

RNA: Primed for Packing?

How does RNA's seemingly subtle 2'-hydroxyl group contribute to its diverse biological functions? Using quantitative structure-activity relationship (QSAR) analysis, Piccirilli and colleagues [1] find that the 2'-OH fosters van der Waals and hydrophobic interactions that play underappreciated roles in RNA structure and molecular recognition.

As the known functions of RNA in cells and viruses continue to expand, detailed images of RNA molecules including the ribosome have confirmed the notion of RNA as a molecular contortionist. In contrast, DNA seems slated to remain RNA's conservative cousin whose chemical properties safeguard genetic information while necessarily limiting functional promiscuity. While this difference between the nucleic acids is well known at a qualitative level, quantitative explanations have been hard to come by.

High-resolution structures of RNA and DNA have highlighted some obvious differences between these chemically similar polymers. The deeper, narrower major groove of RNA duplexes relative to DNA precludes interactions of proteins that might otherwise bind to this part of the molecule. Because the major groove contains many of the available hydrogen-bonding functionalities on the bases, sequence-specific recognition of a canonical RNA duplex is consequently much harder to achieve. RNA solves this problem by utilizing non-Watson-Crick base pairs, which can form because of the propensity of RNA to exist as a single stranded, self-folded structure instead of the uniform double-stranded duplex observed for DNA. In addition to locally widening the major groove, noncanonical base pairings also can create unique structures and ligand binding sites within an otherwise repetitive double helix. (It is interesting to note that despite these structural differences, both RNA and DNA can act as efficient catalysts [2]. This observation suggests that these differences are not as fundamental as the different biological roles of RNA and DNA would make us believe.)

An obvious candidate for generating the observed structural and functional differences between RNA and DNA is the 2'-OH group that decorates the RNA sugar-phosphate backbone. Long the subject of scrutiny, its role in RNA function and structure has remained surprisingly elusive. To shed new light onto the contributions of the 2'-OH, Piccirilli and colleagues focused a keen chemist's eye on a group II ribozyme-catalyzed reaction that mimics the reverse of the second step of self-splicing [1]. Ribozymes have been popular subjects for studying relationships between RNA structure and function, as structural perturbations can be readily and quantitatively assessed by analyzing their effects on catalytic activity. The present study by Gordon et al., which uses

an exceptionally comprehensive set of nucleotide analogs, reveals that at least for this self-splicing related reaction, none of the usual suspects are responsible for the catalytic advantage of the 2'-OH relative to a 2'-H: neither hydrogen bonding nor metal ion coordination nor an inductive effect account for the observed 10-fold effect of substituting the 2'-OH with a 2'-H [1].

To explain their results, the authors turned to quantitative structure activity relationships (QSAR). This analytical technique has been pioneered in pharmacology, where it is frequently applied to help rationalize and optimize the design of bioactive small molecules based on the behavior of an initial set of compounds. In this method, the activity of a series of systematically varied compounds (i.e., enzyme inhibitors or substrates) is plotted in a model-free way against a variety of parameters associated with the inhibitor. These parameters include, but are not limited to, molecular volume, pK_a values, or hydrophobicity.

In the case of the group II ribozyme reaction, QSAR analysis suggests that catalytic activity of different 2'-substituted substrate analogs is a function of the molecular volume of the analog as well as the hydrophobicity of the 2' substituent, with more hydrophilic residues reacting faster. (Although the QSAR analysis quantitatively describes the behavior of a large series of 2'-substituted substrates, the poor reactivity of the 2'-H is not fully explained, as this falls off the lines in Figure 6 in Gordon et al. Perhaps this is because the 2'-H is smaller than all the other 2' substituents, especially if the 2'-F, the next smallest residue, is hydrated.) This dependence on both molecular volume and hydrophobicity is not completely unprecedented. Previous work on DNA polymerase has shown that recognition of the correct incoming nucleotide depends less on hydrogen bonding than on shape complementarity with the base-pairing partner on the template [3]. This observation implied that the cumulative energetic contributions from van der Waals interactions are large relative to those from hydrogen bonds, perhaps due to competition of the latter with water. In another example, quantitative thermodynamic analysis of adenosine-minor groove interactions (the "A-minor" motif) within a group I intron [4, 5] suggested an important role for shape complementarity, which maximizes van der Waals interactions, and assigned rather small thermodynamic contributions to hydrogen bonding (Figure 1). These functional studies are further extended by an analysis of crystal structures of RNAs (and DNAs) showing that, in most cases, interhelical packing maximizes van der Waals contacts [6].

Taken together, these data suggest that at least at the sites where tertiary interactions occur, close packing might be just as important in RNA molecules as it is in proteins. This surprising result would seem to belie the highly charged nature of RNA molecules, which might naturally be expected to result in less close packing—or exclusively metal ion-mediated contacts—due to charge repulsion.

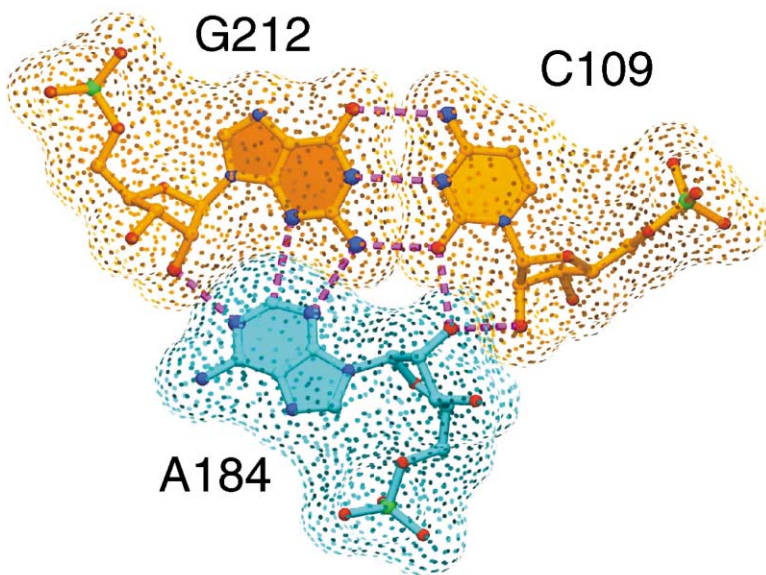


Figure 1. The A-Minor Motif Maximizes van der Waals Interactions

While the electronegative RNA phosphodiester backbone is expected to oppose close packing, it could account for the observed inverse dependence of reaction rates on the hydrophobicity of the 2' substituent. Gordon et al. suggest instead that this observation can be explained by binding of a positioned water molecule (Figure 2). This model is consistent with hydration patterns around 2'-OH groups observed in high-resolution crystal structures. Alternatively, it is possible that in the case of the hydrophobic 2' substituents, the RNA structure rearranges slightly to prevent direct interaction of the hydrophobic group with the highly polar RNA molecule. Such small rearrangements would be consistent with the observation that ethanol but not methanol can partially rescue activity of the 2'-H-substituted substrate (Figure 2), while a 2'-hydroxymethyl substituent reacts

7-fold faster than a 2'-hydroxyethyl-containing substrate.

While the lack of distinct interactions with the cleavage site 2'-OH observed by Gordon et al. might be specific for this system and, indeed, may not even hold for both steps of the group II intron self-splicing reaction, these results underscore the need for a careful analysis of point mutations to reveal functional interactions and their energetic importance. In this regard, RNA has the upper hand over protein: while in vitro production of proteins that specifically incorporate nonnatural amino acids remains challenging, solid-phase synthesis of RNA containing a large variety of nucleotide analogs with different base, sugar, and backbone functionalities is straightforward, provided that the appropriate phosphoramidites are available. Gordon et al. show that

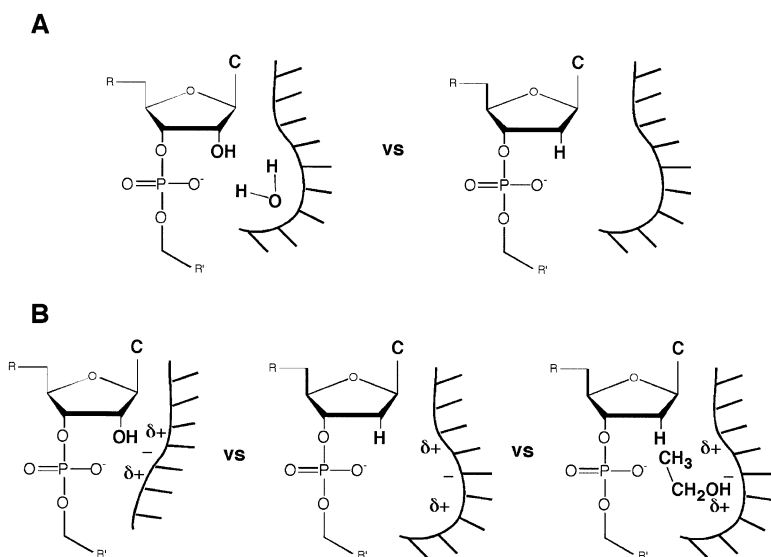


Figure 2. Models to Explain the Inverse Dependence of Reaction Rates on the Hydrophobicity of the 2' Substituent

QSAR analysis of the behavior of such analog-containing RNAs can reveal chemical principles that are not otherwise obvious. Future progress toward total synthesis of proteins containing nonnatural residues will undoubtedly yield similarly surprising insights into the function of protein enzymes.

Acknowledgments

K.K. is a Damon Runyon fellow supported by the Damon Runyon Cancer Research Foundation.

Katrin Karbstein¹ and Jennifer A. Doudna^{1,2}

¹Department of Molecular and Cellular Biology

²Howard Hughes Medical Institute

University of California, Berkeley

Berkeley, California 94720

Selected Reading

1. Gordon, P.M., Fong, R., Deb, S.K., Li, N.S., Schwans, J.P., Ye, J.D., and Piccirilli, J.A. (2004). *Chem. Biol.* *11*, this issue, 237–246.
2. Santoro, S.W., and Joyce, G.F. (1997). *Proc. Natl. Acad. Sci. USA* *94*, 4262–4266.
3. Morales, J.C., and Kool, E.T. (1998). *Nat. Struct. Biol.* *5*, 950–954.
4. Doherty, E.A., Batey, R.T., Masquida, B., and Doudna, J.A. (2001). *Nat. Struct. Biol.* *8*, 339–343.
5. Silverman, S.K., and Cech, T.R. (1999). *Biochemistry* *38*, 8691–8702.
6. Murthy, V.L., and Rose, G.D. (2000). *Biochemistry* *39*, 14365–14370.

Chemistry & Biology, Vol. 11, February, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.02.005

Finding Cinderella after the Ball: A Three-Hybrid Approach to Drug Target Identification

A major bottleneck in drug discovery is identifying the targets of small molecules. The yeast three-hybrid assay extends the two-hybrid approach to screen for protein-small molecule interactions. In this issue of *Chemistry & Biology*, GPC Biotech reports the first application of this promising assay [11].

Drug developers face a problem similar to that of the Prince in the Cinderella story. They have a small molecule that produces the desired physiological response, but they do not know the protein target to which the small molecule binds, the Cinderella who fits the glass slipper. Traditionally, the protein targets of small molecule ligands have been identified using *in vitro* methods such as affinity chromatography and photoaffinity labeling. The ligand is derivatized so that it can be linked to a resin or photoreactive group. The modified ligand is then incubated with a crude cell lysate, and the labeled protein is finally identified by N-terminal sequencing. These methods have been integral to target discovery, but they are laborious and subject to low protein expression levels, protein degradation during cell lysis, or insufficient affinity for the small molecule ligand.

Genomics and proteomics are beginning to provide real alternatives to these traditional methods for drug target identification. Expression profiling using DNA microarrays enabled the identification of the microbial protein and pathway targets for isoniazid, the primary drug used to treat tuberculosis [1], and potential mechanisms of resistance. A complementary approach using a “syn-

thetic lethal” small molecule screen allowed for the identification of proteins and pathways in yeast that are involved in DNA synthesis and repair [2]. Borrowing from a traditional genetic approach, the small molecules camptothecin and hydroxyurea were screened for lethality against a library of single-knockout mutants in yeast. Uncharacterized open reading frames involved in DNA repair were then inferred by comparing the small molecules’ patterns of lethality to those of all double mutants. Alternatively, Snyder and coworkers used protein chip technology to print the >6000 proteins found in *S. cerevisiae* on glass slides and assay them for binding to phospholipids [3].

In 1996, Licitra and Liu extended high-throughput yeast two-hybrid methods to small molecule target identification [4]. In the two-hybrid assay, protein-protein interactions are detected as reconstitution of a transcriptional activator from its DNA binding (DBD) and activation domains (AD) [5]. The assay can be run on a genome-wide scale by creating a library of activation domain-cDNA clones [6]. The three-hybrid assay extends this approach to small molecule-protein interactions by dimerization of two receptor proteins via a bridging heterodimeric ligand [7] (Figure 1). One ligand-receptor pair serves as an anchor, while the other ligand-receptor pair is the small molecule-protein interaction of interest. As a proof of principle, Licitra and Liu used a dexamethasone-FK506 heterodimer to isolate FK506-binding protein 12, the known target of FK506, from a Jurkat cDNA library. This approach allows the cDNA clones to be expressed at uniformly high levels and without the need for purification, and the target protein’s identity to be simply read out from the cDNA-AD sequence at the end of the selection. Despite obvious applications for drug discovery, however, there were no further reports of using the three-hybrid assay for target discovery beyond this proof of principle experiment.

One thought was that the affinity of the ligand-recep-