Modeling RNA−Ligand Interactions: The Rev-Binding Element RNA−Aminoglycoside Complex

Fabrice Leclerc and Robert Cedergren*

Département de Biochimie, Université de Montréal, C.P. 6128, Succursale Centre-ville Montréal, Québec H3C3J7, Canada

Received June 6, 1997[⊗]

An approach to the modeling of ligand−RNA complexes has been developed by combining three-dimensional structure−activity relationship (3D-SAR) computations with a docking protocol. The ability of 3D-SAR to predict bound conformations of flexible ligands was first assessed by attempting to reconstruct the known, bound conformations of phenyloxazolines complexed with human rhinovirus 14 (HRV14) RNA. Subsequently, the same 3D-SAR analysis was applied to the identification of bound conformations of aminoglycosides which associate with the Rev-binding element (RBE) RNA. Bound conformations were identified by parsing ligand conformational data sets with pharmacophores determined by the 3D-SAR analysis. These "bioactive" structures were docked to the receptor RNA, and optimization of the complex was undertaken by extensive searching of ligand conformational space coupled with molecular dynamics computations. The similarity between the bound conformations of the ligand from the 3D-SAR analysis and those found in the docking protocol suggests that this methodology is valid for the prediction of bound ligand conformations and the modeling of the structure of the ligand−RNA complexes.

Introduction

Research over the past several years has amply demonstrated the importance of RNA−ligand complexes in cellular processes. Consequently, there is a growing interest in targeting RNA complexes for therapeutics. A case in point is the aminoglycoside inhibition of the association between the Rev protein and the Rev-binding element RNA (RBE) of HIV-1, which determines the fate of viral mRNA.[1,2] Although the conformation of the RBE bound to the Rev peptide has been predicted by modeling[3] and subsequently confirmed by two NMR studies,[4,5] this structure offers little indication as to how an aminoglycoside could bind to the RNA in a way to prevent the interaction with the Rev protein.

Modeling and docking a ligand to a receptor is computationally complex in general because of the requirement to find mutually complementary sites in two conformationally flexible molecules. But in the case of RNA complexes, more difficulties arise, because many of the computational tools available for studying the structure of molecular complexes have been developed with proteins in mind. Since NMR and X-ray structures of RNA molecules are becoming more commonplace and a method of predicting bound conformations of RNA from low-resolution chemical data and in vitro selection from random libraries[3,6] has been developed in our laboratory, we have recently focused more on methods which determine the bioactive conformation of the ligand and the structure of the complex rather than the bound RNA conformation. Previously, we showed that structure−activity relationship (SAR) analysis of conformationally rigid ligands can be used to obtain information on the three-dimensional structure of the complex between DNA and quinolones[7] and that a docking protocol based on electrostatic and van der Waals energies can be applied to modeling of the Rev peptide−RBE complex.[8]

Here, we first assess the validity and reliability of the SAR method to identify bioactive conformations for the case of conformationally flexible inhibitors, the phenyloxazolines, whose bound structures with human rhinovirus 14 (HRV14) have been determined by X-ray crystallography.[9] Then, the same approach is applied to the study of the RBE-binding aminoglycosides where the bioactive conformation is unknown. A docking protocol that incorporates the binding properties of the aminoglycosides inferred from the 3D-SAR study is then used to predict the binding conformation of the aminoglycosides within the RNA binding site. The final model of the complex is supported by the high similarity between the bioactive conformations of the ligands produced by the SAR study and the docking protocol as well as the ability of the model to rationalize available experimental data from the complex.

Results and Discussion

We have devised a scheme based on a 3D-SAR analysis and a docking protocol to model the bound, bioactive conformation of aminoglycosides and the RBE RNA−aminoglycoside complex (Figure 1). The method involves the prediction of the biologically significant conformations by a 3D-SAR analysis using extensive conformational generation (Figure 1, steps 1−4). The bioactive conformers thus identified are used as starting conformations for a docking and modeling protocol. During docking extensive conformational sampling is used again which, together with molecular mechanics, produces a new series of bioactive conformations (steps 5 and 6). The reliability of models is then evaluated by comparison of the bioactive conformations identified by 3D-SAR and docking (step 7); this step provides a type

* Corresponding author. Tel: (514) 343-6320. Fax: (514) 343-2177. E-mail: ceder@poste.umontreal.ca.

[⊗] Abstract published in *Advance ACS Abstracts,* December 15, 1997.

S0022-2623(97)00372-5 CCC: $15.00
Published on Web 01/15/1998

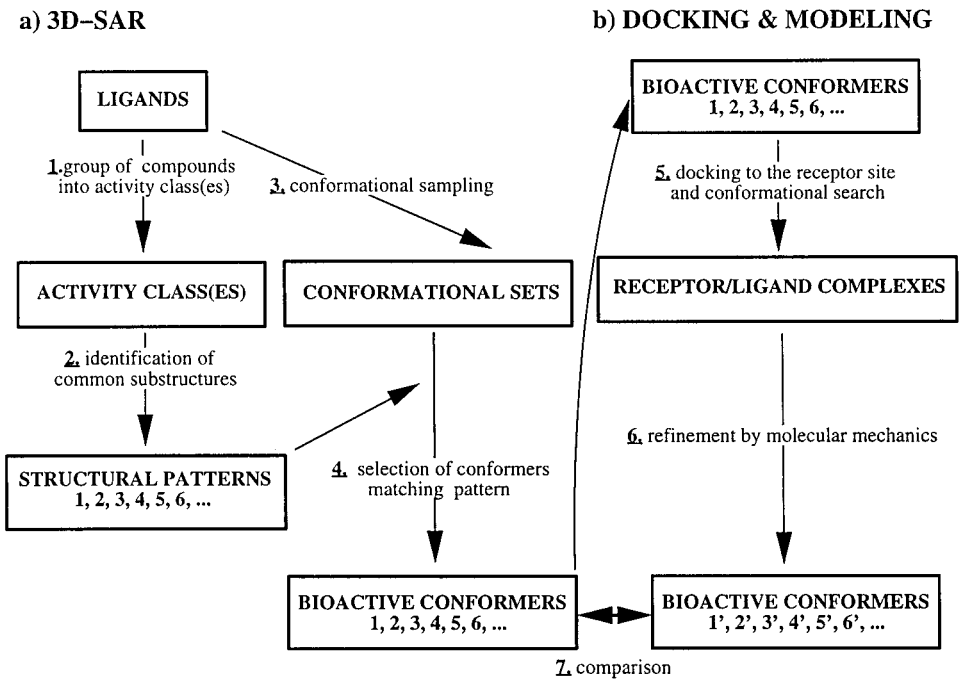
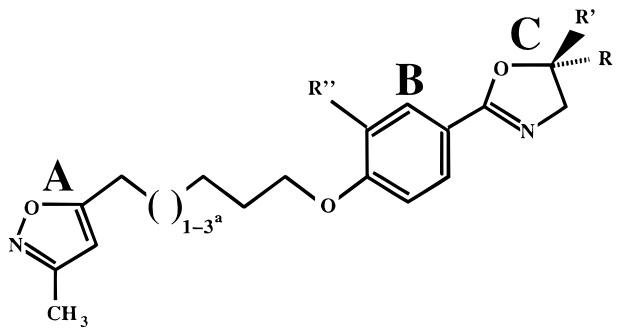


Figure 1. Flowchart for the proposed protocol: (a) 3D-SAR approach (steps 1−4), (b) docking and modeling approach (steps 5, 6) and comparison of the bioactive conformations (step 7).



name	R	R'	R''	PDB	MIC (μM)
I1 [a]	Me	H	H	2RS1	0.03
I2 [a]	H	Me	H	2RR1	0.4
I3 [a]	Me	H	Cl	2RM2	0.2
I4 [a]	Et	H	H	2RS3	0.02
I5 [a]	H	H	H	2R04	0.6
I6	Me	H	H	2RS5	0.6
I7	H	H	H	2R06	0.5
I8	H	H	Cl	2R07	2.4

Figure 2. WIN compounds and their activities: chemical structures and the respective biological activities. The letter x represents three methylene groups in compounds **I1**−**I5** and one in compounds **I6**−**I8**. A represents the oxazole and B, C the phenyloxazole group.

of internal validation of the modeling procedure. The 3D-SAR aspect of the approach was tested first using the inhibitors of HRV14 (the WIN compounds), and subsequently the entire strategy was applied to the study of the RBE-binding aminoglycosides.

Bioactive Conformations of WIN Compounds. The antiviral WIN compounds shown in Figure 2 were chosen for testing since the X-ray crystallographic structures revealed two alternate orientations of the compounds, an indication of high flexibility:[10,11] compounds **I1**−**I4** have the isoxazole, ring A, in the "heel" of the binding pocket composed of amino acids T216, L106, C199, and Y197 (Figure 3a), and compounds **I5**−**I8** have the isoxazole in the "toe" of the binding pocket composed of amino acids P174, V176, and A24 (Figure 3b).

To predict the binding conformation of the WIN compounds, we first combined all eight compounds into a single class in order to find all possible binding orientations and then subsequently determined bioactive conformations for each ligand orientation after classifying them into two groups. The eight WIN compounds produced a number of pharmacophores identified by the APEX-3D program (Table 1); however, the pharmacophores could be divided into two types: one which featured only one ligand orientation with respect to the receptor and the second which allowed two orientations for these groups as found in the crystallographic stuctures. The statistics for the best pharmocophore of each type are given in Table 1.

The detection of two potential binding orientations led us to separate the compounds into two groups, which in fact corresponded to the higher activity compounds (**I1**−**I4**) and the lower activity group, **I5**−**I8** (Figure 2). Table 1 shows the statistically most significant pharmacophores for each group. The bioactive conformation of each compound, shown in Figure 3c,d, was then obtained by parsing the conformational sets generated for each ligand in a group with the pharmacophore identified for the group. These conformations were then compared with the ligand conformation in the crystal structure. The best pharmacophore in group I aligned the ligands very well: the rmsd between the 3D-SAR and crystal structure was 0.0 for compounds **I1**−**I3** and 1.3 Å for **I4** (group I, Figure 3c). The best pharmacophore of the group II compounds predicted bioactive conformations of these compounds within a distance of 1.4 Å rmsd from the crystal structure for **I6**−**I8** and 2.0 Å for **I5** (Figure 3d).

Thus, the simple qualitative SAR model permitted the identification of pharmacophores that correctly describe

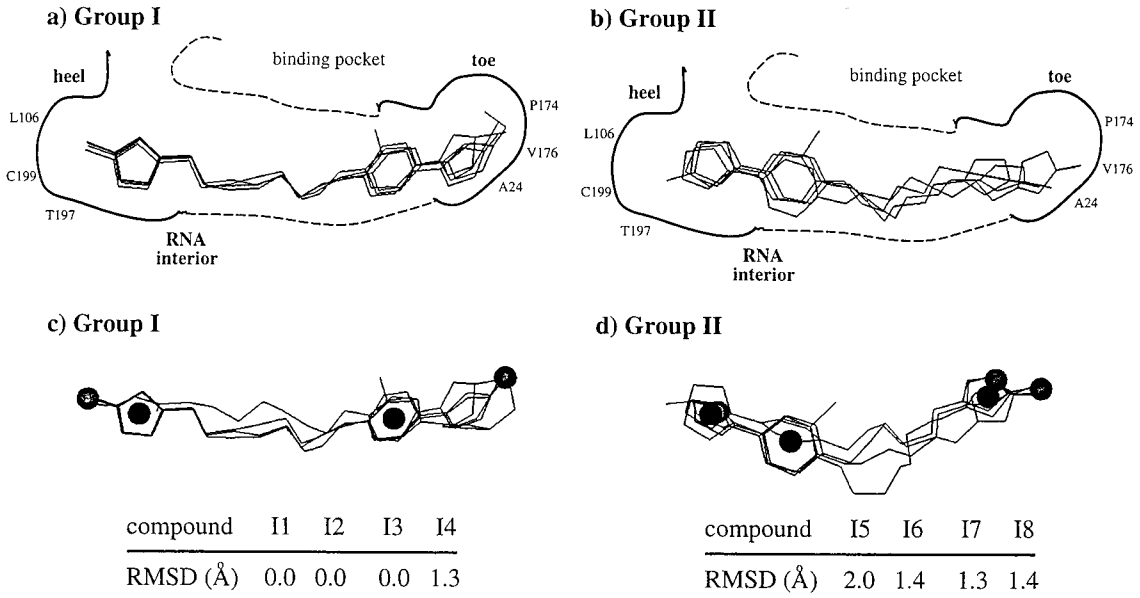


Figure 3. Conformations and binding modes of the WIN inhibitors: (a) crystal structures of inhibitors **I1**−**I4** superimposed according to their position and orientation in the binding pocket, (b) crystal structures of **I5**−**I8** superimposed according to their position and orientation in the binding pocket, (c) alignment of group I compounds using the pharmacophore of Table 1 (rmsd values refer to the differences between the 3D-SAR structure and that of the crystal), (d) alignment of group II compounds using the pharmacophore of Table 1.

Table 1. Pharmacophores of WIN Compounds

data compd	no. of pharmacophores	orientation	probability	reliability	match
all	151[a]	one	0.90	n/a	0.38
		two	0.90	n/a	0.27
group I	55[b]	one	0.83	0.94	0.67
group II	6[b]	one	0.83	0.94	0.36

[a] Number of pharmacophores found for all compounds having a probability ≥0.80 and match ≥0.20. This number included two types of pharmacophores: one having one orientation for all ligands and the other two orientations. [b] Number of pharmacophores found for group I (**I1**−**I4**) and group II (**I5**−**I8**) compounds having a probability ≥0.83. Among these pharmacophores those having the highest match value are indicated.

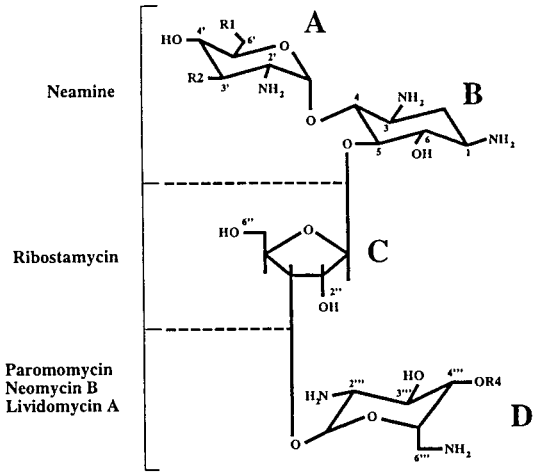
the binding orientation and bioactive conformation of seven of the eight WIN compounds within 1.4 Å rmsd, and this result must be viewed in the light of the fact that the receptor structure was not used during the determination.

Bioactive Conformations of Aminoglycosides. Ten aminoglycosides which inhibit the Rev−RBE RNA complex formation belong to two structural families: the 4,5-diglycosyl-2-deoxystreptamines (4,5-DOS) including neomycin B, ribostamycin, lividomycin A, and paromomycin (Figure 4a) and the 4,6-diglycosyl-2-deoxystreptamines (4,6-DOS) including tobramycin, kanamycin A and B, and gentamicin C1a (Figure 4b). Thus one of the subordinate goals of this study is to determine whether the two structural classes bind to the RNA receptor (Figure 4c) in the same manner. For this reason our analysis focuses on neomycin and tobramycin, the two most active inhibitors belonging to the two structural classes.

The most significant pharmacophores shared by the highly active compounds, neomycin B, lividomycin A, tobramycin, ribostamycin, kanamycin B, and gentamycin C1a, are shown in Figure 5 and summarized in Table 2. Significantly, these patterns show that the core structure defined by rings A and B binds similarly for both the 4,5- and 4,6-DOS compounds. The pharmacophores were then used to parse the conformations and align the compounds. The rmsd calculated between the bioactive conformations predicted by pharmacophores 3 and 4 was 2.3 and 2.9 Å for neomycin and tobramycin, respectively, whereas the rmsd between the conformations predicted by pharmacophores 4 and 5 was 1.4 Å for neomycin. The most similar pair of bioactive conformers was produced by pharmacophores 1 and 4 where the difference for neomycin and tobramycin was 1.2 and 0.6 Å, respectively.

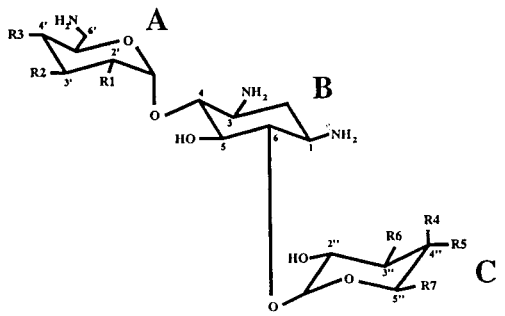
Docking of Aminoglycoside Antibiotics. Docked models were produced for the neomycin B (high activity), tobramycin (high activity), kanamycin B (low activity), paromomycin (inactive) complexes. The initial structure for each complex, models 1−5, was constructed with the five bioactive conformers produced by the pharmacophores (Figure 5). Models were also built using the energy-minimized conformations of the ligands and were identified as models M1 and M2. In all cases, ligands were bound to the previously suggested G46, G47, and G48 region[1] (Figure 4c) of the bound RBE conformation constructed by Leclerc et al.[3] and supported by NMR data.[4,5] Two orientations of the ligand with respect to the RNA binding site were evaluated during docking (see Methods). Only those initial models whose interaction energies were within 50 kcal/mol of the lowest energy model constructed with the energy-minimized ligand conformation were retained for further study (Table 3). Subsequently, searches of ligand conformational space of the retained models were undertaken by exploring the sterically available torsional angles of the ligand until an optimal value for the interaction energy was obtained for each model (see Methods). The significance of the models was evaluated by comparing them with two control models, C1 and C2, which were constructed by docking the aminoglycoside to the region of nucleotides 50, 51, and 52, instead of

a) 4,5–DOS compounds



	R1	R2	R3	R4	[inhibition]	conformers
neamine	NH_2	OH	H	–	> 100 μM	20
ribostamycin	NH_2	OH	H	–	100 μM	104
neomycin B	NH_2	OH	H	H	1 μM	361
paromomycin	OH	OH	H	H	> 100 μM	316
lividomycin A	OH	H	H	mannose	10 μM	304

b) 4,6–DOS compounds



	R1	R2	R3	R4	R5	R6	R7	[inhibition]	conformers
kanamycin A	OH	OH	OH	H	OH	NH_2	CH_2OH	> 100 μM	170
kanamycin B	NH_2	OH	OH	H	OH	NH_2	CH_2OH	100 μM	205
tobramycin	NH_2	H	OH	H	OH	NH_2	CH_2OH	10 μM	214
gentamicin C1a	NH_2	H	H	OH	CH	$NHCH_3$	H	100μM	214
sisomicin	NH_2	H	Δ4'	OH	CH	$NHCH_3$	H	> 100μM	103

c) RBE RNA

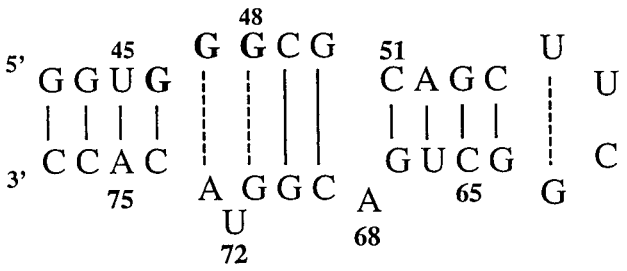


Figure 4. Aminoglycoside inhibitors of the RBE RNA: (a) chemical structures and biological activities of the 4,5-DOS compounds, (b) chemical structures and biological activities of the 4,6-DOS compounds, (c) secondary structure of the RBE RNA. Watson–Crick base pairs are indicated by solid lines and the non-Watson–Crick base pairs by dashed lines. The positions protected from chemical modification by neomycin and tobramycin are indicated in bold.

the G46, G47, and G48 region in two opposite orientations (see Methods). Complexes were solvated and energy-optimized; Table 3 shows the association free energies of the models for the four representative aminoglycosides.

In the case of the 4,5-DOS compounds neomycin and paromomycin, the bioactive conformations selected by pharmacophores 4 and 5 converged to the same model 5 during optimization. Similarly, the bioactive conformations of 4,6-DOS compounds (tobramycin and kanamycin B) from pharmacophores 1 and 4 converged to model 1. The three models with the most favorable association free energies (models 2, 5, and M1) were submitted to molecular dynamics (MD) simulations for further refinement. In these cases, an explicit representation of the solvent which allows displacement of the ligand from its binding site by water molecules was used to evaluate their relative stability. After MD, the geometries of the final structures were similarly optimized, and the association free energies were calculated to determine the degree of accordance between the models and the experimental data. As shown in Table 3, the optimal association free energy is obtained for model 5 in the case of both neomycin and tobramycin supporting the thesis of a common binding mode for 4,5-DOS and 4,6-DOS compounds.

The docking protocol produced bioactive conformations from pharmacophore 2 or the control M1 (models 2 and M1) which significantly deviated from the SAR conformations for both neomycin and tobramycin. However the conformations of model 5 and the SAR-derived structure were quite similar: the heavy atom rmsd calculated between the 3D-SAR bioactive conformation and those obtained after MD is 1.2 Å for both neomycin and tobramycin. We propose model 5 as the best representation of the complex structure based on its consistent results and highest association free energy.

Model of the RBE–Aminoglycoside Complex. The docking and modeling strategy used in this study produced model 5 of the neomycin B or tobramycin and

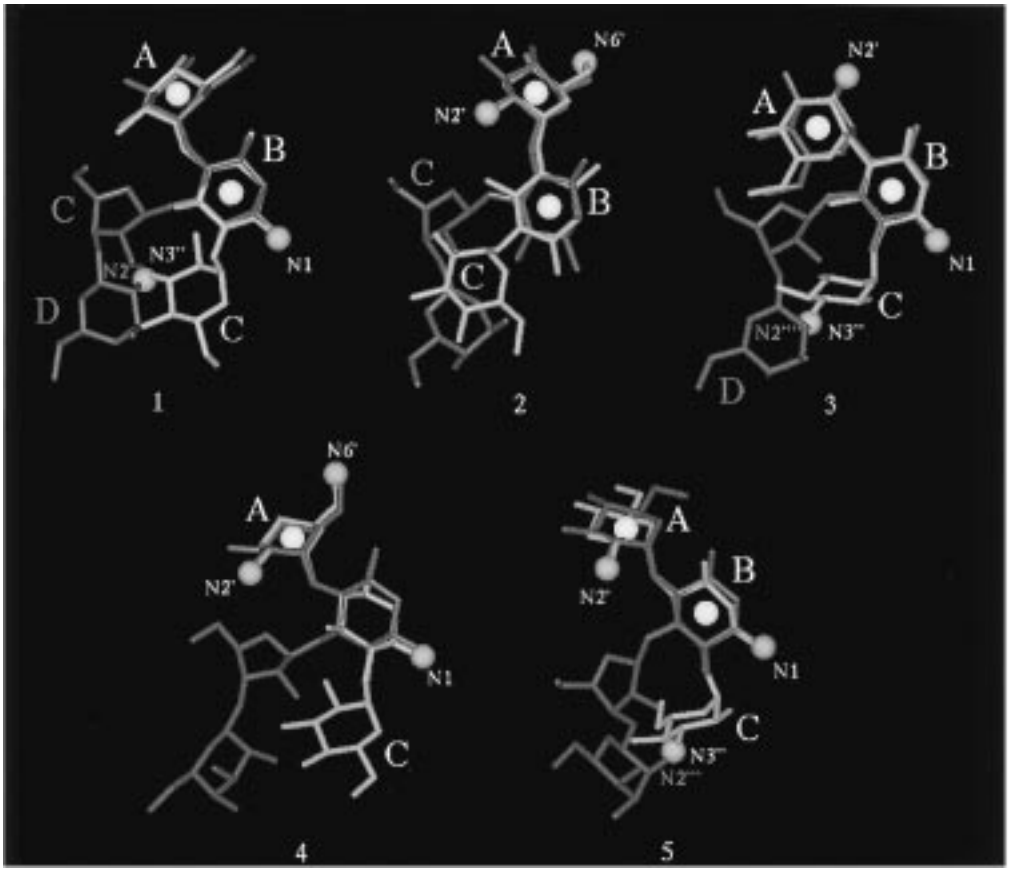


Figure 5. Pharmacophores and bioactive conformations of neomycin and tobramycin. Five pharmacophores were used to align inhibitors illustrated with the neomycin (blue) and tobramycin (orange) bioactive conformations. The pharmacophoric centers are indicated by yellow spheres in the case of ring centers and cyan spheres in the case of atomic centers.

Table 2. Pharmacophores of Aminoglycoside Antibiotics

data compd	no. of pharmacophores	pharmacophore[a]	probability	reliability	match
active	29[b]	1	0.88	0.97	0.40
		2	0.86	0.97	0.57
		3	0.86	0.95	0.55
		4	0.86	0.95	0.52
		5	0.86	0.95	0.47

[a] Five best pharmacophores based on probability. [b] Pharmacophores present in at least five of the six highly active compounds identified having a probability ≥0.85.

the RBE RNA shown in Figure 6 (top and bottom, respectively). These models reveal that the local interactions between both neomycin and tobramycin and the RBE RNA involve all the atomic centers defined by the patterns of pharmacophores 2 and 5. For example, N2′ makes a contact with the Hoogsteen face of G47 and G48 in model 2 and the phosphates of G46 and G47 in model 5. Moreover, the association free energies of this model can rationalize differences in inhibitory strength. The best free energies of complexes reported in Table 3 for the two 4,5-DOS compounds, neomycin B (strong, model 5) and paromomycin (weak, model M1), differ by more than 15 kcal/mol as do the energies between the two 4,6-DOS compounds, tobramycin (strong, model 5) and kanamycin B (weak, model 5).

The models generated for the RBE−aminoglycoside complexes are in agreement with the chemical protection data showing that modifications of the nucleotides G46, G47, and G48 are blocked by neomycin B and tobramycin.[1] The models are also consistent with in vitro selection experiments,[2] where, for example, a decrease in binding affinity of neomycin is observed by substituting the non-Watson−Crick base pair A48:A71 for G48:G71. This observation suggests a direct contact between these residues and the aminoglycoside,[2] a feature of our model (Figure 6).

Although the binding mode is presently unknown, Zapp et al.[1] have proposed an interaction of aminoglycosides with nucleotides G46, G47, and G48 of the RBE RNA. Also, Robinson et al.[12] suggested that the binding involves formation of a bridge across the major groove. Our model 5 supports both these proposals. Even though docking was initiated in the G46−G48 region, modeling did involve extensive translation of the ligands with respect to the RBE, and only suboptimal models were obtained by binding ligands to other regions of the RNA. Parts of the neomycin and tobramycin (rings B, C, and D for neomycin) aminoglycosides bind the region of G46, G47, and G48, and the key nitrogen atom N6′ (ring A) identified in the 3D-SAR study (present in neomycin but absent in the inactive aminoglycoside paromomycin) forms a contact on the opposite side of the major groove with C65 and G66. Recent experimental data demonstrated that aminoglycoside binding to the RBE RNA is stoichiometric and suggested that hydrophobic and/or stacking interactions could occur between the pyrene moieties (ring B; see Figure 6) and the RBE RNA.[13] Although our model did not incorporate this most recent information, model 5 features a B ring buried in the major groove consistent with the experimental data.

Table 3. Binding Energy Contributions for the Different Molecular Models of RBE−Aminoglycoside Complexes after Docking and MD

		association free energies (kcal/mol)[a]					
		after docking			after MD		
aminoglycosides	model[b]	ΔG_{el}	ΔG_{n}	ΔG_{total}	ΔG_{el}	ΔG_{n}	ΔG_{total}
neomycin B	1	−30	−46	−76			
	2	−35	−56	−91	−31	−41	−72
	3	−23	−46	−69			
	5 (4)[c]	−41	−50	−91	−25	−56	−81
	M1	−43	−48	−91	−31	−30	−61
	C1	−34	−48	−82			
	C2	−27	−50	−77			
tobramycin	1 (4)[c]	−25	−44	−69			
	2	−30	−42	−72	−35	−34	−69
	3	−27	−42	−69			
	5	−35	−46	−81	−32	−48	−79
	M1	−27	−38	−65	−28	−30	−58
	M2	−26	−34	−60			
	C1	−28	−44	−72			
	C2	−24	−34	−58			
kanamycin B	1 (4)[c]	−23	−42	−65			
	2	−26	−48	−74	−28	−28	−56
	3	−24	−44	−68			
	5	−30	−42	−72	−27	−33	−60
	M1	−34	−42	−76	−30	−29	−59
	M2	−26	−36	−62			
	C1	−26	−42	−68			
	C2	−20	−40	−60			
paromomycin	1	−20	−46	−66			
	2	−20	−50	−68			
	3	−16	−48	−64			
	5 (4)[c]	−23	−45	−69			
	M1	−29	−50	−79	−15	−42	−57
	C1	−25	−42	−67			
	C2	−18	−50	−68			

[a] The association free energy, ΔG_{total}, was calculated according to the approach used by King and Barford[22] based on the separation of the electrostatic and hydrophobic contributions. The electrostatic contribution, ΔG_{el}, was calculated by the FDPB method as described by Srinivasan et al.[8] The hydrophobic contribution was calculated from the surface-accessible surface area with a proportionality constant of 40 cal/mol Å^2. [b] The models identified by the same root name correspond to a common superposition mode of the core structure among the different aminoglycosides; in parentheses are the pharmacophores which the bioactive conformations led to the same model. [c] The SAR model indicated in parentheses led to the same docked model.

Finally, it is noteworthy that our modeled structure of the neomycin B−RBE RNA complex shares some common structural features with the structure of the complex formed between the 16S rRNA and paromomycin.[14] Both binding regions include an internal loop, featuring non-Watson−Crick base pair interactions and bulged nucleotides, sandwiched between two stems. The similarities between the two complexes reside in the aminoglycoside bioactive conformation and the specificity of the contacts formed with the RNA receptor. In our model, the aminoglycoside adopts an L-form conformation in which rings B, C, and D form a linear array similar to that observed in the 16S rRNA complex. The aminoglycoside exhibits chair conformations for rings A, B (with the amino and hydroxyl groups in equatorial positions), and D with the amino group N6" in an equatorial position and the other substituents (hydroxyl and amino groups) in axial positions, another structural feature retrieved in our model (Figure 6, top).

Rings A and B establish the more specific contacts. In the 16S rRNA, the amino group N1 interacts with the essential nucleotide U1495,[14] whereas in the RBE RNA, it interacts with the nucleotides G47 and G48 chemically protected by neomycin and tobramycin upon binding.[1] In contrast, the C and D rings form only contacts with the phosphate backbone. Similarly, in our model, ring C makes direct contacts with the phosphate backbone at G46 and A68, whereas ring D makes contacts mediated through water molecules with the phosphate backbone at U45 and G70 and the Hoogsteen face (O6 and N7 atoms) of the base at the third chemically protected position G46 (Figure 6, top).

Conclusions

The prediction of bioactive conformations and binding orientations of ligands is not an easy task. Recent X-ray crystallographic studies of ligand−receptor complexes have shown that very similar ligands can adopt different binding modes[15] as in the case of the WIN compounds. Since ligand−receptor interactions may involve high-energy conformations of the individual molecules, bioactive conformations may differ significantly from the unbound conformation. Nevertheless, we show here that it is possible to describe accurately the binding properties of ligands, using a modeling strategy based on 3D-SAR and docking studies, even in the case of small sets of very flexible ligands. The advantage of this protocol is that the computationally complex docking protocol is performed with a greatly reduced, but relevant, conformational library as determined by 3D-SAR. The original conformational space generated for neomycin (maximum rmsd of 5.3 Å) and tobramycin (maximum rmsd of 4.0 Å) was reduced dramatically to rmsd of 2.3 and 2.9 Å for neomycin and tobramycin during docking. Docking and modeling is thus initiated with more likely ligand conformations and refined by an explicit representation of the receptor's geometry. In the final step, MD simulations were used to optimize the interaction with the receptor RNA.

Methods

Data. The data used in this paper were taken from Badger et al.[10] for the phenyloxazolines (the so-called "WIN" compounds) and from Zapp et al.[1] for the aminoglycosides. In the case of the aminoglycosides, since data were collected from experiments carried out at pH 7.9,[1] all amino groups of the 4,5-diglycosyl-substituted 2-deoxystreptamine (4,5-DOS) compounds were assumed to be protonated,[16] with the exception of the 3-amino group of the B ring.

Conformational Libraries. Conformational libraries of all compounds were generated by the Search-Compare module (version 2.3.5; Biosym/MSI Technologies, San Diego, CA) which systematically samples free torsion angles (each single bond of the ligand was rotated through 360° in increments of 120°, unless otherwise indicated). Conformers were clustered into subclasses based on the pairwise root-mean-square deviation (rmsd) determined by superimposing the three-dimensional structures of conformers. A conformational "set" of around 200 conformers for each compound was assembled by taking the lowest energy conformer of each subclass.

Conformational generation was initiated for the WIN compounds using conformations found in the crystal structures of their complexes.[10] Pairwise distances among subclasses of the conformational set were from 0.6 to 1.0 Å rmsd. Starting conformations for the aminoglycosides were those found after structure optimization by molecular mechanics calculations using the CFF93.1 set of force-field parameters[17] or the conformations found in the 3D-SAR study presented here. During conformational generation, a 60° increment of rotation around each single bond was used except for the single nonring connecting bonds which were incremented by 120°. The major pucker forms (C2′-endo, O4′-endo, C3′-endo) for the ribose and the two twist conformers for iodopyranose were also represented in the conformational evaluations. The pairwise

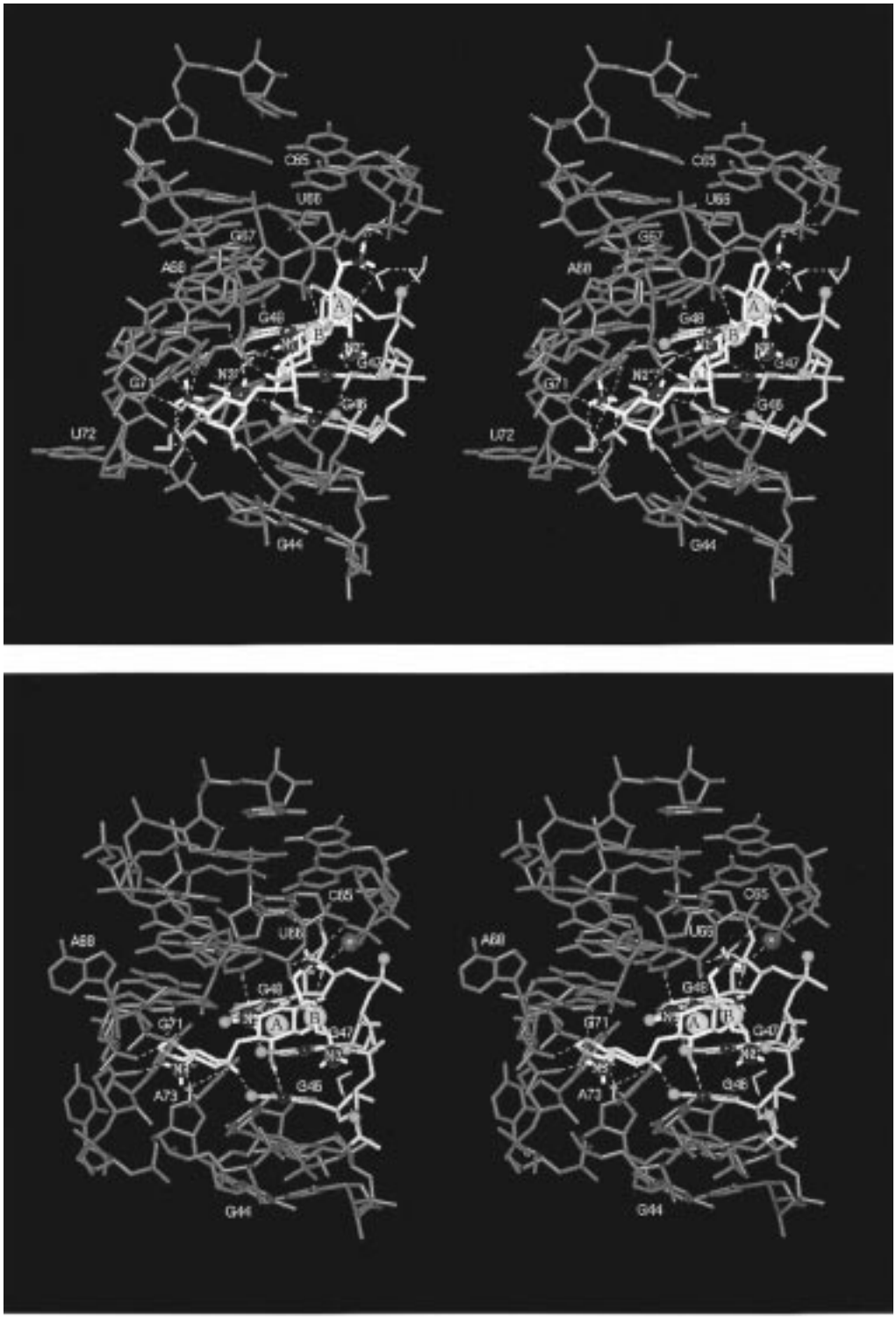


Figure 6. RBE RNA−aminoglycoside model: (top) stereoview of the RBE RNA−neomycin B complex (model 5), (bottom) stereoview of the RBE RNA−tobramycin complex (model 5). The neomycin B and tobramycin are bound in the major groove of the RNA. The nucleotides of the region strongly protected upon binding are indicated in gray (nucleotides G46, G47, G48).[1,2] The nitrogen and oxygen atoms of neomycin B and tobramycin forming molecular contacts are indicated in dark blue (N2′, N1, N2″, or N3″) and red, respectively. The rings corresponding to pharmacophoric centers are indicated by a sphere in gold placed at the center of mass of the ring. The nitrogen atoms corresponding to pharmacophoric centers are in CPK representation as are the O6, N7, and O1P atoms of G47 and G48. Water molecules or counterion (Na^+) involved in phosphate or base contacts between the antibiotic and the binding site are in cyan and purple, respectively. Hydrogen bonds are indicated by white dashed lines.

distances among these conformational subclasses were from 1.0 to 2.5 Å rmsd.

SAR Methodology. The 3D-SAR studies were carried out using the APEX-3D expert system[18] developed by Golender et al.[19] for the classification of compounds sharing a given biological activity (in this case, binding to a receptor). Correlations among atomic and ring centers of the ligand, their associated physical properties, and their activity were determined. Centers and their associated properties were then organized into coherent substructures, the pharmacophores, which represented structural properties common to the largest number of ligands of the activity class. Bioactive conforma-

tions were estimated by examining every conformer of each compound to identify which conformation(s) permitted the best overall three-dimensional superimposition of the pharmacophore. The congruency of the superimposition was estimated by the "match" value calculated from the proportion of centers which overlap. This value varies from 0 to 1 (with 1 as the best-possible fit). The statistical significance of the pharmacophores was estimated by two criteria: (1) the probability that a novel compound, possessing the pharmacophore, would belong to the same activity class (a Bayesian estimate) and (2) the reliability, a measure of a nonchance occurrence of the pharmacophore (binomial probability). By these definitions, when the entire training set belongs to a single class, the reliability is not applicable and the probability is equal for all pharmacophores.

Docking of Aminoglycosides. Aminoglycosides were docked to the Rev-binding RNA in a two-step procedure using starting conformations determined by the SAR analysis or by molecular mechanics calculations of the ligands and the conformation of the Rev-binding RNA previously determined.[3] In the first step, the starting conformations were oriented in either of two ways: one featured the A ring pointing toward the tetraloop and the C and D rings toward the open end of the RBE RNA; the other was oriented in the opposite direction. Different juxtapositions of the ligand with respect to the receptor site were constructed by translations and rotations of the ligand in space with respect to the receptor using increments of 0.1 Å and 0.5°, respectively. All structures were evaluated to find a locally optimal orientation by determination of the van der Waals and electrostatic components of the energy of interaction.[8] In the second step of the procedure, the energy of interaction was optimized by sampling the ligand conformational space within the binding site by rotation about all sterically unhindered single bonds. Final models were obtained after molecular dynamics refinement.

Computer Simulations of the RBE−Aminoglycoside Complexes. The RBE−aminoglycoside complexes constructed in the absence of explicit solvent were subjected to energy minimization until the maximum derivative was less than 5.0 kcal/mol Å. Counterions (Na^+) were placed at 6 Å from the phosphorus atoms along the O−P−O axis. The resulting complex was then solvated with a 6.0-Å thick water layer (approximately 2200 water molecules). To remove the van der Waals conflicts created by hydrating the RNA−aminoglycoside complexes, water molecules were subjected to energy minimization until the maximum derivative was less than 50 kcal/mol Å and then to a 1-ps MD simulation at 300 K. Intermediate evaluations of candidate RBE−aminoglycoside structures were conducted by optimizing their geometry until the maximum derivative was less than 1.0 kcal/mol Å. Models having high binding energies were retained for further analyses. Simulations were completed for these models by first heating the solvent and counterions from 10 to 300 K in steps of 50 K for 3 ps while keeping the RNA−ligand complex fixed in space. In the second heating cycle, the position of the RNA−ligand complex was constrained in space by (1) tethering the first base pair of the RNA stem/loop structure to its initial position with a force constant of 50 kcal/mol Å^2, (2) imposing distance constraints based on quadratic force constants of 20 and 10 kcal/mol Å^2 for Watson−Crick and non-Watson−Crick base pairings, respectively, (3) imposing an equilibrium distance of 2.9 Å between heavy atoms involved in hydrogen bonding, and (4) constraining the angles between the hydrogen donor and acceptor to 160° using quadratic restraints with a force constant of 2.0 kcal/mol Å^2. Finally, the system was submitted to 150 ps of MD simulation under distance constraints only. The nonbonded interactions were treated by the Cell Multipole Method.[20,21] All simulations were performed using the Discover package interfaced to the CFF93.1 force field in the constant-temperature, constant-volume ensemble.

Acknowledgment. The authors would like to thank Erich Vorpagel for his gracious collaboration and helpful comments. R. Cedergren is a Richard Ivey Fellow of the Evolutionary Biology Group of the Canadian Institute of Advanced Research. This work was supported by the Medical Research Council of Canada and Biochem Thérapeutique.

References

(1) Zapp, M. L.; Stern, S.; Green, M. Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell* **1993**, *74*, 969−978.
(2) Werstuck, G.; Zapp, M. L.; Green, M. R. A noncanonical base pair within the human immunodeficiency virus Rev-responsive element is involved in both Rev and small molecule recognition. *Chem. Biol.* **1996**, *3*, 129−137.
(3) Leclerc, F.; Cedergren, R.; Ellington, A. D. A three-dimensional model of the Rev-binding element of HIV-1 derived from analyses of aptamers. *Nature Struct. Biol.* **1994**, *1*, 293−300.
(4) Peterson, R. D.; Bartel, D. P.; Szostak, J. W.; Horvath, S. J.; Feigon, J. ^{1}H NMR studies of the high-affinity Rev binding site of the Rev responsive element of HIV-1 mRNA: base pairing in the core binding element. *Biochemistry* **1994**, *33*, 5357−5366.
(5) Battiste, J. L.; Tan, R.; Frankel, A. D.; Williamson, J. R. Assignment and modeling of the Rev response element RNA bound to a Rev peptide using ^{13}C-heteronuclear NMR. *J. Biomol. NMR* **1995**, *6*, 375−389.
(6) Leclerc, F.; Srinivasan, J.; Cedergren, R. Predicting RNA structures: The model of the RNA element binding *Rev* meets the NMR structure. *Folding Des.* **1997**, *2*, 141−147.
(7) Llorente, B.; Leclerc, F.; Cedergren, R. Using SAR and QSAR analysis to model the activity and structure of the Quinolone/DNA complex. *Bioorg. Med. Chem.* **1996**, 61−71.
(8) Srinivasan, J.; Leclerc, F.; Xu, W.; Ellington, A. D.; Cedergren, R. A docking and modelling strategy for peptide-RNA complexes: applications to BIV Tat-TAR and HIV-RBE. *Folding Des.* **1996**, 463−472.
(9) Diana, G. D.; McKinlay, M. A.; Otto, M. J.; Akullian, V.; Oglesby, C. [[(4,5-Dihydro-2-oxazolyl)phenoxy]alkyl]isoxazole inhibitors of picornavirus uncoating. *J. Med. Chem.* **1985**, 1906−1910.
(10) Badger, J.; Minor, I.; Kremer, M. J.; Oliveira, M. A.; Smith, T. J.; Griffith, J. P.; Guerin D. M. A.; Krishnaswamy, S.; Leo, M.; Rossmann, M. G.; McKinlay, M. A.; Diana, G. D.; Dutko, F. J.; Fancher, M.; Rueckert, R. R.; Heinz, B. A. Structural analysis of a series of antiviral agents complexed with human rhinovirus 14. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3304−3308.
(11) Klebe, G.; Abraham, U. On the prediction of binding properties of drug molecules by comparative molecular field analysis. *J. Med. Chem.* **1993**, *36*, 70−80.
(12) Robinson, H.; Wang, A. H. J. Neomycin, spermine and hexaamminecobalt(III) share common structural motifs in converting B- to A-DNA. *Nucleic Acid Res.* **1995**, *24*, 3700−3706.
(13) Wang, Y.; Hamasaki, K.; Rando, R. R. Specificity of aminoglycoside binding to RNA constructs derived from the 16S rRNA decoding region and the HIV-RRE activator region. *Biochemistry* **1997**, *36*, 768−779.
(14) Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. Structure of the A site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* **1996**, *274*, 1367−1371.
(15) Mattos, C.; Ringe, D. Multiple binding modes. In *3D-QSAR in Drug Design: Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM: Leiden, 1993; pp 226−254.
(16) Botto, R. E.; Coxon, B. Nitrogen-15 nuclear magnetic resonance spectroscopy of neomycin B and related aminoglycosides. *J. Am. Chem. Soc.* **1983**, *105*, 1021−1028.
(17) Potential Energy Functions Consortium PEFC 1994 release; Biosym/MSI Technologies, San Diego, CA.
(18) Golender, V. E.; Vorpagel, E. R. Computer-assisted pharmacophore identification. In *3D-QSAR in Drug Design: Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM: Leiden, 1993; pp 137−149.
(19) Golender, V. E.; Rozenblit, A. B. *Logical and Combinatorial Algorithms in Drug Design*; Research Studies Press: Letchworth, 1983.
(20) Ding, H.−Q.; Karasawa, N.; Goddard, W. A. Atomic level simulations on a million. particles: the cell multipole method for Coulomb and London nonbond interactions. *J. Chem. Phys.* **1992**, *97*, 4309−4315.
(21) Mathiowetz, A. M.; Jain, A.; Karasawa, N.; Goddard, W. A. Protein simulations using techniques suitable for very large systems: the cell multipole method for nonbond interactions and the Newton-Euler inverse mass operator method for internal coordinate dynamics. *Proteins* **1994**, *20*, 227−247.
(22) King, G.; Barford, R. A. Calculations of association free energies. Separation of electrostatic and hydrophobic contributions. In *Molecular Modeling: From Virtual Tools to Real Problems*; Kumosinski, T. F., Liebman, M. N., Eds.; ACS Symp. Series 576; American Chemical Society: Washington, DC, 1994; pp 173−184.

JM970372R