Journal of Computer-Aided Molecular Design **18:** 189–208, 2004.

Validation of an empirical RNA-ligand scoring function for fast flexible docking using RiboDock®

S. David Morley[1,2,*] & Mohammad Afshar[1,3]
[1]*RiboTargets, Granta Park, Abington, Cambridgeshire CB1 6GB, UK;* [2]*Present address: Enspiral Discovery Ltd, P.O. Box 976, Cambridge CB2 4ZT, UK;* [3]*Present address: Ariana Pharma, Pasteur Biotop, 28 Rue Dr. Roux, F-75724 Paris Cedex 15, France*

Received 29 September 2003; accepted in revised form 19 April 2004

Key words: bacterial A-site, cavity detection, enrichment factor, guanidine, high throughput docking, HIV-1 TAR, RiboDock® , RNA aptamer, structure-based drug design, virtual screening

Summary

We report the design and validation of a fast empirical function for scoring RNA-ligand interactions, and describe its implementation within RiboDock®, a virtual screening system for automated flexible docking. Building on well-known protein-ligand scoring function foundations, features were added to describe the interactions of common RNA-binding functional groups that were not handled adequately by conventional terms, to disfavour non-complementary polar contacts, and to control non-specific charged interactions. The results of validation experiments against known structures of RNA-ligand complexes compare favourably with previously reported methods. Binding modes were well predicted in most cases and good discrimination was achieved between native and non-native ligands for each binding site, and between native and non-native binding sites for each ligand. Further evidence of the ability of the method to identify true RNA binders is provided by compound selection ('enrichment factor') experiments based around a series of HIV-1 TAR RNA-binding ligands. Significant enrichment in true binders was achieved amongst high scoring docking hits, even when selection was from a library of structurally related, positively charged molecules. Coupled with a semi-automated cavity detection algorithm for identification of putative ligand binding sites, also described here, the method is suitable for the screening of very large databases of molecules against RNA and RNA-protein interfaces, such as those presented by the bacterial ribosome.

Abbreviations: ACD – Available Chemicals Directory; AMP – adenosine monophosphate; EF – enrichment factor; FMN – flavin mononucleotide; FRET – fluorescence resonance energy transfer; RMSD – root mean square deviation; TAR – trans-activation response element; Tat – transcriptional activator protein.

Introduction

Virtual screening against protein targets, coupled with the subsequent assay of relatively low numbers of prioritised compounds, is seen as a viable alternative to experimental High Throughput Screening whenever structural data is available [1, 2]. The increasing number of available 3D RNA structures gives new opportunities to applyvirtual screening techniques to the discovery of novel RNA binders, by docking large compound libraries to RNA targets [3, 4]. This is especially true with the advent of X-ray crystal structures of the bacterial ribosome in complex with a large number of antibiotics [5, 6]. However, this new perspective requires the development and validation of docking and scoring methods that can accommodate the specific features presented by RNA targets. For example, a key issue when screening against RNA, *in silico* as well as *in vitro*, is that of specificity. In particu-

*To whom correspondence should be addressed. E-mail: d.morley@enspiral-discovery.com

lar, given the negatively charged backbone of RNA, positively charged groups such as guanidine or amine-rich aminoglycosides tend to bind non-specifically. Careful control of electrostatic interactions therefore is critical in any computational approach to prevent indiscriminate selection of charged molecules.

In recent years, there have been several reports focusing on the development or use of *in silico* methods to build models for RNA-ligand complexes [7–10]. Although these studies provide an insight into the structural features that are important for RNA-ligand recognition, their use of computationally intensive simulation methods make them ill suited for screening of large compound databases. The most promising RNA virtual screening procedure reported to date [11, 12] combined a rigid DOCK [13] screen, as a pre-filter, with several rounds of flexible docking using the ICM docking package [14], and an RNA-trained empirical scoring function for the estimation of binding free energies. The system was used to screen the ACD [15] against HIV-1 TAR and identified several novel ligands subsequently confirmed to bind to TAR RNA. The structure of one TAR-ligand complex was solved by NMR [16]. However, the use of rigid ligand docking as a pre-filter is not ideal, and can lead to good binders being rejected early in the virtual screen, as less than 20% of the initial database is carried through to the flexible docking phase.

Our motivation is to design a fast scoring function to describe RNA-ligand interactions that will support automated flexible docking of very large databases of molecules (greater than one million compounds) against RNA targets and RNA-protein interfaces. To achieve this level of performance we built our scoring function on the empirical foundations set by Böhm [17, 18], employed to good effect in several protein-ligand scoring functions such as FlexX [19], Glide [20] and ChemScore [21], and added a number of extensions to the scoring function to encapsulate important RNA-ligand interaction motifs. The new scoring function is implemented in our proprietary high-throughput docking platform, RiboDock® [22].

The method is validated on a number of available RNA-ligand complexes and its ability to differentiate specific from non-specific interactions is measured through cross-docking experiments. The absolute prediction of binding free energies was given a lower priority in the optimisation of the scoring function than the more practical aspects of accurate pose prediction and compound ranking (selection of true binders). A rigorous statistical fitting of the scoring function

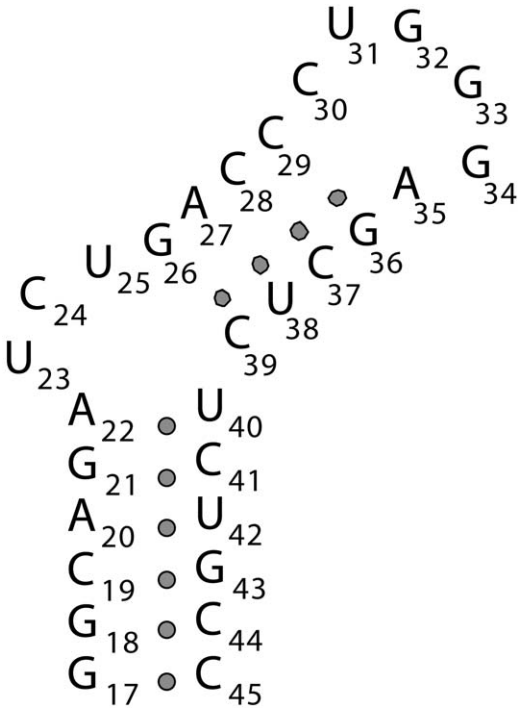


Figure 1. HIV-1 TAR secondary structure.

weights to experimental binding affinities is not possible with such a small training set of RNA-ligand complexes, and is further complicated by such issues as induced fit, in which the RNA undergoes a conformational change on binding to the ligand.

One of the most promising RNA drug targets [4, 12] is the trans-activation response element (TAR) RNA from HIV-1 (Figure 1). Blocking of the binding interaction between transcriptional activator protein (Tat) and TAR has been shown to prevent the synthesis of full-length transcripts and thus reduce the amount of circulating virus. As a further validation of the scoring function, we present the results of compound ranking ('enrichment factor') experiments based around an in-house library of HIV-1 TAR-binding ligands. Such experiments are designed to show the ability of the method to identify subsets of compound libraries that are enriched in active molecules.

Methods

RiboDock

RiboDock® consists of a semi-automated cavity detection algorithm, docking search engines, empirical

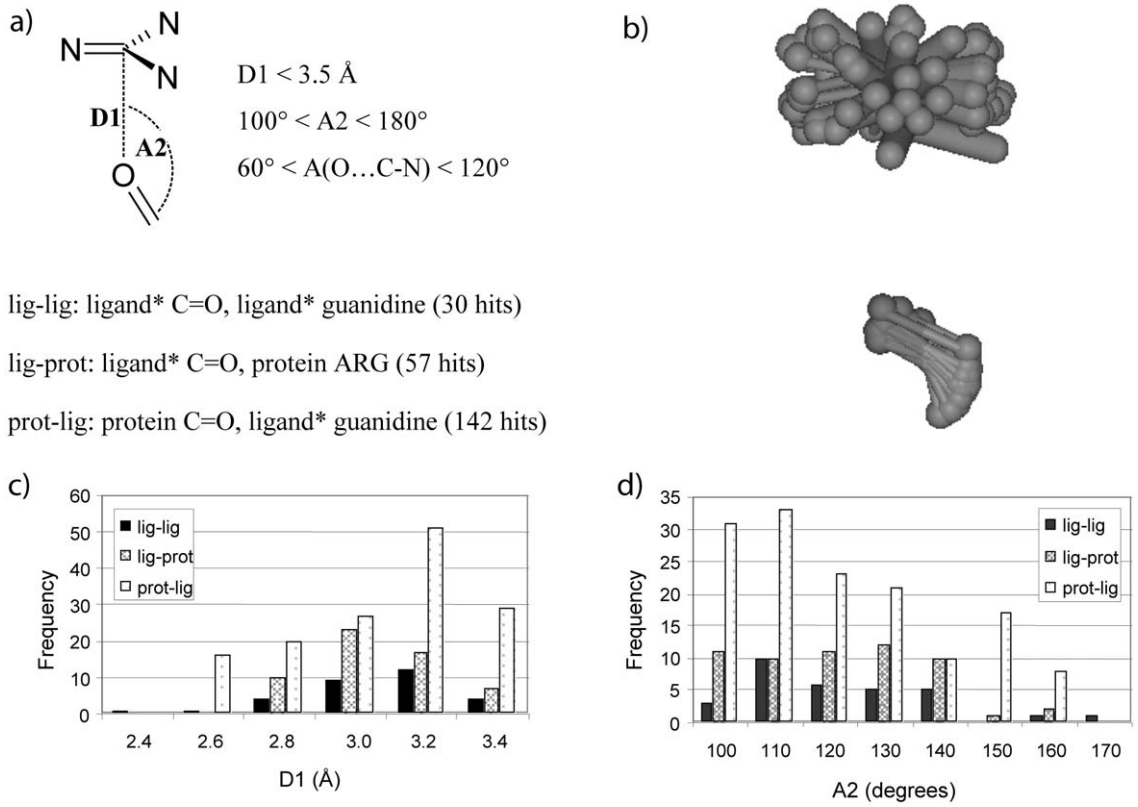
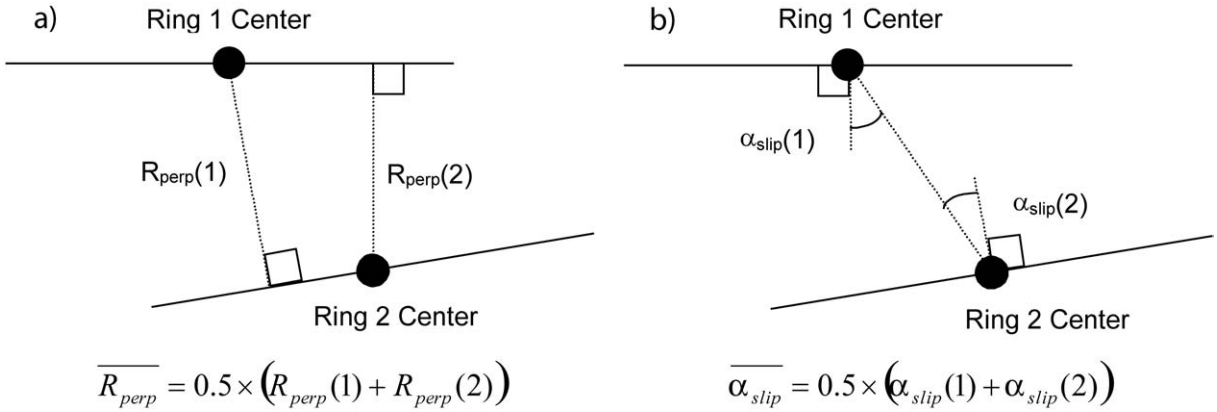


Figure 2. ReliBase analysis of out-of-plane guanidino-carbonyl interactions. (a) ReliBase query. (b) Overlay of 30 nucleic acid-ligand and intra-ligand interactions, aligned on the C=O..N plane. (c) Frequency distribution for D1. (d) Uncorrected frequency distribution for A2. If a geometric correction factor (sinA2) is applied, the optimum value of A2 is 135°. *Relibase considers nucleic acid chains as 'ligand' molecules, therefore the 'ligand-ligand' analysis implicitly includes nucleic acid-ligand interactions.



$$\overline{R_{perp}} = 0.5 \times \left(R_{perp}(1) + R_{perp}(2)\right)$$

$$\overline{\alpha_{slip}} = 0.5 \times \left(\alpha_{slip}(1) + \alpha_{slip}(2)\right)$$

Figure 3. Geometric parameters used to define aromatic interactions. (a) Mean perpendicular distance. (b) Mean slip angle.

scoring functions, and post-docking processing utilities. A strict object-oriented software design supports the rapid implementation of new search algorithms and scoring functions beyond those described here.

RNA target structures are read from industry standard macromolecular file formats such as Charmm PSF/CRD and Tripos MOL2. The scoring function relies purely on knowledge of the element type, hybridisation state, formal charge, connectivity and coordinates. Target coordinates are held rigid during the docking search, including 2′OH groups, the orientation of which must be optimised prior to docking. Formal charges are assigned to RNA atoms based on standard PDB atom and residue names. Phosphate groups carry a conventional formal charge of −0.5 on each terminal oxygen (O1P and O2P). We also found it desirable to set a formal charge of −0.5 on certain nucleotide atoms that are considered to carry high negative partial charges in force field calculations (A:N7, C:O2, G:N7, G:O6, U:O2 and U:O4). No formal positive charges are assigned to RNA.

Ligand structures are read from, and docking poses output to, MDL SD files [15], a format that enables large libraries to be automatically processed and docked with minimal manual intervention. In line with common practice, neutral ionisable centres (amines, guanidines, carboxylic acids etc.) are transformed into the most likely charged state at physiological pH. Formal positive charges are distributed evenly between all the polar hydrogens in any given group, except in amidinium, guanidinium and imidazole groups where the central nitrogen-bonded carbon also receives an equal portion of the charge. Formal negative charges are distributed evenly between all the terminal oxygen atoms in the group. This is similar to the approach taken by the Hammerhead docking program [23].

Empirical RNA-ligand scoring function

Three reusable scaling functions (Equations 1–3) were used to construct a set of interaction pseudo-energy functions (Equation 4) to describe RNA-ligand interactions. The following interaction types are accounted for: hydrogen bonds ($S_{H\text{-}bond}$), attractive lipophilic interactions (S_{lipo}), repulsive steric interactions (S_{rep}), positively charged carbon-acceptor interactions ($S_{posC\text{-}acc}$), aromatic stacking interactions (S_{arom}), donor-donor repulsion ($S_{don\text{-}don}$), acceptor-acceptor repulsion ($S_{acc\text{-}acc}$), and an estimate of the entropic cost of ligand binding. The overall RNA-ligand scoring function, S_{inter}, is the weighted sum of these component terms (Equation 5). Each function is described in detail below.

Table 1. RNA-ligand scoring function parameters.

Term[a]	X^b	X^b_0	ΔX^b_{min}	ΔX^b_{max}
$S_{H\text{-}bond}$	$R_{don\text{-}acc}$	1.9 Å	0.25 Å	0.6 Å
	α_{Don}	180°	30°	80°
	α_{Acc}	150°	30°	70°
$S_{posC\text{-}acc}$	$R_{posC\text{-}acc}$	$(\Sigma R_{vdW}) + 0.6$ Å	0.25 Å	0.6 Å
S_{arom}	R^c_{perp}	3.5 Å	0.25 Å	0.6 Å
	α^c_{Slip}	0°	20°	30°
S_{lipo}	R_{ij}	$(\Sigma R_{vdW}) + 0.6$ Å	0.25 Å	0.6 Å
S_{rep}	R_{ij}	$(\Sigma R_{vdW}) + 0.6$ Å	0.25 Å	0.6 Å
$S_{acc\text{-}acc}$	$R_{acc\text{-}acc}$	$(\Sigma R_{vdW}) + 0.6$ Å	0.25 Å	0.6 Å
$S_{don\text{-}don}$	$R_{don-don}$	$(\Sigma R_{vdW}) + 0.6$ Å	0.25 Å	0.6 Å [a]

[a]Equation 4 (see Appendix 2); [b]Equation 1; [c]Figure 3.

Reusable scaling functions

Function f_1 (Equation 1) describes the deviation of any geometric variable (distance, angle or plane) from an ideal reference value, and was taken directly from the Böhm score [18]. Reference values and tolerances for each occurrence of f_1 in the scoring function are summarised in Table 1. Many of the reference values for distance functions are relative to standard van der Waals radii. We also adopted function f_2 (Equation 2) from Hammerhead [23], which uses the formal charges described above to provide a convenient way of combining neutral and charged interactions within a single polar potential. f_2 ranges from 1.0 for neutral interactions to almost 2 for highly charged interactions. Whilst this favours charged interactions to some extent it is heavily damped compared to a standard Coulombic potential. Function f_3 (Equation 3) [18] is used to distinguish between convex (exposed) and concave (cavity) regions of the receptor surface, and increases the weighting given to buried interactions. f_3 typically ranges from 0.5 to 1.3. This function proved especially useful in controlling the indiscriminate docking of positively charged ligands to exposed phosphate groups.

Polar interactions

Hydrogen bond geometries are defined by four atoms: donor parent–donor hydrogen…acceptor–acceptor parent, e.g. N–H…O=C. The geometric score for a hydrogen bond is a function of the donor-acceptor bond length, the acute angle around the donor

hydrogen, and the acute angle around the acceptor atom. In cases where the acceptor is bonded to more than one parent, the acceptor angle is defined to the mean coordinates of all acceptor parents. Together with the overall weighting of −3.4, functions f_2 and f_3 combine to give a wide range of ideal H-bond interaction scores, from −1.7 for exposed neutral interactions to −8.3 for highly charged interactions in tight concave cavities.

Steric and repulsive interactions

Although originally conceived [17] to promote contact between apolar ligand and target surfaces, we (and FlexX [19]) calculate S_{lipo} and S_{rep} between all target-ligand atom pairs within close contact. S_{lipo} provides a simple description of the steric complementarity between the ligand and the target and can be considered as a gross simplification of an attractive vdW potential. S_{rep} is a short-range linear repulsive potential to prevent atomic overlap during docking, with a weight of 10 times that of the attractive potential. The weight is increased progressively during docking, starting from 10% of its final value, to promote efficient sampling.

$S_{acc\text{-}acc}$ and $S_{don\text{-}don}$ disfavour close acceptor-acceptor and donor-donor contacts, respectively. These terms are less commonly implemented in protein-ligand docking functions but were found to be vital in this work to help discriminate between correct and incorrect docking poses of otherwise similar score. The necessity of these terms is due presumably to the more polar character of RNA relative to most protein active sites, and hence the increased opportunity for non-complementary polar interactions. Donor-donor repulsions are evaluated between all polar hydrogen close contacts, both neutral and formally charged. However, in the acceptor-acceptor term, repulsive interactions between neutral hydrogen bond acceptors are excluded. Including these interactions was found to cause problems with aromatic stacking, as ligands with acceptor atoms in an aromatic ring were prevented from stacking correctly on RNA bases. Thus we only consider formally charged acceptors (those RNA and ligand atoms assigned formal charges as described above) in the repulsive term. A more general solution is to introduce angular dependencies to the functions that will specifically disfavour head-to-head contact. The entropic cost of ligand binding is estimated from the number of ligand rotatable bonds (N_{rot}) plus a constant that represents the loss of translational and rotational freedom of the ligand.

Novel RNA-ligand interactions

It became apparent early in initial testing that certain observed RNA-ligand interactions were not accounted for by the conventional scoring function terms described above. For example, in the arginine aptamer structure (1koc) [24], the binding cavity brings together an array of N7 acceptor atoms from G30, G31 and G35, as well as oxygens from G12, C13, G30, G31 and G35. This electronegative pocket accommodates the guanidinium moiety of the ligand. Recognition between RNA and ligand is mediated through conventional hydrogen bonds in the plane of the guanidinium, but also by specific close contacts between the central CZ ligand atom and carbonyl groups (O6) from G31 and G35 above and below the plane of the guanidinium.

A structural search of the Protein Data Bank (PDB) using ReliBase [25, 26] revealed that similar out-of-plane interactions between guanidinium and carbonyl groups are found in a number of nucleic acid-ligand structures (Figure 2). Examples include other arginine aptamers and complexes (1db6, 1arj, 1aju), and complexes of DNA with bleomycin (1g51) and netropsin (1dne). Whilst these examples are all NMR solution structures, the ReliBase search also identified the interaction in high resolution crystal structures of protein-ligand complexes, both between ligand carbonyl (or carboxyl) groups and arginine residues, and between ligand guanidinium groups and protein backbone carbonyl groups. The average CZ to O=C distance (D1 in Figure 2) is around 3.0 to 3.2 Å. The C=O..CZ angular distribution (A2) is very broad, ranging from 100° or less (the lower limit of the search) to around 150–160°, with a maximum in the distribution (after correcting the frequencies by the geometric factor, sinA2 [27]) at around 135°. Presumably electrostatic in nature, it is debatable whether this should be considered a stacking interaction or an interaction along the lone pair of the carbonyl oxygen. The new term, $S_{posC\text{-}acc}$, was introduced to represent this interaction. It is generalised between ligand carbon atoms given a formal positive charge in the autoprotonation step described above (amidinium, guanidinium and imidazole) and RNA acceptor atoms.

The aromatic term, S_{arom}, favours parallel pi-pi stacking between aromatic rings. As shown in Figure 3, the geometry is described by the average perpendicular distance from each ring centre to the other ring plane (ΔR_{perp}), and by the average slip angle between the rings ($\Delta\alpha_{slip}$). Each ring in a fused ring system is scored independently. Perpendicular stack-

ing geometries are not accounted for. Planar charged pi-systems on the ligand (specifically guanidinium and amidinium) also are included in this term, where the score is calculated to the central charged carbon. There are many examples of guanidinium groups stacking on pyrimidine and purine rings in nucleic acid-ligand and protein-ligand structures (ReliBase analysis, not presented here).

Ligand intramolecular score

The internal energy of each ligand conformation is described by the same combination of terms and weights as for the intermolecular scoring function, with the exclusion of the attractive steric potential. Note that no dihedral potential was used in the results reported here, rather each bond in the ligand is considered as either completely rigid or as a free rotor.

Optimisation of S_{inter}

S_{inter} is used directly during the docking search to rank alternative docking poses of the same ligand, and is also used as the primary means of ranking different ligands at the end of the calculation. The functional form of S_{inter} was optimised empirically based on the ability to reproduce the experimental binding modes of each ligand, and on the ability to discriminate between native and non-native ligands. The latter is demonstrated through the calculation of a full cross-docking matrix of scores, in which each ligand is docked into each target in the training set.

Optimisation began from a simpler base scoring function, S'_{inter} (Equation 6), consisting of only the hydrogen bonding (excluding the formal charge and neighbour density scaling functions) and steric terms. These functions, parameters and weights are essentially unchanged from FlexX [19] and were frozen during the optimisation process. The main experiments conducted to improve the performance of the scoring function were:

(i) presence/absence of f_2 (formal charge) and f_3 (neighbour density) scaling functions,

(ii) presence/absence of formal charges on certain RNA nucleotide atoms (A:N7, C:O2, G:N7, G:O6, U:O2 and U:O4),

(iii) presence/absence of novel attractive potentials ($S_{posC\text{-}acc}$ and S_{arom}),

(iv) presence/absence of novel repulsive potentials ($S_{don\text{-}don}$ and $S_{acc\text{-}acc}$)

The functional form and parameters for f_2 [23] and f_3 [18] were retained intact from their original implementations. The functional form and parameters for the novel potentials are closely analogous to those used for the standard hydrogen bonding potential and were not varied significantly. Weights for the novel potentials are nominal and were varied only within a narrow range of their final values. The final parameters and weights for S_{inter} are listed in Table 1. The performance of S_{inter} and S'_{inter} is compared in the Results section.

Docking site definition

As with most docking methods, the extent of the docking search is limited to a pre-defined volume ('docking site') around the binding site of interest. A two-probe sphere method was implemented to identify the most important cavities for docking (Appendix 1). The docking site consists of a collection of one or more cavities that together define the region the ligand is allowed to explore during docking. A distance restraint is added to the docking scoring function to keep the ligand within the docking site. The penalty function is calculated per atom, and is zero for atoms within the volume of the docking site, increasing linearly with distance from the nearest cavity for atoms outside of the docking site. Final docking poses that have a high residual penalty function are rejected. By tuning the cavity mapping parameters, the docking site can range from a single cavity around a known ligand binding site, to a large collection of cavities covering a broad extent of the RNA surface. The size of the docking site clearly has an impact on the convergence and efficiency of the search.

Docking search method

Each independent docking run starts by placing the ligand within one of the docking site cavities. If multiple cavities are present, an initial cavity is selected at random, with a probability proportional to its size, using the ratio between the cavity volume and the total volume of all cavities. The ligand centre of mass is placed at the geometric centre of the cavity, and the ligand principal axes aligned with the cavity principal axes. All ligand rotatable bonds are randomised. Flexible ring conformations are not sampled during docking.

In the work presented here, the ligand was subjected to Metropolis Monte Carlo sampling after initial placement. The maximum step sizes for rigid translation, rigid rotation and dihedral rotation were optimised to give a Metropolis acceptance rate of approximately 0.5. Monte Carlo sampling was used within a

Table 2. Simulated annealing search protocol.

Phase	Metropolis temperature		Trials	Maximum step sizes		
	T_{start}	T_{final}		Trans.	Rot.	Dihedral
1	1000 K	1000 K	$1250 \times (N_{rot} + 1)$	2.0 Å	30°	30°
2	1000 K	300 K	$1250 \times (N_{rot} + 1)$	2.0 Å	30°	30°
3	300 K	50 K	$1250 \times (N_{rot} + 1)$	0.1 Å	10°	10°
4	10 K	10 K	$1250 \times (N_{rot} + 1)$	0.1 Å	10°	10°

simulated annealing cooling schedule for locating low energy bound conformations (see Table 2 for an example protocol). A schedule was divided into a number of phases, and a fixed number of Monte Carlo trials were performed in each phase. The number of trials was proportional to the number of rotatable bonds in the ligand, such that more flexible molecules were sampled for longer. The sampling temperature was reduced periodically by a constant factor to achieve the desired final temperature (geometric cooling). In addition, the maximum Monte Carlo step sizes were halved if the acceptance rate fell below a threshold, typically 0.25. Each phase began from the best scoring docking pose identified during the previous phase.

The ability of the ligand to effectively sample multiple cavities during a single simulated annealing run is very dependent on the topology of the cavities. If they are well separated then inter-cavity transitions are unlikely. Effective cavity sampling is achieved through multiple runs starting from randomly selected cavities.

Performance optimisation

The rigidity of the target and the short-range nature of the interaction functions were exploited to optimise the scoring function for maximum throughput. Conventional pre-calculated grids were used during the early stages of docking for terms that were functions of interatomic distance only (e.g. S_{lipo}). Such grids were not practical for terms that also include angular dependencies (e.g. $S_{H\text{-}bond}$), and in these cases an indexing grid was used to provide an intermediate level of optimisation. The indexing grid stored lists of all relevant target atoms within a given radius of each grid point, thus minimising the number of interaction distances that were evaluated at each search step. The radius was chosen to just exceed the maximum range of the scoring function term. The S_{lipo} pre-calculated grid was switched to a more precise indexing grid for the final two simulated annealing phases to avoid any grid resolution artefacts. Non-polar hydrogens on the target and ligand were removed and implicitly incorporated into the parameters of their parent carbon atoms to further speed up the calculation of the scoring function.

Average individual run times (Intel PIII 1GHz) were around 5–10 s for typical lead-like molecules ($N_{rot} <= 6$). Run times for more flexible molecules (e.g. aminoglycosides) were much longer (>1 min) but such molecules are less likely to be of interest in small molecule drug discovery. Exhaustive docking was used in the work presented here, with many independent runs per ligand, to ensure that the validation of the scoring function was not limited by sampling effects. Efficient high-throughput protocols, more suitable for virtual screening of large libraries, will be described in a later paper.

Results

Validation experiments were conducted using the published NMR structures of 10 RNA/ligand complexes (Table 3). The data set includes two structures of the 16S ribosomal A-site with different ligands (gentamicin and paromomycin), two different aptamer structures of the same ligand (tobramycin), and a further six aptamers. The ligands cover a wide range of molecular weight, flexibility and formal charges. 3D models for RNA receptor and ligand were built from the first NMR structure in each PDB file, unless the authors had identified a different structure as the most representative of the ensemble. In order to correct some poor geometry and close contacts in the deposited NMR average structures, gentle energy minimization (SD followed by ABNR), in the presence of electrostatic and van der Waals potentials, was carried out on the RNA structures using the package Charmm (version 25 MSI). A harmonic distance restraint was applied to all atoms and released gradually

Table 3. RNA-ligand complexes used for scoring function validation.

PDB code[a]	Ligand	ΔG^{c}_{bind}	Ligand properties[d]			
			MW	N_{rot}	Pos. chg.	Neg. chg.
1am0	AMP	−28.5[e]	346	7	0	−1
1byj[b]	Gentamicin	−45.9[f]	454	13	5	0
1eht	Theophylline	−36.5[f]	180	0	0	0
1fmn	Flavin mononucleotide	−35.9[e]	455	11	0	−1
1koc	Arginine	−24.1[f]	175	6	2	−1
1kod	Citrulline	−23.8[f]	175	6	1	−1
1nem	Neomycin	−39.9[f]	620	22	6	0
1pbr[b]	Paromomycin	−38.2[f]	620	22	5	0
1tob	Tobramycin	−52.2[f]	472	16	5	0
2tob	Tobramycin	−51.2[f]	472	16	5	0

[a]Entries are RNA aptamer-ligand structures solved by NMR, unless otherwise stated.
[b]NMR structure of ligand complexed with 16S ribosomal A-site construct.
[c]Approximate experimental binding free energy (kJ mol^{-1}).
[d]MW = molecular weight, N_{rot} = number of rotatable bonds (including −OH, $-NH3^{+}$), pos. chg. = total formal positive charge, neg. chg. = total formal negative charge.
[e] Extracted from Ref. 42; [f]Extracted from Ref. 12.

as described previously [31]. 2′OH dihedrals were aligned manually where necessary to make optimum interactions with the native ligand.

Cavity mapping was conducted within a sphere of radius 20 Å, centred over the native ligand coordinates, using large and small probe radii of 4.0 Å and 1.75 Å, respectively. In most cases this generated a docking site large enough to allow the ligand to sample docking poses outside of, as well as inside, the recognised binding site.

Cross-docking experiments were also performed for the same complexes in which docking poses and scores were calculated for all pairwise ligand and RNA combinations. The assumption is that each aptamer will show a distinct preference for binding its own native ligand over non-native ligands, and likewise that each ligand will prefer to bind to its own aptamer. Hence the cross-docking experiments test the discrimination power of the scoring function and the docking method. 200 independent docking runs were performed for each complex. Generated poses were sorted into ascending order of intermolecular score (S_{inter}) prior to analysis.

Native dockings

The results of the native ligand docking experiments are summarised in Table 4. The score components and atomic RMS deviation (RMSD) of the ligand are presented for the best scoring (rank 1) docking pose, and, where necessary, for the best scoring 'correct' pose. A correct pose is defined as an atomic RMSD from the ligand NMR structure of less than 2 Å, the generally agreed metric for acceptable docking accuracy. A crude convergence measure is also provided, defined as the number of correct solutions in the top 10 ranked docking poses.

The rank 1 docking pose was correct in 5 out of the 10 test cases (1am0, 1byj, 1fmn, 1pbr, 2tob). With the exception of 1pbr these test cases also showed good convergence, in that over 50% of the top 10 docking poses had an RMSD of less than 2 Å. 1pbr has the largest, most flexible ligand (paromomycin) and there were no further correct solutions identified in the top ten docking poses. This may be indicative of the upper limit of the Monte Carlo search technique for flexible ligands. In a further three test cases, there was a correct pose within the top 10 (1koc, 1nem) or top 25 (1eht) poses. In the remaining two cases (1kod, 1tob), the closest pose to the NMR structure had an RMSD marginally higher than 2 Å, but these poses were still located within the top 20 scores.

The conformation of the arginine side chain in 1koc was reproduced well by the scoring function, with only minor variations in the position and orientation of the guanidinium group amongst the top scoring poses. The position of the exposed amino acid group was less well defined both in the experimental structure and in the docking, which explains the relatively high overall RMSD of 1.9–3.0 Å for the top scoring poses.

Table 4. Native ligand docking results.

PDB code	Convergence N<2A in top 10[a]	Docking pose			Weighted component scores[e]				
		Rank[b]	RMSD[c]	S^{d}_{inter}	Arom	H-bond	Ionic[f]	Lipo	Rep
1am0	5	1	1.8	−34.9	−5.1	−21.1	0.0	−21.5	+0.4
1byj	8	1	1.8	−40.0	0.0	−39.4	0.0	−23.2	+4.5
1eht	0	1	3.6	−26.0	−4.7	−8.5	0.0	−18.5	+0.4
		22	0.5	−22.0	−1.8	−6.2	0.0	−20.5	+1.1
1fmn	7	1	0.8	−34.9	−2.4	−22.0	+0.2	−28.8	+1.7
1koc	2	1	2.7	−37.3	0.0	−25.9	−6.7	−19.9	+3.8
		2	1.9	−36.2	0.0	−24.3	−8.7	−18.4	+3.8
1kod	0	1	3.2	−25.5	0.0	−26.9	+1.2	−14.1	+2.9
		17	2.1	−17.9	0.0	−18.5	0.0	−13.0	+2.2
1nem	1	1	8.7	−37.8	0.0	−54.9	+1.2	−30.7	+19.3
		8	1.9	−32.4	0.0	−39.3	+0.5	−25.8	+4.7
1pbr	1	1	1.8	−30.6	0.0	−32.4	0.0	−27.5	+2.0
1tob	0	1	9.8	−34.0	0.0	−37.7	+0.1	−20.7	+2.8
		15	2.1	−28.7	0.0	−37.7	+0.2	−14.7	+2.0
2tob	6	1	1.5	−34.2	0.0	−39.7	0.0	−25.1	+9.3

[a]Number of docking poses, in the top 10 poses sorted by S_{inter}, with an RMS deviation from the experimental ligand coordinates of less than 2 Å. [b]Rank of docking pose, out of 200 poses sorted by S_{inter}. [c]RMS deviation of docking pose from experimental ligand coordinates. [d]Equation 5 (see Appendix 2). [e]Weighted component scores from Equation 5. [f]Ionic = (posC-acc) + (acc-acc) + (don-don) scores, for tabulation purposes only.

The binding of AMP in the 1am0 NMR structure is dominated by stacking of the ligand adenine between 2 purines (A10 and G11). The interaction is stabilised by hydrogen bonds with the amino groups of A12 and G17, and a hetero purine base pairing of the ligand with G8, forming hydrogen bonds with G8:H21 and G8:N3. There are also several hydrogen bonds between the ligand ribose and RNA backbone. The docking was able to reproduce the bound conformation of the ligand adenine moiety very closely. The placement of the phosphate tail was less precise, consistent with the variation in the deposited NMR structures.

FMN intercalates G10 and G9 in the 1fmn NMR structure. The lower plateau of the binding site is formed by G10, U12 and A25 (base triple). The stacking interactions take place between the benzene ring of FMN and 6-membered rings of G10 and G9, and to a lesser extent between the last cycle of FMN and U12. The edge of FMN base pairs with A26, displaying 2 hydrogen bonds. The tail also makes hydrogen bonding contacts with bases and the RNA backbone. In this case the docking accurately predicted the position and orientation of the FMN ring system and the phosphate tail, reproducing all stacking and hydrogen bonding interactions.

Citruline is an uncharged molecule at physiological pH. Accordingly, the citruline aptamer binding cavity is less negatively charged than the otherwise similar arginine aptamer. The best scoring docking pose had the urea binding group broadly in the correct position, but the flexible amino acid tail deviated significantly from the NMR structure. One issue with this aptamer from a docking perspective is that the binding cavity is overly tight, with a number of marginal steric clashes in the deposited structure that inhibit accurate docking of the urea group. We believe this to be at least partially responsible for the relatively poor performance on this aptamer.

Tobramycin in complex with its aptamer shows many complementary interactions including vdW, hydrogen bonding and ionic contacts, consistent with its tight affinity. The best scoring docking pose reproduced the majority of these interactions, with a reasonable RMSD of 1.5 Å. The orientation of the more solvent-exposed ring C was slightly less well defined than ring A and ring B.

Cross-dockings

The results of the cross-docking experiments are summarised in Table 5. In 5 out of the 10 structures (1am0, 1byj, 1eht, 1fmn, 1koc), the native ligand was ranked first by intermolecular score. In four of these, the bind-

Table 5. Cross-docking results[a].

RNA	Ligand									
	1am0	1byj	1eht	1fmn	1koc	1kod	1nem	1pbr	1tob	2tob
1am0	−34.9	−19.3	−28.9	−26.8	−33.5	−23.1	−22.8	−14.6	−20.0	−20.3
1byj	−21.0	**−40.0**	−14.5	−15.5	−25.9	−23.0	−29.4	−27.3	−30.4	−30.0
1eht	−20.2	−13.4	**−26.0**	−20.6	−25.0	−21.1	−16.5	−12.9	−15.5	−15.4
1fmn	−24.4	−20.8	−20.3	**−34.9**	−33.6	−22.6	−19.1	−15.5	−21.1	−21.8
1koc	−24.7	−29.1	−13.7	−16.7	**−37.3**	−20.1	−25.0	−20.7	−27.0	−24.2
1kod	−19.1	−19.4	−10.5	−18.3	−31.7	**−25.5**	−17.5	−18.8	−19.1	−19.2
1nem	−18.1	−38.3	−13.3	−16.2	−36.3	−24.4	**−37.8**	−32.6	−34.9	−36.0
1pbr	−20.7	−37.0	−15.6	−17.8	−26.7	−21.0	−26.9	**−30.6**	−39.9	−39.1
1tob	−25.7	−37.1	−15.2	−17.4	−33.7	−25.4	−27.6	−33.9	**−34.0**	−38.5
2tob	−23.8	−39.3	−18.5	−18.5	−37.9	−23.0	−28.1	−31.0	−33.8	**−34.2**

[a]S_{inter} (Equation 5) for rank 1 docking pose (200 runs sorted by S_{inter}).

ing mode was correctly predicted also, as described in detail above. Only in the case of 1eht was the native ligand ranked first but with an incorrect binding geometry. If the score of the first correct pose (RMSD < 2 Å) were used in the ranking then theophylline would rank second behind arginine.

With 1tob the native ligand was nominally ranked third, but the top scoring ligand was the alternative tobramycin structure (extracted from 2tob). The difference in behaviour observed for docking of the two tobramycin ligand structures against 1tob can be attributed only to minor variations in valence angles, as ring conformations are identical and all rotatable bonds were freely sampled during docking. Very similar scores were achieved for both tobramycin structures when docked against most other targets in the test set.

In the remaining four cases, the native ligand was ranked second (1kod, 1nem) or third (1pbr, 2tob). Although paromomycin was ranked third against 1pbr, it was beaten only by two other aminoglycosides, gentamycin (from 1byj) and tobramycin (from 1tob and 2tob), both of which are known bacterial A-site binders [28]. In general, there was good separation between the aminoglycosides and the lower molecular weight ligands in the matrix.

A column-based analysis of Table 5 reveals a similar story. Five ligands scored most highly against their native target, and a further three were ranked second or third. This shows that, on this limited dataset at least, the method is able to distinguish between alternative binding sites for a particular ligand.

Comparison with simpler scoring function (S'_{inter})

Table 6 compares the docking performance of the optimised RNA-ligand scoring function (S_{inter}) with the original starting function (S'_{inter}) that does not have any of the novel terms. The percentage of test cases with 1st ranked poses predicted within a ligand RMSD of 2 Å fell from 50% with S_{inter} to 30% with S'_{inter}, with a majority of ligands (60%) giving gross pose prediction errors of greater than 3 Å.

There was a similar reduction in performance in the cross-docking experiments, in particular in the ability of S'_{inter} to identify the native target for a given ligand. To give an example, using S_{inter}, argininamide is the top scoring ligand against its native target (1koc) and ranks its native target second out of the set of target binding sites (data from Table 5). Using the simpler S'_{inter}, the ligand ranking of argininamide against 1koc fell from first to second, and more significantly, argininamide achieved a similar score against all the target binding sites (-21.0 to -26.2 kJ) such that the rank of its native target fell from second to sixth (data not shown). In other words, using the simplified scoring function, argininamide docks non-specifically to a variety of RNA structures and is unable to identify its true binding site.

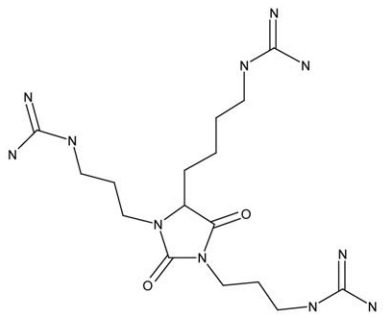
Enrichment factor experiments against TAR RNA

RBT203 (Figure 4) binds to TAR RNA with a K_i of 1.5 μM in an ADP1/TAR FRET binding assay [29, 30], where ADP1 is a Tat-derived peptide reporter. The solution structure of TAR complexed with RBT203 has been solved by NMR [31] and reveals a similar induced TAR conformation to that determined

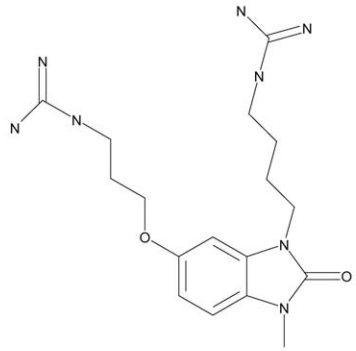
Table 6. Comparison of RNA-ligand scoring function performance.

Scoring function	Docking accuracy (RMSD of 1st ranked pose)			Discrimination (cross-docking)	
	N < 1 Å	N < 2 Å	N < 3 Å	No. of native ligands ranked first against native target[c]	No. of native RNA targets ranked first against native ligand[c]
S_{inter}[a]	1	5	6	6	5
S'_{inter}[b]	2	3	4	4	2

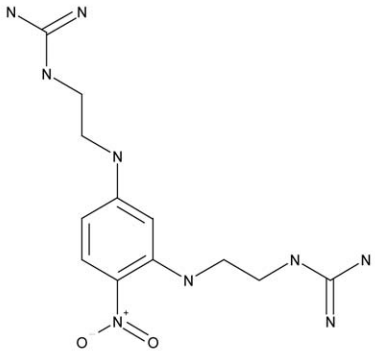
[a]Optimized RNA-ligand scoring function (Equation 5).
[b]Simplified scoring function without novel terms (Equation 6).
[c]Results for alternative tobramycin ligand structures (from 1tob and 2tob) are combined.



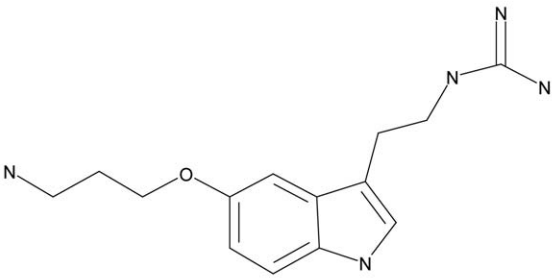
RBT88 K_i=3.6 μM



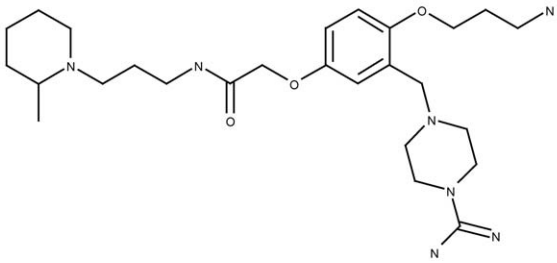
RBT197 K_i=2.7 μM



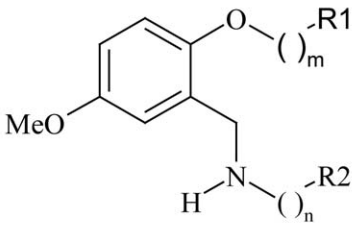
RBT255 K_i=3.9 μM



RBT318 K_i=0.90 μM



RBT338 K_i=4.7 μM



	m	n	R1	R2	K_i (μM)
RBT160	3	2	$NHC(NH)NH_2$	$NHC(NH)NH_2$	4.9
RBT161	3	3	$NHC(NH)NH_2$	$NHC(NH)NH_2$	2.7
RBT162	3	4	$NHC(NH)NH_2$	$NHC(NH)NH_2$	2.6
RBT202	2	3	$NHC(NH)NH_2$	$NHC(NH)NH_2$	3.4
RBT203	2	4	$NHC(NH)NH_2$	$NHC(NH)NH_2$	1.5
RBT392	3	4	NH_2	$NHC(NH)NH_2$	3.2

RBT203 series

Figure 4. Structures and activities of TAR RNA binders.

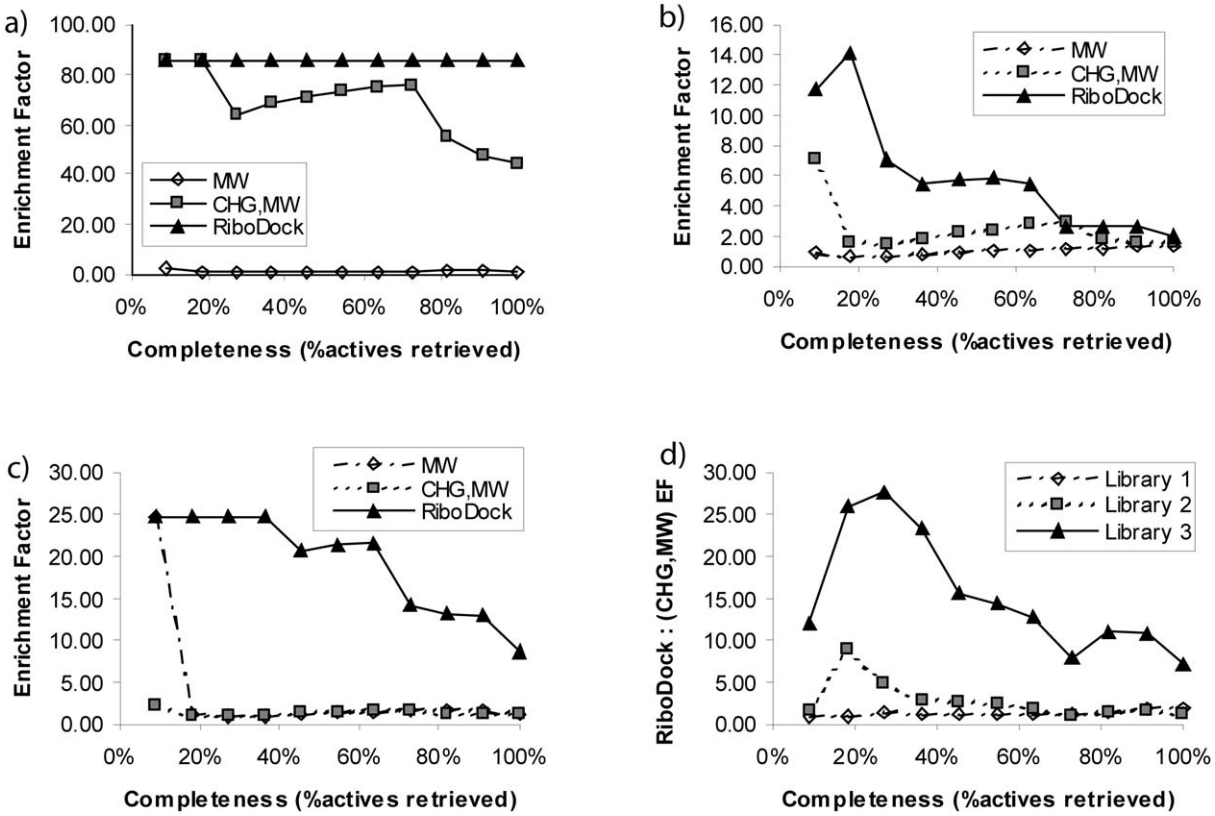


Figure 5. Results of compound ranking experiments for TAR binders 'spiked' into (a) Library 1, (b) Library 2, (c) Library 3. Compounds are ranked by three methods: MW, descending molecular weight; CHG, MW, descending formal charge followed by descending molecular weight; RiboDock, ascending S_{inter}. Enrichment factors (EF, hit rate of true actives in ranked library subset compared to random selection from entire library) are plotted against completeness (% retrieval of true actives) for ranked library subsets of increasing size. Note that maximum EF achievable for Library 1 is 85×, Library 2 is 35×, Library 3 is 25×. (d) Relative RiboDock EF for each library, normalised against EF obtained by ranking by CHG, MW.

for argininamide-bound TAR [32]. A further 10 compounds in this and related series bind to TAR with a K_i less than 5 μM (Figure 4). All have at least two positively charged groups, in most cases guanidinium. Some variability in the alkyl chain length of R1 and R2 in the RBT203 series is tolerated with minimal impact on K_i, but chain lengths shorter or longer than those presented here bind much more weakly. The combination of a flexible RNA, flexible ligands and a flat structure-activity relationship conspire to make this a challenging problem for virtual screening.

Enrichment factors

The 11 active compounds were seeded ('spiked') into three distinct libraries with different property profiles. Library 1 consisted of 928 randomly selected, commercially available 'lead-like' compounds extracted from supplier catalogs [33]. Library 2 consisted of 377 compounds extracted from the RiboTargets corporate compound database, many of which are inactive members of the RBT203 (and related) series. Library 3 consisted of 273 highly charged compounds extracted from the RiboTargets corporate database that are structurally unrelated to the active compounds. All of the compounds in Library 2 were confirmed as non-binders or weak binders ($K_i > 5$ μM) against TAR. The compounds in Library 1 and Library 3 were not assayed but were assumed to be non-binders for statistical purposes, the usual assumption in spiked library experiments.

Table 7 summarises the profiles of each library, in terms of average molecular weight and number of rotatable bonds, formal charge distribution, a diversity metric (number of clusters obtained using MOE [34], MACCS keys, 85% similarity threshold), and two similarity metrics (number of library compounds that share clusters with active compounds, and number

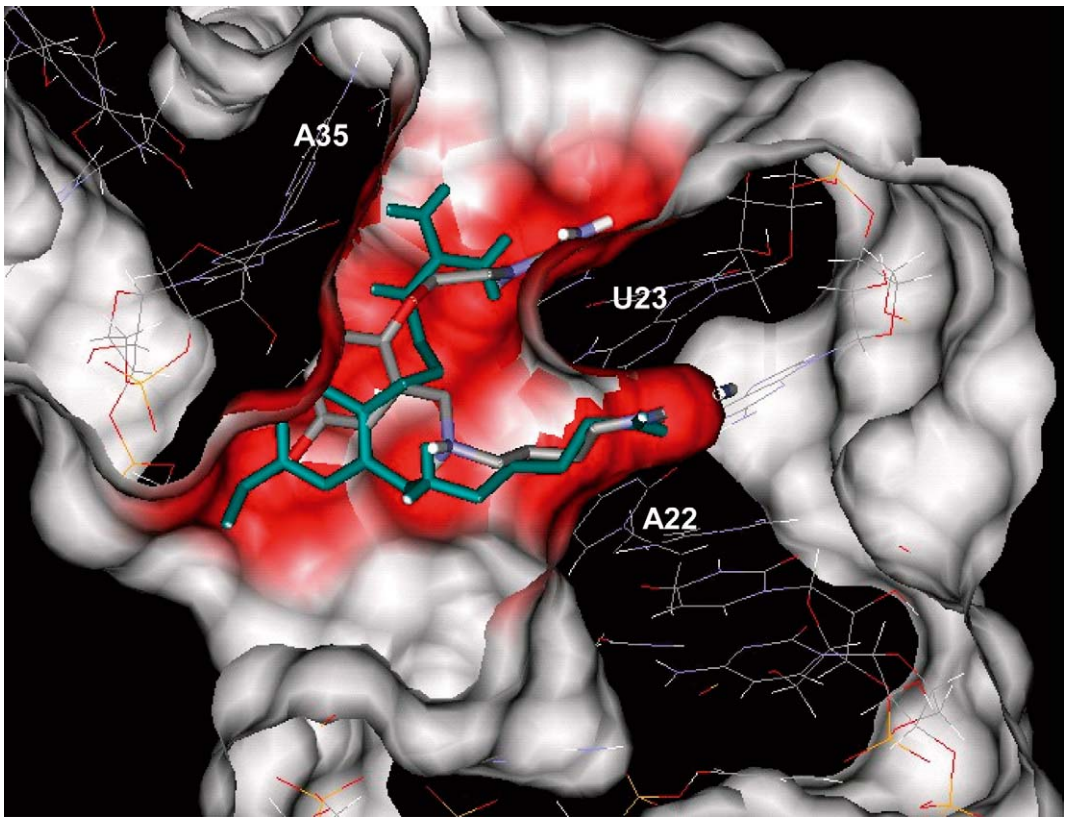


Figure 6. Comparison of RBT203 experimental binding mode (green) and docking pose with lowest RMSD (coloured by atom, $S_{inter} = -33.4$, rank = 3). Red regions of RNA surface indicate close contact with ligand.

of fingerprint bits in common between the library modal fingerprint and the modal fingerprint for the active compounds (MOE GPiDAPH3 graph 3-point pharmacophores [34])). It is clear that Library 1 is predominantly neutral, diverse (low average cluster size) and unrelated to the actives (no shared clusters). Library 2 is less diverse (higher average cluster size), has higher similarity with the actives (79 compounds in shared clusters, 421/432 modal fingerprint bits in common) and is largely positively charged, although to a lesser extent than the actives. Library 3 has a similar charge distribution to the actives but is otherwise unrelated.

The argininamide-bound TAR NMR structure [32] was used for the experiment, and subjected to minor refinement using CHARMm prior to docking [31]. All compounds were exhaustively docked (50 runs) to a large docking site covering the bulge region and the major and minor grooves of the lower and upper stems. Compounds were ranked according to their highest scoring docking pose (S_{inter}) and also, for comparison purposes, by descending molecular weight and by descending formal charge followed by descending molecular weight.

Figure 5 shows the results for each library and ranking method, plotted as enrichment factor (EF) against completeness (%C). For a given selected fraction (%F) of top ranked compounds, %C is defined as the percentage of all true actives retrieved in the fraction, and EF is defined as the observed hit rate of true actives in the fraction relative to the expected hit rate by random selection from the entire library (EF = %C/%F). Enrichment factors were calculated at all distinct completeness levels by selecting appropriately sized fractions required to retrieve each subsequent true active in the ranked order. A plot of EF vs %C gives a clearer visual separation between ranking methods than the alternative plot of %C vs %F.

With Library 1, there was total separation between the 11 true actives and the random compounds using the RiboDock score. This gives a nominal EF over random selection of 85×, the maximum possible with this library size. Unsurprisingly, given the fact that Library

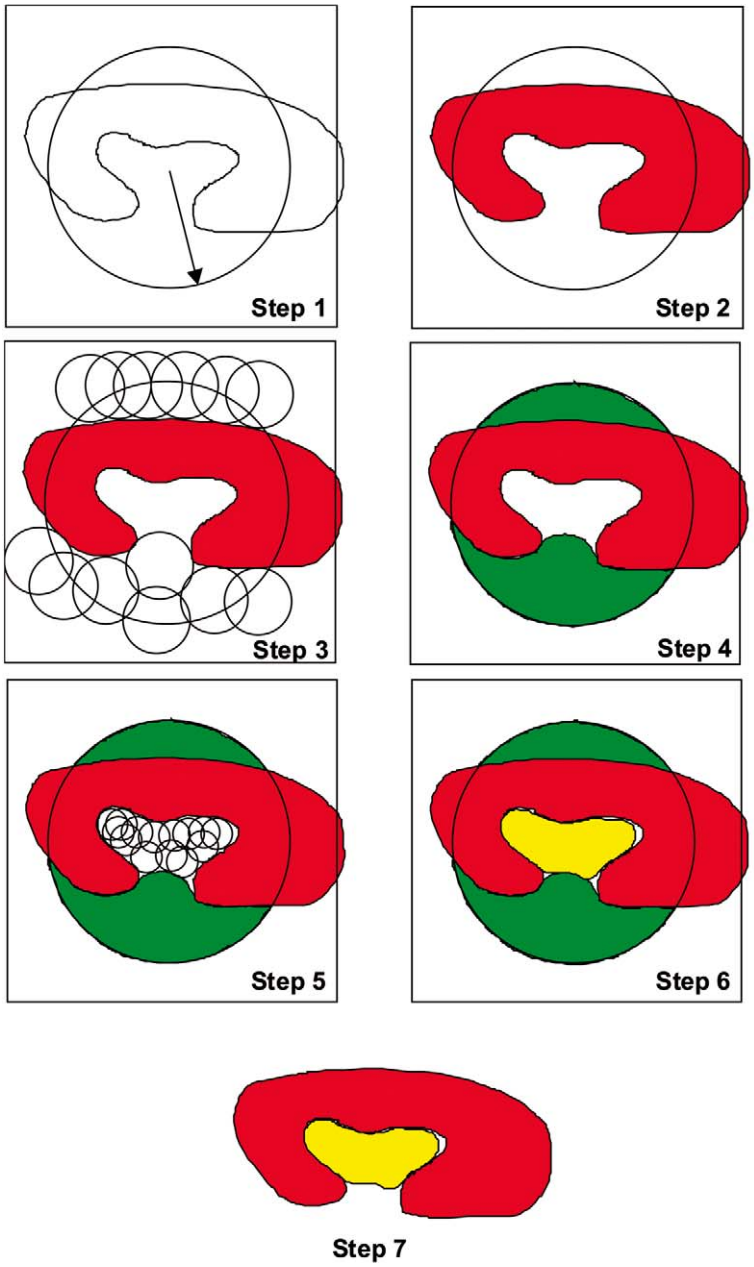


Figure 7. Cavity detection algorithm.

1 is predominantly neutral and that the true actives are all charged, it also possible to achieve similar enrichment by simple ranking by formal charge and molecular weight (Figure 5a), although RiboDock still outperforms at high completeness levels. Ranking by molecular weight alone does not show any enrichment. The interpretation of this result is that this class of positively charged TAR binding ligands can be identified readily by RiboDock although they are unlikely to be well represented in conventional 'lead-like' screening compound collections.

There was less clear-cut separation between actives and inactives with Library 2 (Table 7) but nevertheless significant enrichment was achieved using RiboDock. This type of experiment, in which the true actives are related to the background library, is not often repor-

Table 7. Library profiles for compound ranking experiments.

	Actives	Library 1	Library 2	Library 3
Source	In-house	Commercial	In-house	In-house
Description	TAR binders	Screening compounds	Inactive analogs	Highly charged, unrelated to actives
Size	11	928	377	273
Mean MW (s.d.)	363 (59)	365 (82)	349 (129)	339 (68)
Mean rot. bonds (s.d.)	11.1 (2.0)	4.3 (1.7)	7.9 (4.6)	4.7 (2.3)
#clusters[a]	6	866	116	189
#shared cpds[b]	–	0	79	0
Modal FP[c]	432	5698	2988	3452
#common bits[d]	–	317	421	283
% (CHG= −1)	–	3%	5%	–
% (CHG= 0)	–	84%	14%	2%
% (CHG= +1)	–	11%	15%	2%
% (CHG= +2)	27%	2%	41%	34%
% (CHG= +3)	64%	–	23%	46%
% (CHG= +4)	9%	–	3%	11%
% (CHG= +5)	–	–	–	2%
% (CHG= +6)	–	–	–	2%

[a]No. of clusters for each library, in combination with the active library (MOE, MACCS keys, 85% similarity threshold).
[b]No. of compounds in each library that belong to clusters containing actives.
[c]Total number of fingerprint bits set by each library (MOE, GPiDAPH3 graph 3-point pharmacophore, 7 bins, 8 features).
[d]Number of modal fingerprint bits in common with active library.

ted in the literature and lower enrichment factors are to be expected given the much closer structural similarity and property profiles between the active and inactive molecules. Two true actives were found in the top 5 virtual hits, giving an enrichment factor of 14×. Similarly, six true actives (55% retrieval) were found within the top 36 (EF=6×). Notably, the top four true actives were all members of the RBT203 series. The enrichment factor reduced to around 2–3× at higher completeness levels (>64% retrieval of actives), but given the inherent flexibility of TAR [16, 31] it is unrealistic to achieve full retrieval of actives by docking against a single RNA conformation. RiboDock again outperformed simple ranking by charge and molecular weight with this Library. The results highlight the applicability of the method to the optimisation of initial lead compounds, by providing a means of filtering a library of proposed, related structures to identify a smaller subset likely to be enriched in active molecules.

With Library 3, RiboDock is able to separate clearly the actives from the inactive, highly charged decoys and achieves maximum or near-maximum enrichment factors (20–25×) up to 60% completeness. In contrast, ranking by charge and molecular weight shows little enrichment (<2×) in this case.

In Figure 5d the RiboDock enrichment factor plots for the three libraries are superimposed, normalised by the enrichment obtained with charge and molecular weight ranking. This shows that the highest relative enrichment (>25×) is obtained with Library 3 (selection from random cationic compounds), and that healthy relative enrichment (5–10×) is still possible for selection from closely related analogs (Library 2). RiboDock outperforms the other ranking methods over all three libraries.

Prediction of RBT203 binding mode

In the ensemble of NMR structures [31] the R2 guanidinium of RBT203 is sandwiched between A22 and U23, in a similar position to argininamide. The position of R1 is less well defined although the guanidinium 'head' group may be able to form a similar interaction above the U23 ring. The aromatic scaffold sits in the major groove perpendicular to the aromatic rings of the RNA bases. An examination

of the predicted binding modes for the active compounds, including RBT203, showed that the highest scoring docking poses for each ligand reproduced some, but not all, of the features of the NMR structure of RBT203. A guanidinium group was positioned in the A22/U23 pocket in nearly all cases. The positions of the remaining basic group and the scaffold were more variable, and often only one of the two groups was close to the experimental position in any given pose.

To further probe the ability of the method to adequately predict the binding modes of these ligands, we conducted more detailed docking experiments for RBT203 itself. RBT203 was redocked into its 'native' RBT203-bound TAR structure [31] in place of the argininamide-bound structure used for the database screening experiments. Although similar, the major groove is slightly wider in the RBT203-bound TAR conformation and it was felt this might have an impact on the results. To overcome any Monte Carlo sampling concerns with such a flexible ligand, one guanidinium (chosen at random in each run) was anchored to the experimental binding position in the A22/U23 pocket, and a restricted docking protocol was used to sample the remaining dihedral angles. This eliminated the overall molecular translational and rotational degrees of freedom and gave much better convergence. The use of such restraints is justified when docking closely related molecules for which binding modes can be partially anticipated. Out of 100 runs ranked by S_{inter}, the third ranked pose reproduced the experimental binding mode reasonably well (Figure 6) with an RMSD of 2.3 Å from the most representative member of the ensemble of NMR structures.

Discussion

The issue of overfitting of S_{inter} is a legitimate concern, and it could be argued that we have simply added functions to satisfy the interactions of a relatively small training set. However, we believe our approach is justified for a number of reasons. Firstly, the novel terms that were added to the base scoring function are chemically reasonable and transferable. Secondly, the universe of published high quality structures of RNA-ligand complexes is very small in comparison to the protein-ligand field, and is insufficient to provide training and test sets of statistically significant size in any case. And thirdly, and perhaps most importantly, the scoring function described here (and variations thereof) has been in 'production' use within our company for virtual screening of large compound databases against RNA and protein targets for several years, with some success [35, 36]. In a young field, such as RNA drug discovery, an element of 'bootstrapping' is often necessary to improve the underlying technology. i.e. If improved scoring functions require a larger set of RNA-ligand structures for validation, then this, by definition, implies the discovery of new RNA-binding ligands, and virtual screening provides one route to this goal. Our approach is exemplified by the identification of novel binders to the bacterial ribosome A-site using RiboDock [36].

Another important issue is the extent to which RiboDock is biased towards positively charged molecules. It is well known that most RNA-binding ligands are cationic, but clearly the opposite (that all cationic molecules are RNA binders) is not true. Therefore, although the scoring function should favour positively charged molecules to some extent, this should not be at the expense of steric and polar complementarity. Our experience is that the combination of the formal charge scaling and neighbour density scaling functions work to good effect in limiting the impact of non-specific ligand-phosphate interactions and, in doing so, promote ligand burial. We believe that the results of the cross-docking experiments and, in particular, the compound ranking experiments using TAR RNA binders provide sufficient evidence that the scoring function terms are reasonably balanced. RiboDock performs almost as well at separating the TAR binders from a library of structurally unrelated cationic molecules as it does in the more trivial case of separation from a random library of neutral drug-like molecules. In the most stringent test, of separation from a library of inactive analogs, RiboDock still outperforms naïve ranking by charge and molecular weight.

A comparison of the results presented here with the state-of-the-art in protein-ligand docking accuracy is challenging, for several reasons: (i) the small size of the RNA set makes meaningful comparisons difficult, (ii) the majority of RNA-ligand structures are NMR solution structures, and not high resolution crystal structures, and (iii) docking to RNA may be intrinsically more difficult than proteins. For example, adaptive binding and induced fit are thought to be more significant in RNA-ligand complexes [37], particularly in the aptamers and small RNA structures such as TAR, where in many cases the bound ligand

stabilises a different RNA conformation to the free RNA.

RiboDock achieves a 50% success rate at predicting ligand binding modes to within 2 Å RMSD over the 10 RNA-ligand complexes. For comparison, GOLD successfully predicts 68% of protein-ligand binding modes to within 2 Å RMSD over all 305 entries in the CCDC/Astex test set [38]. This increases to 78% success rate over a more limited high-resolution subset of 92 complexes (crystallographic resolution better than 2 Å). It is worth noting that FlexX, another well-respected protein-ligand docking software package, performs less well (46.5% within 2 Å RMSD over a test set of 200 complexes [39]). Given the less precise nature of NMR structures, and that the docking calculations were performed against a single RNA conformation extracted from the NMR ensemble, we feel that the RiboDock results are acceptable.

A more extensive validation of RiboDock against RNA and protein targets is underway to show that the additional functionality of our scoring function does not detract from its performance in conventional protein-ligand docking. Preliminary results have already been described [40], using a revised scoring function that addresses some of the minor deficiencies noted here, such as the lack of explicit dihedral potential and the workaround for neutral acceptor-acceptor repulsion.

Conclusions

We have designed and validated a fast empirical scoring function suitable for docking large compound databases to RNA targets. The use of cross-docking experiments maximised the information content of a relatively small training set of RNA-ligand complexes. The inclusion of novel terms, beyond those in common usage in analogous protein-ligand scoring functions, improved the quality of predicted binding modes and the discrimination between true and false binders.

The enrichment factors (EF) obtained with selection experiments using a series of HIV-1 TAR ligands are very encouraging. RiboDock performs almost as well at separating the positively charged TAR binders from a library of structurally unrelated cationic molecules (EF = 20–25×) as it does in the more trivial case of separation from a random library of neutral drug-like molecules, in which total separation is achieved. In the most stringent test, of separation from a library of inactive analogs, RiboDock still outperforms naïve ranking by charge and molecular weight, and achieves an enrichment factor of 5–14×. These results show that RiboDock® is capable of identifying RNA-binding ligands from compound screening collections and has a potential role to play in directing the structure-based progression of initial lead compounds in a medicinal chemistry program, by prioritising virtual library subsets for synthesis.

Taking all of the validation results together, we conclude that the RiboDock® scoring function gives a reasonably well-balanced description of steric, polar and charged interactions in RNA-ligand complexes. The method is ideally suited to docking to RNA-protein interfaces, as found in many of the most promising drug targets in the bacterial ribosome, and has been used successfully to identify novel binders to the ribosomal A-site [36].

Acknowledgements

We would like to thank Jonathan Karn for initiating this project, Gabriele Varani and Eric Westhof for many useful discussions and suggestions, Ben Davis for assistance in the comparison of predicted and experimental binding modes for TAR ligands, Szilveszter Juhos for the preparation of Library 1, Philippe Vaglio for the collation of K_i data for the TAR enrichment factor experiments, Nico Baurin for help with MOE, and Rod Hubbard and Fareed Aboul-ela for their help in preparing this manuscript. We would also like to thank the reviewers of the initial draft of this manuscript for their constructive comments that led directly to the inclusion of the Library 3 experiment and the comparison with ranking by charge and molecular weight, and to the comparison with the simpler scoring function, S'_{inter}.

Appendix 1

Cavity detection algorithm

The following procedure is implemented for cavity detection (Figure 7):

1. A grid of resolution GRIDSTEP is placed over the cavity mapping region, encompassing a sphere of radius=RADIUS, centre=CENTER. Cavity mapping is restricted to this sphere. All cavities located will be wholly within this sphere. Any cavity that

would otherwise protrude beyond the cavity mapping sphere will be truncated at the periphery of the sphere. The mapping sphere can be made larger than the target molecule if it is desired to map the entire surface of the target, although this is not usual practice for production virtual screening.

2. Grid points within the volume occupied by the receptor are excluded (coloured red). The vdW radii of the receptor atoms are increased by VOL_INCR in this step.
3. Probes of radii LARGE_SPHERE are placed on each remaining unallocated grid point and checked for clashes with receptor excluded volume. To eliminate edge effects, the grid is extended beyond the cavity mapping region by the diameter of the large sphere (for this step only). This allows the large probe to be placed on grid points *outside* of the cavity mapping region, yet partially protrude *into* the cavity mapping region.
4. All grid points within the cavity mapping region that are accessible to the large probe are excluded (coloured green).
5. Probes of radii SMALL_SPHERE are placed on each remaining grid point and checked for clashes with receptor excluded volume (red) or large probe excluded volume (green).
6. All grid points that are accessible to the small probe are selected (yellow).
7. The final selection of cavity grid points is divided into distinct cavities (contiguous regions). In this example only one distinct cavity is found. The user can define filters for the minimum cavity volume considered appropriate for docking (MIN_VOLUME) and the maximum number of cavities to retain, in descending order of size (MAX_CAVITIES).

Typical parameter values for general use are: GRIDSTEP = 0.5 Å; RADIUS = 10–20 Å; VOL_INCR = 0.3 Å; LARGE_SPHERE = 3.75–5.00 Å; SMALL_SPHERE = 1.20–1.75 Å.

This approach is a faster grid-based implementation of previously published methods [41]. The effect of Step 3 (large sphere) is to eliminate all large cavities and smooth convex regions of the target surface. The radius of the small sphere used in step 5 represents a small ligand fragment or solvent molecule. Taken together, steps 3 to 6 identify regions of the grid corresponding to 'deep' cavities. i.e. those that are accessible to small spheres but not to larger ones.

Appendix 2

Equations

X = any geometric variable (distance, angle, plane angle)
X_0 = ideal value
$\Delta X = X - X_0$
ΔX_{Min} = tolerance on ideal value
ΔX_{Max} = deviation at which score is reduced to zero

$$f_1(\Delta X) = \begin{cases} 1 & \Delta X \le \Delta X_{Min} \\ 1 - \frac{\Delta X - \Delta X_{Min}}{\Delta X_{Max} - \Delta X_{Min}} & \Delta X_{Min} < \Delta X \le \Delta X_{Max} \\ 0 & \Delta X > \Delta X_{Max} \end{cases}$$

Equation 1. Geometric scaling function

c_i, c_j = formal charges on atoms i and j
n = damping factor (0.5)
$f_2\,(c_i, c_j) = (1 + n\,|c_i|)\left(1 + n\,\left|c_j\right|\right)$

Equation 2. Formal charge scaling function

$N(i)$ = no. of non-hydrogen RNA atoms within 5A radius sphere of RNA atom i
N_0 = normalisation factor (average local neighbour density) = 25
α = 0.5

$$f_3(i) = \left(\frac{N(i)}{N_0}\right)^{\alpha}$$

$f_3\,(i)$ is calculated for the relevant RNA atom involved in the interaction (donor or acceptor)

Equation 3. Local neighbour density scaling function

$$
\begin{aligned}
S_{H-bond} &= \sum_{don-acc} f_1(|\Delta R_{don-acc}|).\\
&\quad f_1(|\Delta\alpha_{don}|).f_1(|\Delta\alpha_{acc}|).\\
&\quad f_2(c_{don}, c_{acc}).f_3(i)\\
S_{posC-acc} &= \sum_{posC-acc} f_1(\left|\Delta R_{posC-acc}\right|).\\
&\quad f_2(c_{posC}, c_{acc}).f_3(i)\\
S_{arom} &= \sum_{arom} f_1(\left|\Delta\overline{R_{perp}}\right|).\\
&\quad f_1(\left|\Delta\overline{\alpha_{Slip}}\right|)\\
S_{lipo} &= \sum_{ij} n_i n_j f_1(\left|\Delta R_{ij}\right|)\\
S_{acc-acc} &= \sum_{acc-acc} f_1(\Delta R_{acc1-acc2}).\\
&\quad f_2(c_{acc1}, c_{acc2}).f_3(i)\\
S_{don-don} &= \sum_{don-don} f_1(\Delta R_{don1-don2}).\\
&\quad f_2(c_{don1}, c_{don2}).f_3(i)\\
S_{rep} &= \sum_{ij} n_i n_j f_{rep}(\Delta R_{ij})
\end{aligned}
$$

where

$$
f^{(}_{rep}\Delta R) = \begin{cases} \frac{\Delta R+\Delta R_{Max}}{\Delta R_{Min}-RX_{Max}} & \Delta R \le -\Delta R_{Max} \\ 0 & \Delta R > -\Delta R_{Max} \end{cases}
$$

$$
n_i, n_j = \begin{cases} 2 & \text{methylene } (CH_2) \text{ and methyl } (CH_3) \\ & \text{carbons with implicit hydrogens} \\ 1 & \text{all other atoms} \end{cases}
$$

Equation 4. RNA-ligand component interaction scoring functions

$$
S_{inter} = \begin{cases} -3.4S_{H-bond} - 1.0S_{posC-acc} \\ -1.6S_{arom} - 0.12S_{lipo} \\ +1.2S_{rep} + 1.0S_{acc\text{-}acc} \\ +1.0S_{don-don} + 1.0N_{rot} + 5.4 \end{cases}
$$

Equation 5. RNA-ligand intermolecular scoring function for docking and compound ranking

$$
S'_{inter} = \begin{cases} -3.4 \sum_{don-acc} f_1(|\Delta R_{don-acc}|). \\ f_1(|\Delta\alpha_{don}|).f_1(|\Delta\alpha_{acc}|) - 0.12S_{lipo} \\ +1.2S_{rep} + 1.0N_{rot} + 5.4 \end{cases}
$$

Equation 6. Simplified intermolecular scoring function, without novel terms, for comparison purposes

References

1. Shoichet, B.K., McGovern, S.L., Wei, B. and Irwin, J.J., Curr. Opin. Chem. Biol., 6 (2002) 439.
2. Schneider, G. and Böhm, H.-J., Drug Discov. Today, 7 (2002) 64–70.
3. Afshar, M., Prescott, C.D. and Varani, G., Curr. Opin. Biotechnol., 10 (1999) 59.
4. Drysdale, M., Lentzen, G., Matassova, N., Murchie, A., Aboul-Ela, F. and Afshar, M., Prog. Med. Chem., 39 (2002) 73.
5. Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly B.T. and Ramakrishnan, V., Nature, 407 (2000) 340.
6. Brodersen, D.E., Clemons, W.M. Jr., Carter, A.P., Morgan-Warren, R.J., Wimberly B.T. and Ramakrishnan V., Cell, 103 (2000) 1143.
7. Srinivasan, J., Leclerc, F., Xu, W., Ellington, A.D. and Cedergren R., Folding Design, 1 (1996) 463.
8. Leclerc, F. and Cedergren R., J. Med. Chem., 41 (1998) 175.
9. Leclerc, F. and Karplus, M., Theo. Chem. Acc., 101 (1999) 131.
10. Hermann, T. and Westhof, E., J. Med. Chem., 42 (1999) 1250.
11. Filikov, A.V., Mohan, V., Vickers, T.A., Griffey, R.H., Cook, P.D., Abagyan, R.A. and James, T.L, J. Comput.-Aided Mol. Design, 14 (2000) 593.
12. Lind, K.E., Du, Z., Fujinaga, K., Peterlin, B.M. and James, T.L., Chem. Biol., 9 (2002) 185.
13. Ewing, T.J.A. and Kuntz, I.D., J. Comput. Chem., 18 (1997) 1175.
14. Abagyan, R.A., Totrov, M.M. and Kuznetsov, D.N., J. Comput. Chem., 15 (1994) 488.
15. MDL Information Systems, San Leandro, CA.
16. Du, Z., Lind, K.E. and James, T.L., Chem. Biol., 9 (2002) 707.
17. Böhm, H.J., J. Comput.-Aided Mol. Design, 8 (1994) 243.
18. Böhm, H.J., J. Comput.-Aided Mol. Design, 12 (1998) 309.
19. Rarey, M., Wefing, S. and Lengauer, T., J. Comput.-Aided Mol. Design, 10 (1996) 41.
20. Schulz-Gasch, T. and Stahl, M., J. Mol. Model. [Online], 9 (2003) 47.
21. Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P., J. Comput.-Aided Mol. Design, 11 (1997) 425.
22. UK Registered Trade Mark E1308667.
23. Jain, A.N., J. Comput.-Aided Mol. Design, 10 (1996) 427.
24. Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E. and Famulok, M., Science, 272 (1996) 1343.
25. Hendlich, M., Acta Crystallogr., D54 (1998) 1178.
26. http://relibase.ccdc.cam.ac.uk/
27. We thank reviewer #2 for highlighting the sinA2 correction factor.
28. Wang, Y., Hamasaki, K. and Rando, R.R., Biochemistry, 36 (1997) 768.
29. Murchie, A.I.H., Davis, B., Isel, C., Afshar, M., Drysdale, M.J., Bower, J., Potter, A.J., Starkey, I.D., Swarbrick, T.M., Mirza, S., Prescott, C.D., Vaglio, P., Aboul-ela, F. and Karn, J., J. Mol. Biol., 336 (2004) 625.
30. Karn, J. and Prescott, C.D., U.S. Pat. No. 6,573,045 (2003).
31. Davis, B., Afshar, M., Varani, G., Murchie, A.I.H., Karn, J., Lentzen, G., Drysdale, M., Bower, J., Potter, A.J., Starkey, I.D., Swarbrick, T.M. and Aboul-ela, F., J. Mol. Biol., 336 (2004) 343.
32. Aboul-ela, F., Karn, J. and Varani, G., J. Mol. Biol., 253 (1995) 313.

33. Baurin, N., Baker, R., Richardson, C., Chen, I., Foloppe, N., Potter, A., Jordan, A., Roughley, S., Parratt, M., Greaney, P., Morley, D. and Hubbard, R.E., J. Chem. Inf. Comput. Sci., 44 (2004) 643.
34. Chemical Computing Group, http://www.chemcomp.com/
35. Knowles, D.J.C., Foloppe, N., Matassova, N.B. and Murchie, A.I.H., Curr. Opin. Pharmacol., 2 (2002) 501.
36. Foloppe, N., Chen, I., Davis, B., Hold A., Morley, D. and Howes, R., Bioorg. Med. Chem., 12 (2004) 935.
37. Patel, D.J., Suri, A.K., Jiang, F., Jiang, L., Fan, P., Kumar, R.A. and Nonin, S., J. Mol. Biol., 272 (1997) 645.
38. Nissink, J.W.M., Murray, C., Hartshorn, M., Verdonk, M.L., Cole, J.C. and Taylor, R., Proteins, 49 (2002) 457.
39. Kramer, B., Rarey, M. and Lengauer, T., Proteins, 37 (1999) 228.
40. Barril, X., Hubbard, R.E. and Morley, S.D., Mini Reviews in Medicinal Chemistry, accepted.
41. Afshar, M., Caves, L.S., Guimard, L., Hubbard, R.E., Calas, B., Grassy, G. and Haiech, J., J. Mol. Biol., 244 (1994) 554.
42. Hermann, T. and Patel, D.J., Science, 287 (2000) 820.