Structure-Based Discovery of Ligands Targeted to the RNA Double Helix†

Qi Chen, Richard H. Shafer,* and Irwin D. Kuntz*

Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143-0446

Received April 1, 1997; Revised Manuscript Received July 3, 1997⊗

ABSTRACT: Ligands capable of specific recognition of RNA structures are of interest in terms of the principles of molecular recognition as well as potential chemotherapeutic applications. We have approached the problem of identifying small molecules with binding specificity for the RNA double helix through application of the DOCK program [Kuntz, I. D., Meng, E. C., and Shoichet, B. K. (1994) *Acc. Chem. Res. 27*, 117−123], a structure-based method for drug discovery. A series of lead compounds was generated through a database search for ligands with shape complementarity to the RNA deep major groove. Compounds were then evaluated with regard to their fit into the minor groove of B DNA. Those compounds predicted to have an optimal fit to the RNA groove and strong discrimination against DNA were examined experimentally. Of the 11 compounds tested, 3, all aminoglycosides, exhibited pronounced stabilization of RNA duplexes against thermal denaturation with only marginal effects on DNA duplexes. One compound, lividomycin, was examined further, and shown to facilitate the ethanol-induced B to A transition in calf thymus DNA. Fluorine NMR solvent isotope shift measurements on RNA duplexes containing 5-fluorouracil provided evidence that lividomycin binds in the RNA major groove. Taken together, these results indicate that lividomycin recognizes the general features of the A conformation of nucleic acids through deep groove binding, confirming the predictions of our DOCK analysis. This approach may be of general utility for identifying ligands possessing specificity for additional RNA structures as well as other nucleic acid structural motifs.

Molecular recognition of nucleic acids by small molecules can occur by virtue of binding specificity at the level of primary, secondary, or tertiary structure. By far, the majority of studies to date have focused on the first of these possibilities, i.e., specificity of binding via readout of a particular base or base pair composition and/or sequence in DNA. Some compounds possess a composite specificity involving recognition of both primary and secondary structural features, such as preferential binding to A-T base pairs in the DNA minor groove, with little or no binding to corresponding RNA sequences. Few ligands, however, exhibit the reverse secondary structure specificity, i.e., binding to RNA in one of its grooves but not to DNA.

Because of the many diseases caused by RNA viruses, including AIDS, compounds capable of specific binding to RNA should be considered in developing effective chemotherapy. Bacterial disease sources are also susceptible to this approach to drug treatment. The duplex RNA motif appears in many different contexts. First, a small number of viruses carry a double-stranded RNA genome. Second, some RNA viruses are based on single RNA strands but require synthesis of a complementary RNA strand as part of their life cycle. Third, retroviruses such as RNA tumor viruses and HIV contain single strands of RNA that are often folded into a variety of secondary and tertiary conformations, including local regions of duplex structure, some of which may be distorted due to base mismatches, bulges, etc. Finally, rRNA and mRNA from all sources involve similar motifs of secondary and tertiary folding. Thus, the opportunities for developing therapeutic agents targeted to regular or distorted duplex RNA structures are manifold. For this approach to succeed, it is apparent that such agents should possess specificity for the RNA target in comparison to DNA, in order to avoid unwanted side effects.

There are relatively few studies on small molecules capable of specific RNA binding. Wilson et al. (*1*) examined a wide range of DNA-binding compounds in a study aimed at delineating the effects of ligand structure on binding affinity to RNA and DNA duplexes. While many of these compounds did show significant RNA binding, none showed greater affinity to the RNA duplex than the DNA duplex. Subsequently, Wilson and co-workers synthesized a series of polycationic ligands, some of which demonstrated significant preference for DNA over RNA, as measured by thermal denaturation studies (*2*). Both charge and steric effects were observed to play a role in groove binding. Based on the positive charge on these ligands, along with the fact that the negative charge density is greatest in the major groove of A-like duplexes, it was inferred that binding occurred in the major (deep) groove of the RNA duplex.

Here we describe an approach for discovery of compounds possessing specificity for the RNA double helix based on the unique geometry of its deep major groove. Using the DOCK methodology (*3*), we have identified several aminoglycosides as candidate ligands, characterized by shape complementarity to the RNA groove. We show that one of these compounds not only binds preferentially to RNA over B-form DNA but also facilitates the B to A transition in calf thymus DNA. We also provide preliminary NMR evidence that this ligand binds in the targeted RNA major groove.

†This work was supported by NIH Grants GM51650 (R.H.S.) and GM31497 (I.D.K.), National Institute of General Medical Science. The UCSF Computer Graphics Laboratory is supported by Grant RR01081, Division of Research Resources, NIH.

* Authors to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1997.

S0006-2960(97)00756-3 CCC: $14.00

MATERIALS AND METHODS

Chemicals. Poly(rA), poly(rU), poly(rI)·poly(rC), poly(dA)·poly(dT), poly(dI)·poly(dC), and sonicated calf thymus DNA were purchased from Pharmacia Biotech, and the oligonucleotide d(GCGAATTCGC) was obtained from Genset. All samples were extensively dialyzed against 1 mM sodium phosphate buffer (pH 6.9), except for calf thymus DNA which was dialyzed against water. The oligonucleotides r(GCAGAXCUGC), r(GCAGAUCXGC), and d(GCGAATXCGC), where X = 5-fluorouracil (FU),[1] were purchased from Oligos Etc. and dialyzed against 1 mM sodium cacodylate buffer (pH 6.9) for several hours. All nucleic acid samples were stored at −20 °C in buffer. DNA concentrations (all expressed per mole of nucleotide) were determined by UV absorbance using the following extinction coefficients (M^{-1} cm^{-1}): poly(rA), $\epsilon_{258} = 9800$; poly(rU), $\epsilon_{260} = 9350$; calf thymus DNA, $\epsilon_{260} = 6600$; poly(dI)·poly(dC), $\epsilon_{245} = 5300$; poly(rI)·poly(rC), $\epsilon_{245} = 5450$; poly(dA)·poly(dT), $\epsilon_{260} = 6000$; d(GCGAATTCGC), $\epsilon_{260} = 7400$. The extinction coefficients of the RNA oligomers were estimated according to the nearest-neighbor method (*4*), with the assumption that the modified oligonucleotides had the same extinction coefficients as their unmodified counterparts.

NMR. ^{19}F-NMR measurements were performed on a QE-300 NMR spectrometer operating at 282.8 MHz. The sweep width was ±8 kHz for RNA and ±4 kHz for DNA, with 16K data points acquired. For DNA, 4K transients were collected with a recycle delay of 2.5 s between each pulse. For RNA, 16K transients were collected with a recycle delay of 2.0 s. The FIDs were processed with an exponential multiplication function of 10 Hz prior to transformation. Chemical shifts were reported with respect to 0.25 M KF in buffer containing 10% D_2O as an external reference. All measurements were performed at 20.0 ± 0.5 °C. The final buffer contained 10 mM sodium phosphate, pH 6.9, 0.5 mM EDTA, and NaCl to yield $[Na^+]_{total} = 25$ mM. The DNA and RNA concentrations were 10 and 5.6 mM, respectively, in 400 μL. Measurements in H_2O were done in the presence of 10% D_2O. These samples were then lyophilized in NMR tubes and the original volumes reconstituted with D_2O for measurements in that solvent.

Optical Spectroscopy. Absorbance spectra were acquired on a Varian Cary 3E UV−visible spectrophotometer. Circular dichroism (CD) spectra were recorded on a Jasco J500A spectropolarimeter. UV melting experiments on DNA, RNA, and their complexes were carried out by following the absorbance at 260 nm as a function of increasing temperature, at a heating rate of 0.5 °C/min. Melting temperatures, T_m, were determined as the temperature corresponding to the maximum slope in the melting profile. The reported ΔT_m values were calculated by subtracting the T_m of the free nucleic acid from that of the complex.

The A↔B transition in calf thymus DNA was monitored using CD spectra recorded at 25 °C. A-DNA was prepared by slowly adding ethanol in small aliquots, with vortexing, into an aqueous solution of calf thymus DNA. Final concentrations were 32 μM DNA, 0.4 mM NaCl, and 80% (v/v) EtOH. Attainment of the A-DNA conformation was checked by measuring the molar circular dichroism ($\Delta\epsilon$) at 270 nm as calculated by $\Delta\epsilon = (\theta 4\pi)/(2.303)(180)C$, where θ is the observed ellipticity in degrees and C the molar concentration of DNA in nucleotides. The value of $\Delta\epsilon$ should be ≥8 for the fully formed A conformation (*5*), and a value of ∼8.7 was usually obtained in our preparations. Once the A conformation was formed, the B conformation was induced by addition of aliquots of 0.4 mM NaCl in water to the 80% ethanol solution of A-DNA. CD spectra were recorded after each addition. The fraction, f, of DNA in the A form was calculated using the formula $f = (\Delta\epsilon - \Delta\epsilon_B)/(\Delta\epsilon_A - \Delta\epsilon_B)$, where the subscripts A and B denote A- and B-form DNA. The fraction in the A form was plotted as a function of ethanol concentration to obtain the ethanol concentration, C_{50}, at which half of the DNA would be in the A form. The effect of ligand binding on the A↔B transition was assayed by carrying out the same titration in the presence of 3.2 μM lividomycin ($P/D = 10$).

DOCK Calculations. The structure-based database search software program DOCK 3.5 (*6*) was used to identify ligands that have the potential to bind specifically to the major groove of A-form duplex RNA. The program consists of several modules, including SPHGEN (*7*) and CHEMGRID (*8*). SPHGEN generates clusters of overlapping spheres that describe the solvent-accessible surface of the target site. Each cluster represents a possible binding site for small molecules. A satisfactory cluster of spheres should contain only spheres that evenly cover the whole area of the target site. Typically, good clusters contain at least 20 spheres. Site coverage can be visually examined using computer graphic programs such as MidasPlus (*9*). CHEMGRID precalculates and stores in the grid file the information necessary for force field scoring. This scoring function approximates molecular mechanics interaction energies and consists of van der Waals and electrostatic components (*8*). DOCK 3.5 uses the selected cluster of spheres to orient ligand molecules in the targeted site. Each molecule in the structure database is tested in thousands of orientations within the site, and each orientation is evaluated by the scoring function. Only the best scoring orientation for each compound is stored to the output file, and compounds are ranked in order of their scores. The best candidates are then screened experimentally.

Only a few X-ray structures of A-form RNA duplexes have been published. The recent crystal structure (2.6 Å resolution) of r(UAAGGAGGUGAU)·r(AUCACCUCCUUA) (*10*) was selected as our model structure for the A-form RNA duplex. Its atomic coordinates (File Name 1sdr) were obtained from the Brookhaven Protein Databank (PDB). Each unit cell contains two double-stranded RNA molecules with an rmsd of only 0.69 Å between the two structures. We used the first structure in the PDB file. The target site was focused on the central region of base pairs 4−9 in the major groove. None of the SPHGEN-generated clusters of spheres was large enough to fit well into the whole site, although many spheres lay inside the major groove. All clusters but 1 contained fewer than 10 spheres. The only large one, which contained 21 spheres, had many spheres outside the major groove. We used a reclustering program, CLUSTER (*11*), and manual editing of sphere clusters to find a satisfactory cluster which contained 37 spheres. Twelve critical spheres *(13)* were selected after careful visual examination. They were located around the central bases and deep inside the groove. During DOCK 3.5 runs, the

[1] Abbreviations: FID, free induction decay; FU, 5-fluorouracil; CD, circular dichroism; SIS, solvent isotope shift; rmsd, root mean square deviation.

Table 1: DOCK 3.5 Input Parameters Differing from Defaults

variable	value	variable	value
mode	search	ratio_minimum	0.6
ligand_binsize	0.4	atom_minimum	12
ligand_overlap	0.1	atom_maximum	120
receptor_binsize	0.4	minimize	yes
receptor_overlap	0.1	check_degeneracy	yes
bump_maximum	2	degeneracy_wobble	2
critical_clusters	yes	degenerate_save_interval	25

force field scoring option and energy minimization with degeneracy checking were used (*6*). Default values were used for SPHGEN, CLUSTER, and CHEMGRID as exemplified in the DOCK 3.5 manual. Values of a few DOCK 3.5 parameters that were different from the defaults are listed in Table 1. About 150 000 compounds in the Available Chemicals Database (ACD, Molecular Design Limited Information Systems, San Leandro, CA) were partitioned into 18 subsets according to their formal charges, and the number of compounds in each subset was limited to 10 000. Compounds in different subsets were ranked separately.

In order to differentiate RNA major groove binding from DNA minor groove binding, the crystal structure of the duplex $[d(CCGGCGCCGG)]_2$ (*12*) was used as a "reference" structure (2.2 Å resolution, PDB File Name 1cgc). The same DOCK procedures were performed on the minor groove of the DNA duplex for the top 150 ranking compounds in each subset generated by the RNA docking results, leading to a total of 2700 compounds. The cluster contained 23 spheres, covering the 6 central base pairs. These compounds were further partitioned into five subsets according to their formal charges, i.e., neutral, +1, +2, +3, and ≥+4. The differential score for each compound was calculated by subtracting the score for the DNA minor groove from that for the RNA major groove. The top 50−150 ranking compounds from each new subset (total of ∼400 compounds) were graphically examined using MidasPlus (*9*). Total CPU time was about 450 h on one R4400 Indigo II machine.

Ligand Selection. Several criteria were used to select the compounds that were tested experimentally. Of primary importance, the DOCK-predicted orientation should fit snugly into the major groove, with at most a small part of the molecule projecting into solvent. Such compounds usually have an extended, curved shape dictated by the RNA groove geometry. Rigid planar compounds were discounted to eliminate the possibility of intercalation. Aqueous solubility, chemical and thermal stability, commercial availability, and cost were also considered. The primary experimental assay to test binding involved measuring the effect of ligands on the thermal stability of poly(rA)·poly(rU). Those candidates which stabilized the RNA duplex were examined in more detail.

RESULTS

DOCK Studies. With the implementation of energy minimization on orientation in DOCK 3.5 (*6*), limited computing resources were of major concern in the selection of docking parameters. The combination of critical sphere clustering (*13*) and the high value of the ratio_minimum parameter (Table 1), which screens out ligands much smaller than the binding site size, speeded up the docking search process by at least 5-fold. The target site in the RNA major groove extended over ∼6 base pairs, as defined by the chosen

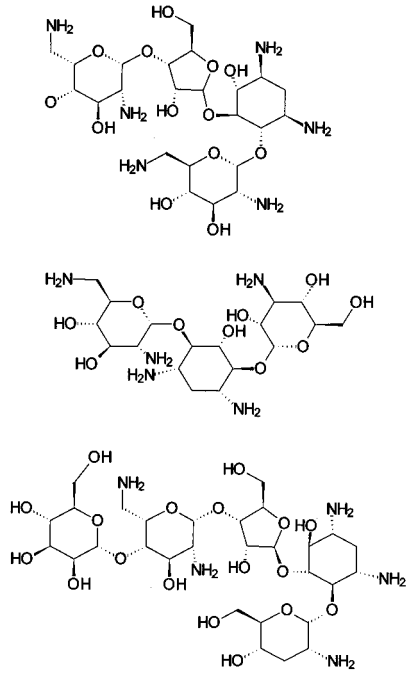

FIGURE 1: Chemical structures of lividomycin (bottom), kanamycin (middle), and neomycin (top).

spheres. The longest distance between receptor spheres was 18.8 Å. The ligand-to-receptor ratio was set at 0.6, ensuring that only compounds covering more than 3 base pairs would be considered. This filtered out almost 60% of the compounds in the ACD. Compared to the DNA minor groove, the RNA major groove is similar in width but much deeper. The critical spheres inside the major groove forced the program to discard those orientations located largely on the edge or close to the outside of the groove.

Among ∼400 top-ranking compounds examined graphically, most of them fit very well into the major groove. Other criteria were then applied along with their DOCK differential score ranking to help select compounds for experimental testing. All neutral compounds were eliminated due to concerns over limited aqueous solubility. Finally, 11 compounds were selected for testing, 3 of which significantly stabilized poly(rA)·poly(rU) against thermal denaturation. These three ligands, lividomycin A, kanamycin B, and neomycin B, all aminoglycosides (Figure 1), were examined in further experiments to more fully characterize their interaction with RNA duplexes. Their DOCK differential scores ranked in the top five both in their subset (i.e., charge ≥+4) and in the entire ACD.

Binding Specificity. Specificity for RNA was determined by measuring the thermal stabilization effect of these three ligands on various RNA and DNA sequences, as illustrated in Figure 2. Very strong preference for binding to RNA [poly(rA)·poly(rU)] over DNA [poly(dA)·poly(dT)] was observed for lividomycin and kanamycin. Both compounds only marginally stabilized the DNA duplex (ΔT_m = 2.3 and 1.4 °C, respectively) while significantly increasing the T_m of the RNA duplex (ΔT_m = 15.2 and 14.3 °C, respectively). Substantial preference was also observed for neomycin, which stabilized both RNA (ΔT_m = 24.7 °C) and DNA (ΔT_m = 13.4 °C). Additional studies carried out with lividomycin, summarized in Table 2, revealed an apparent sequence dependence in its binding to different RNA duplexes. Lividomycin exhibited significantly greater thermal stabiliza-

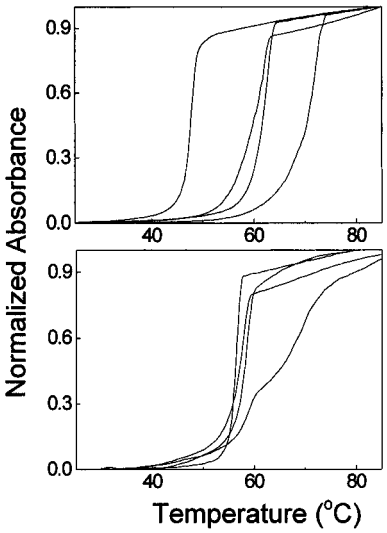

FIGURE 2: Normalized melting curves, monitored at 260 nm, for poly(rA)·poly(rU) and its complexes (top) and poly(dA)·poly(dT) and its complexes (bottom) in 10 mM sodium phosphate buffer, 25 mM $[Na^+]_{total}$, pH 6.9, at $P/D = 10$. From left to right: nucleic acid, 50 μM, alone, and in the presence of kanamycin, lividomycin, or neomycin.

Table 2: Effect of Lividomycin on the Thermal Stability of RNA and DNA Duplexes ($P/D = 10$)

RNA	ΔT_m (°C)	DNA	ΔT_m (°C)
poly(rI)·poly(rC)	38.1	poly(dI)·poly(dC)	3.0
poly(rA)·poly(rU)	15.2	poly(dA)·poly(dT)	2.3
		calf thymus DNA	1.2
$[r(GCAGAUCUGC)]_2$	11.2	$[d(GCGAATTCGC)]_2$	0.0

tion of poly(I)·poly(C) than poly(A)·poly(U). The poor stabilization of the synthetic polydeoxynucleotides by lividomycin was echoed in the results for calf thymus DNA. Thus, the weak binding to DNA is not a peculiarity of the synthetic duplexes. Lividomycin also stabilized the RNA decamer duplex, but to a lesser extent than observed for either polyribonucleotide.

Preference for the A Conformation. Having established that these aminoglycosides bind preferentially to RNA, which is in the A conformation, we sought to determine whether this preference is based on conformational or other differences between DNA and RNA. If the preferential binding of aminoglycosides reflects an underlying specificity for the A conformation of either RNA or DNA, then a direct thermodynamic consequence would be the facilitation of the B to A transition in DNA by these ligands. A ligand which preferentially binds to the A form of DNA should decrease the amount of ethanol needed to reach the midpoint of the transition. Thus, we examined the effect of lividomycin on the ethanol-induced B to A transition in calf thymus DNA (*5*). Results from a CD titration show just this effect, as illustrated in Figure 3. The ethanol concentration required to reach the midpoint of the A−B transition was decreased from 72% in the absence of lividomycin to 64% in its presence. Similar results have been reported for spermine− and spermidine−DNA complexes (*14*).

Binding Site Identification. In order to determine experimentally the binding site of lividomycin, we exploited the sensitivity of ^{19}F-NMR chemical shifts to environmental factors (*15*). When FU is incorporated into nucleic acid duplexes, the fluorine atom lies in the major groove,

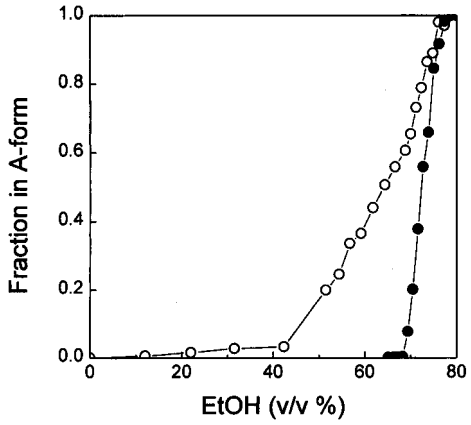

FIGURE 3: CD titration for the ethanol-induced A ↔ B transition. Calf thymus DNA, 32 μM, alone (●) and in the presence of 3.2 μM lividomycin (○).

providing a probe for binding interactions in that groove. The simple observation of chemical shift changes upon ligand binding is not a reliable measure of major groove location, due to the aforementioned sensitivity of ^{19}F resonances. But a more complex experiment, based on changes in ^{19}F chemical shifts of solvent-exposed nuclei upon replacement of H_2O with D_2O (solvent isotope shift, or SIS), can be useful in this regard (*15*). The presence of a ligand tightly bound in the major groove should limit the solvent exposure of major groove atoms and thereby decrease or eliminate the SIS. Measurement of ^{19}F SIS has been used to identify intermolecular contacts in protein−DNA complexes (*16*).

This method of probing the major groove for intermolecular interactions has not been applied to small molecule−nucleic acid complexes. As there are few confirmed small molecules that bind to the RNA or DNA major groove, the feasibility of this approach for small ligands was examined by carrying out a negative control with the DNA minor groove binding agent distamycin. The DNA duplex [d(GCGAAT-FU-CGC)]$_2$ had one peak at 10.88 ppm in H_2O (data not shown). Changing solvent to D_2O shifted the peak upfield to 10.72 ppm, yielding a SIS of 0.16 ppm, similar to SIS values of 0.13−0.17 ppm reported for 5-FU incorporated in B-DNA (*16*). Upon binding distamycin A, the symmetry of the duplex was destroyed, and two new, broader peaks appeared upfield at 10.23 and 8.66 ppm. As expected, the same SIS of 0.16 ppm was observed for the DNA−distamycin complex as for free DNA.

Solvent isotope shift measurements were carried out on the following two self-complementary fluorine-labeled sequences, [r(GCAGA-FU-CUGC)]$_2$ (S1) and [r(GCAGAUC-FU-GC)]$_2$ (S2), each carrying a single 5FU base substitution for U. The effect of lividomycin binding to these RNA duplexes was examined by measuring the ^{19}F chemical shift in the absence and presence of ligand and by determining the SIS for the uncomplexed and complexed oligoribonucleotides. From the spectra shown in Figure 4, it can be seen that addition of the ligand led to a small downfield shift in the case of S1 and a small upfield shift for S2, with substantial broadening in both cases. The SIS for the free duplex S1 was 0.16 ppm and was reduced to 0.04 ppm upon complexation with lividomycin. This substantial decrease in SIS upon ligand binding to S1 provides evidence for lividomycin binding in the RNA major groove, in agreement with the DOCK results (Figure 5). In contrast, the SIS for

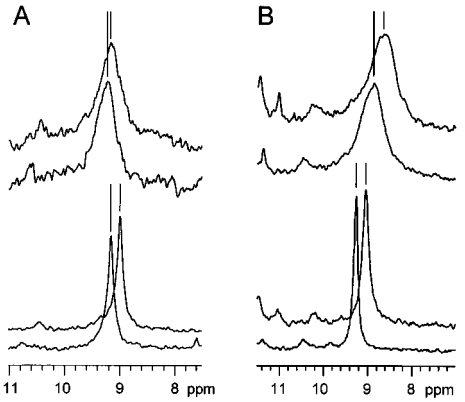

FIGURE 4: ^{19}F-NMR spectra of [r(GCAGAUC-FU-GC)]$_2$ (A) and [r(GCAGA-FU-CUGC)]$_2$ (B), from bottom to top: alone in H_2O, alone in D_2O, complexed with lividomycin in H_2O, and in D_2O, $P/D = 10$; same buffer as Figure 2.

free duplex S2 was 0.22 ppm and remained constant with addition of ligand. This latter result is likely due to end effects. Taken together, these two experiments speak to a possible preference for ligand binding at the central portion of this short duplex.

DISCUSSION

Specific recognition of RNA structures may provide a basis for development of a wide variety of chemotherapeutic agents. In the work described above, we employed the DOCK method to generate ligands capable of binding to the RNA double helix based on the shape and chemical complementarity of the target site and ligand. In order to include specificity for RNA over DNA duplexes, potential RNA binding ligands were also evaluated for their ability to bind to the DNA minor groove, the most common site for nonintercalative DNA binding of small molecules.

Several aminoglycosides were observed to significantly stabilize several RNA duplexes against thermal denaturation while affording only very weak stabilization of the analogous DNA duplexes. This is consistent with recognition of the deep major groove characteristic of the A conformation. These compounds simply do not appear to fit well to B-form DNA. The extent to which other aminoglycosides may behave similarly remains to be determined. While the ACD database contains other structurally similar compounds, the three examined here were the highest scoring of this class. Ethidium displacement experiments indicate that the binding of these aminoglycosides to RNA duplexes occurs with low micromolar equilibrium dissociation constants (data not shown). DOCK computations require a three-dimensional target structure, and hence were carried out for specific RNA and DNA sequences. The extent to which ligand binding occurs for different sequences suggests that the molecular recognition is based on general conformational features rather than specific sequence effects.

Further evidence in support of this interpretation can be found in the observation that lividomycin preferentially binds to the A form relative to the B form of calf thymus DNA, as determined by its lowering of the ethanol concentration corresponding to the midpoint of the A↔B transition. Thus, ligand recognition does not appear to depend directly on the primary structural differences between DNA and RNA but rather indirectly in terms of secondary structure conformation.

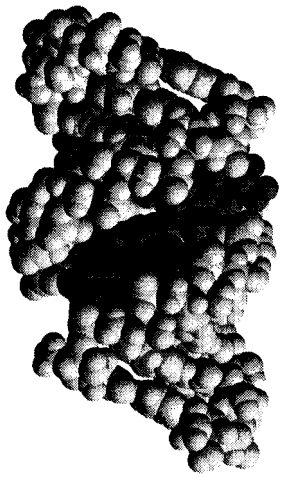

FIGURE 5: Model of lividomycin−RNA complex as predicted by DOCK. Lividomycin is shown in dark shading.

The large differences in ΔT_m (Table 2) between poly(rI)·poly(rC) and poly(rA)·poly(rU) point to a certain degree of sequence specificity in lividomycin binding to RNA. This could arise from hydrogen bonding between the ligands and the major groove groups. For example, hypoxanthine has an oxygen at the 6 position while adenine has an amino group. Similarly, cytosine has an amino group at the 4 position in contrast to the oxygen of uracil. Thus, recognition of the RNA major groove may be primarily determined by shape complementarity but modulated by sequence-specific interactions.

While thermal denaturation and ethanol titration data provide evidence for preferential binding to A-form duplexes, they do not provide information on the binding site location. To address this question, we used FU incorporated into an RNA duplex and determined that the fluorine nucleus, located in the major groove, was protected from solvent by virtue of the presence of bound ligand. This approach assumes that incorporation of FU does not significantly perturb the RNA structure. A recent NMR study has shown this to be the case for the duplex [r(GCGAAUUCGC)]$_2$ (*17*). The structure of the complex between lividomycin and the r(UAAGGAGGUGAU)·r(AUCACCUCCUUA) duplex, as visualized by the DOCK program, is shown in Figure 5. Determination of a detailed binding geometry will require analysis by multidimensional NMR or X-ray diffraction.

Aminoglycosides comprise a large class of antibiotics which interact with prokaryotic ribosomes, probably binding directly to rRNA (*18*). Certain aminoglycosides have been shown to bind to other RNA structures, such as ribozymes (*19*) and the HIV Rev RRE (*20*), with structural evidence in some cases for binding in a distorted major groove (*21*). Little information is available on their binding to regular RNA duplexes and their ability to distinguish RNA from DNA duplexes. There is one report on a neomycin−DNA duplex with the DNA exhibiting an A-like conformation (*22*), consistent with results presented here. Due to the highly cationic nature of these compounds, electrostatic interactions most certainly exert a driving force in their binding to RNA. Our results suggest that their selectivity for RNA over DNA, however, derives from shape complementarity. The ability of these compounds to bind to a wide variety of specific RNA structures may have a common basis in the recognition of the A conformation major groove. Furthermore, the poor

DNA binding described here for several of these compounds may contribute to their therapeutic efficacy by minimizing the potential toxicity due to DNA binding.

In summary, we have demonstrated the ability of DOCK to identify compounds capable of specific recognition of the RNA major groove. Recently, we have described the application of DOCK to the problem of ligand recognition of DNA quadruplexes (*23*). Taken together, these results provide strong evidence that this structure-based method will be a generally useful tool in the discovery of ligands which recognize particular secondary and tertiary nucleic acid motifs.

ACKNOWLEDGMENT

We thank TRIPOS, Inc. and Molecular Design Limited Information Systems for software. We gratefully acknowledge the use of the facilities of the UCSF Computer Graphics Laboratory.

REFERENCES

1. Wilson, W. D., Ratmeyer, L., Zhao, M., Strekowski, L., and Boykin, D. (1993) *Biochemistry 32*, 4098−4104.
2. McConnaughie, A. W., Spychala, J., Zhao, M., Boykin, D., and Wilson, W. D. (1994) *J. Med. Chem. 37*, 1063−1069.
3. Kuntz, I. D., Meng, E. C., and Shoichet, B. K. (1994) *Acc. Chem. Res. 27*, 117−123.
4. Fasman, G. D. (1975) *Handbook of Biochemistry and Molecular Biology*, Vol. I, 3rd ed., CRC Press, Cleveland.
5. Ivanov, V. I., and Krylov, D. Y. (1992) *Methods Enzymol. 211*, 111−127.
6. Gschwend, D. A., and Kuntz, I. D. (1996) *J. Compt.-Aided Mol. Des. 10*, 123−132.
7. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., and Ferrin, T. E. (1982) *J. Mol. Biol. 161*, 269−288.
8. Meng, E. C., Shoichet, B. K., and Kuntz, I. D. (1992) *J. Comput. Chem. 13*, 505−524.
9. Ferrin, T. E., Huang, C. C., and Javis, L. E. (1988) *J. Mol. Graphics 6*, 13−27.
10. Schindelin, H., Zhang, M., Bald, R., Furste, J. P., Erdmann, V. A., and Heinemann, U. (1995) *J. Mol. Biol. 249*, 595−603.
11. Shoichet, B. K., and Kuntz, I. D. (1993) *Protein Eng. 6*, 723−732.
12. Heinemann, U., Alings, C., and Bansal, M. (1992) *EMBO J. 11*, 1931−1939.
13. DesJarlais, R. L., and Dixon, J. S. (1994) *J. Comput.-Aided Mol. Des. 8*, 231−242.
14. Minyat, E. E., Ivanov, V. I., Kritzyn, A. M., Minchenkova, L. E., and Schyolkina, A. K. (1979) *J. Mol. Biol. 128*, 397−409.
15. Rastinejad, F., Evilia, C., and Lu, P. (1995) *Methods Enzymol. 261*, 560−575.
16. Metzler, W. J., and Lu, P. (1989) *J. Mol. Biol. 205*, 149−164.
17. Sahasrabudhe, P. V., Pon, R. T., and Gmeiner, W. H. (1996) *Biochemistry 35*, 13597−13608.
18. Cundliffe, E. (1989) *Annu. Rev. Microbiol. 43*, 207−233.
19. Sage, T. K., Hertel, K. J., and Uhlenbeck, O. C. (1995) *RNA 1*, 95−101.
20. Zapp, M. L., Stern, S., and Green, M. R. (1993) *Cell 74*, 969−978.
21. Fourmy, D., Recht, M. I., Blanchard, S. C., and Puglisi, J. D. (1996) *Science 274*, 1367−371.
22. Robinson, H., and Wang, A. H. (1996) *Nucleic Acids Res. 24*, 676−682.
23. Chen, Q., Kuntz, I. D., and Shafer, R. H. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2635−2639.

BI970756J