Rational Drug Design and High-Throughput Techniques for RNA Targets

Thomas Hermann*[1] and Eric Westhof[2]

[1]*Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, Box 557, New York, NY 10021, USA*

[2]*Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue René Descartes, F-67084 Strasbourg, France*

Abstract: RNA molecules are the only known molecules which possess the double property of being depository of genetic information, like DNA, and of displaying catalytic activities, like protein enzymes. RNA molecules intervene in all steps of gene expression and in many other biological activities. Like proteins, RNAs achieve those biological functions by adopting intricate three-dimensional folds and architectures. Further, as in protein sequences, RNA sequences contain signatures specific for three-dimensional motifs which participate in recognition and binding. In regulatory pathways, RNA molecules exist in equilibria between transient structures differentially stabilized by effectors such as proteins or cofactors. Therefore, RNA molecules display their potential as drug targets on different levels, namely in three-dimensional folds, in structural equilibria and in RNA-protein interfaces. Several examples will be described together with the already available techniques for combinatorial synthesis and high-throughput screening of potential drug and target RNA molecules.

INTRODUCTION

Among the functional components of cells, polypeptides, e.g. protein enzymes, transporters, receptors and ion channels account for the large majority of targets for therapeutic intervention. Effectors directed against ribonucleic acids (RNAs), however, can exhibit both high effectivity and specificity. The rapidly expanding knowledge of the key biological roles RNA molecules play especially in all steps of gene expression has fueled a growing interest in exploiting RNA as a drug target [1-7]. The exhaustive sequencing of expressed messenger RNA (mRNA, expressed sequence tags, ESTs) and complete genomes, in particular the human genome [8,9], allows the systematic exploration of potential RNA targets [1,5,10]. In contrast to proteins, which are end products, RNA molecules participate as intermediate carriers of genetic information (mRNA), as well as functional intermediates of the expression cascade which amplifies single genes into many copies of the encoded proteins. The lower copy number of target RNA molecules per cell as compared to proteins and the absence of cellular repair mechanisms renders RNA-directed therapeutics particularly powerful. Other than purely coding regions of mRNAs, functional RNA molecules and regulatory mRNA domains (for example, 5'- and 3'-untranslated regions, UTRs, of mRNAs, IRES of viral RNAs, catalytic RNAs) share in common with proteins a defined three-dimensional structure which is required for both molecular recognition and functionality. The folding of RNAs gives rise to intricate three-dimensional architectures for which structural details become increasingly available from X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR) [11-14]. Biophysical structure determination methods provide also insight into the molecular basis of specific recognition in complexes between RNA and small molecules [15-17]. These data will be useful for structure-based approaches towards the rational design of RNA-directed effector molecules. Recent progress in the large-scale synthesis and purification of RNA has further rendered possible the application of high-throughput methods for screening compound libraries for high-affinity RNA binders. Combinations of rational structure-based and high-throughput approaches are most promising for the development of RNA-directed therapeutics. This review outlines current efforts of large-scale target prediction, docking, rational design, combinatorial synthesis and high-throughput screening in the search for drugs directed at RNA molecules.

*Address correspondence to this author at the Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, Box 557, New York, NY 10021, USA; Phone: +212-639-7225; Fax: +212-717-3066; E-mail: THermann@sbnmr1.ski.mskcc.org

1386-2073/00 $19.00+.00

RNA STRUCTURE

RNA structure is described in a hierarchical fashion, distinguishing primary (i.e. the sequence of the four nucleotides A, U, C, G), secondary and tertiary structure [18-19]. Hydrogen bonding between complementary bases (A-U, G=C) in Watson-Crick pairs along with G and U in the wobble pair (GoU) define the secondary structure of an RNA molecule. Contiguous stacks of such canonical base pairs give rise to right-handed double helices which adopt the regular A conformation [20]. Distinct alterations of the regular RNA duplex structure are introduced by non-Watson-Crick pairs stabilized by non-canonical hydrogen bonding schemes [21]. Non-Watson-Crick pairs, along with base triples, play pivotal roles in the formation and stabilization of tertiary structure [22] and as recognition elements for small molecules [17], peptides and proteins [23]. Certain arrangements of non-canonical base pairs, triples and interactions involving the RNA backbone are found conserved among three-dimensional RNA architectures [13,14].

For clarity in the following discussion, we will define the terms of "RNA motif" and "RNA fold" (Fig. (**1**)). Both expressions describe the tertiary structure of RNA, albeit on different hierarchical levels. "Motif" will be used in order to address a conserved arrangement (see above) of local tertiary structure elements like non-Watson-Crick pairs and sugar-phosphate backbone reversals. In contrast, "fold" relates to the global tertiary structure or architecture, of an RNA molecule. The hallmark of RNA motifs is the conservation of their topology, whereas their constitutive nucleotides vary following sets of rules which are currently unraveled [21]. Thus, there is usually a number of different sequences (or combinations of nucleotides) that might be arranged to form a given RNA motif. Along with possible exchanges between canonical Watson-Crick base pairs (canonical covariations or compensatory base changes according to Watson-Crick rules), the covariations within non-Watson-Crick pairs ("full covariation" or "isostericity") compatible with the three-dimensional structure of a motif have to be appreciated (Fig. (**2**)) [21,24]. An important consequence for the search for RNA motifs in sequence databases is that it is not sufficient to deal with motif definitions purely based on sequence. Rather, tertiary structure information has to be included, for example, by specifying non-Watson-Crick base pairing rules, the size of secondary-structured regions (helices, single strands, loops) and their approximate mutual distances in the sequence (Fig. (**3**)).

Finally, any comprehensive description of RNA structure has to acknowledge dynamical flexibility in RNA which gives rise to equilibria between different conformers. The transient character of certain structural features in some RNA molecules may pose exceptional challenges for structure-based drug design approaches. Binding of proteins and low-molecular weight compounds might displace conformational equilibria by stabilizing selectively one form of an RNA target. Substrate binding leads to extensive adaptive processes resulting in structural ordering in RNA aptamers obtained from pools of random sequences by *in vitro* selection [17]. Examples for ligand-dependent conformational changes in natural RNAs are provided by two regulatory domains in the RNA genome of the human immunodeficiency virus (HIV) (Fig. (**1a**)), namely the Rev-response element (RRE) and the trans-activating region (TAR) which are recognition sites for the viral Rev (RRE) and Tat (TAR) proteins. In the free RRE, a non-Watson-Crick G•G pair (Fig. (**1a**)), which is a key determinant of the protein binding site, is in the *syn-anti* conformation [25,26], whereas one of the guanines is flipped in the Rev-bound RNA, giving rise to an *anti-anti* G•G pair (Fig. (**4**)) [27,28]. The same G•G pair is also involved in the molecular recognition of RRE by aminoglycoside antibiotics [29]. For the free TAR RNA, an equilibrium between major and minor conformations of the trinucleotide bulge (Fig. (**1a**)) has been observed in solution [30]. Binding of the Tat protein induces a specific conformational change in TAR which leads to coaxial stacking of the stems flanking the bulge [30-32]. Formation of the Tat-TAR complex is non-competitively inhibited by aminoglycosides [33], suggesting that the drugs might lock TAR in a conformer with low affinity for Tat [33,34].

These examples demonstrate that adaptive binding processes have to be taken into account by design strategies for RNA-directed drugs. Conformational flexibility might pose obstacles towards structure-based approaches, requiring knowledge of the three-dimensional structure not only of the free RNA target but also of one or more RNA-drug complexes. The TAR-aminoglycoside interaction shows, however, that conformational equilibria might be exploited by drugs that prevent the formation of functional conformers.

TARGET DEFINITION AND EXPLORATION

Cellular and viral RNA molecules are either mRNAs or non-coding transcripts (ncRNAs)

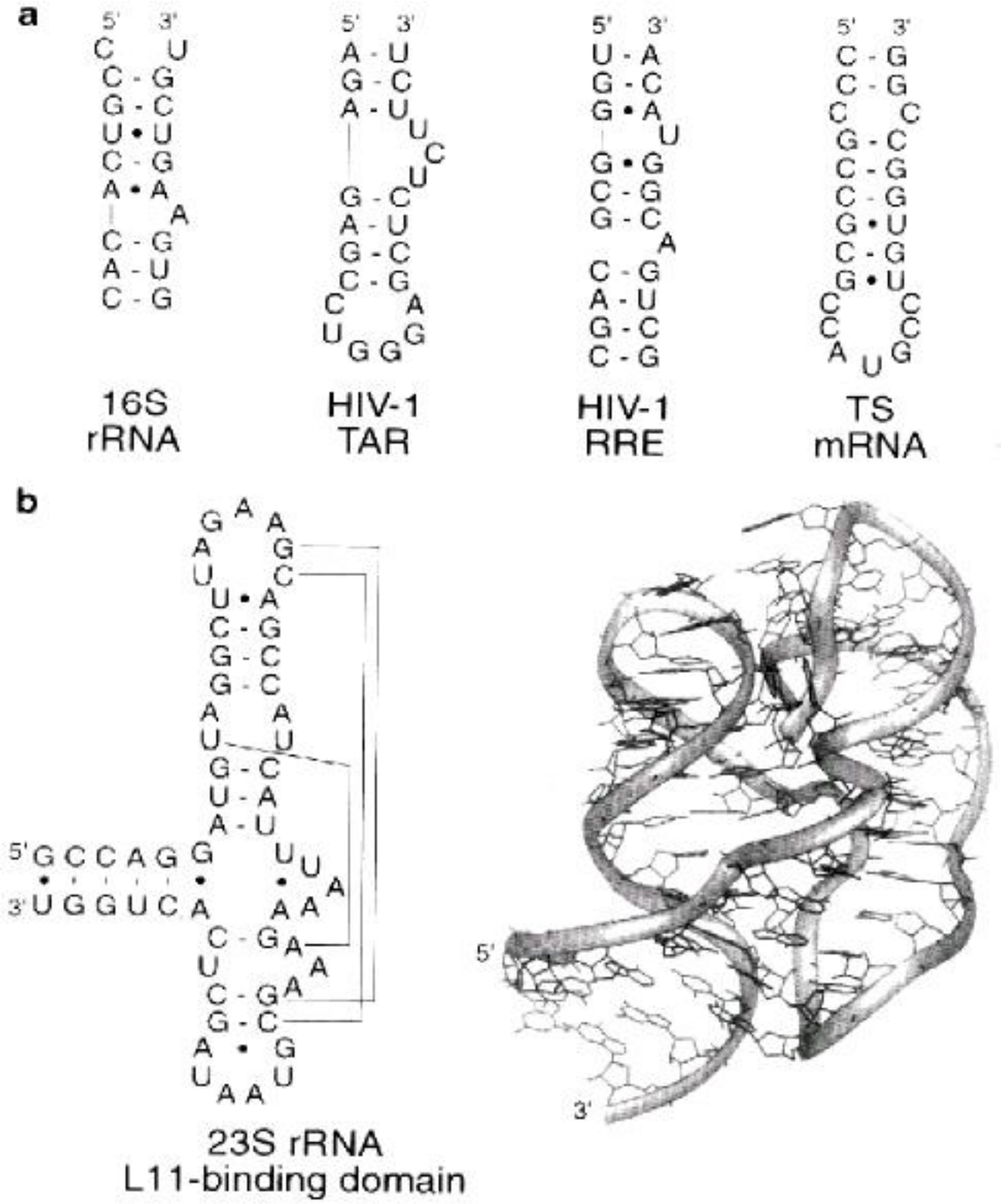

Fig. (1). RNA motifs (**a**) and folds (**b**). (**a**) Four representative examples of RNA motifs are shown, namely the A site of eubacterial 16S rRNA, two regulatory RNA elements of human immunodeficiency virus (HIV-1 trans-activating region, TAR; Rev-response element, RRE), and a regulatory motif in the 5' untranslated region of thymidylate synthase (TS) mRNA. (**b**) The binding site of the L11-protein in 23S rRNA forms an intricate fold stabilized by several tertiary contacts. A secondary structure representation is shown on the left. The right panel depicts the three-dimensional RNA fold determined by X-ray crystallography [57,58].

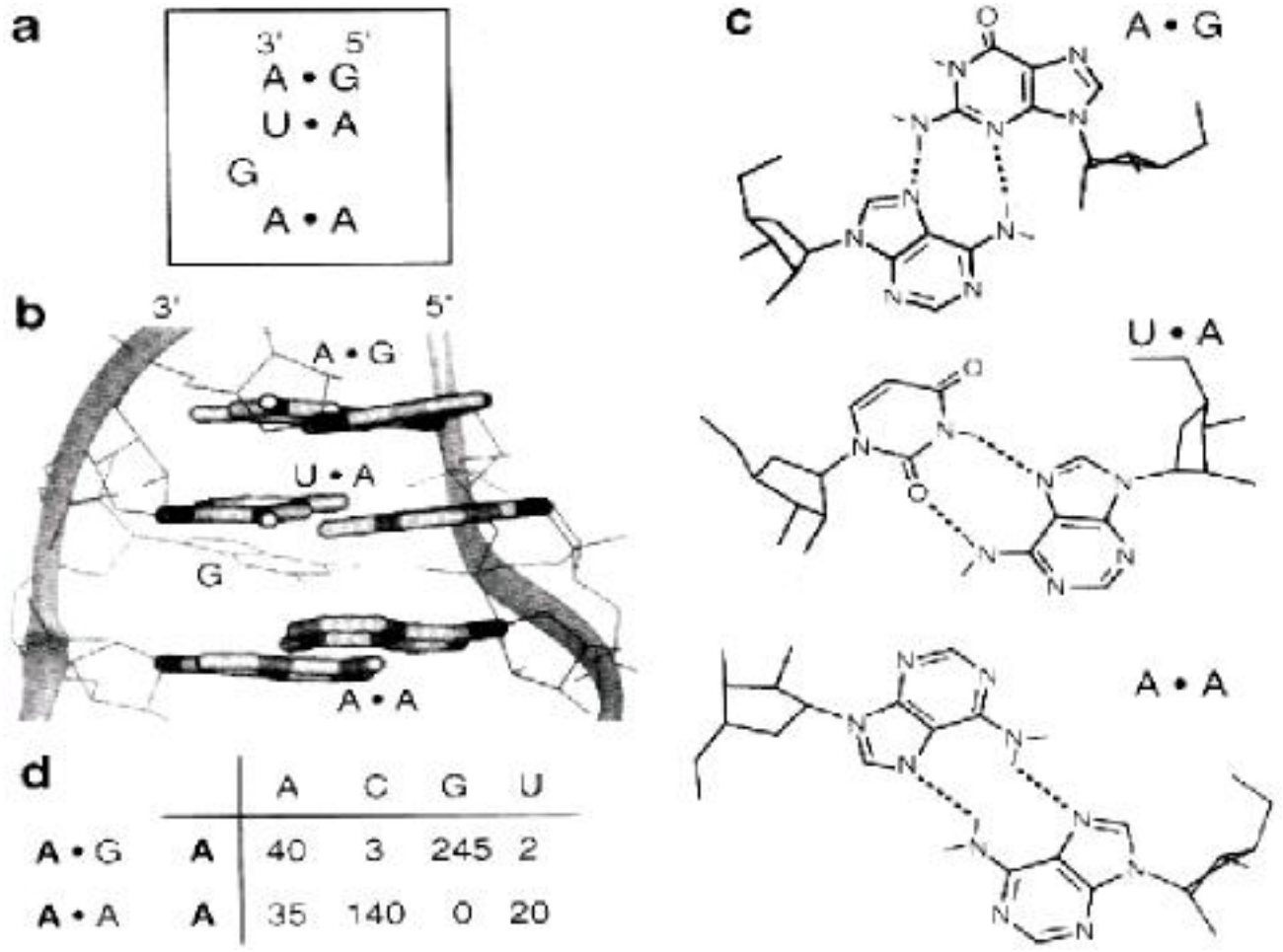

Fig. (2). An RNA motif common to ribosomal RNAs [24]. (**a**) A stack of three consecutive non-Watson-Crick base pairs defines the motif, comprising a sheared G•A, a *trans*-Hoogsteen U•A, and a parallel *trans* A•A. The guanine bulges out and forms a triple with the U of the *trans* Hoogsteen U•A base pair. (**b**) Three-dimensional structure of the loop E motif as determined by NMR analysis of the sarcin/ricin loop of 28S rRNA [127]. An X-ray structure is also available [128] (**c**) Hydrogen bonding schemes in the non-canonical base pairs. (**d**) Motif searches should take into account covariations of isosteric non-Watson-Crick pairs, once the average occurrences of non-canonical pairings in sequence datasets are known and can be compared to three-dimensional structures (here: the sheared G•A and *trans* A•A pairs loop E of eukaryotic 5S rRNAs).

[35,36]. Functional ncRNA molecules of the machinery involved in gene expression, such as tRNA, rRNA, the RNA components of the spliceosomal small nuclear ribonucleoproteins (snRNPs) and a number of catalytic RNAs (ribozymes) adopt complex three-dimensional folds stabilized by many tertiary interactions. The architectures of RNA folds involving common motifs have recently been reviewed [13,14]. For other RNAs, among them probably the majority of mRNAs and the non-coding small nucleolar RNAs (snoRNAs), extensive folding is not required for function or would even be detrimental in case of regions involved in coding or antisense hybridization. Despite lacking complex architectures, mRNAs may contain distinct and transient structural motifs, mostly in their 5'- and 3'-untranslated regions (UTRs), for the recognition by proteins involved in processing, transport, localization,translation, degradation and regulatory functions [37-39]. Recognition elements in mRNAs consist mostly of hairpin motifs with combinations of non-Watson-Crick base pairs, bulges and internal loops within the base-paired stem.

Both RNA folds and motifs provide potential targets for the specific binding of drug molecules which interfere with the function of the RNAs. Possible biologically relevant RNA targets for therapeutic intervention have been comprehensively reviewed recently [1] and only few examples will be discussed here. Outstanding among RNA-directed effectors are the aminoglycoside antibiotics which comprise compounds binding to a number of

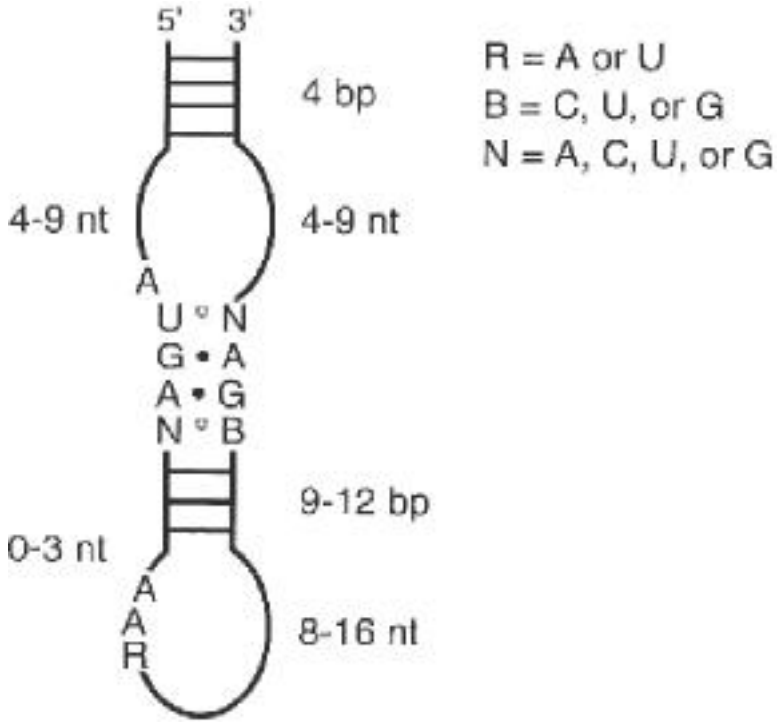

Fig. (3). RNA motifs are encoded in descriptors specifying secondary structure elements, non-Watson-Crick base pairs, their connections and distances. Descriptors, as the one shown here for the selenocysteine insertion sequence (SECIS) (adapted from [70,129]), are used for computational screens of sequence databases.

RNA folds and motifs (Fig. (**5**)) [40,41]. Among the folds containing recognition sites for aminoglycosides are eubacterial 16S rRNA (Fig. (**1a**)) [42], RNase P [43], self-splicing group I introns [44], hammerhead [45] and hepatitis delta virus (HDV) [46] ribozymes. The drugs elicit translational miscoding events in the case of the rRNA target and inhibit the catalytic function of the other RNAs. Besides binding to these RNA folds, aminoglycosides also specifically recognize some motifs such as an internal loop in the 5'-UTR of the human thymidylate synthase (TS) mRNA (Fig. (**1a**)) [47] and regulatory domains in the RNA genome of HIV (Fig. (**1a**)), namely TAR [48] and RRE [49]. Since each of these RNA motifs is a recognition site for a protein (TS itself for the TS 5'-UTR, Tat for TAR, and Rev for RRE), binding of aminoglycosides interferes with the formation of the RNA-protein complex (see above).

The limited number of complex cellular RNA fold topologies and their relatively large size allows straightforward phylogenetic identification in genome sequencing data. Automated search procedures incorporating conserved features of sequence and secondary structure of RNA folds have been used, for example, to scan genomic DNA sequences for tRNAs [50-53], group I introns [54] and hammerhead ribozymes [55]. Drug binding sites in RNA folds, however, are difficult to predict as they may be located at the packing interfaces between individual motifs. This is illustrated by the examples of thiostrepton recognition in the tertiary folded GTPase center of eubacterial 23S rRNA [56-58], recognition of the cyclic peptide antibiotic viomycin by RNA pseudoknots [59], and aminoglycoside binding at the three-way junction of the hammerheadribozyme [45,60,61]. In contrast, the aminoglycoside binding site in the 16S rRNA fold is located at the A site (Fig. (**1b**)) which, as an isolated motif, forms specific complexes with the antibiotics [62].

Similarly, motifs as recognition sites for small molecules in mRNAs are composed of distinct combinations of canonical base pairs, non-Watson-Crick pairs, loops, etc., which define signatures in computational screens for potential RNA targets in genome and EST databases (Fig. (**3**)) [63,64]. Consensus pattern comprising both sequence and secondary structure information have been used to scan databases for small RNA motifs such as iron-responsive elements (IREs) [63,65,66], mRNA instability elements [66,67], Tar-binding TAT and Rev-binding RRE [64,68], catalytic domains [64], aptamer motifs [64], snoRNAs [69] and selenocysteine insertion sequences (SECIS) (Fig. (**3**)) [70,71]. Most algorithms for the screening of

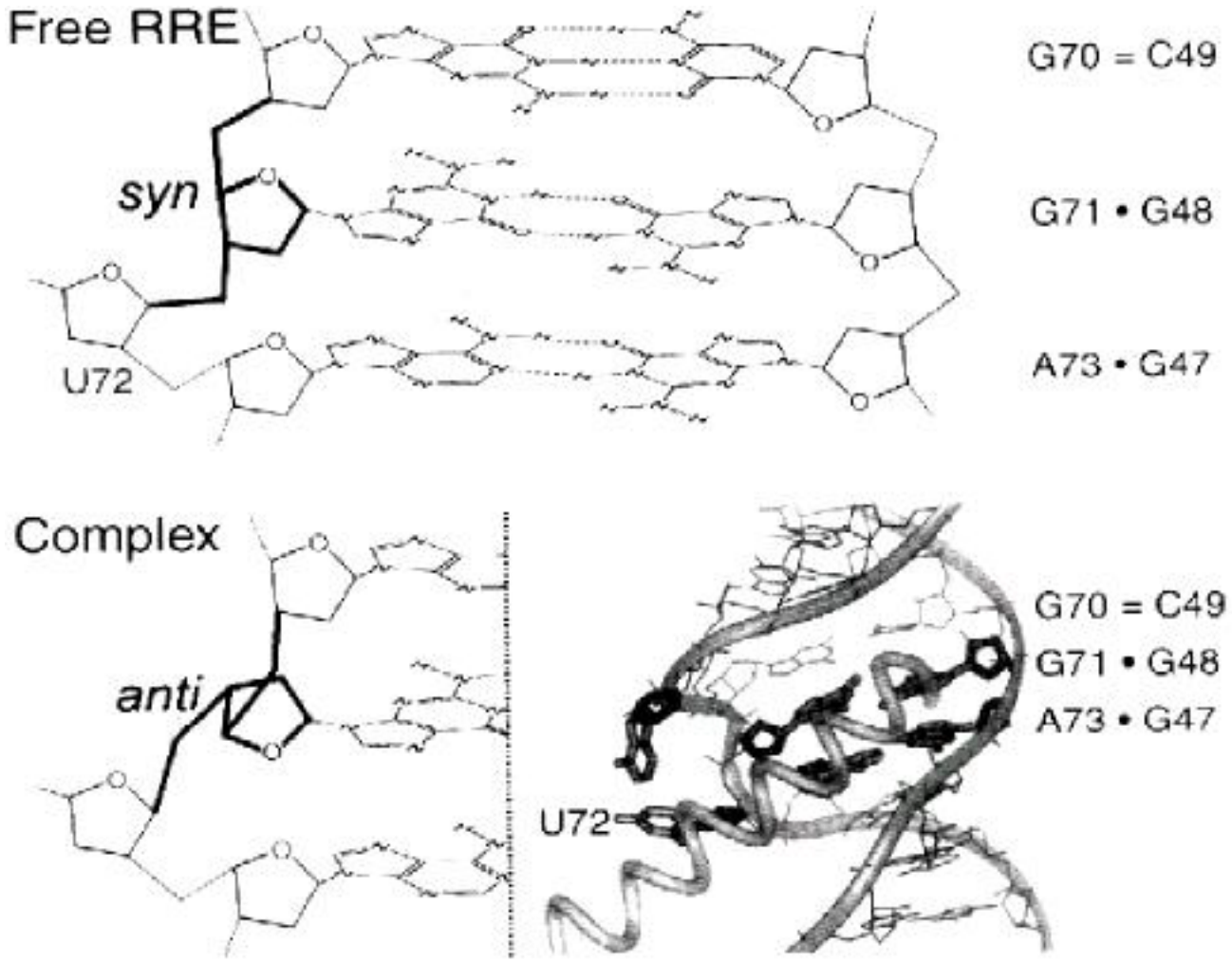

Fig. (4). RNAs may undergo adaptive conformational changes induced by ligand binding. In the HIV-1 RRE RNA, a syn to anti transition of a guanine involved in a G•G base pair (top and bottom left) occurs upon binding of a Rev peptide [25,26]. The peptide binds as an α-helix (bottom right) recognizing a stack of non-Watson-Crick pairs (bold sticks) within the deep groove of the RRE hairpin (see also Fig. (**1a**)) [28].

RNA motifs use probabilistic sequence and secondary structure consensus profiles, sometimes referred to as "covariance models" [51], based on stochastic context-free grammars [52,53]. In order to fully exploit the principle of covariation in the search for RNA motifs, however, the concept of isostericity between non-Watson-Crick base pairs has to be included [21]. Such an approach has been successful in identifying a common motif ("loop E") (Fig. (**2**)) in multi-helix loops of 16S and 23S rRNAs [24].

Whereas many motif searches have been conducted on complete genome sequence data, focused scans are now facilitated by the availability of specialized collections containing mRNA UTRs [72] and non-coding mRNA-like RNAs [36]. In addition to finding RNA motifs, database screening can also reveal whole new ncRNAs as potential targets. Analysis of genome gaps between open reading frames (ORFs) and transcripts following on predicted RNA polymerase III promoters has led to identification of new ncRNAs in yeast [73].

The search for potential targets among RNA motifs may benefit from pharmacogenomics approaches which elucidate the influence of genetic polymorphisms on drug disposition and effects [74,75]. The most common type of genetic variation is due to single-nucleotide polymorphisms (SNPs) which have been demonstrated being able to cause different secondary structure arrangements in two human mRNAs [76]. Based on these findings, it has been suggested that targeting of mRNA structural motifs may be applied to diseases for which allele-specific mRNA down-regulation would provide a therapeutic benefit [76].

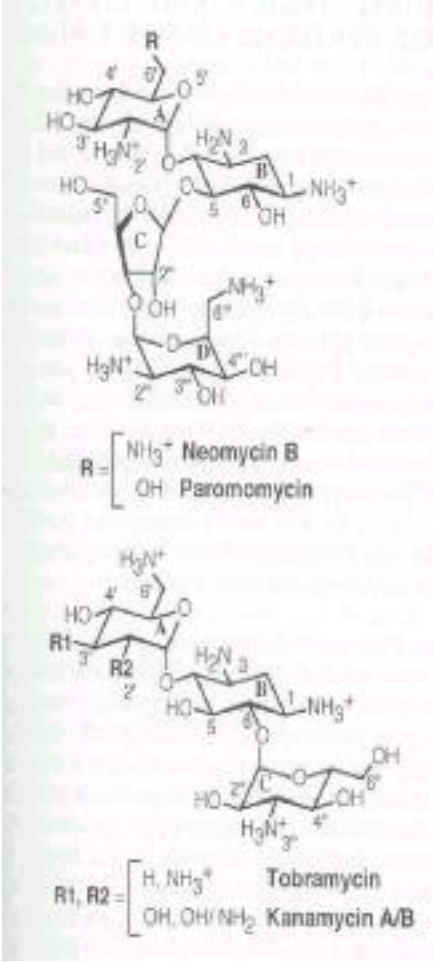

Fig. (5). Structures of natural aminoglycoside antibiotics sharing the neamine core (rings A and B). The protonation states of amino groups, as determined by NMR [130], are shown at neutral pH. All amino groups but position 3 are linked to methylene carbons (6', 6"') or vicinal to electron-drawing OH substituents and are, thus, protonated [99,100].

SCREENING TECHNIQUES

In the search for compounds which interfere with the biological functions of potential therapeutic RNA targets, screening assays have been developed which are based on techniques similar to those used for protein targets, namely filter binding, scintillation, fluorescence, mass spectrometry and surface plasmon resonance. Whereas most of the assays have been tested on a small scale in order to demonstrate the feasibility of the method, reports on screening RNA targets under realistic conditions are scarce. The first published high-throughput screen (HTS) on an RNA target [77] has been conducted to identify small molecule inhibitors of self-splicing group I introns which are found in biologically relevant genes of microorganisms such as *Pneumocystis carinii* but not in human. Scintillation counting of radioactively labelled RNA reaction products immobilized on nitrocellulose membranes was implemented in an assay for robotic screening of ~150,000 compounds on 96-well plates within three months [77,78]. In addition to already known self-splicing inhibitors such as the aminoglycoside antibiotics, more than 1000 small organic molecules were identified to inhibit the *P.carinii* group I introns. As a simplified screen, a scintillation proximity assay (SPA) has been developed to quantify radioactively labelled and biotinylated RNA products immobilized on streptavidin-coated beads which contained an embedded scintillant [78]. SPA assays may produce a considerable number of false positive hits, as non-native RNA structures, immobilized on the beads, may bind ligands in a non-specific manner.

The problematic radioactive labels might be substituted by fluorescence markers on the RNA which has been proven successful in an assay suitable for HTS of RNA targets [79,80]. Alternatively, the use of fluorescently-labelled inhibitors in RNA binding experiments has been suggested [81] which, however, would require an inverse competition assay monitoring the displacement of a labelled high-affinity binder. The latter technique is thus restricted to RNA targets for which at least one specific binder is already available. Fluorescence-based detection strategies in screening for RNA-directed ligands, however, are limited by the potential quenching of the fluorescence probe by amino groups leading to false positive hits.

Both the filter binding and the SPA screen have been successful in identifying RNA-directed inhibitors of the HIV-1 Tat-TAR interaction [82]. Within a 150,000 compound library, about 2000-3000 active compounds have been found, 500 of which were further investigated in a cell-based Tat-transactivation assay [82]. Three families of potent Tat-TAR inhibitor with IC_{50} values in the low μM range were discovered, comprising quinoxaline-2,3-diones, 2,4-diaminoquinozalines and the previously identified aminoglycosides [83].

The formation of the Tat-TAR complex and its interactions with small molecule inhibitors have been further studied by electrospray ionization mass spectrometry (ESI-MS) [48,83,84] suggesting that data from gas phase MS experiments may complement solution studies on inhibitor affinities. MS-based methods for studying RNA targets and

their interactions with small molecules [85,86] may be promising as techniques have been developed for the screening of multiple RNAs simultaneously against compound mixtures [87]. In this new approach towards HTS of RNA-small molecule interactions, the different RNA targets are tagged with neutral mass labels allowing discrimination in the mass spectrum. The technique has been applied to model oligonucleotides representing wild-type and mutant sequences of the A site of eukaryotic and prokaryotic rRNA [88]. Distinct complexes and free RNAs can be sorted out in the mass spectrometer and subjected to a second MS analysis. During the MS-MS experiment, collision induced dissociation (CID) gives rise to different fragmentation pattern for free and complexed RNA targets, indicating sites of ligand binding [88,89]. Whereas MS facilitates screening for compounds that bind to an RNA, it is not a functional assay validating the therapeutic utility of ligands. In a primary high-throughput screen, MS-based assays can identify target-specific ligands among a large set of small molecules. Ideally, MS-based screening precedes functional assays which may be of considerable complexity, as the number of remaining compounds to be screened is expected to be small.

Once HTS approaches have narrowed down on the number of interesting lead compounds, more demanding techniques can reveal insight into the interaction profiles of potential RNA-binding inhibitors. Among these methods, surface plasmon resonance (SPR) has extensively been used to study the binding of aminoglycoside derivatives to HIV-1 RRE [90] and oligonucleotides representing the bacterial 16S rRNA A site [91-94]. In order to perform SPR experiments, the RNA was biotinylated and immobilized on streptavidin-coated chips [90]. Protocols have been described which allow the screening of up to 3 different RNA targets on a single chip against a library of 12 compounds, each at 4 concentrations within one day [91]. The SPR solution approach can be automated and yields direct information on the stoichiometry of ligand binding over a wide variety of solution conditions. For measurements at low ligand concentrations, however, kinetic limitations due to surface transport may require long time frames which are prohibitive for HTS. The relatively low capacity of the SPR method along with stability problems of the coated chips in the presence of dimethyl sulfoxide (DMSO), the commonly used solvent for chemical libraries, have so far prevented the use of SPR in HTS.

RATIONAL DESIGN AND COMBINATORIAL SYNTHESIS OF RNA BINDERS

Rational structure-based design of RNA-targeted effectors requires knowledge of the three-dimensional structures of RNA targets and the principles governing molecular recognition between RNAs and ligands. X-ray crystallography and NMR spectroscopy are the two most important contributors fueling a rapidly expanding repertoire of known RNA three-dimensional motifs and architectures [11-14]. Molecular recognition has been studied for RNA complexes of proteins [95,96], peptides and low-molecular weight compounds [15-17,97]. Strategies for the structure-based rational design of molecules targeting RNA and RNA complexes have recently been outlined [1,98]. Here, we will briefly summarize previous rational and combinatorial synthesis approaches towards target-specific RNA binders.

Core moieties from natural antibiotics (Fig. (**5**)) have been used as building blocks for synthetic aminoglycoside derivatives and mimetics. The high specificity and affinity binding of natural aminoglycosides to a number of RNA targets [40,41] has been attributed to the presence of polar hydroxyl and cationic ammonium groups in the antibiotics, providing a rationale for the design of synthetic RNA binders. Chemically modified aminoglycosides in which amino groups had systematically been added [99] or hydroxyl groups removed one at a time [100] were analyzed for their binding and inhibitory effects on the catalytic hammerhead RNA. Binding affinities of modified aminoglycosides correlate with the number and basicity of cationic ammonium groups [61], supporting a general model for the interaction of cationic antibiotics with RNA based on structural electrostatic complementarity [60]. Similar findings were reported from screening of designed neomycin B analogs targeting the A site of eubacterial 16S rRNA [93]. The A site [94] and HIV-1 RRE [101] were used for binding studies of aminoglycoside mimetics obtained by conventional [94] and combinatorial [101] synthetic derivatization of the neamine core with amino acids [101] and various amines [94]. In an attempt to explore the RNA binding capacity of amino sugar compounds by simplifying the structure of the aminoglycosides, a chemical library was synthesized by adding amino acid and amine side chains to a 2-aminoglucopyranoside synthon [91]. The resulting compounds were screened for binding the eubacterial 16S rRNA A site. In the same system, aminols have been tested as aminoglycoside surrogates [102].

In the different studies of synthetic aminoglycoside analogs, a number of substances were obtained displaying RNA binding affinities in the order of those of the tightest natural binders. The binding specificity and cellular antibiotic activity, however, was generally lower for the structurally simplified aminoglycoside mimetics, underlining the important role of the natural aminoglycoside backbone as a conformationally restricted scaffold for the distinct presentation of polar and cationic groups in space [60]. Semi-synthetic compounds, displaying a higher RNA-binding and -inhibiting potential as compared to natural antibiotics, have been obtained by linking natural aminoglycosides into dimers [103,104] and with intercalating moieties such as acridine [105].

The concept of combining positively charged anchor groups for contacts with the RNA backbone, as they are provided by the ammonium substituents in aminoglycosides, with a flat aromatic moiety for stacking interactions has been realized in the modular In-PRiNts ("inhibitor of protein-ribonucleotide sequences") [106]. They have been synthesized by connecting aliphatic polyfunctional amines via a flexible linker to a number of acridine derivatives. The tripartite In-PRiNts have been designed to inhibit the interaction of the HIV-1 Tat protein with TAR RNA binding site by binding to a bulge and an adjacent stem in the RNA. Several of the compounds displayed *in vitro* activities (CD_{50} of Tat-TAR inhibition) at nanomolar concentrations and were effective in low micromolar concentrations (IC_{50}) in a cell culture-based assay. Structural analyses of the most active inhibitor (designated CGP40336A) in complex with TAR demonstrated that RNA residues of both the bulge and an adjacent stem are involved in drug binding.

The Tat-TAR system was also the target for an approach of isolating a high-affinity TAR-binding compound from a combinatorial library of peptide/peptoid oligomers comprising nine residues [107]. The starting pool of 20^5 ($3.2*10^6$) partially randomized peptide/peptoid substances was deconvoluted by consecutive rounds of binding selection in gel mobility-shift assays. For each randomized position, step by step, the residue that conferred the highest TAR-binding affinity was identified yielding a lead structure (designated CGP64222) which binds to TAR and inhibits Tat binding *in vitro* at nanomolar concentrations. The IC_{50} of CGP64222, determined in a cellular Tat-dependent transactivation assay, is in the low micromolar range. NMR and molecular modelling studies on the CGP64222-TAR complex suggest that the inhibitor recognizes the Tat-binding bulge of the RNA and induces a conformational change which might prevent binding of the protein. Further analyses of the structural determinants of Tat-TAR inhibition by CGP64222 lead to a simplified Tat antagonist CGP74026 with both reduced size and charge [108].

Apart from low-molecular weight RNA-binders, zinc finger peptides have been discussed as prototypes of specific RNA-binding scaffolds [109-111]. Zinc fingers are abundant nucleic acid-binding protein motifs comprising an antiparallel β-sheet of two strands and an α-helix which are tethered together by a coordinated zinc cation [112-114]. In a rational design approach, the RNA-binding helix of the HIV-1 Rev protein has been engineered into the α-helix of a zinc finger motif [110]. The resulting construct displayed highly specific RRE-binding activity leading to inhibition of Rev function *in vivo*. The hybrid zinc finger motif, thus, retains the capacity to recognize a non-Watson-Crick G•G base pair in RRE that is a key determinant for Rev-binding [29]. Zinc finger proteins that recognize a non-Watson-Crick G•A base pair within an RNA helix have been obtained from a phage-displayed zinc finger library [111]. Phage display of zinc finger modules might be useful for rational design of peptides that recognize distinct RNA structural elements such as regulatory motifs in mRNA.

DOCKING STRATEGIES

The rapidly increasing number of RNA complexes for which three-dimensional structures have been elucidated by X-ray crystallography and NMR [16,17,95-97] lays the foundation for rational approaches for the docking of ligands to RNA targets. The investigation of ligand-RNA interactions by computational approaches, outlined in this section, is dependent on the availability of three-dimensional models of RNA targets based on either X-ray, NMR or phylogenetic data. Enormous progress in RNA synthesis and structure determination methods have helped to overcome many of the difficulties in obtaining NMR or crystal structures of RNA [11-14]. Whereas complete structures for large RNAs (rRNA, RNase P, etc.) are not available yet, a number of constitutive smaller RNA motifs, which are potential target sites for the binding of small molecules, has been crystallized or investigated by NMR. Even in the absence of experimentally determined structures, valid three-dimensional models for RNAs can be constructed based on phylogenetic data [115-117].

Careful covariation analyses (see above, Target definition and exploration) have been shown to be powerful tools for predictions beyond Watson-Crick pairing, providing information on non-canonical base pairs, full structural motifs and tertiary interactions [21,24,115-117].

Whereas validated routine procedures are now available for the docking of ligands to protein targets, and even "*in silico* screening" of "virtual libraries" is possible, approaches to dock small molecules to RNA folds are still in their infancy. Docking methods developed for protein targets, such as the DOCK program [118] and 3D-SAR-based docking [119] have been used to investigate ligand binding to RNA double helices [120] and the HIV-1 RRE [121]. The goal of specific molecular recognition of RNA is conceptually close to targeting protein folds. However, the weights for the energetic contributions to complex formation must be modulated. Whereas the roles of H-bonds and interactions maximized by molecular shape complementarity, such as van der Waals, stacking and hydrophobic forces, are comparable in protein and RNA complexes [122], the contributions of electrostatic interactions to RNA binding stand out due to the polyanionic nature of the nucleic acid. Electrostatic interactions, however, are often subject to crude approximations in commonly used docking procedures which were optimized for protein targets.

Although lacking directionality, electrostatic interactions are a major driving force of the binding of cationic compounds to RNA folds. Specificity is achieved via the three-dimensional fold of the RNA and the ensuing highly complex electrostatic fields. Thus, in order to explain the target specificity of the cationic aminoglycoside antibiotics, the concept of structural electrostatic complementarity has been formulated [60], providing a general rationale for the docking of cationic compounds to RNA targets. Due to restricted conformational flexibility, the aminoglycoside ring systems provide scaffolds for distinct spatial orientation of positively charged ammonium groups. The three-dimensional array of positive charges presented by the antibiotics can match electronegative pockets in RNA folds which are occupied by metal cations. The simultaneous docking of several ammonium groups into electronegative pockets allows for specific discrimination of RNA targets by aminoglycoside antibiotics [60,61]. The strength and long-range character of electrostatic forces lead to a straightforward primary *in silico* screening method for docking cationic molecules to RNA folds guided by the positions of either experimentally determined metal ion binding sites or predicted electronegative pockets [34,98]. This method has been used to investigate the inhibition of the self-cleaving hammerhead ribozyme by aminoglycoside antibiotics. Solution conformers of neomycin B and tobramycin have been docked to the hammerhead RNA by displacing catalytic Mg^{2+} ions from their crystallographic positions [60]. Molecular dynamics (MD) simulations of the solvated complexes suggest that intermolecular contacts of the aminoglycosides resemble interactions of metal ions in the electronegative pockets down to an atomic level. Binding studies of rationally modified aminoglycosides support the proposed interactions in the modelled hammerhead-aminoglycoside complexes [61,99,100]. Whereas the presence of several positive charges, and the ensuing problems of drug delivery *in vivo*, renders aminoglycosides less desirable lead compounds, it is safe to assume that any specific ligand for an RNA target will contain at least one cationic functional group along with other positively polarized hydrogen donors. The concepts of electrostatic complementarity and metal ion displacement provide powerful paradigms for the placement of the key cationic group within rationally designed ligands for RNA targets.

For a growing number of potential RNA targets, three-dimensional structures are available from NMR studies or phylogenetic modelling which both do not provide the positions of bound metal ions. In the case of self-splicing group I introns, which are inhibited by aminoglycoside antibiotics and for which molecular models based on phylogeny are available, neomycin has been docked guided by electrostatic complementarity, using modelled positions of metal ions at the active site [123]. In the absence of experimental information on metal ion binding sites, Brownian dynamics (BD) simulations of cation diffusion can be used to predict electronegative pockets guiding the docking of cationic molecules to RNA folds [34,124]. The method has been validated for aminoglycoside complexes of RNA aptamers and the 16S rRNA A site for which predicted electronegative pockets of the RNA coincede with the positions of cationic ammonium groups in the bound aminoglycosides [34].

A general strategy for *in silico* screening of cationic RNA binders has been outlined, combining i) the BD simulation prediction of electronegative pockets, ii) the docking based on electrostatic complementarity and iii) the investigation of intermolecular interactions by MD simulations of solvated RNA-complexes (Fig. (**6**)) [98]. The primary screening step of complementarity-based

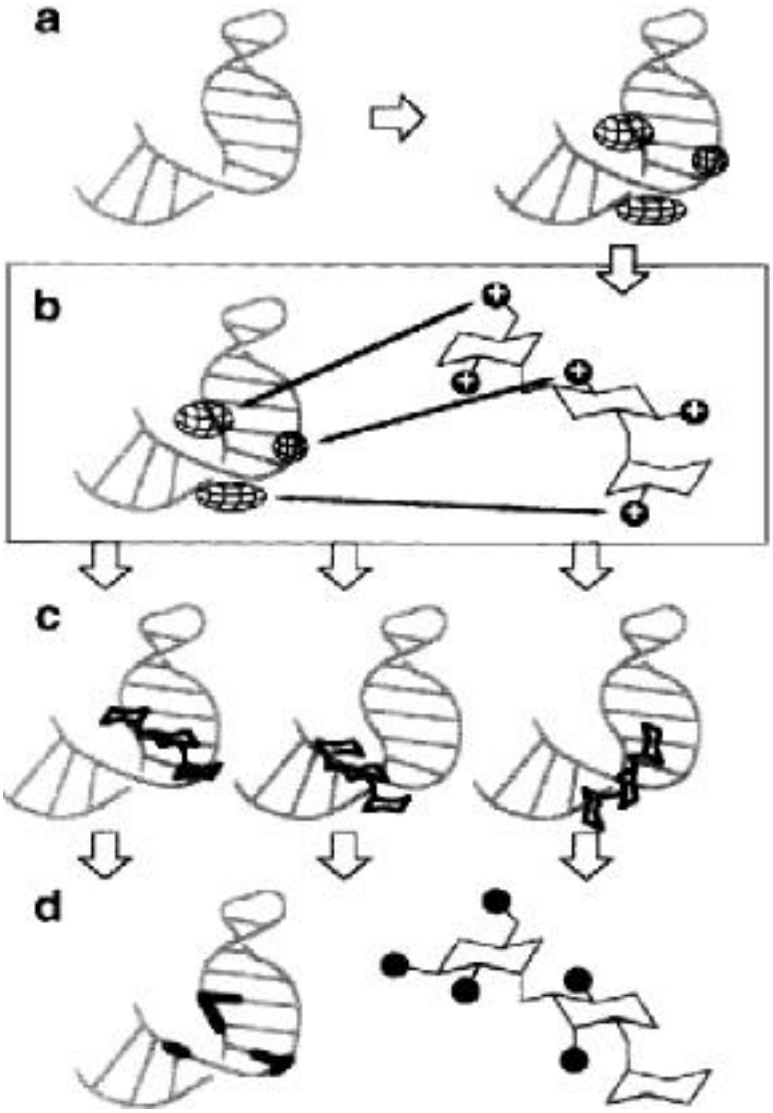

Fig. (6). Strategy for *in silico* screening of ligands for binding to potential RNA targets (adapted from [98]). (**a**) Brownian dynamics simulations of cation diffusion are used to predict electronegative pockets for RNA structural models derived from crystallography, NMR spectroscopy or modelling. (**b**) Structural electrostatic complementarity between the array of positive charges in conformers of the ligands and the arrangement of electronegative pockets in the RNA is exploited to guide the docking of low-molecular weight compounds to the RNA target. (**c**) The modelled complexes are subjected to molecular dynamics simulations in order to investigate possible conserved intermolecular contact sites in both the RNA (indicated by bold lines) and the ligands (circles) (**d**).

docking accounts for the strong and long-ranging electrostatic interactions and steric exclusion. The full range of intermolecular interactions, eventually tuned by solvent and salt ions, is finely probed in the second screening step of the MD simulations which allow the docked ligand to explore a potential binding site in the RNA.

Reliable simulations of highly charged nucleic acids and complexes thereof call for accurate treatment of electrostatic forces by Ewald summation of non-bonded interactions and periodic boundary conditions in the presence of explicit solvent and salt ions [125,126]. The high computational cost for these sophisticated simulations is prohibitive for an application towards the *in silico* screening of large ligand libraries to date. While number crunching power continues to increase, eventually allowing the testing of larger numbers of potential ligands, MD-based approaches to RNA complexes are focusing around lead structures known to bind a distinct RNA

target. The potential to study intermolecular interactions on a short time scale in virtually atomic detail renders MD simulations under realistic solvent conditions the method of choice for structural investigations of RNA complexes, surpassed only by time-, labour- and material-intensive experimental approaches such as X-ray crystallography and NMR spectroscopy.

In addition to high-throughput capacity, a key challenge for docking approaches is the implementation of procedures that deal with conformational flexibility of both the ligand and the RNA target site. Published docking studies on RNA targets have used rigid ligands and RNA [120], which allows for screening large ligand libraries, or rigid RNA and flexible ligand [34,121], which requires additional computation to produce and dock sets of conformers for each compound. While MD simulations of docked complexes [34,60,121] to some extent reveal insight into dynamical changes, it is highly desirable to develop algorithms which take into account flexibility in ligand and target during the docking process. For docking to RNA structures, target flexibility is especially important as adaptive binding processes upon complex formation have frequently been observed with RNA (see above). Unless fast algorithms are available that can also reliably simulate larger conformational rearrangements in RNA targets, a feasible way includes docking to RNA conformers from complexes the three-dimensional structures of which are available from crystallography or NMR spectroscopy.

In summary, future efforts of developing docking strategies for RNA targets will have to deal with two key challenges, namely i) the accurate description of electrostatic interactions including a realistic treatment of the solvent, and ii) the development of algorithms that allow to simulate conformational flexibility in the RNA target.

CONCLUSION AND PERSPECTIVES

Many of the challenges of drug screening and design for protein targets are identical to those approaches intended to exploit RNAs as targets for therapeutic intervention. The successful development of scalable assays for testing target-specific RNA binding of small molecules has demonstrated that it is now possible to routinely perform high-throughput screening for RNA-directed ligands. The RNAs necessary for binding assays are easily accessible in large amounts by *in vitro* transcription or chemical synthesis, without the need for developing elaborate protein expression systems. This will allow for the rapid evaluation of the potential RNA targets which emerge in large numbers from whole genome sequencing projects.

On the side of drug design approaches, our view may still be biased by the protein world. The limitations inherent to the application of methods developed for studying protein targets to RNA world may explain some of the difficulties encountered in the present search for RNA-directed compounds. While the physical laws of intermolecular interactions are the same for proteins and RNAs, the example of electrostatic forces shows, however, that a shifted focus on energetic contributions calls for a much higher level of sophistication in the theoretical treatment of the interactions and thermodynamics of recognition and binding. The wide success of biophysical techniques for studying RNA structures, along with the rapidly increasing computer power, emerging improvements in simulation algorithms and novel molecular docking and design approaches will pave the way for the rational development of new RNA-directed drugs.

ACKNOWLEDGEMENTS

We are both grateful to Dinshaw J Patel for discussions. TH thanks Dinshaw J. Patel for support.

LIST OF ABBREVIATIONS

BD	=	Brownian dynamics
CAD	=	Collisionally activated dissociation
DMSO	=	Dimethyl sulfoxide
DNA	=	Deoxyribonucleic acid
ESI-MS	=	Electrospray ionization mass spectrometry
EST	=	Expressed sequence tag
GTP	=	Guanosine triphosphate
HDV	=	Hepatitis delta virus
HIV	=	Human immunodeficiency virus
HTS	=	High-throughput screening

In-PRiNts = Inhibitor of protein-ribonucleotide sequences

IRE = Iron-responsive element

IRES = Internal ribosome entry site

MD = Molecular dynamics

mRNA = Messenger RNA

MS = Mass spectrometry

ncRNA = Non-coding RNA

NMR = Nuclear magnetic resonance

ORF = Open reading frame

RNA = Ribonucleic acid

RRE = Rev-response element

rRNA = Ribosomal RNA

SAR = Structure-activity relationship

SECIS = Selenocysteine insertion element

snoRNA = Small nucleolar RNA

SNP = Single-nucleotide polymorphism

snRNP = Small nuclear ribonucleoprotein

SPA = Scintillation proximity assay

SPR = Surface plasmon resonance

TAR = Trans-activating region

tRNA = Transfer RNA

TS = Thymidylate synthase

UTR = Untranslated region

REFERENCES

[1] Hermann, T. *Angew. Chem. Int. Ed. Engl.*, **2000**, *39*, 4000-4015.

[2] Wilson, W.D.; Li, K. *Curr. Med. Chem.*, **2000**, *7*, 73.

[3] Afshar, M.; Prescott, C.D.; Varani, G. *Curr. Opin. Biotech.*, **1999**, *10*, 59.

[4] Tor, Y. Angew. Chem. Int. Ed. Engl, **1999**, 38, 1579.

[5] Ecker, D.J.; Griffey, R.H. *Drug Discovery Today*, **1999**, *4*, 420.

[6] Hermann, T.; Westhof, E. *Curr. Opin. Biotech.*, **1998**, *9*, 66.

[7] Pearson, N.D.; Prescott, C.D. *Chem. Biol.*, **1997**, *4*, 409.

[8] Zweiger, G.; Scott, R.W. *Curr. Opin. Biotechnol.*, **1997**, *8*, 684.

[9] Dunham, I. et al. *Nature*, **1999**, *402*, 489.

[10] Lenz, G.R.; Nash, H.M.;Jindal, J.D. *Drug Discovery Today*, **2000**, *5*, 145.

[11] Masquida, B.; Westhof, E. In *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press; Oxford, **1999**, pp.533-565.

[12] Nowakowski, J.; Tinoco, I. jr. In *Oxford Handbook of Nucleic Acuid Structure*; Neidle, S., Ed.; Oxford University Press; Oxford, **1999**, pp.567-602.

[13] Batey, R.T.; Rambo, R.P.; Doudna, J.A. *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*, 2326.

[14] Hermann, T.; Patel, D.J. *J. Mol. Biol.*, **1999**, *294*, 829.

[15] Hermann, T.; Westhof, E. *Biopolymers (Nucleic Acid Sciences)*, **1998**, *48*, 155.

[16] Puglisi, J.D.; Williamson, J.R. In *The RNA World*, 2nd ed.; Gesteland, R.F.; Cech, T.R.; Atkins, J.F., Eds.; Cold Spring Harbor Laboratory Press; New York, **1999**, pp.403-425,.

[17] Hermann, T.; Patel, D.J. *Science*, **2000**, *287*, 820-825.

[18] Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol. Struct.*, **1997**, *26*, 113.

[19] Tinoco, I.; Bustamante, C. *J. Mol. Biol.*, **1999**, *293*, 271.

[20] Saenger, W. *Principles of nucleic acid structure*, Springer-Verlag: New York, 1984.

[21] Leontis, N.B.; Westhof, E. *Quart. Rev. Biophys.*, **1998**, *31*, 399.

[22] Westhof, E.; Fritsch, V. *Structure*, **2000**, *8*, R55.

[23] Hermann, T.; Westhof, E. *Chem. Biol.*, **1999**, *6*, R335.

[24] Leontis, N.B.; Westhof, E. *J. Mol. Biol.*, **1998**, *283*, 571.

[25] Peterson, R.D.; Bartel, D.P.; Szostak, J.W.; Horvath, S.J.; Feigon, J. *Biochemistry*, **1994**, *33*, 5357.

[26] Ippolito, J.A.; Steitz, T.A. *J. Mol. Biol.*, **2000**, *295*, 711.

[27] Peterson, R.D.; Feigon, J. *J. Mol. Biol.*, **1996**, *264*, 863.

[28] Battiste, J.L.; Mao, H; Rao, N.S.; Tan, R.; Muhandiram, D.R.; Kay, L.E.; Frankel, A.D; Williamson, J.R. *Science*, **1996**, *273*, 1547.

[29] Werstuck, G.; Zapp, M.L.; Green, M.R. *Chem. Biol.*, **1996**, *3*, 129.

[30] Long, K.S.; Crothers, D.M. *Biochemistry*, **1999**, *38*, 10059.

[31] Aboul-ela, F.; Karn, J.; Varani, G. *J. Mol. Biol.*, **1995**, *253*, 313.

[32] Aboul-ela, F.; Karn, J.; Varani, G. *Nucleic Acids Res.*, **1996**, *24*, 3974.

[33] Wang, S.; Huber, P. W.; Cui, M.; Czarnik, A. W.; Mei, H.-Y. *Biochemistry*, **1998**, *37*, 5549.

[34] Hermann, T.; Westhof, E. *J. Med. Chem.*, **1999**, *42*, 1250.

[35] Eddy, S.R. *Curr. Opin. Genet. Dev.*, **1999**, *9*, 695.

[36] Erdmann, V.A.; Szymanski, M.; Hochberg, A.; de Groot, N.; Barciszewski, J. *Nucleic Acids Res.*, **2000**, *28*, 197.

[37] Hazelrigg, T. *Cell*, **1998**, *95*, 451.

[38] Gray, N.K.; Wickens, M. *Annu. Rev. Cell Dev. Biol.*, **1998**, *14*, 399.

[39] Pesole, G.; Liuni, S.; Saccone, C. *Gene*, **1997**, *205*, 95.

[40] Schroeder, R.; Waldsich, C.; Wank, H. *EMBO J.*, **2000**, *19*, 1.

[41] Walter, F.; Vicens, Q.; Westhof, E. *Curr. Opin. Chem. Biol.*, **1999**, *3*, 694.

[42] Moazed, S.; Noller, H.F. *Nature*, **1987**, *327*, 389.

[43] Mikkelsen, N.E.; Brannvall, M.; Virtanen, A.; Kirsebom, L.A. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 6155.

[44] von Ahsen, U.; Davies, J.; Schroeder, R. *Nature*, **1991**, *353*, 368.

[45] Stage, T. K.; Hertel, K. J.; Uhlenbeck, O. C. *RNA*, **1995**, *1*, 95.

[46] Rogers, J.; Chang, A. H.; von Ahsen, U.; Schroeder, R.; Davies, J. *J. Mol. Biol.*, **1996**, *259*, 916.

[47] Tok, J.B.H.; Cho, J.; Rando, R.R. *Biochemistry*, **1999**, *38*, 199.

[48] Mei, H.-Y.; Galan, A. A.; Halim, N. S.; Mack, D. P.; Moreland, D. W.; Sanders, K. B.; Truong, H. N.; Czarnik, A. W. *Bioorg. Med. Chem. Lett.*, **1995**, 5, 2755.

[49] Zapp, M. L.; Stern, S.; Green, M. R. *Cell*, **1993**, *74*, 969.

[50] Fichant, G.A.; Burks, C. *J. Mol. Biol.*, **1991**, *220*, 659.

[51] Eddy, S.R.; Durbin, R. *Nucleic Acids Res.*, **1994**, *22*, 2079.

[52] Sakakibara, Y.; Brown, M.; Hughey, R.; Mian, I.S.; Sjolander, K.; Underwood, R.C.; Haussler, D. *Nucleic Acids Res.*, **1994**, *22*, 5112.

[53] Lowe, T.M.; Eddy, S.R. *Nucleic Acids Res.*, **1997**, *25*, 955.

[54] Lisacek, F.; Diaz, Y.; Michel, F. *J. Mol. Biol.*, **1994**, *235*, 1206.

[55] Ferbeyre, G.; Smith, J.M.; Cedergren, R. *Mol. Cell. Biol.*, **1998**, *18*, 3880.

[56] Ryan, P.C.; Lu, M.; Draper, D.E. *J. Mol. Biol.*, **1991**, *221*, 1257.

[57] Conn, G.L.; Draper, D.E.; Lattmann, E.E.; Gittis, A.G. *Science*, **1999**, *284*, 1171.

[58] Wimberly, B.T.; Guymon, R.; McCutcheon, J.P.; White, S.W.; Ramakrishnan, V. *Cell*, **1999**, *97*, 491.

[59] Wallis, M.G.; Streicher, B.; Wank, H.; von Ahsen, U.; Clodi, E.; Wallace, S.T.; Famulok, M.; Schroeder, R. *Chem. Biol.*, **1997**, *4*, 357.

[60] Hermann, T.; Westhof, E. *J. Mol. Biol.*, **1998**, *276*, 903.

[61] Tor, Y.; Hermann, T.; Westhof, E. *Chem. Biol.*, **1998**, *5*, R277.

[62] Fourmy, D.; Recht, M.I.; Blanchard, S.C.; Puglisi, J.D. *Science*, **1996**, *274*, 1367.

[63] Dandekar, T.; Hentze, M.W. *Trends Genet.*, **1995**, *11*, 45.

[64] Bourdeau, V.; Ferbeyre, G.; Pageau, M.; Paquin, B.; Cedergren, R. *Nucleic Acids Res.*, **1999**, *27*, 4457.

[65] Dandekar, T.; Stripecke, R.; Gray, N.K.; Goossen, B.; Constable, A.; Johansson, H.E.; Hentze, M.W. *EMBO J.*, **1991**, *10*, 1903.

[66] Dandekar, T.; Beyer, K.; Bork, P.; Kenealy, M.R.; Pantopoulos, K.; Hentze, M.W.; Sonntag-Buck, V.; Flouriot, G.; Gannon, F.; Schreiber, S. *Bioinformatics*, **1998**, *14*, 271.

[67] Beyer, K.; Dandekar, T.; Keller, W. *J. Biol. Chem.*, **1997**, *272*, 26769.

[68] Ferbeyre, G.; Bourdeau, V.; Cedergren, R. *Trends Biochem. Sci.*, **1997**, *22*, 115.

[69] Lowe, T.M.; Eddy, S.R. *Science*, **1999**, *283*, 1168.

[70] Lescure, A.; Gautheret, D.; Carbon, P.; Krol, A. *J. Biol. Chem.*, **1999**, *274*, 38147.

[71] Kryukov, G.V.; Kryukov, V.M.; Gladyshev, V.N. *J. Biol. Chem.*, **1999**, *274*, 33888.

[72] Pesole, G.; Liuni, S.; Grillo, G.; Licciulli, F.; Larizza, A.; Makalowski, W.; Saccone, C. *Nucleic Acids Res.*, **2000**, *28*, 193.

[73] Olivas, W.M.; Muhlrad, D.; Parker, R. *Nucleic Acids Res.*, **1997**, *25*, 4619.

[74] Bailey, D.S.; Bondar, A.; Furness, L.M. *Curr. Opin. Biotechnol.*, **1998**, *9*, 595.

[75] Evans, W.E.; Relling, M.V. *Science*, **1999**, *286*, 487.

[76] Shen, L.X.; Basilion, J.P.; Stanton, V.P. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 7871.

[77] Mei, H.-Y.; Cui, M.; Sutton, S.T.; Truong, H.N.; Chung, F.Z.; Czarnik, A.W. *Nucleic Acids Res.*, **1996**, *24*, 5051.

[78] Mei, H.-Y.; Cui, M.; Lemrow, S.M.; Czarnik, A.W. *Bioorg. Med. Chem.*, **1997**, *5*, 1185.

[79] Jenne, A.; Gmelin, W.; Raffler, N.; Famulok, M. *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*, 1300.

[80] Llano-Sotelo, B.; Chow, C.S. *Bioorg. Med. Chem. Lett.*, **1999**, *9*, 213.

[81] Hamasaki, K.; Rando, R.R. *Anal. Chem.*, **1998**, *261*, 183.

[82] Mei, H.-Y.; Mack, D.P.; Galan, A.A.; Halim, N.S.; Heldsinger, A.; Loo, J.A.; Moreland, D.W.; Sannes-Lowery, K.A.; Sharmeen, L.; Truong, H.N.; Czarnik, A.W. *Bioorg. Med. Chem.* **1997**, *5*, 1173.

[83] Mei, H.-Y.; Cui, M.; Heldsinger, A.; Lamrow, S.M.; Loo, J.A.; Sannes-Lowery, K.A.; Sharmeen, L.; Czarnik, A.W. *Biochemistry*, **1998**, *37*, 14204.

[84] Sannes-Lowery, K.A.; Hu, P.; Mack, D.P.; Mei, H.-Y.; Loo, J.A. *Anal. Chem.*, **1997**, *69*, 5130.

[85] Crain, P.F.; McCloskey, J.A. *Curr. Opin. Biotechnol.*, **1998**, *9*, 25.

[86] Loo, J.A.; Thanabal, V.; Mei, H.-Y. *Mass Spectrom. Biol. Med.*, **2000**, 73.

[87] Hofstadler, S.A.; Sannes-Lowery, K.A.; Crooke, S.T.; Ecker, D.J.; Sasmor, H.; Manalili, S.; Griffey, R.H. *Anal. Chem.*, **1999**, *71*, 3436.

[88] Griffey, R.H.; Hofstadler, S.A.; Sannes-Lowery, K.A.; Ecker, D.J.; Crooke, S.T. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 10129.

[89] Griffey, R.H.; Greig, M.J.; An, H.; Sasmor, H.; Manalili, S. *J. Am. Chem. Soc.*, **1999**, *121*, 474.

[90] Hendrix, M.; Priestley, E.S.; Joyce, G.F.; Wong, C.H. *J. Am. Chem. Soc.*, **1997**, *119*, 3641.

[91] Wong, C.-H.; Hendrix, M.; Manning, D.D.; Rosenbohm, C.; Greenberg, W.A. *J. Am. Chem. Soc.*, **1998**, *120*, 8319.

[92] Wong, C.-H.; Hendrix, M.; Priestley, E.S.; Greenberg, W.A. *Chem. Biol.*, **1998**, *5*, 397.

[93] Alper, P.B.; Hendrix, M.; Sears, P.; Wong, C.-H. *J. Am. Chem. Soc.*, **1998**, *120*, 1965.

[94] Greenberg, W.A.; Priestley, E.S.; Sears, P.S.; Alper, P.B.; Rosenbohm, C.; Hendrix, M.; Hung, S.-C.; Wong, C.-H. *J. Am. Chem. Soc.*, **1999**, *121*, 6527.

[95] De Guzman, R.N.; Turner, R.B.; Summers, M.F. *Biopolymers: Nucleic Acids Sciences*, **1998**, *48*, 181.

[96] Cusack, S. *Curr. Opin. Struct. Biol.*, **1999**, *9*, 66.

[97] Patel, D.J. *Curr. Opin. Struct. Biol.*, **1999**, *9*, 74.

[98] Hermann, T.; Westhof, E. In *RNA-Binding Antibiotics*; Wallis, M.G.; Schroeder, R., Eds.; Landes Bioscience; Austin TX, **2000**, pp. 148-158.

[99] Wang, H.; Tor, Y. *Angew. Chem. Int. Ed. Engl.*, **1998**, *37*, 109.

[100] Wang, H.; Tor, Y. *J. Am. Chem. Soc.*, **1997**, *119*, 8734.

[101] Park, W.K.C.; Auer, M.; Jaksche, H.; Wong, C.H. *J. Am. Chem. Soc.*, **1996**, *118*, 10150.

[102] Tok, J.B.H.; Rando, R.R. *J. Am. Chem. Soc.*, **1998**, *120*, 8279.

[103] Wang, H.; Tor, Y. *Bioorg. Med. Chem. Lett.*, **1997**, *7*, 1951.

[104] Michael, K.; Wang, H.; Tor, Y. *Bioorg. Med. Chem.*, **1999**, *7*, 1361.

[105] Kirk, S.R.; Luedtke, N.W.; Tor, Y. *J. Am. Chem. Soc.*, **2000**, *122*, 980.

[106] Hamy, F.; Brondani, V.; Flörsheimer, A.; Stark, W.; Blommers, M. J.J.; Klimkait, T. *Biochemistry*, **1998**, *37*, 5086.

[107] Hamy, F.; Felder, E.R.; Heizmann, G.; Lazdins, J.; Aboul-Ela, F.; Varani, G.; Karn, J.; Klimkait, T. *Proc. Natl. Acad. Sci. USA*, **1997**, *94*, 3548.

[108] Klimkait, T.; Felder, E.R.; Albrecht, G.; Hamy, F. *Biotechnol. Bioeng.*, **1998**, *61*, 155.

[109] Friesen, J.W.; Darby, M.K. *Nat. Struct. Biol.*, **1998**, *5*, 543.

[110] McColl, D.J.; Honchell, C.D., Frankel, A.D. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 9521.

[111] Blancaford, P.; Steinberg, S.V.; Paquin, B.; Klinck, R.; Scott, J.K.; Cedergren, R. *Chem. Biol.*, **1999**, *6*, 585.

[112] Lee, M.S.; Gippert, G.P.; Soman, K.V.; Case, D.A.; Wright, P.E. *Science*, **1989**, *245*, 635.

[113] Pavletich, N.P.; Pabo, C.O. *Science*, **1991**, *252*, 809.

[114] Fairall, L.; Schwabe, J.W.; Chapman, L.; Finch, J.T.; Rhodes, D. *Nature*, **1993**, *366*, 483.

[115] Michel, F.; Westhof, E. *J. Mol. Biol.*, **1990**, *216*, 585.

[116] Lehnert, V.; Jaeger, L.; Michel, F.; Westhof, E. *Chem. Biol.*, **1996**, *3*, 993.

[117] Massire, C.; Jaeger, L.; Westhof, E. *J. Mol. Biol.*, **1998**, *279*, 773.

[118] Kuntz, I.D.; Meng, E.C.; Shoichet, B.K. *Acc. Chem. Res.*, **1994**, *27*, 117.

[119] Golender, V.E.; Vorpagel, E.R. In *3D-QSAR in Drug Design: Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM; Leiden, **1993**, pp.137-149.

[120] Chen, Q.; Shafer, R.H.; Kuntz, I.D. *Biochemistry*, **1997**, *36*, 11402.

[121] Leclerc, F.; Cedergren, R. *J. Med. Chem.*, **1998**, *41*, 175.

[122] Nadassy, K.; Wodak, S.J.; Janin J. *Biochemistry* **1999**, *38*, 1999.

[123] Hoch, I.; Berens, C.; Westhof, E.; Schroeder, R. *J. Mol. Biol.*, **1998**, *282*, 557.

[124] Hermann, T.; Westhof, E. *Structure*, **1998**, *6*, 1303.

[125] Sagui, C; Darden, T.A. *Annu. Rev. Biophys. Biomol. Struct.*, **1999**, *28*, 155.

[126] Auffinger, P.; Westhof, E. *Curr. Opin. Struct. Biol.*, **1998**, *8*, 227.

[127] Szewczak, A.A.; Moore, P.B. *J. Mol. Biol.*, **1995**, *247*, 81.

[128] Correll, C.C.; Wool, I.G.; Munishkin, A. *J. Mol. Biol.*, **1999**, *292*, 275.

[129] Walczak, R.; Westhof, E.; Carbon, P.; Krol, A. *RNA*, **1996**, *2*, 367.

[130] Botto, R.E.; Coxon, B. *J. Am. Chem. Soc.*, **1983**, *105*, 1021.

Received 13 March 2000 Accepted 1 May, 2000