Discovery of a thiamin-utilizing α-keto acid decarboxylase ribozyme: Implications for RNA's role in primordial metabolism

by

Paul Cernak

B.Sc., Portland State University, 2007

Thesis Submitted In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department Molecular Biology & Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY

Fall 2013

Approval

Name: **Paul Cernak**

Degree: **Doctor of Philosophy**

Title of Thesis: ***Discovery of a thiamin-utilizing α-keto acid decarboxylase ribozyme: Implications for RNA's role in primordial metabolism***

Examining Committee: **Chair:** Jenifer Thewalt
Professor

Dipankar Sen
Senior Supervisor
Professor

Andrew Bennet
Supervisor
Professor

Edgar Young
Supervisor
Associate Professor

Peter Unrau
Supervisor
Associate Professor

David Vocadlo
Internal Examiner
Professor
Department of Chemistry

Frank Jordan
External Examiner
Professor, Department of Chemistry
Rutgers University

Date Defended: September 19, 2013

Partial Copyright Licence

Abstract

Vitamins are hypothesized to be relics of an RNA World, and likely participants in an RNA-mediated primordial metabolism. If catalytic RNAs could harness vitamin cofactors to aid their function, in a manner similar to enzymes, it would enable ribozymes to catalyze a much larger set of chemical reactions. The cofactor thiamin diphosphate, a derivative of vitamin B1 (thiamin), is used by enzymes to catalyze difficult metabolic reactions, including decarboxylation of stable α-keto acids such as pyruvate. Here I report a ribozyme that uses free thiamin to decarboxylate a pyruvate-based suicide substrate (LnkPB). Thiamin conjugated to biotin was used to isolate catalytic individuals from a pool of random sequence RNAs attached to LnkPB. Analysis of a stable guanosine adduct obtained via digestion of an RNA sequence (clone dc4) showed the expected decarboxylation product. Discovery of a prototypic thiamin-utilizing ribozyme has implications for RNA's role in orchestrating early metabolic cycles.

Keywords: RNA World; ribozyme; thiamin diphosphate; α-keto acid; decarboxylation; coenzyme.

Dedication

I dedicate this thesis to my wife, Swetha Mohan.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). I would like to thank Dr. Robert Young and the members of his group for advice with organic synthesis; Drs. Andrew Bennet, Peter Unrau, Edgar Young, Andrew Lewis, and Mr. Hongwen Chen for their expert and practical advice. I would also like to thank Sen lab, past and present. Special thank you to Dr. Dipankar Sen for support during this project, through thick and thin.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AP	4-Aminopyrimidine
ATP	Adenosine Triphosphate
ATP	Adenosine Triphosphate
C2H	Thiazolium C2 proton
CD	Circular Dichroism
CD	Circular Dichroism
CoA	Coenzyme A
CTP	Cytosine Triphosphate
D_2O	Deuterium Oxide
ddH_2O	Double Distilled Water
DMF	Dimethyformamide
DNA	Deoxyribonuclic Acid
dsDNA	Double Stranded Deoxyribonucleic acid
ESI	Electrospray Ionization
FAD	Flavin Adenine Dinucleotide
FeS	Iron (II) Sulfide
FeS	Iron (II) Sulfide
GTP	Guanosine Triphosphate
HBThDP	C2α-HydroxybenzylThDP
HDV	Hepatitis Delta Virus
HEThDP	C2□-Hydroxyethythiamin Diphosphate
HEThDP	C2□-Hydroxyethythiamin Diphosphate
HPLC	High Performance Liquid Chromatography
IP	1′,4′-imino tautomer of 4′-aminopyrimidine
IP	1′,4′-imino tautomer of 4′-aminopyrimidine
IPA	Isopropanol
LThDP	C2□-Lactylthiamin Diphosphate
LThDP	C2□-Lactylthiamin Diphosphate
NAD+	Nicotinamide Adenine Dinucleotide
NMR	Nuclear Magnetic Resonance
PAGE	Polyacrylamide Gel Electrophoresis

PCR	Polymerase chain reaction
PDHc	Pyruvate Dehydrogenase Complex
RNA	Ribonucleic Acid
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
SFU	Simon Fraser University
StAv	Streptavidin
TB	Thiamin-PEO_3-biotin
TFA	Triflouroacetic acid
Th	Thiamin
ThBi	Thiamin-PEO_3-biotin
ThDP	Thiamin Diphosphate
ThTTDP	Thiathiazolone Diphosphate
TK	Transketolase
UTP	Uridine Triphosphate
VS	Varkud Satelite

Preface

The RNA World hypothesis suggests that life began or transitioned through an era when ribonucleic acid (RNA) was the principal polymer of life. RNA is proposed to have had a dual function, providing both information storage, necessary for transmission of naturally occurring advancements (DNA's role in life today), and catalysis (typically protein enzyme's role in life today), fundamental to the continued flux of energy, necessary to maintain life. Knowledge of how RNA was able to perform catalytic processes necessary to maintain sufficient concentrations and flux of high energy starting materials necessary for even the most minimal requirements of the RNA World, such as self-replication, is severely limited. For the most part, this is because of the incredible diversity of chemical reactions necessary to maintain metabolism. However, due to the relatively homogeneous structure and limited catalytic components of RNA, it is uncertain how many of these reactions it can physically perform. In this regard protein enzymes have a significant advantage over RNA enzymes (ribozymes) because of amino acid diversity, but even protein enzymes use reversibly bound small organic molecules (coenzymes) that aid in their chemical catalysis.

It has been proposed that in an RNA World, RNA may have used these same coenzymes, which were then taken over by protein enzymes (White, 1976). Testing this hypothesis is impeded by the absence of these postulated ribozymes of the RNA World. In this dissertation, I have overcome this limitation and tested this hypothesis by selecting for functional RNAs that perform the same role as a putative metabolic ribozyme of the RNA World. Thiamin diphosphate (ThDP), due to its central role in modern metabolism, catalytic nature and the diversity of reactions it can perform, is a potential coenzyme that could be utilized in the RNA World. I set out to find a ribozyme that uses thiamin in a reaction central to all extant life today, the decarboxylation of an α-keto acid.

1. Introduction

1.1. Heterotrophic hypothesis: Life from a "prebiotic broth"

The heterotrophic origin of life or the heterotrophic hypothesis, first proposed by Oparin and Haldane in the 1920-30s, suggests that life began from a "prebiotic broth" that contained the key building blocks of life (amino acids, purines, pyrimidines, and sugars) which were already present on the early earth, resulting from natural processes. Life began, according to this hypothesis, by the self-organization of these key molecules. Electrical discharge experiments in 1957 by Stanley Miller, of a sealed vessel containing CH_3, NH_3, H_2O and H_2 (simulating a reducing atmosphere), yielded moderately high concentrations of amino acids (25 amino acids were identified) as well as other biologically important metabolites and acids totalling ~15% yield based on carbon (Miller, 1953). It was suggested that many of the compounds were products produced from cyanide ions, made from reactions of methane and ammonia, methane and nitrogen, and/or mixtures of carbon monoxide, nitrogen and hydrogen (Miller, 1955). These experiments suggested that the proposed "prebiotic broth" is not only plausible but also that experiments can be designed to provide limits to what can and can not be synthesized under plausible prebiotic conditions. This opened the field of prebiotic chemistry.

1.1.1 Prebiotic syntheses of nucleobases

Since the Miller experiments, prebiotic syntheses of chemical compounds necessary for ribonucleic acid (nucleotides, nucleosides and ribose sugar) and proteins (many amino acids) have been reported, many of which require cyanide or its derivatives. Notably, synthesis of the purine nucleobase, adenine, by polymerization of concentrated solutions of ammonium cyanide has been shown **(Figure 1.1)** (Oró, 1960; (ORO & Kimball, 1961; 1962). This synthesis was shown to progress more readily when irradiated with a photochemical rearrangement (Sanchez, Ferris, & Orgel, 1967; 1968).

Small molecules such as formaldehyde and other aldehydes catalyze this reaction (Schwartz & Goverde, 1982; Voet & Schwartz, 1983). Other purines (guanine, xanthine, hypoxanthine, isoguanine, and diaminopurine) can also be produced under prebiotic conditions starting with more complex materials, such as aminoimidazole carbonitrile and aminoimidazole carboxamide, reacting with hydrogen cyanide, cyanoacetylene or cyanate (Sanchez et al., 1968; Sanchez, Ferris, & Orgel, 1966b). The prebiotic synthesis of pyrimidines has also been reported using cyanoacetylene, a major product in electric discharge experiments containing methane and nitrogen (Sanchez, Ferris, & Orgel, 1966a) or cyanoacetaldehyde, reacting with cyanate or urea, respectively (Robertson & Miller, 1995). Clearly, cyanide and its derivatives are important molecules for prebiotic syntheses and their presence on the early earth is crucial to the early formation of nucleic acid bases.

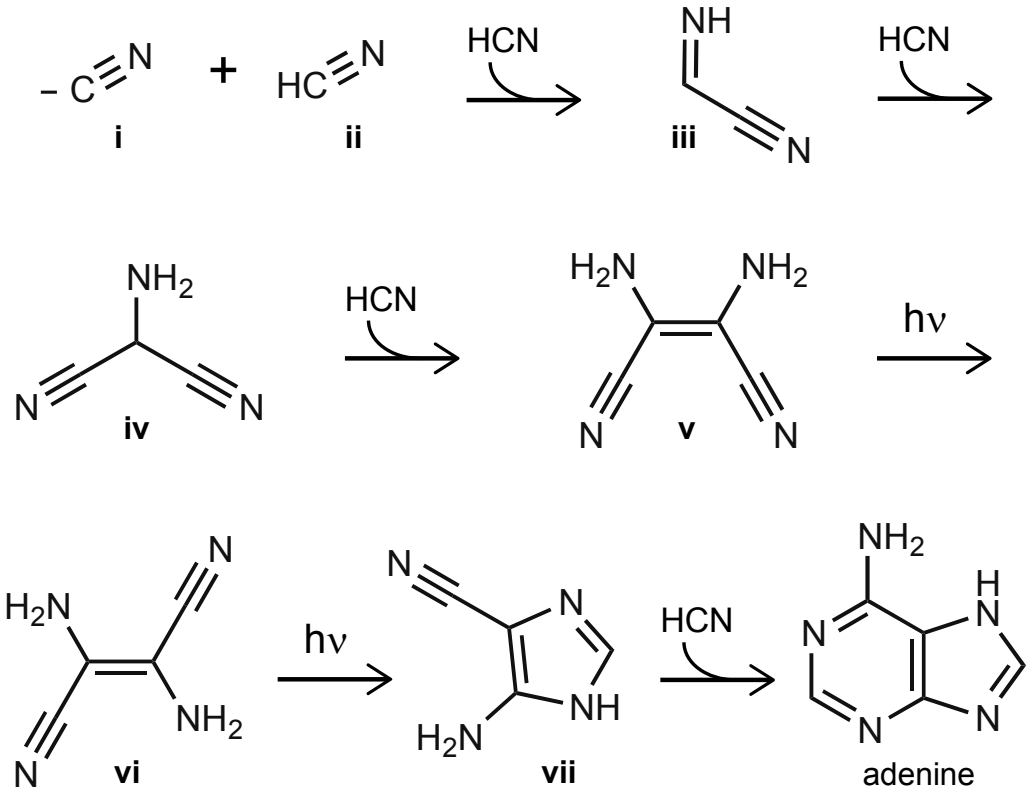

Figure 1.1. Polymerization of ammonium cyanide to form adenine

Structures are (i) cyanide (ii) hydrogen cyanide (HCN) (iii) HCN dimer (iv) aminomalonitrile (v) diaminomaleonitrile (vi) diaminofumaronitrile and (vii) aminoimidazole carbonitrile.

1.1.2 Prebiotic syntheses of ribose

Synthesis of the sugars is a particularly interesting example of the difficulties inherent in prebiotic syntheses. The heterotrophic hypothesis suggests that the necessary components of life were available on an early earth; therefore ribose, a key component of nucleic acids, would most likely be available or present in the "prebiotic broth." Sugars have been known, for over 150 years, to be readily made from the polymerization of formaldehyde, called the formose reaction, using a base catalyst such as calcium hydroxide and high temperatures (Butlerow, 1861). Ribose, the sugar component of nucleic acid, has been produced and characterized from the formose reaction. However, the formose reaction shows very little selectivity and produces more than 40 different sugars (Decker, Schweer, & Pohlamnn, 1982). Clays have been shown to catalyze the reaction with low concentrations of formaldehyde, although with low yields (Gabel & Ponnamperuma, 1967). Moreover, ribose reacts 4-16x faster than other pentoses and hexoses (Dworkin & Miller, 2000), making it particularly unstable with a half-life of only 73 min at 100°C (or 44 years at 0°C) under neutral pH conditions (Larralde, Robertson, & Miller, 1995). This suggests its existence would only be fleeting on an early earth. Springsteen and Joyce found that perhaps ribose's increased reactivity could have worked in its favour, and showed that it preferentially reacts with cyanamide to give a cyclic product that readily crystalizes out of the aqueous solution (Springsteen & Joyce, 2004). It was proposed this crystallization may have helped in enriching the heterogeneous mixture of sugars, produced from the formose reaction, on an early earth.

Recently, John Sutherland and coworkers reported the simultaneous synthesis of the cytosine base, sugar, and phosphate to make pyrimidine nucleotides, circumventing the need to attach the three in subsequent steps. The reaction proceeds through an arabinose amino-oxazoline and anhydronucleoside intermediates **(Figure 1.2)** from the starting materials: cyanamide, glyceraldehyde, glycolaldehyde, cyanoacetylene, and inorganic phosphate. The phosphate was shown to have a surprising role as a nucleophilic and general acid/base catalyst (Powner, Gerland, & Sutherland, 2009).

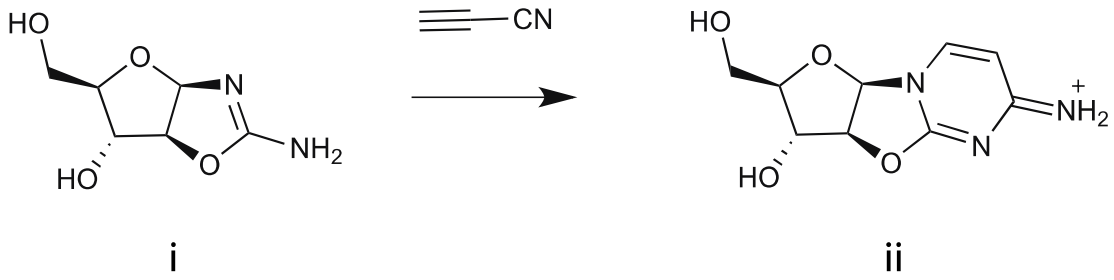

Figure 1.2. Cyclic intermediates in the synthesis of cytidine 2',3'-cyclic phosphate

Reactions of cyanamide, glycoaldehyde and glyceraldehyde produce the amino oxazoline (i) which reacts with cyanoacetylene to give the anhydroarabinonucleoside (ii).

1.2. Role of Metabolism in the Origins of Life

1.2.1 The thioester world

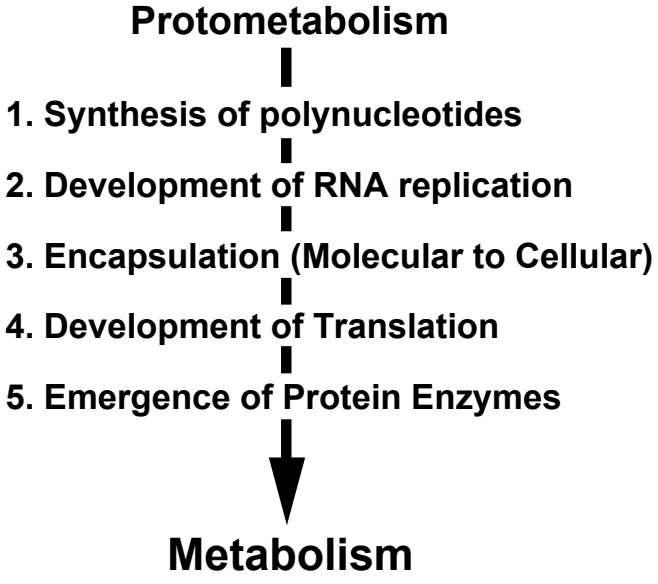

Figure 1.2. Five stages from protometabolism to metabolism proposed by De Duve

Hypotheses on the origins of life often take the assumption that critical elements of extant life are direct descendants of precursor elements, which are derived from the origin of that functionality. In other words, the functional elements (such as thioesters

and high energy phosphate groups) are proposed to be conserved through evolution, and only modest changes have been made to the most fundamental aspects of biology. Exemplifying present day biology, Christian De Duve argues we need to find congruence between our hypotheses on the origins of life and current biological chemistry. He suggests life proceeded through a protometabolism, which transitioned through 5 stages to reach what is known today as metabolism **(Figure 1.2)**. He postulates that the synthesis of polynucleotides, necessary for an RNA World, were most likely the result of a protometabolism, and that extant biology still contains significant vestiges of such a protometabolism, evidenced in present-day metabolic cycles (see White's arguments in **Section 1.5**). He proposes that thioesters, in particular, had a fundamental role in such a protometabolism, based on the central and pervasive roles of thioesters such as acetyl coenzyme A (acetyl-CoA) in present-day metabolism. Acetyl-CoA acts as an acyl donor, bridging the citric acid cycle with glycolysis, gluconeogenesis, the glyoxylate cycle, fatty acid synthesis, as well as serving as a high energy starting material for the synthesis of porphyrins and sterols.

Modern biology uses the energy contained in triphosphate bonds of ATP and other nucleotide triphosphates to push energetically uphill reactions including, notably the polymerization of nucleic acids by protein polymerases. The thioester world hypothesis suggests that since thioesters, which have the same free-energy of hydrolysis as the pyrophosphate bonds of ATP (-50 kJ/mole under physiological conditions), may be high-energy components of ancient metabolic cycles. However, if we use the same congruence argument, life on early earth must have also used high-energy phosphates as modern life does today. This discrepancy can be resolved by the conversion of thioesters to pyrophosphate or triphosphates **(Figure 1.3)**. This conversion has been experimentally demonstrated by mildly heating (50° C) N,S-diacetylcysteamine, at neutral pH, with inorganic phosphate. The reaction was shown to progress with divalent salts and imidazole which presumably serve as catalysts (Weber, 1981). The reaction was further shown to be catalyzed by the calcium mineral hydroxyapatite at ambient temperature with increased yields (Weber, 1982).

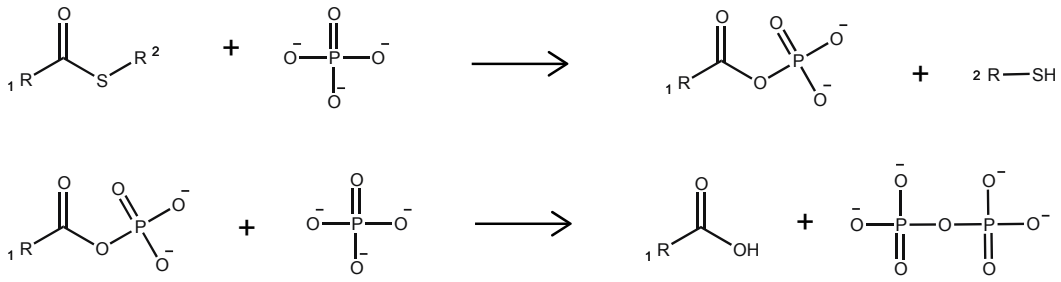

Figure 1.3 Prebiotic synthesis of pyrophosphate from thioesters

The thioester world hypothesis relies on the existence of a prebiotic metabolic cycle that quickly merged into present-day biochemistry, leaving the underlying framework of the precursor functions intact. However, an alternative hypothesis has been proposed, suggesting modern biochemistry dismantled the original framework, leaving no trace whatsoever (Cairns-Smith, 1982).

1.2.2 The iron-sulfur world

Complementing De Duve's "Thioester World," Günter Wächtershäuser offers a similar, if not farther-reaching, hypothesis for the origins of metabolic cycles. Wächtershäuser maintains that existing life (our **only** example of biochemistry) is defined by the flux (not equilibrium) of its chemical cycles. Therefore, life must have begun with metabolism. The machines that drive and control metabolism (RNA and protein) and the information storage polymer (DNA) are ancillary to the cycles producing the building materials. He suggests, like De Duve, that thioesters were probably significant in early metabolism. However, he proposes that thioacids must have been their functional precursors. In this hypothesis, the sulfhydryl groups of modern biochemistry (e.g. coenzyme A) are predecessors to hydrogen sulphide (Wächtershäuser, 1988; 1990).

Wächtershäuser posits that life began with an autocatalytic alpha cycle, and proposes the cycle was likely the precursor to the reductive (reverse) citric acid cycle, driven by the reducing power of FeS/H_2S in the following general reaction: $FeS + H_2S \rightarrow FeS_2 + 2e^- + 2H^+$. It has been suggested that thioacids could have been produced

under volcanic conditions. First carbon dioxide may react with FeS/H_2S to form thiomethane or higher order sulfhydryls, which subsequently react with carbon monoxide to form the thioacid (R-CO-SH). The chemistry has been produced in the laboratory, suggesting the predictions are feasible (Heinen & Lauwers, 1996; Huber, 1997).

These ideas have lead to hypotheses of a chemoautotrophic origin of life. This may have seemed unreasonable at a time when our only known autotrophic life produced food from photosynthesis (a very complex light harvesting technology to fix CO_2—thought by many to be very unlikely in the origins of life). However, the discovery of chemolithoautotrophic archaea living off the H_2, CO_2, S°, and H_2SO_4 produced by the strongly reducing environments of volcanic vents and pools, demonstrate life can exist based on chemical, as well as light energy. In fact, hyperthermophilic prokaryotes dominate short branches near the root of the phylogenetic tree, with few exceptions, suggesting an autotrophic common ancestor (Woese & Fox, 1977; Woese, Kandler, & Wheelis, 1990).

1.3. The RNA World Hypothesis

RNA plays fundamental and central roles in nearly all aspects of present-day biology. Since the proclamation by Frances Crick of the “central dogma of molecular biology,” where RNA was initially given a small role in merely providing a message for the synthesis of protein, RNA has been found to dictate nearly every fundamental aspect of life. In the 50 years since, RNA has been discovered to be the catalytic component of the ribosome for the synthesis of protein (Ban, Nissen, Hansen, Moore, & Steitz, 2000; Cech, 2000), and to play diverse roles in gene regulation by RNA splicing(Cech, Zaug, & Grabowski, 1981; Kruger et al., 1982); performing as a molecular switch (Edwards, Klein, & Ferré-D'Amaré, 2007; Tucker & Breaker, 2005; Vitreschak, 2004; Winkler, Nahvi, & Breaker, 2002); or, at the heart of sophisticated regulation machinery such as in RNAi. Coenzymes, which play critical roles in catalysis by protein enzymes, largely consist of RNA or RNA-like components. Cumulatively, these roles have suggested an RNA-based ancestor, in which RNA performed the tasks of both DNA (information storage) and protein (catalysis and regulation). Therefore, RNA can be thought of manifesting both a genotype and a phenotype.

This hypothesis has been termed the "RNA World," and although speculated about as early as the 1970s, gained recognition with the discovery of catalytic RNA by Thomas Cech and Sydney Altman in 1981 (after their discoveries of group I self-splicing introns (Kruger et al., 1982) and RNaseP's catalysis by RNA(Guerrier-Takada, Gardiner, Marsh, Pace, & Altman, 1983)). Group I ribozymes, variants of the intervening sequence of the group I intron found to perform in a multiple turnover fashion (Zaug & Cech, 1986), have become an excellent example of RNA's facile ability to make and break phosphodiester bonds. These ribozymes perform a disproportionation reaction where they first bind and cleave an RNA oligonucleotide, transferring its 3′-end to the 5′-end of the ribozyme. A second oligonucleotide then binds to the active site and the ribozyme catalyzes the second transesterifcation reaction, ligating the two oligonucleotides. This reaction has been likened to the cellular extant function of recombination and has been used to reassemble ribozymes as products of the reaction and even itself from inactive oligos (Hayden & Lehman, 2006; Hayden, Riley, Burton, & Lehman, 2005; Riley & Lehman, 2003).

This aforementioned capacity of RNA to catalyze transesterifcation reactions has ultimately led to the discovery of a host of naturally occurring ribozymes with self-cleaving capacity, including, the hepatitis delta virus (HDV) (Sharmeen, Kuo, Dinter-Gottlieb, & Taylor, 1988), Varkud Satellite (VS) (Saville & Collins, 1990), *glm*S(Winkler, Nahvi, Roth, Collins, & Breaker, 2004), hammerhead (Prody, Bakos, Buzayan, Schneider, & Bruening, 1986), and hairpin ribozymes. The self-cleavage mechanisms utilize a well positioned Mg^{2+} (other divalent cations are often acceptable) or an internal nucleobase acting as a general acid/base catalyst. The *glm*S ribozyme is a particularly interesting example of a cleaving riboswitch that uses its ligand, glucosamine 6-phosphate, in acid-base catalysis to cleave the mRNA that codes for enzymes responsible for the ligand's synthesis (Winkler et al., 2004).

1.3.1 Systematic evolution of ligands by exponential enrichment (SELEX)

The RNA World hypothesis suggest that RNA played the role of both protein and DNA, therefore, it is presumed to be capable of activities that no longer exist in present-

day biology since these functions have been taken over by their successors. Using readily available polymerases (DNA/RNA polymerase & reverse transcriptase) and synthetic DNA, novel functional activities can be enriched from random sequence RNA libraries containing a large diversity (10^{13}-10^{15}) of unique sequences to provide a proof of principal. *In vitro* selection or SELEX was first demonstrated by Tuerk and Gold (Tuerk & Gold, 1990) and Ellington and Szostak (Ellington & Szostak, 1990), with the selection of RNA sequences that bind ligands (typically small organic molecules) attached to solid supports. A typical *in vitro* selection scheme starts with a dsDNA library consisting of a random sequence, flanked by 5′ and 3′ constant regions (the 5' element containing a T7 RNA polymerase promoter). RNA polymerase enzyme is used to transcribe the DNA library to RNA. The resulting RNAs are folded in a selection buffer containing salts (typically $MgCl_2$) to neutralize the negatively charged RNA backbone, allowing the RNA to fold into more compact structures. These folded RNAs are incubated with the ligand, which has been covalently attached to a solid support. Unbound RNAs are discarded and the bound RNAs are subjected to a series of washes with the selection buffer. The bound RNAs are then eluted off the solid support with a buffer that disrupts RNA structure, such as those containing urea, and reverse transcribed to give cDNA and then amplified by PCR to yield dsDNA. Several selection rounds are performed until a significant proportion of the RNAs bind to the solid support, and the "winning" RNAs are cloned and sequenced. This methodology was later adapted for the discovery of novel ribozymes and DNAzymes (DNA enzymes) as well as for the *in vitro* evolution or directed evolution of mutagenized ribozyme sequences for expanded activities. Examples include variants of the Tetrahymena group I ribozyme (intervening sequence of the group I intron) evolved to have have activity in the presence of calcium ions (Lehman & Joyce, 1993) or to cleave a DNA substrate (Beaudry & Joyce, 1992). The use of *in vitro* selection for the discovery of aptamers for small molecules has become a routine procedure whenever the ligand can be attached to a solid support.

1.3.2 RNA self-replication

RNA has been known to be able to catalyze transesterification reactions since the discovery of catalytic RNA. Therefore, *in vitro* selection studies have focused on the discovery of RNAs performing these essential reactions, such as a kinase ribozyme able

to phosphorylate its 5' end with ATP (Lorsch & Szostak, 1994). At the most minimalistic level, in an RNA World, RNA must be able to self-replicate. A significant step in this direction was accomplished with selection, *in vitro*, for a ribozyme RNA polymerase able to copy a template up to 13 nucleotides (Johnston, 2001). Improved polymerase variants that can extend up to 95 nt have recently been reported (Wochner, Attwater, Coulson, & Holliger, 2011).

1.4. Ribozymes and Metabolic Chemistry

It has become reasonably clear: RNA is capable of catalyzing the necessary chemistry for both its self-replication and its subsequent translation to protein. However, this assumes an RNA World with adequate concentrations and flux of activated substrates that, from many accounts, can only come from a metabolism. Can RNA perform the reactions necessary to orchestrate a primordial metabolism? RNA starts out with a clear disadvantage because it only contains 4 natural bases, all of which are chemically similar and relatively inert, having pKa's far from neutral, while protein enzymes comprise of 20+ amino acids with polar and non-polar, positively and negatively charged, and nucleophilic side chain functionalities.

1.5. Coenzymes and the RNA World

Given the chemical complexity disparity between RNA and protein, the question quickly becomes whether RNA can utilize coenzymes (loosely held prosthetic groups used by enzymes to aid in chemical catalysis) (**Figure 1.4**) to aid in similar chemistry that is performed by present-day metabolic enzymes? The majority of proteinaceous enzymes, in fact, utilize coenzymes and, without such, could not catalyze the diverse reactions that make up modern metabolism. Coenzymes can be classified into those that are truly catalytic in nature, such as biotin, thiamin diphosphate (ThDP), pyridoxal phosphate, and those that are loosely bound *substrates*, such as coenzyme A (CoA), nicotinamide adenine dinucleotide (NAD) and flavin adenine dinucleotide (FAD) (there are exceptions where sophisticated enzymes utilize NAD and FAD catalytically).

Coenzymes, precisely the *catalytic* coenzymes, compensate for functionality that even the diverse amino acid side chains lack, adding versatility to protein catalysis.

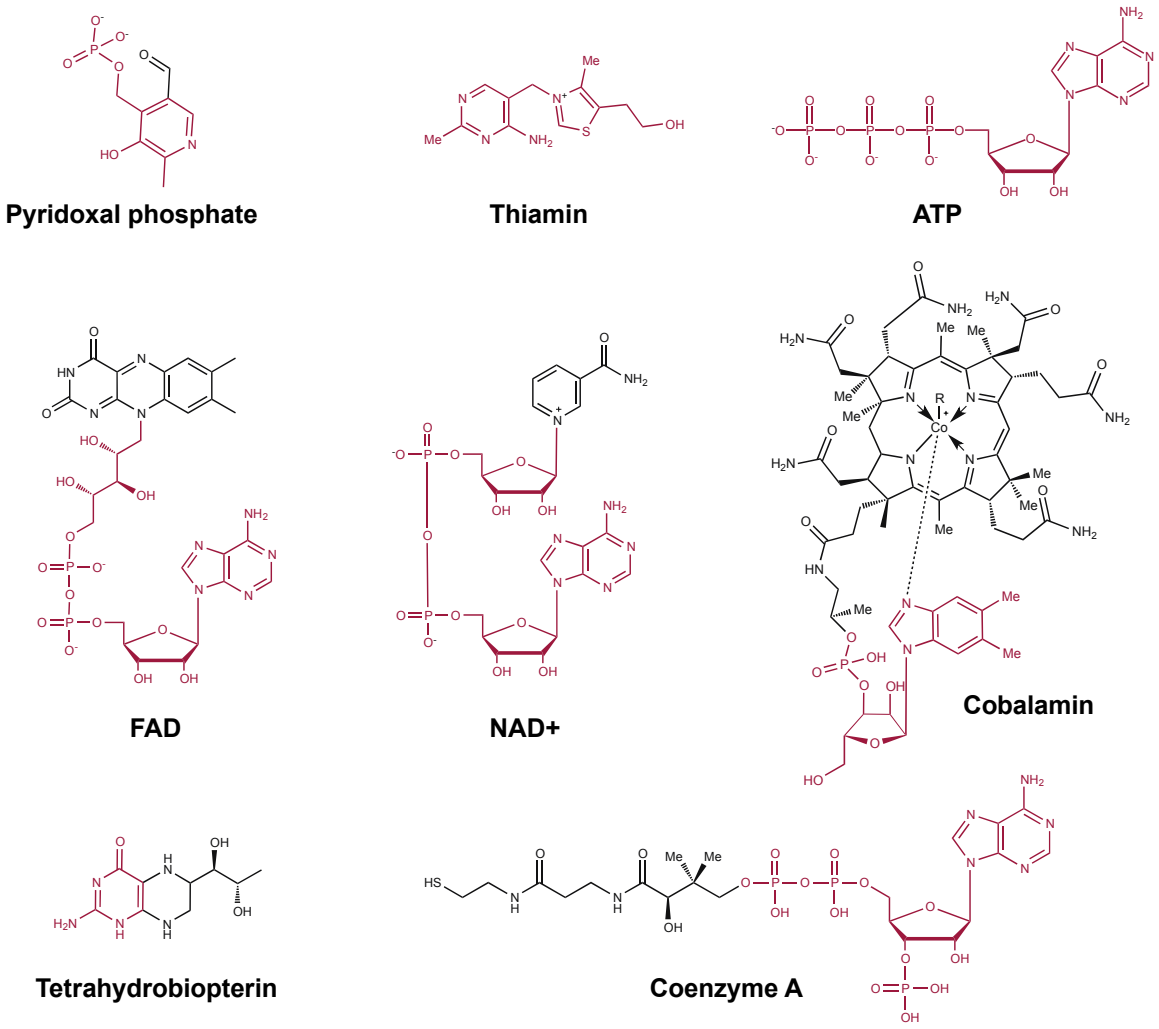

Figure 1.4. Structure of select coenzymes

RNA structural components highlighted in red.

Clues to the origin of coenzymes may be found in the structures of the conenzymes themselves. It has been suggested that coenzymes may be relics from an RNA mediated metabolism. This was first proposed by White in 1976, when observing that many of the coenzymes in present-day biology contain adenosine, ribose, and/or heterocyclic bases resembling RNA in their chemical structures (White, 1976). Sometimes these RNA components have roles in catalysis but often are not directly involved but only provide a scaffold for molecular interactions with the enzyme. The fact

that the RNA components are often ancillary to their catalytic functionality bolsters the hypothesis that protein enzymes borrowed modern-day coenzymes from a time when RNA was the mediator of metabolism and protein enzymes were yet to be established.

The principle of many users

Coenzymes may be exceptionally conserved elements of biology. It has been proposed (Jadhav & Yarus, 2002a) that once coenzymes had a large enough network of utilizing partners, a "critical mass" would be reached where if that coenzyme were to suddenly change, it would be deleterious to the organism because of the number of interactions that would be effected. A complete rearranging of life would have to take place and evolve functionality with the new structure. This effect has been termed the, "principle of many users" and may account for a high level of conservation in the structures of coenzymes.

1.6. In vitro selection for small metabolite chemistry

In vitro selection for RNAs catalyzing small changes to organic molecules has experimental challenges because of the inherent problem that a true catalyst is not changed in the reaction, and therefore cannot be separated away from the inactive sequences in the RNA library. The experimenter must engineering a molecular "handle" to isolate the active sequences from the bulk of the inactive library. This is straightforwardly accomplished when the desired property is phosphodiester self-cleavage and the difference in size/charge of the reacted nucleic acid is significant enough to enable purification of the reacted species on a PAGE gel. When the desired product is a transformation of a small molecule, typically, the random RNA library must be functionalized with the substrate of the desired reaction, while designing a method of enriching the catalytic species (typically using the exceptionally strong streptavidin-biotin interaction). Methods using the interaction of thiolated substrates and mercury PAGE gels have also been successful (Lorsch & Szostak, 1994; Unrau & Bartel, 1998).

To date, there are only a few, yet notable, successes of *in vitro* selections for ribozymes catalyzing reactions with small molecule metabolites. Acetyl-CoA synthase

ribozymes have been isolated from a random library using one reactant attached to the RNA library (acetyl-CoA incorporated at the 5' end of the RNA by a *trans* capping ribozyme) and an exogenous, biotinylated, high-energy acyl-phosphate (Jadhav & Yarus, 2002b). Another thioester synthase ribozyme was reported simultaneously using an almost identical selection scheme (Coleman & Huang, 2002).

Carbon-carbon bond forming ribozymes have been reported. One with Diels Alderase activity (Seelig & Jäschke, 1999) and another with aldolase activity (Fusz, Eisenführ, Srivatsan, Heckel, & Famulok, 2005) using biotinylated maleimide, and a biotinylated aldehyde, respectively. One particularly notable example is the selection of a dehydrogenase ribozyme that acts on an RNA-attached alcohol substrate, forming an aldehyde by converting a freely diffusible NAD^+ to NADH (Tsukiji, Pattnaik, & Suga, 2003). To enrich for catalytic species, the selection used hydrazide attached to biotin, which reacts specifically to the aldehyde product. The reverse of this reaction has also been documented, that is, conversion of an aldehyde to an alcohol using NADH as the hydride source to give NAD^+ as the second product.

2. Thiamin Diphosphate and its Chemistry

2.1. Thiamin Diphosphate decarboxylates α-keto acids like pyruvate

We have been interested in selecting for ribozymes that utilize thiamin (vitamin B_1) for their activity. Thiamin consists of a 4-aminopyrimidine heterocyclic ring attached to a thiazolium ring via a methylene bridge and in the case of the biologically active derivative, thiamin diphosphate (ThDP; see **Figure 2.1a**), a diphosphate functionality. The decarboxylation of pyruvate (an α-keto acid; diagrammed in **Figure 2.1b**, structure ii) using ThDP (**Figure 2.1a**) is a central reaction of modern metabolism, linking glycolysis to the tricarboxylic acid cycle, and this reaction is also important for fatty acid synthesis, anerobic fermentation, and the production of diverse compounds ranging from sugars and amino acids to lipids (**Figure 2.1b**). Chemically speaking, in the absence of thiamin or a comparable catalyst, pyruvate is a compound that remains exceptionally stable to decarboxylation. Unlike the heterolytic decarboxylation of β-keto acids, which undergo facile thermal decarboxylation through a relatively stable enolate intermediate, the analogous decarboxylation of an α-keto acid would transition to an acyl carbanion intermediate (a very high energy species, responsible for negating decarboxylation through this mechanism). Thiamin, or cyanide ion, are able to stabilize the resulting negative charge through delocalization, catalyzing α-keto acid decarboxylation through a covalent, umpolung (polarity inversion), catalytic mechanism.

ThDP utilizing enzymes excel in the decarboxylation of α-keto acids by the activation of thiamine. ThDP utilizing enzymes catalyze a host of other reactions through a common carbanion/enamine intermediate, including, oxidative decarboxylation of pyruvate to form acetate or acetyl-CoA, carbon-carbon bond formation by ketolases or the acylation of CoA in the case of the pyruvate dehydrogenase complex (PDHc). I will

focus mainly on the non-oxidative reaction mechanism of yeast pyruvate decarboxylase (YPDC) that decarboxylates pyruvate to form acetaldehyde.

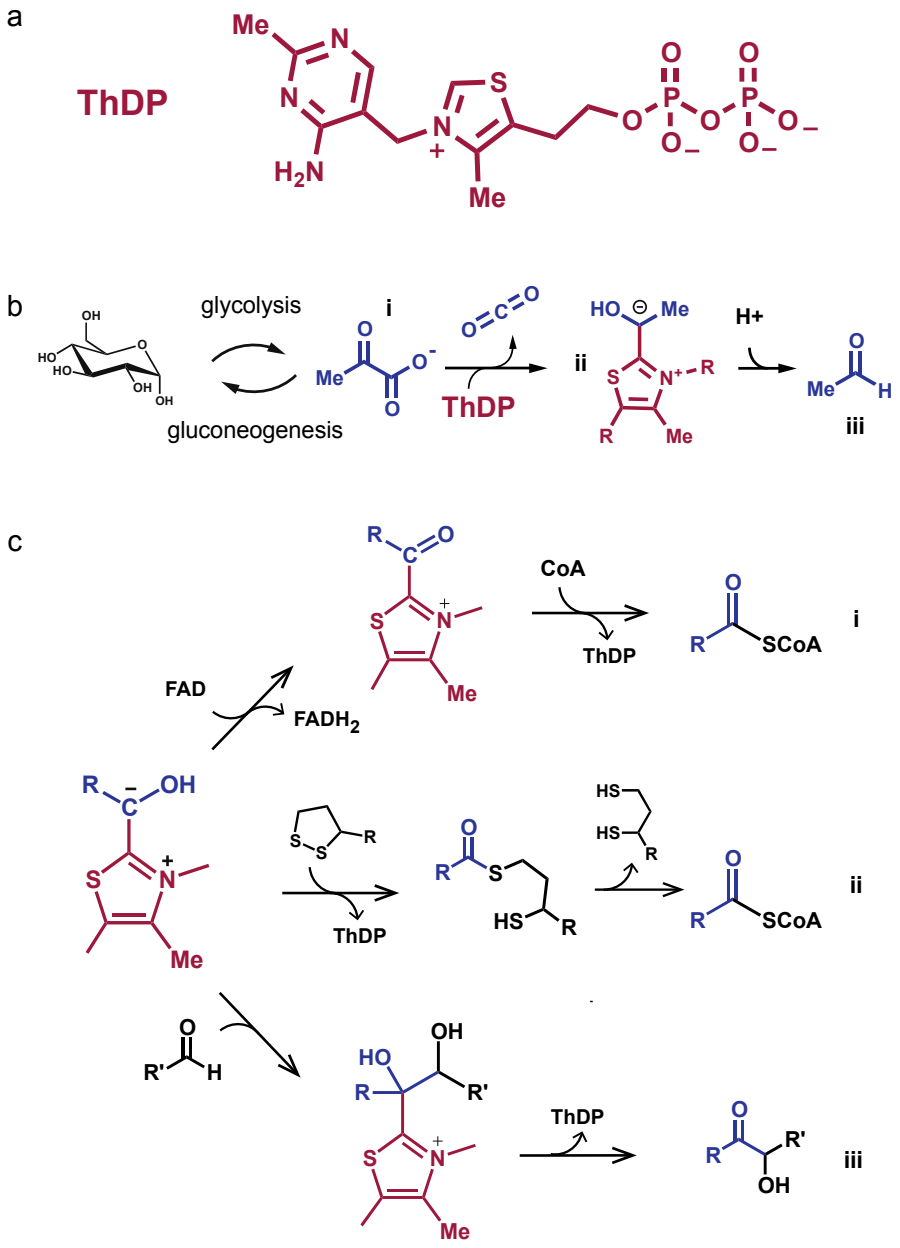

Figure 2.1. Structure of Thiamin diphosphate and its role in modern biochemistry

(a) The structure of thiamin diphosphate (ThDP), the enzymatic cofactor derived from Vitamin B1 (thiamin). It consists of two heterocyclic units, a 2-aminopyrimidine ring and a thiazolium ring, as well as a diphosphate functionality. (b) The roles played by thiamin diphosphate (ThDP) and pyruvate (i) in modern metabolism. Pyruvate is the product and starting compound of major pathways, including glycolysis and gluconeogenesis, essential for cellular energy production.

Thiamin-mediated pyruvate decarboxylation, proceeds through (ii), the central "carbanion/enamine" intermediate. Acetaldehyde (iii) (as catalyzed by pyruvate decarboxyase) is generated as a product. (c) The central carbanion/enamine intermediate is processed by nearly all thiamin enzymes to produce diverse products. (i) acetyl CoA as a produced by pyruvate oxidase. (ii) acetyl CoA as produced by PDHc. (iii) ketol product as catalyzed by ketolases.

Decarboxylation of pyruvate mediated by thiamin or the YPDC holoenzyme, forms a succession of covalent intermediates. The catalytic cycle is schematized in **Figure 2.1c**. First, ThDP (simplified as the thiazolium ring) undergoes deprotonation to form an activated ylide, whose carbanion (located on the C2 position of ThDP's thiazolium ring) attacks the keto functionality of an α-keto acid, in this case pyruvate. α-lactyl-ThDP (LThDP), tetrahedral intermediate initially forms which stabilizes the electron density resulting from decarboxylation by delocalization to the quaternary ammonium to yield the carbanion/enamine intermediate, hydroxyethylidene-ThDP. The carbanion/enamine may be regarded as a "central" intermediate in ThDP catalysis, given that thiamin-utilizing enzymes process it to generate diverse products by ligating, oxidative and nonoxidative reactions. In the case of yeast pyruvate decarboxylase, the carbanion intermediate of ThDP can accept a proton to form hydroxyethyl-ThDP (HEThDP), which decomposes to form acetaldehyde and free thiamine enzyme complex.

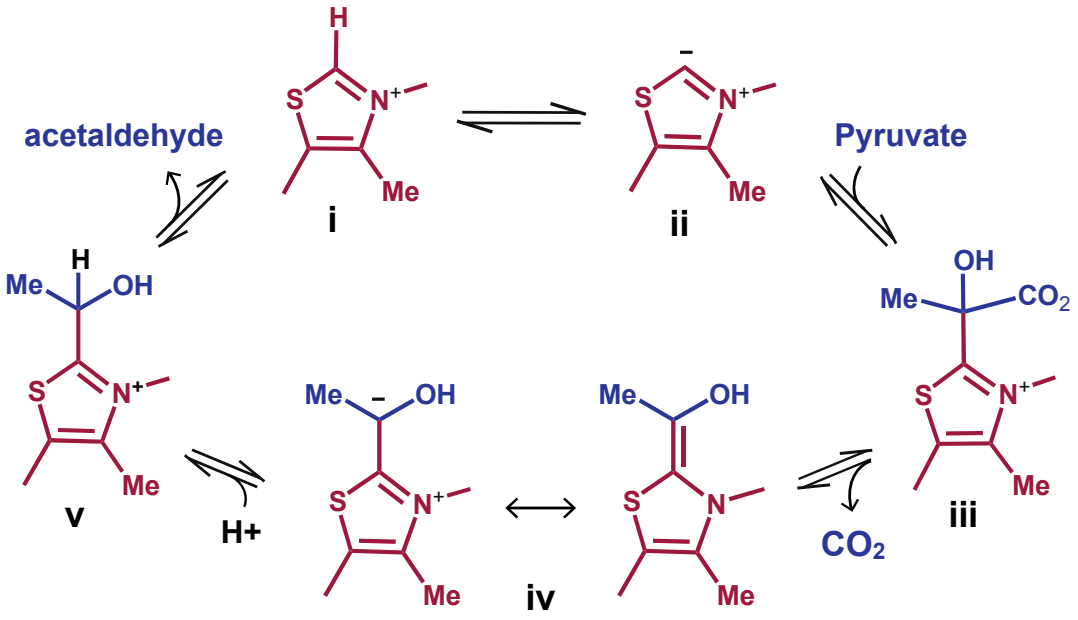

Figure 2.2. Decarboxylation of pyruvate by thiamin or ThDP in PDC

C2 of the thiazolium ring on thiamin or ThDP (i)(red), deprotonates to form the ylide (ii). The ylide form of thiamin performs nucleophilic addition to the carbonyl carbon of pyruvate to form LThDP (iii). Decarboxylation of LThDP results in the carbanion/enamine intermediate (iv). This accepts a proton to from HEThDP which decomposes to acetaldehyde and free thiamin or ThDP.

2.2. Ionization of C2 of the thiazolium ring

Although this mechanism is well known by the biochemistry community and is a rich example of coenzyme catalysis because of the many chemical principals involved, the mechanism of α-keto acid decarboxylation by thiamine and the enzymes that utilze thiamin was a mystery many years after the elucidation of the structure of thiamine in the 1930s. The mechanism was not directly obvious because of the lack of known analogous reactions at the time. It was proposed, that perhaps the thiazolium stabilized a carbanion of the methylene bridge, drawing parallels to cyanide in the benzoin condensation. However, NMR experiments ruled out the exchange of those protons with deuterium oxide (D_2O)(Fry, Ingraham, & Westheimer, 1957). Ronald Breslow, in 1956, while performing NMR on the thiazolium in D_2O and found that the C2 proton (C2H) was particularly acidic for a carbon acid, with a half-life ($t_{1/2}$) of ~20 min. These experiments,

for the first time, strongly suggested C2 of the thiazolum was responsible for catalysis, and inferred a mechanism similar to the benzoin condensation (Breslow, 1957).

pK_a of C2H

Although the C2H was found to be particularly acidic for a proton off a carbon atom, this does not necessarily conclude a low pK_a. To estimate the pK_a of the C2 proton, Kemp and O'Brian used tritiated thiazolium salts (Kemp & O'Brien, 1970). They found the exchange reaction to be general base catalyzed and estimate the Bronsted β value to be nearly 1. This suggests a re-prototonation rate of C2 to be near the solvent diffusion limit and from the respective on/off rates, estimated the pK_a to be from 17 to 20. This suggest that at pH 6, only approximately one hundred in a billion will be in the catalytically active, ylide, state. This demonstrates why ThDP is a noticeably poor catalyst under neutral condition and clearly defines the role of enzymes before the first step of catalysis—stabilize the deprontonated, ylide state of the bound ThDP coenzyme.

2.3. Non-enzymatic catalysis by thiamin

To date, the mechanism of thiamin's entire, non-enzymatic, catalytic cycle has been deciphered, the catalytic rates determined, and all intermediates have been isolated. Deprotonation of C2H of the thiazolium to form the ylide was shown to be general base catalyzed with a β value near 1 (Washabaugh & Jencks, 1988; 1989a; 1989b). After synthesizing and isolating the LThDP intermediate, Kluger et al. (1981) measured the third order rate constant of the base-catalyzed addition of thiamine to pyruvate, reversion of the adduct, and the rate of decarboxylation to form the enamine (Kluger, Chin, & Smyth, 1981). Protonation and deprotonation of the enamine intermediate was determined to be general base catalyzed, with a β value near 1, and the enamine was estimated to have a pKa of ~15 (Washabaugh, Stivers, & Hickey, 1994). The final step, decomposition of the HEThDP, was found to be specific base catalyzed above pH 8. However, at lower pH values, a competing general acid catalyzed cleavage of the ethylene bridge dominates (Kluger et al., 1981; Kluger, Lam, & Kim, 1993). Using these kinetic data, estimates can be made of the rate constants at nearly

any pH by using the Henderson–Hasselbalch equation. These rate constants have played a critical role in determining what function the enzymes play in ThDP catalysis.

2.4. How yeast pyruvate decarboxylase accelerates ThDP catalysis

The high pK_a of C2H makes the non-enzymatic first step of addition to pyruvate rate limiting. Curiously, this step is almost never rate limiting in the enzymes. In efforts to provide evidence for C2 ionization in enzyme-bound ThDP, Kern et al. (1997), measured the H/D exchange of ThDP bound to YPDC and other enzymes, in a manner similar to Breslow's experiments (Kern, 1997). They determined that the enzyme increases the proton exchange rate by 1000 fold over the non-enzymic rate. Additionally, NMR spectra of ^{13}C2 labeled, enzyme bound, ThDP indicated an identical chemical shift to free thiamin. The authors suggested that C2 is largely protonated within the enzyme. However, since they did not decipher any ^{13}C peak corresponding to the C2 ylide within the noisy spectra, no estimation of its pK_a was made.

YPDC accelerates ThDP catalysis a number of ways. The non-enzymatic decarboxylation rate of the LThDP has for some time been known to be influenced by organic solvents, such as ethanol and methanol. This, along with theoretical considerations (based on desolvation effects resulting in the energetically unfavourable transfer of a charged species from a medium with high dielectric constant such as water to a medium with low dielectric constant, reviewed in (Richard & Amyes, 2004)), led to hypotheses that the reduced dielectric constant, non-polar, milieu of the enzyme environment stabilizes the ylide intermediate and may be responsible for the rate acceleration seen from YPDC (Crosby & Lienhard, 1970; Crosby, Stone, & Lienhard, 1970; Kemp & O'Brien, 1970). Reviewed in (Harris & Turner, 2002). Although other contributions are now known, this hypothesis has been supported by preferential solubilisation of the potent inhibitor, thiamin thiazolone (a non-ionic analog of thiamin, weakly soluble in water), to the active site of YPDC (Kluger, Gish, & Kauffman, 1984) and crystal structures of YPDC showing ThDP supported by an isoleucine side-chain. In fact, mutation of this residue to shorter hydrophobic residues decreased the k_{cat} and

increased the pH optimum of the enzyme, suggesting the role of the non-polar milieu of the active site (Guo, Zhang, Kahyaoglu, Farid, & Jordan, 1998).

The Jordan laboratory obtained further evidence for this hypothesis while studying the HEThDP intermediate in the only successful attempt to measure the C2H pK_a of thiamin bound to YPDC (Jordan, Li, & Brown, 1999). For these studies YPDC mutant E91D was used because of its reduced affinity towards ThDP which therefore enabled incorporation of pre-synthesized ThDP intermediates. When C2α-hydroxybenzylThDP (HBThDP) (a well-studied intermediate with a γ_{max} = 380 nm) was added to apo-YPDC, the enzyme was shown to partition not only to the benzaldehyde product but also to the enamine **(Figure 2.3)**. Of the enzyme population containing a ThDP intermediate, nearly all was calculated to be in the enamine, unprotonated form. This suggests a pK_a < 6 of the enzyme bound thiamin C2H, a reduction of approximately 9 pH units from the pK_a in the absence of enzyme (15-16).

How much of the discrepancy between the enzyme bound and unbound thiamin C2H pK_a, can be attributed to the reduced dielectric constant of the enzyme environment? Using thiochrome, an oxidized fluorescent thiamin analogue, Jordan and coworkers attempted to measure the dielectric constant of the YPDC active site. In solvents of lower dielectric constants, thiochrome exhibits a blue-shifted emissions maximum. Using a standard curve made from various alcohols, the emissions maxima of YPDC bound thiochrome diphosphate suggests a dielectric constant in between that of 1-pentanol and 1-hexanol (a dielectric constant of 13-15). It was concluded that this reduced dielectric constant could account for the majority of the pK_a suppression based on free energy calculations (Jordan et al., 1999). Corroboratory evidence was found in the PDHc E1 subunit's ability to also use the HEThDP to form the enamine. The PDHc E1 subunit was shown to increase the rate of enamine formation by a factor of 10^7. Considering that PDHc does not form the HEThDP intermediate in its typical catalytic cycle, this evidence supports the hypothesis that the intrinsic non-polar milieu is responsible for a significant share of the enzymes rate enhancement (S. Zhang, Zhou, Nemeria, Yan, Zhang, Zou, & Jordan, 2005a).

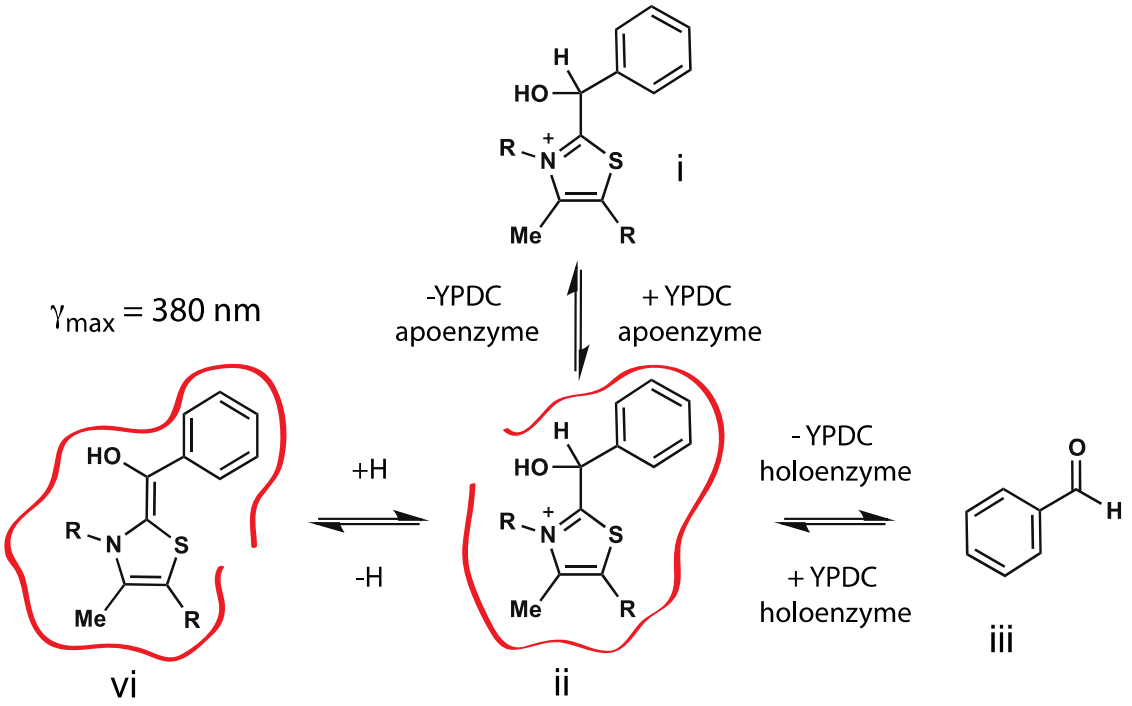

Figure 2.3. Partition of HBThDP to the enamine and benzaldehyde products catalyzed by YPDC

HBThDP (i) incubated with YPDC apoenzyme, forms the enzyme•HBThDP complex (ii) which partitions to the benzaldehyde product (iii) and the enzyme•enamine complex (vi). The enamine has a well defined UV/Vis absorption spectrum with a maximum at 380 nm.

2.5. Role of the 4-aminopyrimidine of ThDP

Although the mechanism of thiamine was a mystery for some time, the function of C2 was evident after the seminal NMR experiments by Breslow, while the isolation of thiamine's catalytic intermediates provided nearly conclusive evidence for the thiazoliums role. This was not true of the other heterocyclic component of thiamine, the 4-aminopyrimidine, which would require a more subtle approach for the elucidation of its catalytic function.

Perhaps the first indication of the role of the 4-aminopyrimidine was the report that N1′-methylthiaminium, a methylated analogue of thiamine (**Figure 2.4**), is an all around better non-enzymic catalyst (Jordan & Mariam, 1978). This analogue is thought to be a mimic of the protonated N1′ form of thiamin and to shift the equilibrium from the

amino tautomeric form toward the 1′,4′-imino ThDP. It was suggested; the N1′-methyl thiaminium's improved catalysis may somehow be attributed to the 4-aminopyrimidine functioning as an acid-base catalyst, abstracting a proton from the C2 position.

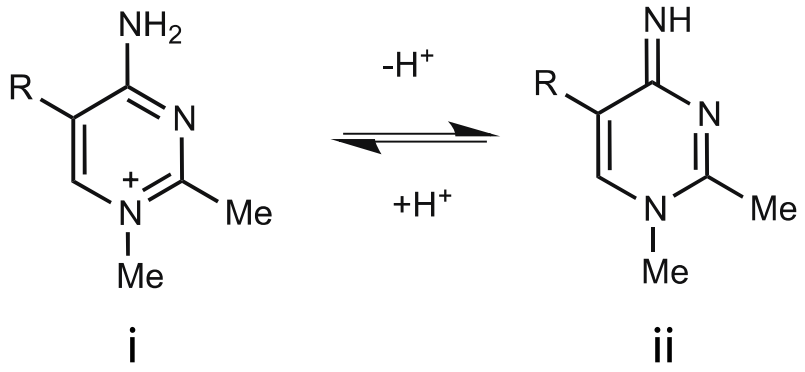

Figure 2.4. N1'-methylthiaminium amino and imino forms

Protonation/deprotonation equilibrium of the amino form of N1'-methylthiaminium (i) and imino form of N1'-methylthiaminium (ii)

The V conformation of ThDP

When the ThDP enzymes, yeast pyruvate decarboxylase, pyruvate oxidase and transketolase, were first crystalized, the bound ThDP was revealed to be bound in a high energy conformation that forms a "V" shape between the two heterocyclic rings (Dyda et al., 1993; Lindqvist, Schneider, Ermler, & Sundström, 1992; Muller & Schulz, 1993; Muller et al., 1993). Crystallized thiamine, in the absence of enzyme, typically arranges to lower energy conformers and almost never this "V" shaped conformation, so why would the enzymes bind ThDP in such a sterically hindered conformation? It was suggested that in this structure lies the modus operandi of ThDP enzymes and how they are able to achieve such an excellent rate enhancement over the non-enzymatic thiamin catalyzed rate. In this so-called V conformation, the exocyclic amine of the 4-aminopyrimidine is intriguingly close to the C2H, about 3.0–3.4 Å away. In fact, this gives little room for the two hydrogen atoms present on the NH_2 and the C2H to be present at the same time. Given these structural details, it was proposed that ThDP enzymes work by an intramolecular proton transfer between the imino tautomer of the 4'-aminopyrimidine and the C2H.

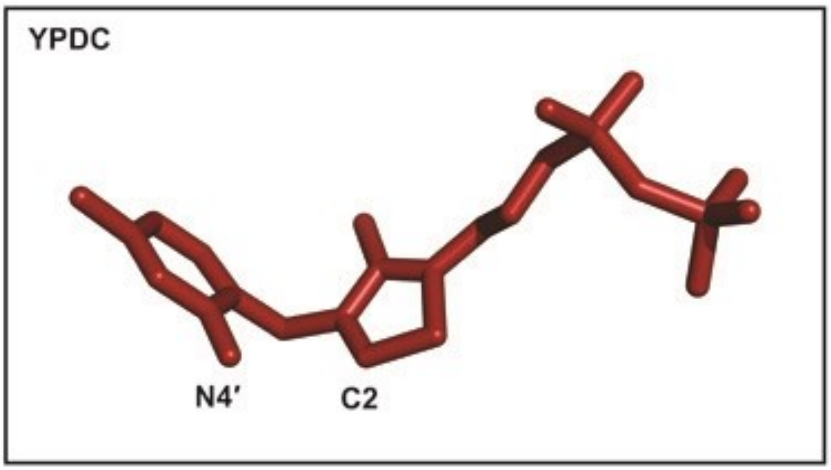

Figure 2.5. V conformation of ThDP bound to YPDC.

Evidence for the 1'-4' iminopyrimidine in ThDP enzymes

Crystal structures of ThDP utilizing enzymes suggest an acid-base role for the 4-aminopyrimidine ring. However, providing sufficient evidence to support this tantalizing hypothesis has been long in the making. The 4-aminopyrimidine UV-Vis spectrum (γ_{max} is ~263 nm under neutral pH conditions) at first appearance does not give a significant "handle" to detect the putative 1′-4′ iminopyrimidine (IP). Any enzyme-induced chromophore in this region of the spectrum would be nearly impossible to detect in the presence of the massive protein absorbance.

Using N1′-methylated 4-aminopyrimidinium salts as a model compound, Jordan and coworkers (Jordan, Zhang, & Sergienko, 2002) found the IP to be stabilized in basic solutions, evidenced by deshielding of the ^{1}H NMR resonances and a red-shifted, UV-Vis spectrum with a new peak appearing at 307 nm. Since the N1′-methylated 4-aminopyrimidinium salts mimic the protonated state of the 4-aminopyrimidine, this suggested enzyme bound ThDP may also have a similar red-shifted absorbance peak if the proposed IP is stabilized by the enzymes.

Circular dichroism (CD) spectroscopy has proven to be exceptionally useful in determining the conformation of the enzyme bound 4-aminopyrimidine. A particularly notable experiment towards elucidation of the putative IP, measuring the CD spectrum of the ThDP analogue, thiamin thiathiazolone diphosphate (ThTTDP), bound to the PDHc

E1 subunit, shows a positive peak at 330 nm (Nemeria et al., 2001). This is where ThTTDP normally absorbs. However, since ThTTDP contains no chiral centre, it was suggested that the enzyme bound “V” conformation is sufficiently chiral to elicit CD signals. This induced chirality of the “V” conformation, along with the assigning of the red-shifted peak of the IP and the use of other model systems, has allowed the detection of the IP using CD difference spectra on a host of ThDP enzymes, with and without bound substrates (Jordan et al., 2002; Nemeria et al., 2004; Nemeria, Chakraborty, et al., 2007a; Nemeria, Korotchkina, et al., 2007b).

3. Selection strategy

I have been interested in providing experimental evidence for support RNAs proposed role in prebiotic metabolic cycles. If RNA were a major orchestrator of prebiotic metabolism in an RNA World, it would require help from catalytic coenzymes to assist difficult reactions that are required to maintain even simple metabolism. ThDP is a particularly interesting example of a catalytic coenzyme, because of its central role in modern metabolism, and its unique chemistry.

To find RNAs that utilize thiamin, I had two ideas. One idea was to use an indirect selection approach by selecting for aptamers (RNA or DNA sequences that bind target molecules) for a transition-state analog. This strategy had been successful in the past for the discovery of a DNAzyme that can perform porphyrin metallation (Y. Li & Sen, 1996). Thiamin thiazolone diphosphate was designed and considered to be a transition-state analog, as it binds with a 10,000x greater affinity to PDHc than does ThDP (Gutowski & Lienhard, 1976). Its high affinity towards the active site was later attributed to desolvation effects and its chemical structure may not be similar to the transition state (Kluger et al., 1984). Such a selection strategy may have been successful nevertheless because, first, thiamin thiazolone crystalizes in the V conformation (supported by a hydrogen bond between the exocyclic amine and the keto oxygen off C2) (Shin & Kim, 1986), and the RNA may have supported this catalytic conformation when bound to thiamin. Second, thiamin thiazolone is uncharged and has limited solubility in water. The RNA aptamer could have provided a hydrophobic interior to accommodate the analog. Such a hydrophobic binding site alone may have lowered the pK_a of a bound thiamine, based on the solvation effect we have seen from PDC.

A direct selection for a post-decarboxylation intermediate was also a possibility, because of the existence of multiple intermediates in the catalytic cycle of pyruvate decarboxylation. An obvious choice was to find a suicide substrate that could trap, perhaps, the carbanion/enamine or the HEThDP intermediate. In an effort to better

define the short-lived carbanion/enamine intermediate, Jordan and coworkers devised an ingenious pyruvate-based suicide substrate (a substrate analog devised to react irreversibly with enzymes or their coenzymes), (E)-4-(4-chlorophenyl)-2-oxo-3-butenoic acid (CPB), designed to stably trap the post-decarboxylation carbanion/enamine intermediate by means of additional resonance stabilization (D. J. Kuo & Jordan, 1983a; 1983b). Indeed, CPB was found to irreversibly inhibit yeast pyruvate decarboxylase, while giving spectroscopic and other evidence for the formation of a stable covalent intermediate with the ThDP of the decarboxylase enzyme.

A design strategy for selecting a thiamin-utilizing ribozyme was chosen that involved, first, covalent attachment of a CPB-derived suicide substrate, *Lnk*PB, to the 5' end of each member of a random sequence RNA pool. *Lnk*PB was so named to indicate the replacement of the chloro group of CPB with an amide-based linker used to attach the α-keto acid substrate to RNA. Thiamin, coupled by its phosphate to a linker and a biotin 'handle' was then incubated with the pool of RNA•*Lnk*PB conjugates in an aqueous buffer solution that encouraged RNA folding (see **Figure 3.1** for structures of selection compounds). The formation of a stable covalent adduct incorporating biotin to active RNA•*Lnk*PB molecules and ThBi could then be monitored, and such ThBi adducts of RNA•*Lnk*PB could be sequestered away from unreacted RNA using streptavidin beads. RNA sequences enriched in this way were amplified by reverse transcription and PCR, followed by a new round of RNA transcription. The resulting RNA was re-conjugated with *Lnk*PB and reacted with ThBi, in iterative fashion, until sufficient enrichment of a putative decarboxylase ribozyme pool was obtained.

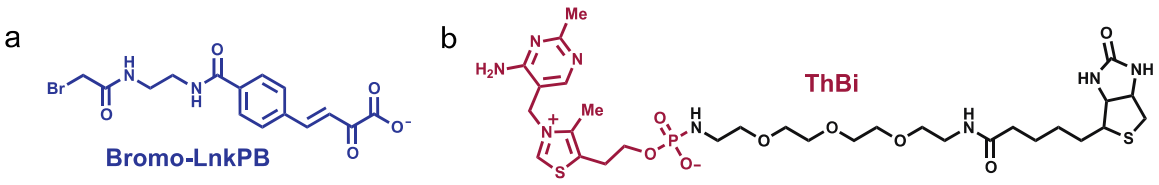

Figure 3.1 Selection compounds

(a) Brominated derivative of a pyruvate-based suicide substrate (Bromo-LnkPB). Bromo-LnkPB is reacted with the 5'-phosphorothioate-RNA to generate the RNA•LnkPB conjugate. (b) Thiamin monophosphate covalently tethered to biotin via a PEG-diamine linker (ThBi).

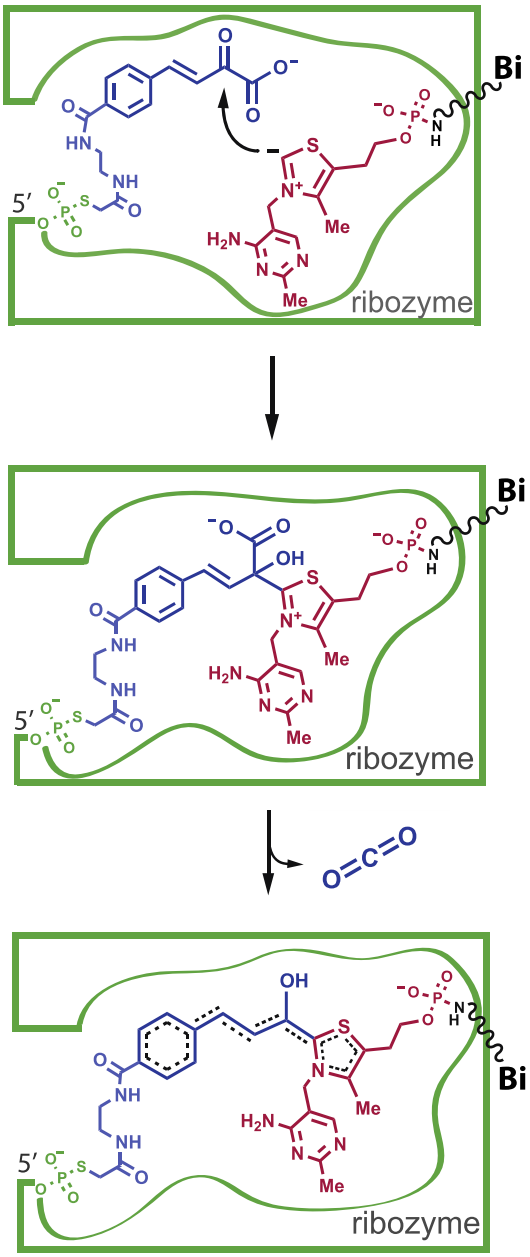

Figure 3.2 ***A strategy for isolating ribozymes with α-keto acid decarboxylase activity.***

Strategy for selection of decarboxylase ribozymes using thiamin-PEG-biotin (ThBi). (top) ThBi carries out nucelophilic addition to a LnkPB suicide substrate attached to the ribozyme (RNA• LnkPB), to form the lactyl intermediate analogue (centre). The lactyl intermediate analogue then undergoes loss of carbon dioxide to form a decarboxylated thiamin adduct, the carbanion/enamine intermediate (bottom), which is not very reactive owing to extended resonance

4. Materials and methods

4.1. Miscellaneous

Reverse-phase (C18) chromatography was performed on Waters 5g C18 Sep-Pac cartridges. Flash chromatography was performed in a glass column using silica gel 60. Compounds were detected using thin-layer chromatography (TLC) on aluminum-backed plates containing Merck Silica gel 60 F254 or using a Nanodrop spectrophotometer. A Bruker 600 MHz NMR spectrometer was used to record all spectra. 2D experiments, ^{1}H-^{13}C HSQC, and ^{1}H-^{13}C HMBC, were performed to confirm chemical assignments. All water was double-distilled, first by evaporative distillation followed by filtration through an ion exchange resin, 0.2 μm filtered and autoclaved. All buffers and salts were 0.2 μm filtered and autoclaved, with the exception of Li-HEPES buffer (0.2 μm filtered only). Thiamin and ThBi concentrations were measured via UV/Vis spectroscopy using the extinction coefficients at 235 nm (ε_{235}= 11,300 M^{-1} cm^{-1}) and 267 nm (8,000 M^{-1} cm^{-1}) in 0.1 M phosphate buffer, pH 7.0. Bromo-*Lnk*PB concentrations were determined using the experimentally determined extinction coefficient of ε_{305} = 3,200 M^{-1} cm^{-1}.

4.2. RNA pool construction

The random-sequence library was synthesized as DNA (5'-GATAA TACGA CTCAC TATAG GGATC GTCAG TGCAT TGAGA-**(N_{70})**-GGTGG TATCC GCAAC GGGTA -3').[1] It, and the forward and reverse PCR primers, 5'-GATAA TACGA CTCAC TATAG GGATC GTCAG TGCAT TGAGA -3' and 5'- TACCC GTTGC GGATA CCACC-3', respectively, were obtained from University of Calgary Core DNA services. The random-sequence library and forward primer were PAGE purified on 5% (w/v) denaturing gels. Eight cycles of large-scale PCR incorporating the forward and reverse primers were used to amplify the ssDNA pool, to generate the full-sized dsDNA pool.[2] *In*

vitro transcription reactions containing the dsDNA pool, 2 mM each of CTP, ATP, UTP and GTP, 4 mM GMPS, 5 µCi [α-^{32}P] UTP, and T7 RNA polymerase, generated an initial RNA pool with a sequence complexity of 4 x 10^{15}.[3] With the above GMPS/GTP ratio, ~70% of the RNA pool was found to be GMPS primed, and no further enrichment of the GMPS-RNA pool was deemed necessary for the selection. The RNA pool was purified by 8% (w/v) denaturing PAGE, and the activated suicide substrate (Bromo-*Lnk*PB **Figure 2.4a**) was attached to the 5' end of the RNAs in reactions that contained 10 µM RNA and 5 mM Bromo-*Lnk*PB in a buffer of 50 mM HEPES, pH 8**,** 1 mM EDTA, 2 mM tris(2-carboxyethyl)phosphine (TCEP), at 22° C for 1 h.[4] The RNA was recovered from the reaction mixture by two successive isopropanol (IPA) precipitations, by the addition of 15 µg glycogen as carrier, 1/10^{th} volume 3 M NaOAc and 1.1 volumes of IPA.[5]

4.3. Selection of decarboxylase ribozymes

Selection was initiated by incubating 10 µM [^{32}P]-RNA•*Lnk*PB together with 10 mM ThBi in Standard Reaction Buffer (SR Buffer): 50 mM Li-HEPES, pH 7.5, 200 mM NaCl, 200 mM KCl, and 40 mM $MgCl_2$. First, the [^{32}P]-RNA•*Lnk*PB was heated, in SR Buffer lacking magnesium, to 95° C for 2 min and cooled for 10 min, following which $MgCl_2$ was added from a stock solution. ThBi was added from a stock to start the reaction, and the reacting solution was incubated for 24 h at 22° C, before quenching by addition of EDTA to 40 mM. The RNA was now IPA-precipitated three successive times to ensure it was clean of excess ThBi. The resulting pellet was dissolved in water, and made up to 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 M NaCl (binding buffer). This solution was added to streptavidin-coupled magnetic beads (Dynabeads)[6] pre-equilibrated in binding buffer, and mixed gently for 30 minutes at room temperature. The beads were washed by fast vortexing with 30 volumes of binding buffer; 30 volumes of 4 M urea; followed by 30 volumes of 50 mM K-HEPES, pH 7.5, and 1 mM EDTA. The bead-bound RNA was reversed transcribed directly on the beads to cDNA, and the resulting crude mixture was used as template for PCR amplification.[7] For negative selections, 5'-phosphorothioate-RNA (no attached *Lnk*PB) was reacted with ThBi followed by incubation with streptavidin beads using a procedure similar to the positive selection. However, after the ThBi-reacted RNA bound to streptavidin beads, the

supernatant's RNA was isopropanol precipitated and subjected to either another negative selection or conjugation to *Lnk*PB for a positive selection. As the rounds of selection progressed, only the reaction time and concentration of ThBi in the reaction were changed (for the positive selection rounds 1-6, 10 mM ThBi, 24 h; rounds 7-8, 5 mM ThBi, 12 h; rounds 9-13, 2.5 mM ThBi, 6 h). Negative selections in rounds 7-8 were performed with 10 mM ThBi for 20 h. In rounds 9-13, for the first and/or second negative selection, 5 mM ThBi and 14 h, were used. The last negative selection in each round had the same conditions as the positive selection to follow, in order to enable a direct comparison between the two.

4.4. Streptavidin gel shift assays

For analysis of RNA transcribed from round 13 clones, GMPS-[^{32}P]RNA (body-labeled with ^{32}P) was purified away from non-GMPS-initiated RNA (to ensure that a quantitative attachment of the *Lnk*PB substrate to the RNA was feasible). This was accomplished by taking advantage of the propensity of GMPS-initiated RNAs toward forming phosphorothioate dimers under non-reducing conditions. The RNA dimer band (approximately 15-30% of total RNA) was separated and purified on 5% a denaturing PAGE gel,[8] eluted in the presence of 2 mM dithiothreitol (DTT), then purified rigorously away from DTT by successive IPA precipitations, resulting in pure GMPS-initiated RNA. Bromo-*Lnk*PB was now conjugated to the GMPS-primed RNA and the resulting [^{32}P]-RNA•*Lnk*PB was reacted with ThBi using the standard reaction protocol. In this, the ThBi-reacted [^{32}P]-RNA•*Lnk*PB was precipitated twice with IPA[9], and the resulting pellet dissolved in 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 mM EDTA, 0.5 mM DTT followed by the addition of 2 µg recombinant streptavidin (ProteoChem), and incubated at 22°C for 10 min. To this solution, an equal volume of streptavidin loading buffer (7 M Urea, 5% glycerol) was added, and the samples were run on a 12% w/v denaturing PAGE containing 7 M urea, at 5 W power. All gels were analyzed by phosphorimagery on a Typhoon Phosphorimager (GE), and analyzed by ImageJ software (NIH).

4.5. Decarboxylation kinetic measurments

All reaction kinetics were measured with 0.1-0.5 μM [^{32}P]-labeled dc4•*Lnk*PB, 10 mM ThBi, in SR buffer, at 30° C. Reactions were initiated as in the selection protocol, unless specified otherwise. Time points from each reaction were collected and quenched using 10 volumes of formamide containing 10 mM EDTA. Each time point sample was precipitated twice with IPA, using glycogen as a co-precipitant, and then analyzed by streptavidin gel-shift. First order kinetics reaction curves were fit to a single exponential equation (GraphPad), from which k_{obs} values were calculated. The kinetics of the Th/ ThBi competition experiments were measured similarly. Th and ThBi were pre-mixed, and the reactions initiated by the addition of the mixture to dc4•*Lnk*PB, to ensure unbiased rate measurements. The formalism $k_{obs} = k' [ThBi] + k'' [Th]$ (where k' is the second order rate constant for reaction of dc4•*Lnk*PB with ThBi, and k'' the rate constant for reaction with Th) was used to determine k' and k''.

4.6. Product isolation and ESI-MS

dc4•*Lnk*PB and dc4•[^{13}C]$_3$-*Lnk*PB, following reaction with ThBi, were purified away from unreacted RNA•*Lnk*PB by the streptavidin-mediated gel shift in a 5% denaturing polyacrylamide gel. The shifted RNA was eluted from the gel into a solution of formamide containing 300 mM NaOAc and 100 μg biotin, by heating to 80° C for 2 min, followed by vigorous vortexing for 5 min. The extracted RNA was recovered from the formamide solution by precipitation with the addition of 1.1 volumes IPA. The resulting RNA was dissolved in water to make up a ~50 μM solution, and was subjected to complete digestion by 0.05 μg/μL P1 nuclease (Sigma) at room temperature for 1 hour followed by treatment with alkaline phosphatase (Roche; 0.05 U/μL) for an additional hour. The crude solution of digested RNA was HPLC purified on a C18 column equilibrated in 5 mM NH_4OAc, pH 6.8, and eluted with a gradient starting with 5 mM NH_4OAc in 100% H_2O to 5 mM NH_4OAc in 1:1 H_2O:ACN. The covalently modified nucleotide product was identified, eluted, put directly on ice and subjected to Electrospray Ionization Mass Spectrometry (ESI-MS), in positive mode, on a Bruker maXis ultra-high performance mass spectrometer. For analysis of the thiamin (Th) adduct of dc4•*Lnk*PB, purification and mass spectrometry were carried out as for the

ThBi reaction; however, the dc4•[^{13}C]-*Lnk*PB was reacted three successive times (with isopropanol precipitation followed by RNA re-folding each time) with 10 mM thiamin (pH 7.5), in SR buffer, for 8h each. Since the Th adduct did not have a biotin label, the total RNA was subjected to P1 nuclease and alkaline phosphatase digestion prior to analysis of the digested products on HPLC.

4.7. Synthesis of thiamine derivatized adduct, thiamine-biotin (ThBi)

Thiamine monophosphate chloride dihydrate (1.38 g, 3.3 mmol; 1.1 equiv.) and N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC) (0.63 g, 3.3 mmol; 1.1 equiv.) were dissolved in in 8 mL of double distilled H_2O (ddH_2O) containing 1-methylimidazole (48 □L, 0.6 mmol; 0.2 equiv.). The solution was brought to pH 6.5 with 1M NaOH, amine-PEG$_3$-biotin (1.25 g, 6.0 mmol; 1 equiv.) was added and the solution was stirred at 4° C overnight. 20% of the reaction was loaded on a 5 g C18 Sep-Pak column equilibrated with H_2O. The product was eluted with a gradient of 0-20% acetonitrile. Product fractions were characterized by thin layer chromatography (H_2O:ACN, 10:1) and stained using ninhydrin. The purification procedure was repeated 5 times. Product fractions were lyophilized to yield a white crystalline solid (1.16 g, 1.5 mmol, 52% yield).

^{1}H NMR (600 MHz, D_2O) δ 8.08 (s, 1H), 5.48 – 5.46 (m, 1H), 5.44 (s, 2H), 4.61 (dd, J = 7.9, 4.8 Hz, 1H), 4.42 (dd, J = 7.9, 4.5 Hz, 1H), 4.04 (q, J = 5.8 Hz, 2H), 3.72 – 3.65 (m, 8H), 3.63 (t, J = 5.4 Hz, 2H), 3.55 (t, J = 5.7 Hz, 2H), 3.39 (t, J = 5.4 Hz, 2H), 3.36 – 3.31 (m, 1H), 3.28 (t, J = 5.6 Hz, 2H), 2.99 (dd, J = 13.1, 5.0 Hz, 1H), 2.95 (dt, J = 9.8, 5.7 Hz, 2H), 2.77 (d, J = 13.0 Hz, 1H), 2.59 (s, 3H), 2.49 (s, 3H), 2.27 (t, J = 7.3 Hz, 2H), 1.75 – 1.53 (m, 4H), 1.46 – 1.35 (m, 2H). ^{13}C NMR (151 MHz, D_2O) δ 179.66, 171.83, 168.08, 164.56, 160.09, 156.46, 145.66, 138.34, 107.17, 72.38, 72.38, 71.60, 65.78, 64.81, 62.98, 58.09, 53.75, 43.29, 42.43, 41.62, 38.17, 30.60, 30.55, 30.42, 27.86, 26.77. ^{31}P NMR (600 MHz, D_2O) δ 9.10; UV/Vis: λ_{max} 235 nm; HRMS (*m/z*): $[M+H]^+$ calculated for $C_{30}H_{50}N_8O_8PS_2^+$,745.2925; found, 745.2932.

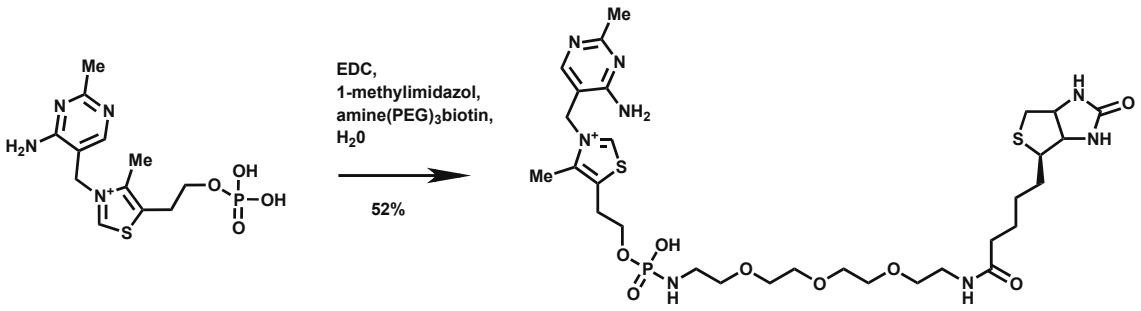

Figure 4.1 Synthesis of ThBi

4.8. Synthesis of tert-butyl 2-(4-formylbenzamido)ethylcarbamate (Boc-*Lnk*-benzaldehyde)

4-carboxybenzaldehyde (4.13 g, 27.5 mmol; 1.1 equiv.) was dissolved in dry DMF (180 mL) in a stirring round bottom flask. A mixture of diisopropylethylamine (DIEA) (4.35 mL, 25 mmol; 1 equiv.) and *tert*-butyl (2-aminoethyl)carbamate (4.00 g, 25 nmols; 1 equiv.) were added dropwise. The reaction was stirred vigorously at room temperature for 30 minutes followed by rotary evaporation to remove the DIEA and DMF. The resulting solid was purified using silica flash chromatography (hexane:EtOAc, 3:7 v/v). Product fractions were characterized using TLC, stained with ninhydrin or 2,4-dinitrophenylhydrazine. The solvent was then removed under reduced pressure to yield a white crystalline product (6.43 g, 22.0 mmol, 88%).

^{1}H NMR (600 MHz, DMSO) δ 10.11 (s, 1H), 8.70 (s, 1H), 8.04 (dd, *J* = 18.4, 8.1 Hz, 4H), 6.95 (s, 1H), 3.40 – 3.30 (m, 2H), 3.21 – 3.11 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (151 MHz, DMSO) δ 193.80 (s), 166.47 (s), 156.67 (s), 140.51 (s), 138.63 (s), 130.24 (s), 128.86 (s), 78.61 (s), 40.67 (s), 40.38 (s), 29.15 (s); HRMS (*m/z*): $[M+Na]^+$ calculated for $C_{15}H_{20}N_2O_4$, 292.1423; found, 292.1423.

4.9. Synthesis of (E)-4-(4-(2-(tert-butoxycarbonylamino)ethylcarbamoyl)phenyl)-2-oxobut-3-enoic acid (Boc-*Lnk*PB) and (Boc-^{13}C*Lnk*PB):

In a 15 mL falcon tube, sodium pyruvate (225 mg, 2.0 mmol; 2 equiv.) or 3-(^{13}C)pyruvate (Cambridge Isotopes, 99.0 atom%) (225 mg, 2.0 mmol; ~2 equiv.) was dissolved in 8 mL of reaction solution (50 mM NaOH in EtOH:H_2O, 1:1 v/v) followed by Boc-Lnk-benzaldehyde (300 mg, 1.0 mmol; 1 equiv.). The solution was vortexed at room temperature for 1 hour. After 15 min, the solution turned a bright yellow/green color. The solution was neutralized with 1N HCl and lyophilized to yield a light yellow/green powder. The powder was taken up in 10 mL ddH_2O, and loaded on a 5 g C18 Sep-Pak column equilibrated with H_2O. The product was eluted with a gradient of acetonitrile/H_2O (0-40% acetonitrile). Product fractions were identified by making absorbance measurements at 310 nm using a Nanodrop spectrophotometer. The product was lyophilized resulting in a yellow powder.

For **Boc-*Lnk*PB** (336 mg, 0.9 mmol, 85% yield). ^{1}H NMR (600 MHz, MeOD) δ 7.90 (d, *J* = 8.1 Hz, 2H), 7.77 (d, *J* = 8.2 Hz, 2H), 7.73 (d, *J* = 16.3 Hz, 1H), 7.08 (d, *J* = 16.4 Hz, 1H), 3.49 (t, *J* = 6.1 Hz, 2H), 3.31 (t, *J* = 6.0 Hz, 2H), 1.45 (s, 9H). ^{13}C NMR (151 MHz, MeOD) δ 197.48 (s), 173.18 (s), 170.49 (s), 159.65 (s), 147.15 (s), 139.99 (s), 138.12 (s), 130.35 (s), 129.86 (s), 127.20 (s), 81.05 (s), 42.33 (s), 41.69 (s), 29.60 (s); UV/Vis: λ_{max} 305 nm; HRMS (*m/z*): $[M+Na]^+$ calcd. for $C_{18}H_{22}N_2O_6$, 362.1478; found, 362.1478.

For the **Boc-^{13}CLnkPB** (0.314 g, 0.8 mmol, 79% yield). ^{1}H NMR (600 MHz, MeOD) δ 7.90 (d, *J* = 8.1 Hz, 2H), 7.78 (d, *J* = 8.2 Hz, 2H), 7.75 (dd, *J* = 14.4, 4.9 Hz, 1H), 7.13 (dd, *J* = 160.5, 16.3 Hz, 1H), 3.49 (t, *J* = 6.1 Hz, 2H), 3.32 (t, *J* = 6.1 Hz, 2H), 1.45 (s, 9H). ^{13}C NMR (151 MHz, MeOD) δ 196.01 (dd, *J* = 66.3, 53.3 Hz), 172.18 (dd, *J* = 66.3, 15.3 Hz), 170.47 (s), 159.66 – 159.64 (m), 147.30 (d, *J* = 71.1 Hz), 139.91 (d, *J* = 6.2 Hz), 138.21 (s), 130.42 (d, *J* = 4.6 Hz), 129.87 (s), 126.82 (dd, *J* = 53.2, 15.4 Hz), 81.05 (s), 42.34 (s), 41.69 (s), 29.60 (s); UV/Vis: λ_{max} 305 nm; HRMS (*m/z*): $[M+Na]^+$ calcd. for $C_{15}[^{13}C]_3H_{22}N_2O_6$, 365.1579; found, 365.1577.

4.10. Synthesis of (E)-4-(4-(2-(2-bromoacetamido)ethylcarbamoyl)phenyl)-2-oxobut-3-enoic acid (Bromo-*Lnk*PB & Bromo-^{13}C*Lnk*PB)

In a 15 mL Falcon tube, Boc-*Lnk*PB or Boc-^{13}C*Lnk*PB (100 mg, 0.25 mmols) was dissolved in 1 mL of anhydrous trifluoroacetic acid (TFA) and vortexed at room temperature for 10 minutes. The TFA was removed under vacuum and the resulting solid was dissolved in ddH_2O (8 mL) and neutralized with 1M NaOH. The solution was brought to a final concentration and volume of reaction solution (12 mL of 0.2 M sodium phosphate buffer, pH 5.3, in H_2O:ACN, 1:1 v/v). Bromoacetic acid N-hydroxysuccinimide ester (BrAcNHS) (64 mg, 0.3 mmols; 1.1 equiv.) was added and the reaction was mixed for 1 hour at room temperature. The solution was concentrated under reduced pressure and loaded onto a 5 g C18 Sep-Pac column equilibrated in H_2O and eluted with a gradient (ACN:H_2O, 0-30% v/v acetonitrile). Fractions were characterized by absorbance measurements at 310 nm and lyophilized to yield a light yellow powder.

For **Bromo-*Lnk*PB** (78 mg, 0.2 mmol, 76% yield). ^{1}H NMR (600 MHz, MeOD) δ 7.90 (d, *J* = 8.4 Hz, 2H), 7.77 (d, *J* = 8.4 Hz, 2H), 7.72 (d, *J* = 16.3 Hz, 1H), 7.07 (d, *J* = 16.4 Hz, 1H), 3.87 (s, 2H), 3.56 (t, *J* = 6.0 Hz, 2H), 3.48 (t, *J* = 6.0 Hz, 2H). ^{13}C NMR (151 MHz, MeOD) δ 197.50 (s), 173.25 (s), 170.81 (s), 170.64 (s), 147.12 (s), 140.05 (s), 138.04 (s), 130.38 (s), 129.87 (s), 127.26 (s), 41.46 (s), 41.36 (s), 29.60 (s); UV/Vis: λ_{max} 305 nm; HRMS (*m/z*): $[M+H]^+$ calcd. for $C_{15}H_{15}BrN_2O_5$, 382.0164; found, 382.0167.

For **Bromo-^{13}C*Lnk*PB** (74 mg, 0.2 mmol, 70%). ^{1}H NMR (600 MHz, MeOD) δ 7.89 (d, *J* = 8.3 Hz, 2H), 7.77 (d, *J* = 8.3 Hz, 2H), 7.72 (dd, *J* = 16.1, 6.5 Hz, 1H), 7.07 (dd, *J* = 160.1, 16.5 Hz, 1H), 3.87 (s, 2H), 3.56 (t, *J* = 6.0 Hz, 2H), 3.48 (t, *J* = 6.0 Hz, 2H). ^{13}C NMR (151 MHz, MeOD) δ 197.49 (dd, *J* = 66.0, 52.3 Hz), 173.47 (d, *J* = 15.1 Hz), 170.81 (s), 170.64 (s), 147.12 (d, *J* = 66.8 Hz), 140.05 (d, *J* = 6.3 Hz), 138.04 (s), 130.38 (d, *J* = 4.6 Hz), 129.87 (s), 127.25 (dd, *J* = 52.3, 15.1 Hz), 41.46 (s), 41.36 (s), 29.60 (s); UV/Vis: λ_{max} 305 nm; HRMS (*m/z*): $[M+H]^+$ calcd. for $C_{12}[^{13}C]_3H_{15}BrN_2O_5$, 385.0265; found, 385.0261.

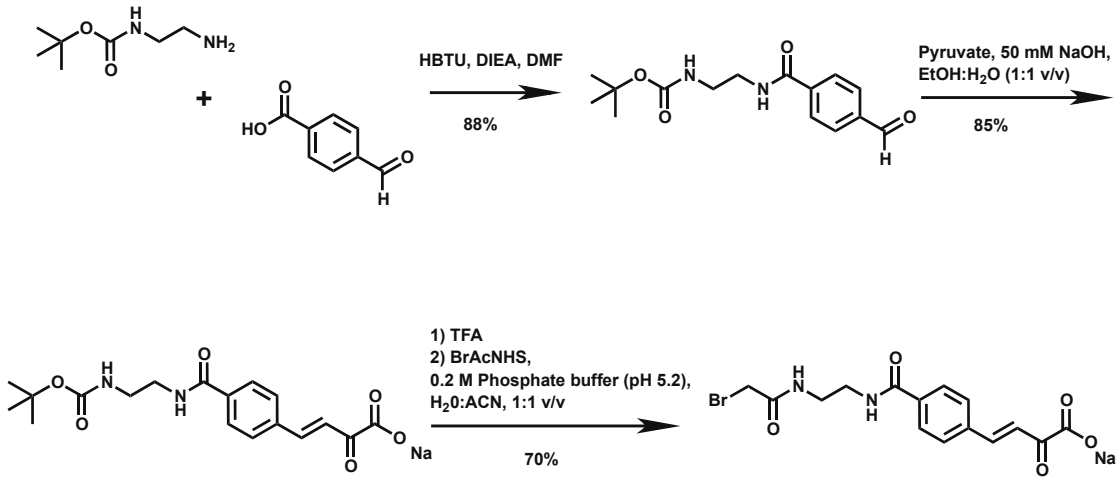

Figure 4.2 Synthesis of Boc-LnkPB & Bromo-LnkPB

5. In vitro selection of TB reactive RNAs with activity *independent* of *Lnk*PB

5.1. Selection conditions and overview

In vitro selection for decarboxylase ribozymes (**Figure 5.1**) was enabled by transcription of the initial random RNA pool from a random sequence dsDNA library using guanosine-5'-monophosphorothioate (GMPS) as transcription initiator. The availability of a thiolate functionality at the 5' end of ~70% of the RNA pool (see Methods) enabled covalent attachment of the RNA to our modified α-keto acid suicide substrate, *Lnk*PB. This modified random RNA pool (*Lnk*PB•5'-RNA) was allowed to fold in the Standard Reaction Buffer (SR Buffer: 50 mM Li-HEPES, pH 7.5, 200 mM NaCl, 200 mM KCl, and 40 mM $MgCl_2$) and made up to 10 mM ThBi. Incubation was carried out for 24 hours at 22°C, following which the RNA pool was purified away from excess unreacted ThBi, and any RNA•*Lnk*PB molecules putatively covalently derivatized with ThBi were pulled down using streptavidin-coated beads. Bead-immobilized RNA was subjected to on-bead reverse transcription, followed by a repetition of the steps described above, for subsequent rounds of selection (**Figure 5.1**).

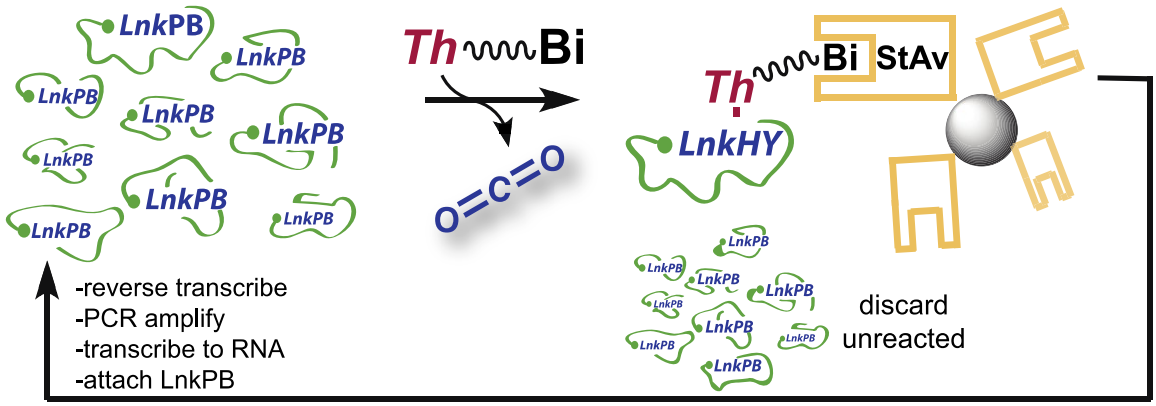

Figure 5.1. In vitro selection of decarboxylase ribozymes

An RNA pool, incorporating a 70-nt random region flanked by known sequences, is first conjugated with *Lnk*PB on its 5' terminus, and then incubated with ThBi in the Standard Reaction (SR) buffer solution. Those RNAs (ribozymes) that are capable of utilizing ThBi to decarboxylate the conjugated suicide substrate (*Lnk*PB) are trapped as an adduct (*Lnk*HY•ThBi), a resonance-stabilized "carbanion/enamine" intermediate. RNAs containing these adducts can be enriched away from the unreacted RNA pool using streptavidin-coated beads.

5.2. *In vitro* selection

Figure 5.2 plots the progression of RNA enrichment (percentage of RNA bound to the beads after repeated washings) for the first 20 rounds of *in vitro* selection. By Rounds 7-8, a significant population of parasitic RNAs was detected, that showed reactivity with ThBi even in the absence of *Lnk*PB conjugation to the RNA. The parasitic RNAs were thought to be reacting with ThBi in an unknown way because they where resistant to denaturants (both 8M urea and 100% formamide), and to heat (95° C for 5 min). The selections were continued without the 5'-GMPS and without attachment of *Lnk*PB. Rounds 10-13 were performed by reducing the ThBi and reaction time by half of the previous rounds conditions, until round 13, in which the reaction time was 1.5 hours and 0.5 mM ThBi. These conditions resulted in the percentage of RNA attaching to beads to reduce to background levels (**Figure 5.2**). These conditions were held constant between 13 and 20. By round 17 the RNA population showed increased binding to the streptavidin beads above background, presumably a result of *in vitro* evolution of sequences with increased tolerance to the lower ThBi concentration, shorter reaction time (faster ribozymes) or both. Round 12 and 20 cDNA was cloned for further analysis.

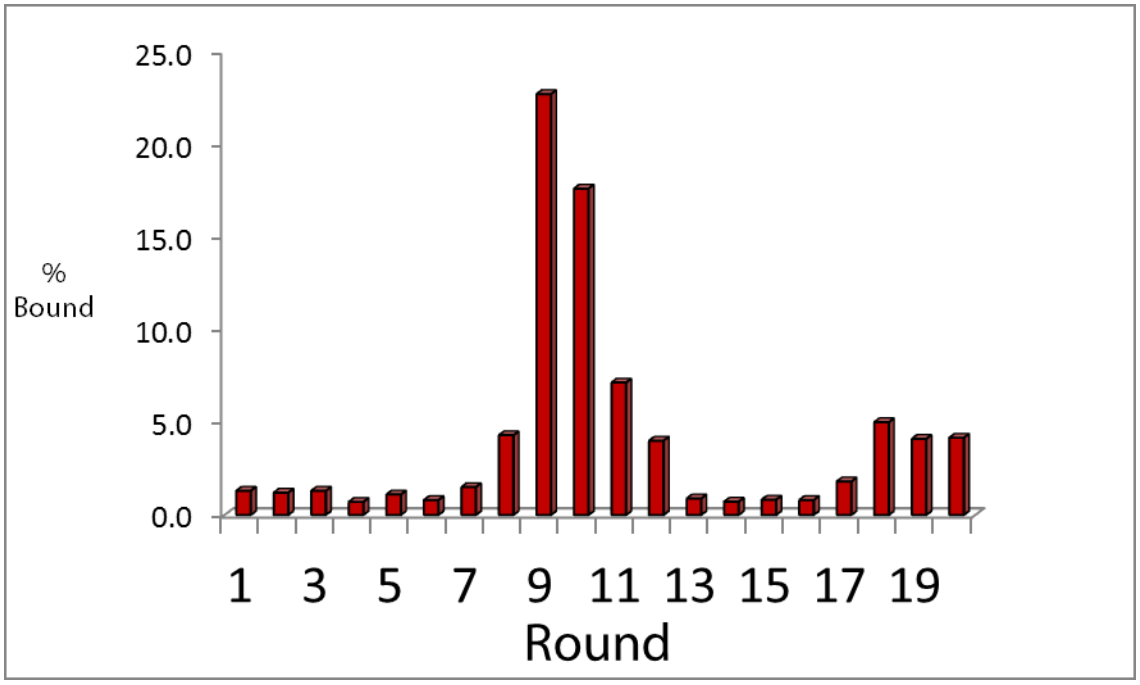

Figure 5.2. In vitro selection of LnkPB independent RNAs

Percentage of the overall RNA pool, in a given round, capable of binding streptavidin beads. For rounds 1-9 *Lnk*PB•RNA was incubated with 10 mM ThBi for 24 hours. Round 10 RNA (-*Lnk*PB) was incubated with 5 mM ThBi for 12 hours, round 11 was incubated with 2.5 mM ThBi for 6 hours, round 12 was incubated with 1.25 mM ThBi for 3 hours and round 13 was incubated with 0.5 mM ThiBi for 1.5 hours. Round 13 conditions were used for all subsequent rounds. Round 20 cDNA was cloned for analysis.

5.3. Analysis of *Lnk*PB-*independent* sequences

27 clones from round 12 and 26 clones from round 20 were aligned using ClustalX software and folded *in silico* using the Mfold algorithm. Sequences were named according to the clone number followed by "x" indicating the *Lnk*PB-independent selection and ending with the round number from which they were cloned (for example: clone one from round 12 of the *Lnk*PB-indpendent selection is called 1x12). The cloned sequences were categorized into 6 classes (class A-F) based on common homology and/or structural motifs.

Class A was the most abundantly represented of all classes and consisted of 20 total sequences (**Figure 5.3**). Class A clones are characterized by a relatively large stem loop (P1), containing multiple, non-conserved bulges, followed by a 18-19 nt stem loop

(P2) and terminating with a stem loop (P3). P3 was particularly interesting because of its near universal conservation and its fully conserved G-A bulge. Although the 3′-end of P3 contains the constant region, there were no evolutionary constraints in creating a fully base-paired P3 stem, therefore, the G-A bulge is was considered an curious structural feature. Class B contained the same 3'-end stem loop as class A, however P1 and P2 were largely heterogeneous, containing additional bulges and, in some cases, stem-loops (**Figure 5.4**). Class C was categorized primarily on their sequence homology. In particular, class C contained the highly conserved 5'-consensus sequence, "GGGTAGTAGGTGGGCAG" among other smaller areas of conservation that makes these sequences distinctive (**Figure 5.5**). Classes E, D and F were represented least of all the classes and therefore few observations can be made into their sequence or structural conservation (**Figure 5.6**).

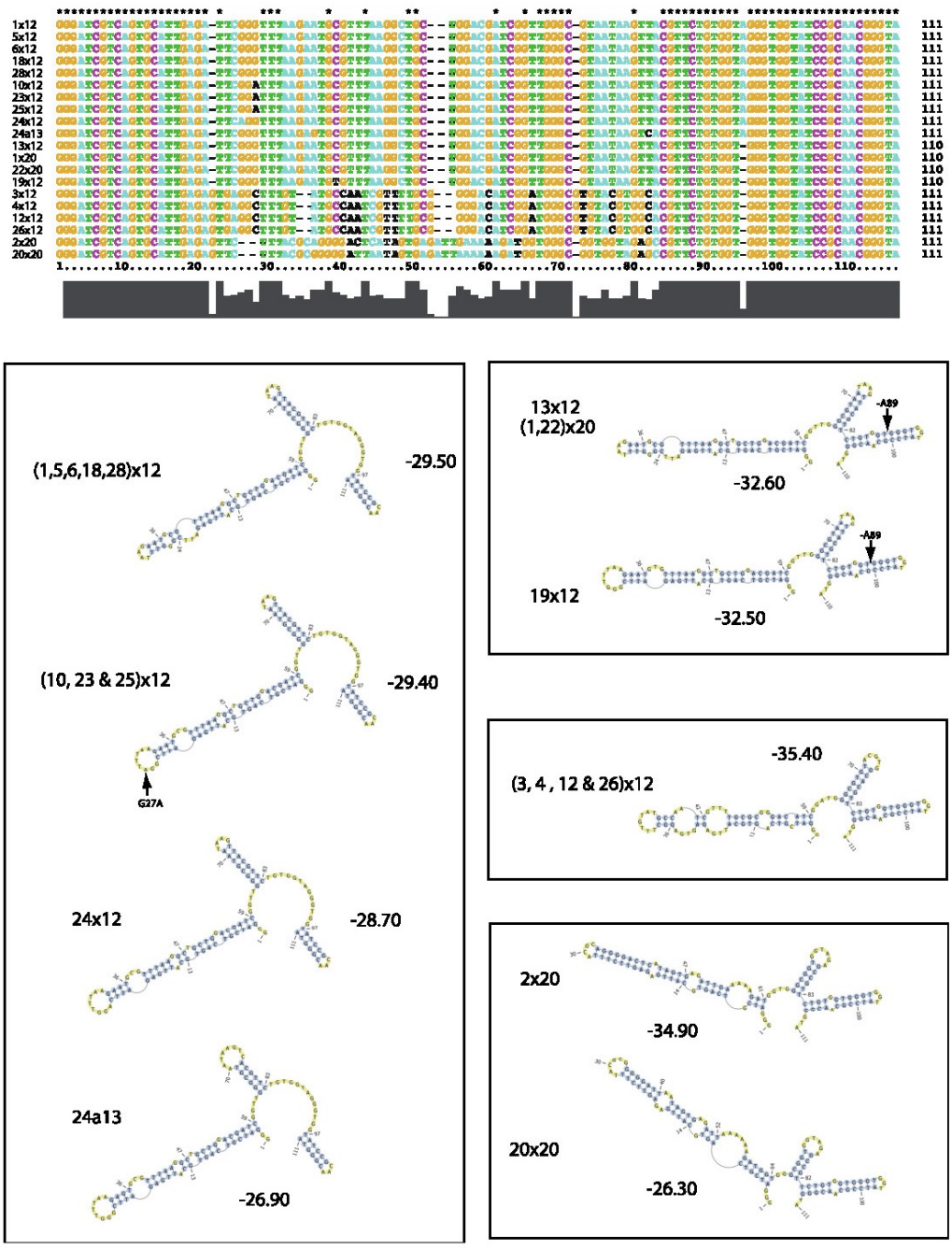

Figure 5.3. Class A sequences

Sequences were aligned using ClustalX software. Gray bars under the sequences represent the level of conservation at each position. Structures were generated using the Mfold algorithm. Negative numbers to the right of sequence names are the Mfold calculated folding energy. Very closely related sequences are grouped in boxes.

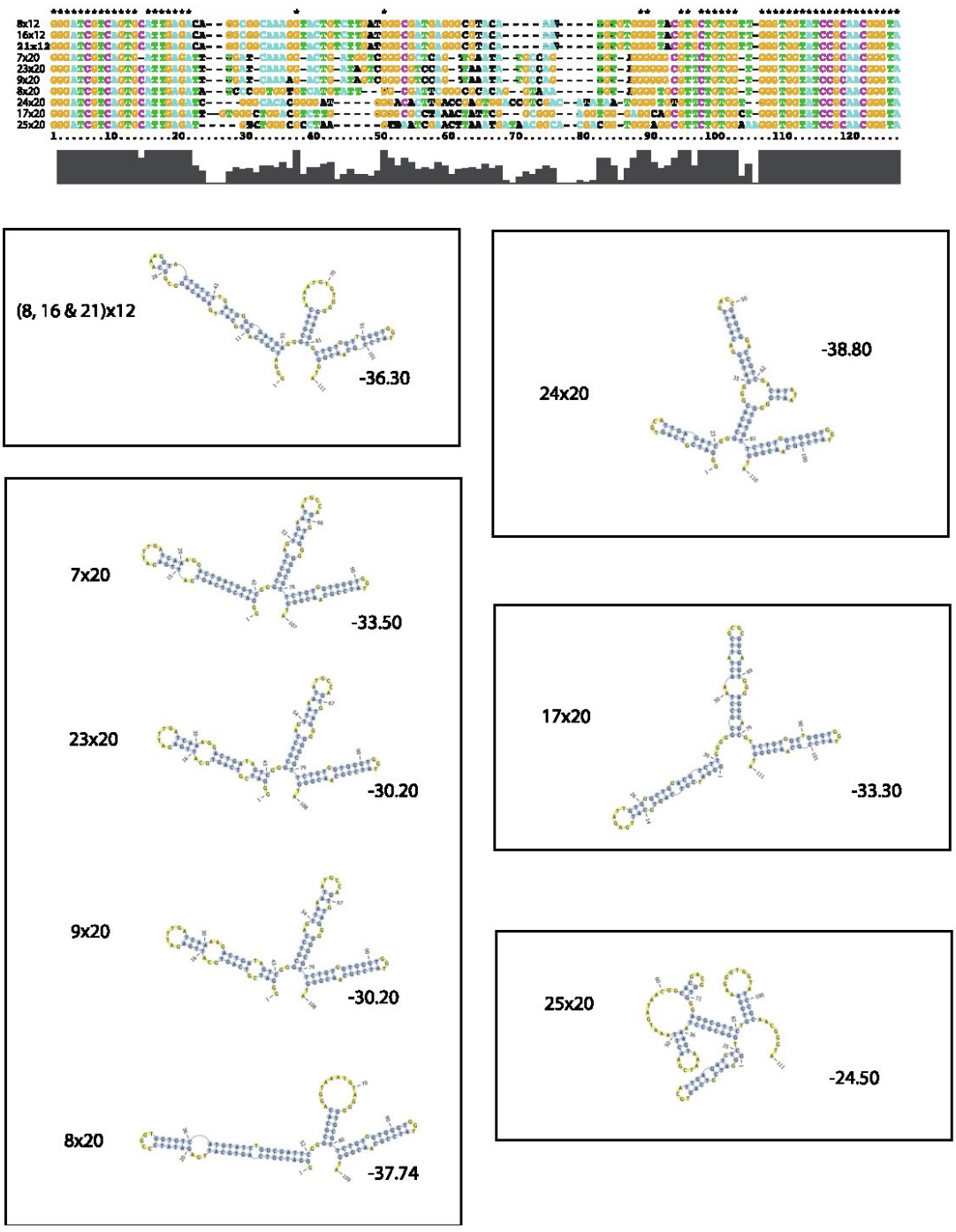

Figure 5.4. Class B sequences

Sequences were aligned using ClustalX software. Gray bars under the sequences represent the level of conservation at each position. Structures were generated using the Mfold algorithm.

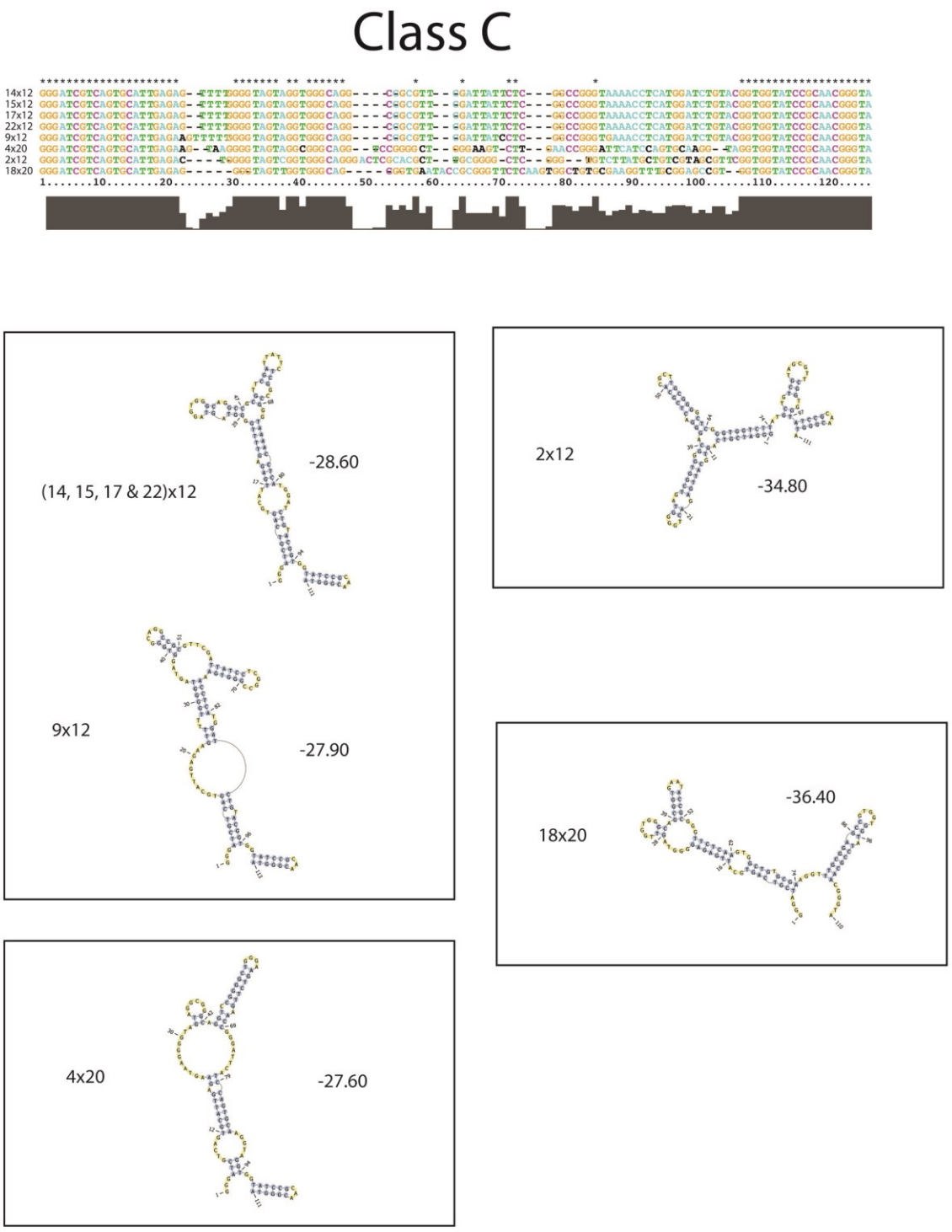

Figure 5.5. Class C sequences

Sequences were aligned using ClustalX software. Gray bars under the sequences represent the level of conservation at each position. Structures were generated using the Mfold algorithm.

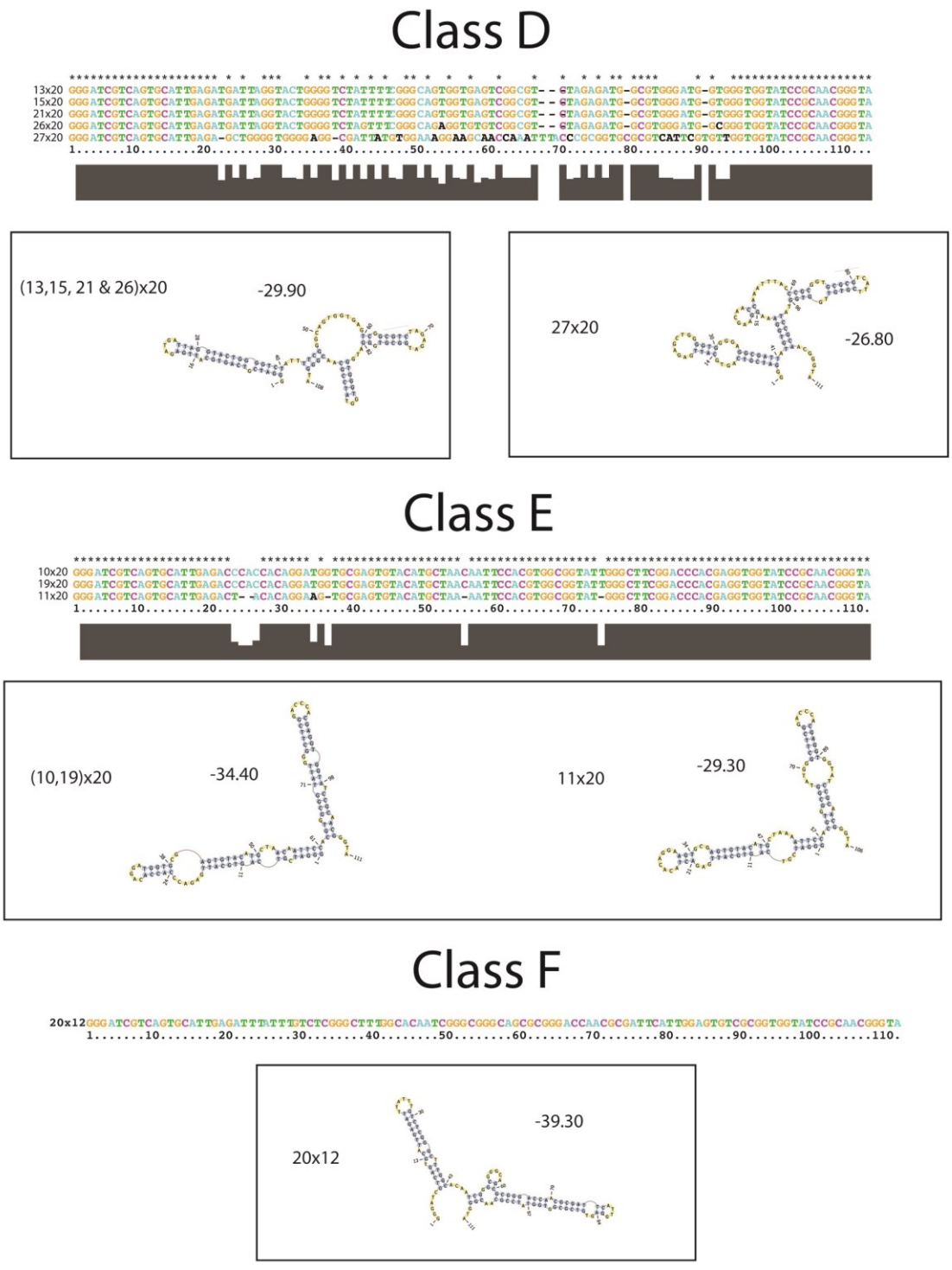

Figure 5.6. Class D, E, & F sequences

Sequences were aligned using ClustalX software. Gray bars under the sequences represent the level of conservation at each position. Structures were generated using the Mfold algorithm

5.4. Analysis of clone 2x20

Clone 2 from round 20 (2x20) was chosen for further analysis. If 2x20 was in fact *reacting* with ThBi in an unknown way, the reacted RNA would show a streptavidin-dependent shift in a PAGE gel. 2x20 RNA incubated in the presence of ThBi (as low as 0.5 mM) for 8-16 hours along with selection buffer, indeed showed a streptavidin-dependent shift in a PAGE gel (**Figure 5.7**). To determine the location of the putative covalent attachment of TB, 5′-end labeled 2x20 was reacted with TB and the reacted RNA streptavidin complex (StAV-TB-RNA) was purified and eluted from a PAGE gel. The reacted RNA was partially hydrolyzed and the fragments resolved to nucleotide resolution on a sequencing gel (6% polyacrylamide, 7M urea). The 5′ labeled fragments containing the covalent attachment resulted in a discontinuity in the sequencing gel. This indicated TB was reacting specifically with adenosine 105, positioned in the 3' constant region, proximal to the G-A bulge of P3.

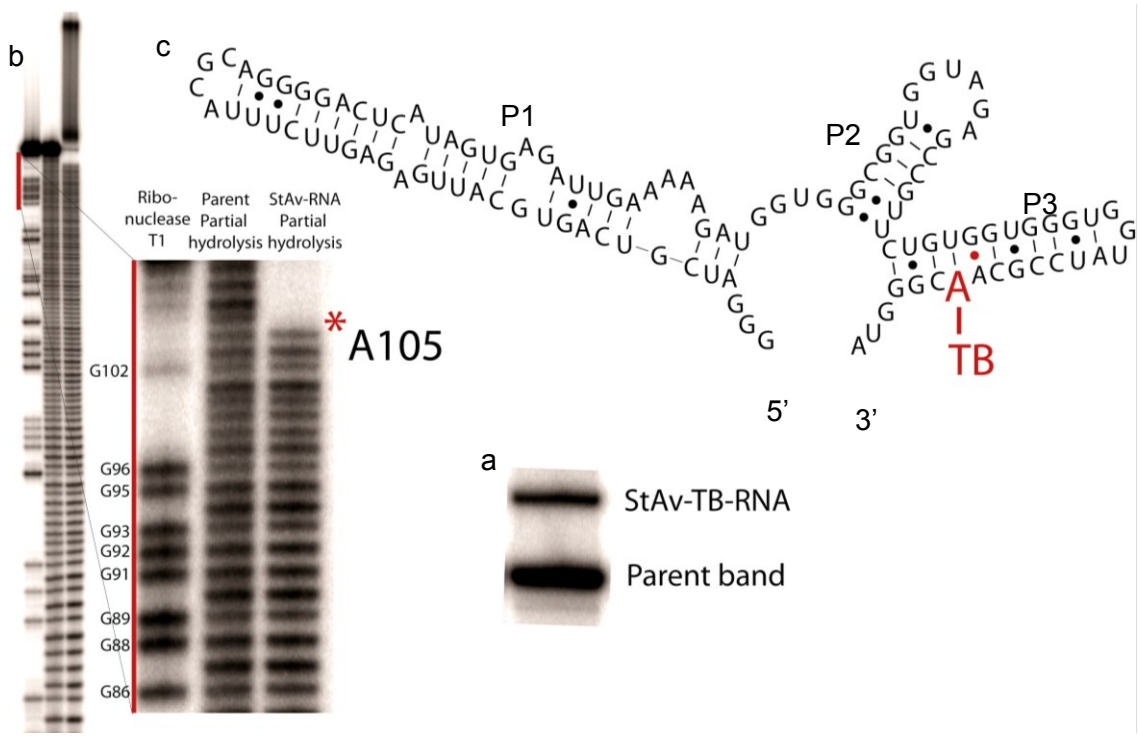

Figure 5.7. Partial hydrolysis footprinting of reacted 2x20 RNA

(a) Streptavidin gel shift assay. Reacted RNA (StAv-TB-RNA) was eluted from the gel. (b) Sequencing gel. A discontinuity is shown in the StAv-RNA partial hydrolysis lane. The red asterisks marks position 105. Ribonuclease T1 and partial hydrolysis of the parent band were used to locate the position of the adduct. (c) MFold structure of 2x20 with adenosine 105 highlighted in red with the attached TB.

5.5. Purification of the 2x20 adenosine adduct

In attempts to characterize the 2x20-TB adduct, the minimal reaction product was first isolated. 2x20 was first reacted with TB in the presence of selection buffer for ~24 hours and the streptavidin-dependent shift was purified from a 4-5% polyacrylamide PAGE gel. The resulting RNA digested using P1 nuclease (sigma) followed by a phosphatase treatment (alkaline phosphatase; Roche) to yield the four nucleosides and the TB-adenosine (A-TB) adduct. This crude reaction mixture was HPLC resolved on a C18 reverse-phase column. **Figure 5.8** shows the HPLC trace of this mixture. Since the unmodified nucleosides represented a large percentage of the mixture these peaks

represent the bulk of the spectrum. However, a small peak was found with a significantly greater retention time than the canonical nucleosides. This peak was collected for mass spectral analysis.

In spite of multiple attempts at ESI-MS (electrospray ionization mass spectrometry) no significant signal was seen close to the expected mass/charge of an adenosine-TB adduct. It was concluded the adduct was breaking down in an unpredictable way, before a mass spectrum could be obtained (Note: it is also feasible that the intact adduct did not properly ionize in the mass spectrometer and/or the signal was suppressed by contaminants in the preparation).

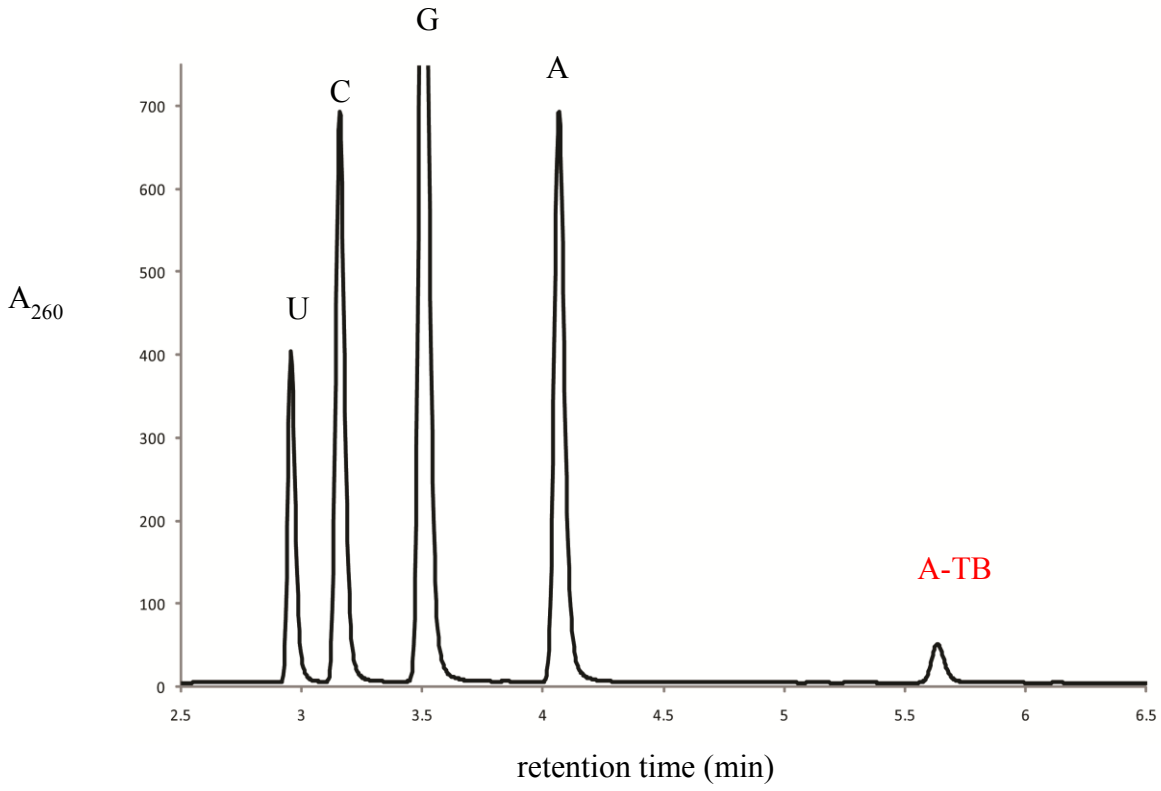

Figure 5.8. HPLC purification of adenosine-TB adduct

HPLC trace of digested 2x20 after purification of reacted RNA. Peaks labeled U, C, G, & A, represent the four canonical nucleosides. Peak labeled A-TB was the putative adenosine-TB adduct and was collected for mass spectral analysis.

5.6. NMR analysis of the adenosine adduct

NMR was thought to be another viable option for characterizing the adenosine-TB adduct of the 2x20 ribozyme. For a ^{1}H-NMR spectrum of the A-TB adduct, ~10-15 nmols was predicted to be sufficient. However, 2D experiments such as heteronuclear single quantum coherence (HSQC) or HMBC would require significantly more material (~50 nmols). To collect enough material for these less sensitive NMR experiments, large scale transcription reaction were performed (15 mL total volume). The gel purified RNA was reacted with 0.5 mM TB in the presence of selection buffer. The TB-reacted RNA was isolated using with a streptavidin-shift gel, nuclease digested using P_1 nuclease/alkaline phosphatase, and the adduct was separated on the HPLC as for the MS studies.

The resulting A-TB was lyophilized from the HPLC buffer (5 mM ammonium acetate, pH 5.5) and reconstituted in D_2O and re-lyophilized to reduce the NMR HOD signal (a high HOD signal would reduce the amount of maximal gain of the spectrophotometer). The sample was transferred to an NMR tube, flushed with argon and the cap was wrapped with parafilm. Initially, ^{1}H-NMR (carbon decoupled) and ^{1}H-^{1}H COSY experiments were performed on the sample. These experiments confirmed that both TB and adenosine were present in the sample. The protons of the thiamin portion of A-TB were assigned using ^{1}H-^{1}H COSY. The chemical shifts measured were very similar to that of a TB standard (**Figure 5.9**). This analysis, of course, is impeded by the heteroatoms of thiamin.

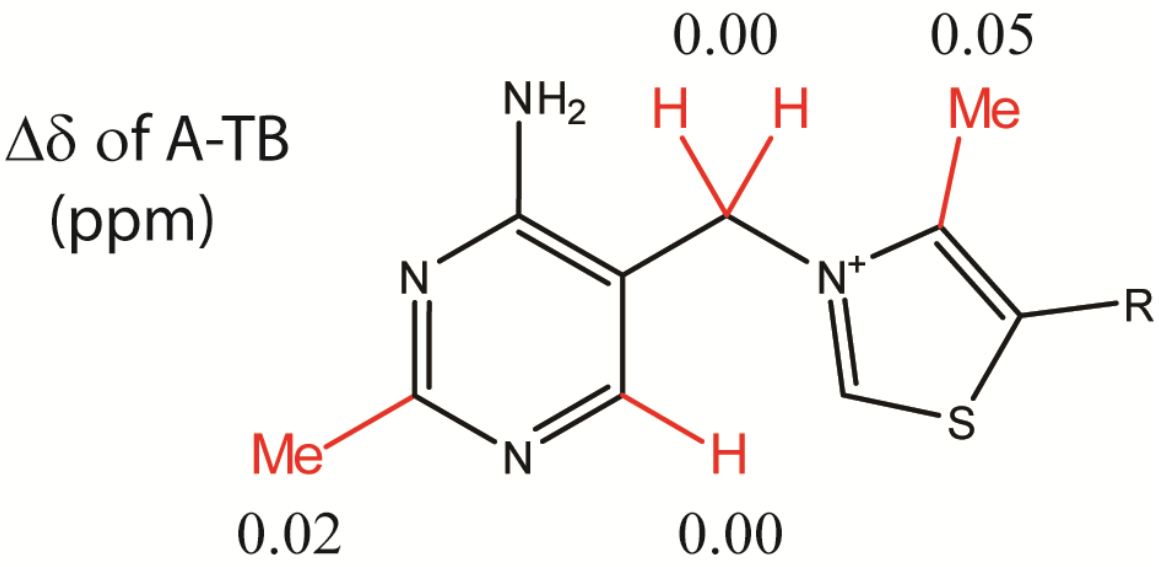

Figure 5.9. Differences in chemical shifts of the thiamin portion of A-TB

Numbers represent the absolute difference in chemical shift (ppm) between the protons of the thiamin portion of A-TB and that of a thiamin HCl standard.

Figure 5.10 compares the proton chemical shifts of the adenosine 3'-phosphate element of A-TB with an adenosine standard. The difference in chemical shift of the adenine base are nearly identical, however, the protons of the ribose elements differ significantly (**Figure 5.10**). Particularly, the chemical shift of the ribose H2′ displays a 0.70 ppm shift up-field. This substantial chemical shift suggests a chemical modification at or near the C2′. The most likely explanation of this data is modification of the 2′-OH (a particularly reactive atom of adenosine), specifically, a modification with an element with electron withdrawing character.

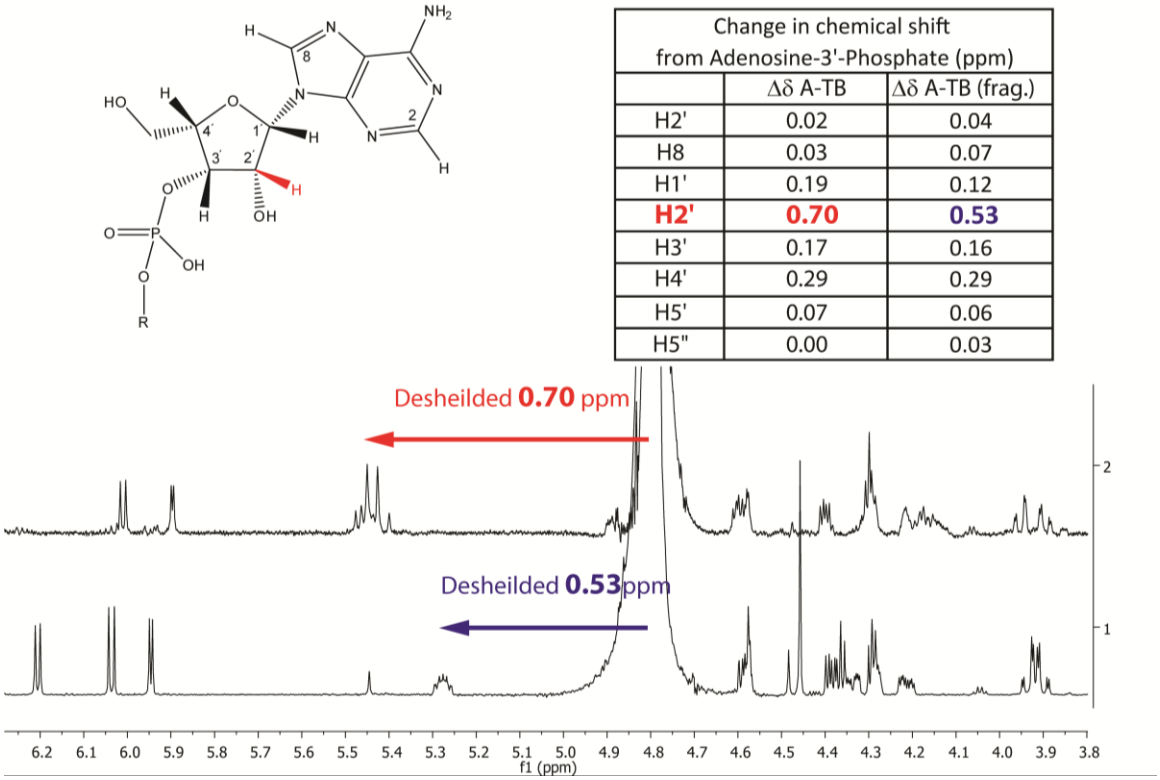

Figure 5.10. Difference in chemical shifts of the adenosine 3'-phosphate portion of A-TB before and after fragmentation

(a) Structure of adenosine 3' phosphate. (b) Table containing the change in chemical shift ($\Delta\xi$), of the A-TB & the fragmented A-TB (A-TB (frag.)) from an adenosine-3'-phosphate standard. (c) NMR spectra of *(1)* A-TB (frag.) and *(2)* A-TB.

Unfortunately, the A-TB sample appeared to degrade or fragment, with an approximate half-life of 1 day, shortly after the initial ^{1}H and ^{1}H-^{1}H COSY spectra were acquired. After ~4 days the degradation appeared to be complete. ^{1}H NMR of the fragmented A-TB (A-TB (frag.)) sample shows a similar chemical shift at the H2' atom as the original material (0.50 ppm) (**Figure 5.10**), suggesting that the observed reaction did not produce the starting materials (adenosine 3'-phosphate and TB). The thiamin portion of the fragmented A-TB, however, showed significant change in the ethylene bridge and thiazolium methyl protons, 0.97 ppm and 0.80 ppm respectively (**Figure 5.11**). This suggests the reaction or fragmentation was the thiamin portion of the A-TB adduct .

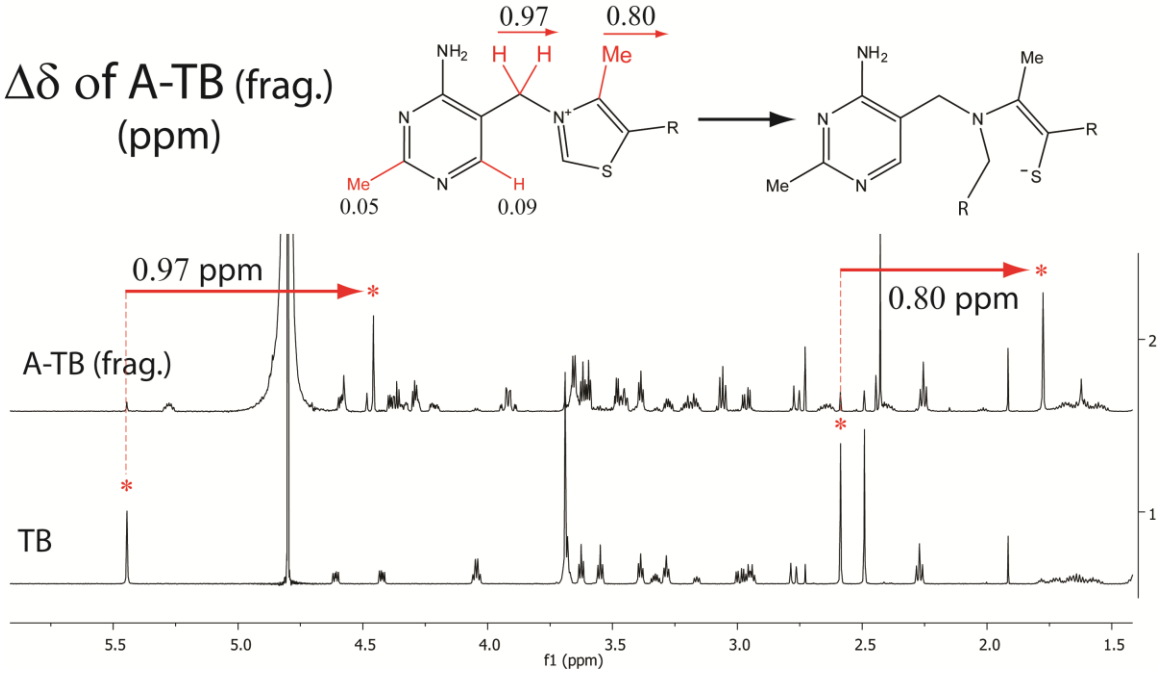

Figure 5.11. Comparison of ^{1}H NMR of fragmented A-TB and ^{1}H NMR of TB

(a) Structure of thiamin with the change in chemical shifts of the indicated protons from a TB standard. (b) NMR spectra of A-TB (frag.) and TB with the ethylene bridge and thiazolium methyl peaks indicated with a red asterisks.

5.7. Conclusion

In attempts to select for α-keto acid decarboxylase ribozymes, a parasitic RNA population was inadvertently selected for that exhibited TB reactivity that was *independent* of the attached suicide substrate (*Lnk*PB). Since both LnkPB and the phosphorothioate were not required, subsequent selections were performed in their absence and the conditions were reduced to 0.5 mM TB and a 1.5 hour incubation, resulting in activity near background. By round 18, the RNA population had partially recovered, with ~4% of the RNA binding to the bead by round 20. Round 12 and round 20 DNA was cloned and sequenced for analysis.

Of the sequences cloned from round 12 and 20, ~50% of the sequences (classes A & B) appear to be from a common ancestor or perhaps have, through convergent evolution found common sequence homology. The most distinctive aspect of these

sequences is near their 3'-end where the constant region is base-paired (partially by the constant region folding back on itself and partially by the randomized sequence) leaving a conserved G-A mismatch. Clone 2 of round 20 (2x20 from class A) was chosen for further analysis.

2x20 RNA was shown to react with TB by exhibiting a streptavidin-dependent shift in a denaturing polyacrylamide gel. To find where TB reacts with the RNA, TB-reacted 2x20 RNA was purified from a streptavidin shift gel and a partial hydrolysis footprinting assay using this RNA shows a discontinuity starting at residue A105 suggesting TB is reacting with this residue.

To further characterize this mysterious reaction, mass spectroscopy was attempted on the minimal adenosine-TB adduct by digesting TB-reacted 2x20 with P_1 nuclease and HPLC purifying the modified nucleoside away from the bulk of the unmodified nucleosides. Although the adduct appeared to be relatively stable, surviving the RNA elution out of the polyacrylamide gel, the P_1 nuclease digestion and HPLC purification, an intact mass spectrum was not obtained. This may have been because the adduct degraded in the mass spectrometer, however, it was not ruled out the adduct was somehow not detected due to ion suppression and/or perhaps poor ionization (although this seems unlikely).

In attempts to elucidate the structure of A-TB, NMR experiments were performed on the adduct. Unfortunately, it proved to be unstable in the NMR tube and degraded with a half-life of ~24 hours. In this time however, a ^{1}H and ^{1}H-^{1}H COSY spectra were obtained of the intact adduct. Although many of the peaks of these spectra were un-assignable (mostly of the PEG_3-biotin portion of the adduct), the adenosine's ribose protons were assigned. Overall, the sugar and base protons have a similar chemical shift to that of an adenosine control, however, the H2' proton had a chemical shift 0.70 ppm downfield. This is indicative of an electron withdrawing group nearby causing desheilding. NMR experiments were also performed on the degraded sample and the chemical shift of the H2' proton changed little, suggesting the reaction did not result in starting materials.

A likely mechanism can be suggested based on these data, by which the C2 hydroxyl attacks the thiazolium ring. This would result in an unstable RNA adduct, however, if the thiazolium C2H were to be oxidized, this would result in a stable adduct (**Figure 5.12**). A reaction of this type can be seen in the vitamin coenzyme NADH, where the hydride transfer restores aromaticity, driving the reaction forward. If this mechanism proves to be correct, this would implicate thiamin in redox chemistry for the first time.

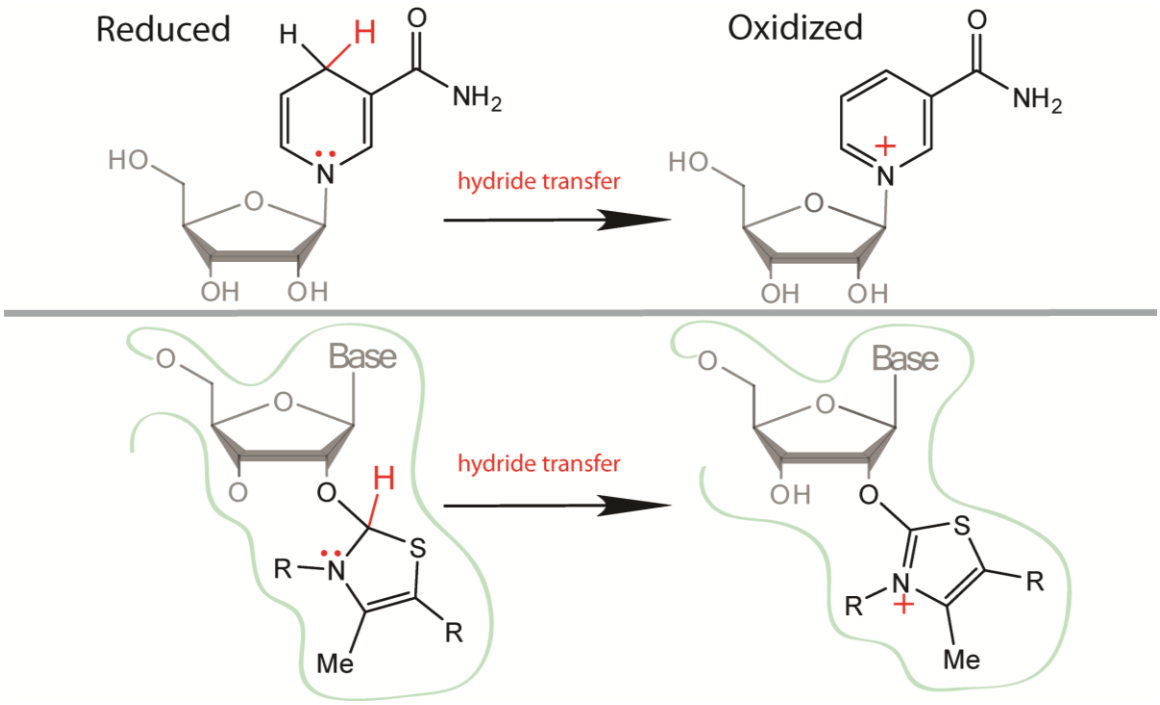

Figure 5.12. **Hydride transfer of NADH to form NAD+ compared to the hypothesized reaction of 2x20**

(top) Through a hydride transfer, the vitamin NADH is oxidized to NAD+. This reaction restores the aromaticity of NAD+. (bottom) hypothesized reaction catalyzed by 2x20. The 2'-OH would first attack the C2 carbon of the thiazolium ring. This adduct would be stabilized by a hydride transfer to restore the aromaticity of the thiazolium ring.

6. Negative selections and charactarization of dc4 ribozyme

6.1. Positive and negative selections for selection of *Lnk*PB *dependent* RNAs

Figure 5.2 plots the progression of RNA enrichment (percentage of RNA bound to the beads after repeated washings) after addition of negative selections for 13 rounds of *in vitro* selection. Negative selections were initiated with round 7 RNA, however, in spite of the negative selection, a significant population of parasitic RNAs were detected by round 8, that showed reactivity with ThBi even in the absence of *Lnk*PB conjugation to the RNA. To enrich for RNA populations that genuinely required *both Lnk*PB and ThBi (TB was renamed ThBi during the positive selections), one or more negative selection steps (**Figure 6.1**, gray bars) were incorporated into each of the selection rounds 9-13. In such negative selections, the GMPS-primed RNA pool, without linking to *Lnk*PB, was reacted with ThBi in SR buffer, and those RNAs that bound to streptavidin beads were discarded from the pool. The residual RNA pool was then conjugated with *Lnk*PB, in order to carry out the subsequent positive selection step. Despite incorporating up to three sequential negative selections in rounds 7-11, significant populations of *Lnk*PB-independent RNAs had to still be systematically eliminated. By round 12, however, the parasitic population had been significantly diminished, as indicated by the emergence of a substantial population of RNAs (red bars in **Figure 6.1**) that rigorously required both *Lnk*PB and ThBi in order to bind to the streptavidin beads. Such RNAs from Round 13, that required both *Lnk*PB and ThBi, were cloned and sequenced.

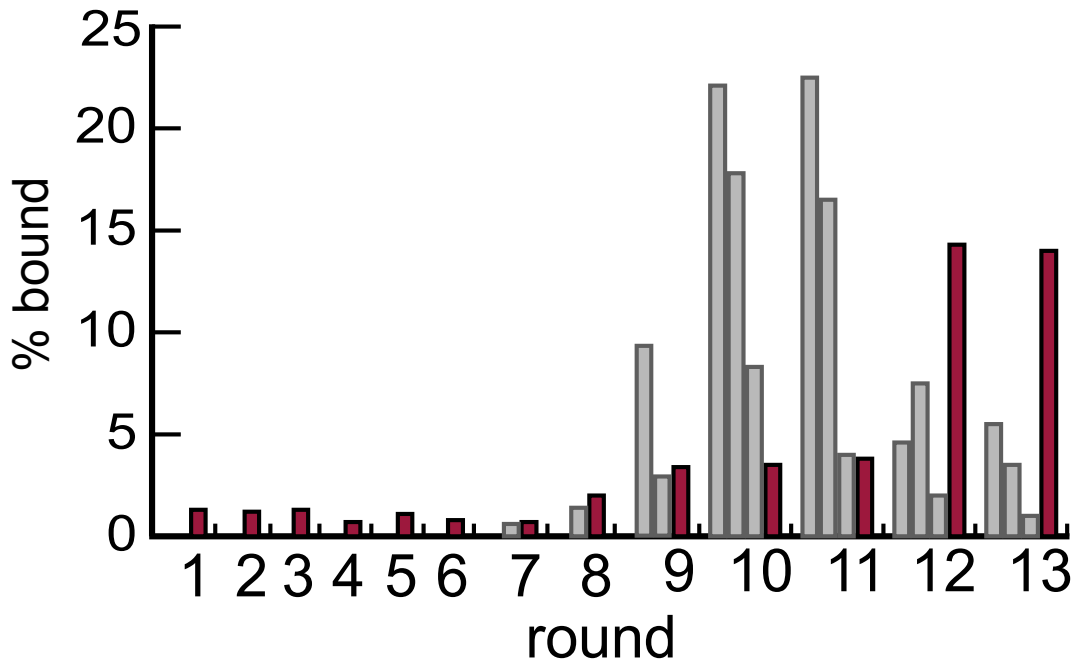

Figure 6.1. The progress of in vitro selection

Percentage of the overall RNA pool, in a given round, capable of binding streptavidin beads. For a positive selection (red), *Lnk*PB•RNA was incubated with 10 mM ThBi (rounds 1-6) and 2.5 mM ThBi (rounds 7-13). Negative selections were performed 0-3 times, in each case by incubation with ThBi prior to linkage of *Lnk*PB to the RNA. The first and second negative selections were performed using 10 mM ThBi incubations. The final negative selection of each round was always performed with 2.5 mM ThBi, to enable direct comparisons with the subsequent positive selection (carried out with 2.5 mM ThBi on RNA conjugated with *Lnk*PB).

6.2. Analysis of positive selection clones

Round 13 DNA was cloned and 25 colonies were sequenced. A search for sequence homology (**Table 6.1**) found little homology between individuals, and only one sequence appeared multiple times in the sampling (clones 4, 14, and 20 had the same sequence). **Figure 6.2.a** shows that RNAs from three of the selected clones (numbers 4, 15 & 19), when tested for their absolute requirement for both *Lnk*PB and ThBi, displayed precisely such a dual requirement in a streptavidin-dependent gel mobility shift (streptavidin gel-shift) assay. Clone 4 RNA, named “dc4”, was chosen for further characterization.

Table 6.1. The nucleotide sequences of the N70 region of 26 clones from Round 13 of the in vitro selection experiment.

Clone #	
1	GGGGCGGCGAUUUGAGAGGUGGUCUUGCGGCGUUAAGGUACGCGCCGGAACAUUGUUUAGGGUAAUGGCC
2	GUAGCUGGAGUGUACUGAGGGCAGUGCUGGCUUAAAGGCCGCGGUGGGAACGGCUACUUCGAUUAAUGC
3	UCGUUUCGAGAGACGGCUCGUUGGGUCUCAUAUAGGCGUGACGAAGUGGAUCGCCCGGUGGUGUGUGUUG
4	UUAUUUGUGAUUGAUAAAUUAUCAAUAAUUAAUGCAGGACUAUGAGCGUGCCAGUCUUCGCUGGAGUGAC
5	UCGUUGUACCGCAAUGGCGGUAUGCCAGAGUGCGGCUCUUUCGCUGAAGUGUGGGGGACUUUAGGCGUG
6	CAUUGGCGUUGGAGCAGGUGUCCUUCGUGUUUUGGCAUGAUGGGUGGGUGGCCGUAAGGGUCCGGGCGUC
7	AACGGUGCGGGCUGGCGGCUAUUUGAACUAGUGGGUGGGUAUUUGUCAUCUAGCACCUGUAUAUGUGC
8	CGCUACGGAUCACAUGCGCGCGACAGCUGAGUGGAUCAGUCGGAGAAUCGCAGCUUAAAGGUGGGAGGCU
9	AUCGGAAAGAAUUGUAAAUCCUUGCAGUGAAUUGUCUGCACUACGAGUGUGGAGCGCGGGAGAGCCAGU
10	UUUACGGUGCGUUGGAUCGAUUUGGUCACCGGGGAGGACCGGACGGGUUCGGGAAUGGGUUGAUGGCUUC
11	CGUUGGUGUCGCGAGCGUUCGUGUACCUUGGGGUGGGGUUUUGUCCGGACGUCGAACAUAACAUAGCACU
12	CAGAUGGGUCAUGGCGUCUGGACUUCAGUGACCGGCGGUGAUGGGAUCUAAGGGCGCGAUAUCGUCGUUGU
13	UUGAGCUAAGAUAGAUAUGCUAAUGCUUUAUGCAUUGUGGUGUAACCGGCUGUGUGAGGAGCGGGUGGCU
14	UUAUUUGUGAUUGAUAAAUUAUCAAUAAUUAAUGCAGGACUAUGAGCGUGCCAGUCUUCGCUGGAGUGAC
15	GAGUAUUUCGAGCGCACGCGCGGUGGUAGGUGGCGGUUAGCUUUGCAUACGAGUGCAGUGAUUCACUGAGU
16	GGGGGUGCUAGGGUUUGCUCUAGUGAGGAGUCUACUCGUGACUUAAGCAGGUGCUGUUGUCUGUUGCGUU
17	UUGCUGGUGUGUCAGAGGUGUGACCGGAGGGCCGGUUUUGUGUGAGAGACCUCUGCACGGGUAGGCACU
18	GUAGAGGCCAAAGUAGGUUUUCGGGCCGUGCUCCCUUGUAAUGGCCAUGUGGAUCAUCGGUGGUGGUGUU
19	AUUAUGUCAGGACGUGAUUUGAGUAUUGAUCGUGUAUACCGCCUAGAGUGUAGGUAGUGGGGCCUAGGUU
20	UUAUUUGUGAUUGAUAAAUUAUCAAUAAUUAAUGCAGGACUAUGAGCGUGCCAGUCUUCGCUGGAGUGAC
21	GCGCCUCGGCCUAAGGGUCUAGGGGGGUGAUGAGCUAAGCGUGCCUGACUGUUGGUUCGUGGGAGAUUGC
22	CACUGCCACAUGGGGCUGUCGUUGGUCGAGCCGUGUGGGACUGUCCUUUGGCUGUGGAGUGGGCGCAUGC
23	GCUGAACAGAAGUAUCCCGGCUUGAUGCAGAUGCGGACUGAUUAGUGAGCACCGGAGGUAGGCGUU
24	UUCGGGUUUAAGAGUGCGUUUAAGGCUGCUGGACGAUCGGUUGGGCGUAAUAAGUCACGUUCUGUGGUAG
25	AGAUGCGGUCUGGGUUAGCUAAAUGCUUCUGGCCGCUCAUUCGGUGAUGCCGGAUCGAAAGUAGGGGC
26	AGGUUUGAGAUGGCGUUAAGCUGGAGUGUGCUCGAGAAGUGUUGUUUUGUAGGCGGGUUUGAGCGGCAU

6.3. Control experiments

The streptavidin gel shift assay was used to evaluate the catalytic properties of dc4 RNA, relative to those of an equivalent amount of RNA from the initial pool used for *in vitro* selection (G0 RNA). In the standard conditions for this assay *Lnk*PB•5'-RNA is allowed to fold in the Standard Reaction Buffer (SR Buffer: 50 mM Li-HEPES, pH 7.5, 200 mM NaCl, 200 mM KCl, and 40 mM $MgCl_2$), made up to 10 mM ThBi, and incubated

at 22°C. The RNA pool is then separated from excess ThBi by successive isopropanol precipitations, re-dissolved in a buffer consisting of 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 mM EDTA, 0.5 mM DTT and 2 µg of recombinant streptavidin, and incubated for 10 min before analysis by denaturing PAGE. **Figure 6.2b** shows a time-dependent reaction of dc4•*Lnk*PB with ThBi, under single-turnover conditions, with a plateau reached within 8 hours. **Figure 6.2c** shows key control experiments, which reveal that with the omission of any of the following: streptavidin, ThBi, or *Lnk*PB, no gel shift of dc4 RNA is observed. Conversely, no absolute requirement is found for magnesium, sodium/potassium chloride, or indeed, the Li-HEPES buffer itself (**Figure 6.2c**). The presence of Mg^{2+}, however, does enhance the reaction rate, with saturation of reaction rate seen as a function of Mg^{2+} concentration (with half-maximal activity reached at ~5 mM Mg^{2+}: **Figure 6.2d**). Thus, magnesium is not strictly required for catalysis, though it likely stabilizes dc4•*Lnk*PB's catalytic conformation. In contrast to the above results, G0 RNA, derivatized with *Lnk*PB and incubated with 10 mM ThBi for as long as 24 hours, does not show any streptavidin-dependent gel shift.

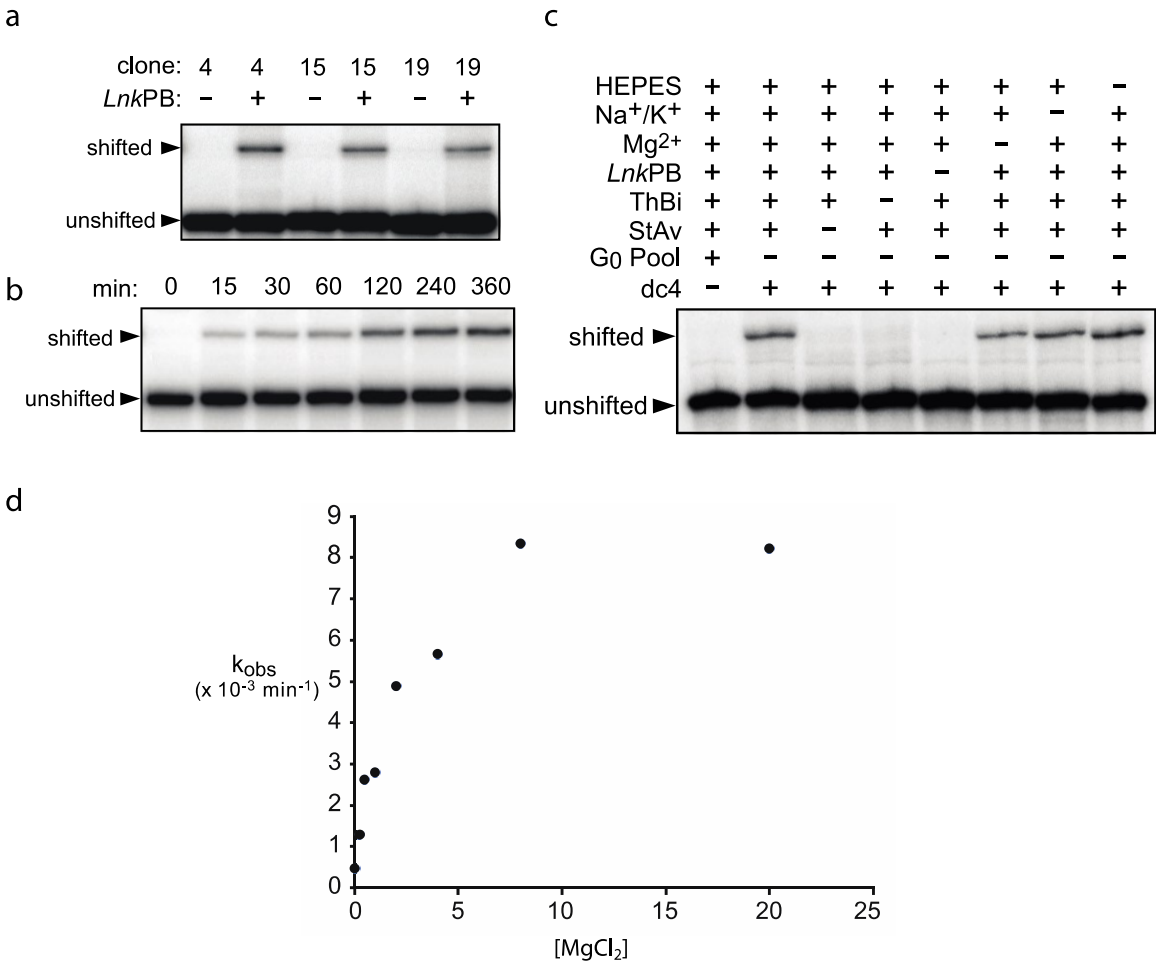

Figure 6.2. Control experiments of select clones

(a) Streptavidin gel-shift of RNA from clones 4 (dc4), 15, and 19 with (+) and without (-) conjugated *Lnk*PB. In the standard conditions for the streptavidin gel-shift assay *Lnk*PB•5'-RNA is allowed to fold in the Standard Reaction Buffer (SR Buffer: 50 mM Li-HEPES, pH 7.5, 200 mM NaCl, 200 mM KCl, and 40 mM $MgCl_2$), made up to 10 mM ThBi, and incubated at 22°C. The RNA is then separated from excess ThBi by successive isopropanol precipitations, re-dissolved in a buffer consisting of 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 mM EDTA, 0.5 mM DTT and 2 µg of recombinant streptavidin, and incubated for 10 min before analysis by denaturing PAGE. (b) A time-course for reaction, under standard conditions, of RNA from dc4 conjugated with *Lnk*PB, with 10 mM ThBi, at 30° C. (c) Control experiments to test the activity requirements of ribozyme dc4•*Lnk*PB. Lane 1: generation 0 RNA. Lane 2: dc4 RNA positive control, containing all selection reagents. Lanes 3-8: various control experiments, as indicated in the figure. (d) Decarboxylation k_{obs} plotted as a function of $MgCl_2$ concentration. dc4•*Lnk*PB was reacted with 10 mM ThBi in SR buffer, except containing varying concentrations of $MgCl_2$ (0, 0.5, 1, 2, 4, 8, 16, 40 mM).

6.4. Mass Spectroscopy of a modified nucleoside reaction product

If, as hypothesized, dc4•*Lnk*PB utilizes the thiamine moiety of ThBi for decarboxylation of the dc4-conjugated *Lnk*PB, generating the stabilized, decarboxylated, carbanion/enamine intermediate described by Jordan (D. J. Kuo & Jordan, 1983b), a characteristic mass for this stable intermediate should be identifiable by mass-spectrometry. First, the adduct of ThBi with dc4•*Lnk*PB RNA was recovered from appropriate streptavidin-shifted gel bands, then subjected to complete digestion by nuclease P1 (to generate 5' nucleoside monophosphates) followed by treatment with alkaline phosphatase (to remove the 5' phosphate from the above nucleotides). Cumulatively, these enzymes would be expected to release a digestion end-product that was a composite of the 5'-most guanosine of the RNA, linked through its 5'-phosphorothioate to the decarboxylated adduct of *Lnk*PB and ThBi (**Figure 6.3.a**). We purified this digestion end-product (marked in red in **Figure 6.3b**) away from unmodified nucleosides using HPLC, and analyzed it using high resolution ESI-MS for its molecular mass. A major peak of 691.7180 m/z (z=2; positive ion mode) was observed (**Figure 6.2c**). This value matched, within 0.4 ppm, the calculated m/z value of our expected product (691.7177). The latter corresponds to a molecular mass of 1,383.4354, which equals the masses of ThBi plus 5'-guanosine•*Lnk*PB minus the mass of carbon dioxide.

6.5. Mass Spectroscopy of ($^{13}C)_3$-*Lnk*PB product

To demonstrate unambiguously that dc4 is indeed a decarboxylase ribozyme, we re-synthesized *Lnk*PB using (^{13}C)-pyruvate (in which all three carbon atoms were ^{13}C) to make a triply ^{13}C- labeled *Lnk*PB substrate (($^{13}C)_3$-*Lnk*PB). If decarboxylation of *Lnk*PB is indeed a reaction being catalyzed by dc4, then the digestion end-product from the dc4•($^{13}C)_3$-*Lnk*PB and ThBi adduct (**Figure 6.3.a**) should retain two of the ^{13}C carbons from the original pyruvate, with loss of one ^{13}C carbon in the form of $^{13}CO_2$. A high-resolution peak of 692.7222 m/z (z=2; positive ion mode) for this product was obtained (**Figure 6.3.c**). The expected value for a decarboxylated product that retains two ^{13}C labels but has lost the third as $^{13}CO_2$ is 692.7211 m/z (z = 2), indicating an error of 1.6 ppm from the measured value. Thus, the digestion end-product in this case was indeed

the decarboxylated product. The measured m/z values (z = 2) of the ^{13}C-containing and ^{12}C-containing decarboxylated products differ by ~1, indicating a difference in molecular mass between them of 2.

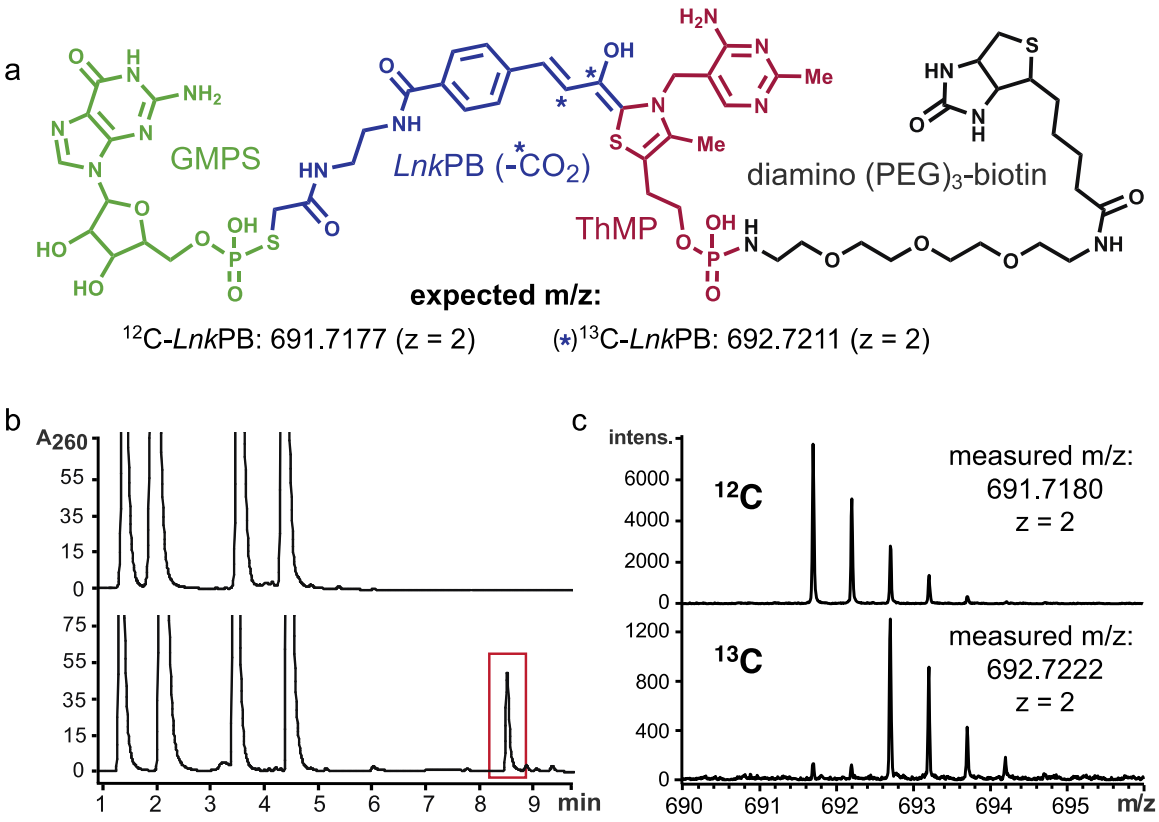

Figure 6.3. Mass spectroscopy of the decarboxylation product.

(a) Expected decarboxylation product after complete digestion by nuclease P and alkaline phosphatase of the adduct of ThBi with dc4•*Lnk*PB. It consists of 5'-GMPS conjugated to the decarboxylated *Lnk*PB, trapped as a stabilized carbanion/enamine adduct with ThBi. Asterisks indicate the ^{13}C atoms when [^{13}C]$_3$-*Lnk*PB is used. (b) HPLC trace of the digestion end products of P1 nuclease- and alkaline phosphatase-treated, ThBi-*unreacted* dc4•*Lnk*PB ribozyme (top). HPLC trace of the digestion end products of ThBi-*reacted* dc4•*lnk*PB, which was first purified from a streptavidin-shift gel (bottom). The four large peaks to the left show the four standard nucleosides found in RNA. The putative decarboxylation product peak (shown within the red box) was collected for ESI-MS analysis. (c) Mass spectra, showing mass/charge (m/z) values corresponding to the exact decarboxylated products (shown in (a)). The result from experiments with non-isotope-labeled *Lnk*PB is shown on top, whereas experiments using ^{13}C-labeled *Lnk*PB (the product retaining the expected two ^{13}C atoms from the original three present in [^{13}C]$_3$-pyruvate) is shown on bottom.

6.6. Thiamin itself is a cofactor for dc4

The dc4•*Lnk*PB ribozyme was *in vitro* selected to utilize a thiamin *derivative* (ThBi) for its decarboxylase activity, the biotin modification being necessary to supply a "handle" to enrich catalytic sequences out of a random RNA library. We were interested to know whether the dc4•*Lnk*PB ribozyme could also utilize the underivatized vitamin, thiamin, for the same task. We initially examined the potential of thiamin in this system by using a kinetic competition assay (**Figure 6.4a**). ThBi-mediated decarboxylation rates (quantitated by the streptavidin gel-shift assay) were measured with an unvarying concentration of ThBi (10 mM) mixed with different concentrations (0-20 mM) of thiamin (Th). **Figure 6.4b** shows that as Th concentration increased, the yield plateau of the ThBi product decreases progressively. This behaviour was consistent with Th competing as a *substrate* with ThBi, rather than inhibiting the ThBi reaction (the competing reactions are schematized in **Figure 6.4b**).

Since under our reaction conditions the concentrations of both ThBi and Th were orders of magnitude higher than the dc4•*Lnk*PB concentration, the curves shown in **Figure 6.4b** could be fit to obtain pseudo-first order rate constants (k_{obs}). The percentage of product reacted in 10 mM ThBi was plotted as a function of time and fit to the first order exponential equation:

$$P = (P_{max})(1-\exp(-k_{obs}t))$$

Where P is percentage of product dc4•*Lnk*HY-ThBi from the total RNA, P_{max} is the maximum product percentage, k_{obs} is the observed psudo-first order rate constant and t is time in minutes. For the reaction of ThBi alone with dc4•*Lnk*PB the k_{obs} value was calculated to be 9.2×10^{-3} min^{-1}.

In order to compare relative rate accelerations, the second order rate constant for the Th reaction, k'', was derived from the relationship: $k_{obs} = k'$ [ThBi] + k'' [Th]. The second order rate constant for the ThBi only reaction was calculated by the equation, $k' = k_{obs}$ / [ThBi]. The rate constant k′ was calculated to be 0.015 ± 0.021 M^{-1} s^{-1} and k'' was 0.023 ± 0.009 M^{-1} s^{-1}. Thus, although the *in vitro* selection for the dc4 ribozyme had been carried out with ThBi, underivatized Thi reacted with dc4•*Lnk*PB ~1.5-fold faster

than ThBi. Both ThBi and Th showed a ~10^5-fold rate enhancement over the non-enzymatic reaction, estimated at 3 x 10^{-7} M^{-1} s^{-1}, between pyruvate and thiamin, corrected to pH 7.5 (Kluger et al., 1981; Washabaugh & Jencks, 1988).

Table 6.2. Competition kinetic data

k_{obs} data is from the mean of three separate experiments (data not shown). The *k*' error was calculated from one standard deviation from the mean of three separate experiments. Error for *k*" was calculated from one standard deviation from the mean of each data set.

[Thiamin] (mM)	**0**	**1.25**	**2.5**	**5**	**10**	**20**
k_{obs} (s^{-1})	**0.0001537**	**0.0001875**	**0.0002442**	**0.000293**	**0.000317**	**0.000415**
k' **(M^{-1} s^{-1})**	**0.015 ± 0.021**					
k" **(M^{-1} s^{-1})**	**0**	**0.02703**	**0.03618**	**0.02776**	**0.01554**	**0.01309**
k" **(M^{-1} s^{-1}) mean**	**0.0239 ± 0.009**					

For further analysis of the reaction of Th with dc4•*Lnk*PB, ESI mass spectroscopy was performed on the presumed adduct between them. Owing to the lack of a biotin "handle" in this adduct, the reacted RNA could not be purified away from unreacted. However P1 nuclease/alkaline phosphatase digestion of the total RNA nevertheless permitted the isolation of the minimal reaction product away from unmodified nucleosides, using HPLC (**Figure 6.5a**). The expected chemical structure of the decarboxylated thiamin-*Lnk*PB•dc4 product is shown in **Figure 6.5b**. ESI mass spectroscopy was performed on this product, and a m/z value (at z = 2) of 452.6320 was indeed observed (**Figure 6.5c**), with a 10.8 ppm error from the expected value of 452.6271.

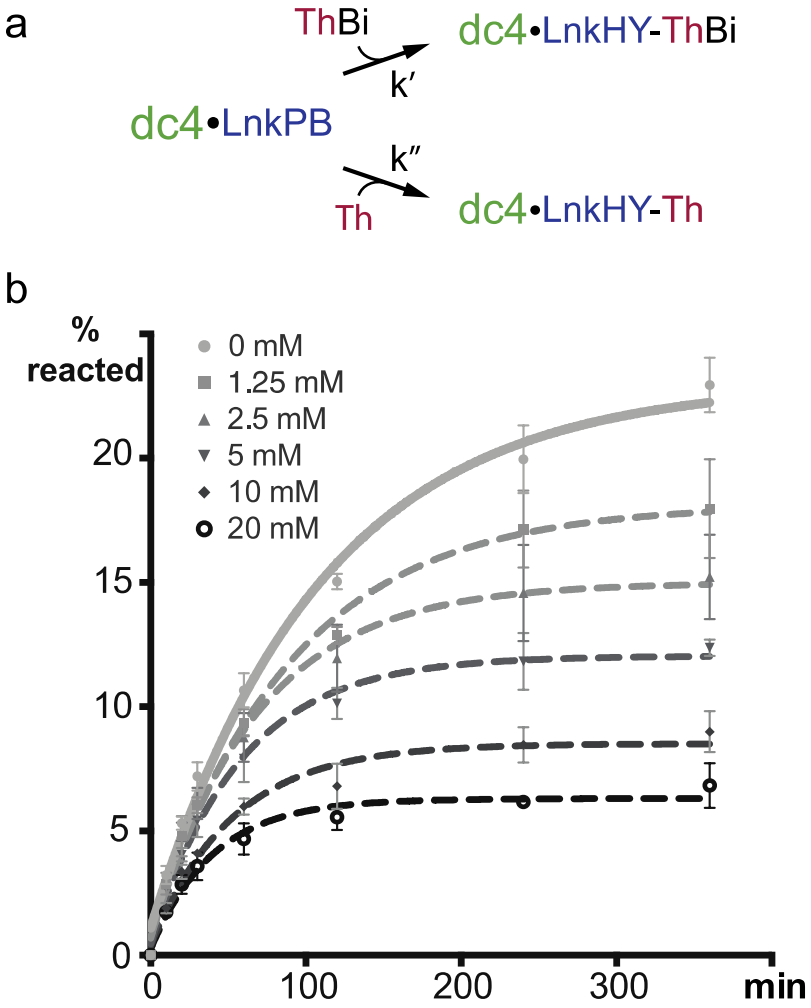

Figure 6.4. Unmodified thiamin is an effective coenzyme

(a) Kinetic scheme showing the competition of thiamin (Th) and ThBi in a mixture reacting with dc4•*Lnk*PB. Error bars were calculated from the standard deviation of the average of three separate experiments. (b) Time-course of yield of dc4•*Lnk*HY•ThBi, determined by quantitation of streptavidin-shift gel electrophoretic bands obtained from reaction with dc4•*Lnk*PB of mixtures of ThBi (10 mM) and thiamin (0-20 mM). Averages of three independent sets of reactions are shown.

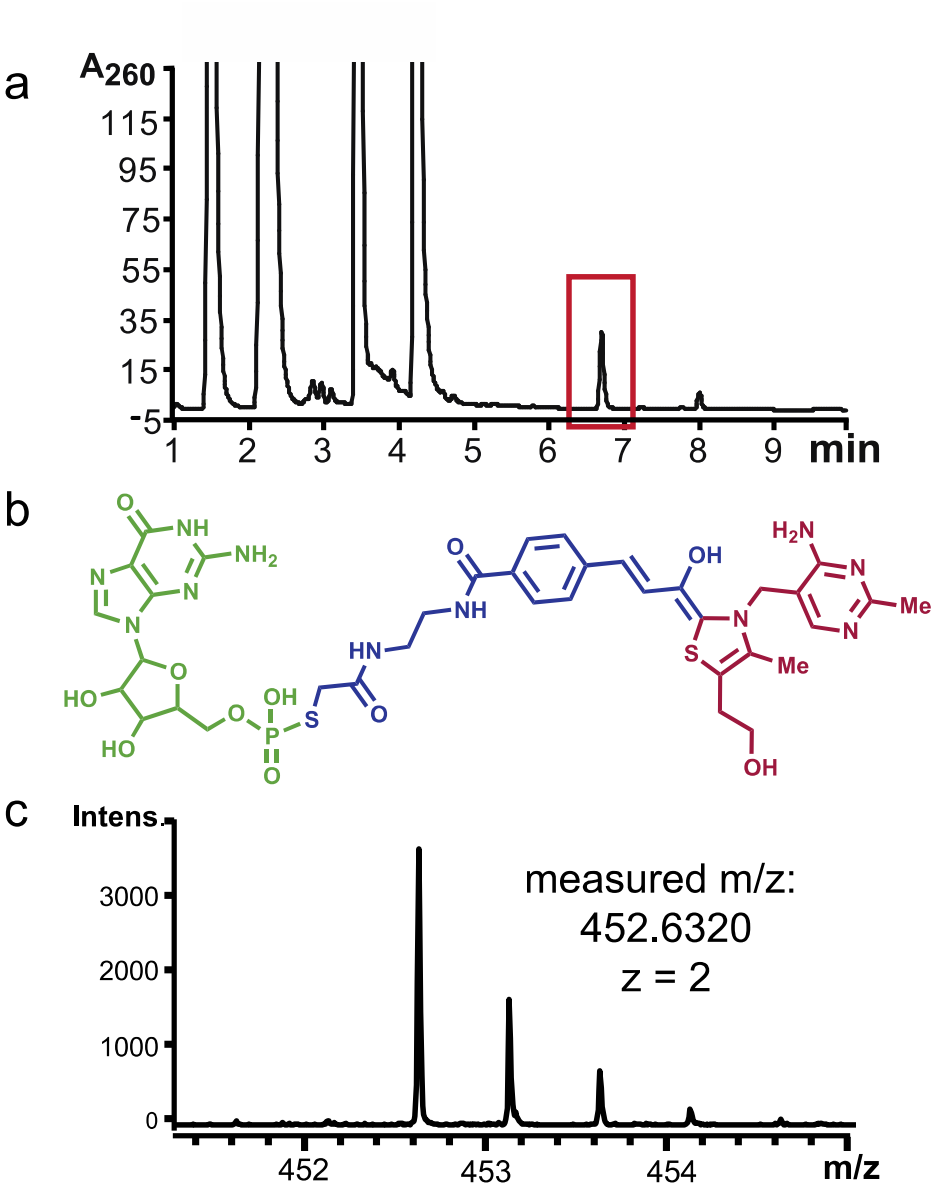

Figure 6.5. **HPLC purification and ESI-MS of the thiamin reacted product**

(a) HPLC trace of the P1-nuclease and alkaline phosphatase digestion end products of the dc4•[^{13}C]$_3$-*Lnk*PB ribozyme, following its incubation with 20 mM thiamin (Th). The four peaks to the left show the four standard nucleosides present in RNA; the peak highlighted in red was collected for ESI-MS analysis. (b) The structure of the digestion end product (GMPS•[^{13}C]$_2$-*Lnk*HY•Th) expected from P1-nuclease and alkaline phosphatase digestion of dc4•[^{13}C]$_3$-*Lnk*PB ribozyme reacted with thiamin (Th). (c) High resolution ESI-MS spectrum of the putatively decarboxylated digestion end product, showing a m/z value (z =2) consistent with the precise molecular mass and formula of GMPS•[^{13}C]$_2$-*Lnk*HY•Th (structure shown in (b)).

6.7. Secondary structure and mutagenesis of the dc4 ribozyme

The program Mfold was used to test for likely secondary structures of dc4 RNA. Two optimal structures, i and ii, of comparable folding energies, were calculated (shown in **Figure 6.6a**). The structures share a 46-nt stem-loop (P2 in structure i; P1 in structure ii) but few other similarities. Based on these two putative folds, nine P1 deletion/mutation constructs of dc4 RNA were made (**Figure 6.6a**) that might be expected to compromise the structure of fold i or ii or both. Mutations **1-5** progressively shortened the major stem present in both folds (stem P2 in structure i and stem P1 in structure ii). Mutation **1** mutated the predicted GAUA terminal loop of this stem to the stable UUCG tetraloop, while the remaining four mutations shortened and UUCG-capped the stem. Each of these mutant RNAs was conjugated with *Lnk*PB and tested for coupling activity with ThBi. It was found that mutants **1-3** retained activity, while mutant **4** completely lost activity (**Figure 6.6b**), suggesting that the shared stem-loop of structures i and ii forms but is largely unnecessary for catalytic activity. Mutations **6** and **7**, which changed terminal loops in structure i to UUCG, maintained activity (**Figure 6.6b**). In fold ii, mutations 6 & **7** might be expected to notably destabilize the entire P2 stem. These mutations, however, retained activity. Therefore, it was concluded that structure i might be the more plausible fold, although this requires further testing. Mutations **8-10**, which shortened and mutated different elements within structures i and ii, all abolished activity**.** Based on the above mutational data, we propose that Mutant **3,** based on structure i, represents a likely 'minimal' fold for this ribozyme (**Figure 6.6c**).

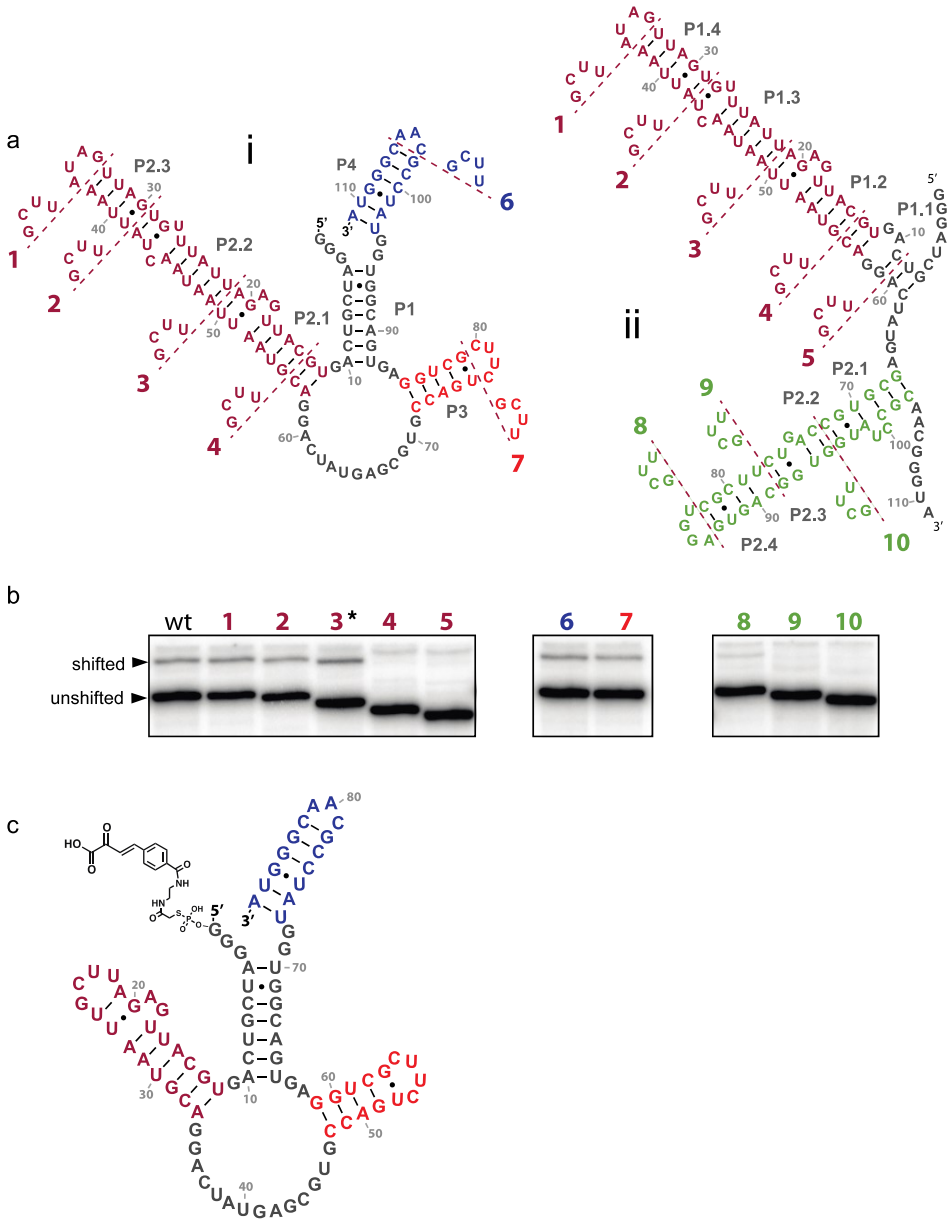

Figure 6.6. Secondary structures for dc4 RNA and its mutants.

(a) Fold A and Fold B for dc4 RNA. Mutants 1, 6, 7, and 8 show mutations of putative terminal loops to UUCG tetraloops. Mutants 2-5, 9, and 10 represent truncations of putative stems, along with such truncations being capped with UUCG tetraloops. (b) Streptavidin gel-shift patterns obtained from the unmodified ribozyme ("wt") and from the various mutant ribozymes. Mutants, containing attached *Lnk*PB, were reacted in SR buffer for 8 hours. The mutant ribozyme, whose activity is shown in lane 3 (labeled with an asterisk) corresponds to the minimal ribozyme, whose structure is shown in panel c. The standard conditions used for the streptavidin gel-shift assay

are described in the legend for **Figure 2**. (c) A minimal structure of folded dc4 RNA, corresponding to Mutant **3**, folded according to Structure I, and shown conjugated to *Lnk*PB.

7. Discussion

7.1. How effective a ribozyme is dc4?

If the assumption is made that the rate-determining step for a nonenzymatic addition of ThBi to *Lnk*PB corresponds to the known rate limiting step of thiamin addition to pyruvate, this ribozyme promotes a ~10^5-fold rate enhancement over the uncatalyzed rate. This would be a relatively strong performance for a natural ribozyme (Doudna & Cech, 2002); however, *in vitro* selected ribozymes and DNAzymes that show substantially higher rate enhancements have been reported recently (Chandrasekar & Silverman, 2013). In its present form dc4 is a single-turnover ribozyme. Instead of turning over substrate it forms a relatively stable adduct with the decarboxylated suicide substrate, *Lnk*PB.

7.2. How does the dc4 ribozyme activate thiamin?

Four factors may contribute to this activation. First, the ribozyme may bring the suicide substrate and thiamin reactants into close proximity. Second, free thiamin is a very weak catalyst under our experimental conditions owing to the very high pK_a,17-19, of its C2 proton (Kemp & O'Brien, 1970). Proteinaceous pyruvate decarboxylase enzymes have evolved to perturb the thiamin thiazolium pK_a within their active sites by providing a non-polar environment (Jordan et al., 1999; S. Zhang, Zhou, Nemeria, Yan, Zhang, Zou, & Jordan, 2005b). From the Born or desolvation effect, it has been hypothesized that the reduction of the dielectric constant within the enzyme mileau, stabilizing the neutrally charged ylide form of the coenzyme or increasing the ΔG of the dipolar LThDP intermediate (destabilizing the ground state). It is possible that dc4 may also provide an environment with a decreased dielectric constant; however, RNAs ionic backbone and lack of truly non-polar active groups may limit RNAs activating potential in

this regard. Third, thiamin bound within proteinaceous pyruvate decarboxylase enzymes is invariably found in a high energy, sterically hindered, "V"-shaped conformation (Frank, Leeper, & Luisi, 2007) rather than a lower energy conformation, such as the extended "H" conformation adopted by free thiamin in solution (Kraut & Reed, 1962). The V conformation facilitates thiazolium C2 deprotonation by way of an intramolecular proton transfer. Key protein residues help stabilize the rare imino tautomer of thiamin's 4-aminopyrimidine, which then serves as a proximally positioned Brønsted base to extract the C2 proton (Jordan et al., 1999; 2003; Kern, 1997). If ribozyme dc4 is binding the "V" conformation, it may be also stabilizing the IP form by hydrogen bond contacts or proton donation to the N1', in a similar way to the proteinaceous enzymes. Deeper studies on this system should help establish whether thiamin associated with dc4•*Lnk*PB does indeed adopt a V rather than a lower energy conformation and if so, whether it stabilizes the IP form of the coenzyme. Fourth, it is important to consider that what is strictly required for catalysis by thiamin is its thiazolium ring. It is conceivable that in dc4•*Lnk*PB, a RNA nucleobase may act as a Brønsted base for the C2 deprotonation, functionally substituting for thiamin's 4-aminopyrimidine moiety. Future structural, mechanistic, and mutagenesis studies should throw light on this tantalizing possibility.

7.3. Does dc4 bind thiamin in the "V" conformation?

This is a particularly interesting question because it would help resolve the origin of the 4-aminopyrimidine. Particularly, if thiamin or similar compound was indeed an important aspect of the RNA World, was the structure completely conserved during the transition from the RNA World to protein-dominated world? Was the 4-aminopyrimidine necessary in an RNA World or was the ribozyme responsible for its role? White proposed that thiamin perhaps started off as a nucleotide, which was incorporated into the RNA sequence. As a consequence, the resulting thiaminium nucleotide (**Figure 7.1**) would acquire properties of the N1′-methylthiaminium, shown by Jordan to be a superior catalyst. Would a thiaminium nucleotide be better than thiamin as a cofactor?

Some of these questions can be answered by further kinetic and binding studies of ribozyme dc4. Thiamin analogs may be used to deduce what parts of thiamin are necessary for its activity and if dc4 likely binds the "V" conformation. **Figure 7.2** shows

thiamin and select thiamin analogs that may be useful in determining how dc4 binds thiamin. If dc4 utilizes the 4-aminopyrimidine in acid-base chemistry, N1′-methylthiaminium would be expected to act as a superior cofactor (assuming the RNA is not inducing the tautomerization through protonating the N1 position). Likewise, oxythiamin would be expected to inhibit dc4 because of the 4-amino group is replaced with a hydroxyl group (oxythiamin is a potent inhibitor of many thiamin enzymes). Thiochrome would be expected to bind (**Figure 7.2, iv**) and perhaps inhibit dc4 only if it binds the "V" conformation. Lastly, if dc4 binds thiamin thiothiazolone similar to the YPDC and other thiamin utilizing enzymes, a positive CD peak would be expected at 330 nm. Together these analoges may be used to deduce how dc4 bind thiamin.

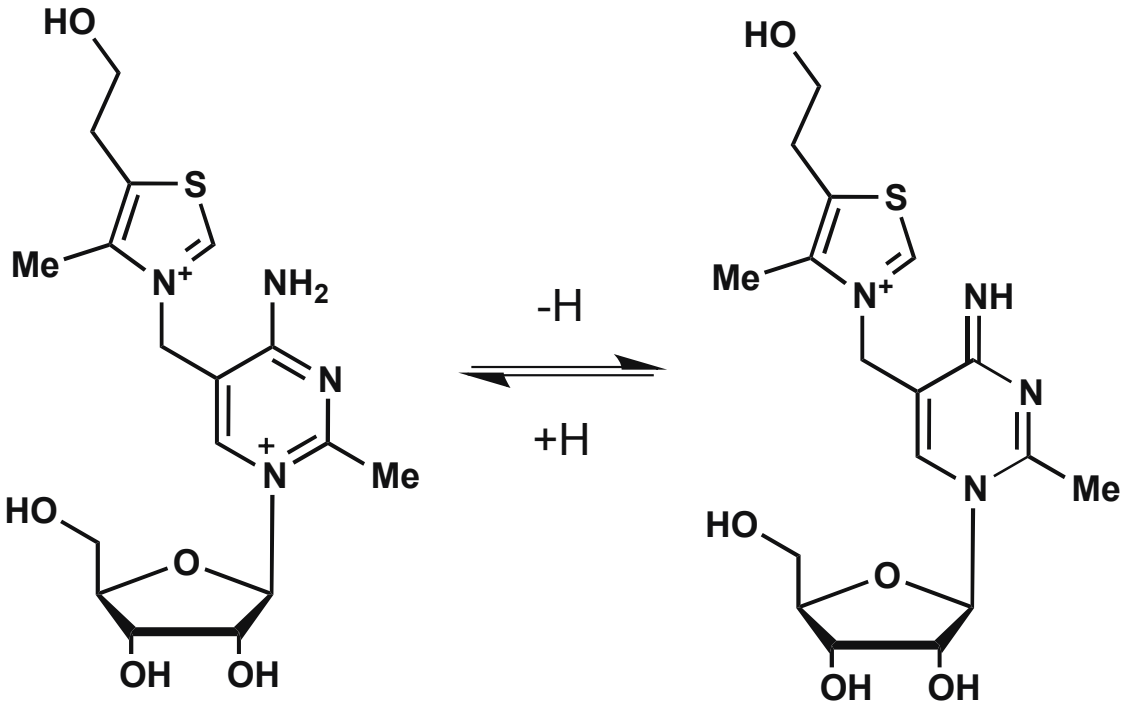

Figure 7.1. Thiaminium nucleotide would exhibit an equilibrium favored towards the imino form of the nucleotide.

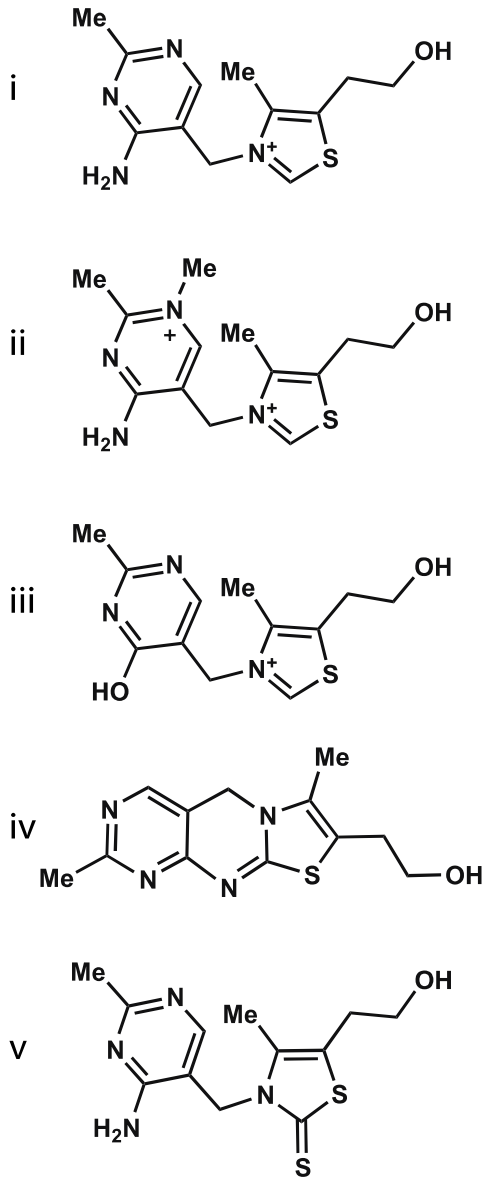

Figure 7.2. Structures of thiamin and its analogs.

The structure of thimain (i) compared to common thiamin analogs. If dc4 binds the "V" conformation, the analog N1'-methylthiaminium (ii) would be expected to be a better cofactor than thiamin alone (assuming the methyl group does not cause significant steric effects). If dc4 utilizes the 4-aminopyrimidine as an acid-base catalyst, as would be the case if it binds the "V" conformation, the analog oxythiamin (iii) would be expected to act as an inhibitor and not a cofactor. Thiochrome (iv) wold be expected to act as a wak inhibitor if dc4 bound thiamin the "V" conformation. Lastly, if dc4 binds the "V" conformation, analog thiamin thiothiazolone (v) would be expected to elicit a CD signal caused by its conformation chirality.

7.4. Comparisons to the ThDP riboswitch

To the best of my knowledge, *in vitro* selection experiments have not yet reported a thiamin binding RNA aptamer. However, ThDP or (TPP)-activated riboswitches have been reported (Winkler et al., 2002), and are among the most widely distributed riboswitches in nature (Tucker & Breaker, 2005). Both the TPP riboswitches and the dc4 ribozyme described here are thiamin-binding RNAs. The discovery of vitamin-binding riboswitches has stimulated discussion of their likely relationship with putative vitamin-utilizing ribozymes from primordial metabolisms (Cochrane & Strobel, 2008). However, high-resolution crystal structures of TPP riboswitches show the TPP ligand bound in an extended (non-catalytic) conformation (Edwards & Ferré-D'Amaré, 2006; Serganov, Polonskaia, Phan, Breaker, & Patel, 2006; Thore, 2006). See **Figure 7.3** for a comparison of the crystal structures of ThDP bound to PDC and the TPP-riboswitch. The riboswitch RNA does not contact the thiazolium ring of TPP. Therefore, no catalytically useful perturbation of the pK_a of bound thiamin (such as might be found in a thiamin-utilizing ribozyme) is anticipated in the TPP riboswitches.

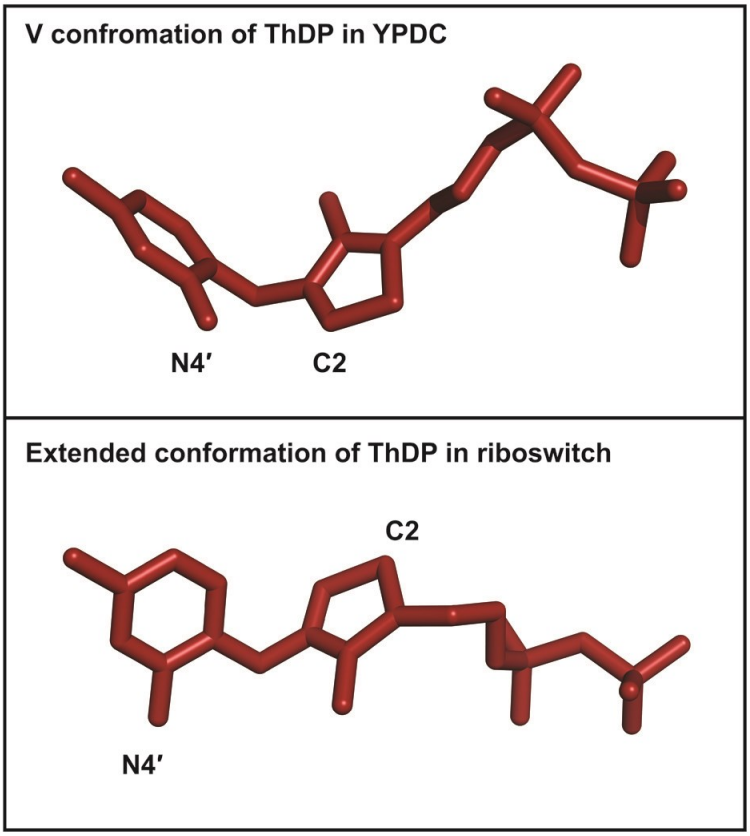

Figure 7.3. Comparison of ThDP bound to PDC (Arjunan et al., 1996) and TPP-riboswitch (Edwards & Ferré-D'Amaré, 2006).

7.5. Cyanide may be the progenitor of thiamin

Cyanide has been known since the 1800s to catalyze the decarboxylation of α-keto acids as well catalyze the benzoin condensation, forming a carbon-carbon bond between two benzaldehyde molecules to create benzoin. These reactions progress through very similar intermediates as the mechanisms of YPDC (YPDC in fact has been shown to produce the acyloin product from two equivalents of acetaldehyde (Chen & Jordan, 1984)). Recently, potassium cyanide was found to catalyze the same reaction as transketolase, a ligation reaction of $_D$-fructose-6-phosphate with glyceraldehyde-3-phosphate to form $_D$-erythrose-4-phosphate and $_D$-xylulose-5-phosphate (Breslow & Appayee, 2013). Cyanide was found to be a better non-enzymatic catalyst than thiamin or its derivatives in this regard owing to its significantly lower pK_a (9.2). Given that

cyanide and thiamin perform the same chemistry (cyanide better in the absence of enzymes and thiamin a much better catalyst with enzymes), it has been proposed that cyanide is a prebiotic catalyst whose role in metabolism was taken over by thiamin enzymes. I will accept this hypothesis; however, thiamin may have had a much earlier arrival than the time when proteins appeared as catalysts. This is supported by the fact that thiamin contains a nucleic acid component and we have now demonstrated RNA can use thiamin in catalysis similar to protein enzymes.

8. Conclusion

RNA may have been an older biopolymer than DNA and protein as posited by the RNA World hypothesis. Thi is thought to be so because RNA can both hold information and performs catalysis. Ribozymes do not intrinsically have the catalytic tools that protein enzymes have but may have use coenzymes to increase their catalytic repertoire.

Thiamin diphosphate is a catalytic coenzyme used by protein enzymes to perform diverse chemical transformations, that are at the core of modern metabolism. Most notably, ThDP utilizing enzymes are responsible for the decarboxylation of α-keto acids in a reaction that links major metabolic pathways including glycolysis, gluconeogenesis and the tricarboxylic acid cycle.

I performed selection *in vitro* for the discovery of ribozymes that utilize thiamin in the decarboxylation of an α-keto acid. A selection strategy was designed, using a pyruvate-like suicide substrate to trap the post-decarboxylation carbanion/enamine intermediate. The suicide substrate *Lnk*PB was designed so it could be attached to the random RNA library. Thiamin monophosphate was derivatized with biotin to make ThBi. RNA sequences able to decarboxylate their attached *Lnk*PB using ThBi were enriched using streptavidin coated beads.

Parasitic RNA sequences were unintentionally enriched for, that outcompeted for decarboxylase sequence space. Negative selections were incorporated to reduce the parasitic population of RNAs and to preferentially enrich for sequences strictly dependent on the *Lnk*PB substrate. Three clones from round 12 show a strict requirement for the attached *Lnk*PB. Clone four, called dc4, was chosen for closer analysis, and found to show a rigorous requirement for *Lnk*PB, ThBi, and streptavidin. However this ribozyme had no absolute requirement for Mg^{2+}.

Analysis of the reaction product was performed by electrospray ionization mass spectroscopy. The reaction product of the 5' terminal guanosine monophosphorthioate was isolated by ribonuclease digestion of the reacted dc4, and HPLC purification of the reaction product. High resolution mass spectra of both the *Lnk*PB and (3-^{13}C) labeled *Lnk*PB reaction adducts show the precise molecular mass to charge ratio of the expected decarboxylation product.

To see if thiamin can act as a cofactor in this reaction, kinetic competition experiments were performed. Holding ThBi concentration constant and titrating thiamin, the plateau of the reaction was shown to be reduced with increasing thiamin concentration while the k_{obs} increased, suggesting thiamin is competing in the reaction. A comparison of the second order rate constants from both the ThBi and Th reactions indicate that thiamin competes favourably, with a slightly higher rate constant. Mass spectroscopy of an adduct purified from dc4, reacted only with thiamin, confirmed thiamin as a coenzyme in the reaction.

This work supports the hypothesis, first proposed within the logical arguments offered by White and others, that coenzymes, particularly thiamin, are relics of an RNA World. Discovery of a thiamin utilizing ribozyme for the decarboxylation of a pyruvate-based suicide substrate suggests that RNA may have orchestrated metabolic processes in the RNA World.

References

Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., et al. (1996). Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast Saccharomyces cerevisiae at 2.3 A resolution. *Journal of Molecular Biology*, *256*(3), 590–600. doi:10.1006/jmbi.1996.0111

Ban, N., Nissen, P., Hansen, J., Moore, P. B., & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, *289*(5481), 905–920.

Beaudry, A. A., & Joyce, G. F. (1992). Directed evolution of an RNA enzyme. *Science*, *257*(5070), 635–641. doi:10.1146/annurev.biochem.73.011303.073717

Breslow, R. (1957). RAPID DEUTERIUM EXCHANGE IN THIAZOLIUM SALTS1. *Journal of the American Chemical Society*, *79*(7), 1762–1763.

Breslow, R., & Appayee, C. (2013). Transketolase reaction under credible prebiotic conditions. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(11), 4184–4187. doi:10.1073/pnas.1301522110

Butlerow, A. (1861). Butlerow: Formation synthétique d'une substance sucrée - Google Scholar. *CR Acad Sci*.

Cech, T. R. (2000). Structural biology. The ribosome is a ribozyme. *Science*, *289*(5481), 878–879.

Cech, T. R., Zaug, A. J., & Grabowski, P. J. (1981). In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell*, *27*(3 Pt 2), 487–496.

Chandrasekar, J., & Silverman, S. K. (2013). Catalytic DNA with phosphatase activity. *Proceedings of the National Academy of Sciences*, *110*(14), 5315–5320. doi:10.1073/pnas.1221946110/-/DCSupplemental/pnas.201221946SI.pdf

Chen, G. C., & Jordan, F. (1984). Brewers' yeast pyruvate decarboxylase produces acetoin from acetaldehyde: a novel tool to study the mechanism of steps subsequent to carbon dioxide loss. *Biochemistry*, *23*(16), 3576–3582.

Cochrane, J. C., & Strobel, S. A. (2008). Riboswitch effectors as protein enzyme cofactors. *RNA*, *14*(6), 993–1002. doi:10.1261/rna.908408

Coleman, T. M., & Huang, F. (2002). RNA-catalyzed thioester synthesis. *Chemistry & Biology*, *9*(11), 1227–1236.

Crosby, J., & Lienhard, G. E. (1970). Mechanisms of thiamine-catalyzed reactions. A kinetic analysis of the decarboxylation of pyruvate by 3,4-dimethylthiazolium ion in water and ethanol. *Journal of the American Chemical Society*, *92*(19), 5707–5716.

Crosby, J., Stone, R., & Lienhard, G. E. (1970). Mechanisms of thiamine-catalyzed reactions. Decarboxylation of 2-(1-carboxy-l-hydroxyethyl)-3, 4-dimethylthiazolium chloride. *Journal of the American Chemical Society*, *92*(9), 2891–2900.

Decker, P., Schweer, H., & Pohlamnn, R. (1982). Bioids: X. Identification of formose sugars, presumable prebiotic metabolites, using capillary gas chromatography/gas chromatography—mas spectrometry of n- …. *Journal of Chromatography A*.

Doudna, J. A., & Cech, T. R. (2002). The chemical repertoire of natural ribozymes. *Nature*, *418*(6894), 222–228.

Dworkin, J. P., & Miller, S. L. (2000). A kinetic estimate of the free aldehyde content of aldoses. *Carbohydrate research*, *329*(2), 359–365.

Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B., & Jordan, F. (1993). Catalytic centers in the thiamin diphosphate dependent enzyme pyruvate decarboxylase at 2.4-. ANG. resolution. *Biochemistry*, *32*(24), 6165–6170.

Edwards, T. E., & Ferré-D'Amaré, A. R. (2006). Crystal Structures of the Thi-Box Riboswitch Bound to Thiamine Pyrophosphate Analogs Reveal Adaptive RNA-Small Molecule Recognition. *Structure*, *14*(9), 1459–1468. doi:10.1016/j.str.2006.07.008

Edwards, T. E., Klein, D. J., & Ferré-D'Amaré, A. R. (2007). Riboswitches: small-molecule recognition by gene regulatory RNAs. *Current Opinion in Structural Biology*, *17*(3), 273–279. doi:10.1016/j.sbi.2007.05.004

Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature*, *346*(6287), 818–822.

Frank, R. A. W., Leeper, F. J., & Luisi, B. F. (2007). Structure, mechanism and catalytic duality of thiamine-dependent enzymes. *Cellular and Molecular Life Sciences*, *64*(7-8), 892–905. doi:10.1007/s00018-007-6423-5

Fry, K., Ingraham, L. L., & Westheimer, F. H. (1957). The Thiamin-Pyruvate Reaction. *Journal of the American Chemical Society*, *79*(19), 5225–5227.

Fusz, S., Eisenführ, A., Srivatsan, S. G., Heckel, A., & Famulok, M. (2005). A Ribozyme for the Aldol Reaction. *Chemistry & Biology*, *12*(8), 941–950. doi:10.1016/j.chembiol.2005.06.008

Gabel, N. W., & Ponnamperuma, C. (1967). Model for origin of monosaccharides. *Nature*, *216*(5114), 453–455.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, *35*(3 Pt 2), 849–857.

Guo, F., Zhang, D., Kahyaoglu, A., Farid, R. S., & Jordan, F. (1998). Is a hydrophobic amino acid required to maintain the reactive V conformation of thiamin at the active center of thiamin diphosphate-requiring enzymes? Experimental and computational studies of isoleucine 415 of yeast pyruvate decarboxylase. *Biochemistry*, *37*(38), 13379–13391. doi:10.1021/bi9807097

Gutowski, J. A., & Lienhard, G. E. (1976). Transition state analogs for thiamin pyrophosphate-dependent enzymes. *Journal of Biological Chemistry*, *251*(9), 2863–2866.

Harris, T. K., & Turner, G. J. (2002). Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. *IUBMB life*, *53*(2), 85–98. doi:10.1080/10399710290038972

Hayden, E. J., & Lehman, N. (2006). Self-assembly of a group I intron from inactive oligonucleotide fragments. *Chemistry & Biology*, *13*(8), 909–918. doi:10.1016/j.chembiol.2006.06.014

Hayden, E. J., Riley, C. A., Burton, A. S., & Lehman, N. (2005). RNA-directed construction of structurally complex and active ligase ribozymes through recombination. *RNA*, *11*(11), 1678–1687. doi:10.1261/rna.2125305

Heinen, W., & Lauwers, A. M. (1996). Organic sulfur compounds resulting from the interaction of iron sulfide, hydrogen sulfide and carbon dioxide in an anaerobic aqueous environment. *Origins of life and evolution of the biosphere : the journal of the International Society for the Study of the Origin of Life*, *26*(2), 131–150.

Huber, C. (1997). Activated Acetic Acid by Carbon Fixation on (Fe,Ni)S Under Primordial Conditions. *Science*, *276*(5310), 245–247. doi:10.1126/science.276.5310.245

Jadhav, V. R., & Yarus, M. (2002a). Coenzymes as coribozymes. *Biochimie*, *84*(9), 877–888.

Jadhav, V. R., & Yarus, M. (2002b). Acyl-CoAs from Coenzyme Ribozymes †. *Biochemistry*, *41*(3), 723–729. doi:10.1021/bi011803h

Johnston, W. K. (2001). RNA-Catalyzed RNA Polymerization: Accurate and General RNA-Templated Primer Extension. *Science*, *292*(5520), 1319–1325. doi:10.1126/science.1060786

Jordan, F., & Mariam, Y. H. (1978). N1'-Methylthiaminium diiodide. Model study on the effect of a coenzyme bound positive charge on reaction mechanisms requiring thiamin pyrophosphate. *Journal of the American Chemical Society*, *100*(8), 2534–2541.

Jordan, F., Li, H., & Brown, A. (1999). Remarkable Stabilization of Zwitterionic Intermediates May Account for a Billion-fold Rate Acceleration by Thiamin Diphosphate-Dependent Decarboxylases †. *Biochemistry*, *38*(20), 6369–6373. doi:10.1021/bi990373g

Jordan, F., Nemeria, N. S., Zhang, S., Yan, Arjunan, P., & Furey, W. (2003). Dual Catalytic Apparatus of the Thiamin Diphosphate Coenzyme: Acid−Base via the 1‘,4‘-Iminopyrimidine Tautomer along with Its Electrophilic Role. *Journal of the American Chemical Society*, *125*(42), 12732–12738. doi:10.1021/ja0346126

Jordan, F., Zhang, Z., & Sergienko, E. (2002). Spectroscopic Evidence for Participation of the 1′,4′-Imino Tautomer of Thiamin Diphosphate in Catalysis by Yeast Pyruvate Decarboxylase. *Bioorganic Chemistry*, *30*(3), 188–198. doi:10.1006/bioo.2002.1249

Kemp, D. S., & O'Brien, J. T. (1970). Base catalysis of thiazolium salt hydrogen exchange and its implications for enzymatic thiamine cofactor catalysis. *Journal of the American Chemical Society*, *92*(8), 2554–2555.

Kern, D. (1997). How Thiamin Diphosphate Is Activated in Enzymes. *Science*, *275*(5296), 67–70. doi:10.1126/science.275.5296.67

Kluger, R., Chin, J., & Smyth, T. (1981). Thiamin-catalyzed decarboxylation of pyruvate. Synthesis and reactivity analysis of the central, elusive intermediate,. alpha.-lactylthiamin. *Journal of the American Chemical Society*, *103*(4), 884–888.

Kluger, R., Gish, G., & Kauffman, G. (1984). Interaction of thiamin diphosphate and thiamin thiazolone diphosphate with wheat germ pyruvate decarboxylase. *Journal of Biological Chemistry*, *259*(14), 8960–8965.

Kluger, R., Lam, J. F., & Kim, C. S. (1993). Decomposition of 2-(1-hydroxybenzyl) thiamin in neutral aqueous solutions: benzaldehyde and thiamin are not the products. *Bioorganic Chemistry*, *21*(3), 275–283.

Kraut, J., & Reed, H. J. (1962). S0365110X62001978. *Acta Cryst (1962). Q15, 747-757 [doi:10.1107/S0365110X62001978]*, 1–11. doi:10.1107/S0365110X62001978

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, *31*(1), 147–157.

Kuo, D. J., & Jordan, F. (1983a). Active site directed irreversible inactivation of brewers' yeast pyruvate decarboxylase by the conjugated substrate analogue (E)-4-(4-chlorophenyl)-2-oxo-3-butenoic acid: development of a suicide substrate. *Biochemistry*, *22*(16), 3735–3740.

Kuo, D. J., & Jordan, F. (1983b). Direct spectroscopic observation of a brewer's yeast pyruvate decarboxylase-bound enamine intermediate produced from a suicide substrate. Evidence for nonconcerted decarboxylation. *Journal of Biological Chemistry*, *258*(22), 13415–13417. doi:10.1002/jcp.24344

Larralde, R., Robertson, M. P., & Miller, S. L. (1995). Rates of decomposition of ribose and other sugars: implications for chemical evolution. *Proceedings of the National Academy of Sciences*, *92*(18), 8158–8160.

Lehman, N., & Joyce, G. F. (1993). Evolution in vitro of an RNA enzyme with altered metal dependence.

Li, Y., & Sen, D. (1996). A catalytic DNA for porphyrin metallation. *Nature Structural Biology*, *3*(9), 743–747.

Lindqvist, Y., Schneider, G., Ermler, U., & Sundström, M. (1992). Three-dimensional structure of transketolase, a thiamine diphosphate dependent enzyme, at 2.5 A resolution. *The EMBO journal*, *11*(7), 2373.

Lorsch, J. R., & Szostak, J. W. (1994). In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature*, *371*(6492), 31–36. doi:10.1038/371031a0

Miller, S. L. (1953). A production of amino acids under possible primitive earth conditions. *Science*, *117*(3046), 528–529.

Miller, S. L. (1955). Production of Some Organic Compounds under Possible Primitive Earth Conditions1. *Journal of the American Chemical Society*, *77*(9), 2351–2361.

Muller, Y. A., & Schulz, G. E. (1993). Structure of the thiamine-and flavin-dependent enzyme pyruvate oxidase. *Science*, *259*(5097), 965–967.

Muller, Y. A., Lindqvist, Y., Furey, W., Schulz, G. E., Jordan, F., & Schneider, G. (1993). A thiamin diphosphate binding fold revealed by comparison of the crystal structures of transketolase, pyruvate oxidase and pyruvate decarboxylase. *Structure*, *1*(2), 95–103.

Nemeria, N., Baykal, A., Joseph, E., Zhang, S., Yan, Furey, W., & Jordan, F. (2004). Tetrahedral Intermediates in Thiamin Diphosphate-Dependent Decarboxylations Exist as a 1‘,4‘-Imino Tautomeric Form of the Coenzyme, Unlike the Michaelis Complex or the Free Coenzyme †. *Biochemistry*, *43*(21), 6565–6575. doi:10.1021/bi049549r

Nemeria, N., Chakraborty, S., Baykal, A., Korotchkina, L. G., Patel, M. S., & Jordan, F. (2007a). The 1′, 4′-iminopyrimidine tautomer of thiamin diphosphate is poised for catalysis in asymmetric active centers on enzymes. *Proceedings of the National Academy of Sciences*, *104*(1), 78–82.

Nemeria, N., Korotchkina, L., McLeish, M. J., Kenyon, G. L., Patel, M. S., & Jordan, F. (2007b). Elucidation of the Chemistry of Enzyme-Bound Thiamin Diphosphate Prior to Substrate Binding: Defining Internal Equilibria among Tautomeric and Ionization States †. *Biochemistry*, *46*(37), 10739–10744. doi:10.1021/bi700838q

Nemeria, N., Yan, Y., Zhang, Z., Brown, A. M., Arjunan, P., Furey, W., et al. (2001). Inhibition of the Escherichia coli pyruvate dehydrogenase complex E1 subunit and its tyrosine 177 variants by thiamin 2-thiazolone and thiamin 2-thiothiazolone diphosphates. Evidence for reversible tight-binding inhibition. *Journal of Biological Chemistry*, *276*(49), 45969–45978. doi:10.1074/jbc.M104116200

ORO, J., & Kimball, A. P. (1961). Synthesis of purines under possible primitive earth conditions. I. Adenine from hydrogen cyanide. *Archives of biochemistry and biophysics*, *94*, 217–227.

ORO, J., & Kimball, A. P. (1962). Synthesis of purines under possible primitive earth conditions. II. Purine intermediates from hydrogen cyanide. *Archives of biochemistry and biophysics*, *96*, 293–313.

Powner, M. W., Gerland, B., & Sutherland, J. D. (2009). Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature*, *459*(7244), 239–242. doi:10.1038/nature08013

Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R., & Bruening, G. (1986). Autolytic processing of dimeric plant virus satellite RNA. *Science*, *231*(4745), 1577–1580. doi:10.1126/science.231.4745.1577

Richard, J. P., & Amyes, T. L. (2004). On the importance of being zwitterionic: enzymatic catalysis of decarboxylation and deprotonation of cationic carbon. *Bioorganic Chemistry*, *32*(5), 354–366. doi:10.1016/j.bioorg.2004.05.002

Riley, C. A., & Lehman, N. (2003). Generalized RNA-directed recombination of RNA. *Chemistry & Biology*, *10*(12), 1233–1243. doi:10.1016/j.chembiol.2003.11.015

Robertson, M. P., & Miller, S. L. (1995). An efficient prebiotic synthesis of cytosine and uracil.

Sanchez, R. A., Ferris, J. P., & Orgel, L. E. (1966a). Cyanoacetylene in prebiotic synthesis. *Science*, *154*(3750), 784–785.

Sanchez, R. A., Ferris, J. P., & Orgel, L. E. (1967). Studies in prebiotic synthesis. II. Synthesis of purine precursors and amino acids from aqueous hydrogen cyanide. *Journal of Molecular Biology*, *30*(2), 223–253.

Sanchez, R. A., Ferris, J. P., & Orgel, L. E. (1968). Studies in prebiotic synthesis. IV. Conversion of 4-aminoimidazole-5-carbonitrile derivatives to purines. *Journal of Molecular Biology*, *38*(1), 121–128.

Sanchez, R., Ferris, J., & Orgel, L. E. (1966b). Conditions for purine synthesis: did prebiotic synthesis occur at low temperatures? *Science*, *153*(3731), 72–73.

Saville, B. J., & Collins, R. A. (1990). A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria. *Cell*, *61*(4), 685–696.

Schwartz, A. W., & Goverde, M. (1982). Acceleration of HCN oligomerization by formaldehyde and related compounds: implications for prebiotic syntheses. *Journal of Molecular Evolution*, *18*(5), 351–353.

Seelig, B., & Jäschke, A. (1999). A small catalytic RNA motif with Diels-Alderase activity. *Chemistry & Biology*, *6*(3), 167–176. doi:10.1016/S1074-5521(99)89008-5

Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R., & Patel, D. J. (2006). Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature*, *441*(7097), 1167–1171. doi:10.1038/nature04740

Sharmeen, L., Kuo, M. Y., Dinter-Gottlieb, G., & Taylor, J. (1988). Antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. *Journal of virology*, *62*(8), 2674–2679.

Shin, W., & Kim, Y. C. (1986). Crystal structure of thiamine thiazolone: a possible transition-state analog with an intramolecular NH. cntdot.. cntdot.. cntdot. O hydrogen bond in the V form. *Journal of the American Chemical Society*, *108*(22), 7078–7082.

Springsteen, G., & Joyce, G. F. (2004). Selective Derivatization and Sequestration of Ribose from a Prebiotic Mix. *Journal of the American Chemical Society*, *126*(31), 9578–9583. doi:10.1021/ja0483692

Thore, S. (2006). Structure of the Eukaryotic Thiamine Pyrophosphate Riboswitch with Its Regulatory Ligand. *Science*, *312*(5777), 1208–1211. doi:10.1126/science.1128451

Tsukiji, S., Pattnaik, S. B., & Suga, H. (2003). An alcohol dehydrogenase ribozyme. *Nature Structural Biology*, *10*(9), 713–717. doi:10.1038/nsb964

Tucker, B. J., & Breaker, R. R. (2005). Riboswitches as versatile gene control elements. *Current Opinion in Structural Biology*, *15*(3), 342–348. doi:10.1016/j.sbi.2005.05.003

Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*.

Unrau, P. J., & Bartel, D. P. (1998). RNA-catalysed nucleotide synthesis. *Nature*, *395*(6699), 260–263.

Vitreschak, A. (2004). Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends in Genetics*, *20*(1), 44–50. doi:10.1016/j.tig.2003.11.008

Voet, A. B., & Schwartz, A. W. (1983). Prebiotic adenine synthesis from HCN—Evidence for a newly discovered major pathway. *Bioorganic Chemistry*, *12*(1), 8–17. doi:10.1016/0045-2068(83)90003-2

Washabaugh, M. W., & Jencks, W. P. (1988). Thiazolium C(2)-proton exchange: structure-reactivity correlations and the pKa of thiamin C(2)-H revisited. *Biochemistry*, *27*(14), 5044–5053.

Washabaugh, M. W., & Jencks, W. P. (1989a). Thiazolium C (2)-proton exchange: general-base catalysis, direct proton transfer, and acid inhibition. *Journal of the American Chemical Society*, *111*(2), 674–683.

Washabaugh, M. W., & Jencks, W. P. (1989b). Thiazolium C (2)-Proton exchange: Isotope effects, internal return, and a small intrinsic barrier. *Journal of the American Chemical Society*, *111*(2), 683–692.

Washabaugh, M. W., Stivers, J. T., & Hickey, K. A. (1994). C (. alpha.)-Proton Transfer from 2-(1-Hydroxybenzyl) oxythiamin: The Unit Bronsted Slope Overestimates the Amount of Bond Formation to the Base Catalyst in the Transition State. *Journal of the American Chemical Society*, *116*(16), 7094–7097.

Wächtershäuser, G. (1988). Before enzymes and templates: theory of surface metabolism. *Microbiological reviews*, *52*(4), 452–484.

Wächtershäuser, G. (1990). Evolution of the first metabolic cycles. *Proceedings of the National Academy of Sciences*, *87*(1), 200–204.

Weber, A. L. (1981). Formation of pyrophosphate, tripolyphosphate, and phosphorylimidazole with the thioester, n, s-diacetyl-cysteamine, as the condensing agent. *Journal of Molecular Evolution*, *18*(1), 24–29.

Weber, A. L. (1982). Formation of pyrophosphate on hydroxyapatite with thioesters as condensing agents. *Bio Systems*, *15*(3), 183–189.

White, H. B., III. (1976). Coenzymes as fossils of an earlier metabolic state. *Journal of Molecular Evolution*, *7*(2), 101–104.

Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., & Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, *428*(6980), 281–286. doi:10.1038/nature02362

Winkler, W., Nahvi, A., & Breaker, R. R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*, *419*(6910), 952–956. doi:10.1038/nature01145

Wochner, A., Attwater, J., Coulson, A., & Holliger, P. (2011). Ribozyme-Catalyzed Transcription of an Active Ribozyme. *Science*, *332*(6026), 209–212. doi:10.1126/science.1200752

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences*, *74*(11), 5088–5090.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, *87*(12), 4576–4579.

Zaug, A. J., & Cech, T. R. (1986). The intervening sequence RNA of Tetrahymena is an enzyme. *Science*, *231*(4737), 470–475.

Zhang, S., Zhou, L., Nemeria, N., Yan, Zhang, Z., Zou, Y., & Jordan, F. (2005a). Evidence for Dramatic Acceleration of a C−H Bond Ionization Rate in Thiamin Diphosphate Enzymes by the Protein Environment †. *Biochemistry*, *44*(7), 2237–2243. doi:10.1021/bi047696j

Appendices

Appendix A.

Experimental details

(1) Synthesized by University of Calgary Core DNA services. The four Phophoramidites were "hand mixed" in equal proportions.

(2) Total volume of PCR was ~15 mL. The 15 mL volume was split into ~30, 500 μL reactions and cycled starting with a 94 °C denaturing step for 1.5 minutes, followed by an annealing step at 50 °C for 1.5 minutes and ending with an elongation step at 72 °C for 5 minutes. The elongation step was later reduced to 1.5 minutes without any noticeable change in yield or product specificity.

(3) ~15 mL of transcription were used for the first round of selection and the volume of the transcriptions were reduced by half each subsequent round until a volume of 125 μL was reached.

(4) TCEP was later found by NMR to react with Bromo-*Lnk*PB irreversibly. The attachment protocol was changed to reflect this finding. The phosphorothioate RNA was first reduced using 2 mM DTT at 90 °C for 2 minutes. The RNA was IPA precipitated away from the DTT because of DTT's reactivity towards the Bromo-*Lnk*PB substrate. The RNA was then reacted like before with exclusion of TCEP. The yield of the coupling reaction was never determined and it was never ruled out that the poor reaction yield of dc4 with ThBi (15-30%) was not because of low *Lnk*PB coupling. More consistent results were often found when the Bromo-*Lnk*PB coupling reaction was performed 2x, IPA precipitating and denaturing before the second reaction.

(5) IPA precipitations were found to be quantitative, selective and very fast. Typically after addition of IPA, the sample could be centrifuged immediately. The centrifugation time was reduced to ~6-8 minutes at full speed without any compromise in yield.

(6) Dynabeads MyOne Streptavidin T1 were used. Beads were first washed 3-4 times with 400 μL NaOH washing buffer containing 100 mM NaOH and 50 mM NaCl, followed by 3-4 400 μL 4M urea washes and finally 4 washes with 400 μL bead binding buffer. 300 μL of bead was used for round 1, 200 μL of beads for round 2, 100 μL of beads for round 3, 50 μL of bead for round 4 and 25 μL of bead for round 5.

(7) Mutant reverse transcriptase, Superscript II (Invitrogen) was used. Reactions were performed according to the manufacturers instructions. The washed beads containing the reacted RNA was reconstituted in ~10 μL of 10 mM HEPES buffer (pH 7.5) so the salt from the bead washing buffer/urea would not inhibit the reaction. The beads were transferred to a 100 μL reverse transcription reaction and annealing to the reverse primer was performed following the manufactures protocol. The reactions were incubated at 42° C for 45 minutes, followed by a 45° C for 45 minutes and ending with a 55° C incubation for an additional 45 minutes.

(8) Transcriptions were typically 50-100 μL (dimmer RNA was inconsistently formed transcriptions were of smaller volume) containing 30 μL uridine 5'-triphosphate [α-^{32}P] (250 μci, 300 ci/mmol) and only 10 mM UTP.

(9) Two IPA precipitations were found to be sufficient to purify the reacted RNA from excess ThBi, however washing was found to be important. The first pellets were washed one time using ~100 μL of 100% ethanol at room temperature. The second pellets were washed two times with ~100 μL of 100% ethanol (making sure to wash the lid by

inversion of the test tubes) followed by a quick table-top centrifugation to collect the excess ethanol. The remaining 5-20 μL was carefully collected away from the pellet and the samples were left open to dry for ~15 minutes. GlycoBlue (Ambion) was found to be critical for this application because of the small size of the pellets and the number of samples.