A Library Approach to the Discovery of Small Molecules That Recognize RNA: Use of a 1,3-Hydroxyamine Motif as Core

Chi-Huey Wong,* Martin Hendrix, David D. Manning, Christoph Rosenbohm, and William A. Greenberg

Contribution from the Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received March 12, 1998

Abstract: A library of compounds based upon an aminoglucopyranoside core has been developed and screened for binding to RNA and specifically to 16S ribosomal RNA. The title molecules simplify the complexity of naturally occurring aminoglycoside antibiotics by embodying a putative recognition motif found within these structures, namely, a 1,3-hydroxyamine. The core pyranoside bearing the hydroxyamine motif was structurally varied at two points through a combinatorial approach utilizing acylation and reductive amination protocols. The aminoglycoside mimetics were screened in an automated assay based upon surface plasmon resonance (SPR), and some were found effective at binding a 27-nucleotide model (AS-wt) of A-site 16S RNA as well as a drug-resistant mutant RNA in the micromolar range.

Introduction

A wealth of information has been discovered in the past 15 years which has fundamentally changed our perspective of the biological role of RNA. It is now clear that RNA plays a central part in many processes of life. In addition to its well-known role as an intermediary in the flow of genetic information, it is crucial to many catalytic processes.[1] It has the ability to cleave itself during the process of splicing or it can act catalytically to cleave other RNA sequences.[2,3] In concert with proteins, it forms the catalytic machinery of the ribosome.[4]

This increased awareness of the central role of RNA has led to the realization that RNA is a potential drug target that has thus far remained largely unexplored.[5] There is a host of natural products targeting the ribosome, and RNA binding seems to be central to their function.[6] Other potential RNA targets include sites of RNA−protein interactions governing transcriptional activation or repression, splicing, nuclear export, or the packaging of RNA virus particles.[7]

Despite the number of RNA targets available for drug intervention, there remains a poor understanding of the principles governing small molecule RNA recognition which hinders the development of new, rationally designed RNA binders. Furthermore, no structural information exists for many possible RNA targets, although RNA structural biology is a rapidly emerging field. In this situation, combinatorial chemistry may be an invaluable tool to discover new lead structures for specific recognition of RNA sequences.[8]

Among the natural products that recognize RNA specifically, the class of aminoglycoside antibiotics stands out. These basic polyamino saccharides have been known for over 50 years as potent antibacterial drugs.[9] Their mechanism of action involves binding to the ribosomal RNA of bacteria, thereby interfering with protein biosynthesis. For one subgroup of aminoglycosides which encompasses the widely used drugs gentamicin, kanamycin, and neomycin B, the target site has been localized to the ribosomal decoding region.[10] Recently, the solution structure of paromomycin bound to a model of this site has been solved and provided the first structural characterization of a specific aminoglycoside−RNA recognition event.[11]

(1) Eckstein, F., Lilley, David M. J., Eds. Catalytic RNA; Vol. 10: Nucleic Acids and Molecular Biology, Springer-Verlag: Berlin, 1996.

(2) (a) von Ahsen, U.; Davies, J.; Schroeder, R. *Nature* **1991**, *353*, 368. (b) von Ahsen, U.; Davies, J.; Schroeder, R. *J. Mol. Biol.* **1992**, *226*, 935. (c) von Ahsen, U.; Noller, H. F. *Science* **1993**, *260*, 1500. (d) Rogers, J.; Davies, J. *Nucleic Acids Res.* **1994**, *22*, 4983.

(3) (a) Stage, T. K.; Hertel, K. J.; Uhlenbeck, O. C. *RNA*, **1995**, *1*, 95. (b) Clouet-d'Orval, B.; Stage, T. K.; Uhlenbeck, O. C. *Biochemistry* **1995**, *34*, 11186. (c) Wang, H.; Tor, Y. *J. Am. Chem. Soc.* **1997**, *119*, 8734. (d) Wang, H.; Tor, Y. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1951.

(4) For a review, see: (a) Green, R.; Noller, H. F. *Chem. Biol.* **1997**, *66*, 679−716. (b) Noller, H. F. *Annu. Rev. Biochem.* **1991**, *60*, 191−227.

(5) Pearson, N. D.; Prescott, C. D. *Chem. Biol.* **1997**, *4*, 409−414.

(6) Reviews: (a) Cundliffe, E. In *The Ribosome: Structure, Function and Evolution*; Hill, W. E., et al., Eds.; American Society for Microbiology: Washington, DC, 1990; pp 479−90. (b) Gale, E. F.; Cundliffe, E. In *The Molecular Basis of Antibiotic Action*; Reynolds, P. E., Richmond, M. H., Waring, M. J., Eds.; John Wiley and Sons: London, 1981. (c) Cundliffe, E. In *Ribosomes: Structure, Function and Genetics*; Chambliss, G. R., et al., Eds.; University Park Press: Baltimore, 1980, pp 555ff.

(7) (a) Zapp, M. L.; Green, M. R. *Nature* **1989**, *342*, 714−716. (b) Daly, T. J.; Cook, K. S.; Gray, G. S.; Maione, T. E.; Rusche, J. R. *Nature* **1989**, *342*, 816−819. (c) Malim, M. H.; Hauber, J.; Le, S.-Y.; Maizel, J. V.; Cullen, B. R. *Nature* **1989**, *338*, 254−257. (d) Malim, M. H.; Cullen, B. R. *Mol. Cell. Biol.* **1993**, *13*, 6180−6189. (e) Fischer, U.; Meyer, S.; Teufel, M.; Heckel, C.; Lührmann, R.; Rautmann, G. *EMBO J.* **1994**, *13*, 4105−4112. (f) Wen, W.; Meinkoth, J. L.; Tsien, R. Y.; Taylor, S. S. *Cell* **1995**, *82*, 463. (g) Fischer, U.; Huber, J.; Boelens, W. C.; Mattaj, I. W.; Lührmann, R. *Cell* **1995**, *82*, 475. (h) Bogerd, H. P.; Fridell, R. A.; Madore, S.; Cullen, B. R. *Cell* **1995**, *82*, 485. (i) Stutz, F.; Neville, M.; Rosbach, M. *Cell* **1995**, *82*, 495.

(8) For some recent combinatorial strategies designed to target RNA, see: (a) Park, W. K. C.; Auer, M.; Jaksche, H.; Wong, C.-H. *J. Am. Chem. Soc.* **1996**, *118*, 10150−10155. (b) Wuonola, M. A.; Powers, D. G. *Use of combinatorial libraries in the discovery and development of novel anti-infectives*; Chaiken, I. M., Janda, K. D., Eds.; Diversity Comb. Chem., Libr. Drug Discovery, Conf.; American Chemical Society: Washington, DC, 1996; pp 284−297. (c) Cload, S. T.; Schepartz, A. *J. Am. Chem. Soc.* **1994**, *116*, 437. (d) Harada, K.; Martin, S. S.; Frankel, A. D. *Nature* **1996**, *380*, 175.

(9) Umezawa, H.; Hooper, I. R., Eds. *Aminoglycoside Antibiotics*; Springer-Verlag: New York, Heidelberg, 1982.

(10) Moazed, D.; Noller, H. F. *Nature* **1987**, *327*, 389−394.

(11) Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. *Science* **1996**, *274*, 1367−1371.

S0002-7863(98)00826-9 CCC: $15.00
Published on Web 08/06/1998

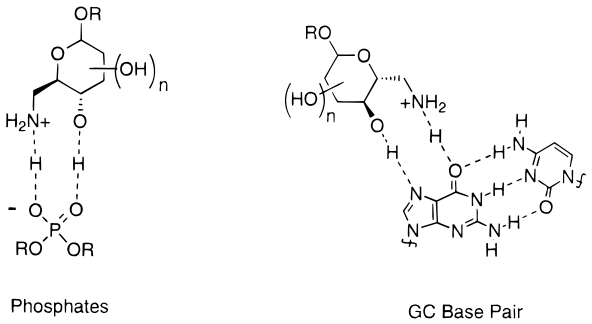


Figure 1. Hydroxyamines: a motif for RNA recognition. Phosphodiesters and the Hoogsteen face of guanosine represent bidentate RNA hydrogen bond acceptors that may be recognized by hydroxyamines.

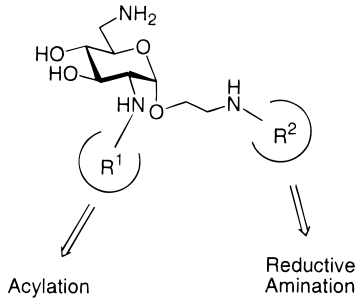


Figure 2. Library design.

We have used the structure of aminoglycosides as inspiration in our search for general motifs that may be suitable for RNA recognition.[12] In this regard, we have identified 1,2- and 1,3-hydroxyamines as an interesting binding motif.[13] Hydroxyamines are bidentate hydrogen bond donors that can interact with bidentate hydrogen bond acceptors found within RNA structure. These include backbone phosphodiesters and the Hoogsteen face of guanosine. Using model systems we have shown that, in particular, 1,3-hydroxyamines with a *gluco* configuration are strong binders for phosphodiesters (Figure 1).

Constructing combinatorial libraries incorporating the hydroxyamine motif may therefore lead to the discovery of compounds that bind to RNA. We designed an approach that would take advantage of the hydroxyamine motif as a central core and would utilize straightforward coupling chemistry to attach readily available building blocks. As shown in Figure 2, the *gluco*-configured 1,3-hydroxyamine is the core structure of our library. Two points of variation are introduced, at the anomeric position and at position 2 of the glucose ring through reductive amination and acylation, respectively.

Although this approach is general, the particular design chosen here includes a structural bias for recognition of the decoding region A-site 16S ribosomal RNA (16S rRNA). Based on the available NMR structure[11] of the aminoglycoside paromomycin complexed with a model of the *E. coli* decoding region A-site RNA, the core of the library may be aligned with the 2-deoxy-2-aminoglucopyranoside ring of paromomycin. Assuming this possible alignment, molecular modeling suggested the α-anomeric substituent would be poised to make hydrogen bonding contacts to acceptors on the bases in the major groove of the RNA, especially G1494 and U1495, bases known to be critical for natural aminoglycoside recognition (Figure 3).[14] The acyl substituent in the 2-position could make contacts with the phosphate backbone of the conserved internal loop.

Results and Discussion

To construct the library, we designed a suitably protected building block (**5**, Scheme 1) which would incorporate a free amine for the introduction of an acyl residue and a masked aldehyde for a subsequent reductive amination. The latter is provided in the form of an allyl group which can be cleaved by ozonolysis to reveal the aldehyde.

A convenient starting material for the synthesis of 2-amino sugars is *N*-acetylglucosamine (**1**). This was converted to the allyl glycoside by Fischer glycosidation.[15] An anomeric mixture (7:1 = α:*β*) is obtained, from which pure **2** can be isolated by peracetylation, chromatographic separation, and subsequent deacetylation. Selective tosylation of the primary alcohol and displacement with sodium azide led to the 6-azido glucoside **4**. The *N*-acetyl group of **4** was cleaved under basic conditions by refluxing with aqueous barium hydroxide to provide the desired building block **5**. For the large scale synthesis of this building block, formation of the allyl glycoside, tosylation, and azide displacement were carried out without the purification of any intermediates. At the stage of azide **4**, chromatography then afforded a clean, anomerically pure compound.

Several approaches exist for the construction of libraries. A strategic choice must be made between the synthesis of individual compounds or compound mixtures. We expected that specific recognition of RNA would be difficult to achieve and were aware that even the selectivity of the naturally occurring aminoglycosides for their decoding region target site is only 5- to 20-fold.[14b,16] This kind of moderate specificity would likely be obscured by overwhelming nonspecific binding when screening mixtures of compounds. We therefore decided to synthesize all library members individually using a parallel solution phase approach.

Scheme 2 shows the series of four combinatorial steps leading from the key intermediate **5** to the final deprotected library compounds. For the initial acylation, four Z-protected amino acid *N*-hydroxysuccinimide esters were used to provide amino acyl derivatives **6**−**9**. These compounds were individually purified and then subjected to the remaining three combinatorial steps. Treatment with ozone and subsequent quenching with dimethyl sulfide led to the clean production of aldehydes **10**−**13** which were reacted with any of six different amines (**a**−**f**). These included four amino acid carboxamides (**a**−**d**) as well as benzylamine (**e**) and mono-Z-protected ethylenediamine (**f**). The reductive amination products **14**−**17a**−**f** were then deprotected by hydrogenolysis in aqueous acetic acid using palladium hydroxide on carbon as a catalyst.

The three-step sequence was carried out without purification of any intermediates. Instead, the final compounds were purified

(12) For other approaches to RNA binders, see: (a) Wilson, W. D.; Ratmeyer, L.; Zhao, M.; Strekowski, L.; Boykin, D. *Biochemistry* **1993**, *32*, 4098−4104. (b) McConnaughie, A. W.; Spychala, J.; Zhao, M.; Boykin, D.; Wilson, W. D. *J. Med. Chem.* **1994**, *37*, 1063−1069. (c) Zhao, M.; Ratmeyer, L.; Peloquin, R. G.; Yao, S.; Kumar, A.; Spychala, J.; Boykin, D. W.; Wilson, W. D. *Bioorg. Med. Chem.* **1995**, *3*, 785−794. (d) Fernandez-Saiz, M.; Schneider, H.-J.; Sartorius, J.; Wilson, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 4739−4745. (e) Mei, H.-Y.; Cui, M.; Lemrow, S. M.; Czarnik, A. W. *Bioorg. Med. Chem.* **1997**, *5*, 1185. (f) Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R. Boytein, D. W.; Wilson, W. D.; Green, M. R. *Bioorg. Med. Chem.* **1997**, *5*, 1149. (g) Perreault, D. M.; Cabell, L. A.; Anslyn, E. V. *Bioorg. Med. Chem.* **1997**, *5*, 1197. (h) Gilbert, B. A.; Sha, M.; Wathen, S. T.; Rando, R. R. *Bioorg. Med. Chem.* **1997**, *5*, 1115. (i) Lim, A. C.; Barton, J. K. *Bioorg. Med. Chem.* **1997**, *5*, 1131.

(13) Hendrix, M.; Alper, P. B.; Priestley, E. S.; Wong, C.-H. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 95−98.

(14) (a) Recht, M.; Fourmy, D.; Blanchard, S. C.; Dahlquist, K. D.; Puglisi, J. D. *J. Mol. Biol.* **1996**, *262*, 421. (b) Wong, C.-H.; Hendrix, M.; Priestley, E. S.; Greenberg, W. A. *Chem. Biol.*, in press.

(15) Lee, R. T.; Lee, Y. C. *Carbohyd. Res.* **1974**, *37*, 193−201.

(16) Alper, P. B.; Hendrix, M.; Priestly, E. S.; Wong, C.-H. *J. Am. Chem. Soc.* **1998**, *120*, 1965−1978.

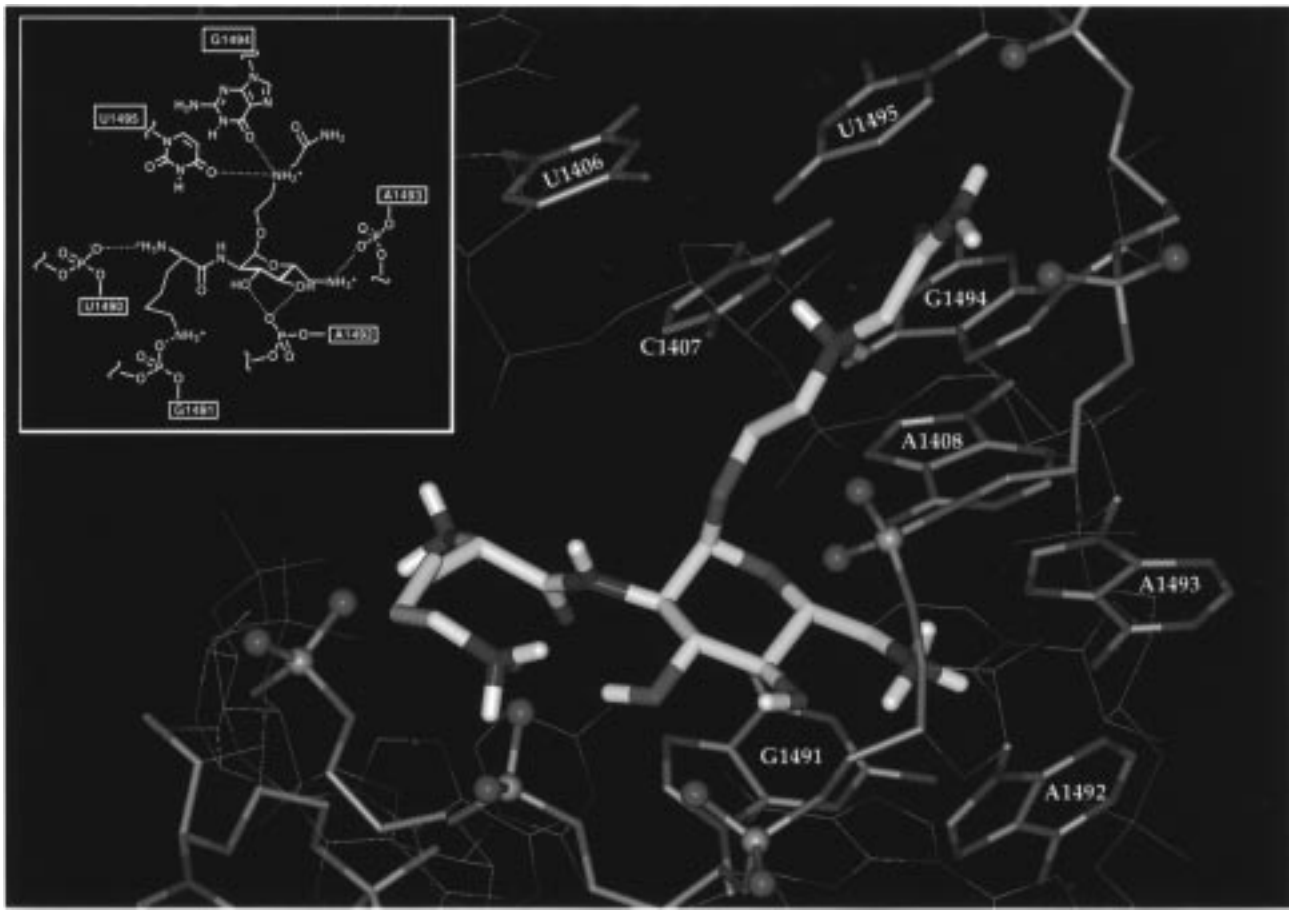
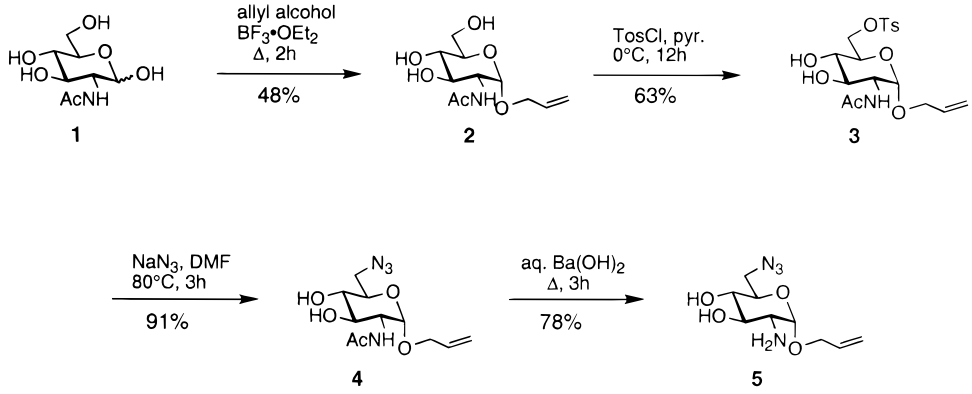


Figure 3. A model of compound **20a** with As-wt.

Scheme 1



by ion-exchange chromatography using Amberlite CG-50 resin (a carboxylic acid containing resin) with aqueous ammonia as the eluent. This procedure afforded final compounds of excellent purity as judged by ^{1}H NMR and ^{13}C NMR. All library members were also characterized by electrospray mass spectrometry, which provided further evidence of the identity of each compound.

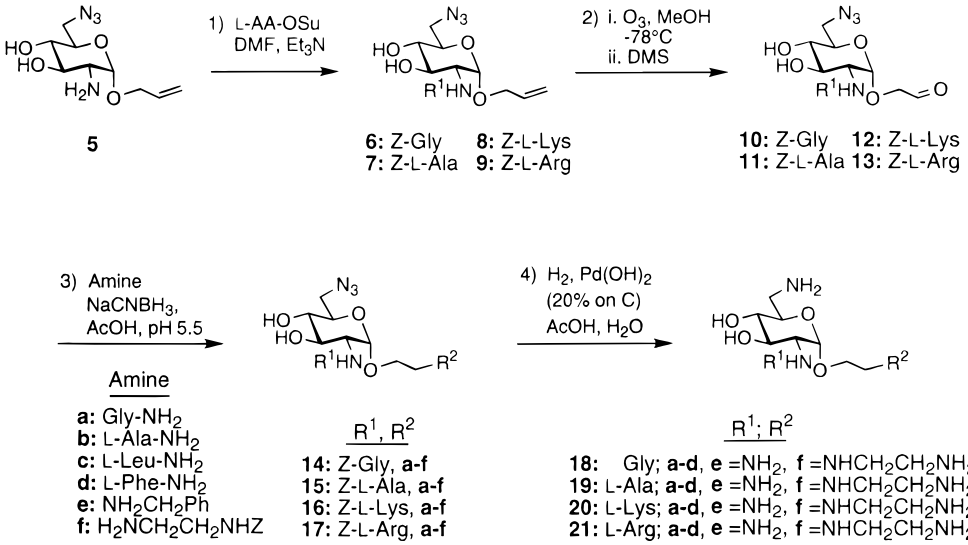
Library Evaluation by SPR

We envisioned two possible ways in which to evaluate the library compounds: a direct testing of their RNA binding affinities to various target RNAs using our surface plasmon resonance (SPR) assay[17] or a screen for biological function, in this case an antibiotic assay.[16] We have chosen the former test as our preferred method of screening at this stage, as it can provide a straightforward assessment of the desired RNA binding properties without being obscured by other issues such as metabolic stability or cellular uptake. Moreover, the SPR assay has been automated so that up to 12 compounds each at four concentrations may be screened against three separate RNA sequences over a 24-h period (Figure 4).

The assay measures the change in SPR angle that occurs upon the binding of an analyte to RNA immobilized on a sensor chip. The RNAs required for immobilization were prepared enzymatically by runoff transcription (Scheme 3). Briefly, RNAs were fabricated by the action of T7 RNA polymerase on an appropriate DNA template primed with guanosine 5′-monophosphorothioate (GMPS).[18] Subsequent alkylation of the product 5′-

(17) Hendrix, M.; Priestley, E. S.; Joyce, G. F.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 3641−3648.

Scheme 2



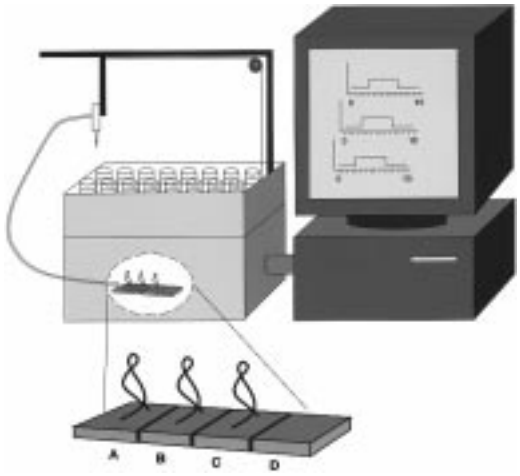


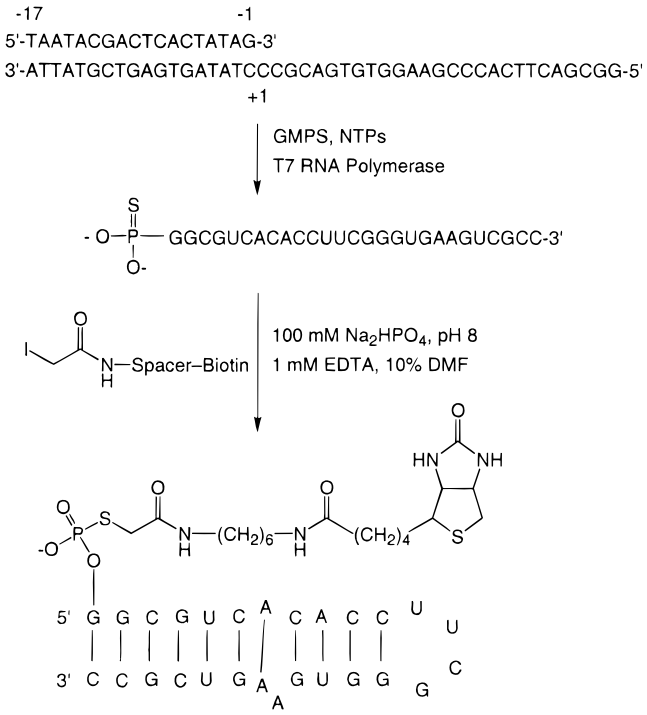
Figure 4. Automated analysis of aminoglycoside mimetics by surface plasmon resonance. Zones A−D represent four distinct regions of the SPR sensorchip that may be individually modified with RNA and simultaneously analyzed for aminoglycoside binding. Twelve aminoglycoside mimetics each at four concentrations may be screened in a 24-h period. For our studies, zones A−C contain immobilized biotinylated RNAs (AS-wt, U1406A, U1495A), while zone D was left unmodified to measure background binding.

phosphorothioate modified transcripts with a biotinylated iodosuccinimide linker affords biotinylated RNAs which when purified by gel electrophoresis were suitable for immobilization on streptavidin-coated sensor chips.

The library compounds were tested for RNA binding with a set of A-site related RNAs (Figure 5).[14,19] Briefly, AS-wt and AS−U1406A are model systems of the aminoglycoside binding site found in the *E. coli* ribosomal decoding region A-site. They form specific complexes with a number of aminoglycosides, particularly the 4,5-linked 2-deoxystreptamine derivatives of the neomycin class. The mutant AS−U1495A no longer possesses this specificity and serves as a negative control for aminoglycoside binding.

Table 1 shows the SPR assay results obtained with AS-wt for library compounds **18**−**21a**−**f**. The library compounds were

Scheme 3



tested at a range of concentrations between 3 and 100 μM. K_d's were derived from a model that includes terms for specific and uniform multisite binding modes (see eq 3, Experimental Section). Representative data from the SPR screen are shown in Figure 6.

There is considerable variation in the affinity of individual library members for the A-site related RNAs. Within a series having the same acyl residue, all compounds have the same number of amines, except for the last entry in each series (**f**, the condensation product with ethylenediamine), thus enabling comparison between compounds of the same charge. In general, compounds **19**−**21a,e** with $NHCH_2CONH_2$ and NH_2 linked to the anomeric center have the highest RNA binding affinity. Compounds with aliphatic hydrophobic residues (alaninamide, leucinamide) show much weaker binding. Interestingly, phenylalaninamide is noticeably better in affinity than leucine in the first two acyl series (i.e., those with glycine or alanine as the acyl residue), but noticeably worse in the third series.

(18) (a) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783. (b) Milligan, J. F.; Uhlenbeck, O. C. *Methods Enzymol.* **1989**, *180*, 51.

(19) Miyaguchi, H.; Narita, H.; Sakamoto, K.; Yokoyama, S. *Nucleic Acids Res.* **1996**, *24*, 3700.

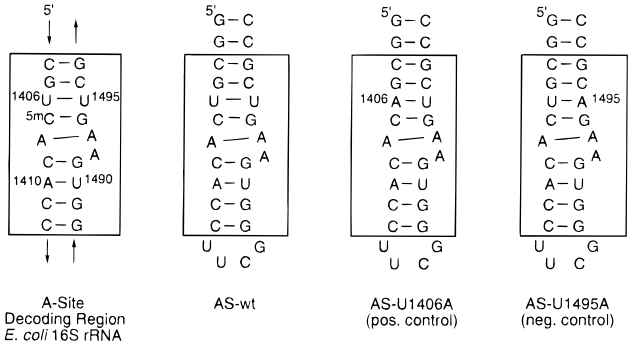


Figure 5. Model RNA sequences and their relationship to wild-type A-site 16s rRNA. AS-wt and AS−U1406A retain the internal loop necessary for specific binding profiles observed in natural aminoglycosides. Mutant AS−U1495A binds aminoglycosides nonspecifically and serves as a negative control for aminoglycoside binding.

Table 1. Screening of the Library against A-Site Related RNAs

	R^1	R^2	As-wt[a]	U1406A[a]	U1495A[a]	Specificity factor[b]
18a	Gly	$GlyNH_2$	130	90	100	0.8
18b		$AlaNH_2$	170	140	200	1.2
18c		$LeuNH_2$	>250	>250	>250	
18d		$PheNH_2$	70	35	70	1
18e		NH_2	100	90	110	1.1
18f		$NH(CH_2)_2NH_2$	50	50	50	1
19a	Ala	$GlyNH_2$	40	30	40	1
19b		$AlaNH_2$	>250	>250	>250	
19c		$LeuNH_2$	>500	>500	>500	
19d		$PheNH_2$	210	170	170	0.8
19e		NH_2	130	130	110	0.8
19f		$NH(CH_2)_2NH_2$	70	70	70	1
20a	Lys	$GlyNH_2$	30	25	40	1.3
20b		$AlaNH_2$	90	70	70	0.8
20c		$LeuNH_2$	>250	190	150	0.6
20d		$PheNH_2$	>250	>250	>250	
20e		NH_2	15	10	10	0.7
20f		$NH(CH_2)_2NH_2$	15	15	15	1
21a	Arg	$GlyNH_2$	60	50	40	0.7
21b		$AlaNH_2$	100	70	60	0.6
21c		$LeuNH_2$	>250	220	200	0.8
21d		$PheNH_2$	130	90	90	0.7
21e		NH_2	50	40	30	0.6
21f		$NH(CH_2)_2NH_2$	10	15	6.0	0.6

[a] K_d values are reported in μM. [b] The specificity factor is the ratio of the K_d for the negative control and the K_d for the wild-type RNA model, K_d(U1495A)/K_d(AS-wt).

The variation observed for the affinities of the RNAs for closely related library members is quite stunning in some cases. Thus, introduction of a single methyl group into compound **19a** to form **19b** led to a greater than 8-fold drop in affinity with the difference to leucine even larger. Thus, small R^2 substituents seem important for high RNA binding affinity in this class of compounds. Introduction of an additional amine into the R^2 substituent was anticipated to increase binding by at least a factor of 10 as is commonly observed for the natural aminoglycosides.[14b] However, only a factor of 2 increase in binding was observed in the **18** and **19** acyl series (mimetics **e** and **f**). For compounds possessing an extra charge (compounds **20** and **21**), the effect was similar, with less than a 10-fold enhancement observed. Possibly the proximity of amines in the ethylenediamine chain perturbs the system so that only one can be significantly protonated at pH 7.4, at which the assay is conducted.[20]

Comparison of the N-acyl side chains among similarly charged derivatives indicates that glycine is slightly more potent than the alanine side chain. For example, glycine derivative **18d** is 5-fold more potent than alanine derivative **19d**. As expected, compounds possessing either a lysine or arginine as the N-acyl side chain are generally more potent than those displaying alanine or glycine, albeit their affinities for the RNA models is similar. These results are presented in a chart shown in Figure 7 to better understand the effects of substituents on binding.

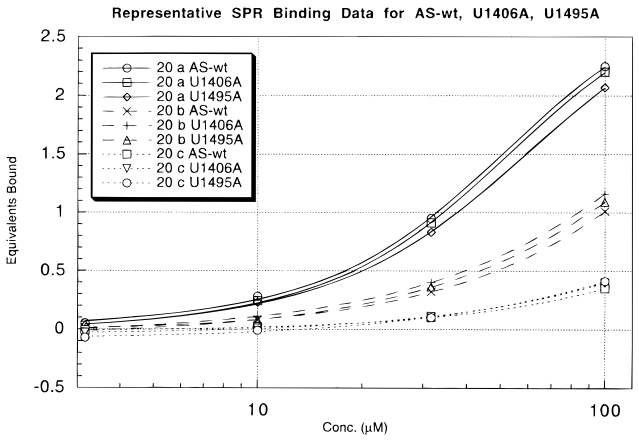


Figure 6. Representative SPR binding data for AS-wt, U1406A, and U1495A.

The A-site related RNA sequences can discriminate between individual library members. Additionally, several of the library compounds (**20a**, **20e**, **20f**, **21f**) have affinities similar to that of the antibiotic neamine which has a dissociation constant of 5 μM[16] as measured by the SPR assay. This can be taken as an indication that the hydroxyamine structure chosen in the scaffold design is an effective core structure to achieve high affinity RNA binding. Whereas the library compounds show generally the same affinity for all three A-site variants, natural aminoglycosides such as neomycin B and synthetic analogues show noticeable differences in affinity for these RNAs, with specificities of 5−20-fold.[14b,16] Molecular modeling suggests these molecules could have specific interactions with conserved bases within the loop region of AS-wt (Figure 3). Perhaps the low specificity observed for the library compounds is related to the flexible nature of side chains R^1 and R^2.

Yeast ribosomal RNA in which the 1409−1491 base pair is disrupted imparts resistance to a wide spectrum of aminoglycoside antibiotics.[21] To address the effectiveness of our aminoglycoside mimetics against drug-resistant strains we examined one mutant (AS-res)[14a] RNA containing a mismatched base pair across the 1409−1451 junction (Figure 8). The affinity for this sequence remained in the micromolar range and was only about 2-fold lower with regard to specificity for AS-wt. Although the mimetics discovered in this study have low specificity for the A-site, they are shown to be effective against the drug-resistant mutant AS-res. It would be of interest to see if the mimetics would bind other RNA sequences, and work is in progress to investigate this specificity issue.

Conclusion

We have demonstrated that a library of small molecule RNA binders can be assembled rapidly in a parallel solution phase approach. An interesting range of RNA binding affinities were obtained with the library compounds, spanning about 2 orders of magnitude. The hydroxyamine core structure is effective

(20) Ethylenediamine: $pK_a^1 = 9.93$, $pK_a^2 = 6.85$. Martell, A. E.; Smith, R. M. *Critical Stability Constants*; Plenum Press: New York, 1974.

(21) Li, M.; Tzagoloff, A.; Underbrink-Lyon, K.; Martin, N. *J. Biol. Chem.* **1982**, *257*, 5921−5928.

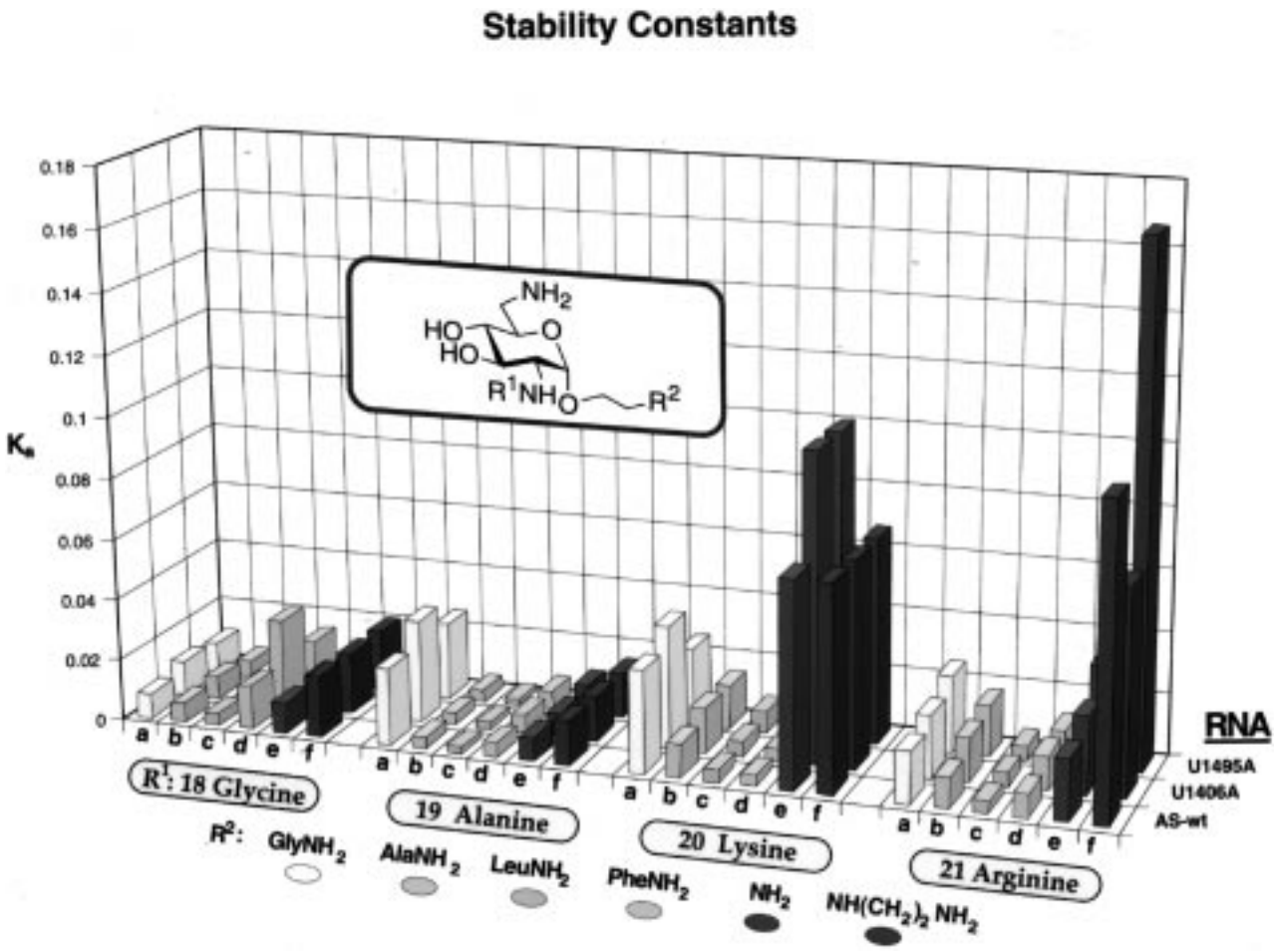
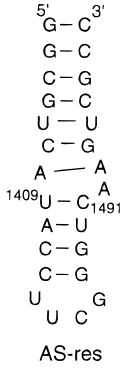


Figure 7. Stability constants for aminoglycoside mimetics versus three RNA sequences: AS-wt, U1406A, U1495A.



	K_d(μM)
19e	190
19f	90
20e	30
20f	30

Figure 8. Mimetic binding to a drug-resistant model. Disruption of the 1409−1491 GC base pair induces broad spectrum aminoglycoside antibiotic resistance in yeast ribosomal RNA.

for the design of high affinity RNA binders, providing hope that with the right choice of substituents specificity may be attained as well.

Experimental Section

General Methods. Reactions requiring anhydrous conditions were carried out in flame dried glassware under argon. NMR spectra were recorded on a Bruker AMX-500 at 500 MHz for ^{1}H and 125 MHz for ^{13}C spectra using TMS or residual solvent as reference standards unless stated otherwise. Mass spectra were recorded on a VG ZAB-ZSE mass spectrometer.

Allyl 6-*O*-Tosyl-α-D-*N*-acetylglucopyranoside (3). Allyl α-D-*N*-acetylglucosaminide (**2**)[15] (200 mg, 0.766 mmol) was dissolved in 3 mL of anhydrous pyridine and cooled to 0 °C in an ice bath. A solution of tosyl chloride (175 mg, 0.918 mmol) was added via syringe. After 1 h, the ice bath was removed and stirring was continued for 12 h. The reaction was quenched with water and concentrated; the residue was purified by column chromatography in ethyl acetate/MeOH (0 to 20%) to afford 200 mg (63%) of **3** as a colorless foam. ^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.78 (d, 2H, J = 8 Hz), 7.32 (d, 2H, J = 8 Hz), 6.14 (d, 1H, J = 8.5 Hz), 5.84 (m, 1H), 5.22 (m, 2H), 4.74 (d, 1H, J = 3.5 Hz), 4.27 (m, 2H), 4.10 (dd, 1H, J = 13, 5 Hz), 4.0 (m, 1H), 3.90 (dd, 1H, J = 13, 6.5 Hz), 3.76 (m, 1H), 3.65 (app dd, 1H, J_1 = J_2 = 10 Hz), 3.46 (dd, 1H, J_1 = J_2 = 10 Hz), 2.42 (s, 3H), 2.03 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 172.0, 144.9, 133.2, 132.8, 129.8, 128.0, 118.1, 96.4, 73.5, 70.6, 69.6, 69.1, 68.3, 53.3, 23.2, 21.6; HRMS *m*/*z* found 548.0355 (M + Cs^+), calculated 548.0367.

Allyl 6-Azido-6-deoxy-α-D-*N*-acetylgludopyranoside (4). A solution of allyl 6-*O*-tosyl-α-D-*N*-acetylglucosaminide (**3**, 190 mg, 0.45 mmol) in DMF containing NaN_3 (300 mg, 4.6 mmol) was heated in an oil bath at 80 °C for 3 h. The solvent was then evaporated and the residue was purified by column chromatography in ethyl acetate/MeOH (0 to 20%) to afford 120 mg (91%) of **4** as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$−10% MeOD) δ = 6.66 (d, 1H, J = 9 Hz), 5.81 (m, 1H), 5.25−5.21 (m, 2H), 4.76 (d, 1H, J = 3.5 Hz), 4.13 (m, 1H), 3.89−3.96 (m, 2H), 3.67 (ddd, 1H, J = 10, 6.5, 1.5 Hz), 3.52 (dd, 1H, J = 10, 9 Hz), 3.46 (dd, 1H, J = 13, 2.5 Hz), 3.40−3.28 (m, 6H), 1.94 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$−10% MeOD) δ = 171.9, 133.2, 117.8, 96.3, 72.2, 71.5, 71.2, 68.2, 53.3, 53.2, 51.4, 22.7; HRMS *m*/*z* found 309.1168 (M + Na^+), calculated 309.1175.

Allyl 2-Amino-2-deoxy-6-azido-6-deoxy-α-D-glucopyranoside (5). A flask fitted with a reflux condenser was charged with allyl 6-deoxy-6-azido-α-D-*N*-acetylglucosaminide (**4**) (960 mg, 3.35 mmol), barium hydroxide octahydrate (14 g), and water (72 mL). The flask was placed into a 120 °C oil bath and the mixture stirred for 3 h. After cooling, the solution was continuously extracted with CH_2Cl_2 for 20 h to afford 639 mg (78%) of **5** as a colorless oil. ^{1}H NMR (CD_3OD) δ = 5.96

(m, 1H), 5.32 (m, 1H), 5.20 (m, 1H), 4.85 (d, 1H, $J = 4$ Hz), 4.24 (m, 1H), 4.04 (m, 1H), 3.72 (ddd, 1H, $J = 9.5, 6.5, 2.5$ Hz), 3.50−3.37 (m, 3H), 3.23 (dd, 1H, $J = 9.5, 9$ Hz), 2.62 (dd, $J = 10, 4$ Hz); ^{13}C NMR (CD_3OD) $\delta = 135.3, 117.7, 99.4, 75.9, 73.3, 72.8, 69.5, 57.1, 52.8$; ESI HRMS *m/z* found 245.1247 (MH^+), calculated 245.1250.

General Procedure for Acylation of Compound 5 with *N*-Hydroxysuccinimide Esters to Afford 2-Acyl Derivatives 6−9. A solution of allyl 2,6-dideoxy-2-amino-6-azido-glucoside **5** (366 mg, 1.50 mmol) in DMF (6 mL) was cooled to 0 °C and treated with triethylamine (0.23 mL, 1.7 mmol) followed by a solution of the succinimide ester (1.5 mmol) in DMF (4 mL). After the reaction was complete (TLC in $CHCl_3$/MeOH = 9:1), the solvent was evaporated, and the residue was taken up in ethyl acetate and washed with saturated $NaHCO_3$ solution. After evaporation, column chromatography afforded the 2-acyl derivatives **6−9** in 75−95% yield.

Compound 6: ^{1}H NMR (400 MHz, CD_3OD) $\delta = 7.36-7.28$ (m, 5H), 5.93 (m, 1H), 5.32−5.09 (m, 4H), 4.83 (d, 1H, 3.5 Hz), 4.2 (dd, 1H, $J = 13$, 5 Hz), 4.02−3.94 (m, 2H), 3.82 (m, 2H), 3.76−3.72 (m, 1H), 3.65 (dd, 1H, $J = 10.5$, 9 Hz), 3.50 (dd, 1H, $J = 13$, 2.5 Hz), 3.40 (dd, 1H, $J = 13$, 6.5 Hz), 3.35−3.29 (m, 1H); ^{13}C NMR (100 MHz, CD_3OD) $\delta = 172.4, 159.0, 138.1, 135.2, 129.5, 129.0, 128.9, 118.0, 73.1, 73.0, 72.7, 69.5, 67.8, 55.2, 52.7, 44.8$; HRMS *m/z* found 568.0821, calculated 568.0808.

Compound 7: ^{1}H NMR (400 MHz, DMSO-d_6) $\delta = 7.75$ (d, 1H, $J = 8$ Hz), 7.43−7.33 (m, 6H), 5.92 (m, 1H), 5.33−5.29 (m, 2H), 5.15 (d, 1H, $J = 10.5$ Hz), 5.03 (m, 2H), 4.90 (d, 1H, $J = 6$ Hz), 4.75 (d, 1H, $J = 3.5$ Hz), 4.19−4.11 (m, 2H), 4.97 (dd, 1H, $J = 9.5, 5.5$ Hz), 3.69−3.61 (m, 2H), 3.52−3.42 (m, 3H), 3.16 (m, 1H), 1.22 (d, 3H, $J = 7$ Hz); ^{13}C NMR (100 MHz, DMSO-d_6) $\delta = 172.9, 155.6, 137.1, 134.3, 128.4, 127.8, 127.7, 117.0, 95.9, 71.6, 70.3, 67.5, 65.3, 53.8, 51.2, 49.9, 18.5$.

Compound 8: ^{1}H NMR (DMSO-d_6) $\delta = 7.76$ (d, 1H, 8 Hz), 7.36−7.25 (m, 6H), 5.86 (m, 1H), 5.30−5.26 (m, 2H), 5.11 (d, 1H, $J = 10.5$ Hz), 5.0 (app s, 2H), 4.98 (app s, 2H), 4.85 (broad s), 4.72 (d, 1H, $J = 3.5$ Hz), 4.12−3.95 (m, 3H), 3.65 (m, 1H), 3.58 (m, 1H), 3.50−3.42 (m, 2H), 3.39 (dd, 1H, $J = 13.5, 6.5$ Hz), 3.12 (m, 2H), 2.94 (m, 2H), 1.60 (m, 1H), 1.58 (m, 1H), 1.45−1.20 (m, 4H); ^{13}C NMR (DMSO-d_6) $\delta = 172.4$ 156.1, 137.1, 134.3, 128.4, 127.8, 127.7, 117.1, 95.9, 71.6, 70.3, 67.5, 65.3, 65.1, 54.6, 53.7, 51.2, 31.9, 29.2, 22.8.

Compound 9: ^{1}H NMR ($CDCl_3$) $\delta = 7.37-7.25$ (m, 15H), 6.32 (d, 1H, $J = 5.0$ Hz), 5.91 (d, 1H, $J = 3.0$ Hz), 5.85−5.75 (m, 1H), 5.28−5.20 (m, 3H), 5.15−5.02 (m, 4H), 4.72 (broad s, 1H), 4.22−4.16 (m, 1H), 4.08 (dd, 1H, $J = 5.0$, 3.1 Hz), 4.00 (t, 1H, $J = 3.1$ Hz), 3.94−3.89 (m, 3H), 3.71−3.67 (m, 1H), 3.53−3.49 (m, 3H), 3.42−3.37 (m, 2H), 1.73−1.62 (m, 8H); ^{13}C NMR ($CDCl_3$) $\delta = 173.0, 162.3, 161.4, 137.8, 134.1, 128.5, 128.3, 128.0, 118.2, 96.8, 73.5, 72.2, 71.4, 68.9, 67.8, 66.5, 55.1, 54.2, 52.8, 44.2, 37.5, 31.2, 25.0$; ESI HRMS *m/z* found 803.3331 (MH^+), calculated 803.3364.

General Procedure for Ozonolysis of Compounds 6−9. A solution of compounds **6−9** (0.15 mmol per reaction) in a total of 7 mL of a MeOH/CH_2Cl_2 mixture (containing only as much CH_2Cl_2 as needed for solubility) was cooled to −78 °C and treated successively with oxygen, then ozone until the solution was blue. After excess ozone had been purged with oxygen, dimethyl sulfide was added (200 μL, 3 mmol) and the solution allowed to warm to ambient temperature (ca. 1 h). The solvent was evaporated and the product dried for 1 h under high vacuum. The crude aldehydes **10−13** were then used in the reductive amination.

General Procedure for Reductive Amination of Compounds 10−13. The aldehydes **10−13** (0.15 mmol) were dissolved in MeOH (1 mL) and treated first with a 1 M solution of the amine in MeOH (0.45 mL), then with a 1 M solution of acetic acid in MeOH (0.5 mL), and finally with a freshly prepared 0.3 M solution of $NaCNBH_3$ in MeOH (0.22 mL). (If an amine hydrochloride salt was used instead of a free amine, water was added to the amine solution as needed for solubility, the amount of AcOH solution was reduced to 0.05 mL, and the amount of $NaCNBH_3$ solution was increased to 0.25 mL.) After 2 h, water (0.5 mL) was added and stirring was continued for 20 min, after which time the mixture was concentrated. The crude products **14−17** were carried on without further purification.

General Procedure for Hydrogenation of Compounds 14−17a−f. A solution of compounds **14−17** (0.15 mmol) in AcOH (3 mL) and water (2 mL) was degassed by evacuating and purging with argon several times. Catalyst (20% $Pd(OH)_2$ on carbon, wet Degussa type, ca. 20 mg) was added; the flask was carefully evacuated and then refilled with hydrogen from a balloon. Hydrogen was bubbled though the solution for ca. 3 min. Then the flask was kept under positive hydrogen pressure with a hydrogen balloon and stirred for 3−12 h until reduction was complete (TLC of the products in MeOH/concentrated NH_3 = 9:1 to 3:1, staining with ninhydrin). The balloon was removed and the solution purged with argon. Water was added (5 mL) and the reaction mixture filtered through a Celite pad, which was washed with water (5 mL). The combined filtrates were concentrated, dissolved in water (2 mL), and applied onto an Amberlite CG-50 column (NH_4^+ form, 16 × 1.5 cm) and eluted with a linear gradient of aqueous ammonia (500 mL total). Gradients of 0 to 1% NH_3(aq) (3 charges, 1 secondary amine), 0 to 5% NH_3(aq) (3 charges, no secondary amine), 0 to 8% NH_3(aq) (4 charges, 1 secondary amine), and 0 to 35% NH_3(aq) (for arginine-containing derivatives) were employed. Fractions (5 mL each) were collected with an automatic fraction collector and the product-containing fractions were pooled and lyophilized. Hydrochloride salts of the final products were prepared by adding excess 1 M HCl and lyophilizing again. All compounds were characterized by ^{1}H NMR, ^{13}C NMR, and electrospray MS.

Compound 18a: ^{1}H NMR (D_2O) $\delta = 4.96$ (d, 1H, $J = 3.5$ Hz), 4.06 (dd, 1H, $J = 10.5$, 3.5 Hz), 4.02−3.99 (m, 1H), 3.95 (app d, 2H, 1 Hz), 3.88−3.85 (m, 3H), 3.76−3.72 (m, 2H), 3.43−3.37 (m, 4H), 3.17 (dd, 1H, $J = 13.5$, 8.5 Hz); ^{13}C NMR (D_2O) $\delta = 169.0, 168.1, 98.1, 72.2, 71.4, 68.9, 63.8, 53.9, 48.4, 47.8, 41.3, 41.1$; ESI LRMS *m/z* found 336 (MH^+), calculated 336.

Compound 18b: ^{1}H NMR (400 MHz, D_2O) $\delta = 4.98$ (d, 1H, $J = 3.5$ Hz), 4.11−4.07 (m, 1H), 3.90−3.87 (m, 3H), 3.79−3.72 (m, 3H), 3.45−3.39 (m, 4H), 3.32−3.28 (m, 1H), 3.19 (dd, 1H, $J = 13$, 8.5 Hz), 1.59 (d, 3H, $J = 7$ Hz); ^{13}C NMR (100 MHz, D_2O) $\delta = 169.8, 99.8, 73.9, 73.1, 70.6, 65.6, 58.5, 55.6, 51.3, 48.2, 43.0, 42.8, 17.9$; ESI, LRMS *m/z* found 350.0 (MH^+), calculated 350.0.

Compound 18c: ^{1}H NMR (D_2O) $\delta = 4.96$ (d, 1H, $J = 3.5$ Hz), 4.08 (d, 1H, $J = 10.5$, 3.5 Hz), 3.99−3.96 (m, 2H), 3.88−3.84 (m, 3H), 3.77−3.71 (m, 2H), 3.44−3.38 (m, 4H), 3.28−3.26 (m, 1H), 3.17 (dd, 1H, $J = 13.5$, 8.5 Hz), 1.86−1.81 (m, 1H), 1.76−1.66 (m, 2H), 0.94 (app t, 6H, $J = 7$ Hz); ^{13}C NMR (D_2O) $\delta = 172.0, 168.0, 98.1, 72.2, 71.5, 69.0, 63.7, 60.3, 53.8, 47.1, 41.3, 41.1, 39.6, 25.0, 22.9, 21.8$; ESI LRMS *m/z* found 392.0 (MH^+), calculated 392.0.

Compound 18d: ^{1}H NMR (D_2O) $\delta = 7.45-7.38$ (m, 5H), 4.96 (d, 1H, $J = 3.5$ Hz), 4.22 (dd, 1H, $J = 9$, 6 Hz), 4.08 (dd, 1H, $J = 10.5$, 3.5 Hz), 4.01−3.98 (m, 1H), 3.88−3.83 (m, 3H), 3.78−3.71 (m, 2H), 3.45−3.41 (m, 3H), 3.36 (dd, 1H, $J = 13.5$, 5.5 Hz), 3.32 (ddd, 1H, $J = 12.5$, 8.5, 3 Hz), 3.21−3.16 (m, 2H); ^{13}C NMR (D_2O) $\delta = 170.9, 168.0, 134.4, 130.3, 129.9, 128.8, 98.1, 72.2, 71.5, 69.0, 63.8, 62.6, 53.8, 47.3, 41.3, 41.1, 36.7$; ESI LRMS *m/z* found 426.0 (MH^+), calculated 426.0.

Compound 18e: ^{1}H NMR (D_2O) $\delta = 4.96$ (d, 1H, $J = 3.5$ Hz), 4.07 (dd, 1H, $J = 10.5$, 3.5 Hz), 3.96 (dt, 1H, $J = 12$, 4 Hz), 3.88−3.83 (m, 3H), 3.75 (dd, 1H, $J = 10.5$, 9 Hz), 3.68−3.64 (m, 1H), 3.44−3.41 (m, 2H), 3.27 (m, 2H), 3.17 (dd, 1 H, $J = 13.5$, 8.5 Hz); ^{13}C NMR (D_2O) $\delta = 168.1, 98.0, 72.2, 71.4, 68.9, 64.7, 53.9, 41.3, 41.1, 39.9$; ESI LRMS *m/z* found 279.0 (MH^+), calculated 279.0.

Compound 18f: ^{1}H NMR (400 MHz, D_2O) $\delta = 4.85$ (d, 1H, $J = 3.7$ Hz), 3.99−3.88 (m, 2H), 3.81−3.71 (m, 3H), 3.69−3.59 (m, 2H), 3.39−3.29 (m, 8H), 3.07 (dd, 1H, $J = 13$, 8 Hz); ^{13}C NMR (100 MHz, D_2O) $\delta = 169.8, 99.8, 73.8, 73.1, 70.6, 65.4, 55.5, 50.1, 46.9, 43.1, 42.8, 37.8$; ESI LRMS *m/z* found 322 (MH^+), calculated 322.

Compound 19a: ^{1}H NMR (D_2O) $\delta = 4.95$ (d, 1H, $J = 3.5$ Hz), 4.08 (q, 1H, $J = 7$ Hz) 4.02−3.99 (m, 2H), 3.94 (app d, 2H, $J = 2.5$ Hz), 3.87−3.83 (m, 1H), 3.77−3.69 (m, 2H), 3.42−3.36 (m, 2H), 3.15 (dd, 1H, $J = 13.5$, 8.5 Hz), 1.52 (d, 3H, $J = 7$ Hz); ^{13}C NMR (D_2O) $\delta = 172.0, 168.9, 98.0, 72.3, 71.2, 69.0, 63.7, 53.9, 50.0, 48.4, 47.7, 41.1, 17.6$; ESI LRMS *m/z* found 250 (MH^+), calculated 250.

Compound 19b: ^{1}H NMR (D_2O) $\delta = 4.93$ (d, 1H, $J = 3.5$ Hz), 4.07−3.95 (m, 4H), 3.85−3.81 (m, 1H), 3.73−3.66 (m, 2H), 3.39−3.33 (m, 3H), 3.25 (ddd, 1H, $J = 12.5$, 8.5, 3 Hz), 3.13 (dd, 1H, $J =$

13.5, 8.5 Hz), 1.53 (d, 3H, J = 7 Hz), 1.47 (d, 3H, J = 7 Hz); ^{13}C NMR (D_2O) δ = 172.8, 172.0, 98.1, 72.4, 71.2, 69.0, 63.9, 56.7, 53.8, 50.0, 46.4, 41.1, 17.6, 16.2; ESI LRMS m/z found 364 (MH^+), calculated 364.

Compound 19c: ^{1}H NMR (D_2O) δ = 4.95 (d, 1H, J = 3.5 Hz), 4.07 (q, 1H, J = 7 Hz), 4.02 (dd, 1H, J = 11, 3.5 Hz), 3.98−3.94 (m, 2H), 3.86−3.82 (m, 1H), 3.74−3.69 (m, 2H), 3.42−3.36 (m, 3H), 3.25 (ddd, 1H, J = 13.5, 9, 3.5 Hz), 3.15 (dd, 1H, J = 13.5, 8.5 Hz), 1.84−1.78 (m, 1H), 1.74−1.65 (m, 2H), 1.51 (d, 3H, J = 7 Hz), 0.92 (app t, 6H, J = 7 Hz); ^{13}C NMR (D_2O) δ = 172.0, 171.9, 98.1, 72.4, 71.3, 69.0, 63.6, 60.2, 53.8, 50.0, 47.0, 41.1, 39.5, 24.9, 22.8, 21.8, 17.6; ESI LRMS m/z found 406 (MH^+), calculated 406.

Compound 19d: ^{1}H NMR (D_2O) δ = 7.40−7.23 (m, 5H), 4.92 (d, 1H, J = 3.5 Hz), 4.18 (dd, 1H, J = 9, 5.5 Hz), 4.05 (q, 1H, J = 7 Hz), 3.99 (dd, 1H, J = 11, 4 Hz), 3.97−3.93 (m, 1H), 3.82 (m, 1H), 3.73−3.67 (m, 2H), 3.39−3.31 (m, 4H), 3.26 (ddd, 1H, J = 13, 8.5, 3.5), 3.13 (m, 2H), 1.48 (d, 3H, J = 7 Hz); ^{13}C NMR (D_2O) δ = 171.9, 170.8, 134.4, 130.3, 129.9, 128.8, 98.1, 72.4, 71.3, 69.0, 63.7, 62.5, 53.8, 50.0, 47.2, 41.1, 36.6, 17.6; ESI LRMS m/z found 440 (MH^+), calculated 440.

Compound 19e: ^{1}H NMR (400 MHz, D_2O) δ = 4.95 (d, 1H, J = 3.5 Hz), 4.06 (q, 1H, J = 7 Hz), 4.00 (dd, 1H, J = 11, 7 Hz), 3.98−3.93 (m, 1H), 3.85 (ddd, 1H, J = 10, 8.5, 3 Hz), 3.72 (dd, 1H, J = 11, 9 Hz), 3.65 (ddd, J = 11, 7, 4.5 Hz), 3.43−3.37 (m, 2H), 3.28−3.25 (m, 2H), 3.16 (dd, J = 13, 8.5 Hz), 1.50 (d, 3H, J = 7 Hz); ^{13}C NMR (100 MHz, D_2O) δ = 173.7, 99.7, 74.0, 72.9, 70.5, 66.3, 55.5, 51.6, 42.8, 41.5, 19.2; ESI LRMS m/z found 293 (MH^+), calculated 293.

Compound 19f: ^{1}H NMR (D_2O) δ = 4.92 (d, 1H, J = 3.5 Hz), 4.10 (q, 1H, J = 7 Hz), 4.00−3.95 (m, 2H), 3.84−3.81 (m, 1H), 3.74−3.66 (m, 2H), 3.44−3.34 (m, 8H), 3.13 (dd, 1H, J = 13, 8.5 Hz), 1.48 (d, 3H, J = 7.5 Hz); ^{13}C NMR (D_2O) δ = 172.0, 98.1, 72.3, 71.2, 68.9, 63.7, 53.8, 50.0, 48.5, 45.2, 41.1, 36.1, 17.7; ESI LRMS m/z found 336 (MH^+), calculated 336.

Compound 20a: ^{1}H NMR (400 MHz, D_2O) δ = 5.00 (d, 1H, J = 3.6 Hz), 4.10−4.06 (m, 3H), 3.99 (app d, 2H, J = 3 Hz), 3.92−3.87 (m, 1H), 3.87−3.75 (m, 2H), 3.47−3.39 (m, 4H), 3.19 (dd, 1H, J = 13, 8.5 Hz), 2.98 (app t, 2H, J = 7.5 Hz), 1.96−1.90 (m, 2H), 1.73−1.65 (m, 2H), 1.52−1.40 (m, 2H); ^{13}C NMR (100 MHz, D_2O) δ = 172.4, 170.6, 99.7, 74.2, 72.7, 70.7, 65.4, 55.6, 55.5, 50.0, 49.4, 42.8, 41.5, 32.8, 28.8, 23.5; ESI LRMS m/z found 407 (MH^+), calculated 407.

Compound 20b: ^{1}H NMR (D_2O) δ = 4.99 (d, 1H, J = 4.0 Hz), 4.11−4.00 (m, 4H), 3.88 (ddd, 1H, J = 9, 9, 3 Hz), 3.79−3.72 (m, 2H), 3.45−3.39 (m, 3H), 3.29 (ddd, 1H, J = 13, 9, 3.5 Hz), 3.18 (dd, 1H, J = 13.5, 9 Hz), 2.97 (app t, 2H, J = 8 Hz), 1.95−1.91 (m, 2H), 1.70−1.66 (m, 2H), 1.59 (d, 3H, J = 7 Hz), 1.48−1.43 (m, 2H); ^{13}C NMR (D_2O) δ = 172.9, 170.7, 98.1, 72.6, 71.1, 69.0, 64.0, 56.8, 54.0, 53.9, 46.4, 41.2, 39.9, 31.2, 27.2, 21.8, 16.2; ESI LRMS m/z found 421 (MH^+), calculated 421.

Compound 20c: ^{1}H NMR (D_2O) δ = 4.95 (d, 1H, J = 3.5 Hz), 4.06−4.02 (m, 2H), 3.98−3.95 (m, 2H), 3.84 (ddd, 1H, J = 9, 9, 5.5 Hz), 3.76−3.70 (m, 2H), 3.42−3.36 (m, 3H), 3.24 (ddd, 1H, J = 11, 9, 5.5 Hz), 3.15 (dd, 1H, J = 13, 8.5 Hz), 2.95 (app t, 2H, J = 7.5 Hz), 1.94−1.88 (m, 2H), 1.85−1.63 (m, 6H), 1.50−1.38 (m, 2H), 0.92 (app t, 6H, J = 7 Hz); ^{13}C NMR (D_2O) δ = 172.0, 170.7, 98.1, 72.6, 71.2, 69.0, 63.7, 60.3, 53.9, 53.8, 47.1, 41.1, 39.8, 39.5, 31.2, 27.2, 25.0, 22.9, 21.8; ESI LRMS m/z found 463 (MH^+), calculated 463.

Compound 20d: ^{1}H NMR (D_2O) δ = 7.44−7.34 (m, 5H), 5.00 (d, 1H, J = 4.25 Hz), 4.25 (dd, 1H, J = 11.5, 7 Hz), 4.10−4.01 (m, 3H), 3.92−3.87 (m, 1H), 3.81−3.75 (m, 2H), 3.48−3.39 (m, 4H), 3.36−3.30 (m, 1H), 3.25−3.18 (m, 2H), 3.00 (app t, 2H, J = 9.5 Hz), 1.99−1.94 (m, 2H), 1.76−1.68 (m, 2H), 1.54−1.43 (m, 2H); ^{13}C NMR (D_2O) δ = 170.8, 170.6, 134.4, 130.3, 129.9, 128.8, 98.1, 72.6, 71.2, 69.0, 63.8, 62.6, 53.9, 53.8, 47.3, 41.1, 39.9, 36.7, 31.2, 27.2, 21.8; ESI LRMS m/z found 497 (MH^+), calculated 497.

Compound 20e: ^{1}H NMR (D_2O) δ = 4.97 (d, 1H, J = 3.5 Hz), 4.07−4.03 (m, 2H), 3.99−3.95 (m, 1H), 3.86 (ddd, 1H, J = 9, 5.5, 2.5 Hz), 3.74 (dd, 1H, J = 10.5, 9 Hz), 3.70−3.65 (m, 1H), 3.45−3.38 (m, 2H), 3.28−3.27 (m, 2H), 3.18 (dd, 1H, J = 13, 8.5 Hz), 2.96 (app t, 2H, J = 7.5 Hz), 1.94−1.90 (m, 2H), 1.71−1.65 (m, 2H), 1.50−1.41 (m, 2H); ^{13}C NMR (D_2O) δ = 170.8, 98.0, 72.6, 71.1, 70.0, 64.7, 53.9, 41.2, 39.9, 31.2, 27.2, 21.8; ESI LRMS m/z found 350 (MH^+), calculated 350.

Compound 20f: ^{1}H NMR (D_2O) δ = 4.99 (d, 1H, J = 3.5 Hz), 4.15 (app t, 1H, J = 6.5 Hz), 4.07−4.02 (m, 2H), 3.90 (ddd, 1H, J = 9.5, 6.5, 3 Hz), 3.80 (dd, 1H, J = 11, 9 Hz), 3.77−3.73 (m, 1H), 3.51−3.39 (m, 8H), 3.19 (dd, 1H, J = 13.5, 8.5 Hz), 2.98 (app t, 2H, J = 7.5 Hz), 1.96−1.92 (m, 2H), 1.71−1.66 (m, 2H), 1.49−1.45 (m, 2H); ^{13}C NMR (D_2O) δ = 170.8, 98.1, 72.6, 71.1, 69.0, 63.8, 54.0, 53.8, 48.7, 45.3, 41.2, 39.9, 36.2, 31.3, 27.2, 21.9; ESI LRMS m/z found 393 (MH^+), calculated 393.

Compound 21a: ^{1}H NMR (D_2O) δ = 4.95 (d, 1H, J = 3.6 Hz), 4.14 (broad s, 1H), 4.04−3.96 (m, 2H), 3.93 (d, 1H, J = 5.9 Hz), 3.86−3.79 (m, 1H), 3.75−3.69 (m, 3H), 3.45−3.35 (m, 3H), 3.20−3.12 (m, 4H), 3.02−2.95 (m, 2H), 1.75−1.65 (m, 2H); ^{13}C NMR (D_2O) δ = 182.0, 171.4, 158.2, 98.7, 73.2, 71.8, 69.7, 65.8, 64.5, 54.5, 49.0, 48.4, 43.6, 41.8, 29.1, 24.1; ESI HRMS m/z found 435.2671 (MH^+), calculated 435.2682.

Compound 21b: ^{1}H NMR (D_2O) δ = 4.96 (d, 1H, J = 3.5 Hz), 4.06−4.00 (m, 4H), 3.90 (app dt, J = 7.0, 2.1 Hz, 1H), 3.76−3.70 (m, 2H), 3.43−3.38 (m, 4H), 3.25−3.15 (m, 5H), 1.74−1.67 (m, 4H), 1.50 (d, J = 6.8 Hz); ^{13}C NMR (D_2O) δ = 181.9, 173.0, 158.2, 98.8, 73.2, 71.8, 69.7, 64.8, 57.4, 54.6, 54.4, 47.0, 41.9, 29.6, 24.9, 24.1, 16.9; ESI HRMS m/z found 449.2827 (MH^+), calculated 449.2836.

Compound 21c: ^{1}H NMR (D_2O) δ = 4.92 (d, 1H, J = 3.5 Hz), 4.06−3.95 (m, 3H), 3.89−3.82 (m, 2H), 3.73−3.65 (m, 3H), 3.48−3.40 (m, 3H), 3.31 (app dt, 1H, J = 10.0, 3.1 Hz), 3.23−3.17 (m, 4H), 1.72−1.62 (m, 5H), 0.95 (t, 6H, J = 9.9 Hz); ^{13}C NMR (D_2O) δ = 182.2, 171.4, 158.6, 98.7, 73.2, 71.9, 69.7, 65.0, 61.0, 54.6, 54.4, 47.7, 41.8, 40.6, 29.7, 25.7, 24.9, 24.4, 23.5, 22.5; ESI HRMS m/z found 491.3289 (MH^+), calculated 491.3305.

Compound 21d: ^{1}H NMR (D_2O) δ = 7.28−7.16 (m, 5H), 4.89 (d, 1H, J = 3.5 Hz), 3.93−3.86 (m, 4H), 3.72−3.68 (m, 2H), 3.62−3.58 (m, 2H), 3.31−3.26 (m, 2H), 3.14−3.02 (m, 7H), 1.50−1.40 (m, 3H); ^{13}C NMR (D_2O) δ = 181.1, 173.3, 171.4, 158.6, 136.2, 130.9, 130.5, 128.1, 98.6, 74.6, 73.1, 69.2, 66.3, 63.9, 55.1, 54.9, 47.5, 41.9, 36.8, 30.2, 25.2, 24.6; ESI HRMS m/z found 525.3133 (MH^+), calculated 525.3149.

Compound 21e: ^{1}H NMR (D_2O) δ = 4.99 (d, 1H, J = 3.9 Hz), 4.09−4.05 (m, 2H), 4.02−3.96 (m, 1H), 3.92−3.85 (m, 1H), 3.77 (t, 1H, J = 12.0 Hz), 3.72−3.66 (m, 1H), 3.47−3.39 (m, 2H), 3.29 (broad s, 2H), 3.24−3.17 (m, 2H), 1.98−1.91 (m, 2H), 1.72−1.66 (m, 2H); ^{13}C NMR (D_2O) δ = 179.7, 172.2, 99.6, 74.2, 72.7, 70.6, 66.3, 55.5, 42.9, 42.8, 41.6, 30.5, 25.9, 23.2; ESI LRMS m/z found 378 (MH^+), calculated 378.

Compound 21f: ^{1}H NMR (D_2O) δ = 4.98 (d, 1H, J = 3.6 Hz), 4.10−4.00 (m, 3H), 3.91−3.86 (m, 2H), 3.82−3.75 (m, 3H), 3.48−3.37 (m, 7H), 3.26−3.19 (m, 3H), 3.10 (t, 1H, J = 4.7 Hz), 1.75−1.66 (m, 2H); ^{13}C NMR (D_2O) δ = 180.9, 176.0, 98.9, 73.2, 71.8, 69.8, 64.5, 59.1, 54.6, 49.2, 42.7, 41.9, 39.2, 29.7, 25.0; ESI HRMS m/z found 421.2899 (MH^+), calculated 421.2887.

Synthesis of Biotinylated RNA. In vitro transcription reactions were performed according to the general procedure reported by Uhlenbeck and co-workers.[18] The DNA templates (for AS-wt, 5′-GGC GUC ACA CCU UCG GGU GAA GUC GCC CCT ATA GTG AGT CGT ATT A-3′; for AS−U1406A, 5′-GGC GAC ACA CCU UCG GGU GAA GUC GCC CCT ATA GTG AGT CGT ATT A-3′; for AS−U1495A, 5′-GGC GUC ACA CCU UCG GGU GAA GAC GCC CCT ATA GTG AGT CGT ATT A-3′; for AS-res, 5′-GGC GUC AUA CCU UCG GGU UAA GUC GCC CCT ATA GTG AGT CGT ATT A-3′) were annealed to a 2-fold excess of 18-mer T7 promoter (5′-TAA TAC GAC TCA CTA TAG-3′) in H_2O by heating to 65 °C and slow cooling to below 37 °C. The 5′-phosphorothioate transcripts were generated by incubating the annealed templates (0.2 μM) in 50 mM Tris (pH 7.5), 15 mM $MgCl_2$, 2 mM spermidine, 5 mM DTT, 2 mM ATP, 2 mM CTP, 2 mM UTP, 0.2 mM GTP, 4 mM guanosine 5′-monophosphorothioate (5′-GMPS) with 5 U/μL T7 RNA polymerase and 0.001 U/μL inorganic pyrophosphatase for 2 h at 37 °C in a total volume of 100 μL. Additional 1-μL aliquots of 20 mM GTP were added at 20-min intervals and 500 U of T7 RNA polymerase were added after 1 h. Reactions were quenched by addition of EDTA,

extracted with phenol, and precipitated with ethanol. Electrophoresis was carried out on 20% denaturing polyacrylamide gels. Full-length transcripts were excised from the gel and eluted into 200 mM NaCl, 10 mM Tris (pH 7.5), 0.5 mM EDTA, and desalted on a Nensorb column (DuPont/NEN). Typically, a 100-μL transcription reaction produced 150−500 pmol of RNA 5′-phosphorothioate. The purified RNA transcripts were resuspended in 18 μL of 100 mM NaH_2PO_4 (pH 8.0) containing 1 mM EDTA and were treated with 2 μL of 20 mM biotin−iodoacetamide in DMF (USB). After 2 h at room temperature (RT), an additional 2 μL of the biotinylation reagent was added and incubation was continued for 1 h. The RNA was ethanol precipitated in the presence of glycogen, gel purified, and desalted as described above.

Immobilization of Biotinylated RNAs. Streptavidin functionalized BIAcore sensorchips were either obtained directly from Pharmacia Biosensor AB (sensorchip SA5) or were prepared from carboxymethylated sensorchips (CM5) by EDC activation followed by injection of streptavidin (Pierce, immunopure grade) in acetate buffer (10 mM, pH 5) over all four flowcells according to the manufacturers recommendations. Prior to immobilization, frozen solutions of biotinylated RNA (1−10 pmol) in 80 μL of buffer (10 mM HEPES, 0.1 mM EDTA, 100 mM NaCl, pH 6.8) were renatured by heating to 80 °C for 2 min followed by slow cooling to RT. Typically, individual flowcells were functionalized by injecting 60 μL of RNA buffer using the QUICK-INJECT command at a flowrate of 2 μL/min, followed by injection of running buffer. Three flowcells were used to immobilize RNA while the fourth remained unmodified to serve as a blank control for matrix affects. Levels of RNA capture and ligand binding were calculated by subtracting response units of the blank flowcell from response units in the RNA functionalized flowcells.

General Procedures for SPR Binding Studies. Surface plasmon resonance measurements were performed using a BIAcore 2000 system from Pharmacia Biosensor AB. Samples were prepared by serial dilutions from stock solutions in RNase free microfuge tubes (Ambion) and were centrifuged at 14 000 rpm for degassing. Unless otherwise noted, all binding studies were carried out using HBS buffer (Pharmacia Biosensor AB) which was used as obtained. Magnesium salts were excluded from the buffer as negligible effects are expected for the binding of aminoglycoside mimetics to the model sequences.[14a] All procedures for binding studies were automated as methods using repetitive cycles of sample injection and regeneration. Typically, buffer was injected in the first two cycles to establish a stable baseline value. Samples were injected at a flowrate of 5−10 μL/min using the KINJECT command. The aminoglycoside mimetics were injected from autoclaved 7-mm plastic vials that were capped with pierceable plastic crimp caps. To minimize carry over, samples were injected in order of increasing concentration. The running buffer was identical with the injection buffer. Mimetic binding to three RNA sequences was monitored at once in three parallel flowcells, with the fourth flowcell containing no immobilized RNA as a control. Compounds were assayed at four concentrations (100, 31.6, 10, 3.16 μM) with injection times 4 min 30 s each. After recording three different series, a control sample of paromomycin was tested to ensure reproducibility. With the available autosampler racks, up to 12 compounds can be screened over a period of 24 h.

Expected values for the equilibrium response of 1 equiv of analyte were calculated from the formula 0.76 × (current response units of RNA bound) × (aminoglycoside molecular weight)/(RNA molecular weight). Binding constants were calculated by fitting the recorded binding isotherm (equivalents bound versus concentration) to a model with two modes of binding. The binding constant K_{ns} describes a general binding mode expected between RNA and aminoglycosides wherein multiple binding events can occur simultaneously. The constant K_a represents a tighter binding specific site, if any, for these ligands. The equation is derived from the third-order general isotherm (eq 1), modified to group higher order binding events with like binding constants ($K_2 = K_3 = K_{ns}$) to give the final eq 3. Grouping K_{ns} binding terms afforded a more robust computational model that does not report negative high-order binding constants in response to small measurement errors. Higher order binding equations ($N = 4$, $N = 5$) which similarly grouped higher order binding constants did not produce significantly more accurate results for conditions in which less than 3 equiv of ligand was bound to the RNA.

$$\text{equivalents bound} = \frac{\sum_{N}^{i=1} i[\mathrm{L}]^i \prod_{i}^{j=1} K_{aj}}{1 + \sum_{N}^{i=1} [\mathrm{L}]^i \prod_{i}^{j=1} K_{aj}} = \frac{K_{a1}[\mathrm{L}] + 2K_{a1}K_{a2}[\mathrm{L}]^2 + 3K_{a1}K_{a2}K_{a3}[\mathrm{L}]^3}{1 + K_{a1}[\mathrm{L}] + K_{a1}K_{a2}[\mathrm{L}]^2 + K_{a1}K_{a2}K_{a3}[\mathrm{L}]^3} \quad (1)$$

$$K_2 = K_3 = K_{ns} \quad (2)$$

$$\text{equivalents bound} = c + \frac{K_a[\mathrm{L}] + 2K_aK_{ns}[\mathrm{L}]^2 + 3K_aK_{ns}^2[\mathrm{L}]^3}{1 + K_a[\mathrm{L}] + K_aK_{ns}[\mathrm{L}]^2 + K_aK_{ns}^2[\mathrm{L}]^3} \quad (3)$$

where c = constant, K_a = equilibrium constant for specific binding, K_{ns} = equilibrium constant for multiple site binding, and [L] = mimetic concentration. A constant c was added to correct for machine baseline inconsistencies. Equation 3 was fit to the data using KaleidaGraph (Abelbeck Software) on a Macintosh computer.

Molecular Modeling. The structure shown is a minimized low energy conformation selected from 100 000 dynamics iterations at 450 K using Insight/Discover. During the dynamics run and the minimization, the coordinates of the 2,6-diaminodideoxy glucopyranoside ring and the RNA were fixed according to the NMR structure of paromomycin bound to AS-wt published by Puglisi and co-workers.[11]

Acknowledgment. This work was supported by Novartis. C.R. thanks Odense Universitet Denmark for financial support. We thank Ian R. Ollmann for assistance in data analysis.

JA980826P