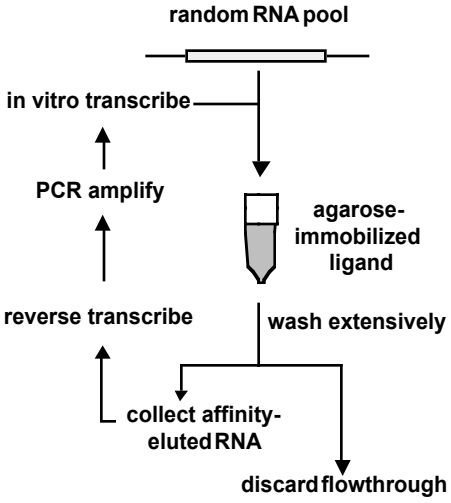
Understanding the basis for RNA folding and molecular recognition: atomic resolution structures of RNA aptamers complexed with their ligands

Django Sussman, Jay Nix, and Charles Wilson

Department of Biology and Center for the Molecular Biology of RNA, University of California, Santa Cruz, California 95064, USA

INTRODUCTION

Our understanding of the molecular basis for a wide variety of biochemical processes has largely come about through the determination and analysis of macromolecular structure. While more than 5000 atomic-resolution protein structures have been determined by these methods, only a handful of RNAs have been studied at the same resolution. The process of *in vitro* selection (outlined in the inset) has made it possible to evolve many different types of small, well-defined RNA motifs with unique binding activities ('aptamers')[1,2]. In a typical SELEX experiment, a pool of random sequence molecules (derived from a random chemical DNA synthesis) encounters a selective pressure that physically separates functional from non-functional molecules (e.g., binding to an affinity column). Repeated rounds of selection followed by enzymatic amplification can be used to isolate RNAs that are highly adapted to specific functions. Crystallographic analysis of such RNAs complexed with the small molecules they bind can potentially provide important insights into the basic mechanisms by which RNAs fold into specific three-dimensional structures and the rules that define how RNAs and small molecules interact. X-ray diffraction experiments carried out at ALS beamline 5.0.2 in September, 1999, provide the first high resolution views of RNA aptamer-ligand complexes.



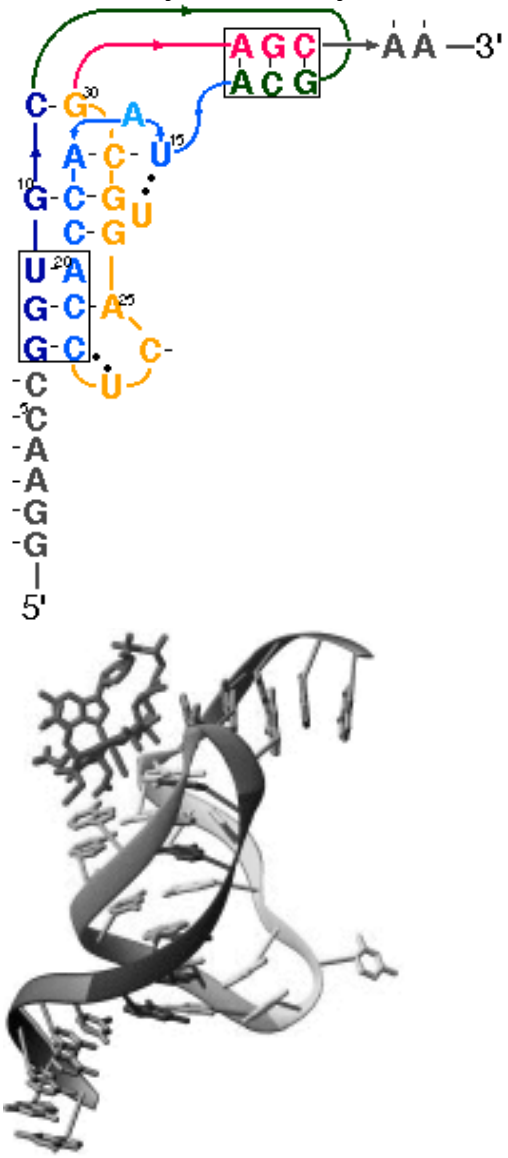
CYANOCOBALAMIN APTAMER

Vitamin B12 and its biologically-active derivatives are large corrin-based cofactors used to catalyze several different classes of enzymatic reactions. Adenosylcobalamin is believed to be one of the most evolutionarily ancient enzyme cofactors[3] and its role in extant ribose reductases has led to the postulate that it played a role in the transition from RNA-based biology to modern biology[4]. Given this background and with the goal of identifying RNAs capable of recognizing large, complex ligands, Lorsch and Szostak used SELEX to isolate RNA motifs that specifically recognize cyanocobalamin[5]. A minimal aptamer containing 31 conserved nucleotides was identified that binds the cofactor with high affinity ($K_D \approx 90$ nM) and specificity (distinguishing between different liganded. 1 M lithium chloride was included during the selection experiment with the knowledge that lithium has the potential to from tight complexes with RNA backbone phosphates[6] in the hope that selected RNAs would be isolated that fold into a highly compact and stabilized structures.

The 3 Å crystal structure of the cyanocobalamin aptamer was determined from a single crystal MAD experiment using the cobalt associated with the corrin ring in the complex as the anomalous scatterer from which crystallographic phases could be deduced. Analysis of the structure reveals that it folds as a pseudoknot whose first helix is embedded within a much larger triplex.

Perpendicular stacking of the triplex and the second helix creates a cleft that functions as the vitamin B12 binding site. Chemical modification studies by Lorsch and Szostak can be interpreted in light of the structure and indicate that the bound and unbound conformations of the aptamer are likely to be fairly similar (most changes in chemical modification upon binding correspond to the protection of residues at the binding interface). Of the 27 nucleotides in the aptamer core (excluding flanking nucleotides required for expression and crystallization), only 4 in helix 2 can be considered to form what would generally be considered conventional secondary structure. The remainder form a variety of interactions that combine to either build the triplex or lock the second helix in position.

Schematic representations of the cyanocobalamin aptamer illustrating its secondary and tertiary structure.



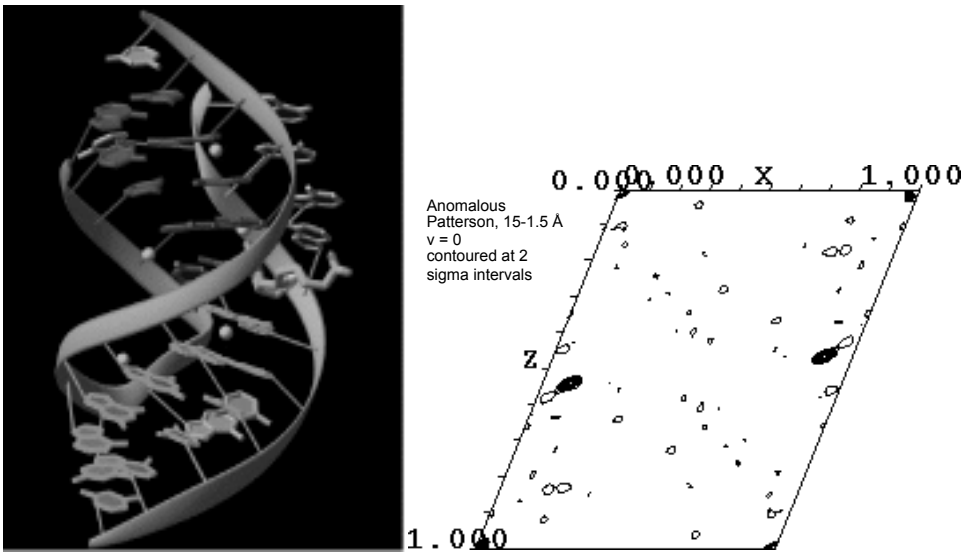
The triplex provides a rich set of tertiary interactions for analysis. Five of the seven tiers in the triplex include base triples stabilized by specific hydrogen bonding. The two triples at the bottom of the triplex (i.e. farthest from the vitamin B12 binding site) are defined by nucleotides packing against the minor groove face of Watson-Crick base pairs. Lying above them in the triplex are two Watson-Crick pairs which mediate a switch in base pairing from strands 1-2 to strands 2-3. As a result of this switch, base triples in the top of the triplex are generally arranged with a nucleotide hydrogen bonding instead to the major groove side of a base pair.

Vitamin B12 binds in the cleft formed at the junction between the triplex and the second helix and makes an extensive set of specific interactions. Methyls protruding axially and equatorially from one side of the corrin ring contact either side of the cleft and make van der Waals contacts with the exposed nucleotide faces of the base triple and base pair that cap the triplex and helix respectively. Supplementing these and additional hydrophobic interactions are a set of four specific hydrogen bonds that bridge the amide side chains of the cofactor and the nucleotides of the RNA.

BIOTIN APTAMER

We previously isolated and characterized an RNA motif that specifically recognizes biotin, a ubiquitous carboxylation co-factor used by protein enzymes[7,8]. A 31-nucleotide pseudoknot, closely resembling the ribosomal frameshifting elements in retroviral mRNAs, is present in all selected RNAs enriched by affinity chromatography. Independent clones containing the pseudoknot bind biotin with relatively high affinity ($K_D \approx 5\ \mu M$) and high specificity. Binding with this avidity is surprising given the absence of charges and aromatic rings on the ligand, functional groups which often provide the driving force for aptamer binding.

The 1.5 Å crystal structure of the biotin aptamer was determined by MAD phasing using the selenium of bound selenobiotin as an anomalous scatterer. An anomalous Patterson map (inset) clearly shows a single corresponding to the bound cofactor. The refined structure reveals a coaxially stacked pseudoknot, fitting well with previous models constructed on the basis of biochemical and genetic constraints. The arrangement of helices in the aptamer is similar to that observed by NMR and crystallography for viral pseudoknots that induce ribosomal frameshifting. Biotin binds at the interface between the two helices in a pocket defined by the base pairs capping each helix and by the 3'-end of an A-rich loop running the length of the minor groove of helix 1.



REFERENCES

1. A. D. Ellington and J. W. Szostak, "In vitro selection of RNA molecules that bind specific ligands," Nature **346** (6287), 818-22 (1990).
2. C. Tuerk and L. Gold, "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase," Science **249** (4968), 505-10 (1990).
3. N. H. Georgopapadakou and A. I. Scott, "On B12 biosynthesis and evolution," J Theor Biol **69** (2), 381-4 (1977).
4. S. A. Benner, A. D. Ellington, and A. Tauer, "Modern metabolism as a palimpsest of the RNA world," Proc Natl Acad Sci U S A **86** (18), 7054-8 (1989).
5. J. R. Lorsch and J. W. Szostak, "In vitro selection of RNA aptamers specific for cyanocobalamin," Biochemistry **33** (4), 973-82 (1994).
6. F.A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry* (John Wiley and Sons, New York, 1988).
7. Charles Wilson, Jay Nix, and Jack W. Szostak, "Functional requirements for specific ligand recognition by a biotin-binding RNA pseudoknot.," Biochemistry **37** (41), 14410-14419 (1998).
8. C. Wilson and J. W. Szostak, "In vitro evolution of a self-alkylating ribozyme," Nature **374** (6525), 777-82 (1995).

This work was supported by a grant from the National Institutes of Health (GM52707), the National Science Foundation (MCB-9876350), and an award from the David and Lucile Packard Foundation.

Principle investigator: Charles Wilson, Department of Biology and Center for the Molecular Biology of RNA, University of California at Santa Cruz. Email: wilson@biology.ucsc.edu. Telephone: 831-459-5126.