loop segments. However, the well-known three-dimensional structure of tRNAs is finalized by the interaction between two of the hairpin loops (the T and C loops). This last step, formation of tertiary structure, contributes only ∼1.5 kcal/mol of free energy.[38]

Those concerned with RNA−ligand interactions, whether small molecule or protein, generally give greater weight to secondary structure, as ligand binding sites typically consist of a single type of secondary structure. Thus, from a targeting perspective, secondary structure becomes the key determinant in defining the drugability of a particular RNA. Furthermore, the notion of disrupting tertiary structure with a small molecule binder is ultimately a question of how to prevent two secondary structures from interacting. As different RNAs from the transcriptome present a multitude of secondary structures, selectivity between RNA targets can be envisioned through recognition of a specific combination of secondary structures.

3. RNA Secondary Structure Creates Binding Pockets for Small Molecules

Which sites within RNA can be targeted? Extensive work in this area has suggested that regions where there is a perturbation of the A-form helix are optimal for RNA targeting. Such perturbations create various classes of secondary structures, such as hairpin loops, internal loops, and bulged regions (Figure 1). Unlike the major and minor grooves of DNA, which permit the binding of small molecules, the major and minor grooves of RNA do not afford optimal binding sites for small molecules (Figure 2).[39] The major groove of RNA is deep and narrow, while the minor groove is shallow; as such, the unique geometry of the fully base-paired A-form helix is less conducive to small molecule binding. It should be noted that in a formal sense the A-form helix does not have "major" and "minor" grooves but rather deep and shallow grooves, although in practice the terms are used interchangeably. The presence of the 2′-hydroxyl on RNA bases results in an alternative puckering of the ribose unit as compared to that observed in DNA duplex regions, resulting in a concomitant change in the helix pitch and tilt of the bases. The major groove of RNA is information rich as all the base pairs project their discriminatory edges into the major groove.[39,40] Also, positively charged

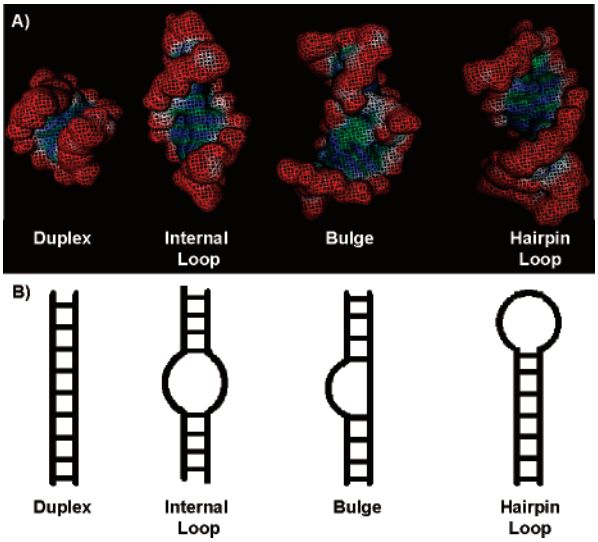


Figure 1. Four general classes of RNA secondary structure. (A) Surface representations of individual secondary structures. Note that the presence of un- or mispaired nucleotides alters the accessibility of the major groove. (B) Schematic representation of RNA duplex, internal loop, bulge, and hairpin loop (also called stem loop) regions of RNA that will be discussed throughout this review.

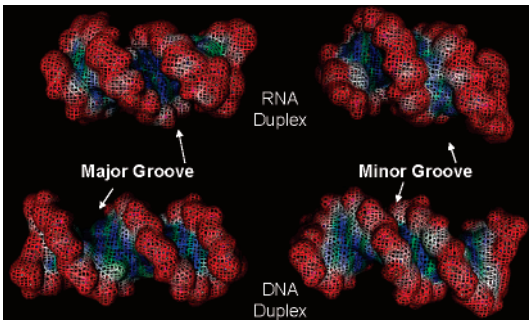


Figure 2. Groove accessibility of RNA and DNA. Surface representations of A-form (top) and B-form (bottom) duplexes. As compared to the DNA major groove, the corresponding groove in an RNA duplex is significantly narrower and deeper. Also, the 'minor groove' of an RNA duplex is shallower as compared to the DNA minor groove. These differences in groove structure help to account for the relative ease by which small molecules can bind a DNA duplex over an RNA duplex.

small molecules should preferentially bind the major groove as computational studies suggest that it is more electronegative than the minor groove.[41] However, some studies have suggested that in the fully base-paired form the major groove of the A-form helix is only 4 Å wide, which sterically precludes small molecule binding.[39] Thus, the deep and narrow major groove of the A-form helix is less accessible to small molecules, and because all of the base pairs project into the major groove, the only defining feature of the shallow minor groove is the constant presence of 2′-hydroxyl groups.[42] It is important to note that the fully base-paired RNA stem does not provide a completely inhospitable environment for small molecule ligands; some ligands do bind RNA duplexes in the major groove.[43−45]

Perturbations of the A-form helix induced by un- or mispaired bases widen the major groove to provide a surface-exposed binding pocket (see Figure 1A). The binding sites formed by such mismatched regions in large part provide the basis for targeting RNA by proteins and small molecules. It has been suggested that the electrostatic surface potentials of the RNAs more closely resemble that of proteins rather than that of DNA.[41] The major difference is that a binding pocket of a protein is likely to present a contour of positive and negative potentials. The analogous pockets found in RNA only present varying degrees of electronegative potentials; that is, the various folds adopted by the different secondary structures offer distinct binding sites that are further characterized by their own electronegative landscape. The secondary structures targeted by small molecules will be discussed in detail throughout this review.

4. Molecular RNA Targets

The targeting of RNA with small molecules has the potential to offer a complementary approach to the targeting of proteins. As mentioned in the Introduction, many newly discovered functions of RNA are regulatory mechanisms in which proteins do not participate. Furthermore, the RNA component of RNA−protein complexes is often essential for their function.[6,46] Thus far, small molecule ligands for RNA have been developed with three major classes of targets in mind: antibacterial, antiviral, and mRNA. Within each of these classes of RNA targets various avenues have been pursued to achieve the desired biological effect, including inhibiting RNA−protein interactions and preventing protein production by binding to a particular mRNA. This section will detail the various RNAs that have been targeted by small molecules to date as well as a few that have not been targeted but appear to be excellent candidates. It should be noted that a great majority of these RNA-binding compounds have been shown only to bind to their RNA target in vitro, and efficacy in cell culture or in vivo has rarely been demonstrated.

4.1. Antibacterial Targets: The Ribosome

The importance of RNA-binding small molecules as antibiotics is unquestioned. To date all clinically approved RNA-targeting drugs exert their effect by binding rRNA. As the topic of antibiotics that bind the ribosome has been covered extensively by many excellent reviews,[47−50] it will only be discussed briefly here. The prokaryote ribosome is formed by two subunits, the small (30S) and large (50S), which are composed of rRNA and ribosomal proteins. The small subunit is responsible for ensuring that only the correct tRNA is allowed to incorporate its amino acid into the growing polypeptide chain. The 50S subunit is responsible for the process of peptide elongation through catalysis of peptide bond formation. The functions performed by each subunit, proof-reading by the 30S and catalysis by the 50S subunits, are done so only with rRNA; no ribosomal proteins directly participate in either process.

Proofreading occurs at the 16S rRNA A site, part of the 30S subunit. The binding of a cognate tRNA induces residues A1492 and A1493 to adopt an extrahelical conformation, while near- or miscognate tRNAs fail to induce the same conformational change (Figure 3).[51] While several classes of antibiotics have been found to bind to the 30S subunit, the aminoglycosides are the most intensely investigated class of RNA-binding small molecules. The aminoglycosides allow incorporation of amino acids from noncognate tRNAs by binding to the A site; upon binding, the 4,5- (Figure 4A) and 4,6-substituted (Figure 4B) deoxystreptamine aminoglycosides cause residues A1492 and A1493 to protrude from the helix in the same fashion as the binding of a cognate tRNA (see Figures 3 and 4C).[52−54] On the basis of the ligand-induced conformational changes, it has been proposed that aminoglycosides lower the energy barrier required to incorporate amino acids from miscognate tRNAs.[51,55]

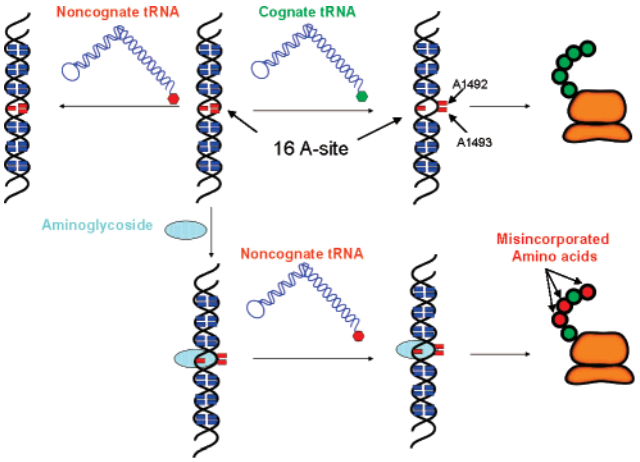


Figure 3. Proofreading by the 16S A-site rRNA. The binding of a cognate tRNA to the ribosome is sensed at the A site by a conformational change in residues A1492 and A1493, which then adopt an extrahelical conformation. Noncognate tRNAs fail to induce the same conformational change in the A site. However, the binding of the aminoglycosides to the A site mimics the binding of a cognate tRNA by inducing a similar conformational change of A1492 and A1493. Thus, in the presence of aminoglycosides noncognate tRNAs are able to incorporate their amino acids into the growing polypeptide chain which ultimately leads to the antibacterial effect.

The oxazolidinones are among the newest classes of antibiotics with linezolid (**1**, Figure 5) representing the first of this class to reach the clinic. Although their exact mechanism of action is currently under debate, a large body of evidence is accumulating that suggests that the oxazolidinones exert their antibacterial effect by binding to the 23S rRNA present within the 50S subunit.[30] For example, the oxazolidinones have been shown to compete with chloramphenicol and lincomycin for binding to the 50S subunit.[56] Also, clinical bacterial isolates that exhibit a linezolid-resistant phenotype acquire mutations within the 23S rRNA, suggesting that linezolid and presumably other oxazolidinones bind to the 23S rRNA.[30] Curiously, various oxazolidinones bind to isolated ribosomes with weak affinity[57] (Figure 5) despite their low micromolar IC_{50} values observed in cell-free in vitro translation assays.[58,59] Recently, it was shown that the oxazolidinones prevent the initiator tRNA from binding to the P site within the 50S ribosomal subunit.[60] Thus, it appears that the oxazolidinones bind near the P site to prevent formation of the first peptide bond, which may rectify the apparent discrepancies between the observed binding affinities for isolated ribosomal subunits and potencies observed for cell-free in vitro translation assays.

4.2. Antibacterial Targets: tRNA

Inhibition of aminoacyl-tRNA synthetases (AaRs) has attracted considerable attention as an antibacterial strategy.[61] AaRs are essential enzymes that are responsible for the coupling ("charging") of amino acids to their cognate tRNAs;[62] only after being charged can a tRNA participate in the initiation and elongation process of protein synthesis. Accordingly, inhibition of AaRs has been demonstrated to be an effective antibacterial strategy as the buildup of uncharged tRNAs leads to inhibition of protein synthesis and ultimately bacterial cell death.[61]

A complementary strategy to inhibiting the aminoacylation reaction is preventing the interaction of AaRs with their respective tRNAs; for this purpose either the AaRs or the tRNA may be targeted. Several aminoglycosides have been shown to bind to $tRNA^{Phe}$;[63,64] from these, neomycin was found to inhibit the lead-mediated cleavage of yeast $tRNA^{Phe}$ with a K_i of ~300 μM (Figure 6A). Neomycin was also equally effective at inhibiting the charging of $tRNA^{Phe}$, K_i of ~300 μM.[63] Other compounds, including non-aminoglycoside small molecules, have been shown to inhibit the charging of tRNAs, although the concentrations of ligand required to do so are often excessively high.[65,66] Interestingly, a portion of the binding site for neomycin on the $tRNA^{Phe}$ is composed of nucleotides located in the variable loop region (Figure 6B); this may suggest that specificity for tRNAs may be gained by the selective targeting of the specific variable loops located on distinct tRNAs.

4.3. Antibacterial Targets: T Box

Expression of a significant percentage of AaRs and other proteins involved in amino acid biosynthesis and transport in Gram-positive bacteria involves a regulatory RNA element, known as a T box, located in their 5′-untranslated region (UTR).[67,68] The term "T box" refers to a 14 nucleotide conserved sequence that is involved in the binding of specific noncharged tRNAs;[69] the T box is able to adopt alternative conformations that dictate whether an antiterminator or terminator stem-loop is formed (Figure 7). Uncharged tRNAs positively regulate the expression of T-box-containing genes for their specific amino acid; that is, uncharged tRNAs directly induce expression of proteins responsible for synthesis of the deficient amino acid. The interaction between tRNAs, charged or uncharged, and the 5′-UTR occurs through the pairing of the anticodon loop of the tRNA with a sequence located in a stem-loop structure at the 5′-portion of the leader sequence (Figure 7). Amino acid availability, as assessed by the ratio of charged to uncharged tRNAs, is determined at the distal portion of the 5′-UTR. As illustrated in Figure 7, the binding of the 3′-end of an uncharged tRNA and the T box promotes the stabilization of the antiterminator structure, which allows translation to proceed. The sequence and secondary structure of the antiterminator stem-loops is highly conserved, containing a seven nucleotide bulged region interrupting two helical segments;[70] the interaction between the 3′-acceptor end of the uncharged tRNA and the bulged region of the antiterminator proceeds sequence specifically by formation of four base pairs between the two RNAs.[71] Under conditions of low amino acid availability, the direct regulation by uncharged tRNAs causes an increase in proteins responsible for biosynthesis of the deficient amino acid. Thus, inhibition of antiterminator T-box−tRNA interaction has been proposed as an antibacterial target.[72]

Molecular insights into the antiterminator structure have been gained from NMR experiments using a model construct of the T box from *B. subtilius*.[73] The overall shape is that of two A-form helices, termed A1 and A2, oriented at ~80° to each other; this abrupt change in direction is caused by a seven nucleotide bulge (Figure 8). The NMR data reveal that the bulged region contains a combination of static and high-mobility residues. The overall root-mean-square deviation of the bulged region was found to be 2.90 Å, suggesting a conformationally unrestrained bulged region; however, the residues responsible for initiating base recognition with the 3′ acceptor portion of the uncharged tRNA account for the bulk of the conformational freedom.[73] The remaining bulged nucleotides exhibit a significant degree of conformational restraint likely caused by extensive stacking interactions.

A collection of aminoglycosides was screened for their ability to bind the T-box antiterminator RNA.[72] The binding

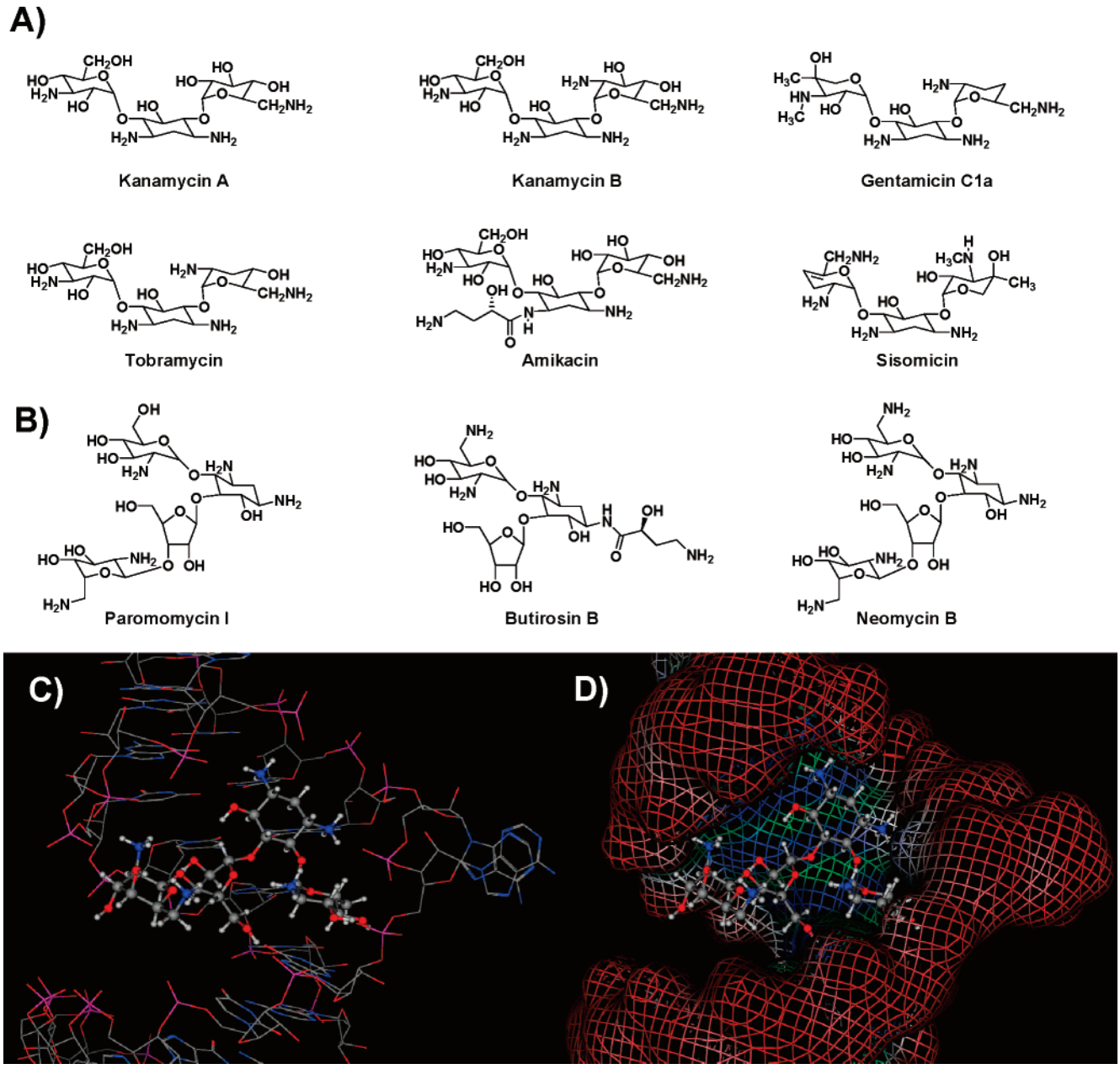


Figure 4. Aminoglycoside−16S A-site interaction. (A) Structure of various 4,6-disubstituted 2-deoxystreptamine aminoglycosides. (B) Structure of various 4,5-disubstituted 2-deoxystreptamine aminoglycosides. (C) Crystal structure of paromomycin bound to the A-site RNA. The binding of paromomycin and various other aminoglycosides induces an extrahelical conformation of A1492 and A1493. This same conformation of the adenines is observed upon binding of a cognate tRNA. (D) Surface representation of the paromomycin−A-site complex.

affinities for various aminoglycosides was determined with dissociation constants ranging from 8.5 (neomycin) to 790 μM (streptomycin); the mid- to high-micromolar affinity of the aminoglycosides for the T-box antiterminator RNA is comparable with tRNAs binding to the bulged region.[71] Future directions for this strategy have included identification of different small molecule ligands.[74] However, determining whether small molecule binders to the bulged region could stabilize the antiterminator structure, which may result in the activation of T-box genes, will need to be addressed.

4.4. Antiviral Targets: Trans-Activating Response RNA

The HIV-1 genome consists of two copies of the ~9 kB genomic RNA.[75] After successful integration into the host cell genome, the initial phase of the viral life cycle begins with transcription from the 5′-end of the genome. Although this step is crucial, RNA polymerase II transcribes poorly from the inefficient viral promoters. The HIV-1 protein Tat functions as an adaptor protein to facilitate efficient transcription; in the absence of Tat only short, non-polyadenylated transcripts are produced, while in the presence of Tat the rate of transcription increases nearly 100-fold to produce the full genomic-length RNA (Figure 9).[75] Because of such stark enhancements in the rate of transcription, Tat has been the subject of intense investigation. Unlike typical transcriptional activators which bind DNA, Tat binds to a specific bulged RNA hairpin loop located at the beginning of viral transcripts. Tat is able to recognize this RNA structure, known as TAR (trans-activating response element), by binding to the bulged region of the RNA.[76,77] After formation of the Tat−TAR complex, cyclin T1 binds to the hairpin loop region of TAR, which further enhances transcription.[78] The exact order of events in formation of Tat−TAR−cyclin

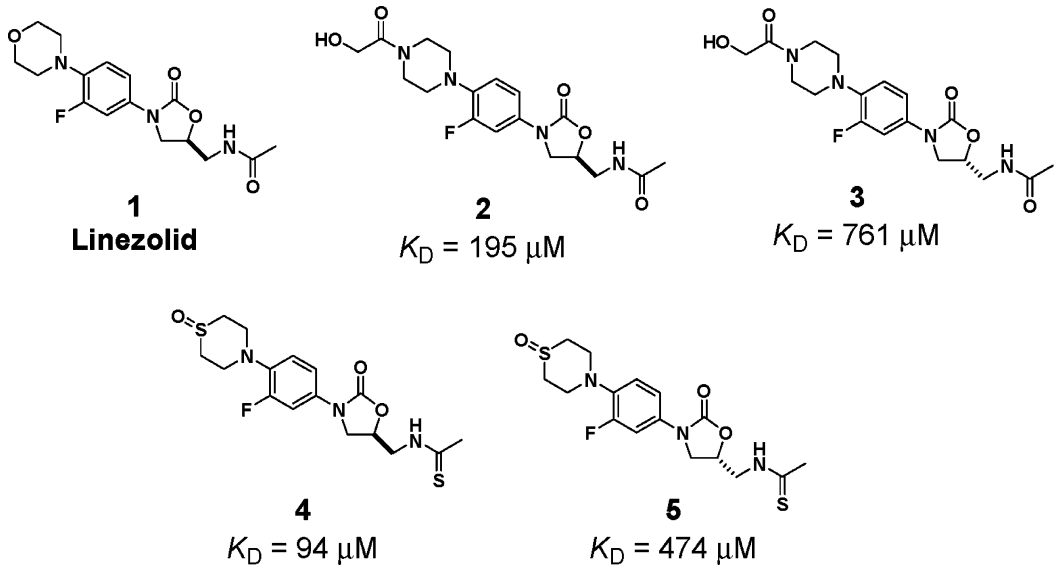


Figure 5. Structures of various oxazolidinone antibiotics and their binding affinity for isolated ribosomes. The binding affinity for linezolid has not been reported.

A)

Aminoglycoside	K_i (μM)
5-*epi* Sisomicin	2.5
Hygromycin	267
Neomycin	342
Sisomicin	440
Paromomycin	2000
Kanamycin A	5000
Tobramycin	8000
Kanamycin B	11,200
Lividomycin	18,000
Ribostamycin	23,300

B)

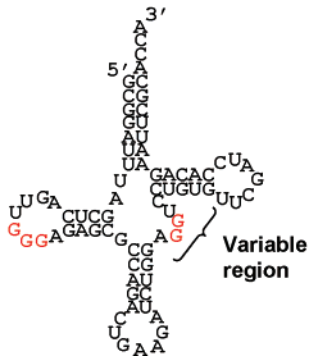


Figure 6. Interaction between tRNAPhe and aminoglycosides. (A) The ability of various aminoglycosides to inhibit lead-mediated cleavage of the yeast tRNAPhe. (B) Secondary structure representation of tRNAPhe. Residues shown in red are those that are involved in the binding of neomycin.

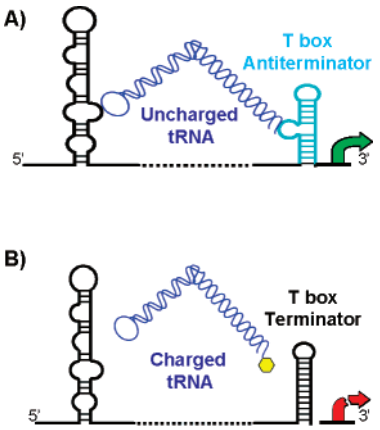
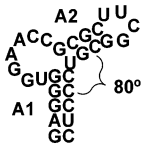


Figure 7. T-box regulation of amino acid biosynthetic operons. Certain tRNAs regulate the synthesis of their cognate amino acids by controlling the translation of proteins responsible for production of the amino acid building block. (A) Under conditions of low amino acid availability, uncharged tRNAs (blue) bind to their specifier sequence through the anticodon loop and the 3′-tail of the tRNA; this binding to the T-box element (cyan) leads to an increase in translation. (B) Charged tRNAs (the charge state is represented by the yellow hexagon) are prevented from binding to the T-box element, allowing formation of a terminator hairpin and thus inhibiting translation.



Aminoglycoside	K_D (μM)
Neomycin B	8.5
Tobramycin	38
Paromomycin	50
Gentamicin C	120
Kanamycin B	150
Kanamycin A	210
Amikacin	760
Streptomycin	790

Figure 8. T-box antiterminator RNA−aminoglycoside interaction. Shown is a secondary structure rendering of the model antiterminator T-box RNA from *B. subtilis* and the binding affinity of various aminoglycosides for this construct.

T1 is debatable as the preformed Tat−cyclin T1 complex binds with greater affinity than Tat alone.[79]

The Tat−TAR interaction is mediated by an arginine-rich segment of Tat (residues 49−57) that binds to the bulged region of TAR.[80] The bulged region of TAR widens the A-form helix to permit Tat binding; the structural recognition elements required for Tat binding are found within the major groove.[81−83] The Tat−TAR interaction proceeds through a two-step process.[84] First, the initial recognition of a single arginine side chain prompts the remodeling of the TAR major groove. Use of a peptide mimetic, argininamide (**6**, Figure 10A), has greatly facilitated the understanding of this early recognition event. Upon ligand binding, argininamide is thought to bind to the Hoogsteen face of G26 and form stacking interactions with U23 (Figure 10A and B). The binding-induced conformational changes reposition the phosphate

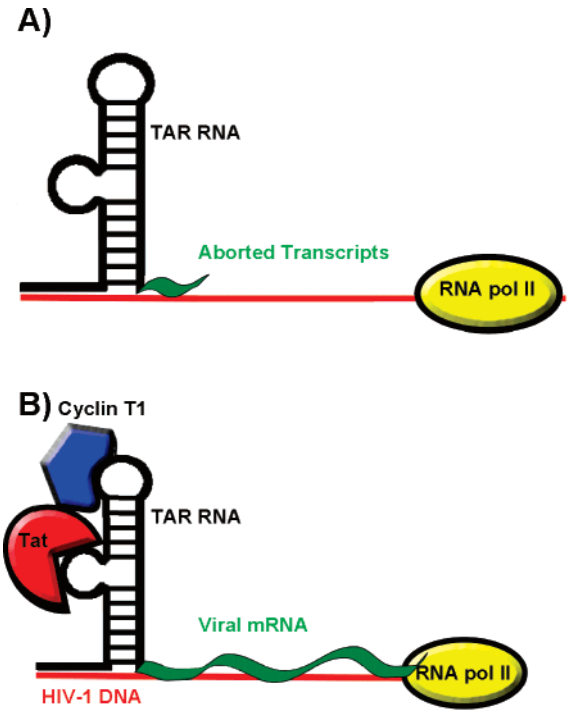


Figure 9. Tat−TAR interaction increases the processivity of RNA polymerase II. (A) In the absence of Tat, RNA pol II (yellow) synthesizes short, non-polyadenylated RNAs (green). (B) The binding of Tat (red) with TAR recruits other proteins, including cyclin T1 (blue), which act as a complex to enhance the efficiency of RNA pol II.

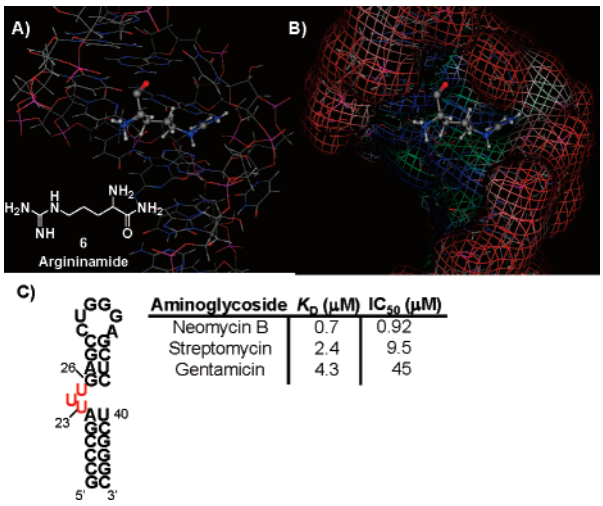


Figure 10. NMR-determined structure of the argininamide−TAR complex. (A) The guanidinium moiety of argininamide disrupts the stacking interaction of A22 and U23. (B) Surface representation of the argininamide−TAR complex. (C) Secondary structure representation of TAR and the binding affinities of various aminoglycosides for this RNA as well as the IC_{50} values for disruption of the Tat−TAR interaction. The bulged nucleotides involved in mediating the Tat−TAR interaction are shown in red.

groups of A22, U23, and U40 into the major groove to form critical contacts with Tat; these ionic interactions have been shown to be important in achieving high affinity, but it has also been suggested that they account for the specificity of the Tat protein for TAR over other bulged RNAs.[85]

Inhibition of the Tat−TAR interaction by binding to TAR has been sought as a potential anti-HIV strategy. Initially a collection of aminoglycosides was screened for their ability to disrupt the Tat−TAR interaction; neomycin emerged as

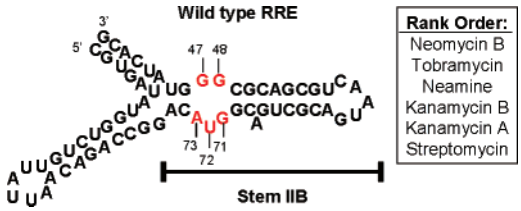


Figure 11. Rev response element (RRE) RNA. Shown is the secondary structure model of the wild-type RRE sequence; the residues in red are the high-affinity binding sites for Rev. Also shown is the rank order (in terms of binding affinity) of various aminoglycosides for RRE, as determined by SPR.

the most potent aminoglycoside, with an IC_{50} of 0.92 μM (Figure 10C).[86] Neomycin binds TAR with a 1:1 stoichiometry as determined by ESI-MS, and in subsequent biochemical experiments the authors demonstrated that the binding site for neomycin is located below the bulged region. This binding site location is consistent with previous data wherein a TAR construct without the bulged nucleotides exhibited nearly identical binding affinities for neomycin.[87] Replacement of the guanines with inosines in the lower part the stem, but not the upper part of the stem, resulted in substantial loss of binding affinity for neomycin, suggesting that neomycin exerts its inhibitory effects by binding the base-paired region below the bulged nucleotides through the minor groove.[88] Attempts to monitor disruption of Tat−TAR complex by ESI-MS resulted in a ternary complex between the three components; this result, coupled with kinetic data, suggested that the neomycin inhibits Tat by a noncompetitive mechanism.[88] From analysis data it has been suggested that neomycin binds to TAR and alters the conformation to increase the off rate of the Tat−TAR complex. Subsequent NMR determination of the TAR−neomycin complex has confirmed the binding site and suggests a molecular rationale for the noncompetitive inhibition.[89] The binding of neomycin to the minor groove induces a conformational rearrangement which repositions key functionalities away from the major groove, thus decreasing the lifetime of the Tat−TAR interaction.

4.5. Antiviral Targets: Rev Response Element RNA

After initiation of transcription of the HIV-1 viral genome, proper splicing and export must occur in order for viral maturation to proceed. The Rev protein is responsible for export of viral RNAs from the host nucleus into the cytoplasm.[75] Rev accomplishes this task by binding to a highly structured RNA segment found in the *env* coding region. The RNA to which Rev binds is called the Rev response element (RRE) RNA and is composed of a series of stem-loop structures (see Figure 11). Within RRE, Rev binds with high affinity to a stem-loop structure termed stem IIB.[90,91] In response to RRE binding Rev oligomerizes within this region of RNA, nucleating from the high-affinity site. This nucleation event facilitates the export of the viral RNA (as Rev contains a nuclear export signal in its C-terminal half) and protects the RNA against further splicing. Thus, it has been proposed that inhibition the Rev−RRE interaction may successfully prevent viral replication and maturation.[92−94]

The interaction between Rev and RRE is mediated by an arginine-rich motif of Rev binding to the asymmetric internal loop of RRE.[95] The overall structure of RRE reveals that the helical segments above and below the internal loop are oriented at a 30° angle to each other.[96] Nucleotides U72 and

A68 maintain extrahelical conformations. Exclusion of U72 from the internal loop allows the remaining residues to form noncanonical base pairs. G47 and A73 pair via their Watson−Crick faces, while the Watson−Crick face of G48 pairs with the Hoogsteen face of G71.[96] Additionally, these noncanonical base pairs are extensively hydrated with a total of seven water molecules bound between them.[96] Upon binding the arginine-rich motif of Rev, both RRE and Rev undergo substantial conformational changes; the position of U72 adjusts to expand the major groove by 5 Å, while Rev adopts a helical conformation. Rev binds deep within the internal loop, allowing the charged arginine residues to dominate the Rev−RRE interaction.[97]

The ability of various aminoglycosides to disrupt the RRE−Rev interaction was first demonstrated by Green and co-workers.[94] After this initial work, neomycin was shown to bind to RRE ($K_D \approx 100$ nM) and disrupt the Rev−RRE interaction with an IC_{50} value of ~ 1 μM.[98−100] Extensive mutational analysis of the residues in and around the RRE internal loop has demonstrated the requirement of a perturbed A-form helix for neomycin binding.[101] Multiple modes of binding analysis have revealed that neomycin binds with a stoichiometry of 3:1.[99,100] Kinetic data coupled with NMR studies have provided a model for neomycin disruption of the Rev−RRE interaction.[99] Initially, neomycin binds to its high-affinity site ($K_D \approx 100$ nM) located in the duplex below the internal loop region. The binding of neomycin to its high-affinity site causes no structural perturbations of the Rev−RRE complex; in fact, in the NMR experiments a ternary complex between neomycin, Rev, and RRE was observed.[99] At higher concentrations neomycin binds to a weaker binding site ($K_D = 1.9$ μM) that overlaps with the binding site of Rev in the internal loop region of RRE and competitively inhibits the Rev−RRE complex. Further titration of neomycin results in a third binding event ($K_D = 41$ μM) which is attributed to nonspecific binding.

4.6. Ribozyme Targets

The discovery of RNAs that form complex 3D folds capable of self-cleavage and splicing has captured the attention of multiple disciplines. These catalytic RNAs, called ribozymes, are believed to provide a window back to the primordial world where RNA may have functioned as both the genetic element and the enzymatic workhorse.[102] The catalytic activity of the ribozymes enables facile analysis of ribozyme function. In addition, the availability of several ribozyme X-ray structures allows one to begin correlating function with structure. Thus, the ribozyme provides a convenient framework from which to begin understanding RNA−ligand interactions.

Aminoglycoside-mediated inhibition of ribozymes was first documented for the group I introns;[103] since this initial discovery various aminoglycosides and other RNA binding antibiotics have been documented to inhibit hammerhead ribozymes,[104] the RNase P ribozymes,[105] and the hepatitis delta virus ribozymes.[106] In each of these cases neomycin was identified as the most potent inhibitor with K_i values ranging from 0.5−28 μM.[103−106] The binding of neomycin to the respective ribozymes was found to be competitive with Mg^{2+} ion binding, as addition of Mg^{2+} was found to reduce the K_i of neomycin. Within the catalytic core of the hammerhead ribozymes there are five Mg^{2+} ions that are responsible for its catalytic activity. The binding of neomycin to this catalytic core is believed to cause the displacement of the Mg^{2+} ions and thus result in inhibition of the hammerhead ribozymes. Molecular modeling studies suggest that the spatial distribution of the amines emanating from the neomycin scaffold are able to simultaneously occupy the binding sites of up to four Mg^{2+} ions. Similar models of Mg^{2+} ion displacement have been invoked for inhibition of other ribozymes by neomycin.[105,107]

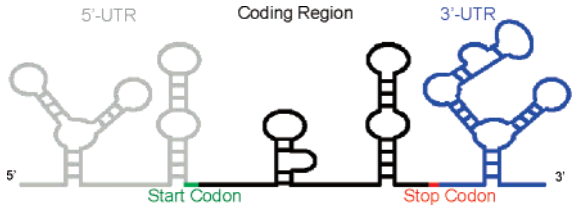


Figure 12. Three segments of a mRNA transcript: the 5′-UTR (shown in gray), the coding region (shown in black), and the 3′-UTR (shown in blue). Present within the 5′-UTR is the binding site for ribosomes, while the 3′-UTR controls the half-life of the mRNA and, in come cases, its subcellular localization.

4.7. Targeting mRNA

The ability to selectively inhibit the translation of a single mRNA transcript within the transcriptome has great therapeutic potential. RNA interference (RNAi) is the process by which small duplex RNAs capitalize on an evolutionarily conserved mechanism that allows for the sequence-selective silencing of a transcript. While clearly a revolutionary tool for biological research, RNAi is just now moving into the therapeutic realm, and the delivery of these RNAs into mammalian cells has proven to be a formidable challenge.[108,109] Therefore, the selective targeting of a transcript with small molecules could combine the benefits of RNAi (translational silencing) with more ideal biological stability and pharmacokinetic properties.

A transcript can be considered as three segments: 5′-untranslated region (UTR), the coding region, and 3′-UTR; each segment offers distinct possibilities for small molecule-mediated translational inhibition. These different regions participate in ribosome binding, translation, and controlling the rate of degradation and subcellular localization of transcripts, respectively (Figure 12).

4.7.1. Targeting mRNA: Riboswitch RNAs as a Proof-of-Concept for the Targeting of mRNAs

Riboswitch RNAs are a class of mRNAs that post-transcriptionally regulate protein translation through the specific recognition of a small molecule metabolite.[24,110] In all documented cases thus far, the proteins under this metabolite-controlled mechanism are directly responsible for the metabolite production; that is, the riboswitch−small molecule interactions act as a feedback loop. Riboswitch mRNAs consist of two domains which can be found primarily in the 5′-UTR of an mRNA: an aptamer domain that binds to its ligand with high specificity and affinity and a platform domain that undergoes conformational change in response to ligand binding and modulates translation (Figure 13). Although not ubiquitous, riboswitches have been observed in various species of bacteria, archaea, fungi, and plants.[111] In fact, bioinformatics studies suggest that $\sim$2% of *B. subtilis* mRNAs are regulated by riboswitches.[112] The general architecture of riboswitches is modular in that ligand-mediated translational control is still maintained when the aptamer domain of a riboswitch is replaced with a different aptamer domain from an unrelated riboswitch.[113−115] Such flexibility has fostered the engineering of artificial ri-

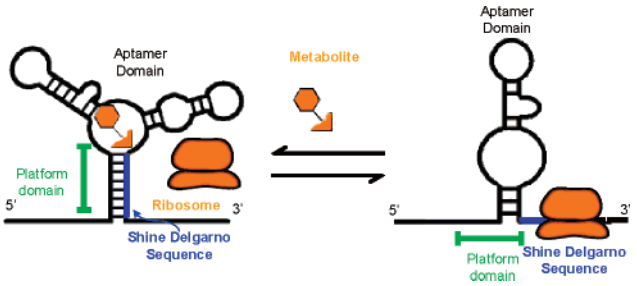


Figure 13. Riboswitch regulation of translation by altering the exposure of the Shine−Delgarno (SD) sequence. In the above example, the riboswitch−metabolite complex inhibits translation by preventing the ribosome from binding to the SD. However, in the unbound state the riboswitch RNA adopts an altered conformation that exposes the SD sequence, thus allowing translation to proceed.

boswitches for applications such as sensors[116] and control of bacterial migration.[117]

The structures of several riboswitch aptamer domains bound to their cognate ligands have recently been determined.[118−123] The binding site for each ligand is formed by the tertiary structure resulting from the association of various secondary structures. Within the binding pocket each small molecule makes an elaborate array of contacts which exquisitely discriminates between cognate and near cognate ligands. For example, the *S*-adenosylmethionine (SAM)-sensing riboswitch exhibits a 75-fold decrease in affinity if a single methylene unit is removed from the methionine side chain.[124] It should be noted that the ligands are substantially more buried in the binding pockets of the riboswitches than other small molecule−RNA complexes observed thus far.[118−123] The bound ligand conformations of the metabolites in their respective binding pockets were nearly identical to their previously determined conformations in solution, suggesting that the RNA (rather than the ligand) undergoes extensive conformational rearrangement.

Breaker and Blount recently proposed targeting riboswitch RNAs that regulate essential genes using small molecules that mimic the metabolite of the riboswitch.[125] Such compounds could be novel antibacterial agents. Critical to their argument for riboswitches as antibacterial targets, the authors note that previously reported metabolite analogs pyrithiamine pyrophosphate, L-aminoethylcysteine, L-4-oxalysine, and roseoflavin, all of which exhibit antibacterial properties, appear to achieve their effect by targeting specific riboswitch RNAs. Resistant strains to each of these metabolite mimics bear mutations in the aptamer domain, which alters the riboswitch−ligand interaction. Furthermore, efforts are underway to identify riboswitch inhibitors by assaying structural mimics of metabolites[126] and through high-throughput screening.[127,128]

4.7.2. mRNA Targeting: Antisense-Mediated Translational Control

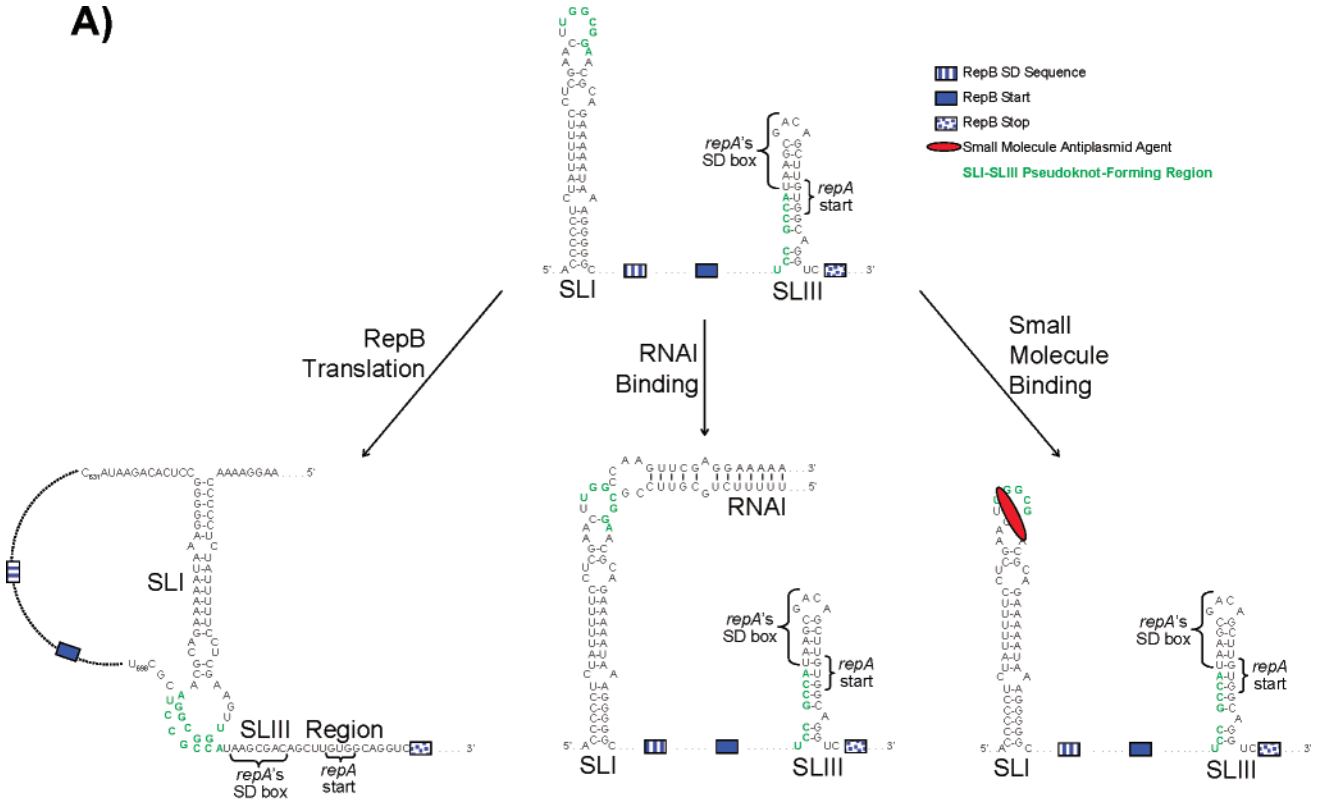
The theme of antisense-mediated translational control is common in bacteria and prevalent in plasmid replication systems.[129] In this post-transcriptional control system the translation of an mRNA (sense RNA) is regulated by the binding of a complementary RNA (antisense RNA). The antisense RNA is typically encoded on the complementary noncoding strand of the gene. Antisense RNAs regulate the translational state of the mRNA by binding via base pair formation to their respective mRNA, thereby inducing formation of secondary structure that prevents translation or causes the degradation of the transcript.[129−131] A bioinformatic investigation revealed that hairpin loops involved in such RNA−RNA interactions often contain a YUNR sequence (where Y = pyrimidine, N = any nucleotide, and R = purine), which is known to form a U-turn motif.[131] This common structural motif allows the hairpin loop to adopt a sharp bend in the phosphate backbone, displaying the bases in the loop for proper Watson−Crick base pairing and lowering the electrostatic repulsion experienced by the two approaching hairpin loops.[132] The ability to modulate such RNA−RNA interactions may provide a new avenue for the development of novel antibiotics or antiplasmid agents.[133]

The demonstration that small molecules can indeed function as antiplasmid agents was recently validated in the IncB plasmid system.[134,135] In this system, the plasmid copy number is ultimately controlled by the level of RepA protein, which acts as a phosphodiesterase to initiate plasmid replication.[136] The RepA mRNA is under stringent translation control by an antisense mechanism. The 5′-UTR of the RepA mRNA forms multiple stem-loop structures; the pairing of one such stem-loop structure (SLI) with its antisense RNA (RNAI) prevents translation as the Shine−Dalgarno (SD) sequence remains base paired in the duplex region of a stem loop (SLIII) (see Figure 14A).[137] However, free SLI binds to SLIII to stabilize the open form of the helix, exposing the SD sequence, allowing translation of the RepA protein and ultimately leading to plasmid replication. It was proposed that small molecules could mimic the function of RNAI by binding to SLI to prevent association with SLIII, thus preventing plasmid replication and leading to plasmid loss.[134] Plasmid elimination strategies may prove useful in the clinic as bacteria often harbor multidrug resistance plasmids;[138] elimination of such plasmids would render the bacteria susceptible to antibiotics to which they had previously been resistant.

In an effort to identify potential antiplasmid agents, a collection of aminoglycosides was screened for their ability to bind SLI. All 4,5- and 4,6-deoxystreptamine aminoglycosides that were tested bound SLI with mid-nanomolar affinity, while hygromycin B and spectinomycin exhibited no binding affinity (Figure 14B).[134,135] Mutagenesis of residues in and around the hairpin loop region revealed the binding site of the aminoglycosides to reside in the symmetric internal loop located closest to the hairpin loop. Results from bacterial cell culture experiments demonstrated that apramycin induced plasmid loss in a dose- and time-dependent manner.[134] The other aminoglycosides that bound SLI in vitro exhibited varying degrees of plasmid loss, while those that failed to bind SLI did not induce plasmid loss.[135] The antiplasmid effect of apramycin was shown to be linked to SLI binding in bacterial cells, as a plasmid bearing a mutant SLI sequence that abolished the apramycin binding site was unable to be eliminated by apramycin.[134]

4.7.3. mRNA Targeting: Protein-Mediated Translation Control

In mammalian cells, proteins can be used in an analogous fashion to the small molecule metabolite- and antisense-mediated translational control mechanisms described above. For example, the binding of the thymidylate synthase (TS) protein to its own mRNA negatively regulates its expression.[139,140] TS binds to two distinct regions of its own mRNA; site 1 resides in the first 188 nucleotides which contains



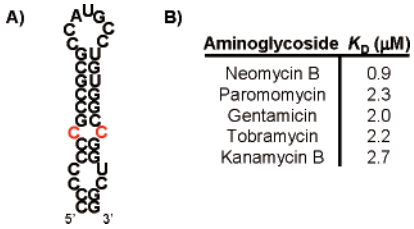
B)

Aminoglycoside	K_D (μM)
Sisomicin	0.03
Paromomycin	0.04
Tobramycin	0.04
Neomycin	0.08
Gentamicin	0.08
Amikacin	0.08
Kanamycin B	0.11
Kanamycin A	0.17
Butirosin B	0.20
Hygromycin B	>4.00
Spectinomycin	>4.00

Figure 14. Antisense-mediated translational control in the IncB plasmid replication system. (A) The translation of RepB opens SLIII, allowing SLI to stabilize the exposed SD sequence. However, the binding of RNAI (or a small molecule) prevents the binding of SLI to SLIII, thus inhibiting translation. (Reprinted with permission from ref 135. Copyright 2005 American Chemical Society.) (B) SLI construct utilized for in vitro binding studies and the corresponding binding constants for various SLI−aminoglycoside interactions. Residues critical for the binding of the aminoglycosides are shown in red.

portions of the 5′-UTR and coding segment, while site 2 is located farther downstream in the coding region.[139] Within site 1 the exact binding site for TS is known to be a 35 nucleotide stem-loop structure, which contains two small symmetric internal loops and a six-membered hairpin loop (Figure 15).[140] Within the hairpin loop resides the start codon for TS mRNA; the presence of the AUG sequence is essential for TS binding as the corresponding AAA mutant does not bind TS.[140] The provocative finding that TS binds to its mRNA in a stem-loop region that contains its start codon has led to the proposal that TS binds to site 1, and specifically the 35 nucleotide stem-loop, as a negative feedback loop to inhibit its own translation.[140]

The thymidylate synthase protein has been in the crosshairs of pharmaceutical R&D programs for decades.[141,142] TS is the sole enzyme responsible for the de novo synthesis of



Aminoglycoside	K_D (μM)
Neomycin B	0.9
Paromomycin	2.3
Gentamicin	2.0
Tobramycin	2.2
Kanamycin B	2.7

Figure 15. Thymidylate synthase mRNA−aminoglycoside interactions. (A) The secondary structure representation of the site I TS construct. Highlighted in red is the proposed binding site of the aminoglycosides. (B) Binding constants of various aminoglycosides for the site I TS construct.

thymine monophosphate, which is then further metabolized into thymine triphosphate for incorporation into DNA.[143] Because of the essential nature of TS, small molecule inhibitors of this enzyme have been developed. However, certain classes of inhibitors result in the loss in the ability of TS to regulate its own translation,[144−146] the net result of which is the constitutive translation of TS. As the concentration of TS increases within the cell, reduced efficacy is observed with TS inhibitors. Thus, in the case of TS there is a need for alternate approaches toward inhibition of the function of this enzyme. A series of aminoglycosides was assayed for binding to the TS site 1 construct (Figure 15).[147] Neomycin was found to bind to the stem-loop construct with K_D of ∼1 μM.[147] Mutational analysis revealed that the small 1 × 1 internal loop, but not the larger 2 × 2 internal loop, is critical for neomycin binding. In addition, the hairpin loop region of TS has no effect on neomycin binding as the sequence of the hairpin loop region of the 16S A site could be used in place of the TS hairpin loop sequence.

The translational inhibition exerted by TS is recapitulated in other systems. For example, a ∼30 nucleotide sequence known as the iron-responsive element (IRE) is found in the 5′-UTR of all ferritin mRNAs and the 3′-UTR of all transferrin receptor mRNAs.[148,149] Regulation of ferritin and transferrin receptors works in synergy in low iron conditions, as the upregulation of transferrin imports iron from the surrounding environment into the cytoplasm and the simultaneous downregulation of ferritin prevents the protein from binding and sequestering iron.[150,151] The levels of ferritin and transferrin receptors are regulated by iron-regulating proteins 1 and 2 (IRP-1 and IRP-2); IRP-1 binds to the ferritin IRE to inhibit translation by preventing initiation factors from binding to the 5′-cap of the ferritin mRNA.[150] IREs are highly conserved in secondary structure, as bioinformatics approaches have defined a consensus sequence and structure of this segment of RNA.[152] The consensus sequence of IREs consists of two helical segments separated by a single bulged cytosine and capped by a hexahairpin loop sequence of CAGUGH (where H is either A, U, or C). The bulged region is believed to act as a molecular hinge which allows for conformational change upon IRP-1 binding;[153] deletion of the bulged region results in a nearly 400-fold decrease in binding affinity.[154] The invariant five base pair region between the bulge and hairpin loop is believed to act as a molecular ruler.[153] Modulation of iron availability by small molecule regulation of IREs has been proposed as an adjuvant therapy for the treatment of sickle cell disease.[155]

4.7.4. mRNA Targets: Inhibition of Ribosome Scanning

Conventional cap-dependent translation in eukaryotic systems proceeds first with the binding of the small ribosomal subunit to a 5′-methylated cap of a mRNA.[156,157] Upon binding the small subunit forms a complex with eIF-2-GTP-Met-tRNA$_i$. The complex then begins the "scanning" process whereby the complex proceeds linearly down the 5′-UTR searching for the first AUG codon. The sequence-specific recognition is caused by the pairing anticodon loop of the Met-tRNA$_i$ and the initiation codon. After this recognition, the complex pauses and waits for the large ribosomal subunit to associate before translation begins.

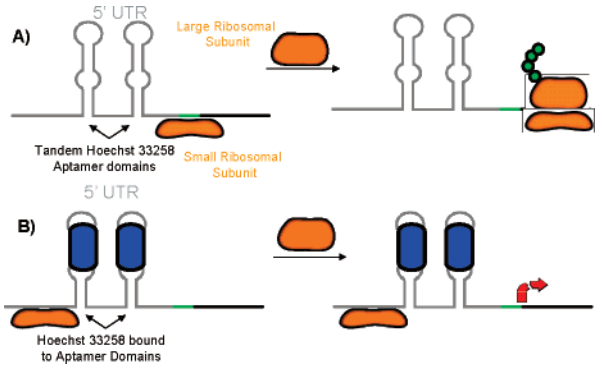


Figure 16. Hoechst 33258 aptamers as proof-of-concept for the selective inhibition of a specific transcript. (A) Insertion of tandem Hoechst 33258 aptamers in the 5-UTR of the β-galactosidase gene had no effect on translation. (B) Addition of Hoechst 33285 (in blue) selectively inhibited the expression of β-galactosidase as determined by the level of its enzymatic activity.

Sufficient secondary structure in the 5′-UTR is known to pause and/or inhibit the scanning process, which results in translation inhibition.[158−161] Insertion of a 17 nucleotide stem-loop structure, with a predicted $\Delta G_{folding}$ of −30 kcal/mol, has been shown to have no effect on translation efficiency, even when the start codon is placed in the stem of the stem-loop structure.[158] However, insertion of tandem 17 nucleotide stem-loops results in nearly complete inhibition of translation.[158] Furthermore, insertion of a single stem-loop structure, with a predicted $\Delta G_{folding}$ of −50 kcal/mol, severely attenuates translation, with only 10−20% translation efficiency observed.[158] Thus, if a small molecule could bind selectively to a highly structured region within the 5′-UTR of a specific transcript, then the small molecule could possibly cause transcript-selective translational inhibition. Evidence that a small molecule can induce the same translational arrest by binding to specific regions within the 5′-UTR has been demonstrated with proof-of-concept experiments.[162] Insertion of tandem Hoeschst 33528-binding aptamers into the 5′-UTR of a β-galactosidase reporter gene was used to validate the concept of small molecule 5′-UTR-mediated translation inhibition (Figure 16). When CHO cells were transfected with a plasmid bearing the tandem Hoescht 33285 aptamers a dose-dependent reduction in β-galactosidase activity was observed upon addition the ligand. Internal controls using luciferase activity determined that effects of Hoescht 33258 were not due to general translation inhibition.

4.7.5. mRNA Targets: Internal Ribosome Entry Site Inhibition

The discovery of a novel cap-independent mechanism of translation in several viruses has prompted intense investigation into the molecular mechanism underlying this distinct mode of translation.[163−167] Viral infection of the host cell can stimulate phosphorylation of elongation factors, inhibiting cap-dependent translation. However, certain viral RNAs

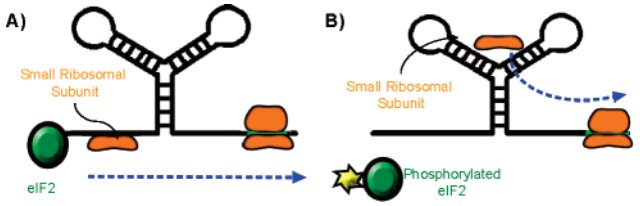


Figure 17. IRES-mediated translation. (A) Under standard cap-dependent translation the binding of eIF2 (green) recruits the small ribosomal subunit, which then proceeds through the typical scanning mechanism of translation initiation. (B) Under conditions where cap-dependent translation is impaired, e.g., phosphorylation (yellow) of eIF2, certain RNA sequences mediate translation by allowing the small subunit to bind independent of eIF2.

are able to recruit the small subunit directly to the 5′-UTR, bypassing the need for binding to the 5′-methylated cap and initiating translation in a cap-independent manner (Figure 17). Because the internal ribosome entry site (IRES) method of translation is distinct from the standard cap-dependent translation, it has been speculated that specific inhibition strategies could be developed which would not affect host cell processes.[164,168,169] Interestingly, a number of pro- and anti-apoptotic proteins have been demonstrated to operate by an IRES-mediated mechanism. For example, Apaf-1, Bcl-2, Reaper, and XIAP appear to utilize an IRES-mediated mechanism of translation.[163,168] The need for cap-independent translation under apoptotic conditions arises from the cleavage of caspase substrates, which include numerous translation initiation factors.[163,168] Thus, the ability to inhibit IRES-dependent translation provides novel opportunities for the discovery of new antiviral and apoptotsis-modulating agents.

The hepatitis C virus (HCV) is a leading cause of illness and mortality worldwide,[170] and treatment options are limited for afflicted individuals.[171,172] The discovery that HCV proteins are translated in an IRES-mediated mechanism has intensified efforts to understand the molecular mechanism behind its unique mode of translation. The 9.5 kb genomic RNA contains a nearly 400 nucleotide long noncoding region; located within this region is the IRES.[173,174] The HCV IRES forms extensive secondary structure which adopts four highly structured domains; domains III and IV are required for IRES translation as they function in binding to the small ribosomal subunit.[175] Although domain II itself fails to bind to the small ribosomal subunit, it is essential for IRES activity.[175] Elucidation of the entire 77 nucleotide domain II stem-loop structure by NMR revealed that the 3D conformation exhibits an L shape, reminiscent of that adopted by tRNAs.[176] Cryo-electron microscopy has determined that domain II interacts with the small subunit in the region of the E site, the site for deacylated tRNAs prior to their exit from the ribosome.[177] As more structural information emerges and the model of HCV IRES-mediated translation becomes more refined, IRES sites are likely to become the new wave of RNA targets.

Despite the tremendous opportunity that the selective targeting of IRES elements presents, there exists significant debate as to the practicality of such an approach and even the in vivo relevance of IRES elements in general.[178] A survey of the IRES literature revealed that there have been 85 cellular and 39 viral IRES elements reported thus far.[163] However, there has been harsh criticism of the assays used to validate the presence of IRES elements in these various studies.[156] Furthermore, no consensus sequence or structure for the various IRES elements has been identified, which brings into question the validity of this seemingly widespread phenomenon. As such, small molecule binders to various IRES elements could play an important role in defining the biological and medicinal relevance of this cryptic RNA element.

4.7.6. mRNA Targets: Sequence and Structurally Distinct Coding Regions

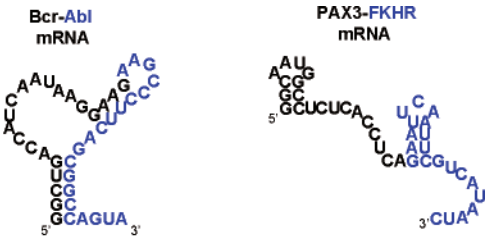
One can envision the selective targeting of an mRNA coding region by identifying sequences and structures unique within the transcriptome in an analogous manner to the targeting of structurally distinct 5′-UTR segments. Such novel transcripts are likely to be specialized cases. The examples presented for translational control from the 5′-UTR, whether antisense RNA, protein, or small molecule mediated, demonstrate that high-affinity ligands can inhibit the initiation of translation. Extending such a targeting strategy to the coding region of an mRNA would be a dramatic first in translation inhibition.

Sites of genetic translocation have attracted considerable attention as novel drug targets, particularly for chronic myelogenous leukemia (CML), where a specific translocation event gives rise to this disease.[179] The fusion of Bcr with Abl in CML results in a constitutively active Bcr-Abl kinase. This translocation event provides a novel target for a small molecule inhibitor, given that the fusion protein only exists in CML and not in normal cells. The targeting of Bcr-Abl kinase by Gleevec (imatinib) has resulted in an extremely effective therapy for CML, with in some cases up to 95% of patients responding to treatment.[179]

At the mRNA level, sites of translocation also afford a unique target that would be present only in disease states. The drugability of the Bcr-Abl and PAX3-FKHR translocation sequences was investigated using various aminoglycosides.[180] The PAX3-FKHR fusion protein, resulting from a translocation of the genes encoding two transcription factors, occurs in the aggressive skeletal muscle cancer alveolar rhabdomyosarcoma.[181] The binding affinities of aminoglycosides to each of these sites of translocation ranged from 1.5 to $>100\ \mu$M, with neomycin exhibiting the tightest binding (Figure 18).[180] No significant degree of selectivity was observed for any of the aminoglycosides for either sequence, although some synthetic derivatives exhibited slightly improved binding affinities and selectivity.[180]

Despite the potential of selectively targeting mRNAs in their coding region, some concerns exist regarding the feasibility of impeding a translating ribosome. It is currently known that the translating ribosome can disrupt secondary structure formation more efficiently than the small subunit alone.[182] For example, insertion of the IRE sequence into the 5′-UTR will inhibit translation upon binding IRP-1.[183] The dissociation constant for this RNA−protein interaction is ∼10 pM.[184] When this same sequence is placed in the coding region of a mRNA, no translation defects are observed;[183] this suggests that the translating ribosome is capable of overcoming significant structural impediments.

From the above example it seems that the most practical way to inhibit a translating ribosome is to target a repeating structural motif in the coding region. Trinucleotide repeat expansion diseases (TREDs) are a group of diseases characterized by the repetitious expansion of a three nucleotide sequence; the most well known of these diseases is Huntington's disease (HD).[185,186] HD is characterized by 34 or more glutamine repeats in the huntingtin protein sequence.[187]



Aminoglycoside	K_D (μM)
Neomycin	1.5
Paromomycin	1.7
Kanamycin B	2.5
Gentamicin	3.0
Tobramycin	3.6
Kanamycin A	10
Amikacin	10
Butirosin	15
Neamine	17
Geneticin	20
Streptomycin	20
Apramycin	20
Ribostamycin	80
Hygromycin B	>100
Spectinomycin	>100

Aminoglycoside	K_D (μM)
Neomycin	1.5
Paromomycin	1.8
Tobramycin	1.8
Kanamycin B	5.0
Gentamicin	6.0
Amikacin	15
Apramycin	20
Kanamycin A	20
Neamine	20
Butriosin	25
Streptomycin	25
Geneticin	30
Ribostamycin	90
Hygromycin B	>100
Spectinomycin	>100

Figure 18. Binding affinities of various aminoglycosides for the Bcr-Abl and Pax3-FKHR sites of translocation. The sequence of the PAX3-FKHR RNA was obtained through a personal communication with the authors.

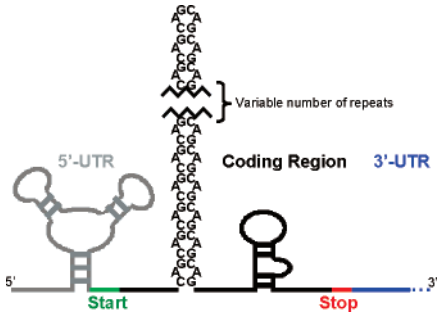


Figure 19. TREDS as structurally unique RNA targets. The repetitious polyQ sequence within the huntingtin protein is encoded by repeating CAG codons. At the mRNA level the repeating CAG sequence gives rise to a well-defined secondary structure.

The net result of this poly glutamine (polyQ) motif is the aggregation of the huntingtin protein, which is believed to be toxic to the cell.

In healthy cells such repeating sequences are exceedingly rare. In fact, only 2% of the entire transcriptome carries a trinucleotide repeat of six or more.[188] Also, because of their long repeating nature, the mRNA coding for TRED proteins has a unique secondary structure. For example, symptoms of HD are manifest when the polyQ expansion reaches greater than 34 glutamine residues, meaning the corresponding mRNA has greater than 34 repeating CAG codons. Biochemical investigation into the secondary structure of CAG repeating RNAs has uncovered a unique repeating structure;[189] CAG repeats have the recurring secondary structure of a 1 × 1 internal loop made up by an A−A mismatch separated by a G−C, C−G base pair (Figure 19). Thus, the long, repeating expansions causing TREDs may provide a novel target for RNA-binding small molecules.

4.7.7. mRNA Targets: 3′-Untranslated Region

The 3′-UTR of mRNAs represent an unexplored area for small molecule targeting in large part because the fundamental biology of many of the systems is not fully understood. As discussed previously (see section 4.7.3), the binding of IRP-1 and IRP-2 to the IRE sequence located in the 3′-UTR of the transferrin receptor mRNA increases its stability in order to increase expression levels of the transferring receptor protein.[190] In addition, the length of the poly(A) tail generally correlates with translation activation with longer poly(A) tails yielding greater translational activation.[191] A model for how a 3′-UTR element may control translation initiation, a 5′-UTR phenomenon, has been proposed to occur by pseudo-circularization of the mRNA by the binding of various factors from the 5′-UTR and 3′-UTR.[192,193] The 3′-UTR has also been demonstrated to determine subcellular localization of various mRNAs.[194] For example, the β-actin mRNA is known to localize to the leading edge of asymmetric cells, such as fibroblasts. The specific subcellular localization of the β-actin mRNA is known to be mediated by the binding of factors to the "zip code" region of the 3′-UTR, a 54 nucleotide AC-rich nucleotide with several repeating regions of ACACCC.[195] Although one might imagine a variety of ways that small molecule 3′-UTR binders could perturb a biological system, additional basic biochemical information is needed before it can be determined if the 3′-UTR is a valid target for RNA-binding compounds.

4.8. MicroRNAs

The discovery of RNAi has revolutionized the selective silencing of a gene product in vivo. MicroRNAs, 20−25 nucleotide sequences that mediate the destructive silencing of transcripts in mammalian cells in a RISC-dependent manner by binding to the 3′-UTR of transcripts in a sequence-specific fashion, appear to be endogenously expressed RNAi molecules.[196] Also, the imperfect pairing of microRNAs and RNA targets can lead to translational inhibition without cleavage of the transcript.[197] Recent estimates suggest that a significant fraction of human genes are regulated by microRNAs.[198] Accordingly, in addition to the regulation of critical processes such as cellular proliferation,[199] development,[200] differentiation,[201] and apoptosis,[202] defining the role of microRNAs in carcinogenesis and sustained progress of cancerous cells is actively being pursued. As microRNAs regulate the translation of target mRNAs, the up- or downregulation of a particular microRNA from its "normal" state can cause the microRNA to act as either a tumor suppressor or an activator.[196] The upregulation of a particular microRNA is likely to lead to a decrease in its targeted protein levels, while downregulation should lead to an increase. Correlating the levels of a particular microRNA with its protein targets is complicated by the observations that multiple microRNAs can regulate one transcript. As the exact molecular details behind microRNA regulation emerge, it is quite likely that many RNA-based targets will begin to surface from these prevalent translational regulators.

5. General Principles of RNA Binding: Lessons from the Aminoglycosides

The aminoglycosides exert their antibiotic effect through binding to the A site of the 16S rRNA, leading to incorporation of noncognate tRNAs and cell death (see section 4.1 and Figure 3).[51,55] These compounds have found widespread utility in studying general facets of small molecule−RNA binding; in fact, the majority of investigations into small molecule−RNA binding have been conducted with ami-

noglycosides. The general affinity of the aminoglycosides for many different RNAs has been both a blessing and a curse: on one hand, when searching for a binder to a novel RNA, it is likely that some aminoglycoside will bind. On the other hand, this promiscuity often leaves investigators with compounds that are useful for in vitro studies but less useful for cell culture or in vivo work. Nevertheless, much data has been collected on the aminoglycoside−RNA interaction, and as summarized in this section, certain basic principles have emerged.

5.1. Binding Sites of the Aminoglycosides

It is now apparent that aminoglycosides and some synthetic small molecules will bind to pockets created by the bases in and around internal loops and bulged regions of RNA. Furthermore, it appears that RNA−RNA and RNA−protein interactions follow these same recognition principles, and aminoglycosides are able to compete for the same binding sites. Thus, for maximal biological effect of a small molecule, the targeting of regions of RNA−macromolecule interactions seems to be a good choice. There are, of course, multiple limitations and caveats placed on the use of aminoglycosides to target RNA. Aminoglycosides are notoriously nonspecific and will bind to many RNA sequences at low micromolar levels.[203] In addition, there are almost no documented examples of aminoglycosides binding to RNA hairpin loops. Finally, the nephro- and ototoxicity of aminoglycosides severely limits their in vivo use at high dosages. That said, the fundamental knowledge gained by close examination of aminoglycoside−RNA interactions will ideally provide a springboard to the next generation to RNA binding small molecules.

5.2. Importance of Electrostatic Interactions

The binding affinities of various aminoglycosides for their target RNAs often correlate with the number of amines present.[72,204] The molecular structure of the aminoglycosides appears to be tuned to modulate the basicity of the amines; the proximal hydroxyl group is believed to lower the pK_a of the amine, thus altering the overall net charge of the compound. The presence of a hydroxyl group proximal to an amine can substantially modulate the ability of various aminoglycosides to bind to the hammerhead ribozymes.[205] Analogous experiments, in which pH was altered or the amine functionalities were replaced with guanidinum groups, demonstrated enhanced aminoglycoside−RNA binding affinities due to an increase in the overall net charge.[206−209]

In an alternative approach, the importance of electrostatic interactions was demonstrated by analyzing aminoglycoside−RNA dissociation constants in the presence of varying concentrations of NaCl; performing binding assays with increasing concentrations of NaCl led to a linear decrease in binding affinity.[208,210] Increasing the salt concentration lowers the electrostatic potential between the negatively charged RNA and positively charged aminoglycosides by shielding the phosphate backbone. Subsequent work has demonstrated that electrostatics account for at least one-half of the total binding energy in certain aminoglycoside−RNA interactions.[207]

Although electrostatic interactions are responsible for the strong aminoglycoside−RNA affinity, they are also responsible for the promiscuity of their RNA binding. In several cases the aminoglycoside:RNA binding stoichiometries were found to be greater than 1:1, suggestive of aminoglycosides bound to multiple regions of RNA secondary structure. Addition of increasing concentrations of NaCl can reduce these weaker affinity interactions, indicating that these lower affinity, nonspecific interactions are primarily mediated by electrostatic interactions.[100] Thus, favorable electrostatic interactions will likely enhance the affinities of synthetic small molecule RNA-binding compounds but at the risk of enhanced promiscuity.

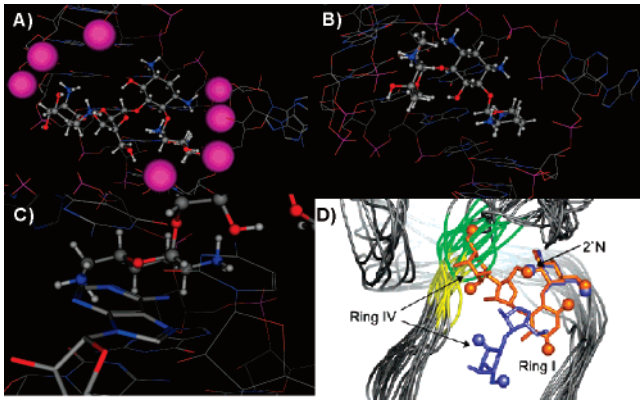


Figure 20. General principles of RNA binding as exemplified by aminoglycoside−RNA structures. (A) The crystal structure of neomycin bound to the 16S A site illustrates the importance of electrostatic interactions as the numerous amines present in neomycin interact favorably with the phosphate-lined binding pocket (pink spheres). (B) The crystal structure of gentamicin bound to the A site demonstrates the importance of nonionic interactions as ring I stacks over top of G1491. (C) Ring I of gentamicin, and other aminoglycosides, forms pseudo-base pair interactions with A 1408. (D) Overlay of the conformationally restrained neomycin derivative (orange) with the solved structure of the neomycin (blue)−TAR (gray) complex. As shown, in order for the neomycin derivative to bind TAR a different conformation must be achieved. (Reprinted with permission from ref 219. Copyright 2005 American Chemical Society.)

5.3. Nonionic Interactions

When comparing the structure of the aminoglycosides (global positive charge) with RNA (negative charge), the electrostatic interactions are immediately obvious; however, nonionic interactions do play an important and often overlooked role in aminoglycoside−RNA interactions. For example, several aminoglycosides have been shown to bury their less polar portions into hydrophobic regions of RNA-binding sites.[53] In the NMR structure of gentamicin C1a bound to the 16S A-site construct, the network of charged moieties interacts with the anionic phosphate backbone and heteroatoms of the bases while ring I stacks above G1491 and packs against A1492 (Figure 20).[211] In their investigation of gentamicin C1a−16S rRNA interaction, Puglisi and coworkers compared the binding orientation of paromomycin with gentamicin C1a. Both ligands bind with comparable affinity to the 16S A-site construct and exhibit similar binding modes for rings I and II. However, as the hydroxyl groups on ring I of paromomycin make contacts with the phosphate backbone, the authors proposed that gentamicin C1a compensates for the loss of these favorable interactions by stacking its hydrophobic portion of ring I against G1491. Crystal structures of aminoglycosides complexed with the 16S A site reveal this same stacking interaction, although in the crystal structures this interaction is more pronounced.[53,212]

5.4. Pseudo-Base Pair Interactions

As discussed above, the binding sites for aminoglycosides (and other small molecules) are generally restricted to areas where the A-form RNA helix is perturbed by noncanonical interactions. Some aminoglycosides appear to mimic base pair contacts with unpaired nucleotides, thus enhancing their recognition by the RNA binding site. Because of the stacking interaction discussed above, ring I of various aminoglycosides is poised to recognize the Watson−Crick face of A1408 (Figure 20C). This pseudo-base-pairing interaction is not specific to the glucosamine moiety of 4,5- and 4,6-substituted deoxystreptamine aminoglycosides because apramycin achieves the same pseudo-base pair recognition using its bicyclic core.[212] Furthermore, apramycin uses its ring III to bind to the C1409−G1491 base pair in a base triplet fashion in the minor groove.[212]

5.5. Water-Mediated Contacts

As expected, the crystallographic structural determination of many RNAs has revealed an extensive hydration shell surrounding various regions. Previous work has established that the interaction between a macromolecule and its solvent can significantly impact the electrostatic surface potential of the macromolecule;[41,213,214] that is, water molecules can enhance or dampen the existing charges within binding pockets. Aminoglycosides can exploit these conserved water molecules to enhance specificity and affinity. In fact, approximately one-third of the interactions formed between the aminoglycosides and 16S A site are made by water-mediated contacts.[26,215,216] Its mobility in the binding pocket allows the water molecules to rearrange upon ligand binding, optimizing geometry and distance of hydrogen-bonding interactions between the ligand and RNA. Furthermore, it has been suggested that such water-mediated contacts serve to reduce the dehydration penalty that the charged aminoglycosides must pay in order to penetrate deeply within the A-site binding pocket.[26]

5.6. Shape Complementarity and Conformational Adaptation

The concepts of shape complementarity (how well a ligand and receptor fit together electrostatically and sterically) and conformational adaptation (binding-induced changes in both the ligand and receptor) are familiar concepts in the field of molecular recognition. However, each concept has a particular emphasis when considering RNA−small molecule interactions. As discussed in section 3, the various folds adopted by the different RNA secondary structures create pockets which differ in architecture and electronegative potential; however, shape complementarity is likely to be dominated by electrostatic interactions as over one-half of the total free energy for aminoglycoside−RNA binding is due to electrostatic interactions.[207] Such a heavy reliance on electrostatic interactions to achieve high affinity makes attaining specificity difficult because of the anionic nature of all RNA binding pockets. For example, modeling studies of solution conformations of neomycin into the hammerhead ribozymes revealed five different orientations that fit the pocket equally well.[217]

In addition, steric complementarity is further reduced as a discriminator because of the inherent flexibility of the various secondary structures, which allow for substantial rearrangement upon ligand binding.[218] Such conformational adaptation has been proposed to be a major contributor to the lack of specificity observed for the aminoglycosides.[218] In attempting to create aminoglycoside derivatives with enhanced specificity for the 16S A site over the TAR RNA, Tor and co-workers synthesized conformationally locked aminoglycosides.[219] By comparing the paromomycin-16S A-site structure with the neomycin−TAR structure, the authors predicted that linkage of the 2′-amine of ring I with the 5″-carbon would provide the correct binding conformation for the 16S A site but not TAR. Despite their well-intended plans, both locked ligands exhibited the same levels of selectivity for the A site over TAR as the unlocked parent aminoglycosides. The authors propose that based on the known flexibility of the TAR RNA binding pocket, the restricted analogues must be achieving a different binding orientation (see Figure 20D); that is, in this case the structural plasticity of the RNA is responsible for the observed promiscuity.

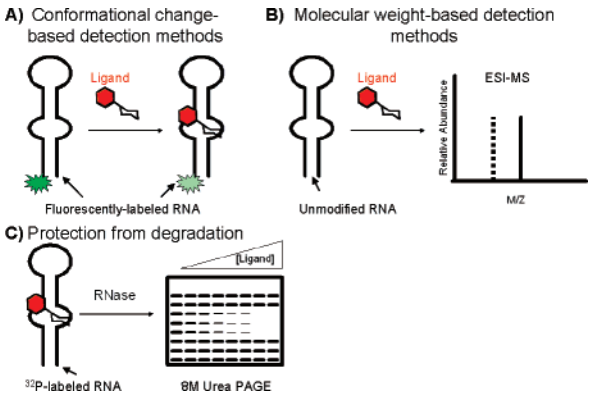


Figure 21. Detection of small molecule−RNA interactions. (A) Binding constants can be determined based on the conformational change of a fluorescently-labeled RNA upon ligand binding. (B) Molecular weight-based methods of detection. ESI-MS allows detection of RNA−ligand interactions by measuring the molecular weight of the ligand-bound complex. The dashed line in the representative MS spectra denotes the molecular weight of the RNA construct, while the solid line represents the molecular weight of the ligand-bound complex. (C) The ability of a small molecule to protect RNA residues from enzymatic or chemical digestion provides another readout of binding.

6. Assays for Evaluating RNA Binding

The in vitro effect of RNA binding ligands typically cannot be evaluated through a direct enzymatic assay as the RNAs being examined do not usually have catalytic activity; ribozyme inhibitors are an obvious exception to this. This lack of enzymatic readout complicates the identification and evaluation of small molecule ligands for RNA. However, several methods are available to the experimentalist looking for a quantitative method to assess small molecule−RNA binding. As robust detection of small molecule−RNA interactions is a problem without a universal solution, the methods that have been developed fall into several categories as delineated further below. Some of these methods are overviewed in Figure 21.

6.1. Methods Utilizing Fluorescently Labeled RNA

The ability to obtain modified RNA oligonucleotides from commercial vendors has facilitated the development of fluorescence-based methods for detection of ligand binding (Figure 21A). In these methods, the RNA oligonucleotide is

made fluorescent by appending a fluorophore to one end or utilizing a fluorescent version of a nucleotide either within or adjacent to the small molecule binding site. In either case, it is generally accepted that the conformational change associated with ligand binding alters the local environment of the fluorophore sufficiently to provide a readout of binding based on a change in fluorescence; conformational changes are often observed for ligand−RNA interactions.[218]

The use of 2-aminopurine as a "site-specific" probe for ligand binding is a popular choice. The naturally fluorescent 2-aminopurine is a versatile probe and has been incorporated into hairpin loops,[220] internal loops,[54,99] bulged regions,[221] and loop junctions.[222] The advantage of this method is its site-specific nature; that is, in an RNA containing multiple secondary structures, the selective incorporation of 2-aminopurine into one of the secondary structures can allow for detection of binding to only the site of interest. However, in practice binding events distal to the location of the 2-aminopurine are also observed.[99,221] Another advantage is that the 5′- and 3′-ends of the RNA are unmodified, allowing for standard follow up assays such as RNase foot printing (discussed below).

A complementary approach to the 2-aminopurine method is the attachment of a fluorophore to either the 5′- or 3′-end of RNA. End-label incorporation of fluorophores has been used to study the hammerhead ribozymes[223] and the 16S A site;[224] in each case the results were similar to those obtained using the 2-aminopurine method. In addition to studying these well-characterized systems, the end-label method has been used in a discovery mode against novel RNA targets.[134,135,225−227] Importantly, because the site of modification is typically on the 5′-end, chemical synthesis of the RNA construct is not a requirement. Rather, standard in vitro transcription assays using guanosine 5′-monophosphorothioate (5′-GMPS) can be used to enzymatically produce a full-length RNA construct in which the GMPS is incorporated as the first nucleotide, allowing it to react readily with iodoacetamide fluorophores or other desired labels.[204,228] A potential drawback of the end-label assay is that the magnitude of change in fluorescence is likely a function of the distance between the binding site and the fluorophore as well as the magnitude of the conformational change. Although the end-label method has been implemented successfully where the fluorophore and binding site were as distant as 11 base pairs away[134,135] and in more conformationally restrained systems such as the 16S A site,[224] success is not guaranteed.[72] Recent work has shown that for deoxystreptamine dimers binding to RNA hairpin loops the binding constants obtained using the end-labeled method are similar to those obtained through isothermal titration calorimetry (ITC).[229]

In an attempt to merge the site-specific nature of the 2-aminopurine method without modifying the base itself, a binding assay was developed based on the incorporation of a pyrene label. Incorporation of commercially available pyrene-labeled uracil (connected through the 2′-position via flexible alkyl linker) into a target RNA potentially removes disadvantages associated with above fluorescence-based methods; that is, the fluorophore is placed directly in the site of interest, and the native nucleobase is used. The pyrene-labeled method has been successfully used to investigate various aminoglycoside−TAR interactions.[230] Limitations of this method include a (current) lack of commercially available nucleotides bearing the pyrene label and the modification of the nucleotide residue in the binding site.

6.2. NMR-Based Methods

The length of many RNA constructs used in biochemical assays fits within the size window suitable for structural determination by NMR.[231] The ability to determine the exact binding site of small molecules, even ligands with affinities in the millimolar range, as well as the conformational changes induced upon ligand complexation are the key advantages of this technique as a screening and validation tool. As a method for the discovery of novel protein ligands, various NMR techniques have been developed that aid in structure−activity relationships and lead generation from fragment libraries.[232] Use of NMR for studying RNA-binding small molecule interactions has not yet reached the level of sophistication where such techniques are routine for screening purposes, although in principle all the tools are in place.

Direct binding of a ligand to a RNA target is monitored by determining imino proton (1D) or pyrimidine H5−H6 chemical shifts (2D), although standard 1D experiments appear to be the method of choice for RNAs with previously determined structures.[231] Other techniques to monitor RNA−ligand interactions measure the resonance of the free ligand as compared to bound species.[233−237] For example, WaterLOGSY (water-ligand observed via gradient spectroscopy), which relies on the bulk magnetization of water to differentiate free from bound ligand, was found to be a superior NMR technique compared to other 2D experiments.[233] A significant downside to NMR is the need to obtain large quantities of RNA, in some cases isotopically labeled. Use of site-specifically labeled ^{19}F nucleotides can compensate for the above-mentioned drawbacks;[238,239] however, such specially labeled constructs are not commonly used likely owing to fact that such constructs can only be accessed by solid-phase synthesis. This is in contrast to more traditional NMR experiments in which in vitro translation is used to obtain the desired construct, including doubly labeled ^{15}N and ^{13}C RNA constructs.

6.3. Electrospray Ionization Mass Spectrometry-Based Methods

Electrospray ionization mass spectrometry (ESI-MS) provides a gentle ionization method that allows for a molecular weight determination of bound receptor−ligand complexes in the gas phase (Figure 21B). ESI-MS provides a unique assay to determine ligand−RNA interactions that is not reliant on conformational changes associated with complexation. This label-free technique can provide the ligand−RNA association constant and stoichiometry of binding by monitoring the shift in molecular weight of the RNA.[240] Griffey and co-workers at Ibis Pharmaceuticals pioneered various approaches for monitoring RNA−ligand interactions via ESI-MS. These authors demonstrated that the gas-phase binding affinities are consistent with those derived from solution-phase experiments for the aminoglycoside−16S A-site interaction.[241] Furthermore, the authors developed techniques to identify the binding site of the small molecule using collisionally activated dissociation MS to fragment the RNA.[242,243] Perhaps where ESI-MS holds the greatest promise is in the area of high-throughput screening, where the authors successfully screened a mixture of RNA targets against a mixture of compounds.[244] The single greatest hurdle in these

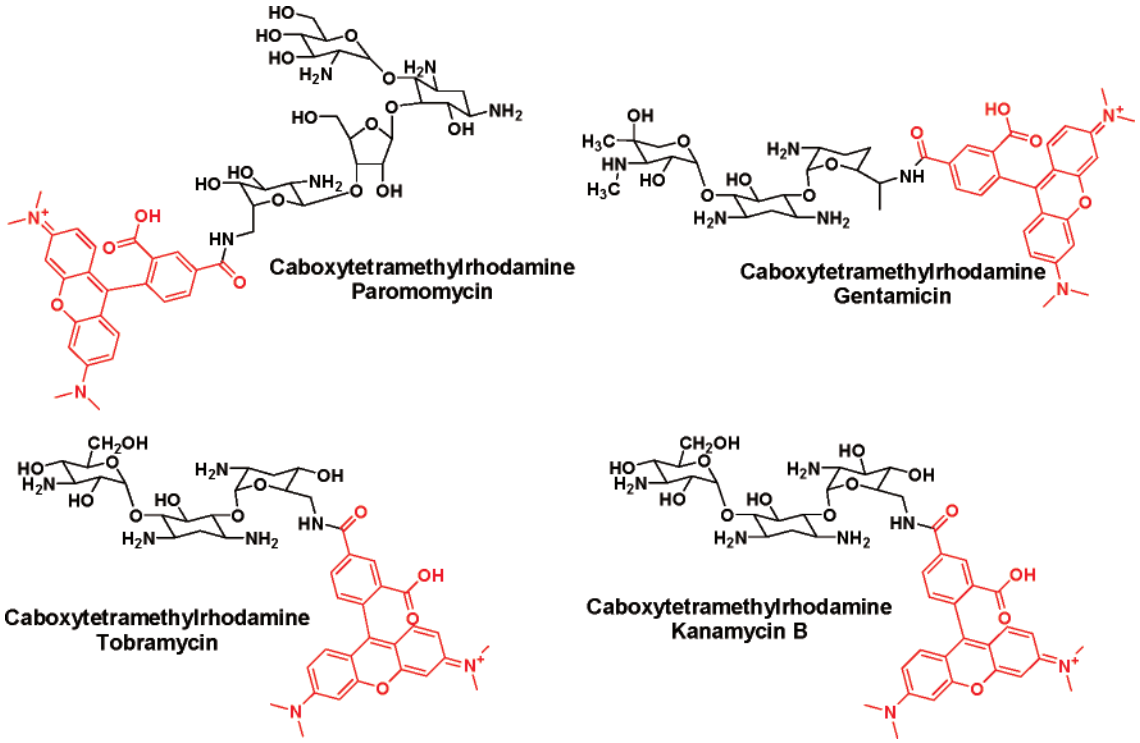


Figure 22. Fluorescently-conjugated aminoglycosides that have been used as ligands in displacement assays.

assays is removal of salt from the phosphate backbone of RNA. As ionic strength is known to have an important contribution to binding energy and selectivity,[208,210] the strength of ligand interaction and binding mode may be different from those observed in an ionic solution-phase experiment. While quite versatile and promising, ESI-MS analysis of small molecule−RNA interactions has yet to enjoy widespread use in discovery and validation platforms, presumably due to the type of specialized instrumentation and expertise required.

6.4. Fluorescent Ligand Displacement-Based Assays

Thus far, the methods that have been discussed monitor changes in the RNA, whether conformational changes or a shift in molecular weight. An alternative approach is the displacement of a fluorescently labeled ligand. Such strategies are robust for high-throughput screening and sufficiently reliable for detailed biochemical and biophysical studies. In addition, displacement of a fluorescently labeled ligand is a pseudo-tagless approach as the RNA and small molecule under investigation are unmodified. However, these methods cannot be used for de novo discovery platforms because they require a known ligand that can be fluorescently labeled.

Fluorescence anisotropy measures the rate of tumbling of a fluorophore in solution, and an unbound fluorescent ligand will tumble faster than the bound fluorescent ligand. For applications to RNA, fluorescently labeled aminoglycosides (Figure 22) are often used as displacement ligands. Displacement of labeled aminoglycosides has been used successfully to identify binders to internal loops;[147,245] however, the targeting of RNA hairpin loops would require a different screening ligand as the aminoglycosides bind poorly to this class of secondary structures.[147,225] Anisotropy assays have also been utilized for identification of compounds that disrupt RNA−protein interactions.[98,246] However, the large size of the ligands, particularly when coupled to a fluorophore, lessens the signal-to-noise ratios typically observed in anisotropy assays and in some cases precludes the use of anisotropy-based methods.[247]

In order to circumvent size limitations placed on anisotropy-based assays, fluorescence resonance energy transfer (FRET)-based techniques have been applied to the identification of small molecule ligands for RNA. In one example, the TAR RNA was end labeled with the donor, while a Tat peptide was labeled with an acceptor, yielding a FRET pair upon binding.[248] In the same TAR−Tat system the Tat peptide was labeled with both FRET pairs; upon binding, the dual-labeled Tat peptide binds in an extended conformation, thus reducing FRET efficiency.[247] Finally, investigation of the T-box antiterminator RNA was performed with both FRET pairs on the RNA, one which was end labeled and the second label located in the hairpin loop. The binding of small molecules altered the FRET efficiency and thus served as a reporter of binding.[72]

6.5. Small Molecule Microarrays

The printing of libraries of small molecules onto functionalized glass slides for high-throughput screening has been successfully used to identify novel protein ligands.[249−252] Small molecule microarrays allow for >10 000 small molecules to be screened simultaneously on a single glass slide, consuming a minimal amount of compound and macromolecule in the process. In its application to RNA, an initial proof-of-concept experiment was performed using a collection of aminoglycosides that were imprinted onto glass slides coated with a tetraethylene glycol linker terminating in succinimidyl succinate.[253] Application of 100 pmol of a fluorescently end-labeled A-site construct revealed that the RNA is able to recognize the ligands displayed from the surface but failed to reproduce the rank order of specificity observed from solution-phase binding assays. These results suggest that the surface, linker, and linker position can substantially affect the binding properties of the compounds,

a result confirmed by a similar study.[254] Although significant progress needs to be achieved for consideration as a screening tool, use of small molecule microarrays has potential for the rapid screening of large compound libraries against a large number of RNA targets.

6.6. Surface Plasmon Resonance and Isothermal Titration Calorimetry

The ability to incorporate a chemical "handle" selectively to either the 5′- or the 3′-end of an RNA construct affords a convenient method for attaching RNA to a surface plasmon resonance (SPR) sensor. In principle, SPR is an ideal format for evaluating binders: low amounts of RNA and compound are required, binding is measured by change in refractive index due to complex association rather than a conformational change, the attachment site can be far removed from any putative binding sites, and a known binder is not needed prior to investigation. SPR can provide both dissociation constants and on and off rates. Wong and co-workers championed the use of SPR for studying ligand−RNA interactions for the 16S A site,[204] RRE,[100] and various protooncogenes.[180] For now, however, the low throughput and high cost of SPR make it more appropriate as a validation tool rather than a discovery tool.

Although not suitable for a high-throughput screen, isothermal titration calorimetry (ITC) is an excellent method to obtain binding constants and thermodynamic parameters for small molecule−RNA interactions. This method has been applied to a wide variety of compounds and RNA constructs.[208,210,229,255,256] The limited aqueous solubility of certain small molecules has limited the application of ITC in some protein−ligand systems. However, most RNA ligands are charged, polar compounds; thus, ITC is an ideal technique for obtaining detailed thermodynamic parameters for small molecule−RNA binding.

6.7. RNase and Chemical Footprinting

Though all of the aforementioned assays provide robust methods for evaluating the strength of ligand−RNA interactions, only NMR methods allow for the precise determination of the ligand-binding site. Biochemical footprinting experiments are the standard follow-up assays for any of the above techniques because they allow for an independent confirmation of binding affinity and provide information about the exact residues that make up the ligand binding site.[257] In these assays the RNA is radiolabeled with ^{32}P on its 5′-end and then incubated with varying concentrations of ligand. After the binding proceeds to equilibrium, the RNA−ligand complex is subjected to degradation, either enzymatic or chemical, and the resulting RNA is analyzed by electrophoresis under denaturing conditions (Figure 21C). The binding site of the ligand is identified by the bases that are protected from cleavage in a dose-dependent fashion; binding affinities can be estimated by densitometry of the bands that appear on the gel. Though footprinting is an excellent validation assay, the time-consuming nature of the assay precludes its use in a high-throughput, discovery mode.

An alternative approach is in-line probing. In these experiments the radiolabeled RNA is allowed to incubate in the presence of varying concentrations of ligand for an extended period of time, on the order of days. Unstructured regions of RNA are inherently more chemically unstable than duplex regions due to spontaneous cleavage of the phosphodiester linkage by transesterification. Ligand binding and conformational changes can alter the spontaneous cleavage rate. This method been used primarily to monitor ligand binding to riboswitch RNAs.[110,258−262]

6.8. Assaying for Selectivity

The examples of aminoglycosides binding to a variety of target RNAs have demonstrated that selectivity, in terms of both secondary structure within a given RNA and one RNA target over others, is an elusive goal. Very few RNA binding compounds demonstrate any reasonable selectivity with the aminoglycosides being notoriously promiscuous.[203] Selectivity is a major issue that must be addressed to make the targeting of RNA with small molecules general and useful.

Assessment of the selectivity of small molecules for various RNA targets is most valuable if competing off-target RNAs are known. As estimates have suggested that ∼15% of the RNA within the cell is comprised of tRNA,[93] tRNAs are often chosen for this purpose. Nearly all of the methods described in the section above (use of fluorescently labeled RNAs, displacement of a fluorescent ligand, and foot-printing assays) are well suited for tRNA competition studies. In these experiments the standard binding assay conditions are repeated using a 100-fold (base) excess of commercially available *E. coli* tRNA.[263] Any deviation from the previously determined binding constant is attributed to off-target binding, thus providing an indication of RNA target selectivity.

Tor and co-workers developed this general semiquantitative selectivity assay while evaluating aminoglycosides and derivatives thereof for their ability to disrupt the Rev−RRE interaction (Figure 23A).[263] The IC_{50} for neomycin disruption of Rev binding was determined to be 7 μM, while addition of a 100-fold (base) excess of tRNA or calf thymus DNA resulted in IC_{50} values of 20 and 8 μM, respectively. The selectivity ratio was defined as the average IC_{50} value in the presence of DNA and tRNA divided by the IC_{50} in the absence of any competitor nucleic acids. These values allow one to infer the relative selectivity of a particular ligand as compared to others tested on the same conditions. For neomycin the specificity ratio is approximately 2.0, and both tobramycin and kanamycin A exhibited similar specificity ratios. Caution must exercised when interrupting the specificity ratio of particular ligand. For example, the affinity of some ligands, like the aminoglycosides as in the above example, are only modestly affected by the presence of competitor tRNA. However, this modest change in affinity when challenged with competitor tRNA is at odds with the ability of various aminoglycosides, particularly neomycin, to bind to a wide variety of RNA targets and secondary structures with approximately equal binding affinity (see section 4). Also, the affinity of some ligands for their respective targets has been shown to be largely *unaffected* by the presence of competitor tRNA;[264] however this does not mean that such ligands are absolutely selective for their respective targets. Despite these caveats, the tRNA competition experiment offers a simple method to quickly assess compound promiscuity.

A second approach for assaying ligand selectivity is to perform binding assays against other common secondary structures; such assays are commonly performed to determine the binding site of a ligand. For example, to determine the structural features important for aminoglycoside binding to RRE, 15 RRE constructs were utilized that varied in presence, size, and sequence of different secondary structural

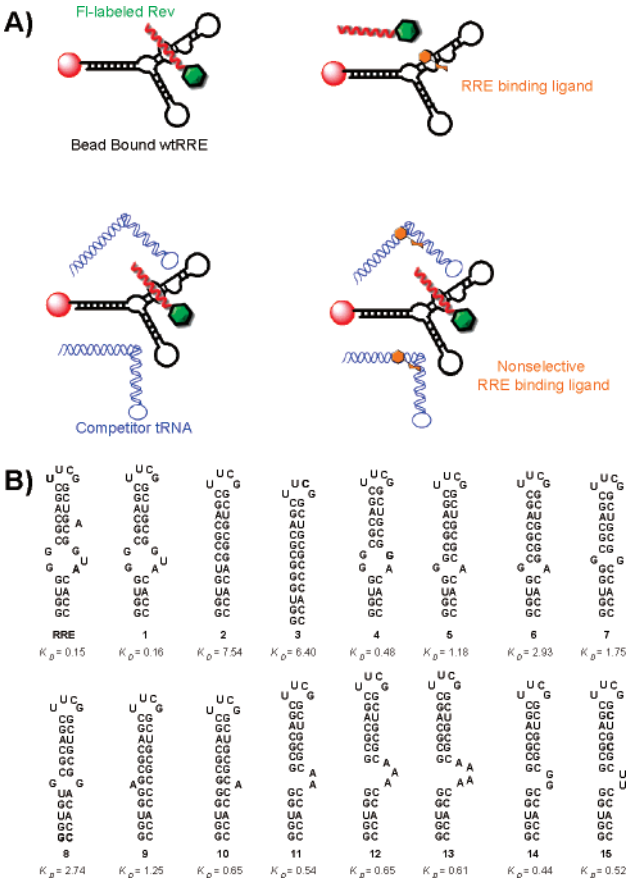


Figure 23. RNA-binding selectivity assays. (A) The solid-phase nucleic acid competition studies for evaluating ligand selectivity developed by Tor and co-workers. In the solid-phase assay, a ligand that disrupts the Rev−RRE interaction results in displacement of fluorescently labeled Rev (top) and an IC_{50} value can be determined. To test for selectivity, the same displacement assay is performed in the presence of a competitor nucleic acid, in this case tRNA (blue). If a compound binds the tRNA, then the ligand will be less effective at disrupting the Rev−RRE, resulting in a higher IC_{50} values. (B) Rando and co-workers systematically examined the structural requirements for the neomycin−RRE interaction. Examination of the binding constants for each RRE construct reveals that neomycin exhibits comparable affinity for a wide variety of RNA structures.

elements (Figure 23B).[101] Of course, given the huge number of possible sequence variations even for a very short RNA oligomer, it is not feasible to individually create every possible RNA sequence and then assess its binding to a compound. Thus, it has been difficult to get a true handle on the selectivity of RNA-binding compounds. While it can be said with some certainty that certain compounds are promiscuous binders, it is very difficult to say that a compound is truly selective for any RNA; it is simply selective within the panel of RNAs it has been tested against.

A recent approach which examines ligand-binding selectivity utilizes small molecule microarrays and in vitro selection of site-specifically randomized RNAs.[265] This approach allows one to simultaneously probe thousands of RNA−ligand interactions. In this proof-of-concept study, Disney and co-workers first attached kanamycin A (derivatized with a terminal alkyne at the 6′-position) to azido-functionalized agraose microarray slides. A library of 3 × 3 internal loop RNAs (labeled at their 5′-ends with ^{32}P) were then assessed for their ability to bind to the displayed ligand. The design of the RNA was such that only the six nucleotides within the 3 × 3 internal loop were randomized while the rest of the RNA structure remained constant. The bound RNAs were revealed by autoradiography, excised from the argrose matrix, and amplified by RT-PCR, and the sequence was obtained through cloning and sequencing. Disney and co-workers identified 16 internal loop sequences which bound the kanamycin A derivative with K_D values ≤22 nM. Interestingly, within the identified internal loop sequences a strong preference was observed (10 of 16) for an adenine to be adjacent to a cytosine. The authors suggest that the results gained from this method provide insights into the preferred RNA motifs for the each individual ligand, and such information can be used to guide target selection experiments or identify alternate RNA targets.

Another new approach which attempts to address the limitations of the above methods is the use of DNA microarrays to probe the entire transcriptome.[266] Two different methodologies have been developed for this purpose. The first method utilized tobramycin covalently attached to a solid support. Commercially available human liver polyA RNA was then incubated with the bead-bound compound. After extensive washing to remove nonspecifically bound RNAs, KOAc was used to elute specifically bound RNAs. After amplification of the eluted RNAs using a standard polyA RNA amplification kit, the identities of the bound RNAs were revealed using commercially available DNA microarrays. Tobramycin conjugated to beads through its 6″ primary amine pulled out 216 transcripts, while tobramycin that was randomly displayed on the beads pulled out 1164 transcripts; only 25 of the 216 transcripts originally identified were not detected by the randomly displayed tobramycin.[266] In a second approach using the same human liver polyA RNA, tobramycin−RNA interactions were identified by the ability of the aminoglycosides to prevent the hybridization process between the RNAs and the DNA probes.[266] This hybridization interference assay revealed only 18 genes with modified intensities; it seems likely that the hybridization interference assay is not reflective of the general RNA binding properties of tobramycin. Also, as shown by the use of aminoglycoside small molecule microarrays,[253] attachment of aminoglycosides (and presumably other RNA-binding small molecules) to solid support can significantly alter their binding properties. Although more work is needed to further validate this approach, it is currently one of the more comprehensive methods for evaluating selectivity of RNA-binding small molecules.

The remaining sections present an overview of efforts that have been made to identify compounds that specifically bind to certain RNA secondary structures, namely, stems, internal loops, bulges, and hairpin loops.

7. Stem Binding Compounds

Although many protein−RNA interactions occur in regions where the A-form helix is disrupted, the RNA stem does mediate some interactions. In fact, proteins which bind to RNA duplex regions can be grouped into superfamilies based on their double-strand RNA binding motif (dsRBM);[267] the various dsRBMs share a common α-*β*-*β*-*β*-α secondary structure.[268−270] The first α-helix and the loop regions connecting the *β*-sheets are responsible for binding to the major and minor grooves; recognition is achieved by binding to the 2′-hydroxyl groups and phosphate backbone.[42] Inhibition of some of these RNA−protein interactions could have therapeutic potential. For example, inhibiting the RNA-binding function of the viral protein RNA-dependent protein

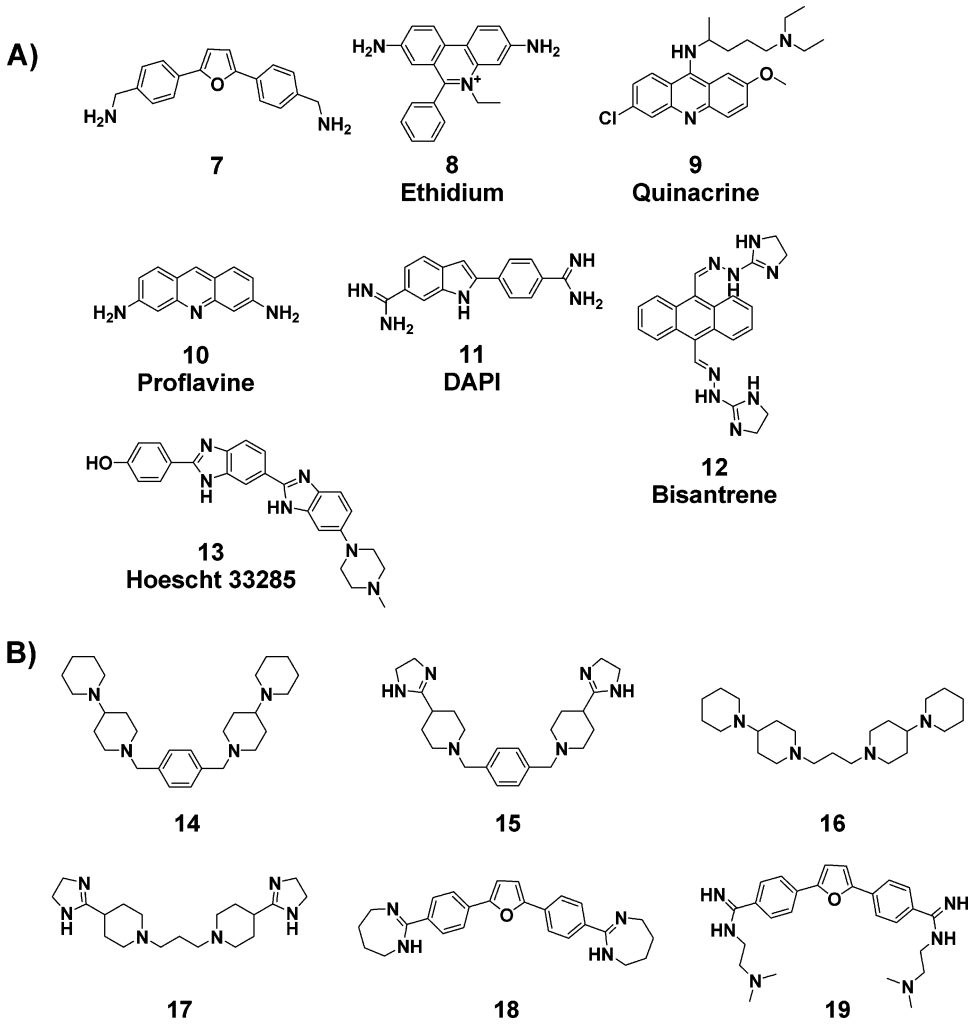


Figure 24. Stem binding compounds. (A) A representative sampling of DNA intercalating agents that were found to bind to RNA duplex regions. Five classes of DNA intercalating agents were assayed: an unfused aromatic scaffold (**7**), classic fused aromatic scaffolds (**8**−**10**), threading intercalators (**12**), and compounds with mixed modes of binding (**11** and **13**). (B) A representative sampling of ligands which bind RNA stems through electrostatic interactions.

kinase (PKR) has been proposed as a possible antiviral strategy.[271] PKR is activated by binding to double-stranded RNA and once activated phosphorylates eIF2 to inhibit host protein synthesis.[272−274] Thus, disruption of the PKR−RNA interaction would abolish kinase activity. However, development of such strategies is made difficult due to the inability of most RNA-binding small molecules to target duplex regions with affinity and specificity. As detailed in the sections below, some progress has been made toward the development of compounds that bind to fully duplex regions of RNA (RNA stems).

7.1. Stem Binding Small Molecules Through Intercalation

Initial investigations toward identifying RNA-binding small molecules began with known DNA-binding compounds. As the binding mode, sequence preference, and strength of interaction were previously determined for DNA, structure−activity relationships to establish properties necessary for RNA binding and selectivity could readily be determined. DNA-binding ligands can be broadly classified as either groove binders or intercalators.[275] While DNA minor groove-binding ligands generally exhibit negligible binding affinities for RNA, many DNA intercalators (compounds **7**−**13**, Figure 24A) are able to associate with RNA, although in nearly all cases binding to DNA is most favorable. Further investigation into the different types of intercalation reveals that not only is threading intercalation possible for RNA duplexes, but threading intercalators can bind with comparable affinities and form longer lived complexes with the RNA duplex than classical intercalators;[276] threading intercalators are discussed in more depth in section 7.3.

7.2. Stem Binding Small Molecules Through Ionic Association

The development of nonintercalative stem binding compounds was spawned from the desire to obtain compounds with specificity for RNA over DNA; such studies focused on identification of ligands which bound more tightly to RNA duplexes than DNA duplexes. Results from these initial experiments revealed that flexible scaffolds with charged centers preferentially bind RNA duplexes over their DNA counterparts (compounds **14**−**17**, Figure 24B).[277] Selectivity was postulated to arise, in part, from the inability of the flexible scaffolds to pack properly in the minor groove of a DNA duplex, a property that is essential for DNA minor groove binding.[277,278] These same properties were then expanded to the diphenylfuran ligands (**18** and **19**), where it was found that large, charged substituents resulted in an ionic rather than an intercalative binding mode (compare **18** and

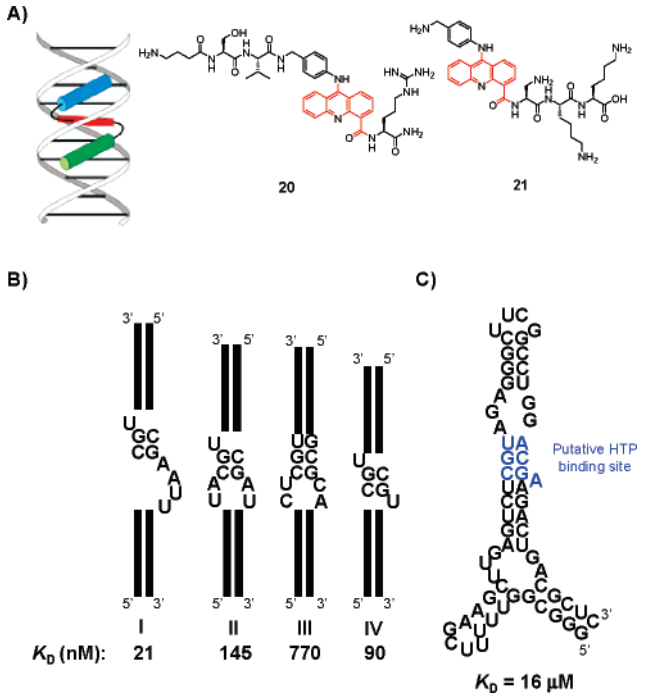


Figure 25. Helix-threading peptides. (A) Illustration of a threading intercalator interacting with RNA, and the structures of two helix-threading peptides (HTPs). Highlighted in red is the core intercalating scaffold. Image courtesy of Professor Peter A. Beal. (B) The proposed binding site for each of the in vitro-selected RNA aptamers is shown with their respective binding affinities for **20** shown below the structures. (C) HTP **21** was chosen to evaluate the ability of HTP ligands to target biologically relevant RNAs, as exemplified by helix 22 of the *E. coli* 16S RNA. Shown in blue is the binding site for **21**.

19 with **7**).[279] In an alternative strategy, Miller and co-workers successfully devised a dynamic combinatorial library strategy capable of generating substantial chemical diversity. In a preliminary study, the authors reported the discovery of one compound (from a theoretical library size of 11 325) that appears to exhibit exceptional selectivity for the HIV-1 frame shift regulatory sequence as compared to other nucleic acids of similar sequence and size, possibly through binding to the RNA stem.[280]

7.3. Stem Binding Small Molecules Through Threading Intercalation

Although the A-form helix does not in an absolute sense prevent recognition of the discriminatory edges of the base pairs that comprise a duplex region, selective recognition of the stem will likely require exploiting more subtle attributes due to the lack of suitable binding pockets present within duplex regions as compared to other secondary structures. Previous work has demonstrated that the rate of base pair opening is enhanced by proximity to other RNA secondary structures;[281] such enhanced conformational dynamics provide an avenue toward binding that threading intercalators may be able to exploit. In threading intercalation the intercalating moiety is stacked within the helix such that substituents attached to the intercalating system are placed on opposite sides of the duplex (see Figure 25A). This unique binding mode, coupled with its longer lifetime of association, provides a tractable avenue for the development of duplex-selective ligands. RNA-binding selectivity could arise from threading intercalators preferentially binding to duplex regions at sites with greater rates of base pair opening. Beal and co-workers explored the potential of threading intercalators using the 9-anilinoacridine-4-carboxamide scaffold with various peptide substituents appended at the 4- and 9-positions.[271,282] Previous work with DNA duplexes has demonstrated that this scaffold does indeed bind with a threading mode of intercalation with the 4-position placed in the major groove,[283,284] although there is conflicting evidence.[285] Using solid-phase synthesis and on-bead screening techniques Beal and co-workers developed an effective platform to obtain and evaluate helix-threading peptides (HTP).[271,282] In order to investigate the preferred RNA binding site of HTPs a representative ligand (**20**, Ser-Val-*Acr*-Arg) was subjected to in vitro selection experiments.[286] The results from these selection experiments defined a consensus sequence that comprises the binding site of **20**, wherein a single bulged uridine or G−U base pair was found directly above a G−C, C−G base pair, which was followed by either an internal loop or bulged region (Figure 25B, RNA constructs **I**−**IV**). The authors propose that the G−C, C−G base pairs likely provide the site of intercalation of the acridine moiety, which is consistent with observations made for DNA duplex binding.[287] It remains to be determined if the differing secondary structures above and below the G−C, C−G base pairs alter the base pair "breathing" rates or simply enhance affinity by interacting with the peptide appendages. However, it is clear that the nature of the secondary structure surrounding the site of intercalation plays a role in defining ligand affinity, as the affinity for **20** varied by >30-fold between the different selected aptamers. This may suggest that different sequences and types of secondary structure likely affect the opening rate of base pairs at the intercalation site differently.

Although the selection experiments were able to provide insight into the preferred binding site of **20**, no information concerning the groove location of the peptide substituents could be gathered. Attachment of EDTA−Fe, a commonly employed hydroxyl radical generating reagent for chemical foot printing, to either the N- or C-terminus of **20** allowed for the determination of each substituent's groove location.[288] From these experiments it was determined that the N-terminus projects into the minor groove, while the C-terminus projects into the major groove of the in vitro-selected RNA **I**. Mutational analysis demonstrated that the size of the bulged structure below the site of intercalation is critical for binding, as a single base bulge experienced an 18-fold drop in binding affinity and deletion of the bulged structure resulted in no observed binding. Interestingly, analogous experiments performed with a C-terminal-modified ligand exhibited an altered binding orientation for the smaller bulge sizes. When the bulge size was decreased to one or two nucleotides a mixed binding mode was observed; that is, the smaller bulge size relieved the preference for inserting the C-terminus in the major groove by allowing the C-terminus to project into either the major or minor groove. These results demonstrate that both ligand and RNA structure dictate the binding orientation and further suggest that subtle differences in RNA structure can be used to discriminate between RNA targets. However, the data also hint that optimizing the substituents emanating from the intercalator scaffold for a particular RNA target may be difficult as continual modification of the ligand may alter its binding orientation.

Various helix-threading peptide ligands have been evaluated for their ability to bind to biologically relevant RNAs. Compound **21** was tested for its binding site specificity to

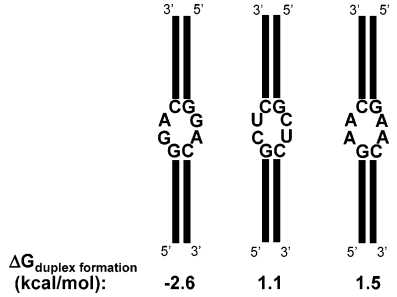


Figure 26. Energetics of internal loop formation. Shown are the three symmetric 2 × 2 internal loops and the impact they have on the free energy of duplex formation.

helix 22 of the *E. coli* 16S RNA.[289] Helix 22 of this RNA features a binding site that is similar to those of the in vitro-selected aptamers except that above the intercalation site a U−A base pair is present rather than a single base U bulge or G−U base pair (Figure 25C). Chemical foot-printing experiments revealed that **21** does indeed bind to the expected region with an affinity of 17 μM, which is 3 orders of magnitude less potent than binding to RNA **I**. The U−A base pair exhibited a significant degree of conformational freedom as inferred from its hyper-reactivity during foot printing. Such conformational freedom is consistent with requirements mandated from the selection experiments. Further support is derived from the loss of binding that is observed when the internal loop above the site of intercalation is replaced with a duplexed region. However, during an investigation of HTP ligands for their ability to inhibit the RNA−PKR interaction, an analogue of HTP **20** (with the terminal arginine replaced by lysine) was found to inhibit this RNA−protein interaction with an $IC_{50} = 37$ μM even though the proposed intercalation site is absent.[271] Although the selectivity of HTP ligands for the consensus binding motif (a G−C, C−G intercalation site with conformationally free secondary structures on either side) remains to be determined, use of threading intercalators to achieve selectivity for duplexed regions is a novel solution for targeting RNA. Recent applications of this work include the synthesis of macrocyclic HTP ligands via ring-closing metathesis and the demonstration that these compounds show enhanced binding affinities over their linear counterparts.[290]

8. Internal Loop Binding Compounds

Internal loops are a ubiquitous secondary structure element in RNA. Significant effort has been put forth to understand the energetic forces that determine stability, structure, and dynamics of internal loops. Detailed investigation of the free energy of formation for internal loops has revealed a strong sequence dependence on internal loop stability. For example, for 2 × 2 internal loops the free energy is known to vary by more than 5 kcal/mol depending on the identity of the nucleotides within the internal loop.[291] Within this variance certain sequences have been shown to stabilize duplex formation (5′-GA-3′ × 5′-GA-3′), while others significantly destabilize duplex formation (5′-CU-3′ × 5′-CU-3′) (Figure 26).[291] In addition to the sequence context of internal loops, the symmetry and size of internal loops further contribute to the variability in the free energy of duplex formation. Asymmetric internal loops are destabilizing toward duplex formation.[292] For example, asymmetric internal loops are typically unstructured in solution, and their inherent flexibility has been evoked to explain their absolute destabilizing nature.[292] Given the known sequence dependence on the free energy of internal loop formation, detailed studies on the size dependence of internal loops are exceedingly difficult. The largest internal loop sizes to be systematically studied thus far are 3 × 3 internal loops.[292] From the limited information available, 1 × 1 internal loops appear to be roughly neutral with regard to free energy of duplex formation,[293] while in larger symmetric internal loops a greater entropic penalty is paid for loop formation. This entropic penalty may be compensated for by enthalpic contributions due to non-Watson−Crick base pair formation.[291,292] This apparent entropic/enthalpic compensation gives rise to the large variability in free energy observed for 2 × 2 and 3 × 3 internal loops. These thermodynamic data could be potentially exploited in the search for loop-size-selective ligands. For example, if a compound is known to bind through an intercalative mode, the targeting of a stabilized internal loop (such as a 5′-GA-3′ × 5′-GA-3′, $\Delta G = -2.6$ kcal/mol) rather than a destabilized sequence (such as 5′-AA-3′ × 5′-AA-3′, $\Delta G = 1.5$ kcal/mol)[291] will likely be preferable. Superficially, both sequences would seem favorable for stacking, but the thermodynamic data suggest formation of G−A base pairs for the stabilized sequence, which may imply a more favorable environment for stacking interactions.

Identification of small molecule ligands for three particular RNA internal loops has been extensively pursued: the 16S A-site RNA, the Rev response element (RRE) RNA, and the thymidylate synthase mRNA. A summary of this work is presented in the sections below. For the biological background on each of these systems, the reader is referred to section 4 of this review.

8.1. Small Molecule Binders to the 16S A-Site RNA

Of all the RNA targets that have been explored for therapeutic intervention, the A-site 16S ribosomal RNA is by far the one with the most success stories. The term 16S refers to the entire helix, while the A-site defines a very specific region. As it is an antibacterial target, compounds that bind the 16S ribosomal RNA have the added advantage of a convenient cell-based assay for assessment of efficacy. Although the aminoglycosides are potent antibiotics, they possess several undesirable properties for clinical use, including the emergence of resistant bacterial strains, poor pharmacological profiles, and off-target effects that lead to oto- and nephrotoxicity.[294] Many groups have attempted to derivatize or create small molecule mimics of the aminoglycosides to circumvent the aforementioned problems. Unfortunately the issue of specificity, in terms of secondary structure or even between RNA targets, is not directly addressed in many of the manuscripts presented in this section; cell culture antibacterial assays are often used as a surrogate for in vitro selectivity assays.

8.1.1. Bifunctional Aminoglycoside Derivatives

One of the main culprits in the decreased efficacy of aminoglycosides is the prevalence of aminoglycoside-modifying enzymes.[295] Bacterial resistance is conferred by the ability of these enzymes to acylate, phosphorylate, or ADP-ribosylate the aminoglycosides as the modified compounds have a decreased affinity for the 16S A-site RNA. To circumvent the problem of aminoglycoside-modifying enzymes, Wong and co-workers synthesized a library of

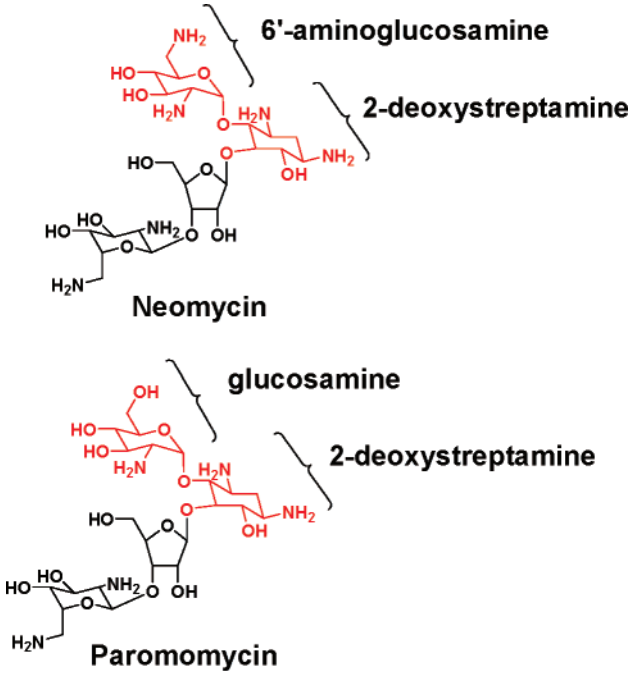


Figure 27. Minimal constructs for A-site binding. Highlighted are the neamine and paromamine ring systems from neomycin and paromomycin, respectively. Both neamine and paromamine have been shown to bind the 16S A-site ribosomal RNA, thus representing a minimal RNA-binding scaffold.

"bifunctional aminoglycosides" by dimerizing neamine.[296] Neamine was chosen as it represents the minimal portion of an aminoglycoside that still retains antibiotic activity (Figure 27). Although it is not entirely clear how such dimerized neamine ligands would preclude susceptibility to aminoglycoside-modifying enzymes, enhanced RNA affinity was anticipated to arise from the ability of the dimeric neamine to bind simultaneously to the internal and hairpin loop regions. Using SPR to characterize the ligand−RNA interactions, neamine was found to bind to the A-site RNA with a 2:1 stoichiometry and a K_D of 10 μM for each site. A series of alkyl-linked neamine dimers of varying tether lengths synthesized through either amide or carbamate linkages was evaluated for binding to the A-site RNA construct (Figure 28). The most potent binder from this alkyl series, compound **22**, displayed a modest improvement in affinity over neamine ($K_D = 0.8$ μM versus $K_D = 10$ μM for neamine) and retained a 2:1 stoichiometry of binding. The authors indicated that the enhanced affinity was due to contributions from the linker rather than the dimeric nature of the compounds. Subsequent efforts led to the design of more flexible, hydrophilic amino alcohol linkers, and compounds incorporating these had substantially enhanced binding affinity. The diaminobutane-linked neamine dimer **23** showed a 250-fold improvement in binding affinity (K_D = 40 nM) and bound with a 1:1 stoichiometry. In vitro characterization of various neamine dimers led to the determination that several dimers exhibited antibiotic activity in bacterial cell culture assays and were indeed poor substrates for a variety of aminoglycoside-modifying enzymes.

Encouraged by the success of the neamine dimers, He and co-workers probed the generality of dimerizing simple sugars by exploring dimers of 2-deoxystreptamine (DOS), 6′-aminoglucosamine, and glucosamine, the core building blocks of the neamine and paromamine ring systems.[297] Homodimers and heterodimers of all three sugars were synthesized with tethers of varying linker length and composition (**24**−**28**, Figure 28). Binding to the 16S A site was evaluated by ESI-MS, from which it was determined that all of the dimerized compounds showed improved affinity as compared to their respective parent monomer. Comparing identically linked compounds, homodimers of 6′-aminoglucosamine (**24**, Figure 28) bound better than homodimers of DOS (**25**), which bound better than homodimers of glucosamine (**26**). Analysis of the DOS heterodimers revealed a significant decrease in binding affinity. With respect to linker length, the binding affinities for 6′-aminoglucosamine−DOS and DOS-glucosamine dimers mirrored the affinities of the homo-DOS and homo-glucosamine dimers, respectively, although one 6′-aminoglucosamine−DOS conjugate showed equal binding affinity relative to the similarly linked homodimer of 6′-aminoglucosamine. The general trend was observed that longer linkers generally provided dimers with better binding affinities. Although none of the compounds reported by He and co-workers surpassed the affinity of the neamine dimers (section 8.1.1), the data presented suggest the general utility of dimerizing compounds with weak binding affinity to give rise to compounds with markedly enhanced affinity.

8.1.2. Neamine Derivatives

Not all aminoglycosides are substrates for the aminoglycoside-modifying enzymes. For example, amikacin is known to be a poor substrate for various aminoglycoside-modifying enzymes; its aminohydroxybutyryl substituent is thought to sterically preclude binding to various modifying enzymes.[298] Mobashery and co-workers devised a computational strategy to identify derivatives of neamine with enhanced binding affinity and reduced susceptibility to aminoglycoside-modifying enzymes.[224] Starting from the NMR structure of paromomycin bound to the A site, rings III and IV were computationally deleted and the resulting structure was used as the template for in silico docking experiments. Sampling the Cambridge Structural Database (CSB) and National Cancer Institute (NCI) 3D database provided >250 000 compounds for the docking experiments; compounds were selected that bound near the N1- and O6-positions of the neamine ring system. The in silico hit compounds were used to guide the design of seven neamine derivatives. As the (*S*)-4-amino-2-hydroxybutyryl side chain of amikacin is essential in resisting the effect of aminoglycoside-modifying enzymes, the N1-position of neamine was equipped with the same side chain as well as aniline-containing side chains. Various aliphatic amine substituents were placed at the O6-position to make contact with the phosphate backbone.

In order to establish the effectiveness of the synthesized compounds, a 5′-fluorescein-labeled A-site construct was used. For the parent neamine, two cooperative binding events were observed: the first cooperative binding site yielded a K_D of 19 μM, while the lower affinity site bound with an affinity of ∼4 mM. In contrast, nearly all synthesized derivates bound to a single site with no signs of cooperative binding, and several exhibited improved affinity over neamine (compounds **29**−**35**, Figure 29A). Addition of the aminohydroxylbutyryl substituent resulted in enhanced binding affinity in all cases. Also, the addition of the diaminoalkane moiety at the O6-position improved the binding affinity as much as 12-fold. Given the enhanced affinity imparted

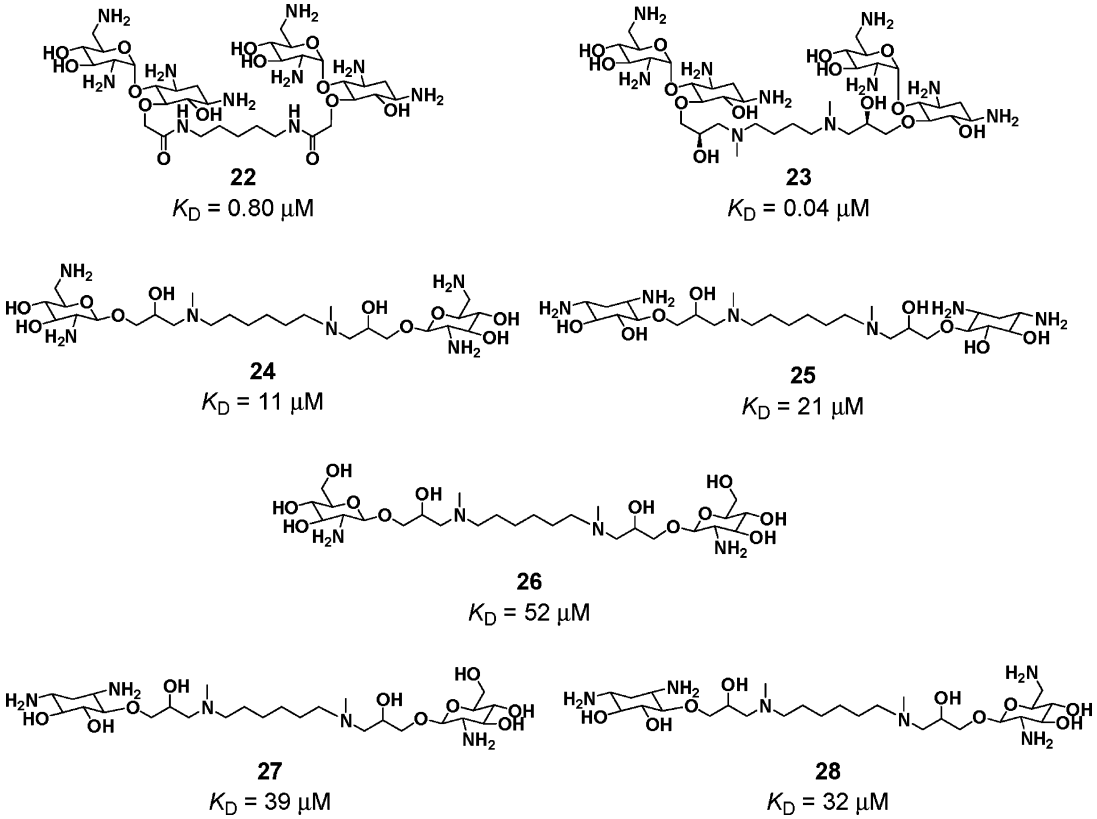


Figure 28. Dimerized aminoglycoside building blocks and their affinity for the 16S A-site RNA.

by the aminohydroxybutyryl substituent, it would have been interesting to compare the binding affinities to amikacin, as many of the derivatives synthesized are more analogous to O6-substituted amikacin compounds than neamine derivatives. As shown in Figure 29B, the crystallographic structural determination of **31** bound to the A-site RNA revealed that the neamine core retains its mode of binding, as intended by the design strategy.[299] Consequently, the aminohydroxybutyryl substituent projects up the helix along the major groove where a series of water-mediated hydrogen bonds are formed with several base pairs in the major groove. The aliphatic amines are also involved in a number of interactions including base-directed contacts between the secondary amine and phosphate backbone contacts by the primary amine. The binding interactions captured in the **31**-16S A-site RNA complex are nearly identical to those of the amikacin−A-site complex which was recently disclosed (Figure 29C).[300] As designed, several of the synthesized neamine derivatives exhibited enhanced antibiotic potency and were poor substrates for a variety of aminoglycoside-modifying enzymes.[224]

8.1.3. Neamine Mimics

Given that the neamine scaffold is highly susceptible to modification by resistance-mediating enzymes, Boons and co-workers sought to develop various disaccharide units which could bind to the A site, thus functioning as neamine mimics.[301] A large number of disaccharides were synthesized from monosaccharides containing a varying number of amine groups in the hopes of discovering a lead compound more amenable to derivatization than neamine. A library of 24 α(1−3)-, β(1−3)-, α(1−4)-, or β(1−4)-linked disaccharides containing 2−4 amino groups was synthesized and evaluated (by ESI-MS) for binding to an A-site construct. Of the 24 disaccharides synthesized nearly all compounds bound to the A-site construct, although with high micromolar affinity; only six disaccharides bound with dissociation constants lower than 50 μM (compounds **36**−**41**, Figure 30). In accordance with the known dependence on electrostatic interactions, all compounds that bound with K_D values below 50 μM contained three or four amines. Other general library trends showed α-linked disaccharides generally bound with tighter affinity than the corresponding β-linked compounds, and 1−4 linkages appear to be preferred to 1−3 linkages. Importantly, the best compound of the disaccharide series, **40** (K_D = 11 μM), bound with comparable affinity to neamine. After adjusting the φ and ψ angles of **40** to match that of rings I and II in the paromomycin-bound structure of A-site RNA followed by superposition the ligands, the authors determined that three out of four amino groups were oriented properly for binding. These results demonstrate that amine-functionalized disaccharides can exhibit comparable binding affinity to the A-site RNA as neamine and possess the capacity to bind in a similar orientation.

Boons and co-workers reasoned that although neamine possesses modest affinity for the A site and weak antibiotic activity, the attachment of various other sugar moieties to generate the aminoglycosides substantially increases the binding affinity and antibiotic activity of neamine.[302] Thus, in an effort to increase the affinity of **37** and **40** for the A-site RNA, a series of trisaccharide derivatives was synthesized by linking either β-D-ribose or 4-amino-6-deoxy-D-glucose to key positions in the initial lead compounds. All trisaccharide derivatives of **37** bound with markedly worse affinity than the parent disaccharide (compounds **42** and **43**, Figure 30). Attachment of β-riboside to the 3-position of **40** also attenuated binding (compound **44**). These results are surprising considering that β-riboside attachment at the 3-position

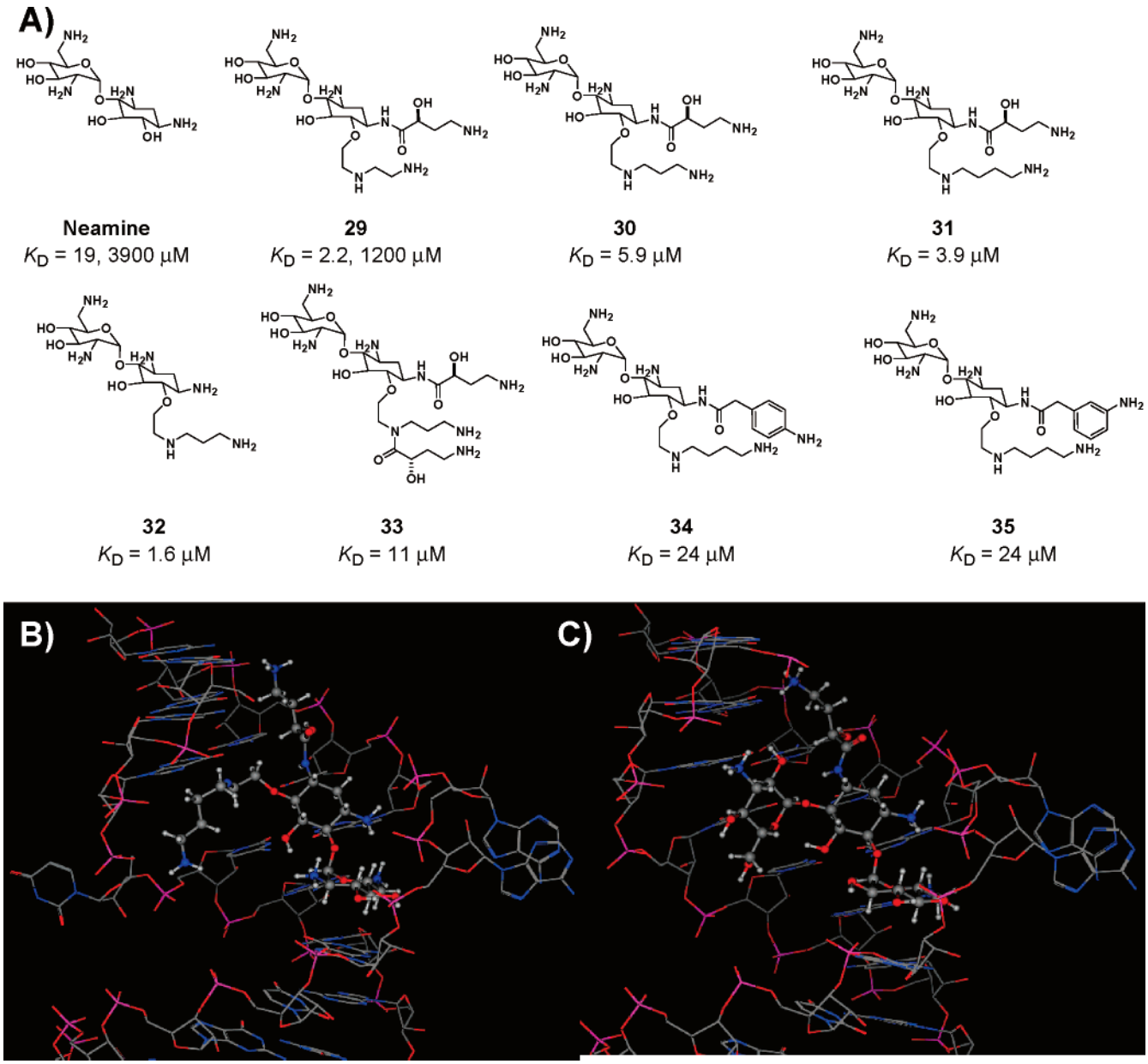


Figure 29. Evaluation of "designer antibiotics" binding to the 16S A-site RNA. (A) Neamine derivatives that were synthesized in the hope of reducing susceptibility to aminoglycoside-modifying enzymes and having enhanced affinity for the A site. Two binding events were detected for neamine and **29**, as such the dissociation constants for both sites are presented. (B) Crystal structure of **31** bound to the A site. (C) Crystal structure of amikacin bound to the A site.

is the natural sugar linkage in neomycin and paromomycin. However, substitution of 4-amino-6-deoxy-D-glucose led to derivative **45** with enhanced binding affinity, $K_D = 2\ \mu M$. The SAR data indicate that α(1−3) and α(1−4) likely exhibit different modes of binding to the A site. Although these results indicate that amino-oligosaccharides can be synthesized with affinities rivaling that of some aminoglycoside−16S A-site interactions, the data also serve to highlight the privileged nature of the neamine scaffold with respect to A-site binding.

8.1.4. Paromomycin Derivatives

Examination of the crystal structure of paromomycin bound to the 16S A site reveals that ring I penetrates deep within the major groove and ring II mediates key interactions with several bases, while rings III and IV mediate more general interactions below the internal loop region (see Figure 20A). On the basis of the paromomycin complex, Westhof, Hanessian, and co-workers determined that the C2″ hydroxyl of ring III was ideally suited for derivatization.[303] Modification of the C2″ hydroxyl to an ether-linked ethylamine (Figure 31A, **46**), propylamine (**47**), or pyridine (**48**) resulted in only minor changes in the observed binding affinity to the A-site RNA. Structural determination of **47** and **48** complexed with the A-site RNA revealed that the neamine ring system retains its normal mode of binding.[303] However, rings III and IV bind in a radically different orientation as compared to the paromomycin complex. As observed in crystal structures of **47** and **48**, ring IV is rotated ∼90° relative to the paromomycin structure (compare Figure 31B and C); this novel orientation is believed to result from the β-D-ribofuranosyl linkage which is rotated by ∼40° and allows an alternate sugar pucker conformation. Although the newly appended substituents project out into solution, mid-

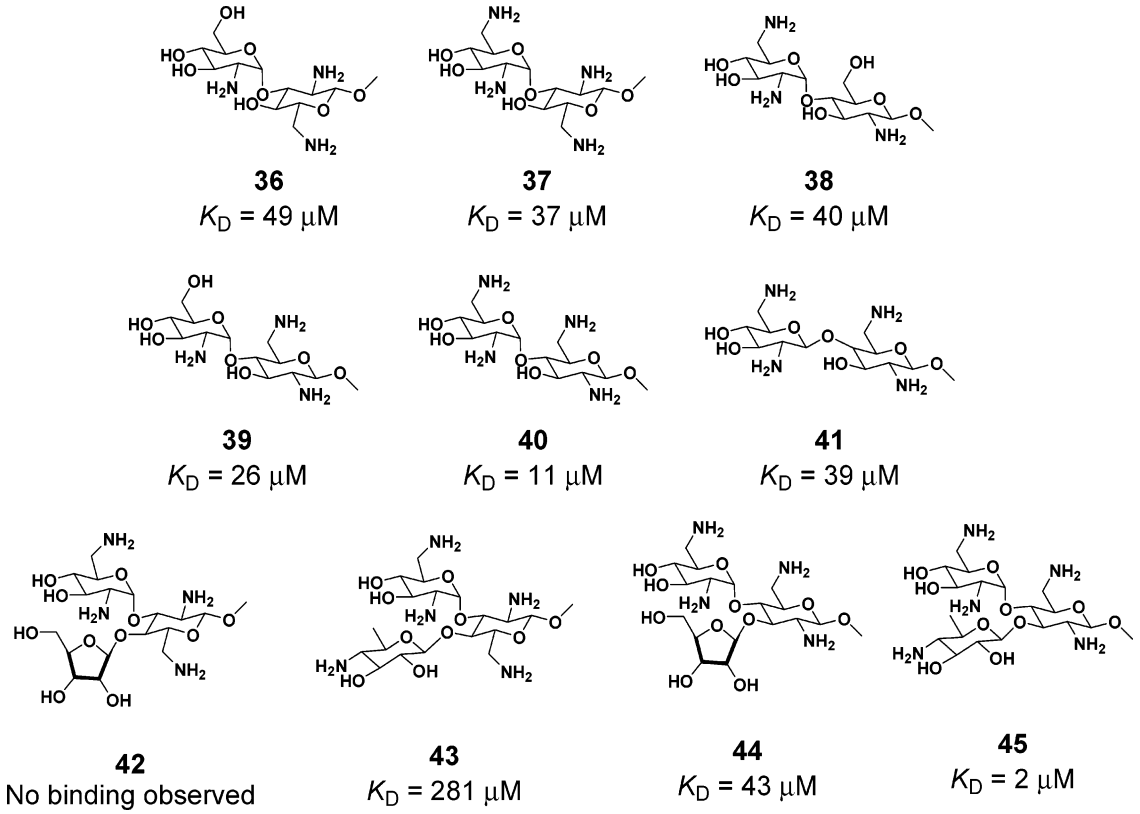


Figure 30. Linked oligosaccharides and their affinities for the 16S A-site RNA.

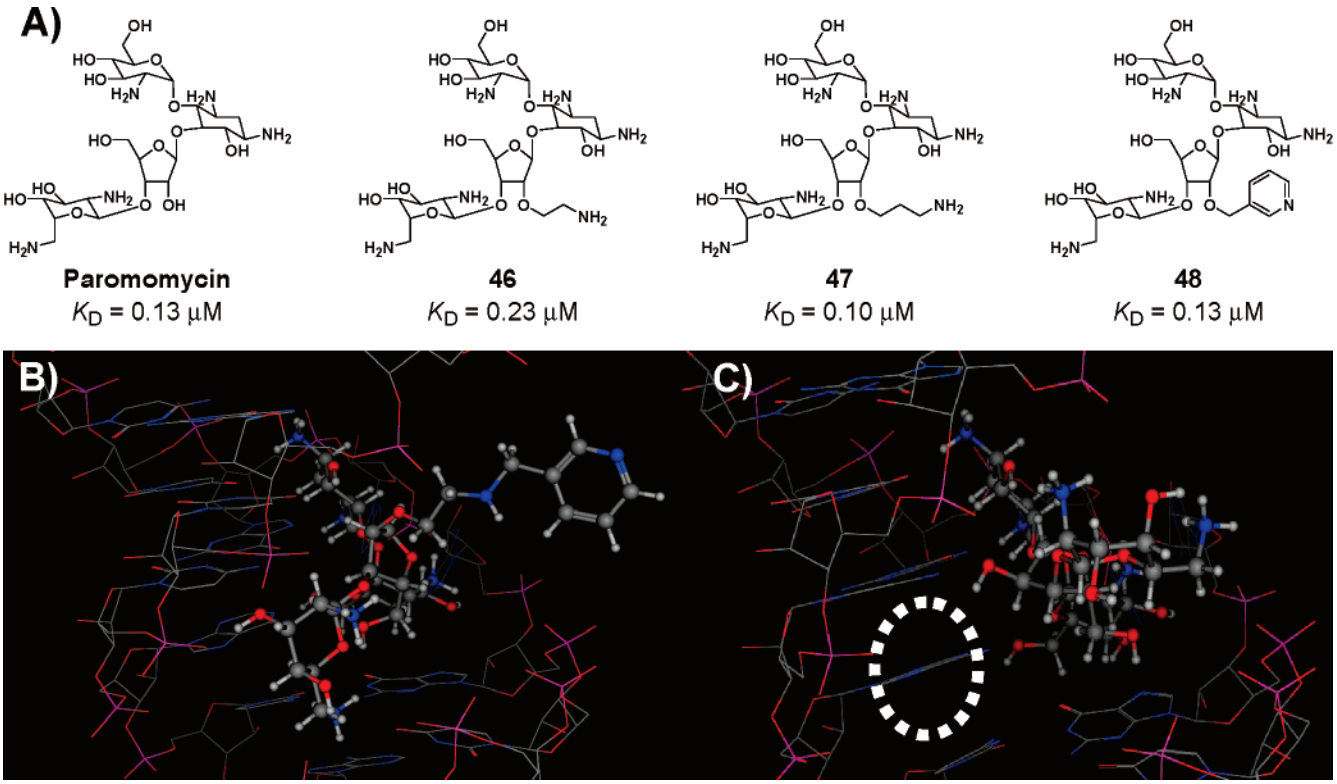


Figure 31. Paromomycin derivatives functionalized through the C2″-position. (A) Paromomycin derivatives and their binding affinities for the 16S A-site RNA. (B) Crystal structure of **48** bound to the A site. (C) Crystal structure of paromomycin bound to the 16S A-site RNA. The white dashed circle indicates where ring IV of **48** projects; this ring is rotated ~90° in the paromomycin complex.

nanomolar binding affinities are still retained by the paromomycin derivatives because of the new network of hydrogen bonds formed by rings III and IV.

In order to identify mimics of aminoglycosides that are more synthetically tractable for chemical optimization, Ding and co-workers synthesized a library of heterocyclic 2-deoxysteptamine (DOS) conjugates substituted at the 4-position of the DOS.[304] Alkylation of a protected form of DOS with various imidazole, triazole, pyrazole, purine, and pyrimidine heterocycles afforded 15 neamine mimics. ESI-MS was used

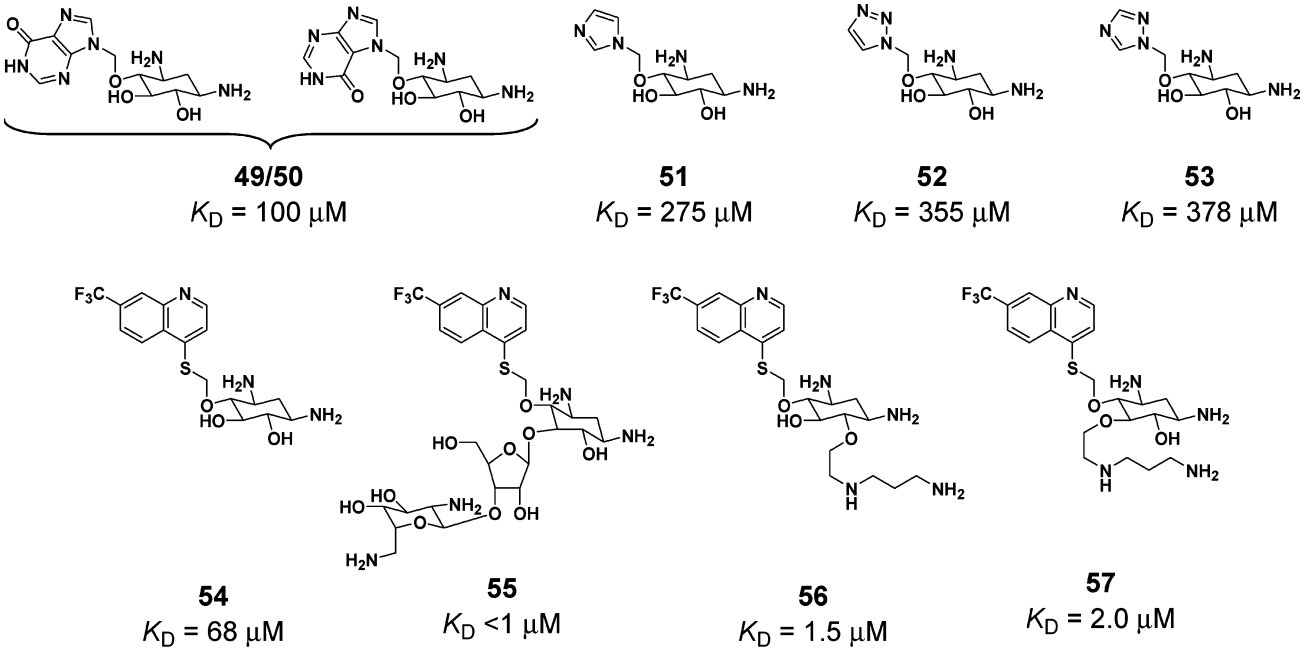


Figure 32. Heterocyclic paromomycin derivatives and their affinity for the 16S A-site RNA.

to determine the binding affinity of the synthesized compounds for the 16S A-site construct. All of the synthesized compounds bound with high micromolar affinity, which is significantly worse than neamine (compounds **49**−**53**, Figure 32). The best of the series was a mixture of two compounds **49**/**50**, which bound with 4-fold weaker affinity than neamine, with a K_D of 100 μM. In a subsequent effort, Ding and co-workers examined the binding affinity of an expanded library of heterocyclic-substituted paromomycin derivatives, again substituted through the 4-position of DOS.[305] As the synthesis of the intended paromomycin derivatives is lengthy, various heterocycle-substituted DOS compounds were evaluated for binding to the A-site construct via ESI-MS. From the 19 compounds evaluated, the [(7-trifluoromethyl)-4-quinolinyl]sulfanyl-substituted DOS conjugate **54** showed the best binding affinity as compared to the other heterocycles tested with a K_D of 68 μM. From these results the heterocycle of **54** was chosen to replace ring I of paromomycin; elaboration of **54** to the corresponding paromomycin derivative (**55**) provided a compound that exhibited an improved binding affinity relative to **54**, K_D < 1 μM. The binding affinity of **55** was weaker than that of paromomycin (K_D = 110 nM) but better than those of several aminoglycosides including apramycin, bekanamycin, and tobramycin (K_D = 2 μM).

Although the above work demonstrates that heterocyclic rings can function as a suitable replacement for ring I of various 4,5-disubstituted aminoglycosides, progress has yet to be made regarding replacement of rings III and IV with more medicinal chemistry friendly pharmacophores. Migawa and co-workers recognized from previous literature reports that simple alkylamines can mimic carbohydrate ring systems.[224,306,307] In an attempt to synthesize "carbohydrate-free" aminoglycoside mimics, alkylamine derivatives of **54** were synthesized and evaluated for binding to the A-site construct by ESI-MS.[308] Addition of an alkylamino substituent at either the 5- or the 6-position resulted in a greater than 30-fold improvement of binding relative to **54** (compounds **56** and **57**, Figure 32).

8.1.5. Acyclic DOS Mimics

With similar goals of discovering aminoglycoside mimics in which one or more of the sugar moieties are replaced with more synthetically tractable pharmacophores, Hermann and co-workers sought to develop alternative scaffolds for the DOS ring structure. Thus, 37 6′-aminoglucosamine derivatives were synthesized where the DOS moiety was mimicked using two separate acyclic scaffolds.[309] In one series it was envisioned that conservation of the amino group at the 1-position may properly orient the scaffold in the binding site, while additional binding affinity would be provided by varying the substituents at R1 and R2 (Figure 33A). The second series of acyclic DOS mimics was intended to probe the tolerance of the 5-position for nonsaccharide moieties, which ideally would replace rings III and IV present in paromomycin and neomycin. In order to prioritize which library members would be subjected to detailed investigation, all compounds were screened in bacterial and mammalian cell-free in vitro translation assays. Inhibitors of bacterial translation with an IC_{50} < 250 μM were considered hits, and their binding affinity to the A-site RNA construct was determined using a previously developed 2-aminopurine assay. Comparison of the derived IC_{50} values for translation inhibition between the bacterial and mammalian systems served as a measure of compound selectivity. On the basis of the biological activity profiles, three groups of compounds emerged. The first group exhibited weak, nonselective translation inhibitory properties and failed to bind the bacterial A-site construct, the second showed nonselective translation inhibition and bound the bacterial A-site construct with low micromolar affinity (K_D = 2.3−4.9 μM), and the final group consisted of one compound which selectively inhibited translation but exhibited no detectable binding affinity for the A site. For the compounds which bound the A-site construct, it appears that intercalation or hydrophobic interactions are likely responsible for the observed low micromolar binding affinities. For example, the biphenyl (**58**) and naphthyl (**59**) substituents show markedly improved binding affinities as compared to the parent benzyl compound **60** (Figure 33A). More generally, replacing rings III and IV

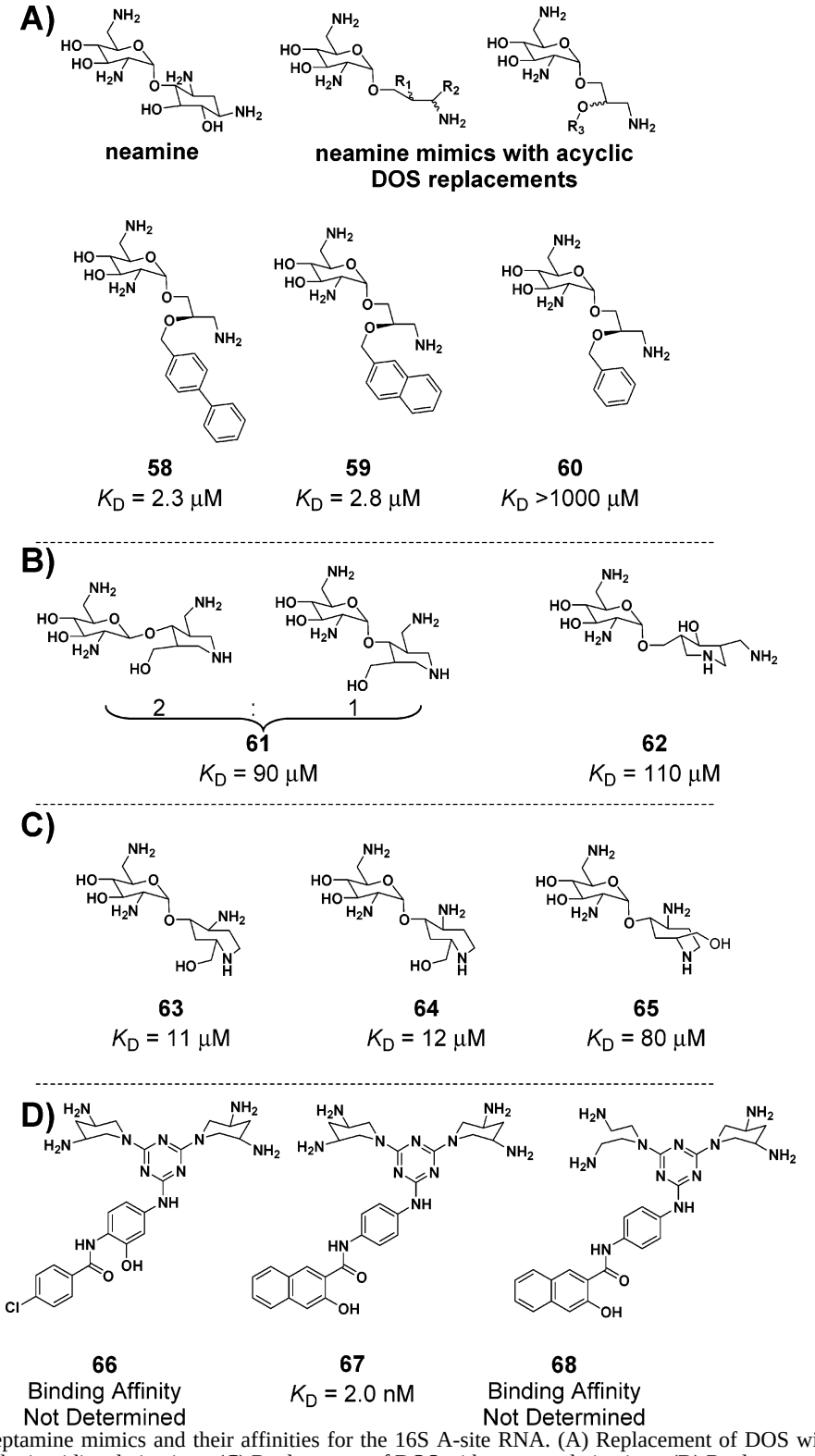


Figure 33. 2-Deoxystreptamine mimics and their affinities for the 16S A-site RNA. (A) Replacement of DOS with acyclic mimics. (B) Replacement of DOS with piperidine derivatives. (C) Replacement of DOS with azepane derivatives. (D) Replacement of DOS with diaminopiperidinyl derivatives.

with nonsaccharide substituents proved to be more successful than mimicking the other half of the DOS ring. The results from this study led the authors to conclude that the rigidity of the DOS scaffold is critical for the potency of aminoglycoside−16S A-site interaction.

8.1.6. Cyclic DOS Mimics

Published simultaneously with the acyclic mimics of DOS, Hermann and co-workers investigated various piperidine-glycosides for their ability to function as aminoglycoside mimics.[310] Guided by the high-resolution structure of paromomycin bound to the 16S A-site RNA, the conformationally restrained 3-(aminomethyl)piperidine scaffold was chosen to mimic the exocyclic equatorial 1,3-diamine motif of DOS. Linking the 6′-aminoglucosamine moiety with 3-(aminomethyl)piperidine in three different substitution patterns provided eight piperidine glycosides. Following their previous platform of evaluation, Herman and co-workers initially

evaluated the compounds in bacterial and mammalian cell-free in vitro translation assays. Compounds **61** and **62** were determined to selectively inhibit translation with IC_{50} values of 35 and 74 μM, respectively. When evaluated for binding to the 16S A site, these compounds bound with dissociation constants of 90 and 110 μM (Figure 33B).

In the continued search for pharmacophores to serve as suitable replacements of the DOS ring system, Hermann and co-workers undertook the investigation of the azepane ring system.[311] Using a crystal structure of the paromomycin bound to the A site as a template, the azepane scaffold was predicted to project a similar (to DOS) radial arrangement of the amine groups. The exocyclic amine at the 5-position of the azepane scaffold was shown to superimpose well with the 1-position amine of DOS, while the ring nitrogen overlaid with the 3-position amine substituent of DOS. A 12-member library of azepane-glycosides was synthesized by linking a variety of substituted azepane scaffolds to 6′-aminoglucosamine. The biological activity of each compound was assessed by cell-free bacterial and mammalian in vitro translation assays, and their binding affinities for the 16S A-site RNA were determined. The results from the panel assays revealed that compounds **63** and **64** were selective inhibitors of bacterial translation and bound with mid-micromolar affinity to the A-site RNA (K_D = 11 and 12 μM). Analysis of the remaining library members revealed that binding to the A site is intolerant to certain isomeric derivatization (e.g., **65**), perhaps consistent with the conformational design of the azepane scaffold. Although the potency of translation inhibition and RNA-binding affinity is significantly less than that of the aminoglycosides, the results are well in line with the values of neamine, suggesting that the azepane scaffold is likely a suitable mimic of DOS.

Next, Hermann and co-workers explored the cis-3,5-diamino-piperidinyl (DAP) scaffold as a structural mimic of the DOS ring system.[255] The authors proposed that DAP is a more ideal structural scaffold because it retains the trademark cis-1,3-diamino motif present within DOS while reducing the stereochemical complexity of 2-DOS. A series of 3,5-diamino-piperidinyl triazine (DAPT) compounds was synthesized and evaluated for their ability to bind the A-site construct. Initial attempts to determine the binding affinity of lead compounds **66**, **67**, and **68** (Figure 33D) for the A-site RNA, as assessed by the 2-aminopurine binding assay, were hampered by interference of these compounds with the emission of 2-aminopurine.[255] To obtain a precise binding affinity, ITC experiments were conducted by adding **67** to an unlabeled A-site construct. Analysis of the binding isotherm led to the determination that **67** binds to a high-affinity binding site (K_D = 2 nM). A second low-affinity site was observed at higher ligand:RNA ratios; such non-specific binding is typically observed in many RNA-binding experiments. The low nanomolar affinity of **67** rivals that of the most potent aminoglycoside binding affinities as measured by ITC.[208,210] Subsequent investigation in a variety of in vitro antibacterial assays determined that the DAPT compounds are ~40-fold less potent than paromomycin in bacteria cell culture models, but the mode of death is consistent with altering the fidelity of the ribosomal decoding site. Furthermore, in vivo mouse models of bacterial infection were used to investigate the efficacy of **68**, which afforded complete protection at dosages of 2.5 and 5.0 mg/kg when administered intraperitoneally.

8.1.7. Small Heterocyclic Compounds

Most small molecules discussed thus far as novel A-site ligands can be at best described as semisynthetic aminoglycoside mimetics. Identification of new chemical scaffolds that bind RNA has been a challenge. To address a bottleneck of drug discovery, Abbott Labs has pioneered an NMR screening platform dubbed "SAR by NMR".[232] Yu and co-workers utilized a similar platform for the identification of novel A-site binding scaffolds.[234] As the imino proton region of the 16S A-site RNA is well resolved, spectral changes within this region upon addition of compound provide a reliable method not only for detecting binding events but also for locating binding sites. To reduce the time and cost of screening a collection of 10 000 compounds, mixtures of 10 compounds (each at 0.5 mM) were added to the A-site RNA, and if a binding event was detected (as indicated by a change in the pattern of chemical shifts) each compound from the mixture would be tested individually. The results from this NMR-based screening campaign identified several drug-like scaffolds. Simple benzimidazole derivatives (**69**−**73**, Figure 34A) were determined to bind with high micromolar affinity to the A-site RNA. An amine substituent at the 2-position (**70**) led to a ~4-fold increase in binding affinity, while the presence or position of methyl substituents had minimal effect on binding affinity (**71**−**73**). Also, several 2-aminoquinoline derivatives were determined to exhibit a wide range of binding affinities (K_D values from 60 to 3180 μM) (compounds **74**−**79**, Figure 34B). In this series of compounds, methyl substitution at the 4-position (**77**) led to ~35-fold increase in binding affinity (K_D = 90 μM) and methylation of the exocyclic amine (**76**) increased binding affinity by ~6-fold relative to 2-aminoquinoline (**74**). Methylation at the 4-position and the exocyclic amine afforded derivative **79**, which showed improved binding affinity (K_D = 60 μM). Further elaboration of the 2-aminoquinoline scaffold led to amide-substituted compounds with low micromolar binding affinity.[234] Finally, 2-aminopyridine scaffolds were also identified from the screening efforts (compounds **80**−**85**, Figure 34C); surprisingly, this simple scaffold exhibited mid-micromolar binding affinity (K_D = 68 μM), which is better than that of the benzimidazole and 2-aminoquinoline series identified from the screen. Methylation at the 6-position (**84**) reduced binding affinity, while methylation at the 4- and 5-position markedly improved binding affinity (**82** and **83**). In fact, compound **82** was the tightest binding compound identified. NMR titration competition experiments with neamine or paromomycin confirmed that **82** competes for binding to the A site. A high-resolution structure of **82** bound to the A site could not be obtained due to the few intermolecular NOE interactions observed. However, a molecular model built from the limited dataset places the pyridine ring stacked between G1491 and A1492, while the terminal amine likely makes electrostatic contacts with the phosphate backbone of G1405. These results suggest the feasibility of obtaining new drug-like lead scaffolds for binding the A site; furthermore, similar NMR-based methods can be extended to any number of RNA targets, although the method is still typically utilized by NMR specialists and requires large amounts of RNA.

Ibis Therapeutics has pioneered the development of "SAR by MS", an ESI-MS platform for ligand discovery analogous to SAR by NMR.[312] The ESI-MS platform is less material intensive, and hits are detected by a shift in molecular weight of the targeted RNA upon complexation of ligand. Analogous

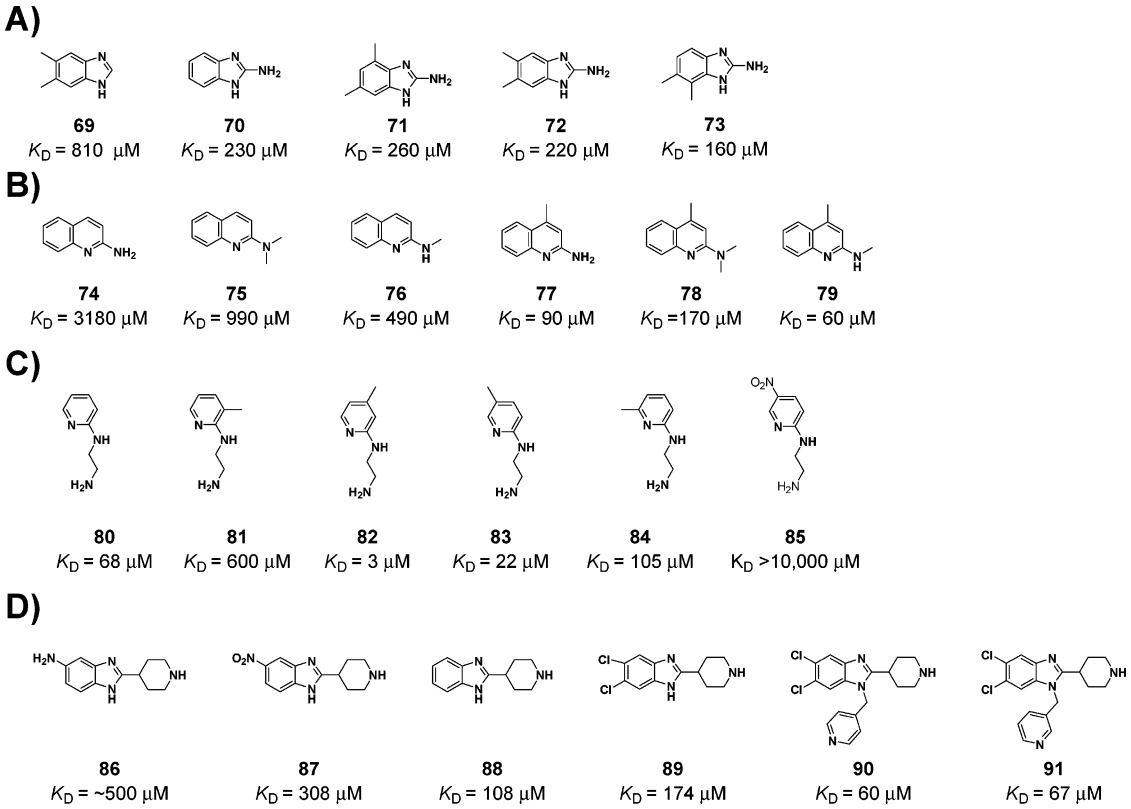


Figure 34. Some small heterocyclic compounds and their affinities for the 16S A-site RNA. (A) Benzimidazole derivatives identified by SAR by NMR. (B) 2-Aminoquinoline scaffolds identified by SAR by NMR. (C) 2-Aminopyridine derivatives identified by SAR by NMR. (D) Benzimidazole **86** was identified as a hit from an ESI-MS screen, while compounds **87**−**91** were evaluated by SAR by MS.

to the NMR platform, compounds can be screened in mixtures, but deconvolution may not be necessary as hit compounds are identified by molecular weight. The primary disadvantage to SAR by MS is that no information is obtained about the binding site; however, as a lead generation and evaluation tool SAR by MS will likely find more general use than SAR by NMR. Using this mass spectrometry-based approach, He and co-workers reported the identification of benzimidazole **86** as a high micromolar affinity hit from a high-throughput screening effort to the A-site RNA with a K_D of approximately 500 μM (Figure 34D).[313] Through competition experiments with glucosamine it was determined that **86** is a competitive binder for the A site, suggesting that tighter binding derivatives of **86** could be lead compounds for the development of new antibiotics. A series of derivatives of **86** was synthesized to probe the tolerance of substituents in the aromatic region. Although the nitro substitution enhanced binding (**87**), the unsubstituted pyridinyl-benzimidazole **88** and halogenated **89** exhibited the better binding affinities for the A-site RNA. Next, substitution of the N1-position was evaluated through the synthesis of alkylated or acylated compounds. The results from ESI-MS binding assays demonstrated that the N1-position is tolerant to a wide variety of substituents. Although the majority of N1-position derivatives showed a decrease in binding affinity, two pyridinyl analogues (**90** and **91**) enhanced binding affinity, $K_D = 60$ and 67 μM, respectively. Interestingly, this study and the SAR by NMR investigation yielded similar hit compounds with comparable binding affinities, demonstrating that SAR by MS is a suitable alternative to SAR by NMR.

Given the number of solved structures of various aminoglycosides bound to the 16S A-site RNA, in silico screening campaigns for identification of novel non-carbohydrate-containing scaffolds have been a popular area of investigation. After inspection of the crystal structure of paromomycin bound to the A site, Johnson and co-workers reasoned that scaffolds with two basic nitrogens could also be strong binders as the nitrogen-containing substituents could mimic the contacts made by DOS and ring IV.[314] After examining an in-house library, compound **92** (Figure 35) was selected as a potential A-site binding candidate compound as the necessary nitrogens were present and the aromatic ring systems could enhance binding by π-stacking or hydrophobic interactions. One-dimensional NMR titration experiments of **92** with the 16S A-site RNA indicated that the binding site is consistent with that of the aminoglycosides. Binding was also evaluated by ESI-MS, from which it was determined that **92** binds with low millimolar affinity (K_D < 15 mM) and 1:1 stoichiometry.

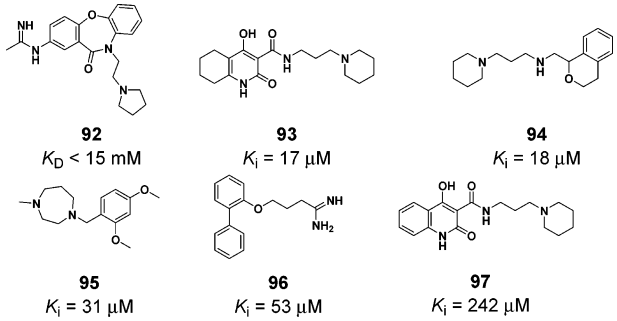


Figure 35. 16S A-site RNA binding compounds identified by in silico efforts. For compounds **93**−**97** the authors report K_i values for the displacement of fluorescently labeled paromomycin from the A-site RNA.

Using a more sophisticated computational screening platform, Foloppe and co-workers undertook the virtual

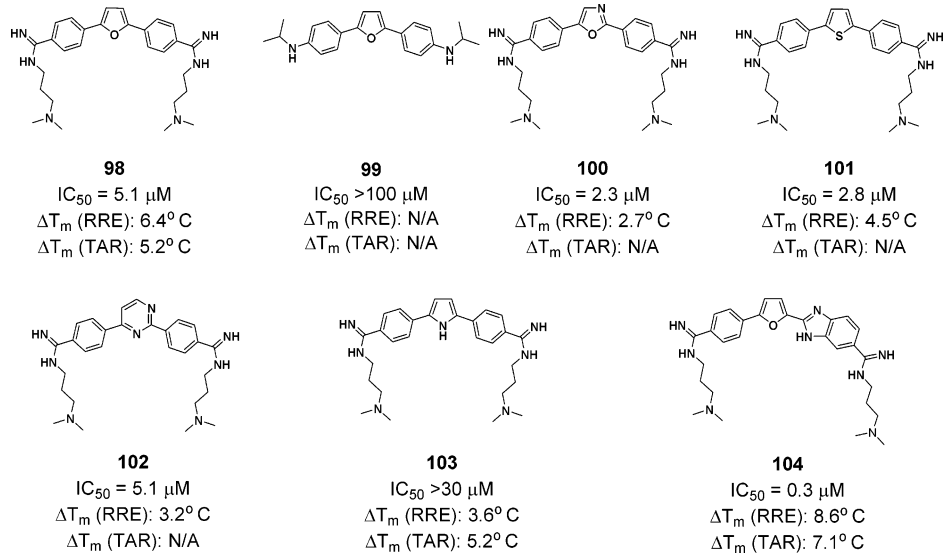


Figure 36. Diphenyl furans, their IC_{50} values for disruption of Rev−RRE, and their interaction with RRE and TAR RNAs (as assessed by changes in T_m).

screening of ∼1 million compounds to identify novel small molecule binders to the A-site RNA.[315] The ∼1 million-member library was initially prefiltered to remove compounds with greater than seven rotatable bonds and whose molecular weights were not in the range of 250−550 Da. The resulting ∼900 000 compounds were docked into the A site where the docking volume was set to match that of the neamine ring system from the crystal structure of paromomycin complexed with the A-site RNA using RiboDock. The best-ranking 2000 compounds were visually inspected for compounds that formed stacking interactions with G1491 and hydrogen bonds with A1408; the resulting 129 compounds were then evaluated for RNA binding as assessed by a FRET-based displacement of a fluorescent-paromomcyin ligand from the A site and NMR titration experiments. Of the 129 selected compounds, 34 were active in the displacement assay with K_i values ranging from 17 to 426 μM (**93**−**97**, Figure 35). Compound **93** emerged as the most potent competitor of the paromomycin−16S A-site RNA interaction. Interestingly, the aromatic derivative **97** was more than 10-fold less potent in the competitive displacement assays. The authors suggest that, based on the structural data from various aminoglycosides bound to the A site, the bulkier aliphatic ring system is a better fit for the binding pocket of the internal loop. Importantly, NOE data from 2D NMR experiments are consistent with **93** binding to the internal loop region by stacking with G1491 and contacting A1408, A1492, and A1493.

8.2. Small Molecule Binders to the Rev Response Element (RRE) RNA

8.2.1. Diphenylfurans and Derivatives as RRE Ligands

Building from their work with small molecule binders to RNA duplex regions, Zapp, Wilson, and co-workers tested various diphenylfurans and other aromatic heterocyclic compounds for their ability to compete for Rev binding to RRE.[316] To evaluate binding, a gel-shift competition assay using ^{32}P-labeled wild-type RRE (wtRRE) and Rev was used as a screen. The binding of Rev to ^{32}P-labeled wtRRE alters the migration rate of wtRRE through the gel; that is, the Rev−wtRRE complex migrates at a higher molecular weight than wtRRE alone, thus providing a convenient but low-throughput readout. Small molecules that compete for binding, either through interaction with wtRRE or Rev, would cause wtRRE to migrate faster than wtRRE−Rev complex. In total, 30 heterocyclic ligands, the majority of which contained the diphenylfuran core, were screened, and compound **98** was identified as the most potent ligand at preventing the Rev−wtRRE interaction (Figure 36). The IC_{50} of **98** was determined to be 0.1 μM, approximately 10-fold better than neomycin, although in a subsequent publication **98** was determined to have an IC_{50} of 5.1 μM.[317] General trends that emerged from the collection of ligands were that the central furan scaffold appeared to be required for binding and dicationic groups on both sides of the molecule fared better than monocationic groups. More detailed investigation through chemical foot-printing experiments with wtRRE revealed that **98** bound to duplex regions in areas where the base pairs experience greater flexibility, such as at the ends of a helix, near a loop junction, or other secondary structures. On the basis of overlapping protection patterns of Rev and **98**, it was proposed that **98** binds to the G−C base pair just below the critical internal loop to exert its inhibitory effects. Attempts to address the specificity of **98**−wtRRE interaction were made by use of an in vitro pre-mRNA splicing assay. In this assay, a ^{32}P-label pre-mRNA is incubated with mammalian nuclear extracts; as pre-mRNA splicing is composed of numerous RNA−RNA and RNA−protein interactions, any inability to form the mature RNA in the presence of compound is deemed to be due to off-target binding of the compound. Previously, the neomycin−RRE interaction was demonstrated to be "selective" using this assay;[94] **98** was found to be equally selective using the same assay.

In the continued pursuit of RRE ligands the structure−activity relationship of **98** and RRE was determined.[317] Three key structural features were addressed: the identity of the central heterocycle, the substitution pattern of the dicationic groups radiating from the phenyl rings, and the identity of cationic substituents. To determine the efficacy of ligand binding, melting temperature and gel-based competition assays against Rev were performed. Interestingly, the SAR data of the central heterocycle suggested that five- and six-member rings are permitted, provided that a hydrogen-bond-accepting group was positioned in the center. For example,

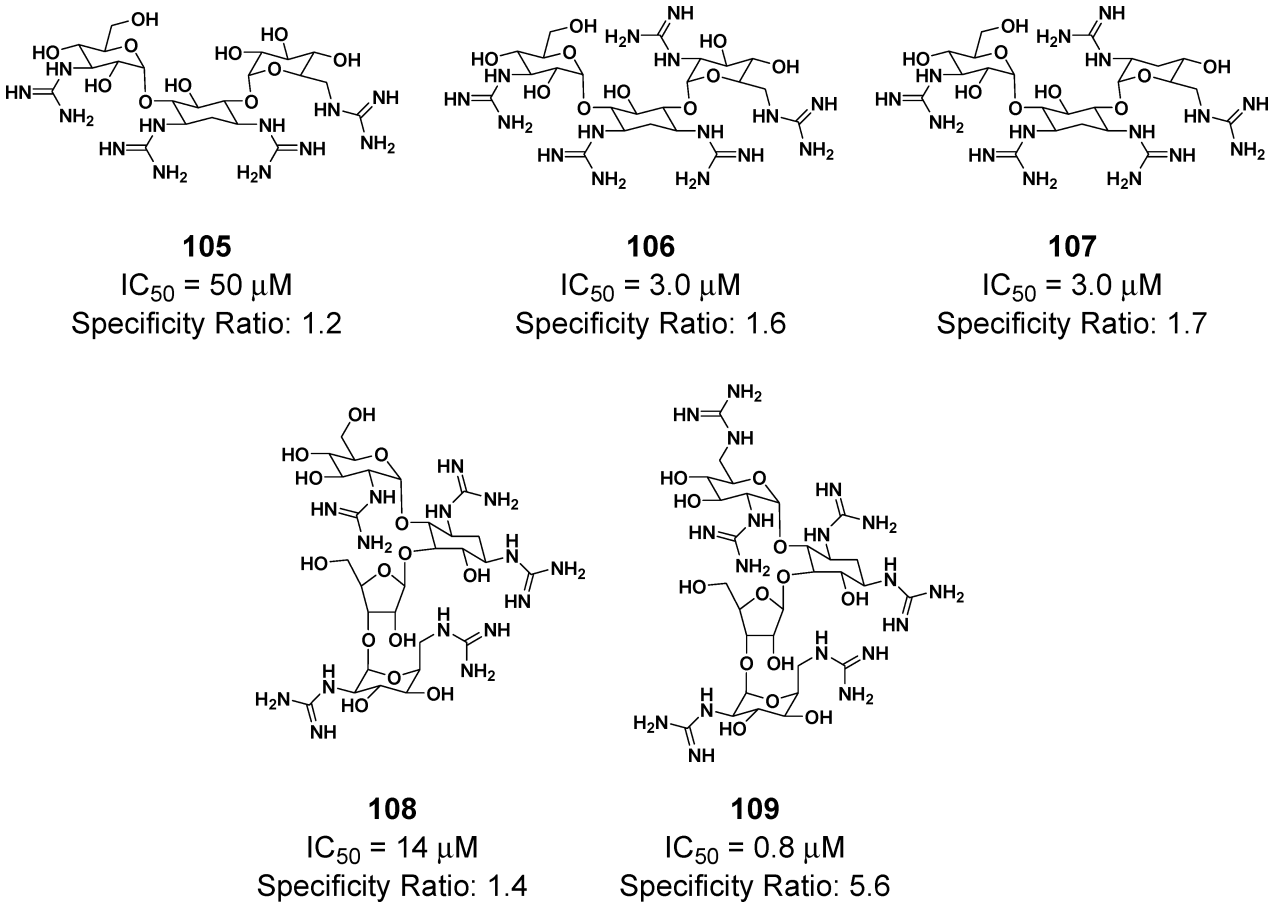


Figure 37. Guanidinoglycosides and their ability to prevent the Rev−RRE interaction. Higher specificity ratios indicate compounds whose binding is appreciably affected by the presence of competing nucleic acids.

furan (**98**, but not **99**), oxazole (**100**), thiophene (**101**), and pyrimidine (**102**) scaffolds were found to compete efficiently with Rev binding; however, pyrrole (**103**) or *N*-methyl pyrrole scaffolds experienced decreased affinity for RRE and failed to compete for Rev binding (Figure 36). Consistent with previous findings, monocationic substituents often failed to compete with Rev binding, suggesting that dicationic groups are required. It is important to note that the SAR data was used to identify features important for competing with Rev−RRE interaction and not necessarily binding to RRE. The general correlation can be made that compounds which bound RRE better, as assessed by melting temperature studies, are more likely to prevent Rev binding; however, there is no clear correlation between degree of stabilization and IC_{50} of Rev competition. From these studies **104** emerged as the most potent RRE binding ligand ($\Delta T_m = 8.6$ °C) and the most potent Rev competitor ($IC_{50} = 0.3$ μM). In order to evaluate selectivity, analogous melting temperature and competition assays were performed using the Tat−TAR system. As the TAR RNA represents a different ligand-binding site, "selective" compounds for RRE would fail to bind to TAR and compete for Tat binding. Although it was concluded that several ligands were indeed selective for RRE because such compounds failed to disrupt the Tat−TAR interaction, in all cases the ligands tested were found to bind TAR with comparable affinity to RRE, as determined by comparing the ΔT_m values. Subsequent NMR studies have revealed that **104** binds to the internal loop of RRE through the minor groove as a dimer with a K_D of ~7 nM.[318] The binding sites are proposed to be highly cooperative, as SPR studies and experiments monitoring the quench in fluorescence of **104** upon binding to RRE produced binding curves consistent with a single-site mode of binding. SPR experiments conducted with increasing salt concentrations could not resolve the individual binding events, further confirming that two ligands bind with high cooperativity.

8.2.2. Aminoglycosides Derivatives as RRE Ligands

In an alternative strategy, Tor and co-workers sought to enhance the affinity and specificity of the aminoglycoside−RRE interaction by tuning various parameters known to impact binding. Earlier work established that the general affinity between the RNA and aminoglycoside arose primarily through electrostatic contact of the charged ammonium groups with the phosphate backbone.[100,205,319] Tor and co-workers hypothesized that replacement of the amine substituents with guanidinium groups would result in greater affinity and selectivity.[209] The envisioned "guanidinoglycosides" were anticipated to have enhanced affinity owing to their greater basicity, while specificity would be conceivably endowed by directional hydrogen-bonding resulting from the planar configuration adopted by guanidinium groups. Using a solid-phase displacement assay to monitor the Rev−RRE interaction, guanidinylation of kanamycin A (**105**), kanamycin B (**106**), tobramycin (**107**), paromomycin (**108**), and neomycin (**109**) resulted in ligands that were as much as 10-fold more potent at preventing the association of Rev−RRE than the corresponding parent aminoglycoside (Figure 37). Importantly, a general method was developed to directly assess the relative specificity of a given RNA−ligand interaction. By performing the same solid-phase displacement assay in the presence of competitor nucleic acids (either

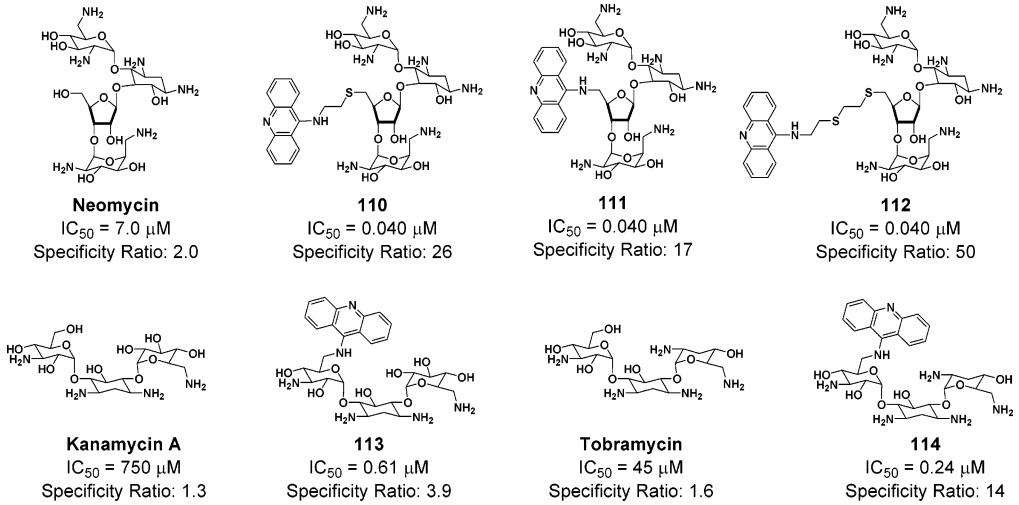


Figure 38. Acridine-conjugated aminoglycosides. The IC_{50} values listed correspond to those derived from the solid-phase Rev−RRE disruption assay. The specificity ratio is the average IC_{50} in the presence of DNA and $tRNA^{mix}$ divided by the IC_{50} in the absence of these nucleic acids.

polyA−polyU or DNA) a realtive assessment of nonspecific binding can be made by comparing the IC_{50} values performed in the presence and absence of competitor nucleic acids. From this type of analysis, the guanidinoglycosides were generally found to be more selective than their parent aminoglycoside when challenged with competitor RNA, and none of the ligands tested were significantly affected by the presence of competitor DNA. Interestingly, in a subsequent investigation Tor and co-workers examined the cellular uptake of fluorescently labeled amino- and guanidinoglycosides.[320] The results of this investigation found that **107** and **109** penetrated cells 10−20-fold more efficiently than their parent aminoglycosides. The mechanism of cellular uptake appears to be similar to that of poly-Arg peptides, as the unlabeled guanidino-neomycin could effectively compete with labeled poly-Arg cellular uptake.

Various intercalators have been shown to exhibit significant affinity for RNA structures. However, as intercalators are generally thought to be nonspecific, use of intercalating agents to target specific RNA structures has been largely avoided. Tor and co-workers recognized that intercalator−aminoglycoside conjugates may lead to potent inhibitors of the Rev−RRE interaction as the aminoglycoside portion could provide the specificity for binding to RRE, while the intercalator could enhance affinity.[98] Initial efforts to create a neomycin−acridine conjugate linked through the 5″-position (**110**) proved successful as the binding affinity between RRE and the neomycin−acridine conjugate rivaled that of the Rev peptide itself, with a K_i of 1.5 nM for **110** and 1.0 nM for the Rev peptide (Figure 38). Importantly, as intended by the design of aminoglycoside−intercalator conjugates, the gel-shift-derived IC_{50} values for 9-aminoacridine, neomycin, and the neomycin−acridine conjugate conclusively demonstrate that the neomycin−acridine conjugates exhibit substantially enhanced potency of either component individually. However, subsequent work to investigate the selectivity of **110** for RRE determined that **110** is less selective than neomycin or **109** (Figure 38), although the extent of off-target binding is somewhat dependent on the type of competitor nucleic acid.[263] Additionally, modification of the linker length and composition between neomycin and acridine (compounds **110**, **111**, and **112**) revealed an inverse relationship between linker length and specificity for RRE over the $tRNA^{mix}$. While all three compounds disrupt the Rev−RRE interaction with equal potencies, shorter linker length gave rise to compounds with enhanced specificity, while longer linkers were less selective. The generality of aminoglycoside−acridine conjugates was probed by the development of kanamycin A−acridine (**113**) and tobramycin−acridine (**114**) conjugates. In both cases the conjugates remarkably improved the efficacy of disrupting the RRE−Rev interaction, 1200- and ∼200-fold, respectively. Unfortunately, these aminoglycoside−acridine conjugates were equally susceptible to nonspecific binding, as in both cases evaluated the parent aminoglycosides were relatively unaffected by the presence of competing nucleic acids, whereas **113** and **114** displayed specificity ratios of 3.9 and 14, respectively.

8.2.3. Acridine Derivatives as RRE Ligands

Recognizing the potential of acridine-like compounds, that is charged, polycyclic, aromatic small molecules, Marino and co-workers performed a similarity search for such compounds and selected a subset for evaluation in the Rev−RRE system.[321] Using their previously developed 2-aminopurine RRE binding assay,[99] 12 compounds were initially screened for their ability to induce a change in fluorescence of 2-aminopurine (compounds **9**, **10**, and **115**−**124**, Figure 39). Acridine orange (**120**) and proflavin (**10**) were selected for further investigation as these ligands exhibited the greatest change in fluorescence upon binding. Compounds **120** and **10** bound to RRE with low micromolar affinity and fit well to a single-site model with K_D values of 1.53 and 1.15 μM, respectively. Further characterization by ^{1}H NMR revealed an apparent **10**:RRE binding stoichiometry of 2:1. Furthermore, only a single set of imino proton signals was observed during the course of the titration, implying that the two molecules of **10** bind to the same site with high cooperativity. NOE experiments between the bound ligands suggest the molecules are stacked above each other in the RRE binding site within the internal loop region of RRE. Curiously, competition experiments with the Rev peptide revealed that **10** competes for binding to RRE with an IC_{50} of 0.1 μM; the results from the initial dose-dependent change in fluorescence of 2-aminopurine upon **10** binding are apparently the result of a higher affinity site that is not competitive with Rev binding. Thus, **10** is proposed to inhibit the binding of Rev by binding as a dimer directly to RRE in the critical

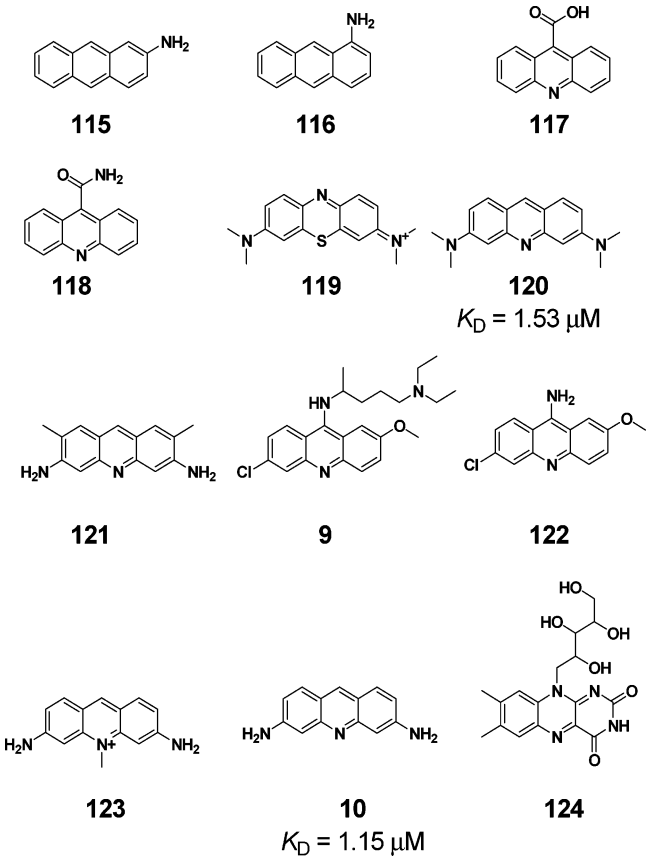


Figure 39. Charged, polycyclic aromatic small molecules examined for binding to RRE. Compounds **10** and **120** were determined to induce the greatest amount of change in fluorescence upon binding to a 2-aminopurine-labeled RRE construct, and thus, their affinities for RRE were determined.

internal loop region; upon binding to RRE, **10** induces many of the same stacking and base-pairing interactions as the Rev peptide.

Although the above-mentioned studies have succeeded in generating tight binding ligands by enhancing general binding properties such as electrostatic contacts or hydrophobic effects, the possibility of selectivity also exists through exploitation of the Watson−Crick face of the exposed bases. Working primarily in the DNA molecular recognition field, Nakatani, Saito, and co-workers have sought the development of small molecule ligands to selectively recognize regions of mismatched DNA. Key to their design was the hypothesis that unpaired, exposed bases could be captured by intercalating agents possessing the properly displayed hydrogen-bond-donating and -accepting groups complementary to the intended base.[322] Initial efforts toward the development of this paradigm utilized 2-aminoacylamino-1,8-naphthyridine (**125**) for the selective recognition of guanine (Figure 40A).[322] Quantitative DNase I foot-printing experiments and ESI-MS determined that **125** binds with 1:1 stoichiometry to single guanine bulges with a K_D of ~30 μM. Importantly, no binding was observed for a DNA duplex containing a single adenine bulge; these results were further corroborated by CD experiments. Continued effort has led to the identification of azaquinolone (**126**) as a small molecule suitable for recognition for adenine.[323] Dimeric versions of **125**, **126**, or hybrid **125**−**126** dimers have strong binding affinities, K_D values of 0.1−0.5 μM, and selectively recognize G−G, A−A, or G−A mismatches within DNA duplex regions, respectively.[323−325]

This use of **125** and **126** has recently been extended to the selective recognition of RNA bases. Tok and co-workers evaluated the generality of the dimers of **125** for G−G mismatch recognition in a variety of nucleic acids (Figure 40B).[227] Using the fluorescent end-labeled binding assay, **127** was evaluated for binding to DNA duplex, DNA−RNA hybrids, and RNA duplex featuring a single G−G mismatch. Compound **127** bound the DNA construct most efficiently with a K_D of 1.4 μM while exhibiting slightly reduced affinity for the RNA construct, K_D of 3.2 μM. Also, **128** was determined to bind with comparable affinity to single guanine bulges regardless of the composition of the nucleic acid.[227] Nakatani and co-workers used dimers of **125** and **125**−**126** conjugates to target RRE,[326] as structural analysis of RRE demonstrated that the 2 × 3 internal loop forms G−G and G−A noncanonical base pairs. Using SPR, **129** and **130** were shown to exhibit a strong response in the presence of RRE, indicative of tight binding; however, no binding affinity or kinetic data was reported. On the basis of the magnitude of the SPR response upon binding to RRE, **129** was determined to bind with greater affinity than **130**. Selective recognition was assessed by performing SPR experiments with an RRE construct lacking the internal loop, TAR RNA, and a single-strand unstructured RNA sequence; for all compounds tested, minimal off-target binding was observed.

8.3. Small Molecule Binders to the Thymidylate Synthase (TS) mRNA

Previously, the aminoglycosides were shown to bind the 1 × 1 internal loop of the TS mRNA with low micromolar affinity (section 4.7.3). In order to identify ligands with higher affinity, Cho and Rando screened various DNA intercalators (quinicrine, proflavine) and minor groove-binding compounds (Hoechst 33258, DAPI, distamycin A) for their ability to bind the TS mRNA;[327] an anisotropy-based displacement assay featuring rhodamine-labeled paromomycin was utilized. All DNA binding compounds tested were able to compete for the paromomycin-binding site of the TS mRNA, suggesting that these compounds bind to the 1 × 1 internal loop region. In general, the groove-binding ligands bound the TS mRNA construct with greater affinity than the intercalators (average K_D of 286 nM versus 925 nM). Hoechst 33258 (compound **13**, Figure 24) was the most effective at displacing paromomycin from the internal loop with a K_D of 60 nM. Mutational analysis of the TS mRNA construct revealed that tight binding was independent of the sequence of the internal loop or the surrounding bases, suggesting that the presence of an internal loop is the lone requirement for binding. Footprinting analysis with Hoechst 33258 determined that nucleotides in and around the internal loop are protected with high concentrations of ligand; however, the presence of a broad footprint suggests that the binding site may not be confined to just the internal loop region. Previously, Hoechst 33258 was demonstrated to increase fluorescence upon binding to DNA.[328] By monitoring the fluorescence of Hoechst 33258 upon binding a TS mRNA construct lacking the 1 × 1 internal loop a nonsaturating binding curve was obtained, suggesting that Hoechst 33258 binds the TS mRNA duplex in a nonselective fashion, possibly accounting for the broad footprint observed.

9. Bulge Binding Compounds

A bulged region within a RNA duplex occurs when one or more unpaired bases are present on only one strand. Unlike

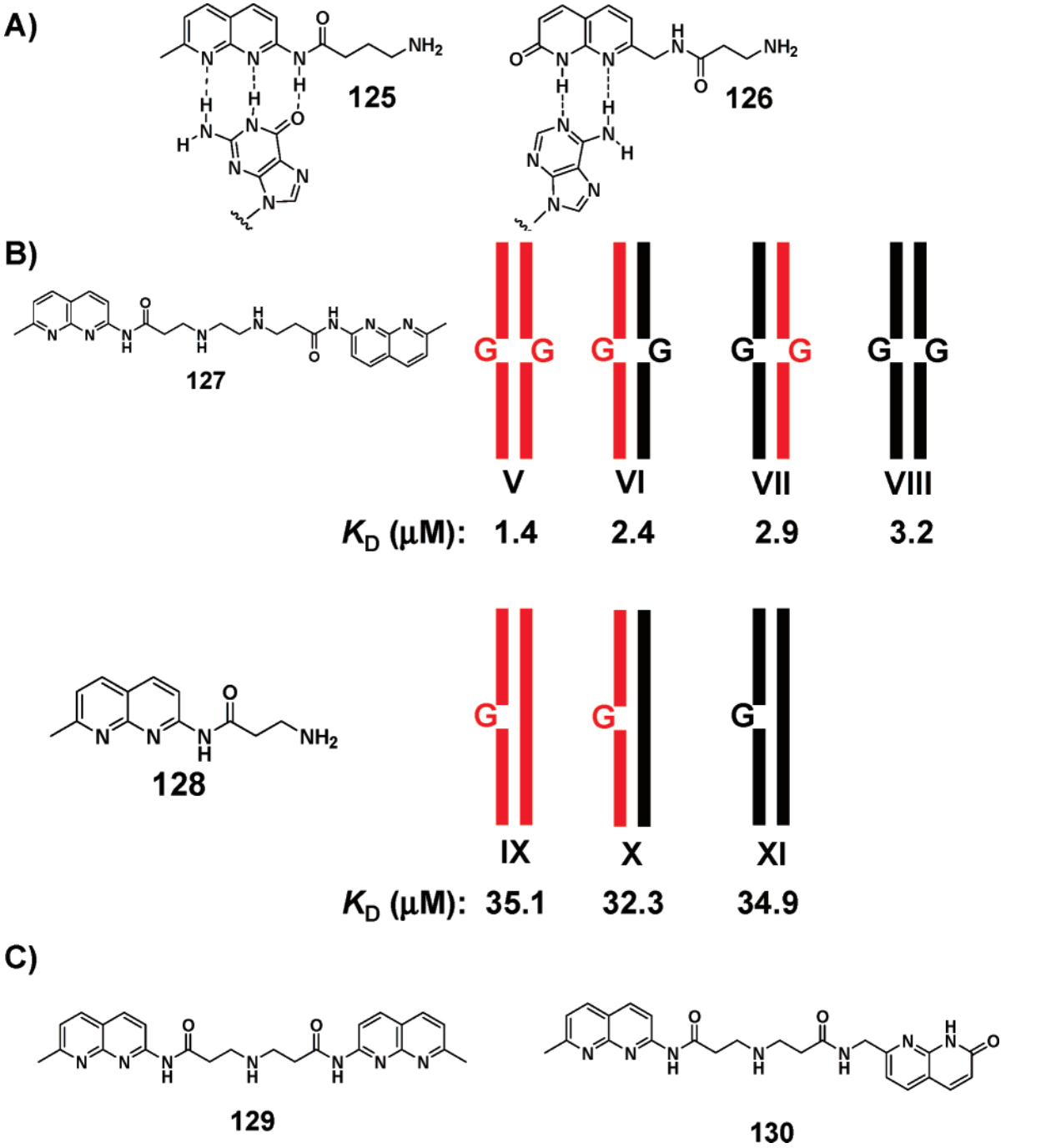


Figure 40. Selective base recognition by naphthridine derivatives. (A) Base recognition mode of **125** and **126**. (B) Evaluation of bulged guanine recognition by dimeric and monomeric naphthyridine derivatives. Red designates DNA strands; black designates RNA. (C) Compounds tested for sequence-dependent binding to RRE.

internal loops, formation of bulged regions is always destabilizing toward duplex formation and becomes more destabilizing with increasing bulge size.[329,330] However, despite the apparent lack of stabilizing interactions within bulged regions, structural determination of numerous bulged regions indicates that these regions can adopt defined folds. For example, single nucleotide bulges can be stacked into or intercalate within the duplex, participate in recognition of adjacent base pairs, or protrude outward into solution;[40,329] these "looped-out" bulges often serve to increase the flexibility of the RNA backbone.[40] The structure and dynamics in bulged regions of RNA are less well investigated than internal loops; however, it is likely that bulged regions present a different binding environment than internal loops. The most commonly targeted RNA bulges are the trans-activating region (TAR) RNA, T-box RNA, and iron response element (IRE) RNA. Compounds found to bind to these regions are described in the sections below; for background on the biological function and importance of each of these RNAs, see section 4 of this review.

9.1. Small Molecule Binders to the Trans-Activating Region (TAR) RNA

9.1.1. Intercalators

The bulged region of TAR is one of the best-studied and most targeted RNAs. As part of their initial efforts to identify non-aminoglycoside small molecule ligands for RNA targets,

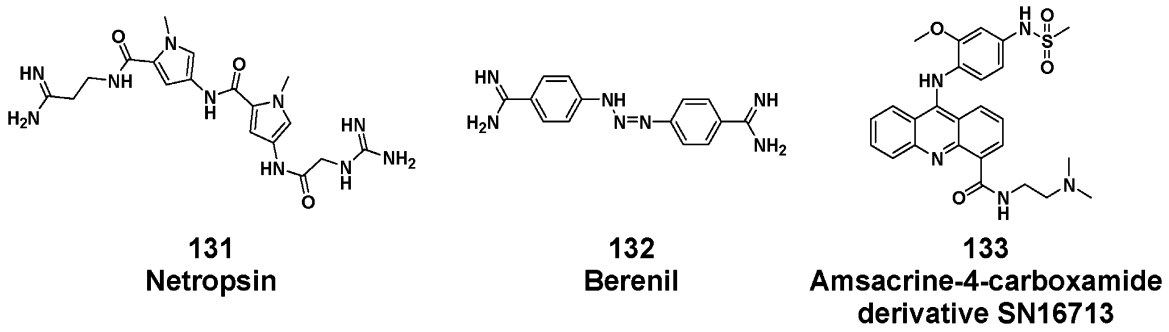


Figure 41. DNA binding agents examined for binding to TAR RNA.

Wilson and co-workers discovered that ethidium (**8**, Figure 24) efficiently bound TAR as assessed by UV melt studies.[331] This initial result spurred the examination of other intercalators as suitable ligands for the TAR RNA and more generally examination of intercalation as an operative mode of binding. Bailly and co-workers undertook the examination of several classic DNA intercalating agents (ethidium and proflavine, Figure 24), DNA minor groove-binding compounds (DAPI and Hoescht 33285, Figure 24; netropsin (**131**) and berenil (**132**), Figure 41), and a threading intercalator (amsacrine-4-carboxamide derivative SN16713 (**133**)) for their binding to TAR. Through electric linear dichroism (ELD) these authors determined that all compounds tested bound to TAR through an intercalative mode; the only exception was **131**, which did not bind TAR.[332] Further detailed characterization of the interactions between Hoescht 33285 and TAR versus a bulgeless TAR construct by UV melt, circular dichroism (CD), ELD, and RNase footprint revealed that Hoescht 33285 preferentially binds to the bulged region of TAR.[333] Both UV melt and CD experiments demonstrated that Hoescht 33285 has a greater affinity for TAR over the bulgeless construct; however, ELD experiments established that Hoescht 33285 binds to both RNA constructs through an intercalative mode of binding.

Hamy and co-workers extended the use of intercalators in the design of "In-PRiNts" (inhibitor of protein−ribonucleotide sequences) compounds.[334] In-PRiNts were designed as modular compounds consisting of three components: a polyaromatic/heterocyclic scaffold for stacking interactions, a positively charged moiety to enhance affinity by interacting with the phosphate backbone, and a linker region connecting the intercalating domain with the electrostatic domain (see Figure 42A). In-PRiNts compounds, such as **134**, which featured 6-chloro-2-methoxy acridine as the intercalating moiety and spermidine as the electrostatic enhancer proved to be superior at disrupting the Tat−TAR interaction relative to either entity alone. The generality of this approach was further demonstrated using mitonafide linked to spermidine (**135**), which exhibited improved inhibition of the Tat−TAR complex as compared to mitonafide alone. By varying the linker length, composition, and length of the electrostatic moiety, features important for the disruption of the Tat−TAR complex were unveiled. It was demonstrated that the linker position and chemical composition of the linker had a substantial impact on the effectiveness of the compound. For example, substitution at the 9-position of the acridine scaffold was well tolerated, while 4-position substituents were not. Furthermore, all attempts to link the electrostatic moiety through an amide linkage (**136**) resulted in less active derivatives compared to their amine-linked counterparts. Compound **134** proved to be the most potent compound at disrupting the Tat−TAR interaction (IC_{50} = 22 nM) by binding the bulged region of TAR, as determined by gel-shift assays, RNase foot printing, and NMR titration experiments. The NMR experiments provided greater insight into the binding mode as the acridine scaffold was determined to stack between A22 and U23 while the secondary amine at the 9-position forms hydrogen bonds with the G26−C39 base pair (Figure 42C), which may explain the deleterious effects of an amide linkage at the same position.

The results of the NMR study were used to guide biochemical validation experiments examining the binding of **134** to TAR mutants.[335] In order to confirm the interaction between the acridine scaffold and the G26−C39 base pair, constructs featuring a G26A−C39U and G26C−C39G mutation were utilized (Figure 42D). Quantitative RNase foot-printing experiments were used to determine that **134** binds to the wild-type TAR sequence with an apparent K_D of 150 nM. Analogous experiments conducted with the G26A−C39U base pair resulted in an apparent K_D of ∼7 μM, a nearly 50-fold difference in binding affinity. Reversing the order of the base pair (G26C−C39G) or altering the bulged nucleotide (U23C) resulted in a 14- and 11-fold decrease in binding affinity. Analogous to the work conducted by Nakatani and co-workers, Bailly and co-workers demonstrate that simple, yet properly functionalized aromatic scaffolds can afford selective recognition of base pair elements. Several In-PRiNts were evaluated for their ability to disrupt the Tat−TAR complex in vivo by use of a Tat−TAR reporter gene system. The association of Tat with TAR drives the expression of β-gal, thus providing a convenient functional readout within the background of the complex bacterial transcriptome. Several In-PRiNts demonstrated the ability to block the expression of β-gal. A general correlation was observed between the potency observed in the gel-shift assay and reduction in β-gal activity; however, several compounds that failed to disrupt the Tat−TAR complex in vitro still reduced the expression of β-gal.

Rather than attempt to enhance or direct the specificity of general intercalators for the bulged region of TAR, Condom and co-workers sought the design of a modular ligand that would exploit two separate binding sites within TAR by creating ethidium−arginine conjugates.[336] By tethering compounds that bind to distinctly different sites, this modular design of bifunctional ligands attempts to enhance the specificity and affinity of TAR binding ligands (Figure 43A). Ethidium−arginine conjugates were synthesized with varying linker lengths and assayed for anti-HIV activity using in vitro cell culture models. From these initial assays, **137** and **138** (Figure 43B) exhibited anti-HIV activity with no apparent toxicity to noninfected cells. Subsequent in vitro characterization confirmed that **138** binds to TAR, as determined by UV melt experiments. In order to deconvolute the binding sites of each module in the tethered compounds, RNase

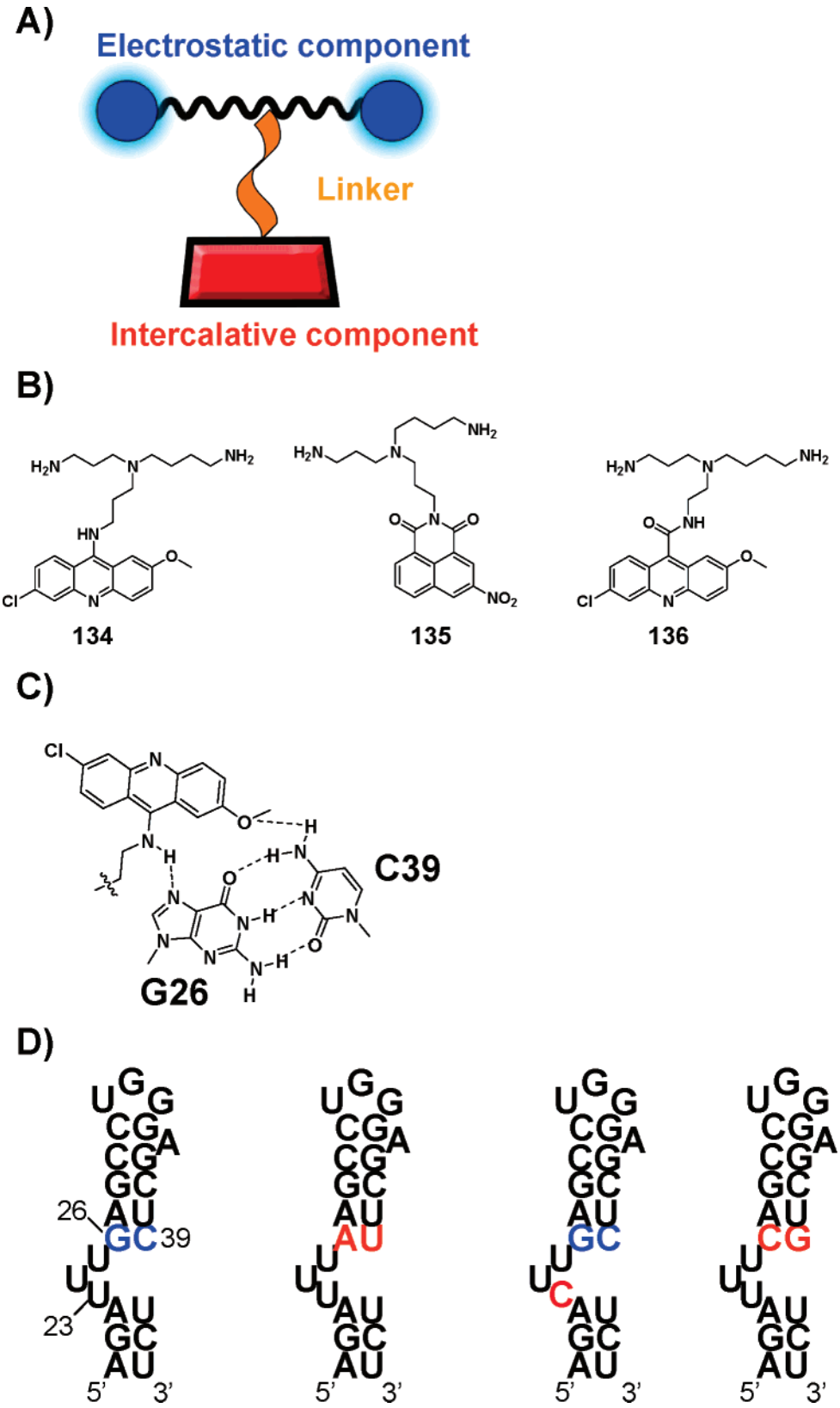


Figure 42. Modular design of TAR binding compounds. (A) Schematic of the In-PRiNts design. (B) InPRiNTs ligands. (C) Schematic binding model of **134** complexed with TAR as determined by NMR. (D) TAR constructs utilized to confirm the mode of binding as determined from the NMR data. Shown in blue is the binding site, while red denotes mutations from the wild-type construct.

footprints were conducted. These experiments demonstrated that residues in and around the bulged region and the guanine residue of the C18−G44 base pair were protected with increasing concentrations of **138**. Additionally, molecular modeling of the various synthesized ligands with TAR confirmed that **137** and **138** contained linkers of appropriate length to span the individual binding sites, thus corroborating the foot-printing data. However, the results do not conclusively demonstrate that each portion of the conjugate binds to its intended location as key controls of ethidium

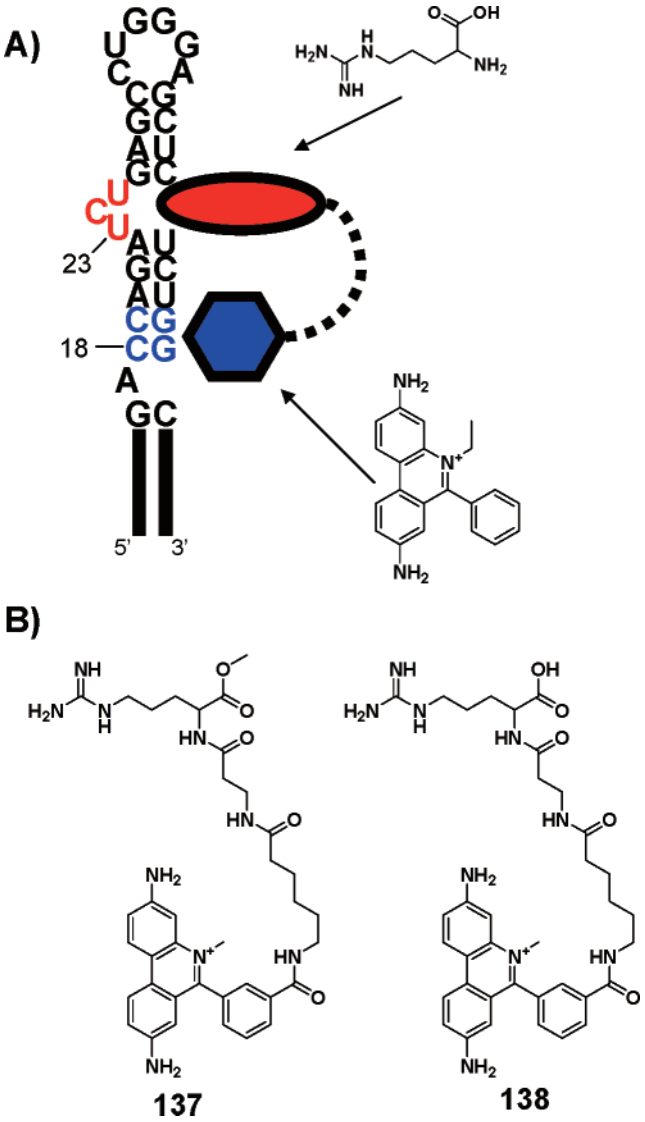


Figure 43. Design of ethidium−arginine conjugates. (A) Schematic of a modular ligand based on the known specificities of ethidium for the C−G, C−G base pairs (shown in blue) and arginine for the bulged region (shown in red). (B) Synthesized ethidium−arginine conjugates which were characterized in vitro.

and arginine conjugated with only the linker were not presented.

Molecular modeling has been extensively used to guide library design of 16S A-site RNA-binding compounds. However, a notable dearth of in silico screening campaigns led James and co-workers to develop a computational platform for such efforts.[238,337] The chemical entities that obeyed Lipinski's rules of five from the Available Chemical Database, which contains >180 000 compounds, were screened in silico for binding to TAR in a multitiered platform. The first tier utilized a rapid docking procedure (DOCK) to dock flexible ligands into a rigid TAR structure. The top 30 000 compounds were then redocked using a different scoring function (ICM), which takes into account hydrogen bonding and computationally intensive terms such as solvation. The resulting top 5000 compounds were reanalyzed by allowing flexibility in ligand and RNA. After visual inspection and consideration of factors such as drug likeness, cost, and availability, 50 compounds were tested for their ability to inhibit Tat−TAR association; 10 of these caused complete inhibition between 0.1 and 1.0 μM (compounds **139**−**148**, Figure 44A). From these the studied phenothiazines (**140**, **143**, and **147**) emerged as a class of compounds which inhibited the Tat−TAR interaction with reasonable potency.[337] The phenothiazine core structure consists of a nonplanar heterocycle which prevents a purely intercalative mode of binding.[338] NMR titration experiments determined that **143** induces chemical shifts of residues in and around the bulged region; increasing titrations to form a 5:1 ratio of ligand to RNA failed to cause shifts in nucleotides in the stem or hairpin loop region. Subsequent work to define the generality of phenothiazines as novel RNA-binding ligands resulted in results in conflict with the original report. Despite the ability of **143** to disrupt the Tat−TAR interaction with an $IC_{50} < 1$ μM,[337] NMR-derived binding constants revealed a weak binding interaction between **143** and TAR, K_D = 270 μM (Figure 44B).[338] The interaction between **143** and TAR was shown to be nonselective as weak binding interactions were detected for the internal loop region of the 16S A-site RNA (K_D = 360 μM), a two nucleotide bulged region of the coxsackie virus B3 RNA (K_D = 330 μM), and the hairpin loop region of a polio virus loop B construct (K_D = 1800 μM).[338]

Detailed characterization of the **143**−TAR complex revealed structural perturbations not typically seen with other TAR binding compounds.[339] The central ring system was found to insert between base pairs G26−C39 and A22−U40 in a pseudo-threading intercalating fashion. The unsubstituted aromatic ring is placed in the major groove where it is "sandwiched" between U25 and U40.[339] The substituted aromatic ring was determined to project through the major groove and into the minor groove, occupying a space below G26, which may be stabilized by partial stacking interactions. The aliphatic tail is also placed in the minor groove and interacts with the G26−C39 base pair,[339] reminiscent of the neomycin−TAR complex.[89] Surprisingly, the stacked conformation U25 and U40 resembles the structure of TAR alone; structural determination of TAR complexed with other compounds capable of preventing association with Tat (arginamide,[340] Tat−peptide,[84] or neomycin[89]) showed this stacked conformation of nucleotides to be disrupted. Consistent with these findings, Al-Hashimi and co-workers determined that acetylpromazine (**143**) failed to arrest the global flexibility of TAR,[341] a property associated with the binding of arginine and neomycin. The residual dipolar coupling experiments showed that upon binding **143** the conformational freedom of the **143**−TAR complex was nearly identical to that of TAR alone. In stark contrast, the primarily electrostatically driven ligand neomycin induces conformational arrest of TAR global flexibility, presumably by interacting with the phosphate backbone. As conformational arrest of TAR is thought to be important for preventing association with Tat, the results from the residual dipolar coupling experiments suggest that binding to the bulged region of TAR is necessary but not sufficient to halt conformational flexibility.[341] This inability to induce conformational arrest coupled with the high micromolar binding affinity observed for TAR suggests that **143** and likely the entire phenothiazine class disrupt the Tat−TAR interaction by a mechanism other than simply binding directly to TAR.

9.1.2. Guanidinylated Compounds

The design of novel 16S A-site binding ligands has been greatly aided by the determination of various structures of ligand-bound complexes. For example, structural determi-

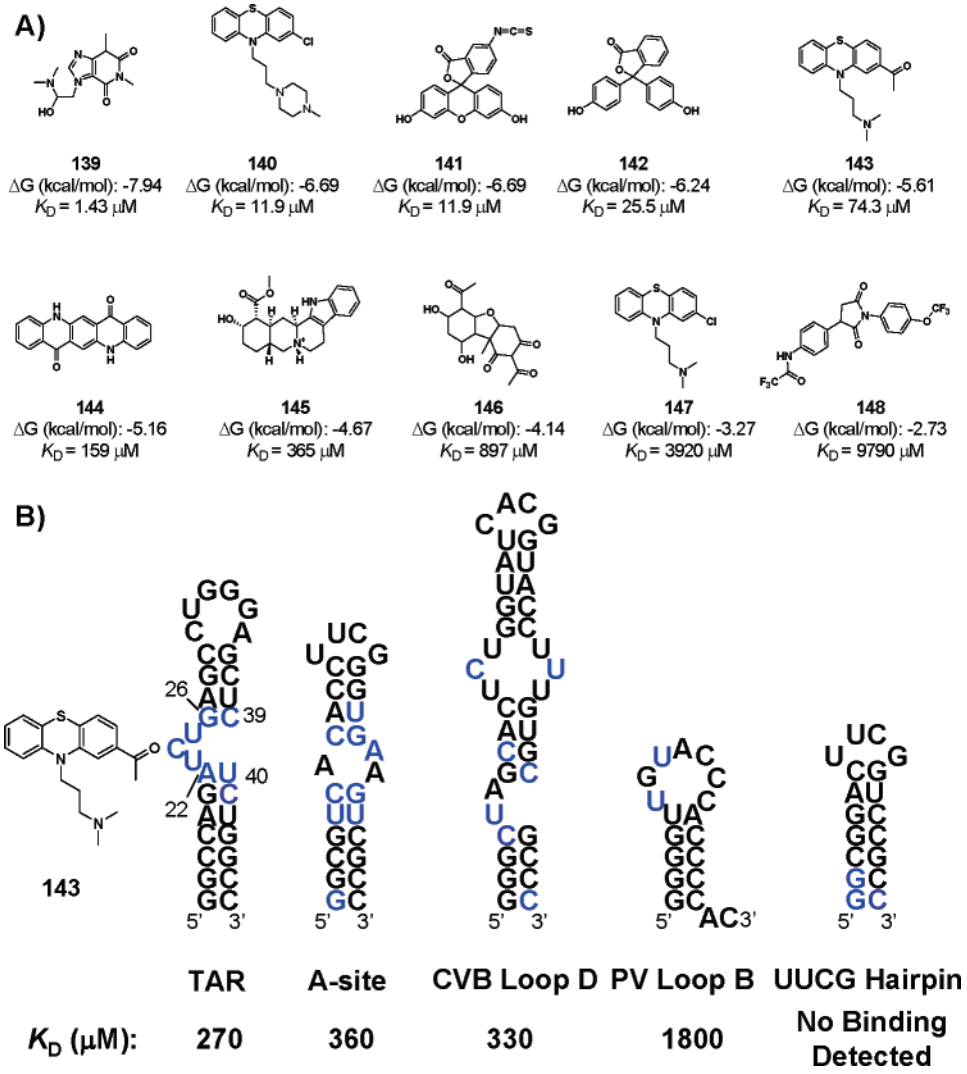


Figure 44. In silico screen for TAR binding compounds. (A) Ligands identified from the computational screen that were determined to disrupt the Tat−TAR interaction by gel-shift assay. The computationally determined free energies of binding are provided, from which the corresponding dissociation constant was calculated from $\Delta G = -RT \ln(K_a)$. (B) NMR-derived dissociation constants for **143** with TAR, A site, coxsackie virus B3 (CVB), polio virus (PV), and a UUCG hairpin RNA. Residues in blue are those that were found to interact with the ligand.

nation of various aminoglycosides bound to the 16S A-site RNA has captured the "functional" conformation that must be induced in order for this class of compounds to exert their antibiotic effect. As discussed in the previous paragraph, the majority of ligand-bound TAR complexes induce several key changes in the bulged region of TAR. Aboul-ela and co-workers recognized that the ability of various ligands to induce similar conformation within the TAR bulge likely signified a "functional" conformation for inhibiting Tat complexation.[342] In order to guide the development of a synthetic library, a Poisson−Boltzmann electrostatic surface potential was calculated for TAR from the previously determined structure bound to arginamide (with the arginamide ligand excluded) (see Figure 10). The results of the calculations were used to demonstrate that in the bound conformation two electronegative "hotspots" exist within the bulge, which result from an asymmetric distribution of charge within the bulged region.[342] Thus, the authors proposed that the proper placement of positively charged substituents directed toward the electronegative hotspots could lock TAR in a nonproductive conformation for Tat binding.

With this in mind, a series of flexible mono- and bis-guanidinylated ligands was synthesized and evaluated for disruption of the Tat−TAR complex by FRET (**149**−**159**, Figure 45Aand B).[342] Several compounds were identified with K_i values in the low micromolar range, with **149** emerging as the most potent competitive binder (K_i = 1.54 μM) (see Figure 45A). Although the length of the alkyl chain can be varied with little consequence on activity, an apparent minimum length is required. Substitution of the ether-linked guanidinum group with a primary amine generally resulted in loss of activity, although **153** produced similar K_i values as **151**. Through SPR experiments two binding events were detected for **149** binding to a biotinylated−TAR construct, one in the low micromolar range and the other at concentrations exceeding 100 μM. Analysis of the **149**−TAR complex by NMR determined that, upon binding, **149** induces many of the same conformational changes as arginamide. The scaffold was shown to bind in the major groove with the ether-linked guanidinium moiety positioned underneath U23, similar to the conformation of arginamide, where it is poised to form cation−π interactions with either U23 or A22 (see

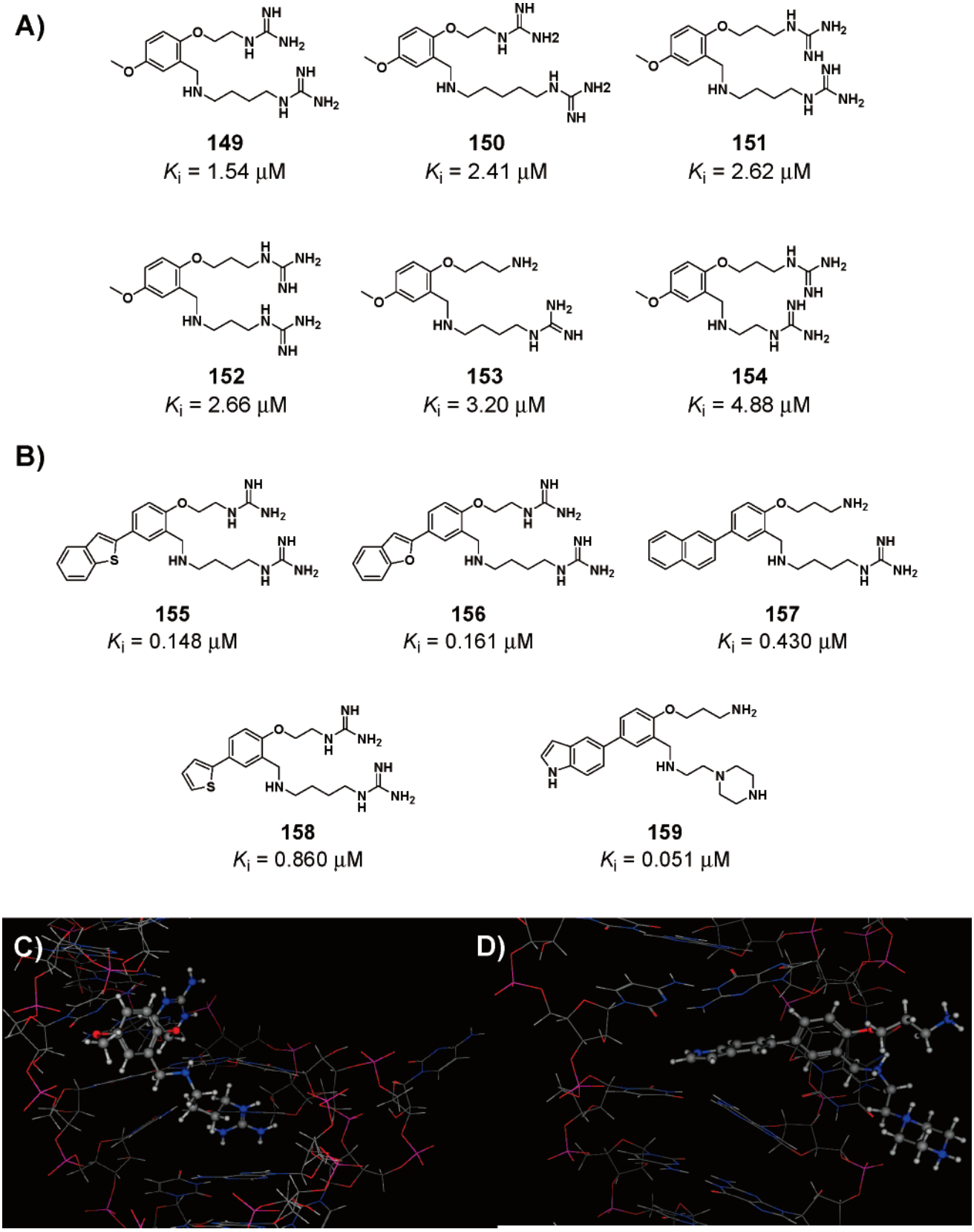


Figure 45. TAR electrostatic "hotspot" binding ligands. (A) Bis-guanidinylated series of ligands. (B) Second-generation hotspot binding ligands. (C) NMR structure of the **149**−TAR complex. (D) NMR structure of the **159**−TAR complex.

Figure 45C). The amine-linked guanidinium projects above U23 and may form similar cation−π interactions with U23. Importantly, as intended, each guanidinium group projects toward the electronegative "hotspots" within the binding pocket of the bulge.

Although **149** served to validate the "hotspot" hypothesis, the moderate binding affinities observed and less than drug-like properties of the bis-guanidinylated ligand series led Aboul-ela and co-workers to synthesize second-generation derivatives where the methoxy group was replaced with

various aryl substitutents.[343] On the basis of the structure of the **149**−TAR complex, aryl substituents were predicted to have reduced binding affinities as the methoxy group is in close contact with the phosphate backbone. However, several biaryl derivatives exhibited substantially enhanced potency in the FRET assay used to monitor the disruption to the Tat−TAR interaction (see Figure 45B). NMR experiments were conducted to examine the mode of binding of **156**. The chemical shift pattern differed from that of the parent compound **149** and other TAR−ligand complexes, suggesting that the biaryl ligands have a unique binding mode within the bulged region of TAR. As the binding mode of these biaryl derivatives is distinct, the requirement for the guanidinium groups was re-examined; previously, derivatives of the parent compound demonstrated a nearly absolute requirement for the presence of both guanidinium moieties.[342] Several compounds lacking both guanidinium substituents were determined to have only slightly reduced potency in the Tat−TAR disruption assay, and one compound from this series, **159**, exhibited a ~30-fold K_i improvement over **149**. Structural determination of **159** bound to TAR revealed that the indole ring intercalates between the A22-U40 and G26-C39 base pairs, while backbone contacts are mediated by the amine substituents (see Figure 45D). The authors report nearly 50 compounds with submicromolar K_i values belonging to the biaryl ligand series, highlighting the importance of base stacking in the Tat−TAR system.

9.1.3. Derivatives of the Neocarzinostatin Chromophore

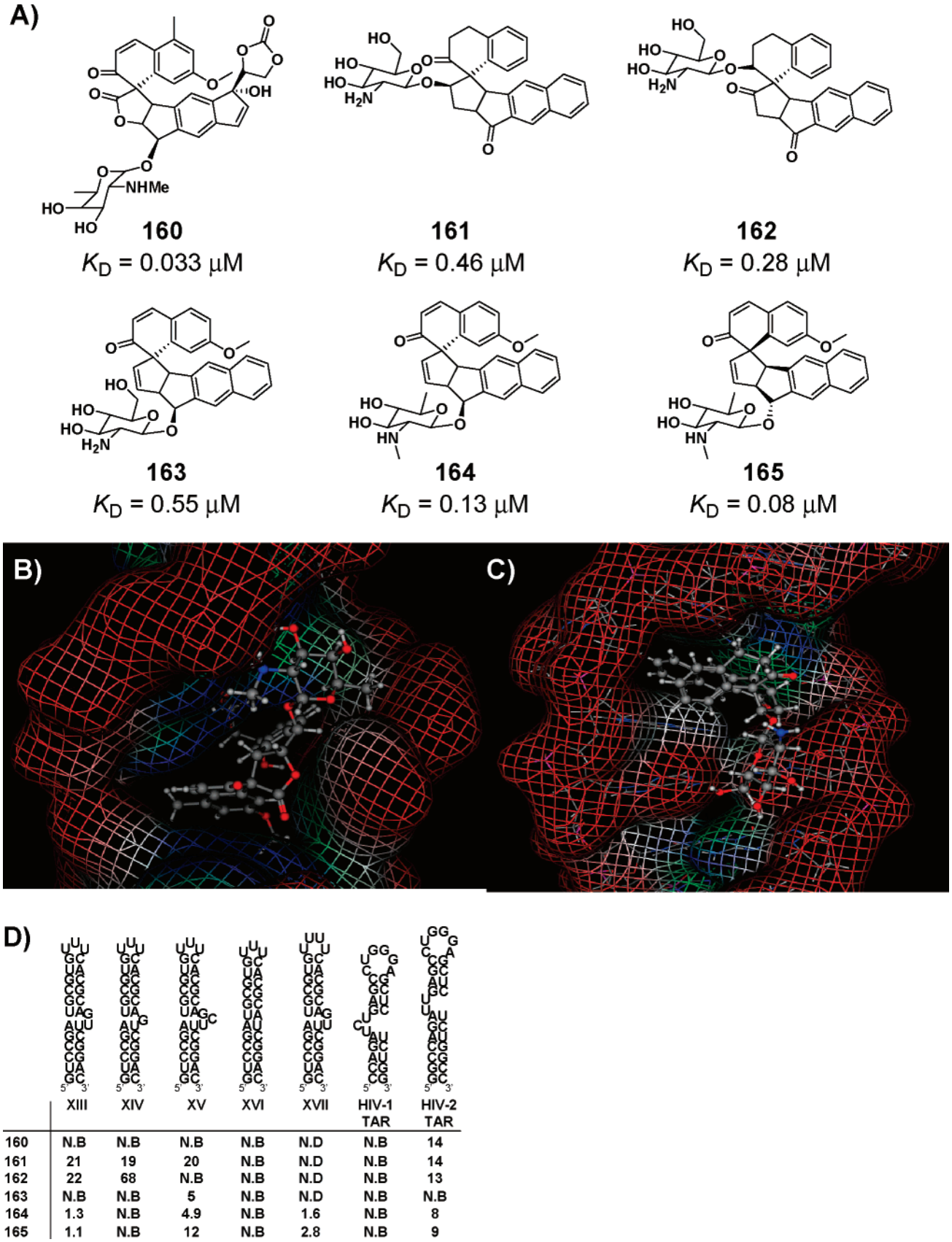
Thus far, design strategies for small molecule binders to the TAR bulge have attempted to achieve selective recognition by tuning intercalative and electrostatic binding modes. However, each class of secondary structure is likely to adopt a unique shape that may serve as a basis for selective recognition. For example, the thiol-independent breakdown product of neocarzinostatin chromophore (**160**, Figure 46A) features two aromatic systems joined by a spirocyclic junction;[344,345] the spirocyclic junction allows the two aromatic systems to stack on top of each other resulting in ~35° twist in the molecule. Goldberg and co-workers determined that this unique wedge-shaped compound exhibits affinity for a variety of DNA bulges with the greatest preference for two base bulges.[346] Structural investigation of **160** bound to a DNA duplex containing a two-base bulge revealed that **160** induces the two bulged nucleotides to protrude outward into solution, allowing the wedged scaffold to fill the "triangular prism pocket" previously occupied by the bulged nucleotides (see Figure 46B).[344−346] The two aromatic ring systems are proposed to mimic the geometry of the helical bases as **160** has the appropriate twist angle and stacks with the base pairs above and below the bulged region. However, due to the relative difficulty of synthesizing **160** and derivatives thereof, synthetically tractable mimics of **160** were sought. The design of the DDI ("double-decker intercalator") scaffold proved successful as several compounds (**161**−**165**) have been synthesized that exhibit mid-nanomolar binding affinities with a similar preference for two-base bulges[346−349] and binding geometries within DNA bulges (see Figure 46C).[350,351] Such results suggest that matching the 3D shape of ligands to complement the potentially unique 3D pockets formed by the various secondary structures may provide an unexplored avenue in the design of RNA binding compounds.

Recently, Goldberg and co-workers determined the effectiveness of **160** and various DDI derivatives for binding to various RNA bulges, including TAR (Figure 46D).[352] In all cases quenching of the native fluorescence from aromatic spirocyclic scaffolds was used to measure the binding affinity to various RNA bulges. The results from the binding analysis showed that for all compounds tested the binding affinities observed were at least 10-fold weaker as compared to a two-base DNA bulge. Compound **160** exhibited measurable binding affinity only to a two-base RNA bulge but with nearly 500-fold reduced affinity (K_D = 14 μM for RNA, versus K_D = 0.033 μM for DNA). DDI **161** proved to be the most general RNA bulge-binding compound with an average K_D of ~20 μM for one-, two-, and three-base bulges. Compounds **164** and **165** exhibited the tightest binding affinities for all compounds tested, K_D = 1.3 and 1.1 μM, respectively, for a two-base RNA bulge. These general selectivity trends were observed when comparing the binding affinities for the HIV-1 TAR bulge and HIV-2 TAR bulge; no binding was observed to the HIV-1 TAR bulge, while an approximately 10 μM affinity was observed for HIV-2 TAR bulge. These results serve to highlight that lessons learned from DNA-binding compounds may not directly translate to RNA binding due to the structural differences between the two forms of nucleic acids.

9.1.4. Aminoquinilones and Related Structures

In addition to the "rational" design strategies for binding to the bulged region of TAR, several discovery-based initiatives have been put forth. Mei and co-workers performed a high-throughput screen from which three compounds were identified that disrupt the Tat−TAR interaction.[87] Neomycin was one of the three compounds identified and discussed previously (section 4.4). Characterization of the remaining structures revealed that compound **166** (Figure 47A) inhibited the Tat−TAR interaction (IC_{50} = 1.3 μM) by binding to the bulged region of TAR. ESI-MS experiments further confirmed that **166** binds to TAR with an apparent 1:1 stoichiometry. Chemical foot-printing experiments with increasing concentrations of **166** revealed significant protection of C24 in a dose-dependent fashion, consistent with binding to the bulged region. However, foot printing with RNase V1, which cleaves duplex or stacked nucleotides, demonstrated that **166** exhibits a large footprint, extending well beyond the bulged region. To address concerns of selectivity, the same chemical and RNase foot printings were performed in the presence of either unlabeled tRNA or calf thymus DNA. In these experiments **166** showed a reduced ability to protect the RNA from degradation, indicating that nonspecific binding is a significant problem for compound **166**. The binding of **166** was markedly affected by the presence of both types of competitor nucleic acid, suggesting **166** is not selective for TAR or RNA in general.

As part of the development of anti-infective agents, Palu and co-workers determined that several 6-aminoquinolones (compounds **167**−**170**, Figure 47B) were active in an anti-infective HIV-1 in vitro cell culture assay.[353,354] Subsequent experiments suggested that **167** exerts its effects through inhibition of transcription, which prompted an investigation into the possibility of an interaction between **167** and the TAR RNA.[354] The binding of **167** to TAR was observed as a quench in fluorescence of the quinolone ring system. Using this assay, aminoquinolone **167** binds tightly to TAR (K_D = 19 nM);[354] there appears to be some variability in the reported binding affinity as Palu and co-workers also reported a



	XIII	XIV	XV	XVI	XVII	HIV-1 TAR	HIV-2 TAR
160	N.B	N.B	N.B	N.B	N.D	N.B	14
161	21	19	20	N.B	N.D	N.B	14
162	22	68	N.B	N.B	N.D	N.B	13
163	N.B	N.B	5	N.B	N.D	N.B	N.B
164	1.3	N.B	4.9	N.B	1.6	N.B	8
165	1.1	N.B	12	N.B	2.8	N.B	9

Figure 46. General bulge binding compounds. (A) Various DNA bulge-selective compounds. The binding constants shown are for DNA two-base bulges. (B) NMR structure of **160** bound to a two-base bulge within DNA. (C) NMR structure of **161** bound to a two-base bulge within DNA. (D) Binding affinities (in μM) of the DNA bulge-binding compounds to various RNA bulges. N.B indicates that no binding was detected. N.D. indicates that the binding affinity was not determined.

binding affinity of 200 nM.[355] In order to define the binding site of **167** on TAR, binding assays were performed using TAR constructs lacking either the hairpin loop or bulged region. The binding affinity for the TAR construct lacking

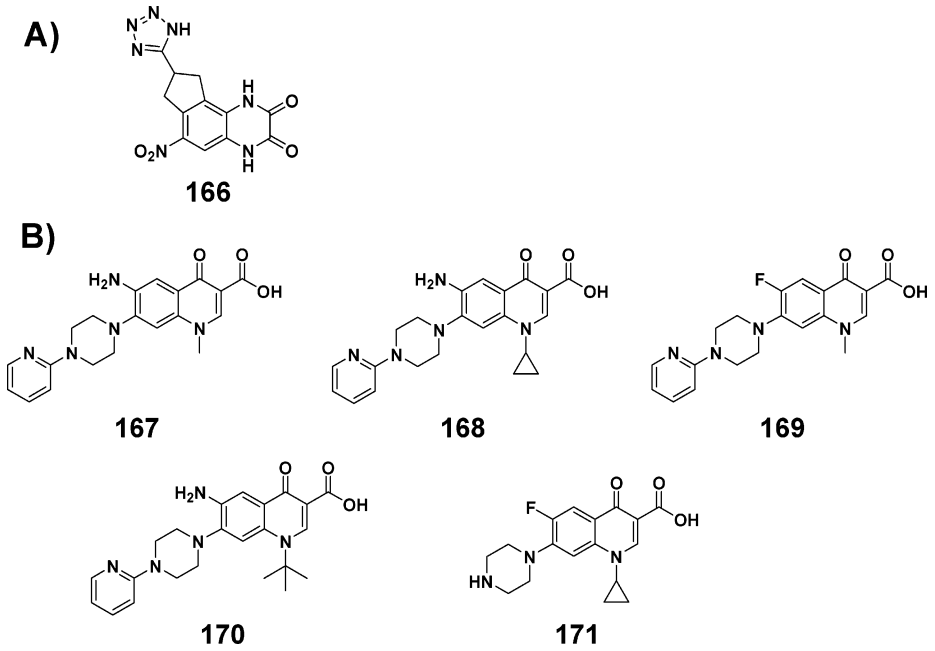


Figure 47. TAR binding ligands from discovery-based efforts. (A) **166** was one of three ligands identified from a high-throughput screen of a corporate library. (B) Several substituted 6-aminoquinolones (**167**−**170**) were discovered to possess anti-infective properties using in vitro cell culture models of HIV-1 infection. Compound **167** was the only compound identified from the original screen that was tested for and shown to bind to TAR.

the hairpin loop was essentially unchanged. However, no binding was observed for the TAR construct lacking the bulged region nor was any binding observed to tRNA, single-stranded DNA, or double-stranded DNA.[355] Standard gel-shift assays used to monitor the disruption of the Tat−TAR complex determined that **167** disrupts the complex with a K_i of 3.5 μM. Interestingly, ciprofloxacin (**171**), a structurally related quinolone, failed to disrupt the Tat−TAR complex, suggesting which functional groups might be important for the activity of **167**.[355]

9.2. Small Molecule Binders to the T Box RNA

Electrostatic interactions between small molecules and RNA are key to binding affinity, while at the same time they contribute heavily to nonselective binding. This duality has fueled the search for compounds with reduced reliance on electrostatic interactions for their binding to RNA. The oxazolidinone class of antibiotics appear to derive their in vivo efficacy by binding to the 23S rRNA subunit.[30] Although there is an apparent lack of structure−activity relationship concerning the oxazolidinone−RNA binding interactions, Bergmeier and co-workers noted that the less charged nature of the oxazolidinones may afford more selective RNA binding compounds.[74] A small library of 3,4,5-trisubstituted oxazolidinones was synthesized and evaluated for their ability to differentially bind the antiterminator T-box RNA or a C11U construct. Previous biochemical experiments suggested that the C11U mutation induces an alternative fold in the bulged region (Figure 48A);[71] thus, compounds that bind to the antiterminator T-box RNA but not the C11U T-box RNA likely recognize the specific conformation of the bulged region of the antiterminator T-box RNA. Characterization of the antiterminator structure revealed that residues A9−C12 adopt a stacked conformation.[73] As the binding of ligands to the bulged region is likely to induce a conformational change, replacing A9 with 2-aminopurine provided a convenient fluorescence-based binding assay given that the fluorescence intensity of 2-aminopurine is sensitive to change in its local environment. Through screening 27 3,4,5-trisubstituted oxazolidinones for binding to both the T-box RNA and C11U constructs (at 1 mM ligand concentration), it was determined that most compounds exhibited no preference for either RNA; however, a few compounds did exhibit some degree of preference between the two RNA constructs (compounds **172**−**176**, Figure 48B). Selectivity was assessed by comparing the change in relative fluorescence induced by a compound between the two RNA constructs. Caution should exercised when comparing normalized changes in fluorescence between different RNA constructs at a single ligand concentration because the extent of conformational change upon ligand binding is likely to be construct dependent and thus may not track with binding affinities; that is, a large charge change in fluorescence is not necessarily indicative of a strong binder. As **173** exhibited the greatest degree of selectivity (a ∼4-fold preference for the T-box RNA over the C11U mutant construct) compound **173** was examined further. Using a previously developed FRET assay, where the T-box or C11U construct is end labeled and hairpin loop labeled with FRET pairs,[72] addition of **173** caused a dose-dependent change in FRET efficiency from which a binding constant for the T-box RNA and the C11U mutant construct was determined, with K_D values of 3.4 and 25 μM, respectively (Figure 48C). Importantly, the 7-fold selectivity observed in the FRET assay is consistent with the results of the 2-aminopurine screen. Evaluating other derivatives of **173** revealed that removal of the methoxyphenyl (**177**) reduced the binding affinity by ∼3-fold but enhanced the selectivity (Figure 48C). Also, dramatic changes to the ester side chain resulted in substantially altered binding affinities and selectivity profiles.

9.3. Small Molecule Binders to the Iron Response Element (IRE) RNA

Most research in the small molecule−RNA-binding arena has concentrated on the inhibition of translation, either

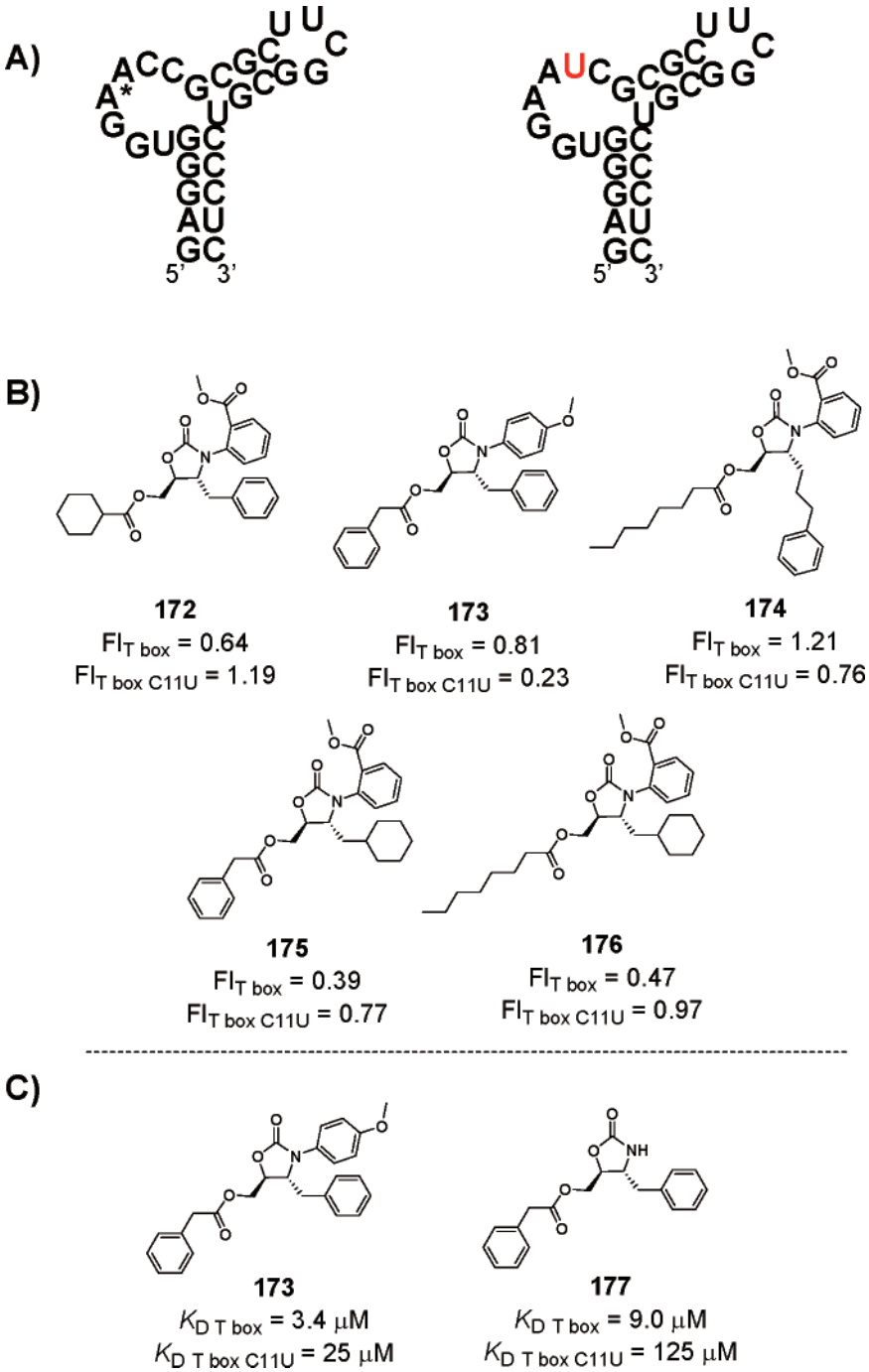


Figure 48. Oxazolidinones as T-box binding ligands. (A) Secondary structure model of the T-box RNA and the C11U (shown in red) mutant construct. A* indicates the location of the 2-aminopurine residue. (B) Selectivity profile of compounds with a preference for the T-box RNA construct. (C) Dissociation constants for **173** and **177** with the two RNAs.

transcript specific[134,162] or by inhibiting the ribosome. Realizing the therapeutic potential of small molecule *activators* of translation, Thorp and co-workers examined ligand-binding effects in the ferritin IRE−IRP system.[155] Under typical cellular conditions the ferritin mRNA is repressed by the binding of IRP to the ferritin IRE in the 5′-UTR. Thus, a small molecule that could prevent the binding to IRP, or otherwise increase translation by binding to ferritin IRE (fIRE), would serve to increase the levels of ferritin protein. Using chemical footprinting methods to monitor ligand binding, the natural product yohimbine (**145**) was determined to bind to residues above and below the bulged region of the fIRE, which the authors interpreted as being indicative of selective binding to the internal loop region. However, protection of weakly cleaved bases between the bulge and hairpin loop region was observed, suggesting that yohimbine binds multiple secondary structures within fIRE. Utilizing the native fluorescence of yohimbine, titration of fIRE resulted in the decreased fluorescence of yohimbine from which a K_D of 3.9 μM was calculated (Figure 49). Examining the affinity of yohimbine for a variety of bulged RNAs yielded binding affinities in the low micromolar range (average $K_D = 7.9$ μM). Having established that yohimbine binds to the IRE sequence, the ability of yohimbine to inhibit the interaction between fIRE and IRP was evaluated by gel-shift assays. Addition of yohimbine at 40 μM caused an 8% increase of free fIRE; given the tight binding interaction between IRP and the fIRE ($K_D = 90$ pM) and the modest

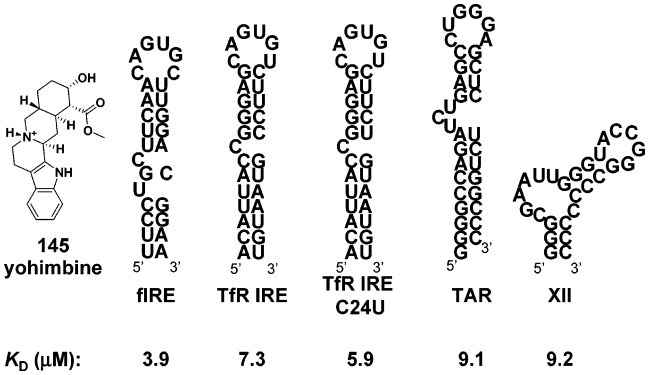


Figure 49. Binding affinity of yohimbine for a variety of RNA targets.

binding affinity of yohimbine for IRE (K_D = 3.9 μM) the low level of disruption observed is not surprising. Unexpectedly, preincubation of 40 μM yohimbine with full-length ferritin mRNA followed by standard eukaryotic cell-free translation assay showed an increase in translation of the ferritin mRNA by 40%. Furthermore, cloning the fIRE upstream of a luciferase reporter gene produced a similar increase in translation. Such dramatic increases in translation coupled with the weak competitive binding of yomhibine for fIRE suggests that decreased IRP binding together with unwinding of the IRE structure by yohimbine results in the increased translation observed. These results point to the exciting prospect of targeting RNA to selectively increase translation of a target protein.

10. Hairpin Loop Binding Compounds

RNA hairpin loops occur when a sequence region folds back on itself to form a duplex that is linked through single-strand nucleotides; the resulting structures are often called stem-loops. In terms of abundance of secondary structure, it is estimated that RNA hairpin loops are second only to duplex regions.[356] The predominance of hairpin loops is matched by their functional significance as they provide sites of nucleation for RNA folding and participate in RNA−protein and RNA−RNA interactions.[357] The absolute prevalence of various RNA hairpin loop sizes and sequences is unknown; however, it seems reasonable that thermodynamic stability would correlate with prevalence in vivo. Solely on the basis of hairpin loop size, hexa- and heptaloops have been determined to be the most thermodynamically stable as six to seven nucleotides is the ideal length for spanning the A-form helix.[357] Formation of larger hairpin loops is penalized by unfavorable entropy, and the loops tend to be less stable than their smaller counterparts. However, the sequence of a given hairpin loop can contribute greatly to its overall stability. In fact, entire classes of "exceptionally stable" hairpin loops have been discovered which are considerably smaller than hexaloops. Tetraloops consisting of UUCG, GNRA (where R = purine), and YNMG (where Y = pyrimidine and M = adenine or cytosine) sequences are the primary examples of hairpin loop sequences dictating stability.[356,358,359]

10.1. Binders to the TAR Hairpin Loop

Despite the biological precedent for hairpin loops to participate in interactions with proteins and RNA, few small molecules have been shown to bind with appreciable affinity to hairpin loops of any size or sequence. The first small molecule with demonstrated affinity for hairpin loops was identified from the screen performed by Mei and coworkers.[87] In total three compounds were identified that disrupted the Tat−TAR interaction as assessed by gel-shift assays. Two compounds were determined to bind to nucleotides in and around the bulged region of TAR. However, the third hit compound (Figure 50, **178**) was shown to bind the hairpin loop region of TAR. Chemical and RNase foot printing confirmed that A35 and G36 of the TAR construct depicted in Figure 44B were protected by **178** in a dose-dependent fashion. Interestingly, the same RNase footprinting experiments failed to detect protection of guanine residues in the apex of the hairpin loop. Further confirmation of hairpin loop binding was provided by ESI-MS experiments. Incubation of **178** with either TAR or a TAR construct wherein the hairpin loop was replaced with a PEG linker resulted in a mass shift only for TAR, while no binding was observed for the PEG-linked TAR construct.

10.2. Binders to the U1A snRNA Hairpin Loop

RNA hairpin loops are common binding sites for proteins. Noting the dearth of investigations focused on identification of small molecule disruptors of RNA−protein interactions other than the Rev−RRE and Tat−TAR interactions, Baranger and Gayle initiated the search for binders to stem loop 2 of the U1A snRNA, which is bound in vivo by the U1A protein and controls splicing of eukaryotic pre-mRNAs.[360] The U1A protein binds to stem loop 2 by making critical contacts in the hairpin loop itself. The authors reasoned that because the tight binding interaction between U1A and stem-loop 2 required the GC closing base pair, ligands with a demonstrated capacity to bind to similar binding sites might prove to be a suitable starting point for identification of disruptors. Comparing the secondary structure of the stem-loop 2 RNA with ligand specificity reported in the literature, **134** and **178** (Figure 50A) were selected for evaluation as disruptors of the U1A−stem-loop 2 interaction. These ligands were chosen because previous investigations demonstrated that both ligands disrupted the Tat−TAR interaction by binding to CG or GC base pairs and the adjacent secondary structure. Using a gel-shift assay, **134** inhibited the U1A−RNA interaction with an IC_{50} = 1.0 μM; use of **178** at concentrations up to 10 mM failed to disrupt the complex. Acridine, spermidine, or the closely related quinacrine (**9**) were equally ineffective, consistent with previous reports suggesting that both components of **134** are required for activity.[334] Using the native fluorescence of acridine, titration experiments with stem-loop 2 determined that the ligand:RNA stoichiometry was 2:1. RNase footprint experiments revealed that in addition to protection of residues C8 and U11, the cleavage rate for most of the nucleotides in the hairpin loop was altered upon binding. However, the authors note that due to the inherent flexibility of this large hairpin loop the effects of direct binding or ligand-induced conformational changes cannot be differentiated.

10.3. Binders to the GNRA Tetraloop

Continuing along the theme of targeting biologically relevant RNA hairpin loops, Baranger and Yan examined the ability of commercially available RNA-binding compounds to bind a GNRA tetraloop.[220] GNRA tetraloops mediate tertiary contacts critical for group I introns, hammerhead ribozymes, and the ribosome.[361−363] As mentioned

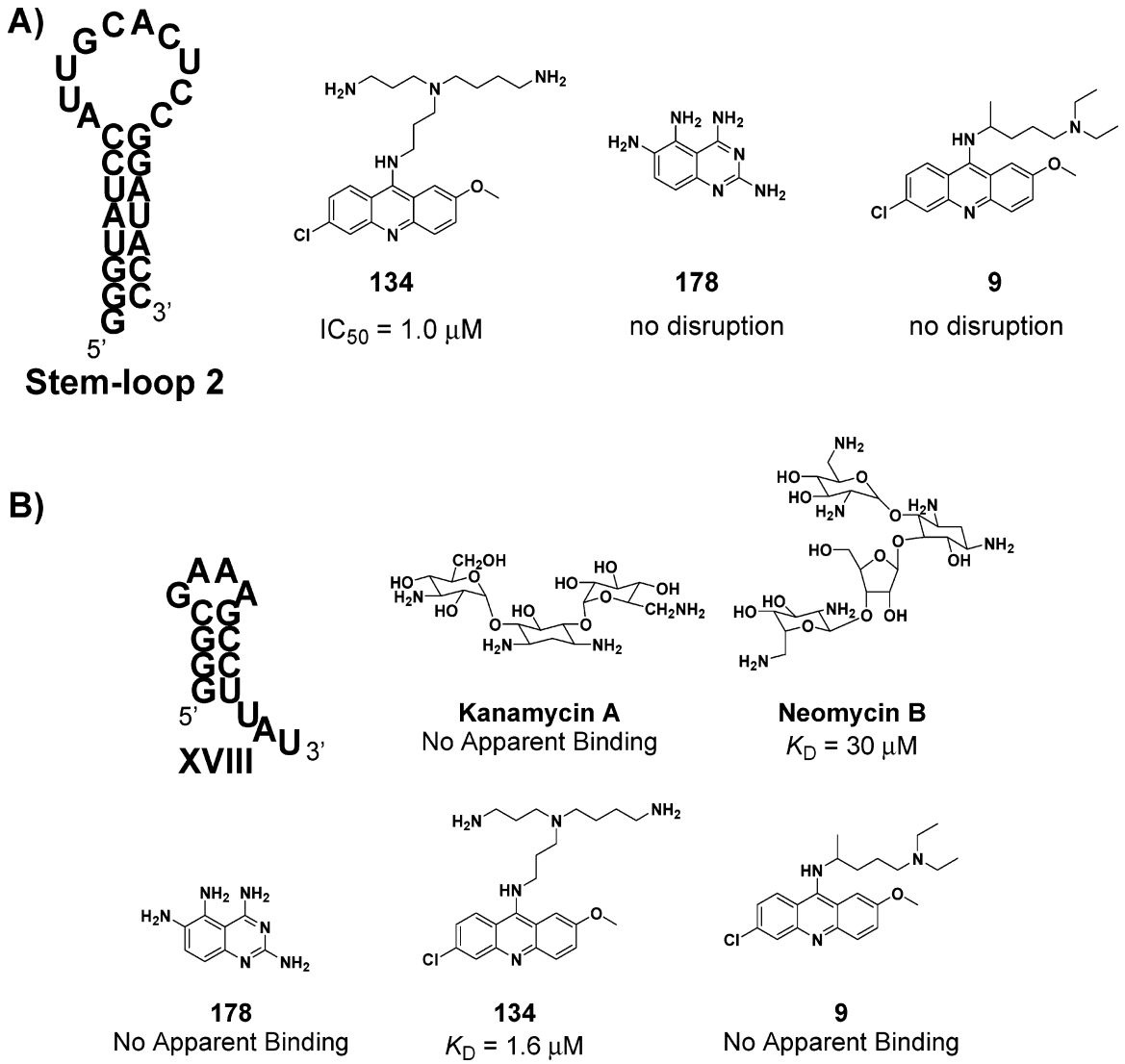


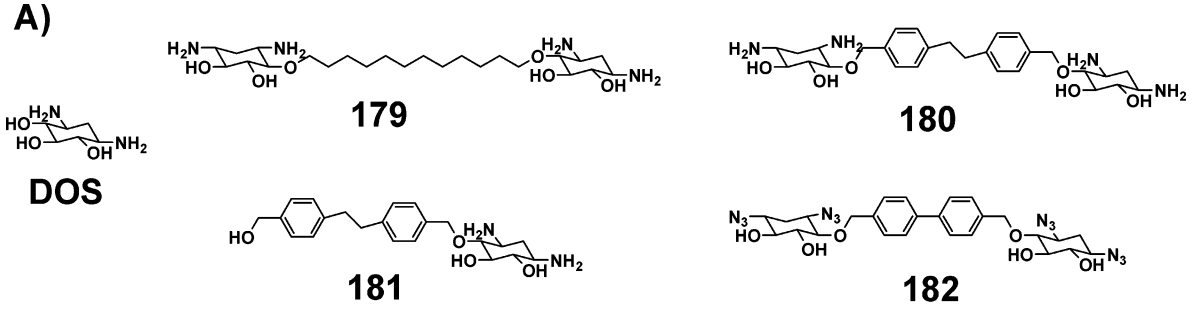
Figure 50. Compounds that bind to RNA stem loops. (A) Compound **134** was found to bind stem-loop 2 of the U1A snRNA and disrupt the U1A−RNA interaction with an IC_{50} of 1.0 μM, whereas compounds **178** and **9** had no effect. (B) Several RNA-binding compounds were assessed for their ability to bind to a GNRA tetraloop construct (**XVIII**); compound **134** was the best binder in this study.

previously, GNRA tetraloops are part of the class of exceptionally stable tetraloop sequences. The enhanced stability arises from the conformation of nucleotides in the hairpin loop; the first and fourth nucleotides of the hairpin loop form a noncanonical G−A base pair, while nucleotides two through four form stacking interactions. This array of hydrogen-bonding and stacking interactions may prove suitable for tight, selective recognition by small molecule binders. Of the compounds tested only **134** bound to the GRNA hairpin loop with low micromolar affinity, $K_D = 1.6$ μM (Figure 50B). Further characterization of the ligand−RNA interaction revealed a binding stoichiometry of 1:1. Footprinting experiments were consistent with binding of **134** to the hairpin loop region, although the extent of protection and exact location of the binding site was dependent on the method of cleavage (chemical versus RNase). Using a computational platform, Baranger and co-workers recently disclosed a derivative of **134** that exhibits greater selectivity for the GNRA tetraloop sequence over duplex or single-stranded RNAs.[364] Also, in a separate endeavor Baranger and co-workers discovered a quinolone derivative which binds the 3′-terminal nucleotides of their GNRA construct.[256]

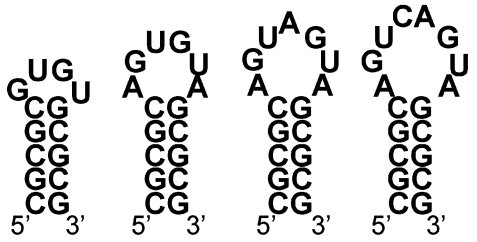
10.4. Deoxystreptamine (DOS) Dimers as RNA Hairpin Loop Binders

RNA hairpin loops participate in numerous macromolecular interactions; thus, the lack of compounds with sufficient affinity for hairpin loops makes a large number of RNA-mediated interactions resistant to small molecule approaches. Hergenrother and co-workers sought to develop general hairpin loop binding compounds by synthesizing dimers of 2-deoxystreptamine (DOS). Monomeric DOS was previously demonstrated to bind to areas of perturbed secondary structure, including hairpin loop regions, with millimolar affinity.[365] Reasoning that the linking of DOS monomers via a tether of appropriate length would provide suitable affinity and flexibility for targeting hairpin loops, a small collection of DOS dimers was synthesized. By varying the linker length and composition, weak to moderate binding affinities were observed for tetra-, hexa-, hepta-, and octa hairpin loops RNA constructs with compounds such as **179** and **180** (Figure 51A and B).[225] In cases when binding was observed similar binding affinities were obtained across the entire panel of RNA hairpin loops, establishing the DOS dimers as a class of general RNA hairpin loop binding compounds. The DOS

A)



B)

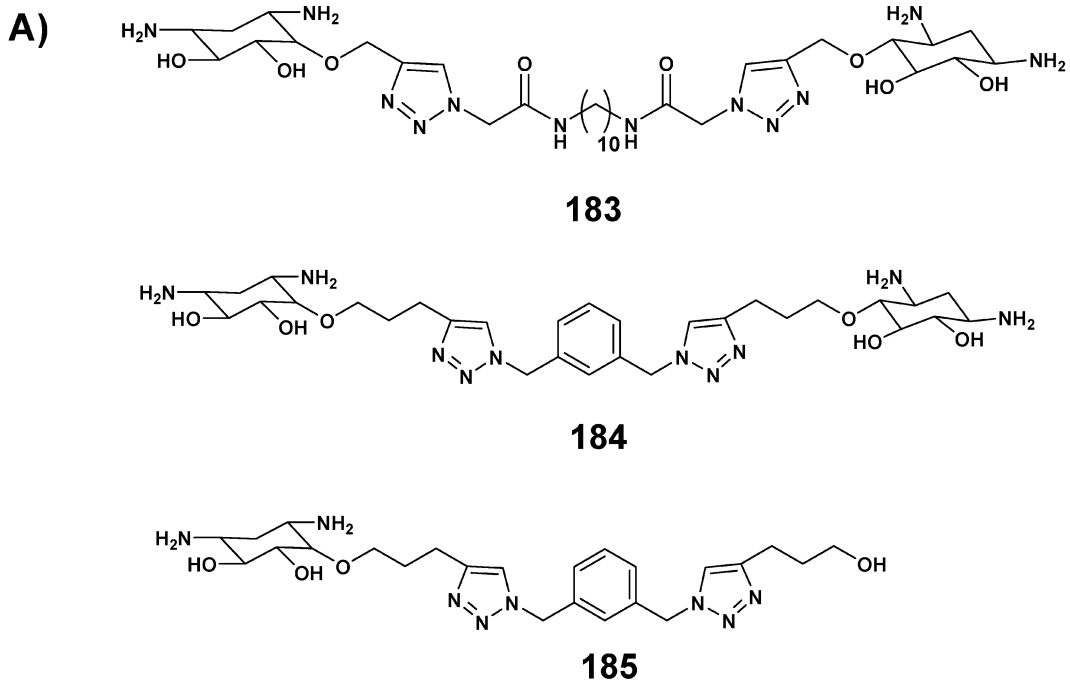


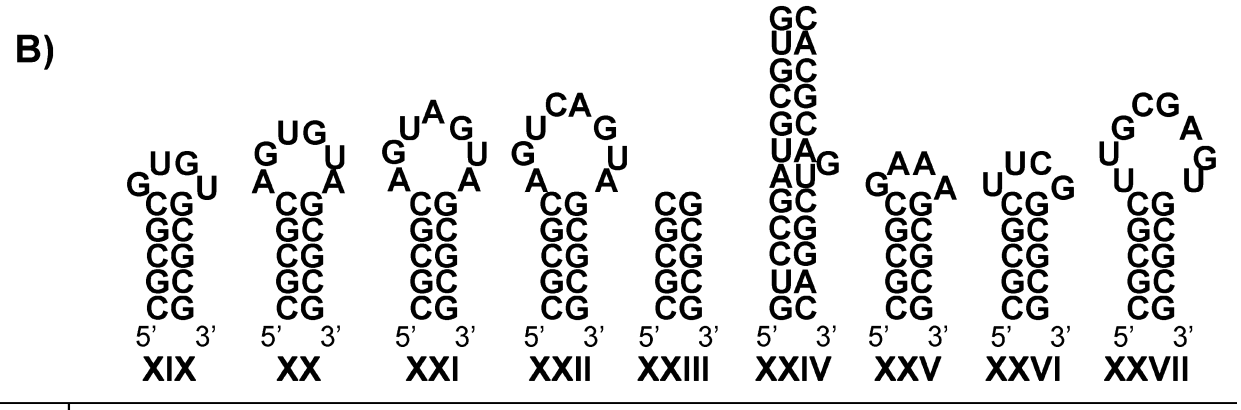
compound	XIX	XX	XXI	XXII
DOS	>1000	>1000	>1000	>1000
179	34	11	11	8
180	16	6	6	6
181	>1000	>1000	>1000	>1000
182	>250	>250	>250	>250

Figure 51. DOS dimers as a general class of RNA hairpin loop binding ligands. (A) The structure of 2-deoxystreptamine (DOS), two DOS-dimers (**179** and **180**), and two control compounds (**181** and **182**). (B) DOS dimers bind to RNA hairpin loops. Shown are the dissociation constants (in μM) of the various compounds for the hairpin loop structures.

dimer **180** proved to be the most potent binder in this initial study, with a 6 μM binding affinity for RNAs **XX**−**XXII**, and slightly reduced affinity for the RNA **XIX** (K_D = 16 μM).[225] Critical control compounds demonstrated the necessity of the dimeric nature of the compounds and the importance of electrostatic interactions as the mono-linked DOS moiety (**181**) and the tetraazide **182** exhibited no detectable binding affinity.[225] Binding to the hairpin loop region was confirmed by RNase footprinting, which revealed that several residues within the hairpin loop are protected upon binding. Also, binding appears to be independent of sequence as **180** demonstrated equal binding affinity for a variety of heptaloop sequences, further establishing the generality of binding.

In a continued effort to establish RNA hairpin loop binding "modules",[229] Hergenrother and co-workers synthesized a combinatorial library of DOS dimers utilizing the facile copper-catalyzed variant of the Huisgen 1,3-dipolar cycloaddition of alkynes and azides. Linker length was varied by the synthesis of DOS monomers functionalized with alkynes of differing length, while diverse functionality of the linker was generated by a set of 35 diazides. The 105-membered compound library was screened for dose-dependent binding to RNAs **XIX**−**XXII** using the end-label method (Figure 52).[226] The results of the screen led to identification of five compounds which appeared to be selective for RNA **XXII**, while two compounds were found to be selective for RNA **XIX**. Subsequent investigation determined that **183** was the most selective for tetraloops and **184** was the most selective for octaloops, exhibiting at least 5- and 30-fold selectivity, respectively, over the hairpin loops assayed (Figure 52A and B).[226] In addition to achieving hairpin loop size selectivity, **183** and **184** bound their respective hairpin loops with substantially enhanced binding affinity as compared to the original DOS dimers with K_D values of 0.30 and 0.32 μM, respectively.[226] Furthermore, consistent with the initial design of the DOS dimers, the dimeric nature of the compounds is essential as the mono-linked versions (such as **185**) exhibited a substantially altered binding profile. Binding to the hairpin loop region was confirmed by modest protection observed during RNase footprinting. As a further measure of selectivity, **183** and **184** showed little to no binding to a simple RNA duplex or single base bulge. Selectivity was further evaluated by competition binding assays using tRNA as the competitor nucleic acid; the binding of **183** and **184** to their respective RNAs was unaffected by the 100-fold (base) excess of tRNA, suggesting both ligands may have suitable binding properties for use inside the cell. The generality of tetra- and octaloop binding was probed by assaying **183** and **184** for binding to biologically relevant hairpin loop sequences. DOS dimer **183**





	XIX	XX	XXI	XXII	XXIII	XXIV	XXV	XXVI	XXVII
183	0.30	6.49	1.60	4.82	>25.0	5.54	0.10	0.07	>25.0
184	>25.0	10.0	9.87	0.32	>25.0	>25.0	1.36	1.54	0.15
185	>25.0	>25.0	>25.0	>25.0	1.05	9.60	>25.0	4.34	>25.0

Figure 52. DOS dimers as size-specific ligands for RNA hairpin loops. (A and B) Compound **183** has specificity for RNA tetraloops, while compound **184** specifically binds to RNA octaloops. Shown are the dissociation constants (in μM) for the compounds with the indicated RNA.

bound a UUCG and GNRA tetraloop (RNAs **XXV** and **XXVI**) with ~100 nM affinity, while **184** bound the octaloop of the Hepatitis Delta Virus C (HepC) IRES (RNA **XXVII**) with a K_D of 150 nM; importantly, selectivity was also retained as **184** was found to be at least 13-fold more selective for RNA **XXVII**, and **184** also appeared to have minimal off-target binding (Figure 52).[226]

Subsequent in vitro characterization of **184** with an RNA octaloop revealed that the electrostatic contribution to binding is substantially lower than that of the aminoglycosides.[229] Only 20% of the total binding energy for **184** was determined to be due to electrostatic interactions, as compared to the aminoglycosides which derive >50% of their free energy of binding from electrostatic interactions.[229] This reduced reliance on electrostatic interactions was confirmed by isothermal titration calorimetry. In these experiments the determined binding affinity was inversely correlated with the enthalpic contribution to binding, suggesting that often overlooked entropic factors such as desolvation or cation release may be critical components for achieving selective recognition for RNA binding small molecules.[229]

11. Conclusion

Considerable progress has been made toward the goal of selectively targeting RNA with small molecules. Screening followed by synthetic optimization is likely to provide small molecules that bind in vitro to the intended RNA target with low- to submicromolar binding affinities. A variety of RNA-binding assays are available to determine the strength of a

RNA−ligand interaction in addition to biophysical experiments that provide additional information about the binding site and the kinetics/thermodynamics of binding (e.g., foot printing, SPR, and ITC). Thus, investigators are equipped with a battery of primary and secondary assays to evaluate and characterize novel RNA−ligand interactions in vitro.

However, significant hurdles must be overcome in order to turn the therapeutic potential of RNA ligands into reality. Below are listed four challenges for the next 10 years of small molecule−RNA research.

(1) Wanted: High-throughput screens. Currently a bottleneck for targeting RNA is the limited number of suitable small molecule ligands. Additional general, high-throughput methods for identification of RNA−ligand interactions would be tremendously enabling. Conformational change associated with ligand binding to fluorescently labeled RNA constructs is a widely used method to determine the binding constants of RNA−ligand interactions and can (in some cases) be used in a high-throughput screen. NMR[234] and ESI-MS[313] techniques have been used in screening efforts, although specialized equipment and expertise are typically required. Displacement of a fluorescent ligand is another useful screening method, although it requires known small molecule ligands for the target RNA. However, a simple and general high-throughput screening method would greatly facilitate the discovery of novel RNA-binding ligands.

(2) Wanted: Detailed characterization of small molecule−RNA interactions. Even though scores of compounds have been identified that bind their RNA targets with high affinity, in many cases little work has been done to characterize these interactions. To address the lack of quantitative insight into ligand−RNA interactions, Pilch and co-workers characterized the binding of various aminoglycosides to the 16S rRNA A site by ITC.[206−208,210] At the time of their investigations it was well understood that electrostatic interactions contributed substantially to the strength of RNA−aminoglycoside interactions. However, Pilch and co-workers were able to quantify the contribution of electrostatic interactions to the total free energy of binding; it is now known that >50% of the total free energy of binding by the aminoglycosides to the A site is contributed by electrostatic interactions at physiological pH.[210] Previously, Hergenrother and co-workers demonstrated that electrostatic interactions are essential for several DOS dimers that bind to various RNA hairpin loops;[225,226,229] however, analysis by ITC found that electrostatic interactions accounted for only ∼20% of the total free energy of binding.[229] Thus, it appears that even though electrostatic interactions are a crucial feature of RNA binding, such interactions can contribute differently to the overall free energy of binding depending on the system. Because of the lack of quantitative data on many RNA−ligand interactions, it remains unclear whether the aminoglycosides reliance on electrostatic interactions is the exception or the rule. In order to develop and evaluate the rules of RNA binding, detailed biochemical and biophysical investigations need to be performed on multiple RNA−ligand systems.

(3) Wanted: Exploitation of additional RNA targets. Despite the potential of many types of RNA targets, the 16S A site, RRE, and TAR are the most popular targets and have become somewhat of a proving ground for new RNA-binding compounds. The examples presented in the molecular targets section of this review (section 4) show that there are many additional RNA targets where the secondary structure of the RNA is known, and a site for ligand binding has been defined. There is no shortage of RNA targets, and by branching out from the standard targets the likelihood for novel discoveries and greater impact increases.

(4) Wanted: RNA ligands with activity in cell culture and in vivo. Given the substantial challenges in simply finding a ligand that binds to RNA in vitro, it is rare that these compounds are carried forward into cell culture or in vivo assays (the antibacterial compounds are an obvious exception to this). Until RNA binding compounds can be shown to consistently hit RNA targets inside the cell and exert their desired biological effect, the field of small molecule−RNA binding will not have a substantial impact outside of the chemical biology realm. In short, with the obvious potential and the great number of targets, small molecule ligands for RNA could indeed become the next wave of drug discovery. At some point, however, this potential must be turned into tangible results, and in vitro studies must be translated into cell culture and in vivo results.

Small molecule−RNA binding has progressed significantly since the last *Chemical Reviews* article on this subject. Although even greater progress is required before RNA targets are given the same consideration as protein targets in the drug discovery realm, the proper tools and techniques are currently in place. Hopefully within the next 10 years multiple in vivo success stories with RNA ligands will emerge.

12. Acknowledgments

The authors are grateful to the National Institutes of Health (NIHGMS R01-GM68385) and the Office of Naval Research (N00014-02-1-0390) for supporting research in the area of small molecule−RNA binding.

13. References

(1) Bayne, E. H.; Allshire, R. C. *Trends Genet.* **2005**, *21*, 370.
(2) Jovanovic, M.; Hengartner, M. O. *Oncogene* **2006**, *25*, 6176.
(3) Sen, G. L.; Blau, H. M. *FASEB J.* **2006**, *20*, 1293.
(4) Manche, L.; Green, S. R.; Schmedt, C.; Mathews, M. B. *Mol. Cell. Biol.* **1992**, *12*, 5238.
(5) Sledz, C. A.; Holko, M.; de Veer, M. J.; Silverman, R. H.; Williams, B. R. *Nat. Cell. Biol.* **2003**, *5*, 834.
(6) Fedor, M. J.; Williamson, J. R. *Nat. Rev. Mol. Cell. Biol.* **2005**, *6*, 399.
(7) Drews, J. *Nat. Biotechnol.* **1996**, *14*, 1516.
(8) Drews, J.; Ryser, S. *Nat. Biotechnol.* **1997**, *15*, 1318.
(9) Hopkins, A. L.; Groom, C. R. *Nat. Rev. Drug Discovery* **2002**, *1*, 727.
(10) Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. *Nucleic Acids Res.* **2006**, *34*, D668.
(11) Imming, P.; Sinning, C.; Meyer, A. *Nat. Rev. Drug Discovery* **2006**, *5*, 821.
(12) Zheng, C.; Han, L.; Yap, C. W.; Xie, B.; Chen, Y. *Drug Discovery Today* **2006**, *11*, 412.
(13) Zheng, C. J.; Han, L. Y.; Yap, C. W.; Ji, Z. L.; Cao, Z. W.; Chen, Y. Z. *Pharmacol. Rev.* **2006**, *58*, 259.
(14) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. *Nat. Rev. Drug Discovery* **2006**, *5*, 993.
(15) Fletcher, S.; Hamilton, A. D. *Curr. Opin. Chem. Biol.* **2005**, *9*, 632.
(16) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284.
(17) Dervan, P. B.; Doss, R. M.; Marques, M. A. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 373.
(18) Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* **2004**, *11*, 1583.
(19) Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 3313.
(20) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12063.
(21) Wagner, E. G.; Altuvia, S.; Romby, P. *Adv. Genet.* **2002**, *46*, 361.

(22) Sevignani, C.; Calin, G. A.; Siracusa, L. D.; Croce, C. M. *Mamm. Genome* **2006**, *17*, 189.
(23) DeJong, E. S.; Luy, B.; Marino, J. P. *Curr. Top. Med. Chem.* **2002**, *2*, 289.
(24) Winkler, W. C.; Breaker, R. R. *ChemBioChem* **2003**, *4*, 1024.
(25) Schroeder, R.; Barta, A.; Semrad, K. *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 908.
(26) Vicens, Q.; Westhof, E. *ChemBioChem* **2003**, *4*, 1018.
(27) Tor, Y. *Biochimie* **2006**, *88*, 1045.
(28) Brisson-Noel, A.; Trieu-Cuot, P.; Courvalin, P. *J. Antimicrob. Chemother.* **1988**, *22* (Suppl B), 13.
(29) Brodersen, D. E.; Clemons, W. M., Jr.; Carter, A. P.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. *Cell* **2000**, *103*, 1143.
(30) Bozdogan, B.; Appelbaum, P. C. *Int. J. Antimicrob. Agents* **2004**, *23*, 113.
(31) Chow, C. S.; Bogdan, F. M. *Chem. Rev.* **1997**, *97*, 1489.
(32) Turner, J. J.; Fabani, M.; Arzumanov, A. A.; Ivanova, G.; Gait, M. J. *Biochim. Biophys. Acta* **2006**, *1758*, 290.
(33) Turner, J. J.; Jones, S.; Fabani, M. M.; Ivanova, G.; Arzumanov, A. A.; Gait, M. J. *Blood Cells Mol. Dis.* **2007**, *38*, 1.
(34) Paulasova, P.; Pellestor, F. *Ann. Genet.* **2004**, *47*, 349.
(35) Rao, S. T.; Rossmann, M. G. *J. Mol. Biol.* **1973**, *76*, 241.
(36) Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113.
(37) Tinoco, I., Jr.; Bustamante, C. *J. Mol. Biol.* **1999**, *293*, 271.
(38) Shelton, V. M.; Sosnick, T. R.; Pan, T. *Biochemistry* **1999**, *38*, 16831.
(39) Weeks, K. M.; Crothers, D. M. *Science* **1993**, *261*, 1574.
(40) Hermann, T.; Patel, D. J. *J. Mol. Biol.* **1999**, *294*, 829.
(41) Chin, K.; Sharp, K. A.; Honig, B.; Pyle, A. M. *Nat. Struct. Biol.* **1999**, *6*, 1055.
(42) Carlson, C. B.; Stephens, O. M.; Beal, P. A. *Biopolymers* **2003**, *70*, 86.
(43) Jin, E.; Katritch, V.; Olson, W. K.; Kharatisvili, M.; Abagyan, R.; Pilch, D. S. *J. Mol. Biol.* **2000**, *298*, 95.
(44) Chen, Q.; Shafer, R. H.; Kuntz, I. D. *Biochemistry* **1997**, *36*, 11402.
(45) Arya, D. P.; Xue, L.; Willis, B. *J. Am. Chem. Soc.* **2003**, *125*, 10148.
(46) Autexier, C.; Lue, N. F. *Annu. Rev. Biochem.* **2006**, *75*, 493.
(47) Ogle, J. M.; Ramakrishnan, V. *Annu. Rev. Biochem.* **2005**, *74*, 129.
(48) Hermann, T. *Curr. Opin. Struct. Biol.* **2005**, *15*, 355.
(49) Poehlsgaard, J.; Douthwaite, S. *Nat. Rev. Microbiol.* **2005**, *3*, 870.
(50) Hermann, T. *Biopolymers* **2003**, *70*, 4.
(51) Yoshizawa, S.; Fourmy, D.; Puglisi, J. D. *Science* **1999**, *285*, 1722.
(52) Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. *Nature* **2000**, *407*, 340.
(53) Francois, B.; Russell, R. J.; Murray, J. B.; Aboul-ela, F.; Masquida, B.; Vicens, Q.; Westhof, E. *Nucleic Acids Res.* **2005**, *33*, 5677.
(54) Shandrick, S.; Zhao, Q.; Han, Q.; Ayida, B. K.; Takahashi, M.; Winters, G. C.; Simonsen, K. B.; Vourloumis, D.; Hermann, T. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 3177.
(55) Ogle, J. M.; Brodersen, D. E.; Clemons, W. M., Jr.; Tarry, M. J.; Carter, A. P.; Ramakrishnan, V. *Science* **2001**, *292*, 897.
(56) Lin, A. H.; Murray, R. W.; Vidmar, T. J.; Marotti, K. R. *Antimicrob. Agents Chemother.* **1997**, *41*, 2127.
(57) Zhou, C. C.; Swaney, S. M.; Shinabarger, D. L.; Stockman, B. J. *Antimicrob. Agents Chemother.* **2002**, *46*, 625.
(58) Shinabarger, D. L.; Marotti, K. R.; Murray, R. W.; Lin, A. H.; Melchior, E. P.; Swaney, S. M.; Dunyak, D. S.; Demyan, W. F.; Buysse, J. M. *Antimicrob. Agents Chemother.* **1997**, *41*, 2132.
(59) Barbachyn, M. R.; Hutchinson, D. K.; Brickner, S. J.; Cynamon, M. H.; Kilburn, J. O.; Klemens, S. P.; Glickman, S. E.; Grega, K. C.; Hendges, S. K.; Toops, D. S.; Ford, C. W.; Zurenko, G. E. *J. Med. Chem.* **1996**, *39*, 680.
(60) Bobkova, E. V.; Yan, Y. P.; Jordan, D. B.; Kurilla, M. G.; Pompliano, D. L. *J. Biol. Chem.* **2003**, *278*, 9802.
(61) Jarvest, R. L.; Berge, J. M.; Berry, V.; Boyd, H. F.; Brown, M. J.; Elder, J. S.; Forrest, A. K.; Fosberry, A. P.; Gentry, D. R.; Hibbs, M. J.; Jaworski, D. D.; O'Hanlon, P. J.; Pope, A. J.; Rittenhouse, S.; Sheppard, R. J.; Slater-Radosti, C.; Worby, A. *J. Med. Chem.* **2002**, *45*, 1959.
(62) Ibba, M.; Soll, D. *Annu. Rev. Biochem.* **2000**, *69*, 617.
(63) Mikkelsen, N. E.; Johansson, K.; Virtanen, A.; Kirsebom, L. A. *Nat. Struct. Biol.* **2001**, *8*, 510.
(64) Kirk, S. R.; Tor, Y. *Bioorg. Med. Chem.* **1999**, *7*, 1979.
(65) Walter, F.; Putz, J.; Giege, R.; Westhof, E. *EMBO J. J* **2002**, *21*, 760.
(66) Kirillov, S.; Vitali, L. A.; Goldstein, B. P.; Monti, F.; Semenkov, Y.; Makhno, V.; Ripa, S.; Pon, C. L.; Gualerzi, C. O. *RNA* **1997**, *3*, 905.
(67) Henkin, T. M. *Mol. Microbiol.* **1994**, *13*, 381.
(68) Grundy, F. J.; Henkin, T. M. *Cell* **1993**, *74*, 475.
(69) Henkin, T. M.; Glass, B. L.; Grundy, F. J. *J. Bacteriol.* **1992**, *174*, 1299.
(70) Grundy, F. J.; Moir, T. R.; Haldeman, M. T.; Henkin, T. M. *Nucleic Acids Res.* **2002**, *30*, 1646.
(71) Gerdeman, M. S.; Henkin, T. M.; Hines, J. V. *Nucleic Acids Res.* **2002**, *30*, 1065.
(72) Means, J. A.; Hines, J. V. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2169.
(73) Gerdeman, M. S.; Henkin, T. M.; Hines, J. V. *J. Mol. Biol.* **2003**, *326*, 189.
(74) Means, J.; Katz, S.; Nayek, A.; Anupam, R.; Hines, J. V.; Bergmeier, S. C. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3600.
(75) Frankel, A. D.; Young, J. A. *Annu. Rev. Biochem.* **1998**, *67*, 1.
(76) Jakobovits, A.; Smith, D. H.; Jakobovits, E. B.; Capon, D. J. *Mol. Cell. Biol.* **1988**, *8*, 2555.
(77) Selby, M. J.; Bain, E. S.; Luciw, P. A.; Peterlin, B. M. *Genes Dev.* **1989**, *3*, 547.
(78) Richter, S.; Cao, H.; Rana, T. M. *Biochemistry* **2002**, *41*, 6391.
(79) Strebel, K. *AIDS* **2003**, *17* (Suppl 4), S25.
(80) Delling, U.; Roy, S.; Sumner-Smith, M.; Barnett, R.; Reid, L.; Rosen, C. A.; Sonenberg, N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6234.
(81) Roy, S.; Delling, U.; Chen, C. H.; Rosen, C. A.; Sonenberg, N. *Genes Dev.* **1990**, *4*, 1365.
(82) Weeks, K. M.; Crothers, D. M. *Cell* **1991**, *66*, 577.
(83) Weeks, K. M.; Crothers, D. M. *Biochemistry* **1992**, *31*, 10281.
(84) Aboul-ela, F.; Karn, J.; Varani, G. *J. Mol. Biol.* **1995**, *253*, 313.
(85) Churcher, M. J.; Lamont, C.; Hamy, F.; Dingwall, C.; Green, S. M.; Lowe, A. D.; Butler, J. G.; Gait, M. J.; Karn, J. *J. Mol. Biol.* **1993**, *230*, 90.
(86) Mei, H. Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. *Bioorg. Med. Chem.* **1997**, *5*, 1173.
(87) Mei, H. Y.; Cui, M.; Heldsinger, A.; Lemrow, S. M.; Loo, J. A.; Sannes-Lowery, K. A.; Sharmeen, L.; Czarnik, A. W. *Biochemistry* **1998**, *37*, 14204.
(88) Wang, S.; Huber, P. W.; Cui, M.; Czarnik, A. W.; Mei, H. Y. *Biochemistry* **1998**, *37*, 5549.
(89) Faber, C.; Sticht, H.; Schweimer, K.; Rosch, P. *J. Biol. Chem.* **2000**, *275*, 20660.
(90) Pollard, V. W.; Malim, M. H. *Annu. Rev. Microbiol.* **1998**, *52*, 491.
(91) Mann, D. A.; Mikaelian, I.; Zemmel, R. W.; Green, S. M.; Lowe, A. D.; Kimura, T.; Singh, M.; Butler, P. J.; Gait, M. J.; Karn, J. *J. Mol. Biol.* **1994**, *241*, 193.
(92) Luedtke, N. W.; Tor, Y. *Biopolymers* **2003**, *70*, 103.
(93) Tor, Y. *ChemBioChem* **2003**, *4*, 998.
(94) Zapp, M. L.; Stern, S.; Green, M. R. *Cell* **1993**, *74*, 969.
(95) Heaphy, S.; Finch, J. T.; Gait, M. J.; Karn, J.; Singh, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7366.
(96) Ippolito, J. A.; Steitz, T. A. *J. Mol. Biol.* **2000**, *295*, 711.
(97) Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R. *Science* **1996**, *273*, 1547.
(98) Kirk, S. R.; Luedtke, N. W.; Tor, Y. *J. Am. Chem. Soc.* **2000**, *122*, 980.
(99) Lacourciere, K. A.; Stivers, J. T.; Marino, J. P. *Biochemistry* **2000**, *39*, 5630.
(100) Hendrix, M.; Priestley, E. S.; Joyce, G. F.; Wong, C. H. *J. Am. Chem. Soc.* **1997**, *119*, 3641.
(101) Cho, J.; Rando, R. R. *Biochemistry* **1999**, *38*, 8548.
(102) Dworkin, J. P.; Lazcano, A.; Miller, S. L. *J. Theor. Biol.* **2003**, *222*, 127.
(103) von Ahsen, U.; Davies, J.; Schroeder, R. *Nature* **1991**, *353*, 368.
(104) Stage, T. K.; Hertel, K. J.; Uhlenbeck, O. C. *RNA* **1995**, *1*, 95.
(105) Mikkelsen, N. E.; Brannvall, M.; Virtanen, A.; Kirsebom, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6155.
(106) Rogers, J.; Chang, A. H.; von Ahsen, U.; Schroeder, R.; Davies, J. *J. Mol. Biol.* **1996**, *259*, 916.
(107) Earnshaw, D. J.; Gait, M. J. *Nucleic Acids Res.* **1998**, *26*, 5551.
(108) Racz, Z.; Hamar, P. *Curr. Med. Chem.* **2006**, *13*, 2299.
(109) Lieberman, J.; Song, E.; Lee, S. K.; Shankar, P. *Trends Mol. Med.* **2003**, *9*, 397.
(110) Winkler, W. C.; Breaker, R. R. *Annu. Rev. Microbiol.* **2005**, *59*, 487.
(111) Sudarsan, N.; Barrick, J. E.; Breaker, R. R. *RNA* **2003**, *9*, 644.
(112) Mandal, M.; Boese, B.; Barrick, J. E.; Winkler, W. C.; Breaker, R. R. *Cell* **2003**, *113*, 577.
(113) Soukup, G. A.; Breaker, R. R. *Trends Biotechnol.* **1999**, *17*, 469.
(114) Soukup, G. A.; Breaker, R. R. *Curr. Opin. Struct. Biol.* **2000**, *10*, 318.
(115) Silverman, S. K. *RNA* **2003**, *9*, 377.
(116) Breaker, R. R. *Nature* **2004**, *432*, 838.
(117) Topp, S.; Gallivan, J. P. *J. Am. Chem. Soc.* **2007**, *129*, 6807.
(118) Batey, R. T.; Gilbert, S. D.; Montange, R. K. *Nature* **2004**, *432*, 411.
(119) Serganov, A.; Yuan, Y. R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A. T.; Hobartner, C.; Micura, R.; Breaker, R. R.; Patel, D. J. *Chem. Biol.* **2004**, *11*, 1729.

(120) Thore, S.; Leibundgut, M.; Ban, N. *Science* **2006**, *312*, 1208.
(121) Serganov, A.; Polonskaia, A.; Phan, A. T.; Breaker, R. R.; Patel, D. J. *Nature* **2006**, *441*, 1167.
(122) Montange, R. K.; Batey, R. T. *Nature* **2006**, *441*, 1172.
(123) Edwards, T. E.; Ferre-D'Amare, A. R. *Structure* **2006**, *14*, 1459.
(124) Winkler, W. C.; Nahvi, A.; Sudarsan, N.; Barrick, J. E.; Breaker, R. R. *Nat. Struct. Biol.* **2003**, *10*, 701.
(125) Blount, K. F.; Breaker, R. R. *Nat. Biotechnol.* **2006**, *24*, 1558.
(126) Blount, K. F.; Wang, J. X.; Lim, J.; Sudarsan, N.; Breaker, R. R. *Nat. Chem. Biol.* **2007**, *3*, 44.
(127) Mayer, G.; Famulok, M. *ChemBioChem* **2006**, *7*, 602.
(128) Blount, K. F.; Puskarz, I.; Penchovsky, R.; Breaker, R. R. *RNA Biol.* **2006**, *3*, 77.
(129) Brantl, S. *Plasmid* **2002**, *48*, 165.
(130) Praszkier, J.; Pittard, A. J. *Plasmid* **2005**, *53*, 97.
(131) Gerdes, K.; Gultyaev, A. P.; Franch, T.; Pedersen, K.; Mikkelsen, N. D. *Annu. Rev. Genet.* **1997**, *31*, 1.
(132) Franch, T.; Gerdes, K. *Curr. Opin. Microbiol.* **2000**, *3*, 159.
(133) DeNap, J. C.; Hergenrother, P. J. *Org. Biomol. Chem.* **2005**, *3*, 959.
(134) Denap, J. C.; Thomas, J. R.; Musk, D. J.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2004**, *126*, 15402.
(135) Thomas, J. R.; DeNap, J. C.; Wong, M. L.; Hergenrother, P. J. *Biochemistry* **2005**, *44*, 6800.
(136) Praszkier, J.; Murthy, S.; Pittard, A. J. *J. Bacteriol.* **2000**, *182*, 3972.
(137) Praszkier, J.; Pittard, A. J. *J. Bacteriol.* **2002**, *184*, 5772.
(138) Moritz, E. M.; Hergenrother, P. J. In *Antimicrobial Resistance in Bacteria*; Horizon Bioscience: Norfolk, 2007; p 25.
(139) Chu, E.; Koeller, D. M.; Casey, J. L.; Drake, J. C.; Chabner, B. A.; Elwood, P. C.; Zinn, S.; Allegra, C. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 8977.
(140) Chu, E.; Voeller, D.; Koeller, D. M.; Drake, J. C.; Takimoto, C. H.; Maley, G. F.; Maley, F.; Allegra, C. J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 517.
(141) Chu, E.; Callender, M. A.; Farrell, M. P.; Schmitz, J. C. *Cancer Chemother. Pharmacol.* **2003**, *52* (Suppl 1), S80.
(142) Gmeiner, W. H. *Curr. Med. Chem.* **2005**, *12*, 191.
(143) Carreras, C. W.; Santi, D. V. *Annu. Rev. Biochem.* **1995**, *64*, 721.
(144) Chu, E.; Koeller, D. M.; Johnston, P. G.; Zinn, S.; Allegra, C. J. *Mol. Pharmacol.* **1993**, *43*, 527.
(145) Van der Wilt, C. L.; Pinedo, H. M.; Smid, K.; Peters, G. J. *Cancer Res.* **1992**, *52*, 4922.
(146) Keyomarsi, K.; Samet, J.; Molnar, G.; Pardee, A. B. *J. Biol. Chem.* **1993**, *268*, 15142.
(147) Tok, J. B.; Cho, J.; Rando, R. R. *Biochemistry* **1999**, *38*, 199.
(148) Aziz, N.; Munro, H. N. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8478.
(149) Hentze, M. W.; Caughman, S. W.; Rouault, T. A.; Barriocanal, J. G.; Dancis, A.; Harford, J. B.; Klausner, R. D. *Science* **1987**, *238*, 1570.
(150) Gray, N. K.; Hentze, M. W. *EMBO J.* **1994**, *13*, 3882.
(151) Binder, R.; Horowitz, J. A.; Basilion, J. P.; Koeller, D. M.; Klausner, R. D.; Harford, J. B. *EMBO J.* **1994**, *13*, 1969.
(152) Theil, E. C. *Biochem. J.* **1994**, *304* (Pt 1), 1.
(153) Addess, K. J.; Basilion, J. P.; Klausner, R. D.; Rouault, T. A.; Pardi, A. *J. Mol. Biol.* **1997**, *274*, 72.
(154) Jaffrey, S. R.; Haile, D. J.; Klausner, R. D.; Harford, J. B. *Nucleic Acids Res.* **1993**, *21*, 4627.
(155) Tibodeau, J. D.; Fox, P. M.; Ropp, P. A.; Theil, E. C.; Thorp, H. H. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 253.
(156) Kozak, M. *Gene* **2005**, *361*, 13.
(157) Kozak, M. *Gene* **2002**, *299*, 1.
(158) Kozak, M. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 2850.
(159) Goossen, B.; Hentze, M. W. *Mol. Cell. Biol.* **1992**, *12*, 1959.
(160) Kozak, M. *Mol. Cell. Biol.* **1989**, *9*, 5134.
(161) Wang, L.; Wessler, S. R. *Plant Physiol.* **2001**, *125*, 1380.
(162) Werstuck, G.; Green, M. R. *Science* **1998**, *282*, 296.
(163) Baird, S. D.; Turcotte, M.; Korneluk, R. G.; Holcik, M. *RNA* **2006**, *12*, 1755.
(164) Dasgupta, A.; Das, S.; Izumi, R.; Venkatesan, A.; Barat, B. *FEMS Microbiol. Lett.* **2004**, *234*, 189.
(165) Spriggs, K. A.; Bushell, M.; Mitchell, S. A.; Willis, A. E. *Cell Death Differ.* **2005**, *12*, 585.
(166) Otto, G. A.; Puglisi, J. D. *Cell* **2004**, *119*, 369.
(167) Gallego, J.; Varani, G. *Biochem. Soc. Trans.* **2002**, *30*, 140.
(168) Holcik, M. *Curr. Cancer Drug Targets* **2004**, *4*, 299.
(169) Jubin, R. *Curr. Opin. Investig. Drugs* **2003**, *4*, 162.
(170) Neyts, J. *Antiviral Res.* **2006**, *71*, 363.
(171) Manns, M. P.; McHutchison, J. G.; Gordon, S. C.; Rustgi, V. K.; Shiffman, M.; Reindollar, R.; Goodman, Z. D.; Koury, K.; Ling, M.; Albrecht, J. K. *Lancet* **2001**, *358*, 958.
(172) Fried, M. W.; Shiffman, M. L.; Reddy, K. R.; Smith, C.; Marinos, G.; Goncales, F. L., Jr.; Haussinger, D.; Diago, M.; Carosi, G.; Dhumeaux, D.; Craxi, A.; Lin, A.; Hoffman, J.; Yu, J. *N. Engl. J. Med.* **2002**, *347*, 975.
(173) Bukh, J.; Purcell, R. H.; Miller, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4942.
(174) Han, J. H.; Shyamala, V.; Richman, K. H.; Brauer, M. J.; Irvine, B.; Urdea, M. S.; Tekamp-Olson, P.; Kuo, G.; Choo, Q. L.; Houghton, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1711.
(175) Kieft, J. S.; Zhou, K.; Jubin, R.; Doudna, J. A. *RNA* **2001**, *7*, 194.
(176) Lukavsky, P. J.; Kim, I.; Otto, G. A.; Puglisi, J. D. *Nat. Struct. Biol.* **2003**, *10*, 1033.
(177) Spahn, C. M.; Kieft, J. S.; Grassucci, R. A.; Penczek, P. A.; Zhou, K.; Doudna, J. A.; Frank, J. *Science* **2001**, *291*, 1959.
(178) Kozak, M. *Nucleic Acids Res.* **2005**, *33*, 6593.
(179) Ren, R. *Nat. Rev. Cancer* **2005**, *5*, 172.
(180) Sucheck, S. J.; Greenberg, W. A.; Tolbert, T. J.; Wong, C.-H. *Angew. Chem., Int. Ed. Eng.* **2000**, *39*, 1080.
(181) Barr, F. G. *Oncogene* **2001**, *20*, 5736.
(182) Takyar, S.; Hickerson, R. P.; Noller, H. F. *Cell* **2005**, *120*, 49.
(183) Paraskeva, E.; Gray, N. K.; Schlager, B.; Wehr, K.; Hentze, M. W. *Mol. Cell. Biol.* **1999**, *19*, 807.
(184) Barton, H. A.; Eisenstein, R. S.; Bomford, A.; Munro, H. N. *J. Biol. Chem.* **1990**, *265*, 7000.
(185) Koeppen, A. H. *Cerebellum* **2005**, *4*, 62.
(186) Di Prospero, N. A.; Fischbeck, K. H. *Nat. Rev. Genet.* **2005**, *6*, 756.
(187) Gusella, J. F.; Macdonald, M. E. *Trends Biochem. Sci.* **2006**, *31*, 533.
(188) Jasinska, A.; Michlewski, G.; de Mezer, M.; Sobczak, K.; Kozlowski, P.; Napierala, M.; Krzyzosiak, W. J. *Nucleic Acids Res.* **2003**, *31*, 5463.
(189) Broda, M.; Kierzek, E.; Gdaniec, Z.; Kulinski, T.; Kierzek, R. *Biochemistry* **2005**, *44*, 10873.
(190) Erlitzki, R.; Long, J. C.; Theil, E. C. *J. Biol. Chem.* **2002**, *277*, 42579.
(191) Wickens, M.; Anderson, P.; Jackson, R. J. *Curr. Opin. Genet. Dev.* **1997**, *7*, 220.
(192) Wilkie, G. S.; Dickson, K. S.; Gray, N. K. *Trends Biochem. Sci.* **2003**, *28*, 182.
(193) Mazumder, B.; Seshadri, V.; Fox, P. L. *Trends Biochem. Sci.* **2003**, *28*, 91.
(194) Chabanon, H.; Mickleburgh, I.; Hesketh, J. *Brief Funct. Genomic Proteomic* **2004**, *3*, 240.
(195) Ross, A. F.; Oleynikov, Y.; Kislauskis, E. H.; Taneja, K. L.; Singer, R. H. *Mol. Cell. Biol.* **1997**, *17*, 2158.
(196) Garzon, R.; Fabbri, M.; Cimmino, A.; Calin, G. A.; Croce, C. M. *Trends Mol. Med.* **2006**, *12*, 580.
(197) He, L.; Hannon, G. J. *Nat. Rev. Genet.* **2004**, *5*, 522.
(198) Rajewsky, N.; Socci, N. D. *Dev. Biol.* **2004**, *267*, 529.
(199) Cheng, A. M.; Byrom, M. W.; Shelton, J.; Ford, L. P. *Nucleic Acids Res.* **2005**, *33*, 1290.
(200) Karp, X.; Ambros, V. *Science* **2005**, *310*, 1288.
(201) Chen, C. Z.; Li, L.; Lodish, H. F.; Bartel, D. P. *Science* **2004**, *303*, 83.
(202) Xu, P.; Guo, M.; Hay, B. A. *Trends Genet.* **2004**, *20*, 617.
(203) Verhelst, S. H.; Michiels, P. J.; van der Marel, G. A.; van Boeckel, C. A.; van Boom, J. H. *ChemBioChem* **2004**, *5*, 937.
(204) Wong, C. H.; Hendrix, M.; Priestley, E. S.; Greenberg, W. A. *Chem. Biol.* **1998**, *5*, 397.
(205) Wang, H.; Tor, Y. *J. Am. Chem. Soc.* **1997**, *119*, 8734.
(206) Kaul, M.; Barbieri, C. M.; Kerrigan, J. E.; Pilch, D. S. *J. Mol. Biol.* **2003**, *326*, 1373.
(207) Kaul, M.; Barbieri, C. M.; Pilch, D. S. *J. Mol. Biol.* **2005**, *346*, 119.
(208) Pilch, D. S.; Kaul, M.; Barbieri, C. M.; Kerrigan, J. E. *Biopolymers* **2003**, *70*, 58.
(209) Luedtke, N. W.; Baker, T. J.; Goodman, M.; Tor, Y. *J. Am. Chem. Soc.* **2000**, *122*, 12035.
(210) Kaul, M.; Pilch, D. S. *Biochemistry* **2002**, *41*, 7695.
(211) Yoshizawa, S.; Fourmy, D.; Puglisi, J. D. *EMBO J.* **1998**, *17*, 6437.
(212) Han, Q.; Zhao, Q.; Fish, S.; Simonsen, K. B.; Vourloumis, D.; Froelich, J. M.; Wall, D.; Hermann, T. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 2694.
(213) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.
(214) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins* **1986**, *1*, 47.
(215) Vicens, Q.; Westhof, E. *Biopolymers* **2003**, *70*, 42.
(216) Vicens, Q.; Westhof, E. *Structure* **2001**, *9*, 647.
(217) Hermann, T.; Westhof, E. *J. Mol. Biol.* **1998**, *276*, 903.
(218) Hermann, T. *Biochimie* **2002**, *84*, 869.
(219) Blount, K. F.; Zhao, F.; Hermann, T.; Tor, Y. *J. Am. Chem. Soc.* **2005**, *127*, 9818.
(220) Yan, Z.; Baranger, A. M. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5889.
(221) Bradrick, T. D.; Marino, J. P. *RNA* **2004**, *10*, 1459.
(222) Kirk, S. R.; Luedtke, N. W.; Tor, Y. *Bioorg. Med. Chem.* **2001**, *9*, 2295.
(223) Llano-Sotelo, B.; Chow, C. S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 213.

(224) Haddad, J.; Kotra, L. P.; Llano-Sotelo, B.; Kim, C.; Azucena, E. F., Jr.; Liu, M.; Vakulenko, S. B.; Chow, C. S.; Mobashery, S. *J. Am. Chem. Soc.* **2002**, *124*, 3229.
(225) Liu, X.; Thomas, J. R.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2004**, *126*, 9196.
(226) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2005**, *127*, 12434.
(227) Tok, J. B.; Bi, L.; Saenz, M. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 827.
(228) Wong, C. H.; Liang, F. S. *Methods Enzymol.* **2003**, *362*, 340.
(229) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *Biochemistry* **2006**, *45*, 10928.
(230) Blount, K. F.; Tor, Y. *Nucleic Acids Res.* **2003**, *31*, 5490.
(231) Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. *ChemBioChem* **2005**, *6*, 1492.
(232) Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Q. Rev. Biophys.* **1999**, *32*, 211.
(233) Johnson, E. C.; Feher, V. A.; Peng, J. W.; Moore, J. M.; Williamson, J. R. *J. Am. Chem. Soc.* **2003**, *125*, 15724.
(234) Yu, L.; Oost, T. K.; Schkeryantz, J. M.; Yang, J.; Janowick, D.; Fesik, S. W. *J. Am. Chem. Soc.* **2003**, *125*, 4444.
(235) Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. *J. Biomol. NMR* **2001**, *21*, 349.
(236) Qin, P. Z.; Dieckmann, T. *Curr. Opin. Struct. Biol.* **2004**, *14*, 350.
(237) Zidek, L.; Stefl, R.; Sklenar, V. *Curr. Opin. Struct. Biol.* **2001**, *11*, 275.
(238) Mayer, M.; James, T. L. *Methods Enzymol.* **2005**, *394*, 571.
(239) Kreutz, C.; Kahlig, H.; Konrat, R.; Micura, R. *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 3450.
(240) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377.
(241) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264.
(242) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10129.
(243) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discovery* **2006**, *5*, 585.
(244) Hofstadler, S. A.; Sannes-Lowery, K. A.; Crooke, S. T.; Ecker, D. J.; Sasmor, H.; Manalili, S.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436.
(245) Hamasaki, K.; Rando, R. R. *Anal. Biochem.* **1998**, *261*, 183.
(246) Wang, Y.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1997**, *36*, 768.
(247) Matsumoto, C.; Hamasaki, K.; Mihara, H.; Ueno, A. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1857.
(248) Hwang, S.; Tamilarasu, N.; Kibler, K.; Cao, H.; Ali, A.; Ping, Y. H.; Jeang, K. T.; Rana, T. M. *J. Biol. Chem.* **2003**, *278*, 39092.
(249) MacBeath, G.; Koehler, A. N.; Schreiber, S. L. *J. Am. Chem. Soc.* **1999**, *121*, 7967.
(250) Duffner, J. L.; Clemons, P. A.; Koehler, A. N. *Curr. Opin. Chem. Biol.* **2007**, *11*, 74.
(251) Hergenrother, P. J.; Depew, K. M.; Schreiber, S. L. *J. Am. Chem. Soc.* **2000**, *122*, 7849.
(252) Kuruvilla, F. G.; Sternson, S. M.; Shamji, A. F.; Hergenrother, P. J.; Schreiber, S. L. *Nature* **2002**, *416*, 653.
(253) Disney, M. D.; Seeberger, P. H. *Chem. Eur. J.* **2004**, *10*, 3308.
(254) Bryan, M. C.; Wong, C.-H. *Tetrahedron Lett.* **2004**, *45*, 3639.
(255) Zhou, Y.; Gregor, V. E.; Sun, Z.; Ayida, B. K.; Winters, G. C.; Murphy, D.; Simonsen, K. B.; Vourloumis, D.; Fish, S.; Froelich, J. M.; Wall, D.; Hermann, T. *Antimicrob. Agents Chemother.* **2005**, *49*, 4942.
(256) Yan, Z.; Ramisetty, S. R.; Bolton, P. H.; Baranger, A. M. *ChemBioChem* **2007**, *8*, 4096.
(257) McPike, M. P.; Goodisman, J.; Dabrowiak, J. C. *Methods Enzymol.* **2001**, *340*, 431.
(258) Nahvi, A.; Sudarsan, N.; Ebert, M. S.; Zou, X.; Brown, K. L.; Breaker, R. R. *Chem. Biol.* **2002**, *9*, 1043.
(259) Soukup, G. A.; Breaker, R. R. *RNA* **1999**, *5*, 1308.
(260) Soukup, G. A.; DeRose, E. C.; Koizumi, M.; Breaker, R. R. *RNA* **2001**, *7*, 524.
(261) Winkler, W.; Nahvi, A.; Breaker, R. R. *Nature* **2002**, *419*, 952.
(262) Winkler, W. C.; Cohen-Chalamish, S.; Breaker, R. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15908.
(263) Luedtke, N. W.; Liu, Q.; Tor, Y. *Biochemistry* **2003**, *42*, 11391.
(264) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2005**, *127*, 12434.
(265) Childs-Disney, J. L.; Wu, M.; Pushechnikov, A.; Aminova, O.; Disney, M. D. *ACS Chem. Biol.* **2007**, *2*, 745.
(266) Liang, F. S.; Greenberg, W. A.; Hammond, J. A.; Hoffmann, J.; Head, S. R.; Wong, C. H. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12311.
(267) Fierro-Monti, I.; Mathews, M. B. *Trends Biochem. Sci.* **2000**, *25*, 241.
(268) Nanduri, S.; Carpick, B. W.; Yang, Y.; Williams, B. R.; Qin, J. *EMBO J.* **1998**, *17*, 5458.
(269) Kharrat, A.; Macias, M. J.; Gibson, T. J.; Nilges, M.; Pastore, A. *EMBO J.* **1995**, *14*, 3572.
(270) Bycroft, M.; Grunert, S.; Murzin, A. G.; Proctor, M.; St Johnston, D. *EMBO J.* **1995**, *14*, 3563.
(271) Carlson, C. B.; Spanggord, R. J.; Beal, P. A. *ChemBioChem* **2002**, *3*, 859.
(272) Katze, M. G.; Wambach, M.; Wong, M. L.; Garfinkel, M.; Meurs, E.; Chong, K.; Williams, B. R.; Hovanessian, A. G.; Barber, G. N. *Mol. Cell. Biol.* **1991**, *11*, 5497.
(273) de Haro, C.; Mendez, R.; Santoyo, J. *FASEB J.* **1996**, *10*, 1378.
(274) Clemens, M. J.; Bommer, U. A. *Int. J. Biochem. Cell. Biol.* **1999**, *31*, 1.
(275) Wilson, W. D.; Ratmeyer, L.; Zhao, M.; Strekowski, L.; Boykin, D. *Biochemistry* **1993**, *32*, 4098.
(276) Zhao, M.; Janda, L.; Nguyen, J.; Strekowski, L.; Wilson, W. D. *Biopolymers* **1994**, *34*, 61.
(277) McConnaughie, A. W.; Spychala, J.; Zhao, M.; Boykin, D.; Wilson, W. D. *J. Med. Chem.* **1994**, *37*, 1063.
(278) Zimmer, C.; Wahnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31.
(279) Zhao, M.; Ratmeyer, L.; Peloquin, R. G.; Yao, S.; Kumar, A.; Spychala, J.; Boykin, D. W.; Wilson, W. D. *Bioorg. Med. Chem.* **1995**, *3*, 785.
(280) McNaughton, B. R.; Gareiss, P. C.; Miller, B. L. *J. Am. Chem. Soc.* **2007**, *129*, 11306.
(281) Nonin, S.; Jiang, F.; Patel, D. J. *J. Mol. Biol.* **1997**, *268*, 359.
(282) Carlson, C. B.; Beal, P. A. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1979.
(283) Todd, A. K.; Adams, A.; Thorpe, J. H.; Denny, W. A.; Wakelin, L. P.; Cardin, C. J. *J. Med. Chem.* **1999**, *42*, 536.
(284) Adams, A.; Guss, J. M.; Collyer, C. A.; Denny, W. A.; Wakelin, L. P. *Biochemistry* **1999**, *38*, 9221.
(285) Malinina, L.; Soler-Lopez, M.; Aymami, J.; Subirana, J. A. *Biochemistry* **2002**, *41*, 9341.
(286) Carlson C. B., Vuyisich M., Gooch B. D., Beal P. A. *Chem. Biol.* **2003**, *10*, 663.
(287) Adams, A. *Curr. Med. Chem.* **2002**, *9*, 1667.
(288) Gooch, B. D., Beal, P. A. *J. Am. Chem. Soc.* **2004**, *126*, 10603.
(289) Gooch, B. D.; Krishnamurthy, M.; Shadid, M.; Beal, P. A. *ChemBioChem* **2005**, *6*, 2247.
(290) Krishnamurthy, M.; Simon, K.; Orendt, A. M.; Beal, P. A. *Angew. Chem., Int. Ed. Engl.* **2007**, *46*, 7044.
(291) Schroeder, S. J.; Burkard, M. E.; Turner, D. H. *Biopolymers* **1999**, *52*, 157.
(292) Chen, G.; Znosko, B. M.; Jiao, X.; Turner, D. H. *Biochemistry* **2004**, *43*, 12865.
(293) Kierzek, R.; Burkard, M. E.; Turner, D. H. *Biochemistry* **1999**, *38*, 14214.
(294) Begg, E. J.; Barclay, M. L. *Br. J. Clin. Pharmacol.* **1995**, *39*, 597.
(295) Azucena, E.; Mobashery, S. *Drug Resis. Update* **2001**, *4*, 106.
(296) Sucheck, S. J. W., A. L.; Koeller, K. M.; Boeher, D. D.; Draker, K-A.; Sears, P.; Wright, G. D.; Wong, C.-H. *J. Am. Chem. Soc.* **2000**, *122*, 5230.
(297) Wu, B.; Yang, J.; Robinson, D.; Hofstadler, S.; Griffey, R.; Swayze, E. E.; He, Y. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3915.
(298) Kotra, L. P.; Haddad, J.; Mobashery, S. *Antimicrob. Agents Chemother.* **2000**, *44*, 3249.
(299) Russell, R. J.; Murray, J. B.; Lentzen, G.; Haddad, J.; Mobashery, S. *J. Am. Chem. Soc.* **2003**, *125*, 3410.
(300) Kondo, J.; Francois, B.; Russell, R. J.; Murray, J. B.; Westhof, E. *Biochimie* **2006**, *88*, 1027.
(301) Venot, A.; Swayze, E. E.; Griffey, R. H.; Boons, G. J. *ChemBioChem* **2004**, *5*, 1228.
(302) Rao, Y.; Venot, A.; Swayze, E. E.; Griffey, R. H.; Boons, G. J. *Org. Biomol. Chem.* **2006**, *4*, 1328.
(303) Francois, B.; Szychowski, J.; Adhikari, S. S.; Pachamuthu, K.; Swayze, E. E.; Griffey, R. H.; Migawa, M. T.; Westhof, E.; Hanessian, S. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 6735.
(304) Ding, Y.; Hofstadler, S. A.; Swayze, E. E.; Griffey, R. H. *Org. Lett.* **2001**, *3*, 1621.
(305) Ding, Y.; Hofstadler, S. A.; Swayze, E. E.; Risen, L.; Griffey, R. H. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 3409.
(306) Greenberg, W. A.; Priestley, E. S.; Sears, P.; Alper, P. B.; Rosenbohm, C.; Hendrix, M.; Hung, S. C.; Wong, C. H. *J. Am. Chem. Soc.* **1999**, *121*, 6527.
(307) Hanessian, S.; Tremblay, M.; Kornienko, A.; Moitessier, N. *Tetrahedron* **2001**, *57*, 3221.
(308) Wang, X.; Migawa, M. T.; Sannes-Lowery, K. A.; Swayze, E. E. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4919.
(309) Vourloumis, D.; Winters, G. C.; Takahashi, M.; Simonsen, K. B.; Ayida, B. K.; Shandrick, S.; Zhao, Q.; Hermann, T. *ChemBioChem* **2003**, *4*, 879.
(310) Simonsen, K. B.; Ayida, B. K.; Vourloumis, D.; Winters, G. C.; Takahashi, M.; Shandrick, S.; Zhao, Q.; Hermann, T. *ChemBioChem* **2003**, *4*, 886.

(311) Barluenga, S.; Simonsen, K. B.; Littlefield, E. S.; Ayida, B. K.; Vourloumis, D.; Winters, G. C.; Takahashi, M.; Shandrick, S.; Zhao, Q.; Han, Q.; Hermann, T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 713.
(312) Swayze, E. E.; Jefferson, E. A.; Sannes-Lowery, K. A.; Blyn, L. B.; Risen, L. M.; Arakawa, S.; Osgood, S. A.; Hofstadler, S. A.; Griffey, R. H. *J. Med. Chem.* **2002**, *45*, 3816.
(313) He, Y.; Yang, J.; Wu, B.; Robinson, D.; Sprankle, K.; Kung, P. P.; Lowery, K.; Mohan, V.; Hofstadler, S.; Swayze, E. E.; Griffey, R. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 695.
(314) Maddaford, S. P.; Motamed, M.; Turner, K. B.; Choi, M. S.; Ramnauth, J.; Rakhit, S.; Hudgins, R. R.; Fabris, D.; Johnson, P. E. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5987.
(315) Foloppe, N.; Chen, I. J.; Davis, B.; Hold, A.; Morley, D.; Howes, R. *Bioorg. Med. Chem.* **2004**, *12*, 935.
(316) Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W.; Wilson, W. D.; Green, M. R. *Bioorg. Med. Chem.* **1997**, *5*, 1149.
(317) Xiao, G.; Kumar, A.; Li, K.; Rigl, C. T.; Bajic, M.; Davis, T. M.; Boykin, D. W.; Wilson, W. D. *Bioorg. Med. Chem.* **2001**, *9*, 1097.
(318) Li, K.; Davis, T. M.; Bailly, C.; Kumar, A.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **2001**, *40*, 1150.
(319) Tor, Y.; Hermann, T.; Westhof, E. *Chem. Biol.* **1998**, *5*, R277.
(320) Luedtke, N. W.; Carmichael, P.; Tor, Y. *J. Am. Chem. Soc.* **2003**, *125*, 12374.
(321) DeJong, E. S.; Chang, C. E.; Gilson, M. K.; Marino, J. P. *Biochemistry* **2003**, *42*, 8035.
(322) Nakatani, K.; Sando, S.; Saito, I. *J. Am. Chem. Soc.* **2000**, *122*, 2172.
(323) Hagihara, S.; Kumasawa, H.; Goto, Y.; Hayashi, G.; Kobori, A.; Saito, I.; Nakatani, K. *Nucleic Acids Res.* **2004**, *32*, 278.
(324) Nakatani, K.; Sando, S.; Kumasawa, H.; Kikuchi, J.; Saito, I. *J. Am. Chem. Soc.* **2001**, *123*, 12650.
(325) Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C. *Nat. Chem. Biol.* **2005**, *1*, 39.
(326) Nakatani, K.; Horie, S.; Goto, Y.; Kobori, A.; Hagihara, S. *Bioorg. Med. Chem.* **2006**, *14*, 5384.
(327) Cho, J.; Rando, R. R. *Nucleic Acids Res.* **2000**, *28*, 2158.
(328) Loontiens, F. G.; Regenfuss, P.; Zechel, A.; Dumortier, L.; Clegg, R. M. *Biochemistry* **1990**, *29*, 9029.
(329) Znosko, B. M.; Silvestri, S. B.; Volkman, H.; Boswell, B.; Serra, M. J. *Biochemistry* **2002**, *41*, 10406.
(330) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911.
(331) Ratmeyer, L. S.; Vinayak, R.; Zon, G.; Wilson, W. D. *J. Med. Chem.* **1992**, *35*, 966.
(332) Bailly, C.; Colson, P.; Houssier, C.; Hamy, F. *Nucleic Acids Res.* **1996**, *24*, 1460.
(333) Dassonneville, L.; Hamy, F.; Colson, P.; Houssier, C.; Bailly, C. *Nucleic Acids Res.* **1997**, *25*, 4487.
(334) Hamy, F.; Brondani, V.; Florsheimer, A.; Stark, W.; Blommers, M. J.; Klimkait, T. *Biochemistry* **1998**, *37*, 5086.
(335) Gelus, N.; Hamy, F.; Bailly, C. *Bioorg. Med. Chem.* **1999**, *7*, 1075.
(336) Peytou, V.; Condom, R.; Patino, N.; Guedj, R.; Aubertin, A. M.; Gelus, N.; Bailly, C.; Terreux, R.; Cabrol-Bass, D. *J. Med. Chem.* **1999**, *42*, 4042.
(337) Lind, K. E.; Du, Z.; Fujinaga, K.; Peterlin, B. M.; James, T. L. *Chem. Biol.* **2002**, *9*, 185.
(338) Mayer, M.; James, T. L. *J. Am. Chem. Soc.* **2004**, *126*, 4453.
(339) Du, Z.; Lind, K. E.; James, T. L. *Chem. Biol.* **2002**, *9*, 707.
(340) Brodsky, A. S.; Williamson, J. R. *J. Mol. Biol.* **1997**, *267*, 624.
(341) Pitt, S. W.; Zhang, Q.; Patel, D. J.; Al-Hashimi, H. M. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 3412.
(342) Davis, B.; Afshar, M.; Varani, G.; Murchie, A. I.; Karn, J.; Lentzen, G.; Drysdale, M.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T.; Aboul-ela, F. *J. Mol. Biol.* **2004**, *336*, 343.
(343) Murchie, A. I.; Davis, B.; Isel, C.; Afshar, M.; Drysdale, M. J.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T. M.; Mirza, S.; Prescott, C. D.; Vaglio, P.; Aboul-ela, F.; Karn, J. *J. Mol. Biol.* **2004**, *336*, 625.
(344) Stassinopoulos, A.; Ji, J.; Gao, X.; Goldberg, I. H. *Science* **1996**, *272*, 1943.
(345) Gao, X.; Stassinopoulos, A.; Ji, J.; Kwon, Y.; Bare, S.; Goldberg, I. H. *Biochemistry* **2002**, *41*, 5131.
(346) Xi, Z.; Hwang, G. S.; Goldberg, I. H.; Harris, J. L.; Pennington, W. T.; Fouad, F. S.; Qabaja, G.; Wright, J. M.; Jones, G. B. *Chem. Biol.* **2002**, *9*, 925.
(347) Xiao, Z.; Kappen, L. S.; Goldberg, I. H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2895.
(348) Kappen, L. S.; Lin, Y.; Jones, G. B.; Goldberg, I. H. *Biochemistry* **2007**, *46*, 561.
(349) Lin, Y.; Jones, G. B.; Hwang, G. S.; Kappen, L. S.; Goldberg, I. H. *Org. Lett.* **2005**, *7*, 71.
(350) Hwang, G. S.; Jones, G. B.; Goldberg, I. H. *Biochemistry* **2004**, *43*, 641.
(351) Hwang, G. S.; Jones, G. B.; Goldberg, I. H. *Biochemistry* **2003**, *42*, 8472.
(352) Xiao, Z.; Zhang, N.; Lin, Y.; Jones, G. B.; Goldberg, I. H. *Chem. Commun.* **2006**, 4431.
(353) Cecchetti, V.; Parolin, C.; Moro, S.; Pecere, T.; Filipponi, E.; Calistri, A.; Tabarrini, O.; Gatto, B.; Palumbo, M.; Fravolini, A.; Palu, G. *J. Med. Chem.* **2000**, *43*, 3799.
(354) Parolin, C.; Gatto, B.; Del Vecchio, C.; Pecere, T.; Tramontano, E.; Cecchetti, V.; Fravolini, A.; Masiero, S.; Palumbo, M.; Palu, G. *Antimicrob. Agents Chemother.* **2003**, *47*, 889.
(355) Richter, S.; Parolin, C.; Gatto, B.; Del Vecchio, C.; Brocca-Cofano, E.; Fravolini, A.; Palu, G.; Palumbo, M. *Antimicrob. Agents Chemother.* **2004**, *48*, 1895.
(356) Varani, G.; Cheong, C.; Tinoco, I., Jr. *Biochemistry* **1991**, *30*, 3280.
(357) Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379.
(358) Abramovitz, D. L.; Pyle, A. M. *J. Mol. Biol.* **1997**, *266*, 493.
(359) Proctor, D. J.; Schaak, J. E.; Bevilacqua, J. M.; Falzone, C. J.; Bevilacqua, P. C. *Biochemistry* **2002**, *41*, 12062.
(360) Gayle, A. Y.; Baranger, A. M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2839.
(361) Scott, W. G.; Finch, J. T.; Klug, A. *Cell* **1995**, *81*, 991.
(362) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678.
(363) Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4899.
(364) Yan, Z.; Sikri, S.; Beveridge, D. L.; Baranger, A. M. *J. Med. Chem.* **2007**, *50*, 4096.
(365) Yoshizawa, S.; Fourmy, D.; Eason, R. G.; Puglisi, J. D. *Biochemistry* **2002**, *41*, 6263.

CR0681546