Studies Towards a Chemical Origins of RNA:

Divergent Prebiotic Synthesis of Pyrimidine and 8-oxo-Purines Ribonucleotides

A thesis submitted to University College London (UCL)

for the degree of

DOCTOR of PHILOSOPHY

in the Faculty of Mathematical and Physical Sciences

2018

Arif Nikmal

Chemistry Department

20 Gordon St

London, WC1H 0AJ

Declaration & Copyright

I, Arif Nikmal confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:_____

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the "Copyright") and he has given University College London (UCL) the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

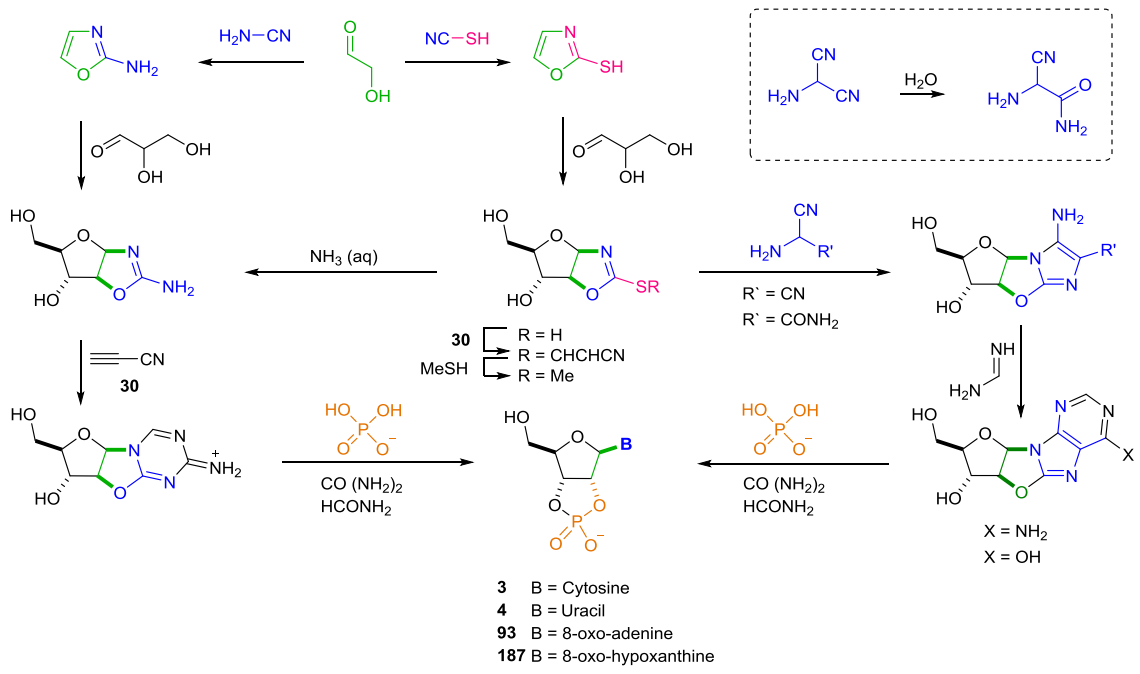
Further information on the conditions under which disclosure, publication and exploitation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it may take place is available from the Head of Chemistry Department, UCL.

Those parts of this thesis having previously been published at the time of writing:

Stairs, S.; Nikmal, A.; Bučar, D.-K.; Zheng, S.-L.; Szostak, J. W.; Powner, M. W., Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides, *Nat. Comm.*, **2017**, 8, 15270. DOI:10.1038/ncomms15270

Abstract

Elucidating a prebiotic plausible nucleotide synthesis is considered one of the greatest challenges in the origins of life research. Although separate synthesis that account for the pyrimidine and purine ribonucleotides have been developed, no divergent synthesis from a single precursor has yet been developed that accounts for the synthesis of both pyrimidine and purine ribonucleotides. Furthermore, the reported prebiotic synthesis of pyrimidine and purine nucleotides are not mutually compatible and could not co-operate under similar conditions. This thesis proposes a divergent synthesis of both anhydro-pyrimidine and anhydro-purine ribonucleotides on a regiospecific sugar scaffold, followed by concurrent phosphorylation and chemical inversion to yield pyrimidines (**3** and **4**) and 8-oxo-purines (**93** and **187**) in the natural *β*-ribo-stereochemistry.



This concomitant approach employs only feedstock small molecules that are consider to be prebiotically plausible and condition that are consider to be reasonable under early Earth geothermal conditions. Moreover, given that 8-oxo purines and pyrimidines ribonucleotides share a common prebiotic synthetic route, it is hypothesised that 8-oxo purines could have predated natural purines in biology and that they could have played an important role in the transition of chemistry to biology and facilitated the modern Watson-Crick base pairing system.

Contents

Acknowledgments

I am grateful to my supervisor Dr Matthew W Powner for giving me the opportunity to study for a PhD degree at UCL. The long discussions and meetings with Matt has been instrumental to get the best of my PhD and his unconditional help and assistance was always there. It was a pleasure working with everyone at Powner lab, there was always a lot fun, support and I will dearly miss everyone. In particular I would like to thank Lello for extensively proof reading my thesis and giving valuable suggestions. I would like to thank Shaun for proof reading my experimental section and for his enormous support throughout my time at UCL. I would not have made it in the hot summer days at Powner lab without Shaun's eccentric dry ice air conditioning system.

Studying for a PhD was not easy and without my family and friends I would not have been able to make it to the end. I cannot thank enough my parents for tirelessly pushing us to study and always placing our education before everything. I would like to thank my uncles Ajmal and Abid for their exceptional support and encouragement for studying a PhD degree.

And finally I would like to thank my wife Liza Nikmal for her patience, company and reassurance during hard times. Thank you for believing in me. She is the star of my life and without her I would not have made this difficult journey.

Abbreviations

A	adenine
Ac	acetyl
AICA	5-amino-imidazole-4-carboxamide
AICN	5-amino-imidazole-4-carbonitrile
AMP	adenosine monophosphate
AmTP	amidotriphosphate
Ar	aryl
ATP	adenosine triphosphate
aq.	aqueous
APCI	atmospheric pressure chemical ionisation
B	nucleic acid base
t-Bu	*tert*-butyl
Bn	benzyl
Bz	benzoyl
°C	degrees Celsius
C	cytosine
c	concentration
ca.	*circa* (Latin: about)
calcd	calculated
cAMP	adenosine-3′,5′-cyclic phosphate
cat.	catalytic
CI	chemical ionisation
cm^{-1}	wavenumber
conc.	concentrated
COSY	correlated spectroscopy (NMR)
δ	chemical shift
Δ	heat

DAMN	diaminomaleonitrile
DCM	dichloro methane
DHA	dihydroxyacetone
DIAD	*N*,*N*′-diisopropylazodicarbodiimide
DMA	*N*,*N*-dimethylacetamide
DMAP	4-(dimethylamino)-pyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
d-NDP	deoxynucleotide diphosphate
DOSY	diffusion ordered spectroscopy (NMR)
Dowex®	Dowex® ion exchange resin
Eds.	editors
ES	electrospray
Et	ethyl
et al.	*et alia* (Latin: and others)
ether	diethyl ether
eq.	equivalent(s)
FT	Fourier transform
ΔG	Gibbs free energy
G	guanine
GC	gas chromatography
GMP	guanosine monophosphate
GTP	guanosine triphosphate
h	hour(s)
hν	electromagnetic irradiation (UV)
HMBC	heteronuclear multiple-bond correlation multiple-quantum correlation
HPLC	high pressure liquid chromatography

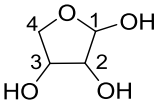
HR	high resolution
Hz	Hertz
i	*iso*
inc.	including
IR	infrared
J	NMR coupling constant measured in Herts
LA	Lewis acid
LHF	liquid hydrogen fluoride
lit.	literature (reference)
m	milli
M	molar
Me	methyl
MHz	megahertz
min(s)	minute(s)
mL	millilitre(s)
mmol	millimole(s)
m.p.	melting point
MS	mass spectrometry
µl	microlitre
µM	micromolar
m/z	mass / charge ratio
NADH	nicotinamide adenine dinucleotide
NDP	nucleotide diphosphate
NMP	nucleotide monophosphate
NTP	nucleotide triphosphate
NMR	nuclear magnetic resonance
op.	open chain
PG	protecting group

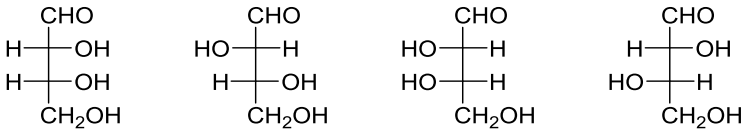
Ph	phenyl
Pi	inorganic phosphate
PNA	peptide nucleic acid
PPi	inorganic pyrophosphate
ppm	parts per million
Pr	propyl
p-RNA	pyranosyl ribonucleic acid
PRPP	5-phosphoribosyl-1-pyrophosphate
pyr.	pyridine
quant.	quantitative yield
R	unspecified group
rac-	racemic mixture
RNA	ribonucleic acid
rt	room temperature
sca-	scalemic mixture
sat.	saturated
soln.	solution
S&O	Sanchez and Orgel
t	tertiary
tert	tertiary
T	thymine
$t_{1/2}$	half life
TBDMS	*tert*-butyldimethylsilyl
TFA	trifluOró acetic acid
THF	tetrahydrofuran
THFo	tetrahydrofolate
TIPS	tetraisopropylidene disiloxane
TLC	thin layer chromatography

TNA	*L*-α-threofuranosyl (3'-2') nucleic acid
U	uracil
UV	ultraviolet

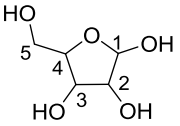
Numbering and Nomenclature

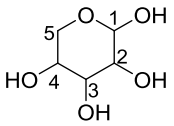
Tetrose sugars

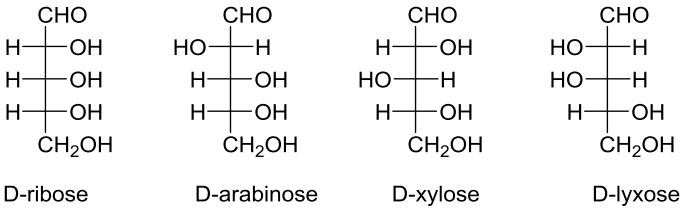




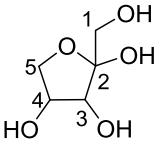
Pentose sugars (Fisher projections in natural D-series)

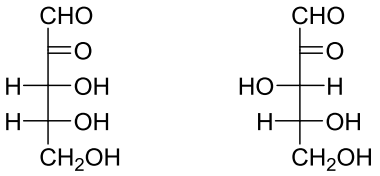




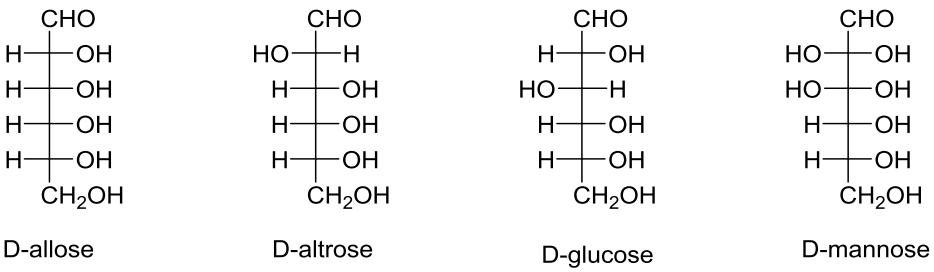


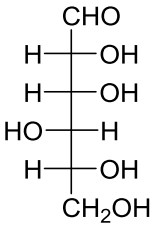
Pentulose sugars (Fisher projections in natural D-series)



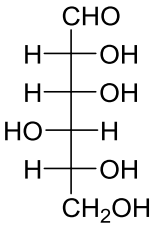


Hexose sugars (Fisher projections in natural D-series)

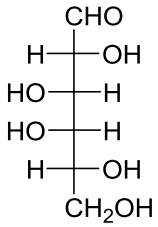




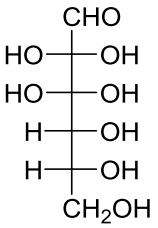
D-gulose



D-idose

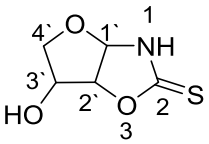


D-galactose

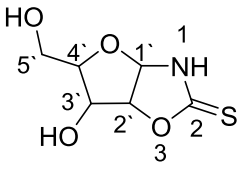


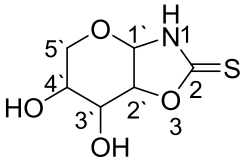
D-talose

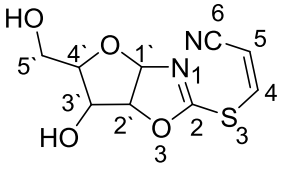
Tetrose oxazolidinone thione

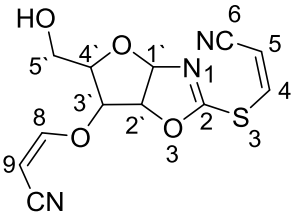


Pentose oxazolidinone thione

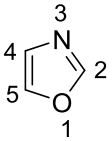




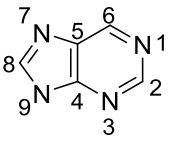


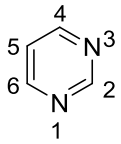


Oxazoles and thiooxazoles

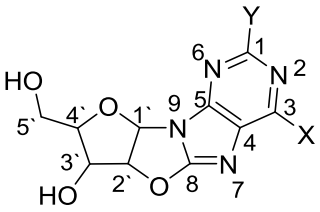


Pyrimidines and Purines



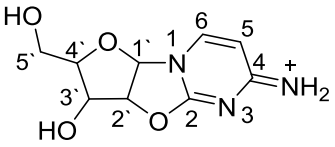


Anhydronucleosides

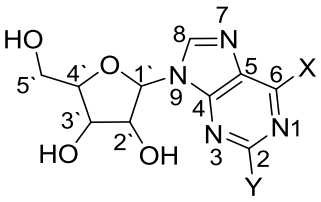


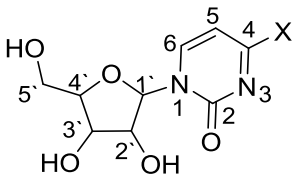
Y = H/NH_2

X = NH_2/OH

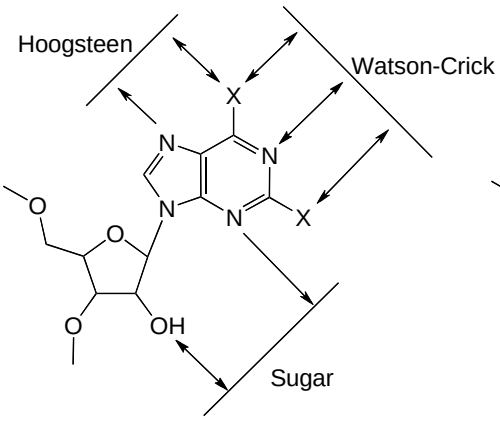


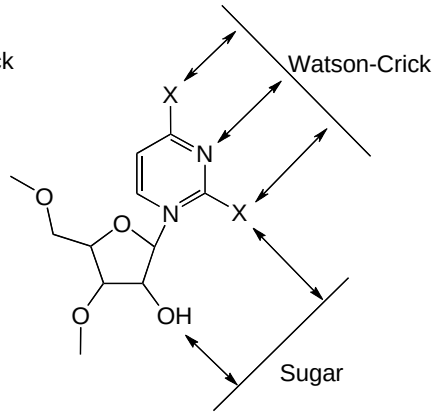
Nucleosides





Base-pairing interactions





"Read: In the name of thy Lord Who createth, Createth man from a clot. Read: And thy Lord is the Most Bounteous, Who teacheth by the pen, Teacheth man that which he knew not."

Quran 96:1-5

"Don't only practice your art, but force your way into its secrets, for it and knowledge can raise men to the divine."

Ludwig van Beethoven

"We not only want to know how nature is (and how her transactions are carried through), but we also want to reach, if possible, a goal which may seem utopian and presumptuous, namely, to know why nature is such and not otherwise."

Albert Einstein

1. Introduction

1.1. What is life?

There are an approximately 8.7 million different species thriving on Earth.[1] All living species are cellular, and all characterised species have been classified by biologist into the Tree of Life (Figure 1.1).[2]

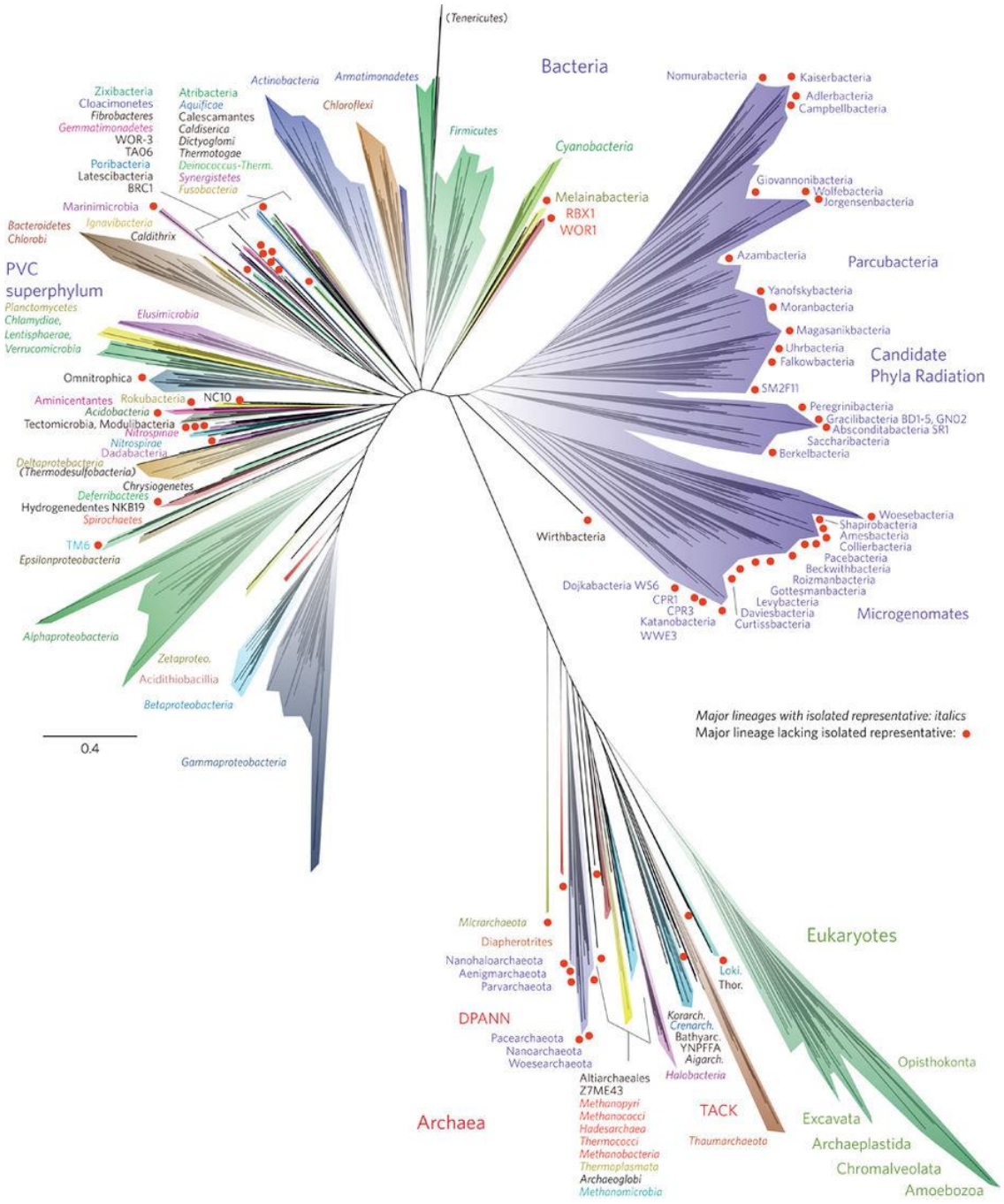


Figure 1.1: *A new metagenomic view of the tree of life by Hug et al.*[2]

Life can be found in almost all environments on Earth. Organisms have been found at the bottom of the ocean surviving chemotropically at deep-sea vents and in frozen lakes isolated for millions of years under the Antarctic ice. Even within rocks as deep as 11000 meters living species can be found.[3] Though life has now come to dominate our planet it is self-evident that life must not have been present initially. The Earth and indeed the universe are known to be only 4.5 and 14 billion years old respectively. Life cannot predate the universe. However, fossil evidence for prokaryotes have been discovered in rocks dating to 3.5 billion years.[4] In 2016, researchers in Australia discovered fossilised bacteria dating back 3.7 billion years (Figure 2).[5] Further isotopic fractionation evidence suggest that life was present on Earth as long as 3.8 billion years ago.[6] However, the small sample size of geological data for this period of Earth history,[7] make it difficult to say with any degree of certainty the nature of the first few hundred thousand years of Earth history. By June 2017, NASA's Kepler satellite had discovered 219 planets bringing the total number of exoplanets to 4034, 10 of which are presumed to be rocky and Earth-like. However, until now, life outside the boundaries of Earth has not been discovered.[8] This makes the Earth the only sure candidate planet on which life originated. Therefore it is fundamental to understand the universe, and specifically our planet and our own history, to elucidate the mechanism by which life came to exist on our Earth, exploiting what is known about our own planet (and life on it) for the investigation of the origins of life.

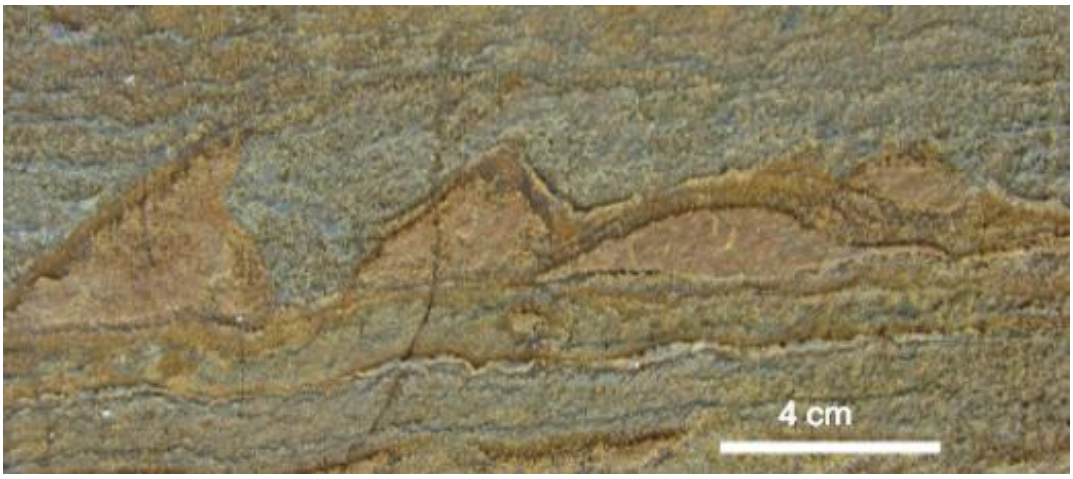


Figure 1.2: *Plausible 3.7 billion year old fossilised bacteria (zigzag patterns) embedded within rocks in Australia.*[5]

Many scientist and philosophers have tried to define life[9] and there are no less than twenty definitions of "life" found in the Webster dictionary.[10] In 1994 NASA arranged an international panel of experts to attempt to define a consensus definition of life, and the committee proposed that life is:

"*self-sustaining chemical system capable of Darwinian evolution*" [11]

However, whether defining a catch all definition of life is valuable to understanding the origins of life is debated. Luisi, in a recent review developed the reductive suggestion that in the field of prebiotic chemistry researchers need to ignore the most complex properties of life (such as intelligence, consciousness and ethics) and concentrate on building *minimal life*.[12] These simplest form of life might give us a more realistic approach to provide experimental evidence toward the chemical origins of life.[13] But how simple could life be? Others, including Nobel Laureate Jack Szostak, have argued that any attempts to define life will not help to understanding the origins of life, rather what is important in field of origin of life research is developing a physical model to understand the transitions that led from chemistry to biology.[14] This thinking leaves behind the philosophical aspect of defining life and demands an entirely experimental approach to investigating how life originated from simple organic molecules.

Given the complexity and diversity of life, and the inaccessibility of irrefutable geological evidence for the earliest periods of Earth history there are some, including Albert Eschenmoser, who have suggested that the origins of life cannot be discovered, and that it must be reinvented. To quote Eschenmoser: "*the aim of an experimental etiological chemistry is not primarily to delineate the pathway along which our (natural) life on Earth could have generated, but to provide decisive experimental evidence, through the realisation of model systems ('artificial chemical life') that life can arise as a result of the organisation of the organic matter*".[13]

The work outlined in this thesis will specifically explore this principle outline by Eschenmoser aiming to experimentally delineate a pathway to the canonical structures that unpin biology through the reaction of prebiotically plausible chemicals.

1.2. Origins of life

The origins of life is one of the greatest unsolved mysteries of all time.[13, 15] There is evidence to suggest that (minimal) life is sustained by replication and so exploring how the replicable element of life could have chemically arisen must be key to understanding the origins of life. Nucleic acids RNA (ribonucleic acid) and DNA (deoxyribonucleic acid) are amongst the most fundamental, essential and universally conserved elements of living organisms.

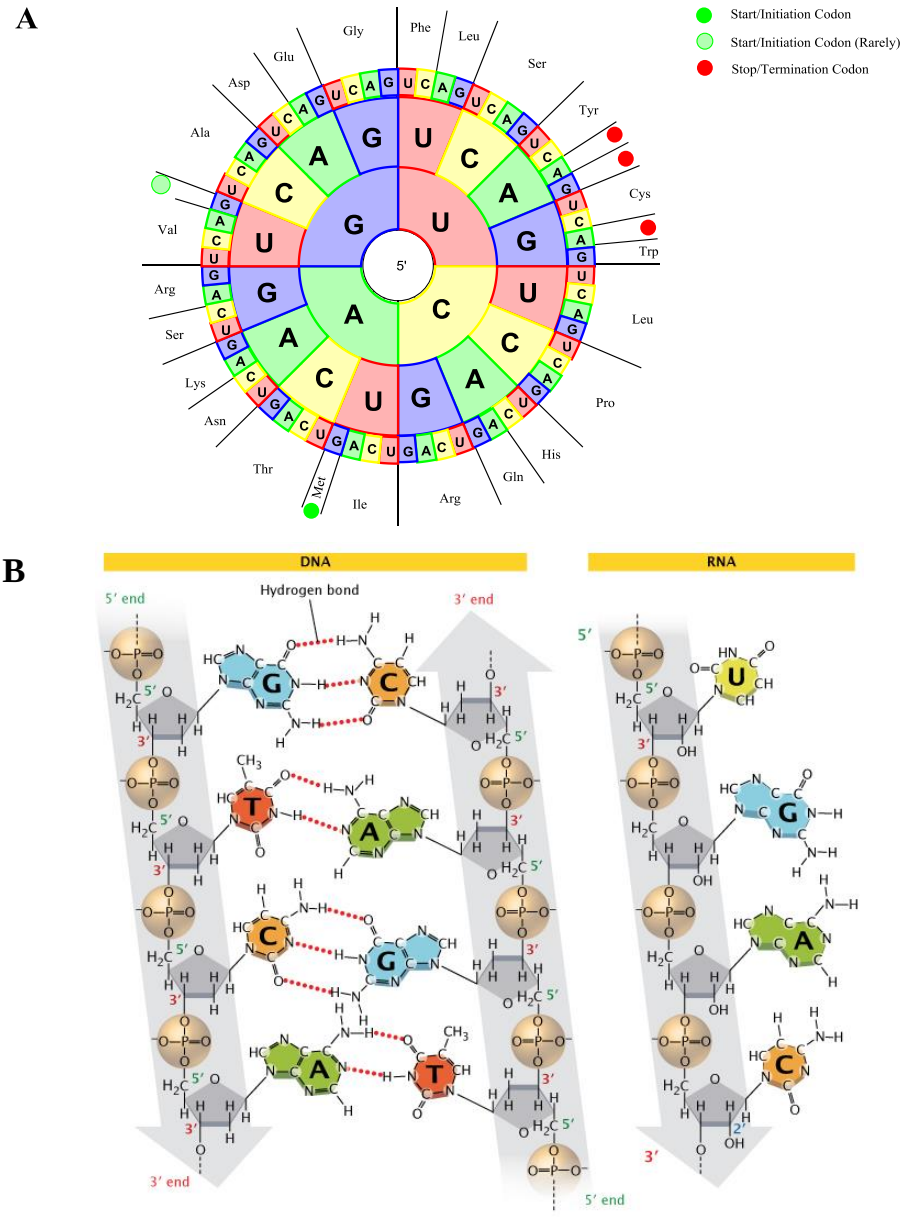


Figure 1.3: A) *Universal genetic code. The amino acids specified by each mRNA codon. Multiple codons can code for the same amino acid.* ***B)*** *Schematic representation for Watson-Crick base pairing of purine (guanine, adenine) with pyrimidine (cytosine, thymine) nucleobases in RNA. (Credit: Nature Education)*[16]

There are 5 canonical nucleotides that are used by all living cells to store and transfer information by Watson-Crick base pair (Figure 1.3B), which is the simple, predictable coding system that underpins the universal genetic code (Figure 1.3A). Before considering the prebiotic origins of information transfer, it is worth considering information transfer in extant biology. Proteins are made of amino acids, which can fold into highly specific secondary and tertiary structures. The basic building blocks of all proteins are the 20 proteinogenic amino acids (Figure 1.3A), which are joined in different sequences to a make variety of polypeptides, which are folded into proteins, and are used to perform different (programmable) tasks in living systems and accordingly (usually) express the phenotype of a given genetic element. To appreciate the complex and intricate process of protein assembly (in living cells), a brief overview of transcription and translation in living cells will be given in the next section.

1.2.1. Transcription and translation of the genetic code

The library of living information required by modern living cells is stored in the DNA located in the nucleus of cell. When a protein is required to be synthesised, the genetic information for that protein has to be decoded from the DNA, a process known as transcription. During transcription RNA polymerase (green) and its associated transcription factors binds to the DNA helix and unwind the promoter region with the correct code ready to be transcribed by RNA polymer (Figure 1.4A). RNA polymerase (green) aligns against the single DNA strand and begins moving down the DNA template strand in the 5´→3´ direction. As the polymerase moves along it polymerises complementary nucleotides (pink) and in this way the new strand of mRNA is prepared which will code for a specific protein (Figure 1.4B). However before the mRNA strand can be translated into proteins it has to be released from the DNA template. The final step of transcription is achieved when RNA polymerase reaches the termination sequence on the mRNA strand, causing RNA polymerase and the mRNA strand to fall off from the DNA template (Figure 1.4C). Now that information required for protein synthesis is extracted from DNA into the RNA transcript, the mRNA has to make a complex journey from the nucleus to the cytoplasm where the process of translation can begin.

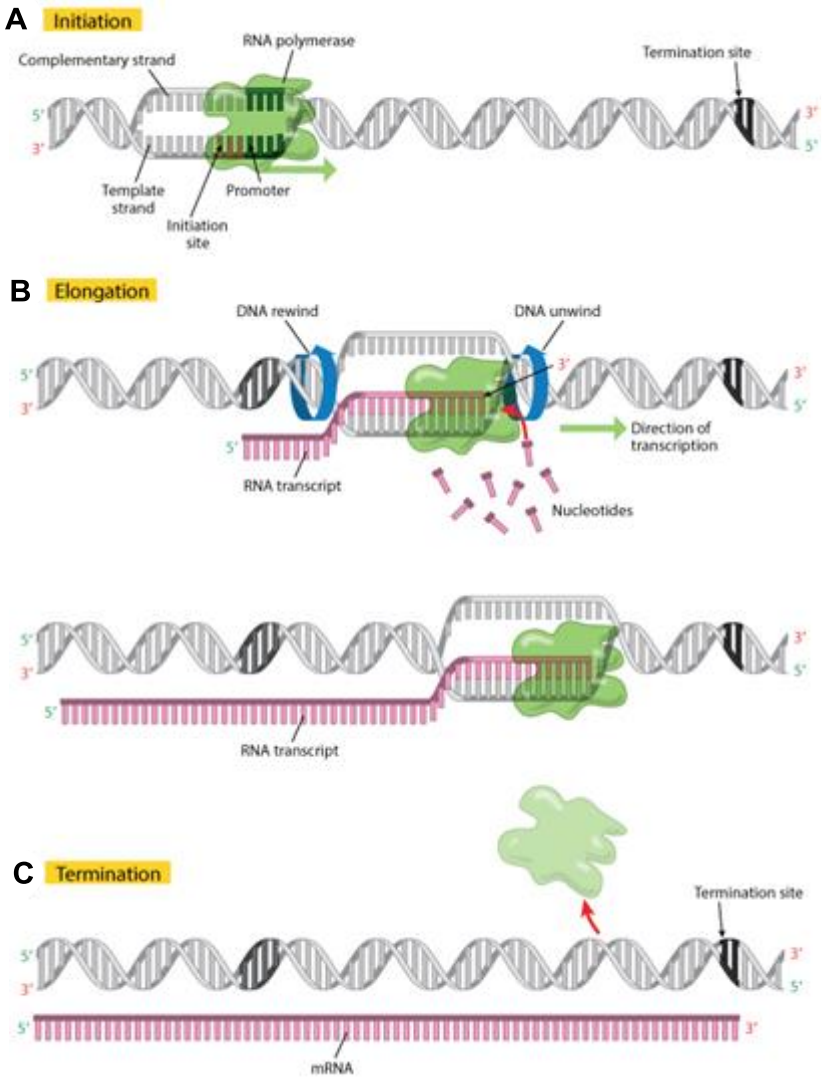


Figure 1.4: *Showing steps in transcription.* ***A)*** *RNA polymerase (green) binds to the DNA helix and unwind the promoter region (initiation).* ***B)*** *The DNA double helix unwinds and RNA polymerase reads the template DNA strand and adds nucleotides to the end of a growing RNA transcript (elongation).* ***C)*** *Transcription is terminated and the mRNA transcript and RNA polymerase are released from the complex (termination). (Credit: Nature Education)*[16]

Transfer RNAs (tRNAs) are responsible for matching the amino acids with the complementary mRNA docked at the ribosome. The tRNA molecule has two binding sites, one for the amino acid and the other for binding to the mRNA codon. Once the ribosome (grey) docks on the mRNA molecule, and charged tRNAs bring in amino acids by binding with the complementary mRNA codons. The tRNA mediated translation of genetic information is coordinated within the ribosome, and only two tRNAs can be next to each other within ribosome at any one time. Therefore, as a third

tRNA binds the mRNA codons, the first tRNA molecule is displaced, leaving its amino acid in the growing peptide chain. In this way, different length of specific peptides are synthesised as required by living cells (Figure 1.5).

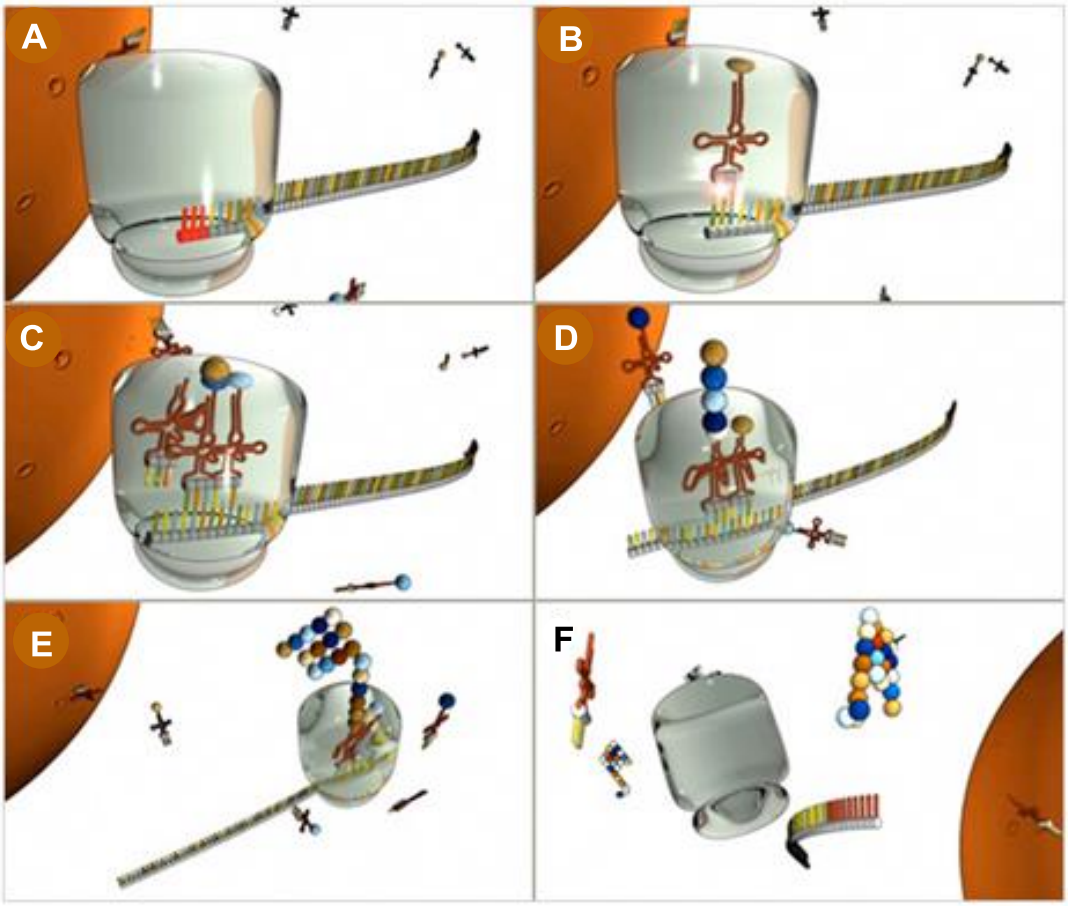


Figure 1.5: *Steps in translation.* ***A)*** *Ribosome (grey) docks on a start codon (red) of an mRNA molecule.* ***B)*** *The tRNA molecules carrying amino acids (spheres) dock at the corresponding triplet codon sequence on the mRNA molecule. (**C**, **D**, and **E**). This process is continuously repeated, connecting successive amino acids into a growing peptide chain of specific length as initiated in the transcription process.* ***F)*** *Translation is terminated when ribosome encounters a stop codon, releasing the polypeptide for use in the cell. (Credit: Nature Education)*[16]

1.2.2. The Central Dogma of Molecular Biology (CDMB)

Proteins are made of one or more different chains of amino acids folded into highly specific secondary and tertiary structures. Proteins and nucleotides, are both essential to almost every biochemical processes in a living organism (Figure 1.4 and 1.5). Interestingly, despite the power of accuracy of protein catalysis, it is of note that

perhaps the most important catalytic function in every cell: the synthesis of sequence specific peptide bonds, and consequently the construction of primary structure of proteins is carried out by the ribosome. This is of particular interest because the ribosome is an RNA catalyst (Section 1.2.1); therefore, it can be said that nucleotides not only store and transfer genetic information, but are also responsible for the most essential and universally conserved catalytic function/structure in cells. By examining the structure of these universally conserved organelles and their relationship across the phylogenetic tree, raises some simple questions that need to be answered to understand the route from chemistry to biological organisation. DNA and RNA are essential to protein synthesis, however protein (e.g. DNA and RNA polymerases) are essential to DNA and RNA synthesis and replication. Therefore, which of the essential biopolymers DNA or RNA or proteins, was synthesised first? Given the simplicity of nucleotide-nucleotide information transfer, it seems reasonable that nucleotides would have predated peptides in biology (or at least in information transfer in biological systems). However, this raises more questions such as, is it possible to abiotically synthesise DNA or RNA? Are there simple robust mechanisms to transform abiotically plausible material into any (or indeed all) of the essential biopolymers of life?

It is highly possible that we will never be able to fully answer these questions. However, these are the questions at the heart of understanding the origins of biology and life on Earth. Therefore, through incremental investigation of the molecular self-assembly of biopolymers we seek to better understand their origins in biology. Whilst developing new derivatives and analogues of the natural nucleotides or mechanisms to synthesise nucleotides, countless applications in diagnostics[17], medicine[18] and antiviral drugs[19] have also been found.

1.3. The RNA World Theory

The highly coordinated, systematic and optimised chemistry that DNA/RNA/proteins perform in the cell, make it ostensibly difficult to envisage how a living system could function without their interdependence. However, the RNA World Theory[1] provides a mechanism to by-pass this “chicken and egg paradox” at the heart of biology (Central Dogma of Molecualr Biology).[20] It postulates that before life was dependent

on an interrelationship between DNA/RNA/protein, there was an earlier stage of evolution in which life was entirely dependent upon RNA as the only biopolymer (Figure 1.6).

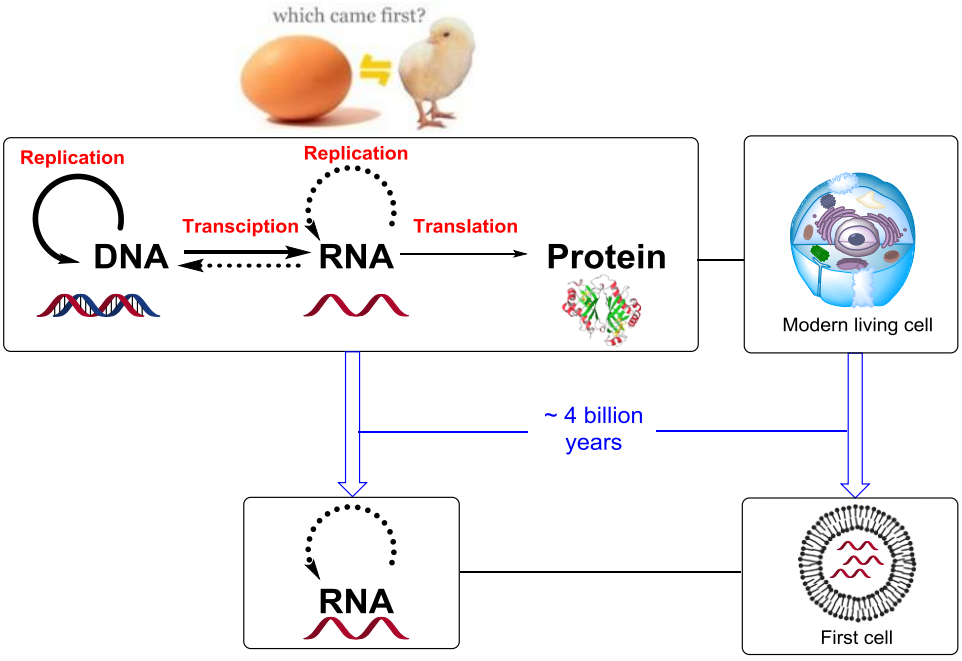


Figure 1.6: *The Central Dogma of Molecular Biology (CDMB). This figure shows schematically the flow of information starting from DNA→RNA→Protein. The interrelationship of DNA/RNA/Protein in the modern cell makes it difficult to envisage origins of life if life must rely on this triad of polymers. To simplify the problem, it has been postulated that RNA alone was used to store information and achieved programmable catalysis at the origins of life (the RNA World Hypothesis).*

The RNA World poses a scientific challenge to chemists to elucidate an abiotic synthesis of all four components (Figure 1.7) of RNA from plausible prebiotic feedstock of molecules, without these feedstocks the RNA World could never have initiated.

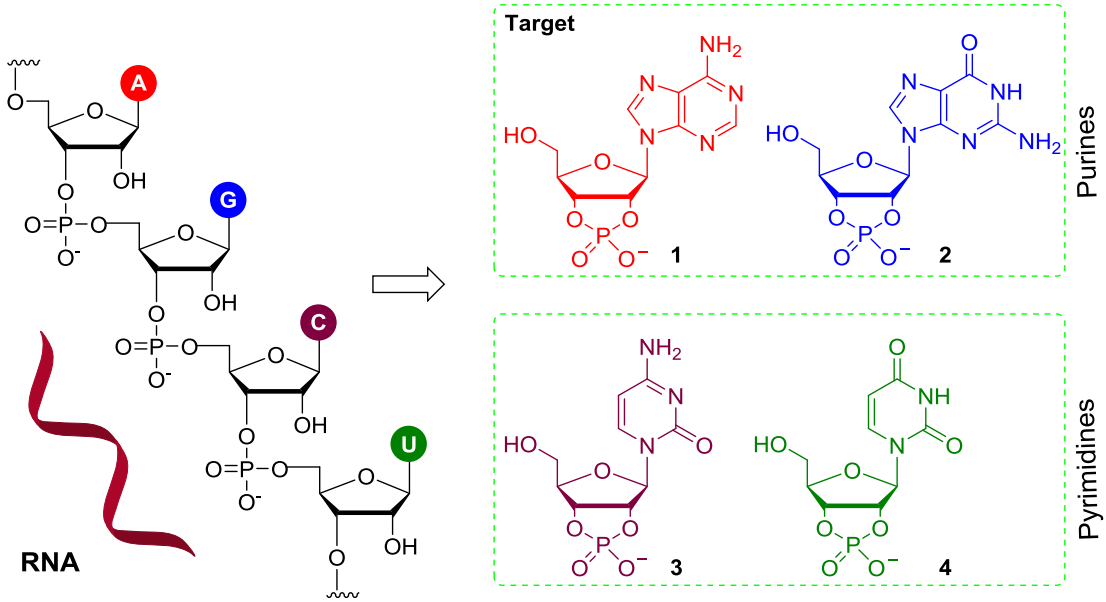


Figure 1.7: *Retrosynthetic analysis of ribonucleotide (RNA). Prebiotic synthesis of **1** and **2** are the target of this thesis, the pyrimidine ribonucleotides **3** and **4** have already been synthesised.*[21, 22] *G (guanine), A (adenine), C (cytosine), U (uracil).*

1.3.1. What makes RNA a better prebiotic candidate than DNA in the quest of the origin of life chemistry?

A number of lines of evidence suggest that RNA is the earliest biopolymer. All genes are transcribed into mRNA, tRNA or rRNA before the biological phenotype is accessed from a gene. Furthermore, strikingly some plant pathogens such as the tobacco mosaic virus have genomes entirely composed of RNA.[23] The biosynthesis of deoxyribonucleotides requires the enzymatic reduction of ribonucleotides, but the latter are synthesised in *de novo* biosynthetic pathways (Section 1.4), and many enzyme cofactors are derivatised ribonucleotides.

However, the discovery of ribozymes in 1981 by Cech[24] and Altman[25] provided substantial evidence for the potential role of RNA at the origin of life (and the first real validation of the catalytic process of RNA, essential to the RNA World hypothesis). Their discovery of catalytic RNA, or ribozymes, gave real credence to the proposed RNA world hypothesis. This work was recognised by awarding the Nobel Prize in Chemistry to Sidney Altman and Thomas Cech in 1981. Subsequent studies have also shown that certain RNAs can catalyse the synthesis of a complementary strand RNAs on a RNA template without any enzymatic intervention.[26]

RNA strands isolated from the genome of Tetrahymena thermophila have shown catalytic activity for the reactions of nucleotidyl,[27] phosphoryl-transfer,[28] hydrolysis of an aminoacyl ester[29] and the synthesis of complementary RNA strands.[30] The additional 2`OH group present in RNA, but missing in DNA, plays a significant role in the secondary structure of RNA molecule. For example, the 2` OH of RNA often plays an important role in the catalytic activity of ribozymes, due to its dual functionality (as a proton donor and acceptor, as well as provide a site for metal ion coordination) that facilitates a ribozyme to achieve stable 3D structures through increase intrastrand non Watson-Crick stabilising interactions, as well as supporting catalytic activity. It has also been shown that ribosomes (which are largely composed of RNA) can retain some catalytic activity, in spite of complete protein extraction.[31]

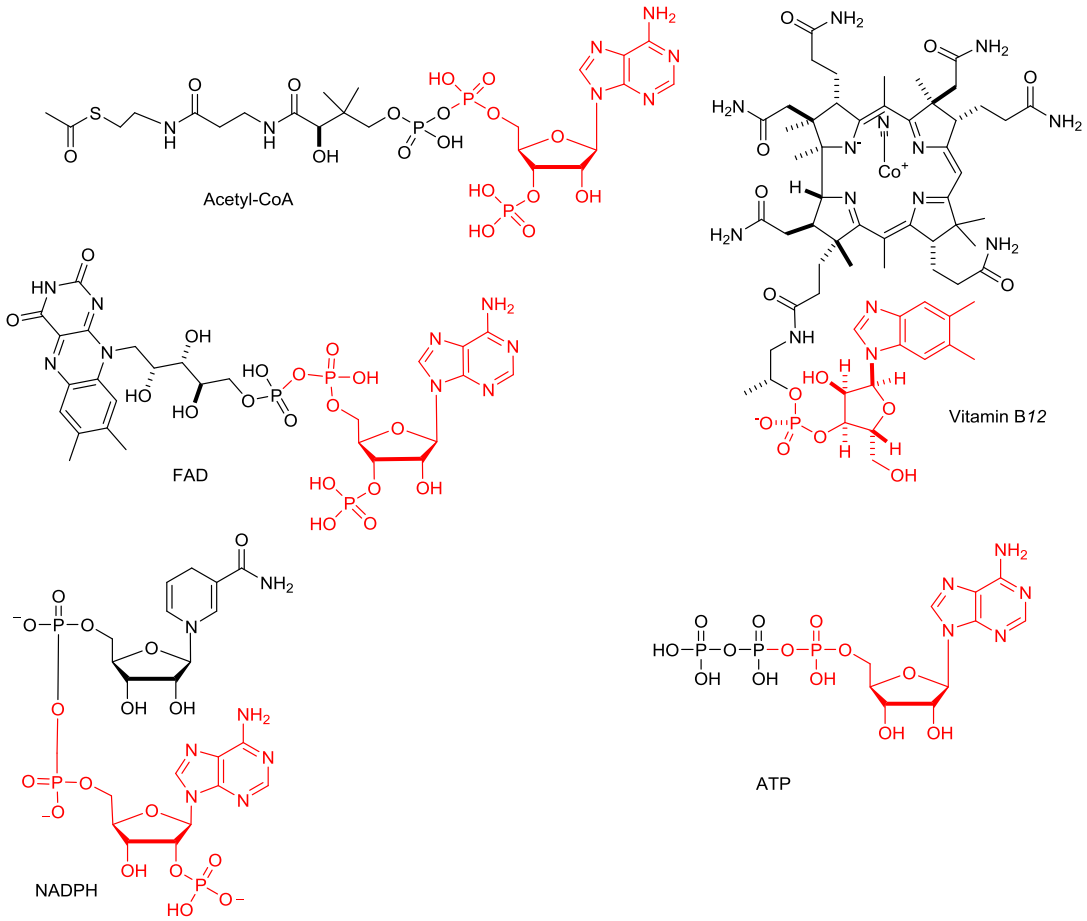


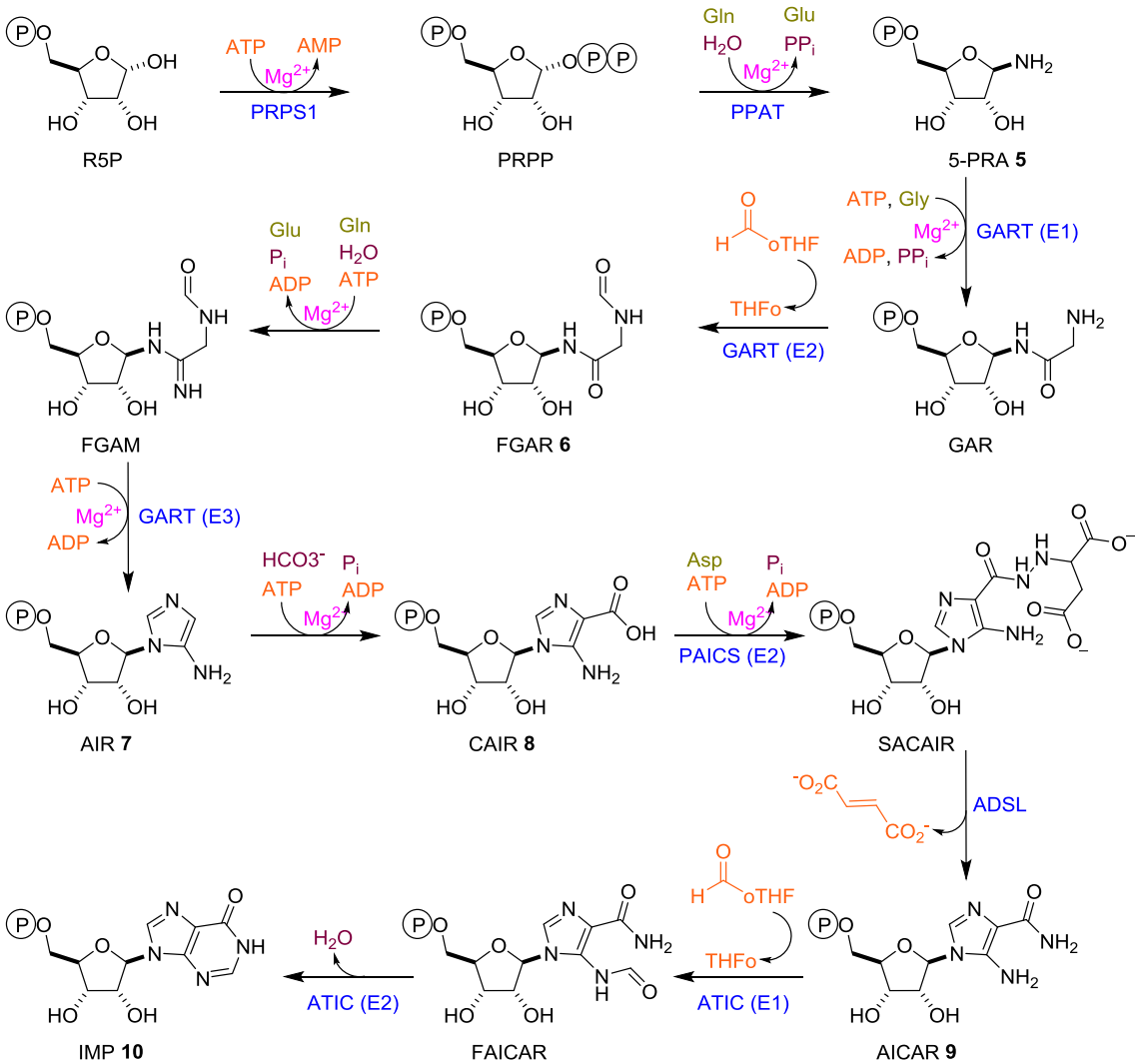
Figure 1.8: *Supposed molecular fossils of an RNA world. Nucleotides (red) are attached to different (bio)chemical structures.*

Many vital components of the modern cell are mostly composed of RNA or contains RNA derived structures such as enzyme cofactors (ATP, Acetyl-CoA, NADPH, FAD, Vitamin B-*12*). Coenzymes are complex organic structures, crucial to many enzyme

catalysed reactions, and more than half of the enzymes discovered so far require a coenzyme for full biochemical functioning.[32] It has been proposed that the RNA component of nucleotidic coenzymes are molecular fossils that are vestigial from the RNA world (Figure 1.8).[32]

1.4. *De novo* biosynthesis of purines

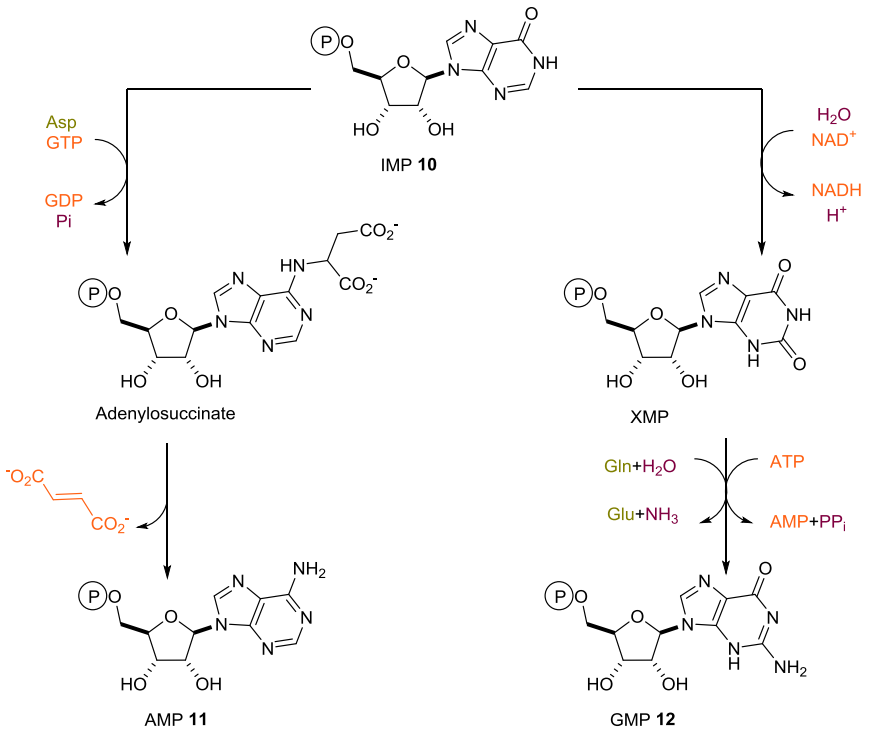
Purines are synthesized *de novo* through biosynthesis, catalysed by a complex set of enzymes in biology (Scheme 1.1).[33] Inosine monophosphate **10** is an intermediate in the synthesis of the canonical purine ribonucleotides (adenosine monophosphate **11** and guanosine monophosphate **12)** (Scheme 1.2).



Scheme 1.1: *De novo biosynthesis of inosine monophosphate **10** as purine nucleotide precursor of adenine and guanosine monophosphates **10** and **11**. The colour scheme is as follow: enzymes, coenzymes, substrate names, metal ions, inorganic molecules.*

The biosynthesis of purines starts by the displacement of pyrophosphate from PRPP by ammonia (NH_3) to furnish 5-phosphoribosyl-1-amine (5-PRA, **5**). The aminated ribotide **5**, reacts with glycine, formylated by tetrahydrofolate (THFo) and is then converted to amidine (FGAR, **6**). ATP induces cyclisation and amination in **6**, followed by dehydration to give 5-aminiimidazole ribonucleotide (AIR, **7**), which reacts with activated bicarbonate, followed by rearrangement to produce carboxyaminoimidazole ribonucleotide (CAIR, **8**). Aminolysis of the carboxylate moiety in CAIR is achieved, by reaction with aspartate coupled by conversion of ATP to ADP, followed by loss of fumarate to give AICAR **9**. AICAR then undergoes formylation by tetrahydrofolate (THFo), followed by cyclisation and dehydration to give inosine monophosphate **10**.

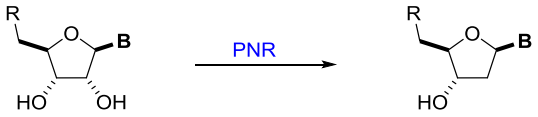
IMP **10** is then transformed into AMP **11** and GMP **12** in a series of enzyme-catalysed biochemical steps (Scheme 1.2).



Scheme 1.2: *Biosynthesis of purine nucleotides (AMP **11** and GMP **12**) from common canonical purine intermediate, inosine monophosphate **10**.*

The ribonucleotide reductase (RNR) enzyme catalyses the conversion of ribonucleotides into deoxyribonucleotides (Scheme 1.3). This is thought to be further evidence to show that RNA predates DNA in biology, as from a chemical perspective

conversion of DNA to RNA would be a difficult transformation to achieve without enzymatic control.[34]



Scheme 1.3: *Biosynthesis of deoxyribonucleotides (RNA) catalysed by ribonucleotide reductase (PNR).*

1.5. Prebiotic chemistry

The origins of life has been studied for over 60 years, and has engaged researchers from many different scientific disciplines. One of the leading theories for origins of life is based on the assumption that the evolution of RNAs was instrumental in the transition from chemistry to biology (Figure 1.9).

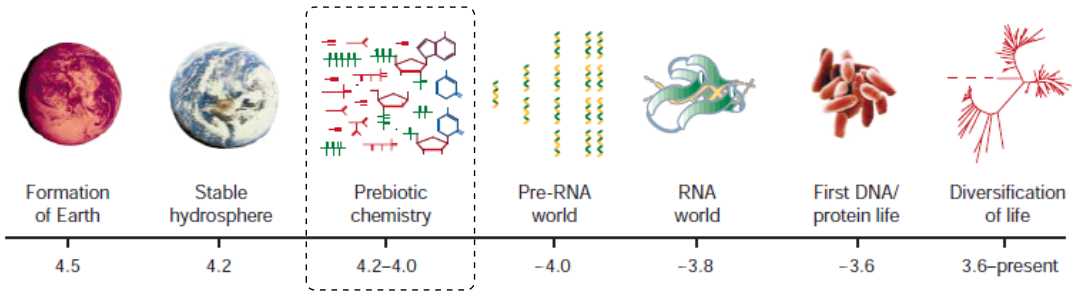
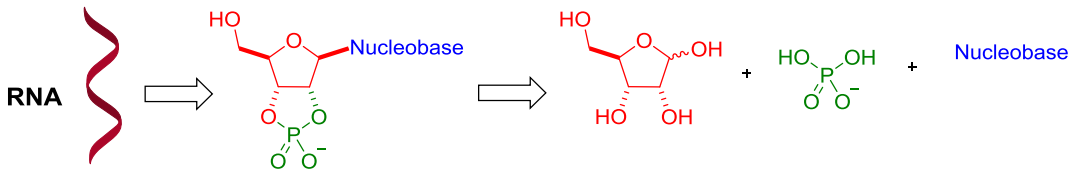


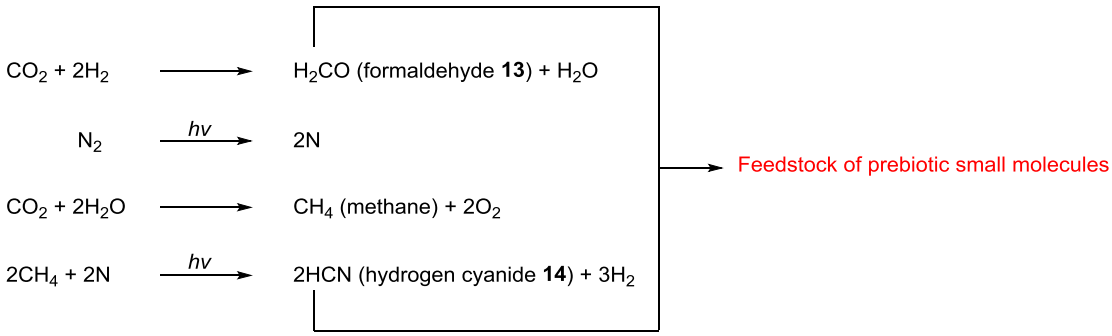
Figure 1.9: *A time line of origins of life based on RNA as the first molecule of life. The dates are in billions of years. (Credit: Joyce)*[35]

Examination of a ribonucleotide, points to a simple approachable synthesis (Scheme 1.4) through two condensation reactions. Accordingly, it was hypothesised that incubation of ribose, phosphate and the nucleobase would yield ribonucleotides and a 4000 million years old problem of ribonucleotide synthesis will be ostensibly solved. However, for such a simple retrosynthetic approach to work, the prebiotic feedstock for these nucleotide precursors (ribose, nucleobases) are required.



Scheme 1.4: *Retrosynthetic analysis of RNA.*

The nature of the early Earth's atmosphere being reducing or oxidising has been highly debated, but it is now believed that the primitive atmosphere was weakly reducing,[36] which can give rise to plausible reactions acting as feedstock of small prebiotic molecules (Scheme 1.5). However within the primitive atmosphere other localised environments would have co-existed, giving rise to highly oxidising or more reducing localised conditions.[37]



Scheme 1.5: *Plausible reactions in weakly reducing primitive atmosphere yielding major prebiotic small molecules.*

In an attempt to simulate the (then supposed to be) strongly reducing conditions of primitive Earth, a pioneering prebiotic chemistry experiment was conducted by Harold Urey and Stanley Miller in 1953.[38] Urey and Miller observed that a range of organic molecules were synthesised upon passing an electric discharge through water vapour, ammonia, methane and hydrogen (Figure 1.10). Remarkably, in these spark discharge experiments proteinogenic amino acids (α-alanine, glycine and aspartic acid) were detected, along with other molecules of prebiotic importance including non-proteinogenic amino acids such as β-alanine.[38]

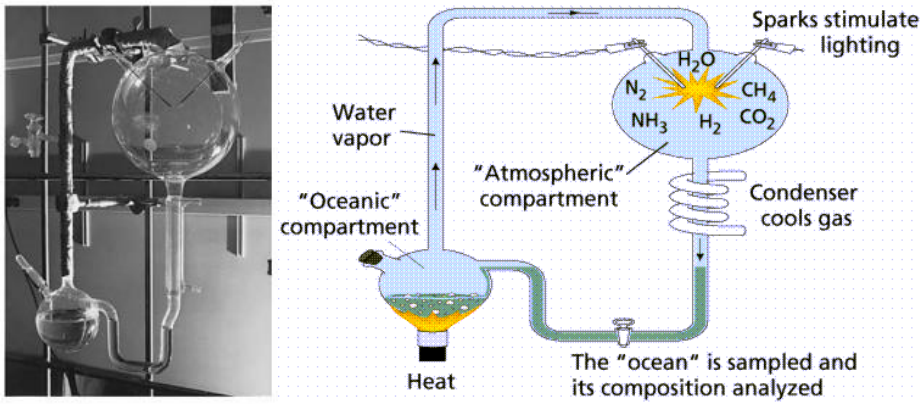
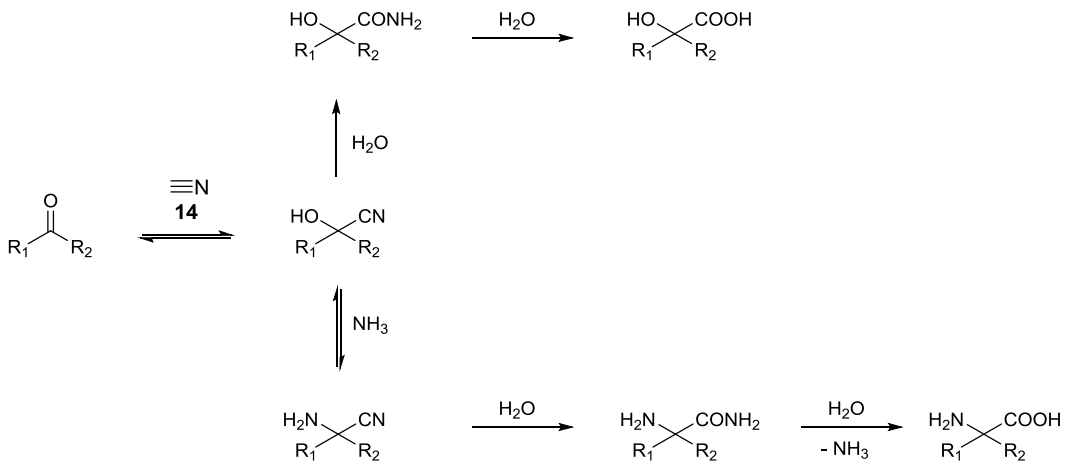
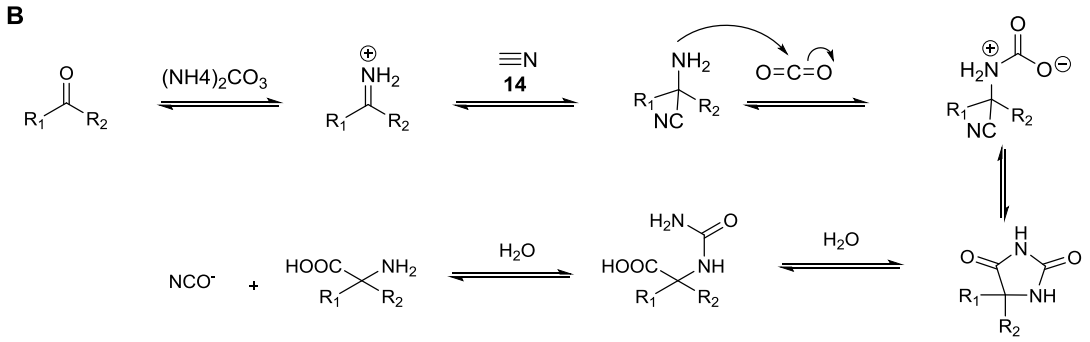


Figure 1.10: *The apparatus used by Urey and Miller to synthesise amino acids and other organic compounds in a reducing atmosphere, the apparatus is made entirely of glass with the exception of electrodes which are made from tungsten (left).*[38] *A diagrammatic representation of Urey Miller's apparatus (right).*[25]

Perhaps, after the controversial synthesis of urea (an organic molecule) from ammonium cyanate (an inorganic molecule) in 1828 by Friedrich Wohler,[39] Urey-Miller's synthesis of amino acids by spark discharge were the most exciting prebiotic chemistry experiments to have been conducted at the time. For the first time, it was shown that the basic building blocks of the most complicated biological machines (proteins) can be obtained by bombarding gaseous carbon, oxygen and nitrogen matter with simple spark discharge. Detailed inspection of Urey-Miller experiments are now known to produce up to 40 different amino acids and amines, mainly via the Strecker or Bucherer-Bergs mechanisms (Scheme 1.6).[40]

A





Scheme 1.6: *Prebiotic synthesis of amino acids.* ***A)*** *The Strecker mechanism for the formation of amino acids, and complementary synthesis of hydroxyacids.* ***B)*** *The Bucherer–Bergs mechanism for the synthesis of amino acids, which employs CO_2 to promote synthesis for (Strecker product) aminonitriles.*

Following in the footsteps of Urey-Miller, others have shown that hydrogen cyanide **14** can be synthesised by spark discharge through an atmosphere of methane, nitrogen and carbon monoxide gases,[41] and similar laboratory experiments have shown to produce acetylene, cyanoacetylene **30** [42] and cyanogen **25**.[43]

Many small molecules have also been detected in the interstellar dust clouds and meteorites,[44] these molecules are assumed to be non-biological in origin, their non-biological syntheses have also been confirmed by laboratory experiments to mimic the primordial environment of Earth and meteorite formation (Figure 11).[45]

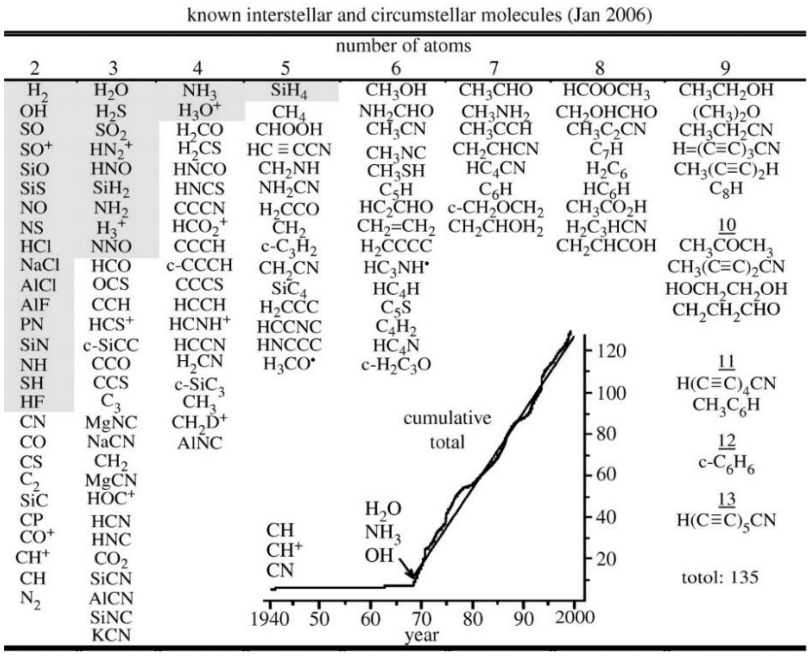
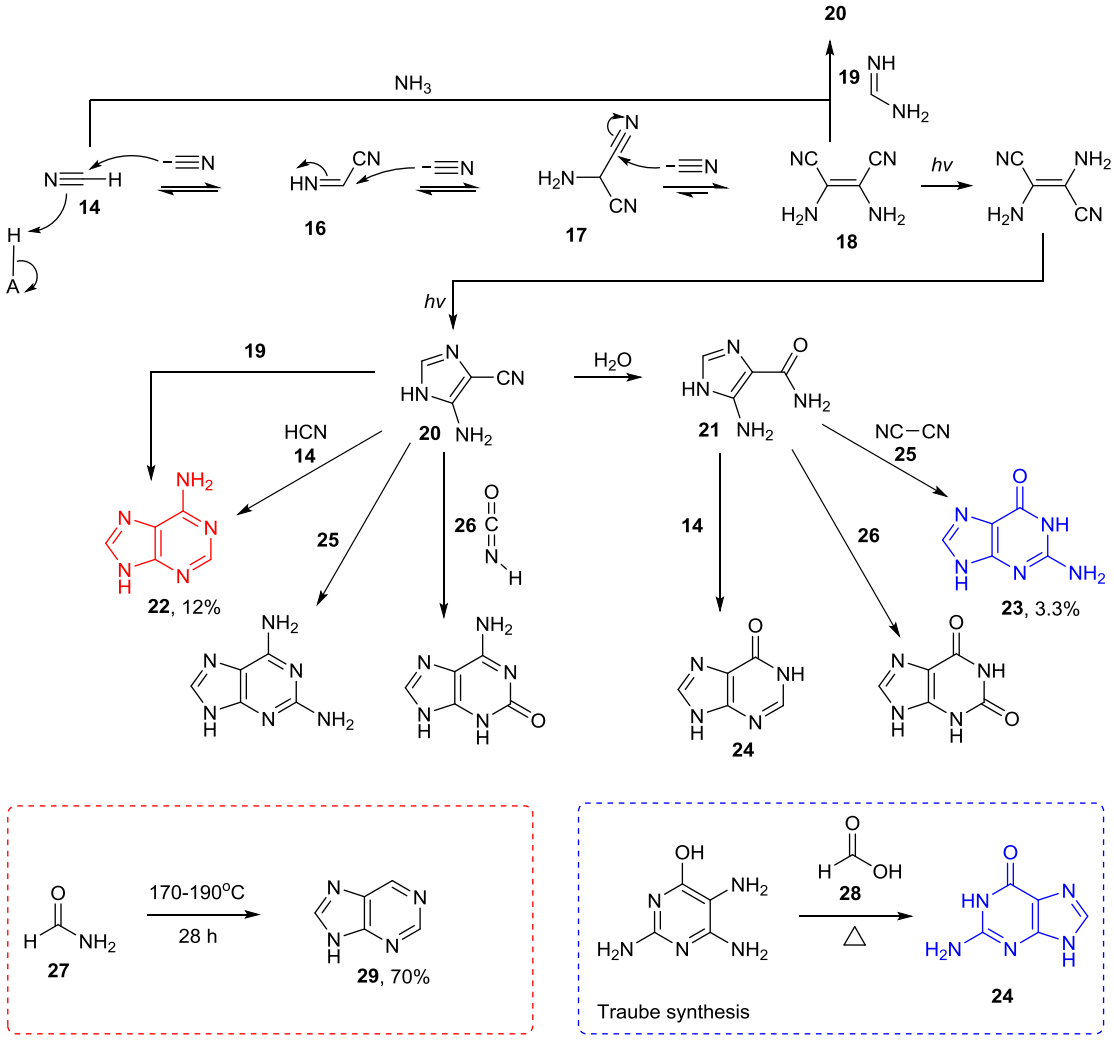


Figure 1.11: *Prebiotic feedstock of molecules observed in the interstellar cloud. (Credit: Thaddeus)*[45]

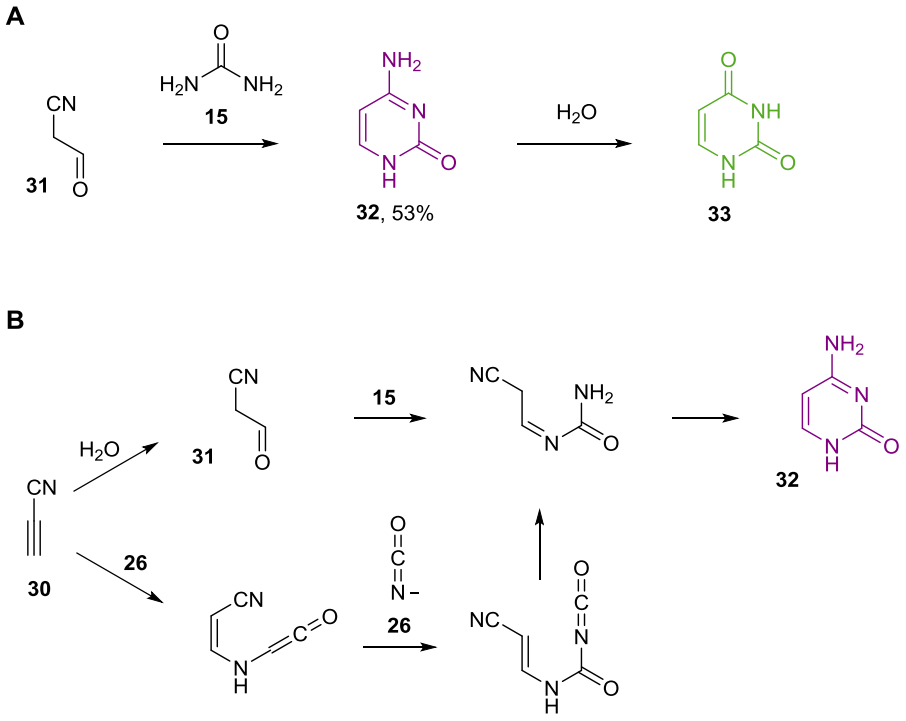
1.5.1 Prebiotic synthesis of ribonucleotide nucleobases

Hydrogen cyanide **14** derivatives are known to be plausible precursors of purine nucleobases.[38] Oró first demonstrated that the oligomerisation of **14** can lead to the synthesis of adenine **22**. Adenine **22** ($C_5H_4N_5$) is formally a pentamer of **14**, and can be synthesised along with other important prebiotic compounds from **14**-oligomerisation.[46-49] Ferris and Orgel meticulously studied Oró's HCN oligomerisation reactions in order to seek plausible mechanisms for the formation of **22** (Scheme 1.7).[50-54]



Scheme 1.7: *Prebiotic synthesis of purine nucleobases via hydrogen cyanide **14** oligomerisation. Adenine **22** (12%) and guanine **23** (3.3%) is obtained under best possible prebiotic conditions. The photochemical reaction of HCN tetramer **18** was discovered by Orgel and Ferris. Heating of formamide **27** yielding 70% purine **29** (inset, red). Traube's synthesis of purines (inset, blue).*

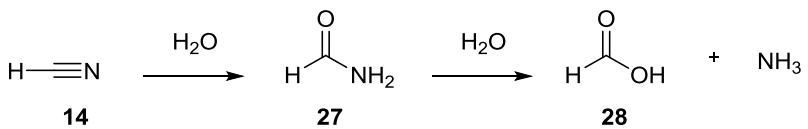
Incubation of a concentrated (~11M) mixture of hydrogen cyanide **14** yields HCN-dimer **16**, which can react with further HCN **14** to form aminomalononitrile **17**, followed by another HCN **14** to give diaminomalononitrile **18**. The tetramer **18** then undergoes intramolecular photochemical rearrangement (or reacts with formamidine **19** synthesised from **14** and NH_3) to furnish 5-aminoimidazole-4-carbonitrile (AICN) **20** in near quantitative yield. Subsequent hydrolysis of **20** gives 5-aminoimidazole-4-carboxamide (AICA) **21** (Scheme 1.7). Interestingly, during the formation of diaminomalononitrile **18**, other molecules with prebiotic significance were also observed, these include cyanamide **68**, urea **15**, and cyanogen **25**. Formylation and cyanation of **20** and **21** by various reagents leads to formation of several purine nucleobases including adenine **22**, guanine **23** and hypoxanthine **24** (Scheme 1.7). However, the best yield achieved so far for adenine **22** and guanine **23** synthesis are 12% and 3.3% respectively under plausible prebiotic conditions.[55, 56] Other notable purine synthesis (with prebiotic implications) include polymerisation of formamide **27** at elevated temperatures to give yield up to 70% purine **29**,[57] and Traube's purine synthesise which yields purines by the reaction of derivatised diaminopyrimidines with formic acid **28** (Scheme 1.7).[58-60]



Scheme 1.8: A) *prebiotic synthesis of cytosine* ***32*** *and uracil* ***33*** *by reaction of cyanoacetaldehyde* ***31*** *and urea* ***15****.* ***B)*** *Plausible mechanism for the prebiotic synthesis of pyrimidines as postulated by Miller and Cleaves.*[61]

Pyrimidine bases have not received much attention as the purines in the prebiotic literature, but the highest yielding prebiotic conditions for synthesise cytosine **32** was reported by Miller *et al.* Miller and co-workers demonstrated that incubation of highly concentrated cyanoacetaldehyde **31** with urea **15** yielded up to 53% cytosine **32** (Scheme 1.8A).[62] The other pyrimidine nucleobase, uracil **33**, can be obtained via hydrolysis of **32** (Scheme 1.8B).[61]

The pyrimidine and purine nucleobases (assumed to be) required for the synthesis of ribonucleotides have been synthesised, but with pitfalls. The synthesis relies on concentrated HCN **14** pools which if uncontrolled can lead to the most deleterious pool of prebiotic chemistry. Upon synthesising such a diversity of compounds it is not clear how the "essential" biological compounds could be selected from such complex mixtures. Adenine **22** requires HCN **14** concentration of at least 0.01M.[63] However, detailed modelling has estimated HCN **14** concentrations in the primitive oceans would have been 6×10^{-16} M, 7×10^{-13} M and 2×10^{-6} M at 200°C, 100°C and 0°C respectively.[55] In dilute concentrations hydrogen cyanide **14** hydrolyses to formamide **27**, which then hydrolyses to formic acid **28** and ammonia (Scheme 1.9). Therefore, the plausibility of HCN **14** polymerisation taking place in the warm primitive oceans is rejected, in favour of eutectic freezing conditions or localised small bodies of water that have become concentrated in HCN **14**. However, this still does not provide a solution for the uncontrolled HCN chemistry which predominantly yields polymers and non-biological aromatics. Moreover, the feedstock of cytosine **33**, acetaldehyde **31** (pKa = 8.1) readily undergoes homoaldol reaction, implying that the controlled synthesis of cytosine **33** would be prebiotically difficult to orchestrate.[64, 65]



Scheme 1.9: *Hydrolysis of hydrogen cyanide **14** under dilute conditions.*

1.5.2. Prebiotic synthesis of sugars

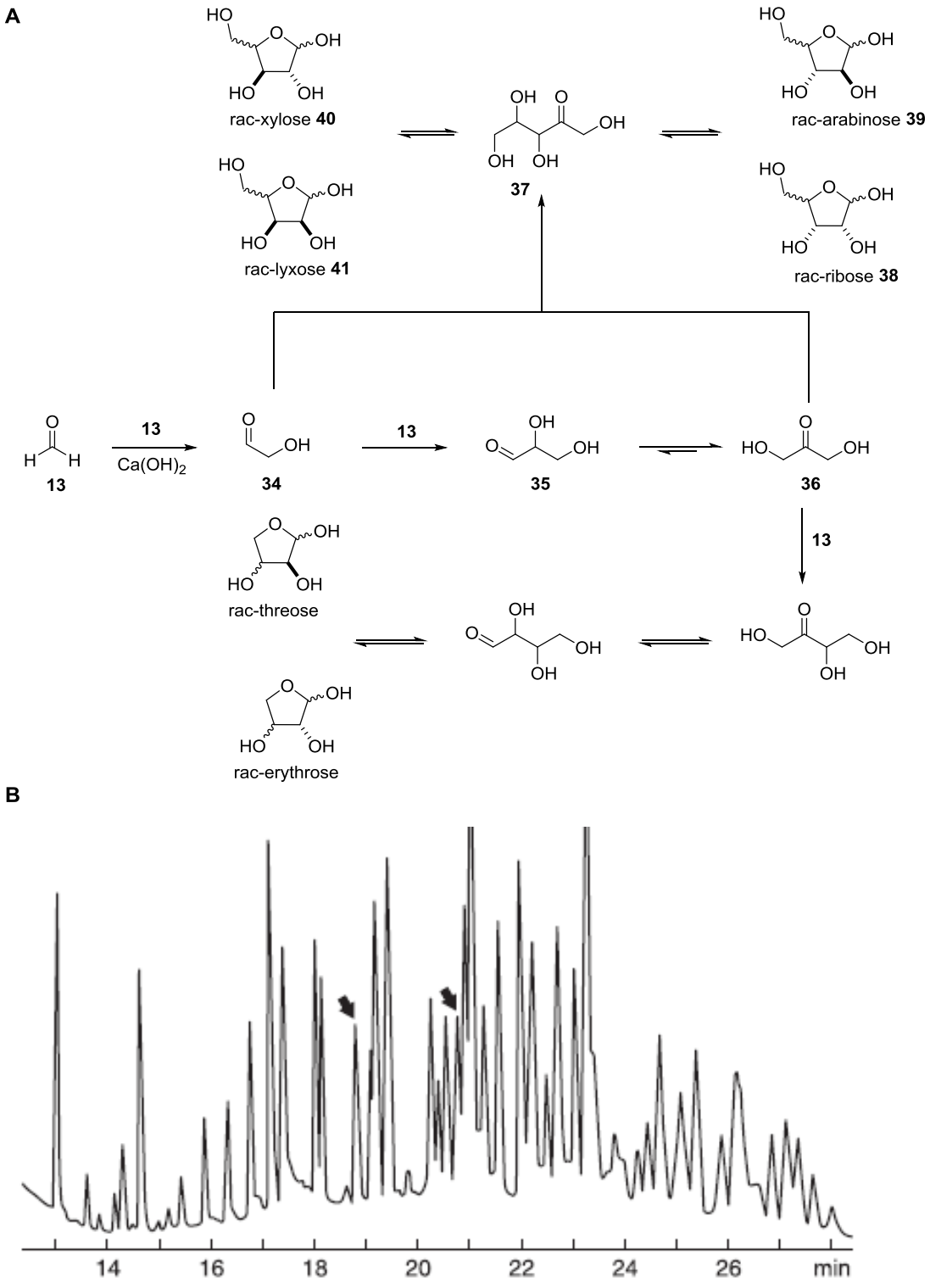
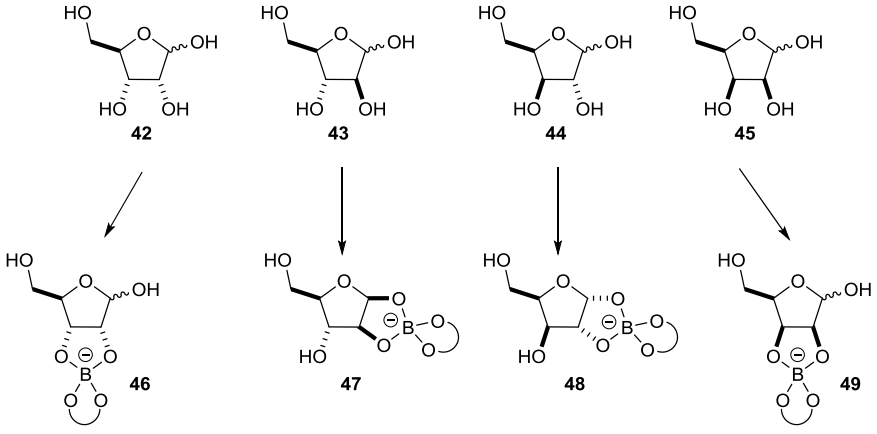


Figure 1.12: A) *Simplified schematic representation of Butlerow's formose reaction.* ***B)*** *Gas chromatogram of derivatives of sugars arising in the formose reaction. The arrows are pointing to two ribose* ***38*** *isomers. (Credit: Decker et al.)*[66]

The formose reaction, discovered by Butlerow in 1861, is the polymerisation of formaldehyde **13** in the presence of divalent metal ion catalysts such as $Ca(OH)_2$ in highly alkaline conditions to yield sugars and other small sugar-like molecules (Figure 1.12).[67, 68] The exact details of Butlerow's sugar synthesis have still not been fully understood more than a centrury after it's discovery, however it is thought to preceed via an autocatalytic process in which glycolaldehyde **34** acts as an autocatalyst for its own formation from formaldehyde **13** (following two aldol reactions with formaldehyde and a retro-aldol dissociation of tetrose sugars into two molecules of glycolaldehyde **34**). Formose reaction lead to the synthesis of a plethora of sugars, yielding mixtures of trioses, tetroses, pentoses, and hexoses alongside many other related compounds.[66, 69] Although the formose reaction leads to ribose **38** and related aldopentoses, it is debatable whether it could have led to the accumulation of these particular sugars on the early Earth, since ribose **38** is an extremely minor component (1% yield) of the sugars synthesised through the formose reaction and given that ribose is particularly an unstable ($t_{1/2}$ = 5 h, pH 12.5) product of the reaction.[70-75] To add to the "ribose problem", it has been shown that ribose **38** equilibrates to yield a mixture of pentose sugars under mild aqueous conditions (pH 7, 25°C), where arabinose **39** (75%) predominates over ribose **38** (19%) at equilibrium.[70]

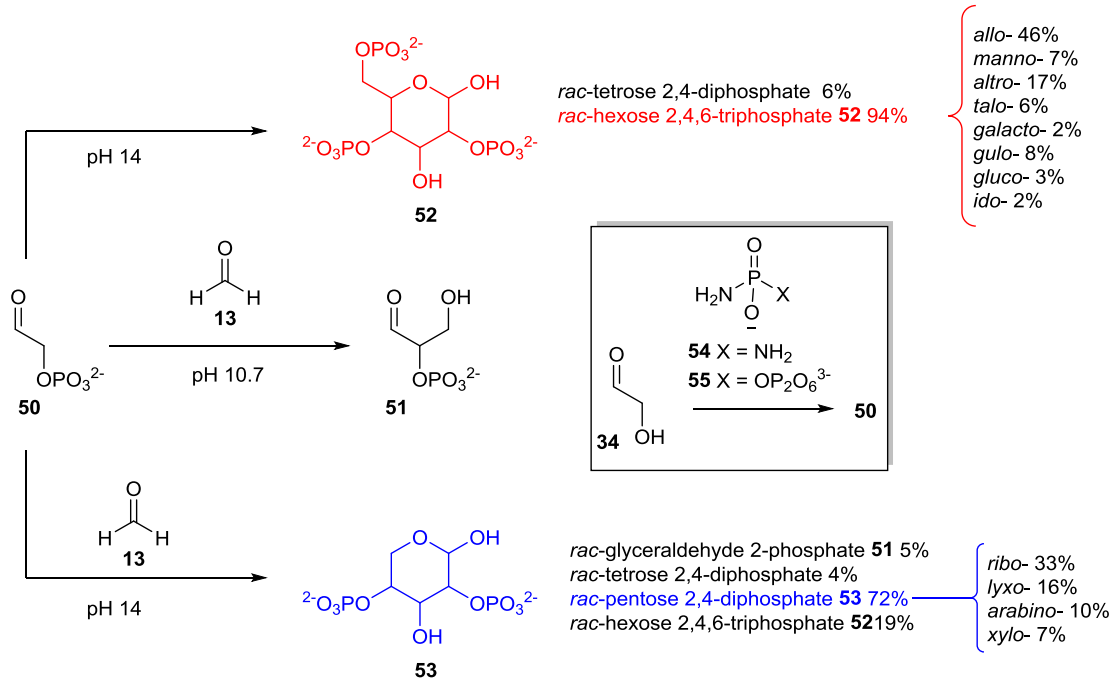


Scheme 1.10: *Stabilisation of pentose sugars by borate incorporation under formose reaction conditions.*

Many attempts to simplify the pattern of products synthesised by the formose reaction have been made, one such method exploits stabilisation of sugars by chelation to borate (Scheme 1.10).[76] It has been found that complexation with borate minerals stabilizes pentoses sugars (ribose **42**, arabinose **43**, xylose **44** and lyxose **45**,) under

formose reaction conditions.[77, 78] Interestingly, the ribose-borate sugar complex **46** exhibits greater stability relative to the other aldopentoses (**47**, **48** and **49**), however, under formose reaction conditions this stability is modest ($t_{1/2}$ = 45 hr, pH 12.5).[72, 79] It is also important to note, that ribulose **37** has shown a greater degree of stabilization by borate than ribose **42**.[72]

Eschenmoser and colleagues have shown that phosphorus incorporation into sugars provide greater selectivity and stability of sugars (Scheme 1.11).[80] Glycolaldehyde phosphate **50**, formed by the selective *α*-phosphophosphorylation of glycolaldehyde **34** by amidophosphates (**55** or **54**) in near quantitative yields, undergo homoaldol condensation to produce hexose 2,4-6-triphosphates **52** in a remarkable 94% yield (Scheme 1.11).[81, 82] Glycolaldehyde phosphate **50** can also react with formaldehyde **13** at pH 10.7 to give glyceraldehyde-2-phosphate **51**. The crossed aldol condensation reaction of glycolaldehyde phosphate **50** and glyceraldehyde-2-phosphate **51** give a mixture of phosphorylated pentoses containing 72% pentose-2,4-diphosphates **53** in which ribose **38** is the major component (33%).

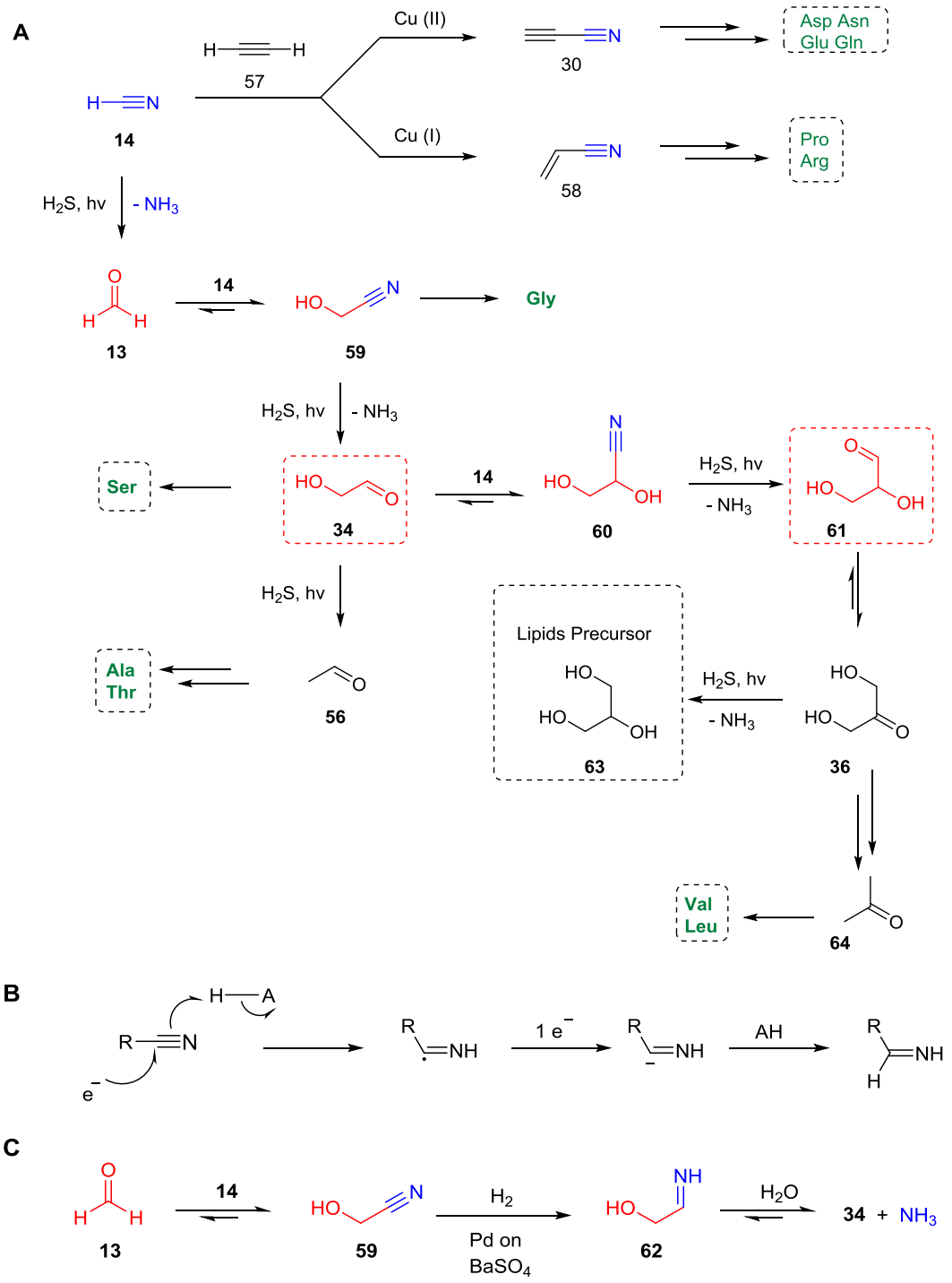


Scheme 1.11: *Phosphorylation of sugars providing selectivity and milder Formose reaction conditions. Glycolaldehyde **34** undergoes selective α-phosphorylation by amidophosphates (**54** or **55**) to furnish glycoldaldehyde phosphate **50** in excellent yield (inset). Homoaldol condensation of **50** produces a mixture of rac-hexose 2,4,6,-triphosphates **52** (94%). Adapted from Islam and Powner.*[83]

In comparison with the formose reaction, the phosphate groups incorporated in the sugars, prevent the rearrangement of trioses, tetroses, and pentoses, therefore inhibiting the uncontrollable formation of sugars (tars) that is the characteristic of the formose sugar mixture. However, these reactions are still not completely disastereoselective, where a mixture of sugar phosphates mixture is still produced. Furthermore prebiotic nucleobase glycosylation with these phosphorylated sugars have not been reported.

Until recently, the formose reactions and its variants were thought to be the prebiotic source of sugars.[77, 84] However, these are extremely unselective and give a plethora of sugar.[70, 85] Sutherland and colleagues have found a different solution to the sugar problem by exploiting the Kiliani-Fischer homologation (Scheme 1.12C)[86, 87] of formaldehyde to find a prebiotic route for the synthesis of glycolaldehyde **34** and glyceraldehyde **35**.[88] Their rather simple system, which generates relatively complex organic biomolecule precursors, is termed by Sutherland "cyanosulfidic protometabolism" (Scheme 1.12A). These pathways form glycolaldehyde **34** and glyceraldehyde **35** (essential for the prebiotic synthesis of ribonucleotides, *vide infra*),[22] the hydrophilic moiety **63** of glycerophospholipid membranes, and twelve of the proteinogenic amino acid precursors in high yielding reaction step.[89] The key step of the reaction is the remarkable reduction step achieved by photochemical irradiation of cyanide **14** with hydrogen sulfide (H_2S). The photochemical reduction of hydrogen cyanide **14** with H_2S produces formaldehyde **13** and ammonia, **13** is rapidly trapped by **14** to generate glycolonitrile **59**, which can undergo further nitrile reduction to yield glycolaldehyde **34** (Scheme 1.12A).

The formation of cyanohydrins such as **59** has been a major drawback of formose and HCN oligomerisation, as the formation of a cyanohydrin between formaldehyde **13** and cyanide **14** is supremely favoured towards cyanohydrin **59**.[70, 90, 91]. However, Sutherland and colleagues have utilised the highly efficient cyanohydrin formation to their advantage, since glycolonitrile **59** is irreversibly reduced to form glycolaldehyde **34** by the photochemical irradiation in the presence of H_2S.[88, 92] The C_3 sugar glyceraldehyde **35** is also obtained, by Kilinai-Fischer reaction of glycolaldehyde **34**.



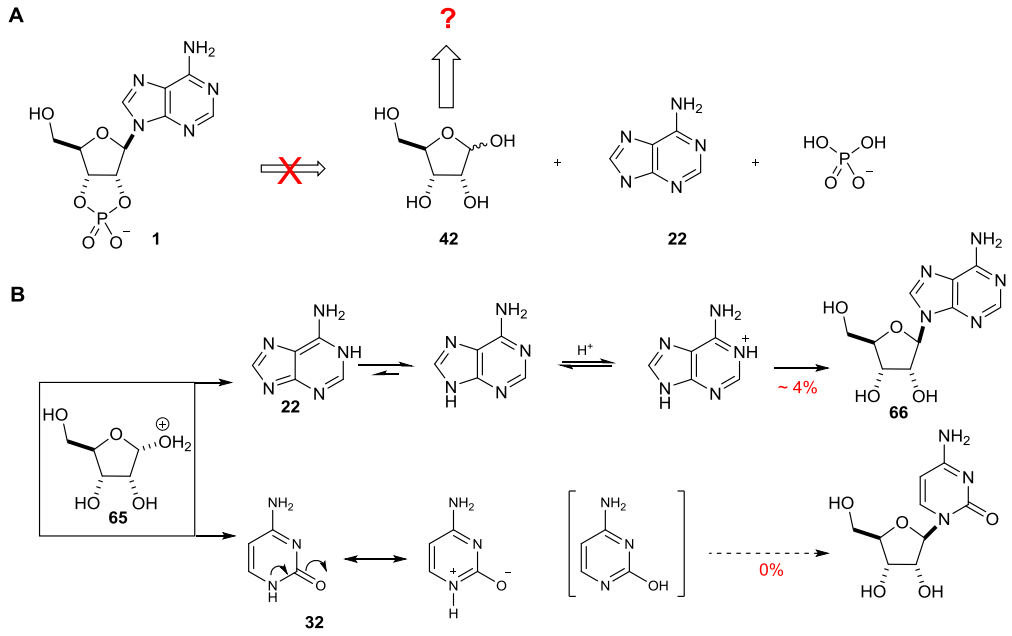
Scheme 1.12: *Cyanosulfidic protometabolism.* ***A)*** *Photochemical reduction of HCN* ***14*** *with H_2S and/or cyanocuprates leading to a network of systems chemistry furnishing glycolaldehyde* ***34*** *and glyceraldehyde* ***35****, glycerol* ***63*** *(lipid precursor) and 12 proteinogenic amino acid precursors in high yields.* ***B)*** *Plausible mechanism for photochemical reduction of cyanide to form iminium species.* ***C)*** *Kiliani-Fischer homologation, where the reduction of the nitrile moiety is achieved using a poisoned palladium catalyst. Adapted from Islam and Powner.*[83]

Furthermore, upon close inspection of glycolonitrile **34** reduction with H_2S, acetaldehyde **56** is also detected as the photochemical deoxygenation product of glycolaldehyde **34**. Interestingly acetaldehyde **56** is the precursor of amino acids threonine (Th) and alanine (Al). This led to the realisation that cyanohydrins **59** and **60** could be converted to amino acid precursors as well as glycolaldehyde **34** and glyceraldehyde **35** (ribonucleotide precursors), establishing a prebiotic synthetic association between protein and ribonucleotides under "cyanosulfidic" conditions.

Formaldehyde **13** and glycolaldehyde **34** are the Strecker aldehydes of glycine (Gly) and serine (Ser), and their cyanohydrins **59** and **60** are readily converted to the corresponding aminonitriles in the presence of ammonia. Moreover, dihydroxyacetone **36**, the phosphate-catalyzed isomerisation product of glyceraldehyde **35**, can undergo deoxygenation by irradiated H_2S to yield acetone **64**, which provides the branched backbone of valine (Val) and leucine (Leu). It is remarkable to observe these facile deoxygenation reactions with H_2S and photons, given the levels of intricacy required to achieve desired and controlled chemical transformation under prebiotic condition. To complete the remarkable collection of 12 different amino acids, copper-mediated cross-coupling of cyanide **14** and acetylene **57** is used to produce precursors for asparagine (Asn), aspartate (Asp), glutamine (Gln), glutamate (Glu), proline (Pro), and arginine (Arg). To establish that the same network of reactions might have been used by biology to build membranes, glycerol **63** provides a direct connection to phospholipids is obtained by the reduction of dihydroxyacetone **36**. Although Sutherland's cyanosulfidic protometabolism at this stage is preliminary in the sense that such a simple system gives rise to an ostensibly complex set of bimolecular chemicals. Further investigation into realising this chemistry in systems is required. Nevertheless the 'cyanosulfidic protometabolism' scenario is rather remarkable for the ease with which ribonucleotide, protein, and lipid precursors are formed. It has also reduced the invoking of notoriously messy and uncontrollable chemistry of the formose reactions and HCN oligomerisation, and spark-discharge experiments,[46, 63, 70, 93, 94] and has opened up a more cohesive approach to tackling the origins of life chemistry.[95]

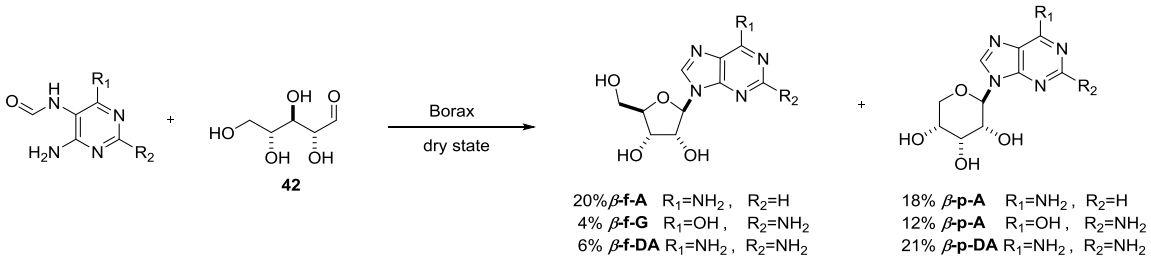
1.5.3. Condensation of sugars and nucleobases: Conventional synthesis of nucleotides

As discussed in the last section, prebiotic synthesis of ribose **42** has many pitfalls and is generally considered impractical and implausible to have existed on the primitive Earth.[70, 83, 85] Nevertheless, many continue to assume that a plausible pathway to ribose **42** will be found and have continued with exploring glycosylation of canonical nucleobases with ribose **42**. This, however, has proven very difficult to achieve (Scheme 1.13A). Ribosylation of canonical pyrimidines does not work at all, and incubation of adenine **22** with ribose **42** only gives 4% β-adenosine **66** in a solid-state reaction involving a large excess of ribose and a high concentration of Mg^{2+}, followed by heating with concentrated ammonium hydroxide.[96, 97] The reasons for the poor reactivity of ribose **42** and the canonical nucleobases are mainly due the unavailability of the right isomer of the ribose **65** and correct (nucleophilic) tautomer of each base (Scheme 1.13B). The sugar mixture consists mainly of β- and α-pyranose isomers with only 7% of the desired α-furanosyl isomer **65**.[98] In pyrimidines, the lone pair of the desired endocyclic nitrogen is delocalised into the ring in favour of aromaticity therefore no reaction is observed with **65** (Scheme 1.13B). In case of purines again the required tautomer (unprotonated *N*-9) is available at low levels at equilibrium, leading to reduced reactivity with the ribose (Scheme 1.13B).[99]



Scheme 1.13: *A) Retrosynthetic analysis of ribonucleotides.* ***B)*** *The difficulties of selective ribosylation of adenine* ***32*** *and cytosine* ***33****, due to the unavailability of the correct tautomer. The ideal isomer of ribose,* ***65*** *for ribosylation of nucleobase is a minor product (7%) in aqueous media (inset).*

The most recent synthesis of purine nucleoside, reported by Carell and co-workers, exploiting a combination of this dry-state glycosidation strategy with Traube purine synthesis to overcome the problem of glycosidation (Scheme 1.14), however this synthesis remains low yielding, lacking regio- and stereo-selectivity, and does not (yet) provide a solution for selective incorporation of phosphorus to get purine nucleotides (**1** and **2**).[59]



Scheme 1.14: *Synthesis of purine nucleosides by dry state glycosidation of Traube purines yielding both furanosyl and pyranosyl nucleosides. A (adenine), G (guanine) and DA (diaminopurine).*

1.6. Pre-RNA World

Given the synthetic complexity of ribonucleotides (and the difficult history that prebiotic chemist have had synthesising them), a pre-RNA World has also been suggested, where earlier "simpler nucleotides" where the first biopolymers and then RNA "appeared" later on and was incorporated into biology in a genetic takeover event.[100]. There are viable questions regarding nature's selection of RNA. Why does RNA use the 4 canonical nucleobases (adenine, guanine, cytosine, uracil) from a possible selection of at least 36 similar purines and pyrimidines?[13] Why did RNA assume the pentose sugar backbone instead of a hexose or tetrose sugar? Why is it *ribo*-furanose, instead of *ribo*-pyranose? Why are the nucleobases in β configuration instead of α? Why D sugar and not L? Why is the ionisable linker that connects RNA monomers a phosphate? And why is the 3`,5` phosphodiester linkage favoured rather than the 2`,5` phosphodiester linkage (Figure 1.13). In an attempt to answer some of these questions, plausible alternative candidates to ribonucleic acids, which are closely related to ribonucleic acid, have been chemically synthesised and their ability to undergo Watson-Crick pairing has been explored.

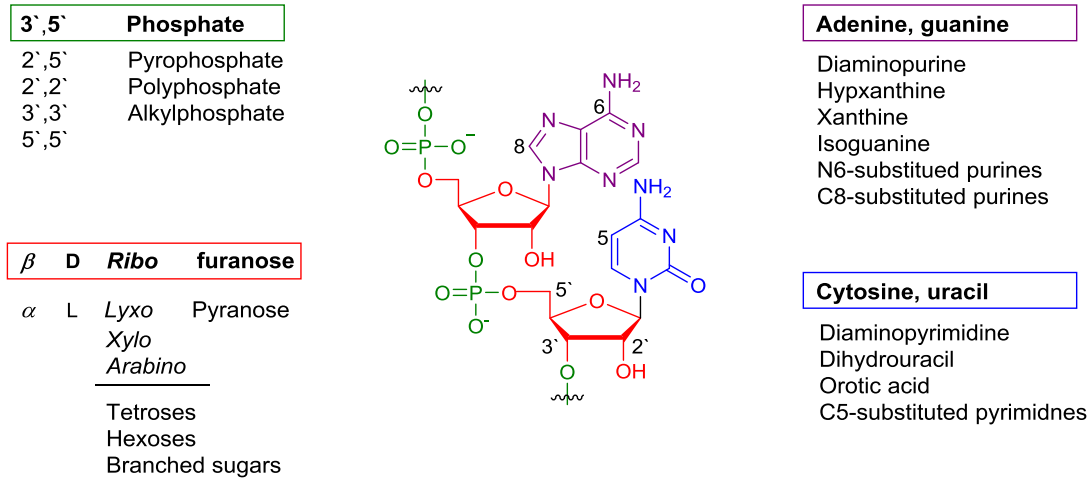
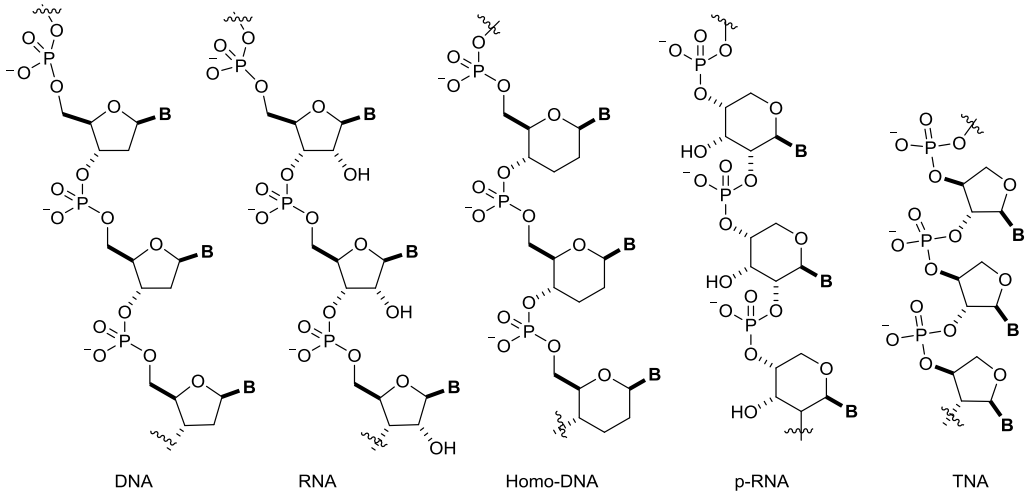


Figure 1.13: *Nature's selection of RNA from a clutter of plausible chemistry. The natural selected chemistry of RNA is shown in the coloured boxes, under each box are the possible routes that RNA could have taken during each individual assembly stages. (Adapted from Joyce)*[35]

1.6.1. Homo-DNA, p-RNA and TNA

Eschenmoser and his colleagues have undertaken a ground-breaking systematic approach to studying the etiology of nucleic acids. They have chemically synthesised a broad spectrum of unnatural nucleic acids that are "in the chemical neighbourhood of RNA" i.e nucleic acids with canonical nucleobases with sugar-phosphate backbones that may have arisen by the same (or similar) prebiological chemical pathways that would have resulted in the ribose phosphate backbone of RNA. They have undertaken a thorough investigation to compare the Watson-Crick base pairing strength, the phosphodiester linkage, chemical properties, non-enzymatic self-replication, and structural and functional properties of different analogues of DNA and RNA.[13, 85, 101-103] Initially Eschenmoser and co-workers investigated the effects of replacing the furanosyl in DNA with a pyranosyl ring, this new pyranosyl nucleic acid was called homo-DNA (Scheme 1.15).

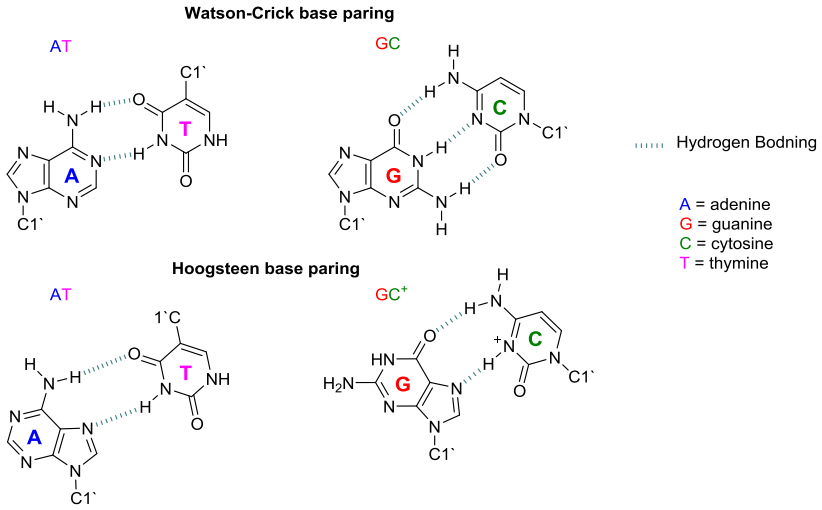


Scheme 1.15: *Comparison of DNA and RNA with Eschenmoser's unnatural nucleic acids (Homo-DNA, p-RNA and TNA).*

Interestingly, homo-DNA displayed increased melting temperatures, and therefore stronger Watson-Crick base pairing, than either DNA or RNA.[104] This increased strength was proposed to be due to the "greater rigidity in pyranose than furanose rings resulting in a preorganization of the strands backbone."[101] and consequently a smaller entropic penalty for duplex formation. These results were extremely interesting, since for the first time it was shown that Watson-Crick pairing was no longer a DNA or RNA phenomenon. Homo-DNA was a strictly model system (model system in the sense that a potential prebiotic synthesis of it seems highly unlikely due to its unusual reduced state of the sugar.) Members of the hexo-pyranosyl family (*β*-allo, *β*-altro and *β*-glucopyranosyl) were also explored. These hydroxylated pyranosyl nucleotides were not observed to form Watson-Crick base pairs with complementary strands, as a result of increased intrastrand steric clashes between additional hydroxyl groups and the nucleobases.

Next, Eschenmoser's group synthesised pyranosyl RNA (p-RNA). This polymer is of particular interest in the series studied as it is clearly most closely related to the natural RNA (Scheme 1.15). The pyranosyl oligonucleotide, p-RNA displayed a much stronger Watson-Crick base pairing than RNA or DNA and showed greater pairing specificity. It pairs exclusively via Watson-Crick pairing whilst completely avoiding Hoogsteen interactions (Scheme 1.16).[103] A whole family of pento-pyranosyl nucleic acids (*β-ribo-*, *β-xylo-*, *β-lyxo-*, *α-xylo-* and *α-arabino*-pyranosyl)[105] were synthesised and all of these showed enhanced Watson-Crick base pairing than RNA. In fact the *α-*

arabino-pyranosyl is the strongest Watson-Crick base pairing system yet to have been encountered.[106] Although unselective cross pairing was observed between all members of the pento-pyranosyl oligonucleotides, none of the p-RNA isomers cross-paired with RNA, and so it is difficult to envisage how sequence information could be passed over to nascent RNA polymer. As these systems do not 'crosstalk' with RNA there is no direct mechanism for genetic transfer to take place. Therefore, it was concluded from these studies that nature did not choose RNA or DNA due to maximal pairing strength, as this would have stopped chemical evolution into RNA genetic system. Rather, it may have chosen RNA as it was the optimal base-pairing system.



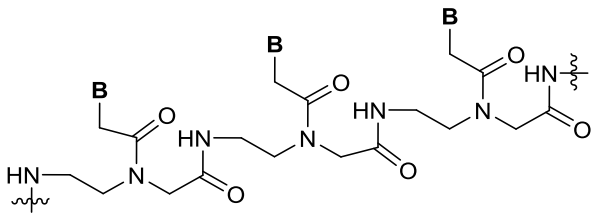
Scheme 1.16: *Comparison of Hoogsteen and Watson-Crick pairing systems.*

Eschenmoser's etiological studies of nucleic acids were not quite finished at this point, and it is a testament to the thoroughness of his group's studies that they found that threose nucleic acid (TNA) displayed extraordinary properties. Threose nucleic acid (TNA) is a nucleic acid with a tetrafuranosyl sugar instead of the pentose sugar in oligonucleotides (Scheme 1.15).[107, 108] Tetrose nucleotide were thought to be "simpler" than pentose nucleotides because the core sugar moiety could result from a $C_2 + C_2$ homo aldol reaction, unlike the five carbon sugar that appear to require crossed aldol reaction, leading to a plethora of other undesired sugars.[70] Furthermore, the formation of phosphodiester linkages will be more selective in TNA due to absence of one hydroxyl group. Remarkably, it was found that TNA can not only efficiently and specifically self-base pair (TNA-TNA), but also capable of cross-pairing with DNA (DNA-TNA) and RNA (RNA-TNA), a phenomenon that had not been observed before with other unnatural oligonucleotides. To unravel chemical diversity of

nucleobases, the Swiss group exchanged adenine with 2,6-diaminopurine in TNA, leading to enhanced pairing.[109] Eschenmoser's mind opening functional studies of oligonucleotides have raised more questions: Why is there a need for a sugar backbone in nucleic acids, can it be negated? Are there even simpler nucleotides capable of genetic transfer and cross-pairing?

1.6.2. Peptide nucleic acids (PNAs)

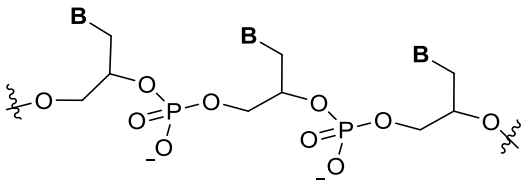
PNA is another important example of a nucleic acid derivative that can Watson-Crick base-pair, but does not have a sugar-phosphate backbone and has found wide use in antisense technologies (Scheme 1.17).[110] It can efficiently Watson-Crick base pair with itself and DNA (duplex and triplex) and has largely been explored for its potential use in medicinal chemistry.[111] It has been suggested, based largely on Urey-Miller experiments, that PNA is more prebiotically plausible than ribose or deoxyribose nucleotides. A number of authors suggest the possibility that PNA was synthesised before RNA, hence there was a "peptide nucleic acid world"[110] prior to the "RNA world". It has already been shown that genetic information can be transferred from PNA to PNA, PNA to RNA (or DNA) and vice versa.[112] However, the synthesis or oligomerisation of PNA monomers, despite nearly 20 years having passed since its design, remains to be demonstrated. Furthermore, it is important to note that monomers of PNA cyclise when they are activated, which will make prebiotic oligomerisation difficult to achieve.



Scheme 1.17: *Structure of peptide nucleic acid (PNA).*

1.6.3. Glycol nucleic acid (GNA)

Another novel nucleic acid is glycol nucleic acid (GNA) (Scheme 1.18).[113] Compared to other potential prebiotic nucleic acids it is structurally simpler (and achiral) but still supporting stable duplex formation. The sugar backbone is at a low oxidation state, so a nucleobase can be stably attached to the backbone without the need for a cyclic acetal linkage, but how this ligation could be achieved prebiotically is difficult to envisage. The GNA oligonucleotide has an acyclic phosphodiester junction, which is capable of Watson-Crick base pairing with itself to form GNA duplex (GNA-GNA). However just like PNA, GNA monomers also cyclise when activated and have only been copied in trimer form.[114]



Scheme 1.18: *Structure of glycol nucleic acid (GNA).*

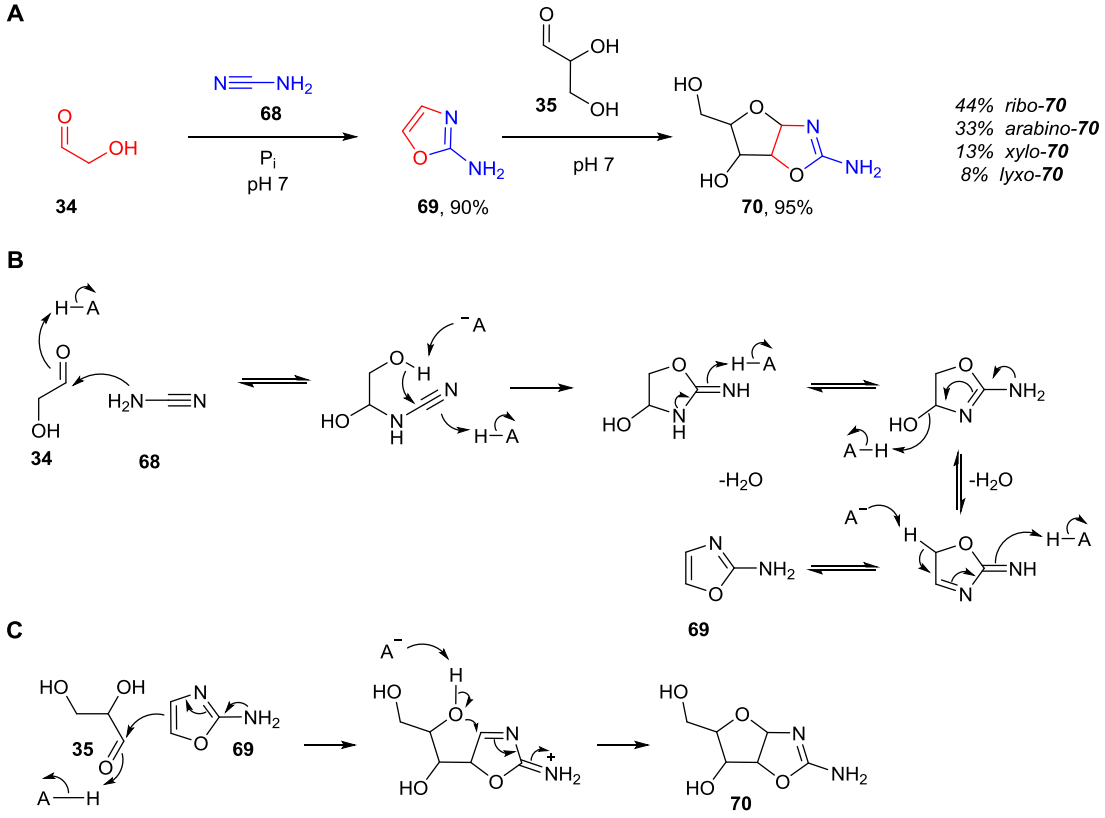
The ethological studies of nucleic acids by Eschenmoser and others does not solve the problem of why nature chose RNA, by finding an alternative to RNA that could perform RNA's role. They all ended up falling back on "evolution" as the only possible answer to how RNA was selected by biology, but provide no real evidence for this transition. There is no clear evolutionary driving force which could explain the mechanism of evolution that leads from a precursor XNA to RNA (if both monomeric unit were not available).[115] Therefore, perhaps RNA was nature choice, because there was no choice (or at least the choice was highly restricted) and that RNA is the simplest molecule to synthesise, oligomerise and replicate under prebiotic constraints.

1.7. Synthesis of Activated Ribonucleotides by avoiding Ribose and Nucleobases

The problem of prebiotic ribosylation of nucleobases has persisted more than 60 years, and in 2004 Leslie Orgel stated that: "*there is at present no convincing, prebiotic total synthesis of any nucleotide. Many individual steps that might have contributed to the formation of nucleotides on the primitive Earth have been demonstrated, but few of the reactions give high yield of products, and those that do tend to produce complex mixtures of products.*"[100] Since then, however there have been significant gains in prebiotic synthesis of ribonucleotides via a completely different methodology: a systems chemistry approach.[83, 116-119] It is upon the system chemistry advance that this thesis will specifically build.

The successful synthesis of activated *β*-cytidine-2`,3`-cyclic phosphate **3** and *β*-uridine-2`,3`-cyclic phosphate **4** nucleotides under plausible prebiotic conditions were initiated by a systems chemistry approach to tackle the problem of prebiotic synthesis of ribonucleotides.[21, 22, 118, 119] The relatively futile conventional method of reacting the unstable and elusive ribose **42**, with pre-formed purines (adenine **22** and guanine **23**) or pyrimidine (cytosine **32** and uracil **33**) and phosphate, is by-passed utilising the prebiotically available C_2 and C_3 sugars glycolaldehyde **34** and glyceraldehyde **35** rather than ribose **42** (Scheme 1.19).[88, 89, 92] The synthesis starts by reacting the C_2 sugar glycolaldehyde **34** with cyanamide **68** at pH 7 in the presence of inorganic phosphate to give 90% 2-amino-oxazole **69.** Oxazole **69** then reacts with the C_3 sugar glyceraldehyde **35** to furnish ribose pentose aminooxazoline **70** in near quantitative yield. The reaction exhibits remarkably diastereoselectivity for *ribo*-**70** and *arabino*-**70**, which make up 74% of the total composition of **70**. The formation of aminooxazoline **70** with complete furanosyl selectivity is a crucial step of the synthesis as the desired pentafuranosyl backbone is made by by-passing free pentose sugars. More importantly, pentose aminooxazoline **70** is synthesised from prebiotically available molecules (glycolaldehyde **34**, glyceraldehyde **35** and cyanamide **68**) and orthophosphate is used throughout the reaction as acid base catalyst. It is also of note that *ribo*-**70** ($t_{1/2}$ > 1 week, pH 10, 55°C) have shown 70 fold

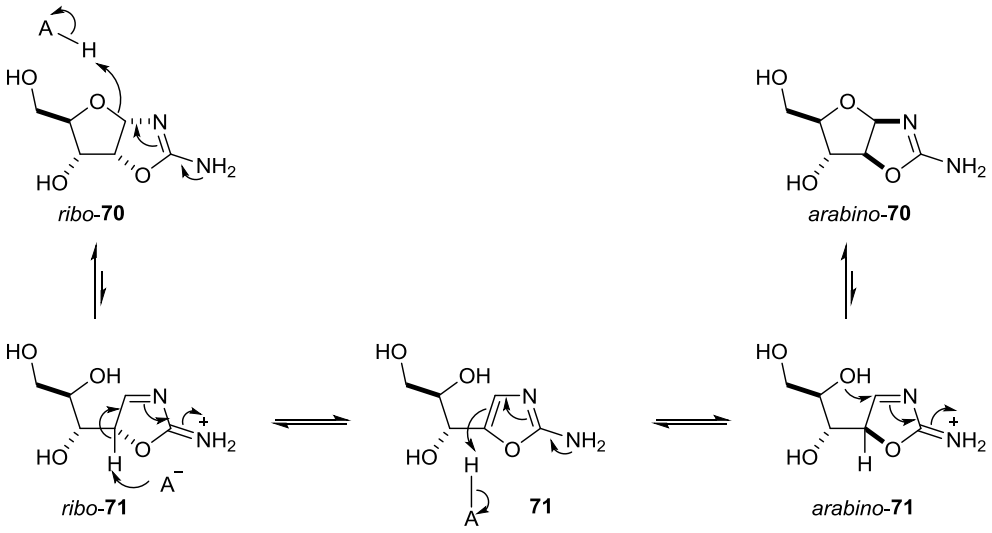
greater stability than ribose **42** ($t_{1/2}$ < 3h, pH 10, 55°C) under comparable aqueous conditions and ribose aminooxazoline *ribo*-**70** has been considered as a stable "organic mineral" that may have been exploited to direct prebiotic chemistry towards biological structures.[71, 120]



Scheme 1.19: *Systems chemistry approach to the synthesis of the pentafuranosyl backbone of ribonucleotides that avoids free pentose sugars.* ***A)*** *The synthesis of pentose aminooxazolines* ***70*** *utilising a combination of mixed nitrogenous and oxygenous chemistry to produces 2-aminooxazole* ***69*** *from the reaction of C_2 sugar glycolaldehyde* ***34*** *and cyanamide* ***68.*** *Oxazole* ***69*** *then reacts with glyceraldehyde* ***35*** *to produce pentose aminooxazolines* ***70*** *with the highest ribo- and arabino-selectivity observed (so far) in prebiotic chemistry.* ***B)*** *Postulated mechanism for acid base catalysed coupling of* ***34*** *with* ***68*** *to furnish* ***69****.* ***C)*** *Mechanism for acid-base catalysed reaction of* ***69*** *with* ***35*** *to give* ***70****.* P_i *= inorganic phosphate; H-A = general acid;* A^- *= general base.*

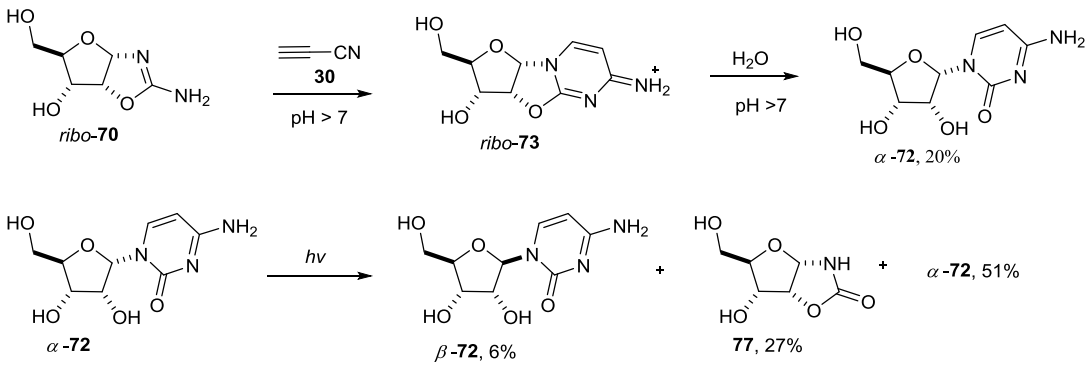
The differential solubility of pentose aminooxazolines **70**, leads to purification of *ribo*-**70** from the crude mixture of all four aminooxazolines via spontaneous crystallisation.[121] Interestingly, it has also been observed that reaction of scalemic glyceraldehyde **34** (60% enantiomeric excess) with 2-aminooxazole **69** can lead to spontaneous crystallisation of pure homochiral D-*ribo*-**70**, which will have greater

implications for the origins of prebiotic homochirality in ribonucleotides.[122] Phosphate is also exploited in the interconversion of *ribo*-**70** to *arabino*-**70**, via a general acid-base catalysed C2` epimerisation (Scheme 1.20). Crystallisation of *ribo*-**70**, leaves behind *xylo*-**70** and *lyxo*-**70** minor products in the crude mixture of **70**. Furthermore *xylo*-**70** and *lyxo*-**70** are then not assessable (requiring C3` epimerisation) from ribo-**70** crystals by phosphate catalysed epimerisation.[123] The interconversion reaction of *ribo*-**70** and *arabino*-**70** can be used to permanently eliminate *xylo*-**70** and *lyxo*-**70** aminooxazolines, due to the privileged crystallinity of *ribo*-**70**, and an acidic C2` proton.



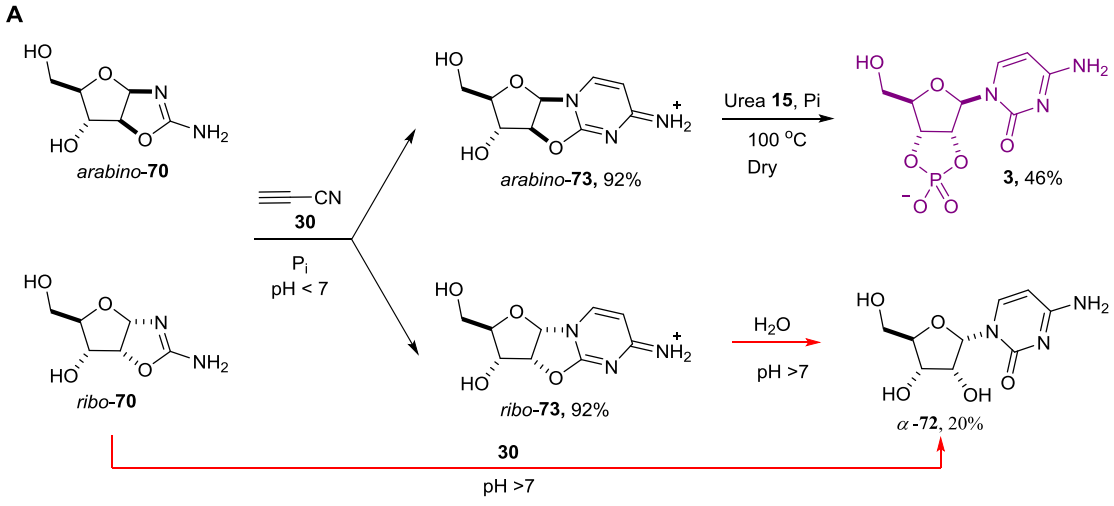
Scheme 1.20: *Plausible mechanism for phosphate catalysed interconversion between ribo-**70** and arabino-**70**. H-A, general acid; A⁻, general base.*

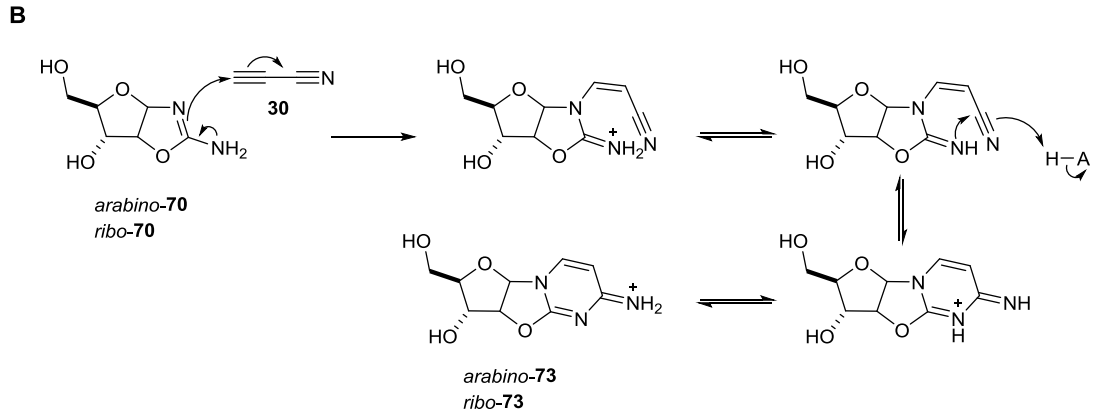
The next step of the synthesis involves formation of the pyrimidine ring on the sugar scaffold via cyanovinylation. The reaction of *ribo*-**70** with excess cyanoacetylene **30** gives 20% *α-ribo*-cytidine *α*-**72**, which can photoisomerise to yield *β*-**72** *albeit*, under the best photoanomerization conditions only giving 6% *β*-**72** (Scheme 1.21).[124, 125] Furthermore, it is observed that a pH increase (>pH 7) during the cyanovinylation reaction leads to the hydrolysis of anhydropyrimidine, *ribo*-**73** into *α*-**72** and then a clutter of compounds, including multiple cyanovinylated products most of which have not been identified, but including *α*-**72**.[22, 124]



Scheme 1.21: *Unbuffered (> pH 7.0) cyanovinylation of aminooxazoline ribo-**70** giving 20% α-cytidine α-**72**, which undergoes poor photoanomerization only yielding 6% β-cytidine β-**72**.*

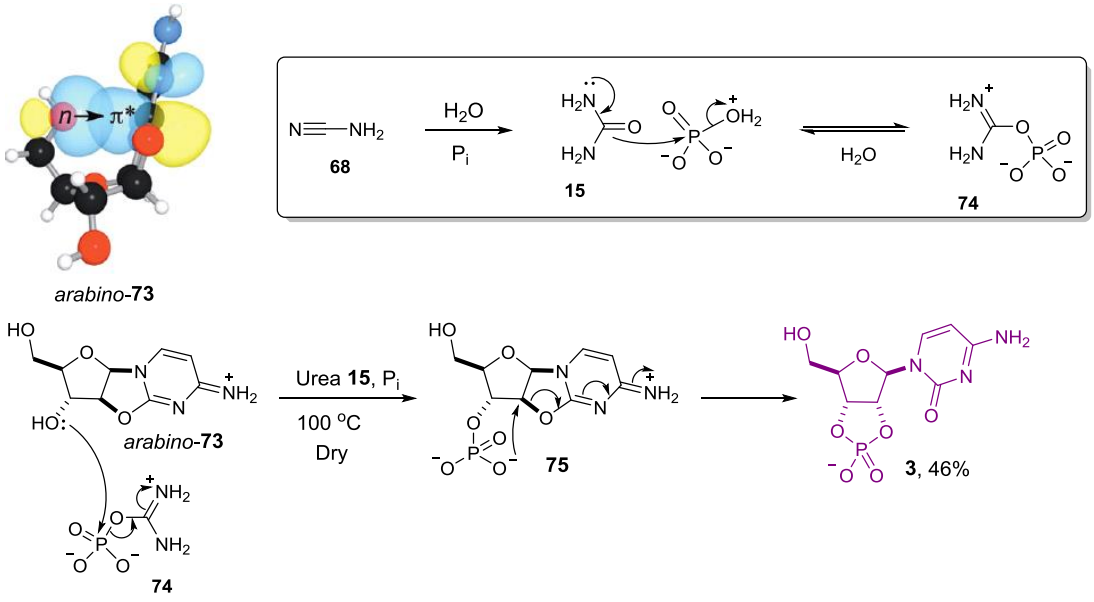
The underlying problem of the unselective chemistry observed between *ribo*-**70** and cyanoacetylene **30** was resolved by employing inorganic phosphate to buffer the system at pH 7. The incorporation of inorganic phosphate in the aqueous reaction leads to remarkable chemoselectivity for the reaction of cyanoacetylene **30** with aminooxazoline *ribo*-**70** and *arabino*-**70** to furnish *ribo*-**73** and *arabino*-**73** in excellent yields (Scheme 1.22). The inorganic phosphate plays a dual role: it is used as an acid base catalyst for example to facilitate the coupling of glycolaldehyde **34** with cyanamide **68** to give 2-aminooxazole **69** (Scheme 1.28) and it also acts as a buffer to maintain the pH at 7 to stop hydrolysis of anhydrocytidine *arabino*-**73** and *ribo*-**73**. Furthermore phosphate is also ultimately then required as a reagent for the final phosphorylation step to form 2`,3`-cyclic phosphates from **73**.





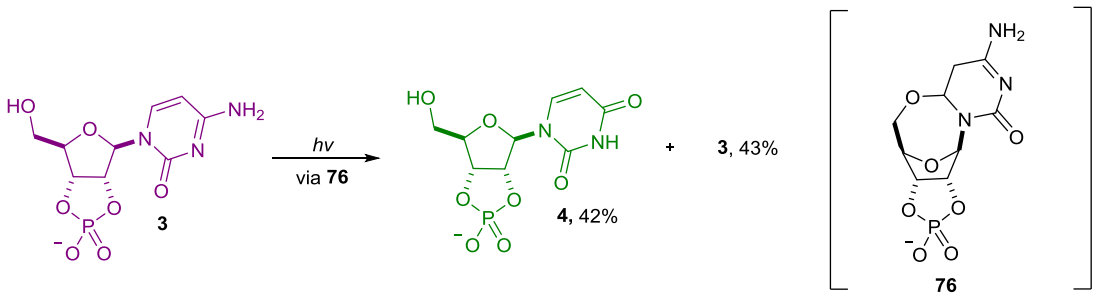
Scheme 1.22: *Conversion of pentose aminooxazoline **70** to anhydrocytidines **73** and their conversion to pyrimidine ribonucleosides **72** and nucleotides **3**. Aminooxazolines, arabino-**70** and ribo-**70**, can be cyanovinylated in up to 92% yield with cyanoacetylene **30** to furnish anhydronucleoside, arabino-**73** and ribo-**73** respectively. Arabino-**73** proceeds further en route to ribonucleotides by being converted to activated pyrimidine **3** via phosphorylation. Destructive chemistry (red): unbuffered reaction of ribo-**73** with cyanoacetylene **30** and/or its hydrolysis yields α-cytidine α-**72. B)** Plausible mechanism for N1 cyanovinylation of aminooxazoline **70** with cyanoacetylene **30**.*

The prebiotic phosphorylation of *arabino*-**70** proceeds via the secondary C3´-OH, instead of the ostensibly less hindered primary C5´-OH (Scheme 1.22). The key to this remarkable chemoselectivity was revealed by X-ray crystallographic studies of *arabino*-**73**, showing an intramolecular stereoelectronic effect that reduces reactivity of the C5´-OH by an n→π* interaction of the non-bonding lone pair of electrons on the C5´ hydroxyl with the antibonding orbital on the C2 carbon atom of the *arabino*-**73**.[126] The interaction of C5´-OH into C2 carbon also provides the needed protection from hydrolysis by blocking phosphate attack on C2 moiety in *arabino*-**73**. Interestingly, urea **15**, the product of phosphate-catalysed hydration of cyanamide **68** plays the crucial role of acting as a solvent in molten form and also activates inorganic phosphate (P_i) by forming the uredo-phosphate transient intermediate **74** (Scheme 32). The C3´-phosphorylated anhydronucleoside **75** then cyclises to form *β*-cytidine-2´,3´-cyclic phosphate **3,** via the phosphates attack at C2´ carbon, breaking the C2´-O-C2 linker, to give the correct sugar and the nucleobase stereochemistry as required for natural nucleotides (Scheme 1.23).



Scheme 1.23: *X-ray structure of arabino-**73** reveals an n→π* interaction between the C5´-OH and C2, which provides selective phosphorylation of C3´-OH in favour of the primary C5´-OH, followed by intramolecular cyclisation of **75** to give 2´,3-cyclic phosphate **3**. Urea **15**, the phosphate-catalysed hydrolysis product of **69**, activates phosphate by forming a phospho-uredo compound **74** (inset). DFT for arabino-**73** reported by Choudhary et al.*[126]

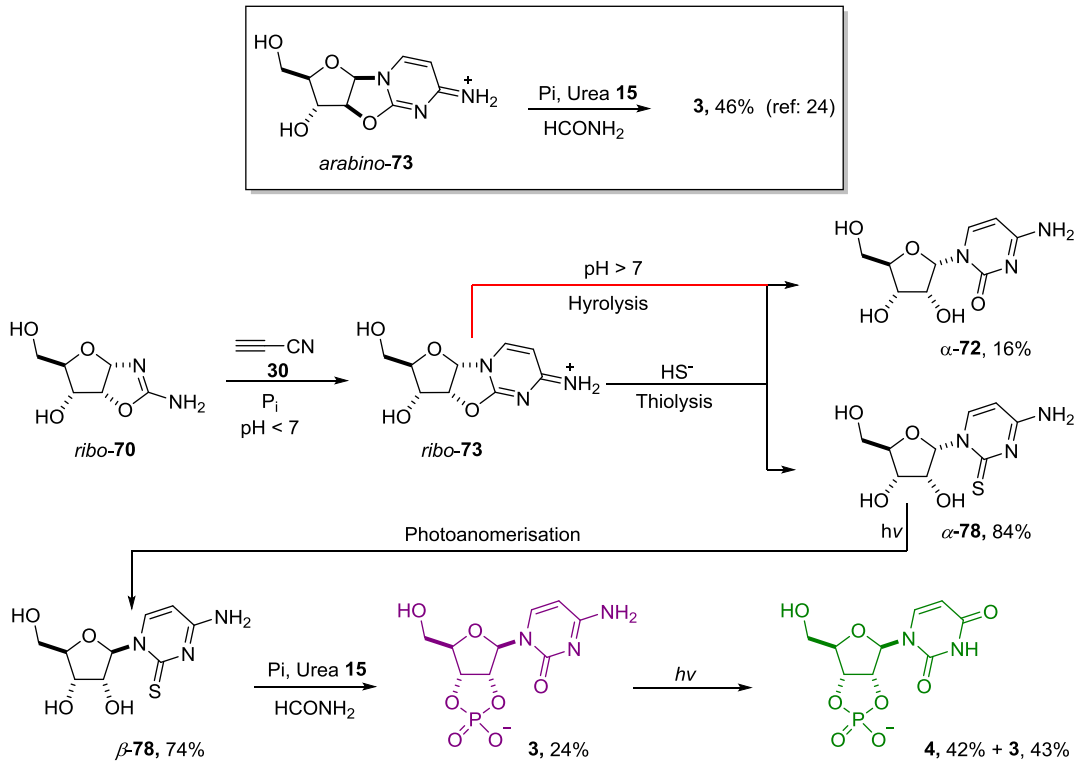
In order to obtain the second pyrimidine nucleotide, *β*-uridine-2`,3`-cyclic phosphate **4**, **3** is irradiated for 3 days in aqueous solution with UV light (254 nm) which gives **4** in good yields with minimal destructive photochemistry (Scheme 1.24). A protective mechanism to conserve the activated pyrimidine ribonucleotides (**3** and **4**) was observed. This is due to the locking the 2`,3`-cyclic phosphate in 'East-West' conformation instead of the more conversionally predominate 'North-South' confirmations (found in RNA). The East-West conformation brings the C5`-OH in close proximity to the C5 carbon atom of the cytosine nucleobase. The C5` OH can undergo the intramolecular cyclisation to form 5,5`-bicyclic structure **76** upon irradiation that protect these specific nucleotide analogues from photochemical degradation. The C5-O5` bond provides pyrimidine nucleotides enhanced stability during photochemistry which destroys the other isomeric by-products and (potentially) permits the activated pyrimidines (**3** and **4**) to move on to oligomerisation phase and can be photochemically purified from other similar isomers.



Scheme 1.24: *Photochemical hydrolysis of β-2´,3´-cyclic phosphate cytidine **3** to β-2´,3´-cyclic phosphate uridine **4**. The cyclic phosphate locks the sugar in West-East confirmation (not North-South), bringing close the C5´-OH of the sugar in close proximity to C5 atom of the pyrimidine nucleobases (cytosine, uracil) to from a bicyclic structure **76**. The photoirradiation is also thought to proceed via **76** which provides protective mechanism for **3** and **4**, whilst destroying other deleterious chemicals in the crude mixture.*

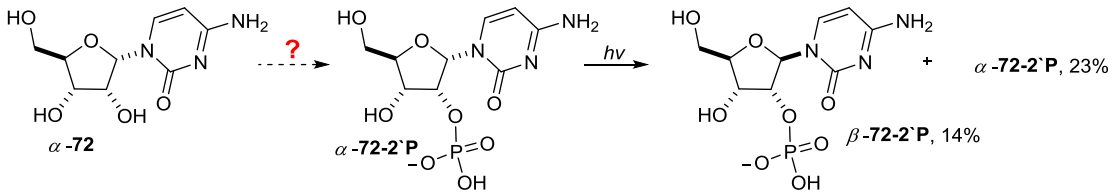
Recently, a second route to the synthesis of **3** and **4** has been published by the Sutherland group providing further proof for the intermediacy of pentose aminooxazolines **70** *en rout*e to ribonucleotides, and perhaps further validating the systems chemistry approach in ribonucleotide synthesis. This pathway proceeds via thioribocytidines.[21] In their new synthesis, Sutherland and co-workers found a solution to the low yield of photoanomerization observed in *α*-**72** by positioning a sulfur atom at the C2 position of the cytidine nucleobase, following the cyanovinylation and thiolysis of *ribo*-**73** (Scheme 1.25). This sulfur promoted remarkable photoanomersation, therefore demonstrated that both *ribo*-**73** and *arabino*-**73** can be used for selective ribonucleotide synthesis allowing even greater yields of **3** and **4** *en route* to ribonucleotides (via **73**).[21, 22] Their reaction is a continuation of their cyanosulfidic protometabolism studies which used photoreductive hydrogen sulfide to find a network prebiotic synthesis of C_2, C_3 sugars and precursors for proteinogenic amino acids and lipids.[88, 89, 92] The reaction of *ribo*-**73** with hydrosulfide (HS^-) yields the corresponding *α*-thioribocytidine *α*-**78** (84%) and *α*-ribocytidine *α*-**72** (16%). However, under photoirradiations conditions, *α*-**72** is degraded whilst *α*-**78** undergo remarkable *α*→*β* photoanomerization to furnish *β*-**78** in excellent yields. The final step of the synthesis involves phosphorylation to get the activated pyrimidines, *β*-**78** is dry phosphorylated in the presence of urea to furnish **3** in good yields, which can then be irradiated under the same photochemical conditions

as the first synthesis to obtain **4** by hydrolysis/deaminiation of the cytosine nucleobase.



Scheme 1.25: *Synthesis of β-ribopyrimidines via thiolysis and photoanomerization reactions. Thiolysis of ribo-**73** yields α-thioribocytidine α-**78** by hydrosulfide. Irradiation of α-**78** furnishes 74% β-**78** via α→β photoanomerization, whilst α-**72** is eliminated from the mixture by undergoing destructive photochemistry. Urea-mediated phosphorylation converts β-**78** to β-2`,3`-cyclic phosphate cytidine **3**. Destructive chemistry of ribo-**73** in unbuffered solution leads to α-**72**, which is relatively impervious to C1`-stereoinversion chemistry and is mostly destroyed in the mixture to give oxazolidinone **77** (red). As shown earlier arabino-**72** also furnishes **3** (46%) via chemical inversion induced by cyclo-phosphorylation (inset).*

There had been limited success to convert α-**72** to β-**72** via photoanomerization, for this C1´-stereoinversion to occur α-**72** had to undergo C2´-phosphorylation to form α-cytidine-2´-phosphate α-**72-2´P** , which is envision to be difficult and unselective (Scheme 1.26).[127] Hence, the efficient (74%) photoanomerization in α-**78** due to the presence of sulfur at C2 carbon atom of cytidine nucleobase, removes the requirement of the unnatural C2´-OH phosphorylation to control photoanomerization in pyrimidine ribonucleosides and opens up a new chapter for the role of sulfur in prebiotic stereochemistry.



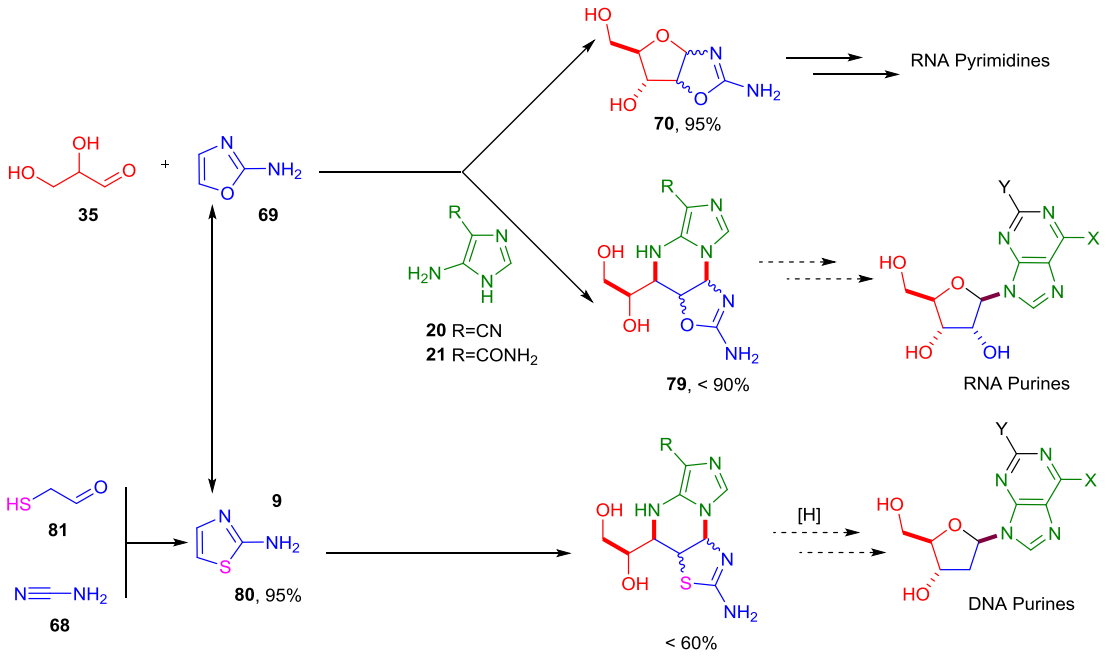
Scheme 1.26: *C1´ photoanomerization of α-cytidine-2´-phosphate α-**72-2´P** into β-cytidine-2´-phosphate β-**72-2´P**. The nucleotide α-**72-2´P** was prepared employing conventional synthetic chemistry as a plausible prebiotic route to selectively synthesis α-**72-2´P** from α-**72**is not known.*

These two (potentially consecutive) synthesis of canonical pyrimidines via *arabino*-**73** and *ribo*-**73**, have undoubtedly led to a radically different strategy of tackling the prebiotic synthesis of ribonucleotides, however its compatibility with the synthesis of the remaining purine ribonucleotides still needs to be demonstrated.

1.7.1 Multicomponent Prebiotic Chemistry

In an attempt to find a congruent pathway for the synthesis of both purines and pyrimidines, preliminary studies have shown promising results based on the current prebiotic synthesis of activated pyrimidines.[128] The three component reaction of 2-aminooxazole **69**, glyceraldehyde **35** and HCN cyclic oligomers (AICA **21** or AICN **20**) have furnished the purine ribonucleotide precursor **79** in excellent yields (Scheme 1.27). Interestingly this multicomponent chemistry exhibits divergent pH dependence, which can be fine-tuned to control purine or pyrimidine selectivity. The mixed, three component reaction at pH 7 predominantly gives two component product aminooxazoline **70** (pyrimidines precursor), whist the purine precursor **79** is the prevailing product between pH 4 and 5, whereas a mixture of purine and pyrimidine precursors is observed between pH 5 and 6.5. Although this multicomponent reaction proceeds with absolute regiospecificity and high diastereoselectivity, unfortunately **79** has yet to be fully elaborated into purine ribonucleotides, therefore a one pot prebiotic synthesis of purine and pyrimidine is still required. Interestingly, exchange of 2-aminooxzole **69** with 2-aminothiazole **80** (synthesised prebiotically from **81** and **68**) could potentially lead to the synthesis of DNA nucleotides, by C2` carbon atom

desulfurization in the final stages.[129] These innovative studies clearly shows the inherent association of HCN oligomers with native purine nucleotide synthesis. However more investigations will be required to demonstrate the completion of a purine synthesis by this strategy.

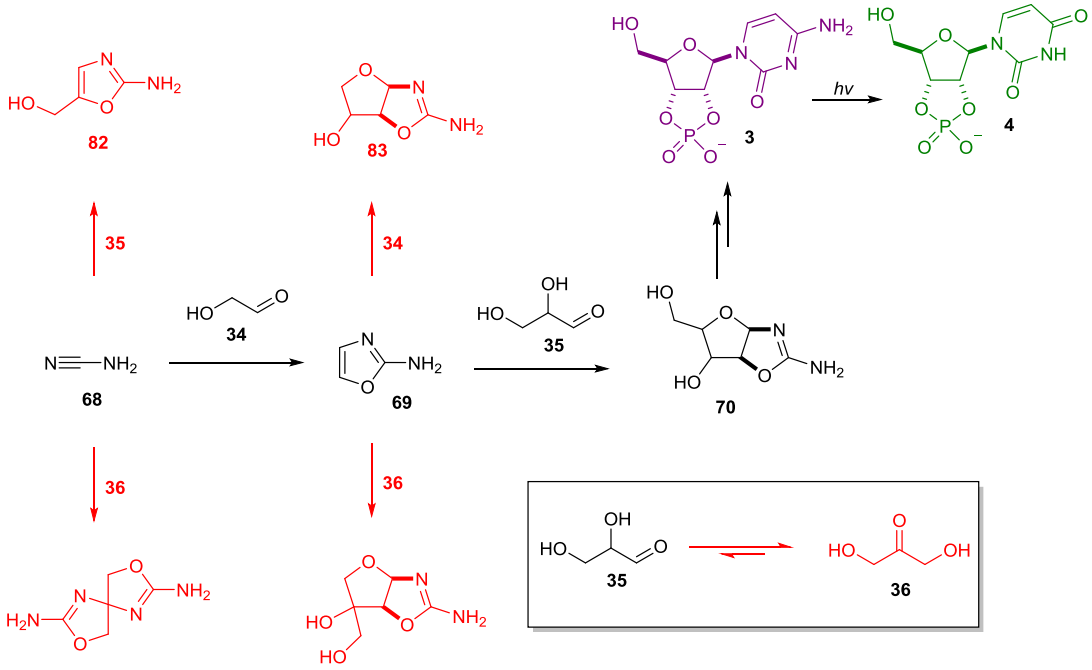


Scheme 1.27: *Attempted one pot synthesis of purine and pyrimidines ribonucleotides from a mixture of glyceraldehyde* ***35****, 2-aminooxazole* ***69*** *and AICA* ***21*** *or AICN* ***20*** *compounds. The purine precursor* ***79*** *is obtained in excellent yields, however it has failed to give the canonical purine nucleotides so far. A plausible prebiotic link between DNA and RNA can also be established by exchanging* ***69*** *with 2-aminithiozole* ***80****, which can be prebiotically synthesised from the reaction of β-mercapto acetaldehyde* ***81*** *and cyanamide* ***68.*** *Solid arrows, reported chemistry; dashed arrows proposed chemistry.*

1.7.2. The problem of sequential addition and controlled chemistry in ribonucleotide synthesis

The two concurrent synthesis of activated pyrimidines **3** and **4** (Scheme 1.25) are the most complete models of prebiotic pyrimidines ribonucleotides synthesis. However, this synthetic route has received some criticism mostly due to the use of pure glycolaldehyde **34** and glyceraldehyde **35** and the sequential addition of the components to control chemistry.[115, 130-132] The synthesis of pyrimidine precursor,

aminooxazoline **70** is reliant on the reaction of 2-aminooxazole **69** (synthesised from phosphate catalysed reaction of **34** and cyanamide **68**).



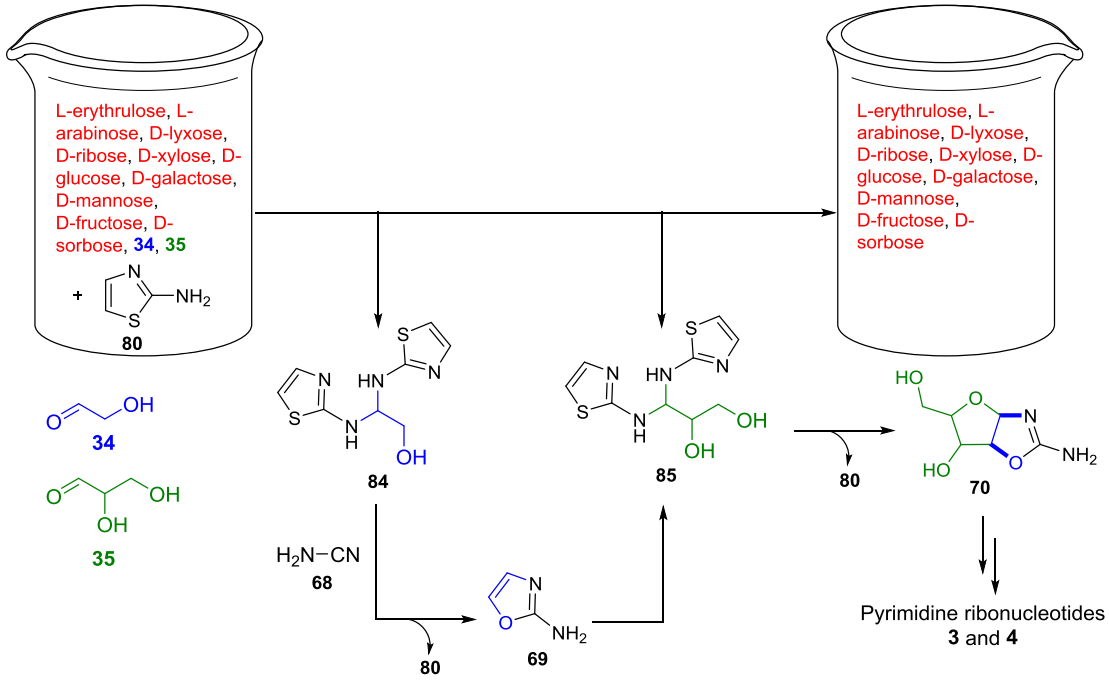
Scheme 1.28: *The dependence of pyrimidine synthesis in the controlled, sequential addition of C_2 and C_3 sugars in the system (black). In the absence of the ordered assembly of sugars (addition of **34** followed by **35**, each equimolar), the synthesis gives rise to a series of undesired products, reducing the overall yield of the incremental synthesis. To add to the increasing number of problems, glyceraldehyde **35** readily isomerises to dihydroxyacetone **36** in the presence of phosphate (inset). Adapted from Islam and Powner.*[83]

However formation of **69** in the presence of glyceraldehyde **35** gives the unwanted aminooxazole **82**. Also, the reaction of **69** in the presence of glycolaldehyde **34** leads to the formation of tetrose and erythrose aminooxazolines **83**. The co-existence of dihydroxyacetone **36**, the predominant triose isomer at equilibrium, can cause further problems by reacting with cyanamide **68** and 2-aminooxazole **69** to furnish undesired aminooxazolines in the system (Scheme 1.28).[133]

It is thought that the chemical clutter in this mixture could be decreased by extracting **69** from the mixture via sublimation and then rained into a separate pool containing glyceraldehyde **35**. However this method still exposes **69** to **34**, **35**, dihydroxyacetone **36** and the prospect of finding an exclusive pool of glyceraldehyde **35** is minimal due to its synthetic reliance on glycolaldehyde **34** as a precursor.[88, 89, 92] Furthermore the outlook of excluding phosphate from the mixture to eliminate dihydroxyacetone **36**

cannot be considered, as the pyrimidine synthesis requires phosphate throughout the system as a catalyst, buffer and reagent.[70, 100]

Islam *et al.* have found a valid and robust solution for all the questions raised, augmenting the credibility of the plausible prebiotic synthesis of pyrimidines.[134] The synthesis employs 2-aminothiazole **80**, which selectively forms aminal (**84** and **85**) with glycolaldehyde **34** and glyceraldehyde **35** in the presence of complex mixture of sugars (aldoses and ketose), including dihydroxyacetone **36** (Scheme 1.29).



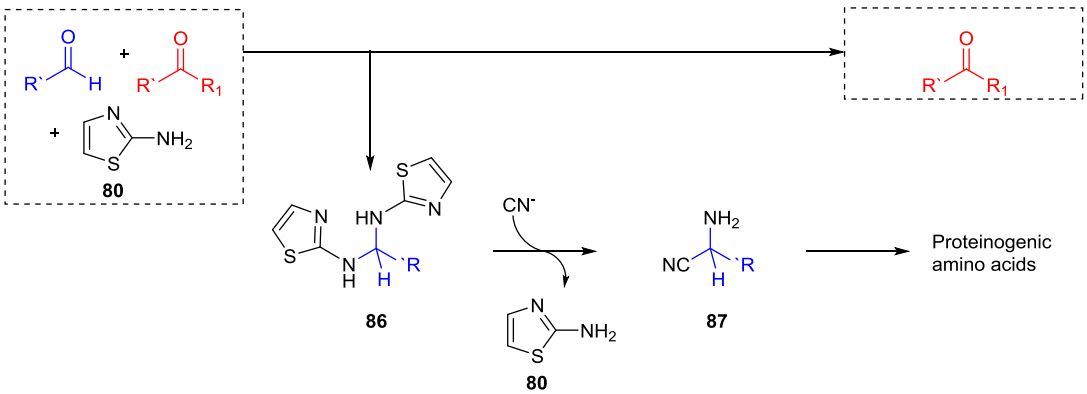
Scheme 1.29: *Prebiotic selection and assembly of natural nucleotides from complex mixtures. Aminal-induced selection of natural nucleotide precursors from a highly complex C_2, C_3, C_4, C_5, and C_6 aldose and ketose sugar mixture via **80**-induced crystallisation.*

The differential solubility of the C_2 and C_3 aminals was observed to lead to the spontaneous crystallisation of **84** from even complex mixture of sugars, and the mixture was completely devoid of glycolaldehyde **34**. Sequestration of the glyceraldehyde **35** is slower, because it exists as a minor component in equilibrium with the thermodynamically more favoured ketose isomer dihydroxyacetone **36**, which allows enough time for **84** to sequestered form the mixture, but also by physically separated from aminal **85**. Remarkably, even after prolonged incubation

none of the C_4, C_5 or C_6 aldoes/ketoses sugars crystallises from the mother liquor, therefore 2-aminithiazole **80** is completely selective to C_2 and C_3 aminal precipitation.

To solve the problem of dihydroxyacetone **36** in the mixture, **36** is incubated with **80**, and amazingly only the C_3 aminal **85** is formed instead of the dihydroxyacetone **36** aminal despite it being the predominant C_3 sugar at equilibrium. 2-Aminothiazole only sequesters aminals of aldehydes, and not ketones (likely due to the steric constrain of positioning two aminothiazoles at a quaternary centre). After selective sequestration of the aminals from the complex mixture, **84** reacts with cyanamide **68** in phosphate buffer to give 2-aminooxazole **69**. The canonical pyrimidine precursor, **70** is obtained by the reaction of C_3 aminal **85** with **69** solving the inherent thermodynamic isomerisation problem of dihydroxyacetone **36**.

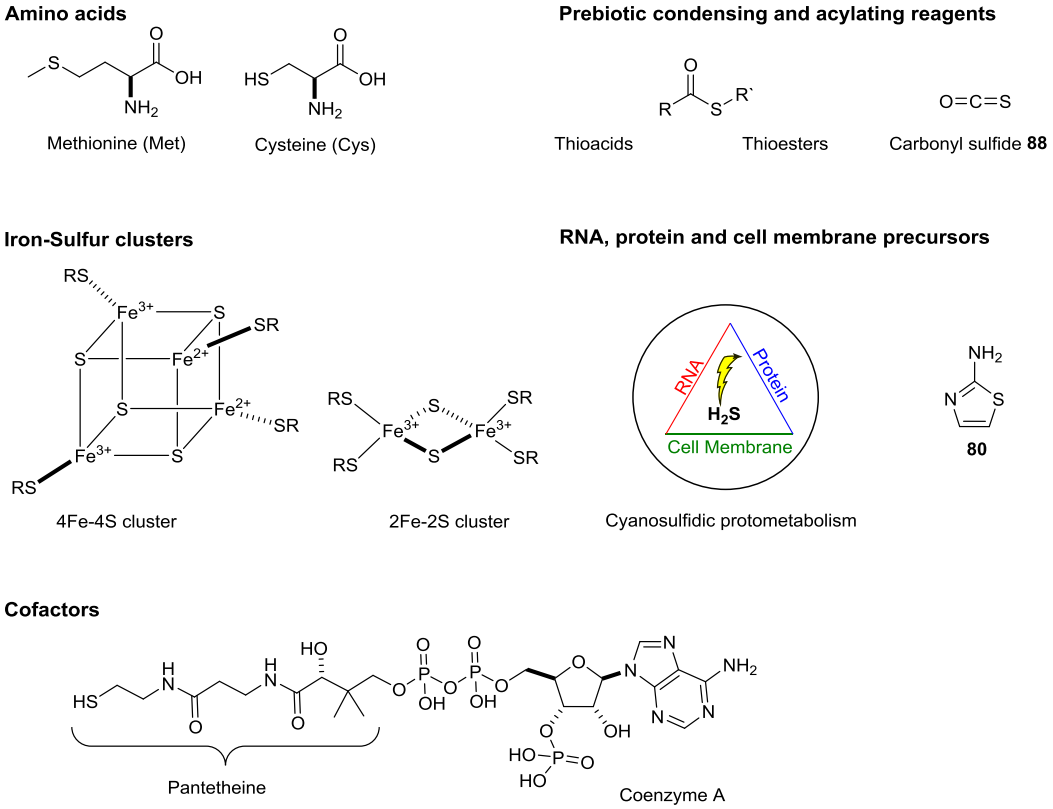
Given the importance of the cyanosulfidic protometabolism pointing towards a systems chemistry associated with nucleotides, proteins and lipids, the physiochemical role of **80** in the synthesis of proteinogenic amino acids has also been investigated.[89] Interestingly, **80**-induced crystallization also provides an absolute chemical selection for the Strecker aldehydes of the natural amino acids from a complex mixture (Scheme 1.30). The ability of **80** to form aminals with C_2 and C_3 sugars, and the reaction of these aminals leading to the synthesis of aminooxazoline **70** is an extremely powerful tool, providing sequential controlled chemistry for the synthesis of pyrimidine nucleotides.



Scheme 1.30: *Prebiotic selection and assembly of proteinogenic amino acids from complex mixtures. Aminal-induced crystallization of Strecker aldehydes (blue) from a complex mixture containing closely related prebiotic ketones (red). The Strecker aldehyde-aminal* ***86****, then reacts with cyanide to form α-aminonitriles* ***87****.*

1.8. Sulfur in Prebiotic Chemistry

Sulfur is an essential element for life. Among the many organic building blocks of biochemical structures, sulfur is found in two of the proteinogenic amino acids, methionine (Met) and Cysteine (Cys). Sulfur is also incorporated in many cofactors including coenzyme A, which is used in the synthesis and oxidation of fatty acids and in the oxidation of pyruvate in the citric acid cycle. Interestingly, the pantetheine moiety of coenzyme A (CoA), has also been synthesized under plausible prebiotic conditions by Miller and Keefe, albeit in low yield.[135] Sulfur also form clusters with iron, e.g. ferrodoxins (iron-sulfur proteins) which play an important role in the oxidation-reduction of mitochondrial electron transport.[136] It is also suggested that iron-sulfur clusters might have played an important prebiotic role in the chemistry to biology transition phase.[137, 138] Recently, Bonfio *et al.* have successfully shown a photochemical prebiotic synthesis of iron-sulfur clusters.[139] Clusters, [2Fe-2S] and [4Fe-4S] have been successfully synthesised by the photooxidation of ferrous ions and the photolysis of organic thiols. These clusters have been stabilised by a wide range of cysteine containing peptides in model protocells. Sulfur adducts have a crucial role in the non-enzymatic polymerisation of amino acids. Thioesters have been extensively used in the synthesis and ligation of peptides.[140, 141] Recently chemoselective acetylation of ribonucleotides with thioacetate have been used in the prebiotically plausible synthesis of oligoribonucleotide, enriched in 5`-3` phosphodiester linkages.[142] Thioacids have also been used in amide bond formation and lipid synthesis.[143, 144] Furthermore, Orgel *et al.* have used carbonyl sulfide **88**, a simple volcanic gas as a prebiotic condensing agent in the synthesis of peptides from amino acids in mild aqueous conditions.[145] The cyanosulfidic protometabolism (Section 1.5.2), in which glycolaldehyde **34**, glyceraldehyde **35**, 12 proteinogenic amino acids and glycerol **63** (phospholipid precursor) has been prebiotically synthesised, further point to the importance of sulfur in prebiotic chemistry.[89] This network of systems chemistry relies on the prebiotic photochemical reduction with hydrogen sulfide (H_2S). Lastly, Islam *et al.* (Section 1.7.2) have demonstrated that 2-thiooxazole **88** can be used in the sequestration and delivery of canonical ribonucleotides and proteinogenic amino acids.[134] The importance of sulfur in prebiotic chemistry should not be underestimated and clearly warrants further future investigation (Scheme 1.31).

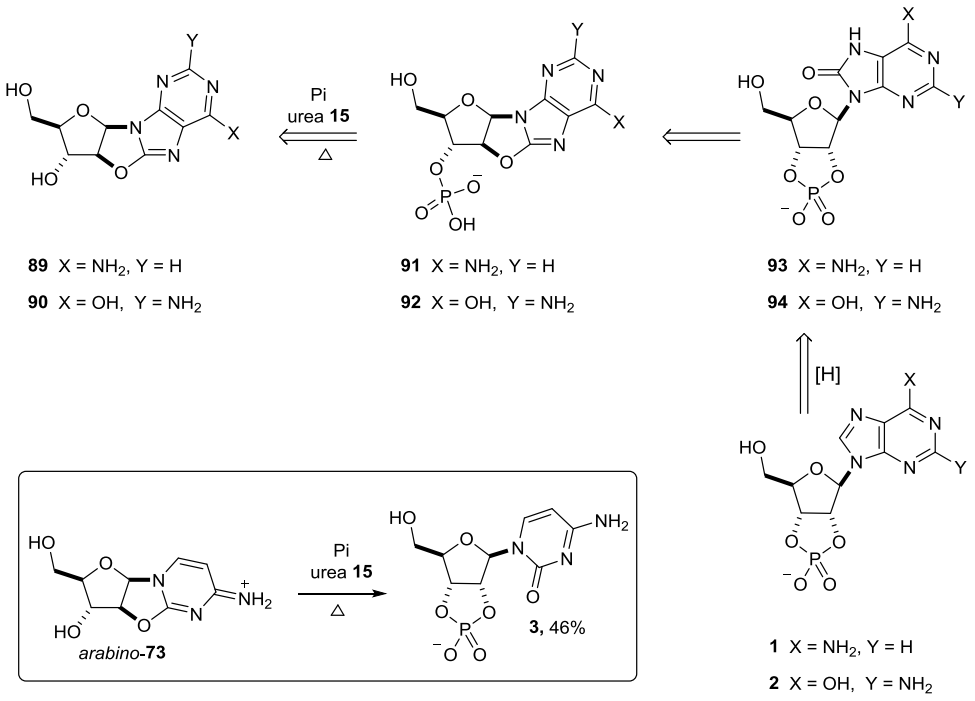


Scheme 1.31: *Prebiotic application of sulfur in different (bio)chemical systems. Cofactors work together with enzymes to achieve optimum biological functioning. Amino acids are the building blocks of proteins. Iron sulfur clusters act as electron shuffles in living cells and are considered to be ancient providers of metabolism for early life. Thioacids and thioesters have been employed in the prebiotic polymerisation of protein and oligoribonucleotide. Sutherland's cyanosulfidic protometabolism utilised photochemical hydrogen sulfide (H_2S) to synthesise glycolaldehyde **34**, glyceraldehyde **35**, precursors for 12 amino acids and phospholipids. 2-Aminothioozole **80** has been used in the purification and delivery of ribonucleotide and proteinogenic amino acids by Islam et al.*

1.9. Aim of the project

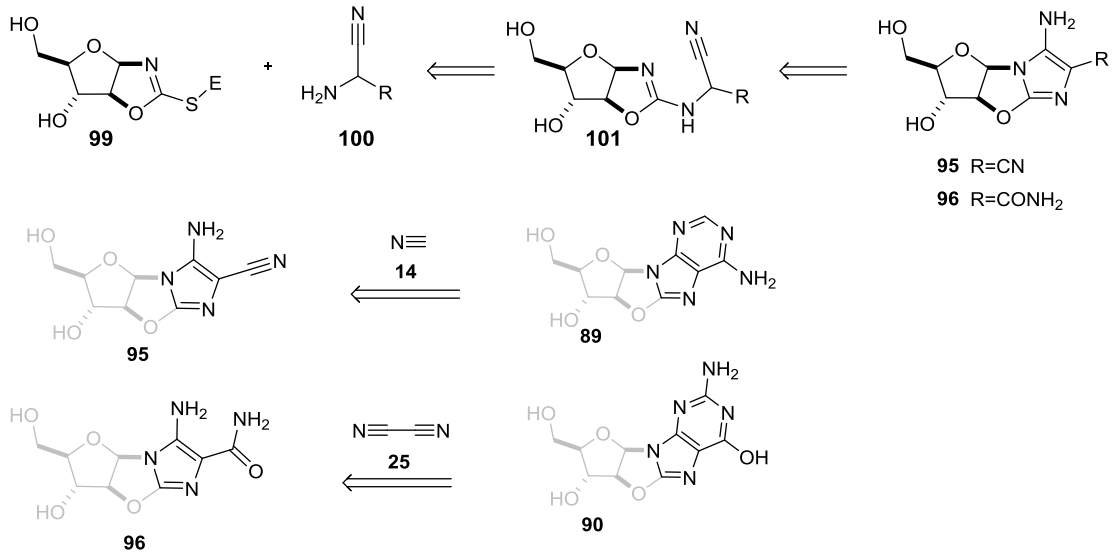
Whilst the RNA world theory provides a plausible solution to the central dogma of molecular biology (Section 1.2.2), there remain many other challenges to overcome in order to explain spontaneous emergence of life.[115] The core aim of this thesis is to develop a new methodology and new chemistry to establish a chemo- and diastereoselective synthesis of β-2`,3`-cyclic phosphates (**1** and **2**). We specifically

seek to establish a route that is compatible with pyrimidine synthesis (as described in Section 1.7) by developing a pathway with parity to pyrimidine synthesis via 2,2`-anhydrocytidine *arabino*-**73**. We anticipated that an anhydronucleoside ligation between the (purine) nucleobase and the sugar would once again be a key element in the synthesis, and could be accessed via a structure similar to *arabino*-**73**. Specifically we envisaged opening up the 8,2`-anhydro-*arabino*-cycloadenine **89** and 8,2`-anhydro-*arabino*-cycloguanine **90** via C3`-OH phosphorylation (**91** and **92**) to yield *β*-8-oxo-2`,3`-cyclic phosphate adenine **93** and *β*-8-oxo-2`,3`-cycliic phosphate guanine **94** nucleotides respectively (Scheme 1.32). With direct analogy to 2,2`-anhydro-*arabino*-cytidine *arabino*-**73**, **89** and **90** have *iso*-amide moiety embedded in the anhydronucleotide linkage that we expect could be exploited to induce intramolecular rearrangement to furnish the correct sugar structure of nucleotides by C2` inversion.[146]



Scheme 1.32: *Retrosynthetic analysis for urea-mediated phosphorylation of* ***89*** *and* ***90*** *to yield* ***93*** *and* ***94*** *respectively (via* ***91*** *and* ***92****) and its subsequent prebiotic reduction furnishing β-2´,3´-cyclic phosphates* ***1*** *and* ***2*** *respectively. Reported phosphorylation and chemical inversion of 2,2´-anhydrocytidine arabino-****73*** *furnishing 46% β-2´.3´-cyclic phosphate adenine 3 (inset).*

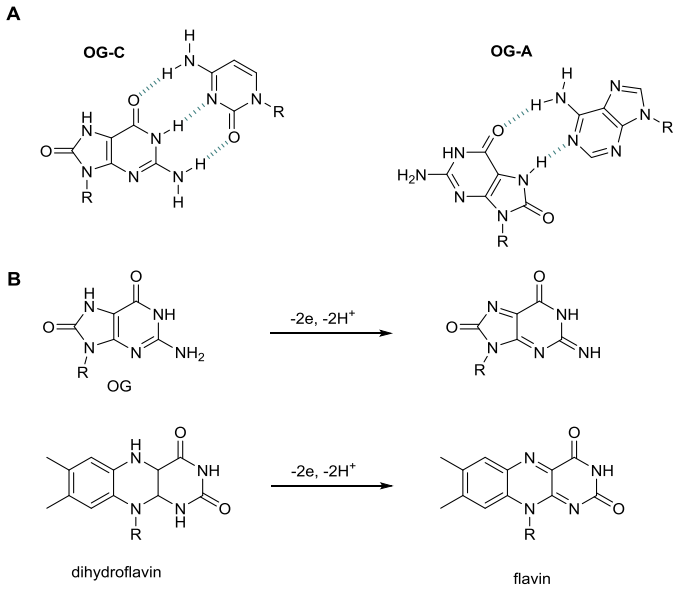
Developing this idea further we then anticipated the established relationship between HCN **14** and purine nucleobases,[47] would be exploited to yield amino nitrile nucleophiles **100** capable of synthesising the core structure of purines **89** and **90** (Scheme 1.33) upon nucleophilic addition to an "activated" oxazoline scaffold. We suspect that position sulfur at C2 would lead to facile activation. We anticipated that weak sulfur-carbon orbital overlap and the precedence for electrophilic thioacids and thioamide in prebiotic chemistry could be exploited during the nucleophilic attack of **100** at the C2 carbon atom of an oxazoline scaffold (e.g. **99**)



Scheme 1.33: *Retrosynthetic analysis for furnishing 8,2´-anhydro-arabino-cyclonucleosides **89** and **90** via cyanation with HCN **14** and cyanogen **25** of **95** and **96** respectively.*

We hypothesised that the imidazole moiety **95** and **96** could be built by nucleophilic addition of an amino nitrile **100** to the electrophilically activated C2 carbon atom of a pentose oxazolidinone **99** to give aminooxazoline **101**, following cyclisation to give imidazoles **95** and **96**. Subsequent electrophilic annulation of pyrimidine moiety of the purines **95** and **96** by HCN **14** and cyanogen **25** (known precursors of adenine and guanine nucleobases) respectively would then be investigated to furnish 8,2`-anhydro-arabino-cyclonucleosides **89** and **90**. We recognised that our strategy would involve C2` inversion of **89** and **90** to yield 8-oxo-nucleotides **93** and **94**, rather than the canonical purines **1** and **2**. It is possible that subsequent reduction at the C8 position

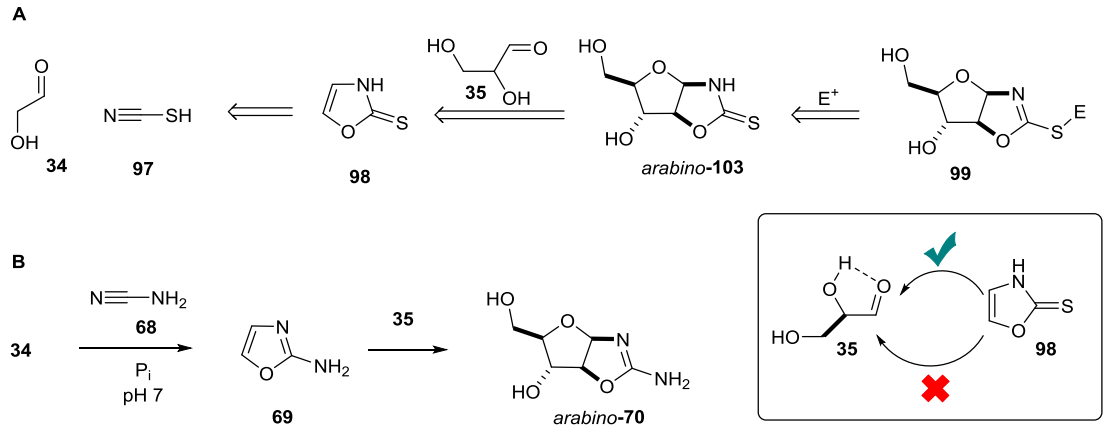
could furnish the canonical purine moiety, and such reduction would certainly merit investigation. However, it is also of note that C-8 oxidation in purines is well known and 8-oxo-guanine (OG) is reported to be the main cause of mutation in DNA due to G-T and C-A transversion via DNA Hoogsteen pairing.[147, 148] Furthermore, there are numerous reports demonstrating genetic information transfer via OG-A and OG-C base pairing systems (Scheme 1.34A).[149-154] Interestingly, the prebiotic role of 8-oxyguanine (OG), as a flavin mimic, has also been investigated, in which OG is used to photorepair pyrimidine bases (Scheme 1.34B). It is also important to note that RNA is more susceptible to C8-oxidation than DNA, as RNA is more distributed in the cell and is found as often as single strand form.[155] Therefore, we hypothesis that 8-oxo purines may have played a vital role in the initial stages of RNA World and further studies to investigate its prebiotic function leading to the modern Watson-Crick base pairing in biology is required.



Scheme 1.34: A) *Showing stable Watson-Crick base pairing in OG-C and OG-A.* ***B)*** *Prebiotic role of 8-oxo-guanine (OG) as a flavin mimic to photorepair damaged bases.*

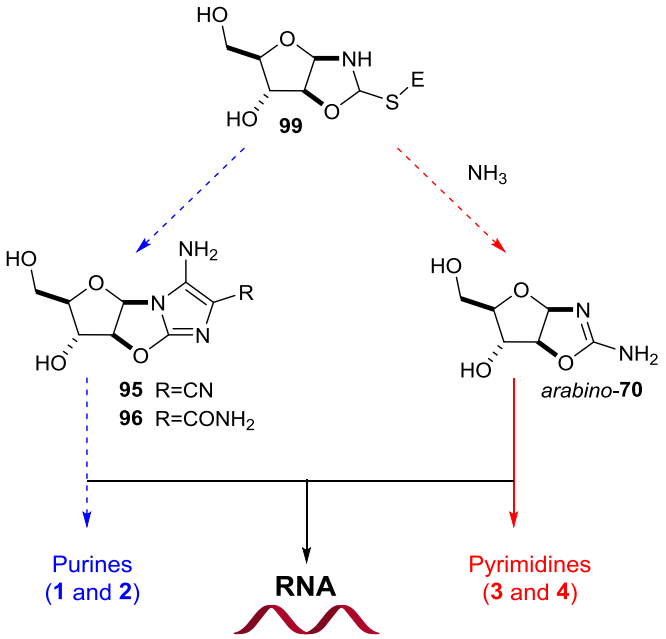
We propose to investigate the hypothesis that oxazolidinone thione **103** could have been a key intermediate *en route* to the formation of purine ribonucleotides, and that by positioning sulfur at the C2 position would give access to a species that could be selectively (electrophilically) activated at this carbon (where nucleobase elaboration

must occur). Here we anticipated that sulfur would allow the nucleophilic elaboration of the purine bases, rather than the previously described electrophilic elaboration of the pyrimidine structure (by cyanoacetylene **30**) upon pentose aminooxazoline *arabino*-**70**. We envisage a stereo-selective synthesis of *ribo*/*arabino* **103** selectivity upon the masked aldol reaction of glyceraldehydes **35** and 2-thiooxazole **98** (Scheme 1.35A), assuming that this selectivity would be controlled by internal H-bonding in glyceraldehydes **35** by direct analogy to the reaction of **35** with 2-aminooxazole **69** (Scheme 1.35B). The C2 carbon atom of oxazolidinone thione **103** would then be activated electrophilically to facilitate nucleophilic construction of the purine nucleobase moiety.



Scheme 1.35: A) *Proposed retrosynthetic analysis for prebiotic formation of oxazolidinone thione* ***103*** *and its subsequent C2 electrophilic activation to form activated thione* ***99****. Proposed hydrogen-bonding in glyceraldehydes* ***35*** *that controls stereochemical addition of oxazole nucleophiles (inset).* ***B)*** *Reported synthesis of aminooxazoline arabino-****70*** *from the reaction of 2-aminooxazole* ***69*** *with glyceraldehyde* ***35*** *is with direct analogy to our proposed reaction of 2-thioooxazole* ***98*** *with* ***35****.*

And finally to establish a concomitant purine and pyrimidine ribonucleotide prebiotic route leading to the synthesis of RNA, we envisaged the addition of ammonia (NH_3) to activated thione **99** may also yield aminooxazoline *arabino*-**70**, the key precursor of canonical pyrimidine ribonucleotides (Scheme 1.36).

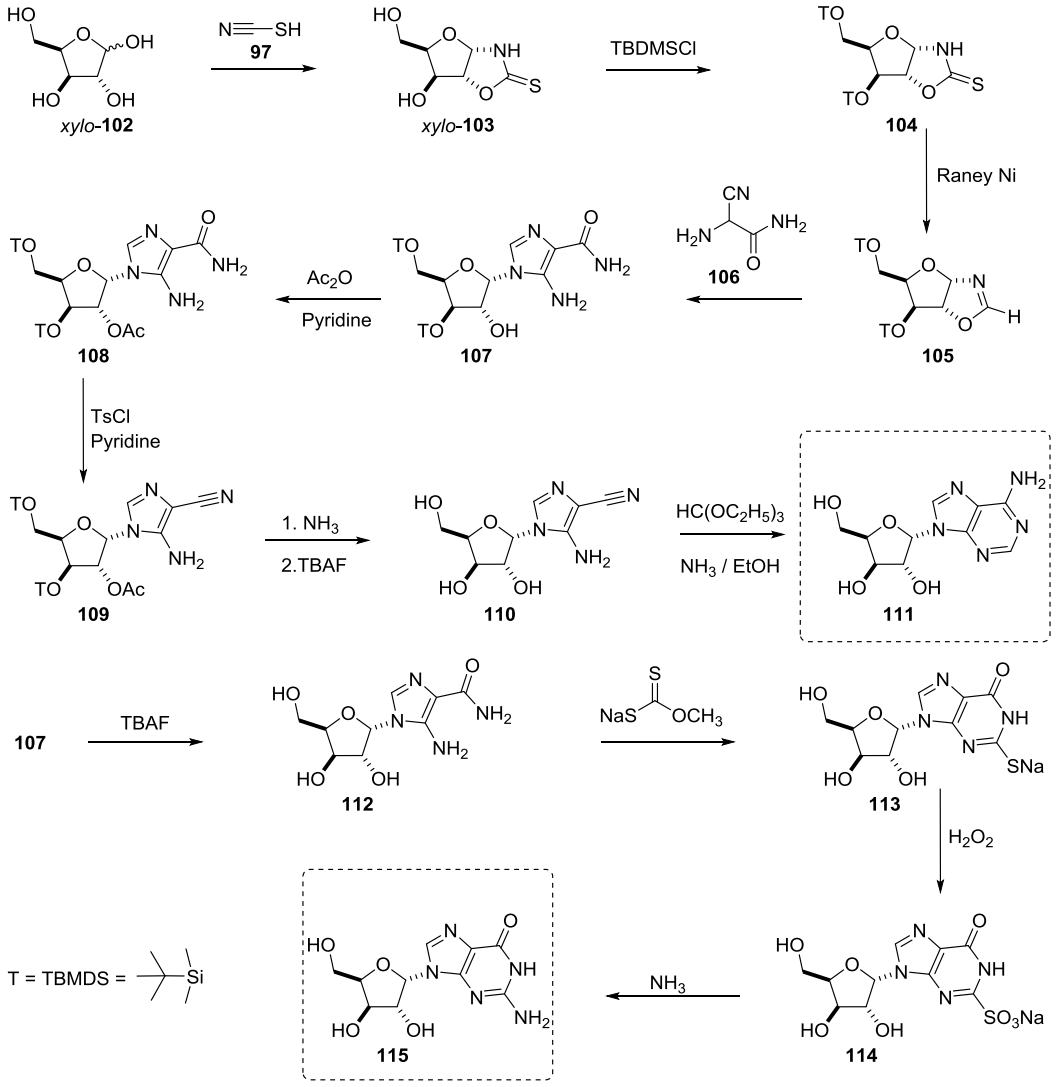


Scheme 1.36: *Plausible synthesis of aminooxazoline arabino-**70** via the reaction of activated thione **90** with ammonia. Bold arrows, reported chemistry and dashed arrows, proposed chemistry.*

2. Synthesis and Activation of Oxazolidinone Thione

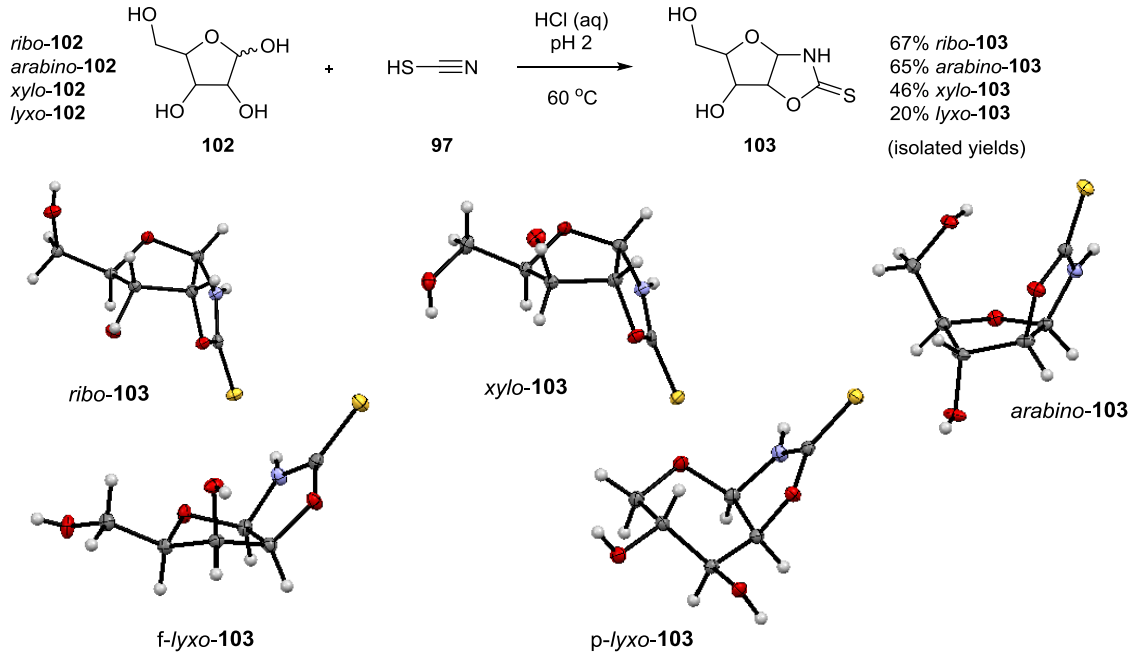
2.1. Using free pentose sugars to form oxazolidinone thiones

Oxazolidinone thione *arabino*-**103**, the key intermediates of our proposed prebiotic purine synthesis, bears a significant structural relationship to aminooxazolines *arabino*-**70** (Scheme 1.19), which suggests that a similar synthetic strategy may be applied to its synthesis. Indeed outside the realms of prebiotic chemistry, Imbach *et al*. have exploited *xylo*-**103** as an intermediate in the synthesis of purine nucleosides. (Scheme 2.1).[156]



Scheme 2.1: *Imbach's synthesis of α-xylofuranosyl purine nucleosides* ***111*** *and* ***115*** *via oxazolidinone thione xylo-****103****. The synthesis relies on organic regents and classical protecting strategies that are outside the constraints imposed by reasonable prebiotic conditions.*

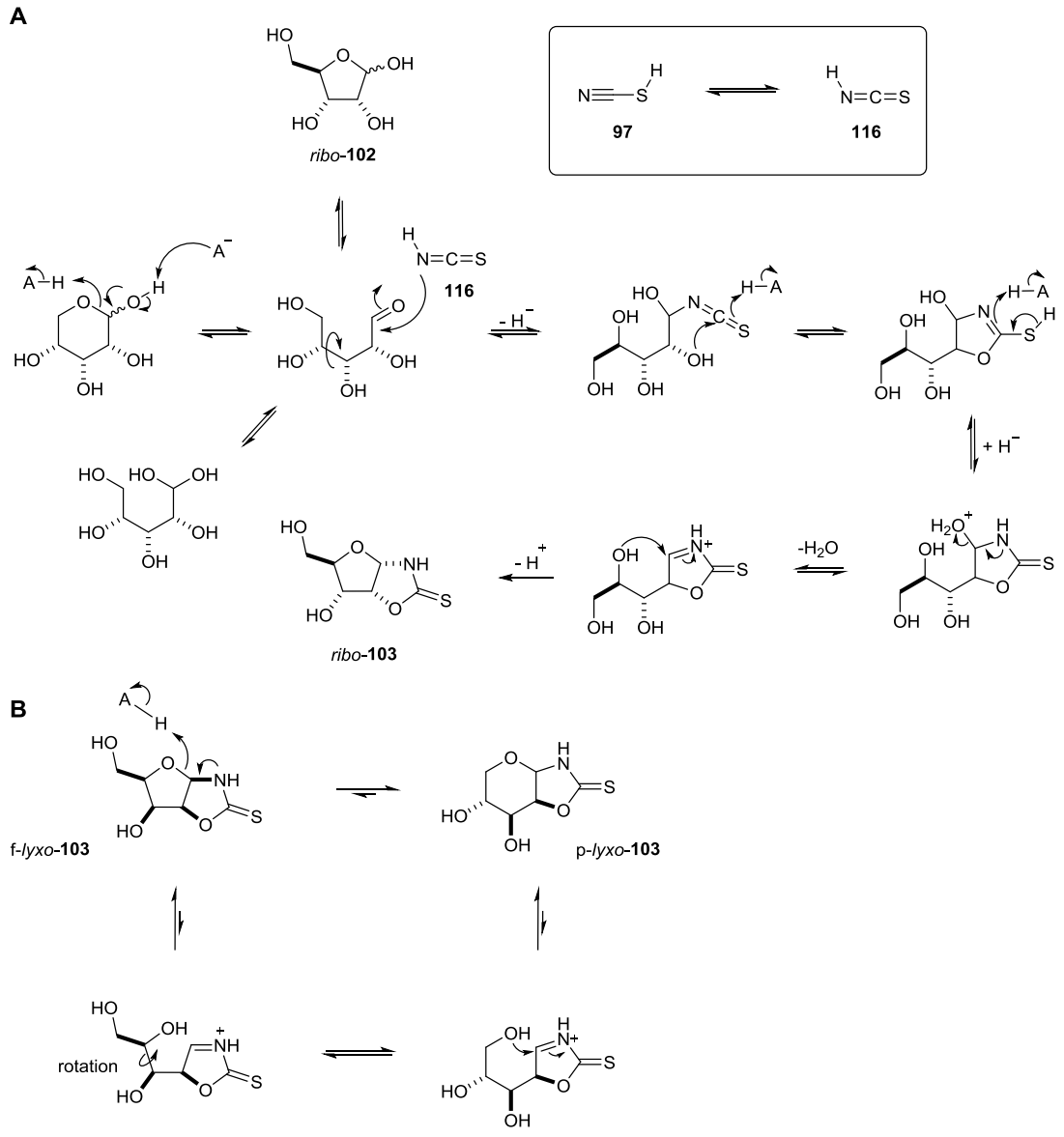
It is of note that Imbach *et al.* synthetic strategy is directly comparable to Orgel's method of the synthesis of aminooxazoline *arabino*-**70**.[124] Specifically, *arabino*-**70** and *xylo*-**103** have been synthesised by direct reaction with ribose *ribo*-**102** with either cyanamide **68** or thiocyanic acid **97**, respectively. Girniene *et al.* have also extensively explored the use of oxazolidinone thione *ribo*-**103** for the synthesis of quinazolinedione nucleoside analogues.[157-159] To increase our understanding of the behaviour of oxazolidinone thiones, and their potential role in prebiotic chemistry we started our project by the synthesis of thiones **103** from their respective free sugars under acidic (pH 2) aqueous conditions (Scheme 2.2).[159] We found that mineral acid catalysed formation of *ribo*-**103**, *arabino*-**103**, *xylo*-**103** and *lyxo*-**103** was readily achieved from their respective sugars (*ribo*-**102**, *arabino*-**102**, *xylo*-**102** and *lyxo*-**102**) with acceptable isolated yields (*ribo*-**103**, 67%; *arabino*-**103**, 65%; *xylo*-**103**, 46% and *lyxo*-**103**, 20 %). To ensure correct assignment of furanosyl selectivity (in the *ribo*- and *arabino*-isomers) we then acquired crystal structures of all the thiones synthesised (Scheme 2.2).



Scheme 2.2: *Synthesis of oxazolidinone thione **103** via free pentose **102**. X-ray structures of all oxazolidinone thiones showing the correct stereochemistry.*

It was found that all of the pentose oxazolidinone thione **103** were synthesised with complete furanosyl selectivity, except *lyxo*-isomer. Importantly this means that the

core furanosyl sugar moiety of ribonucleotides can be established during the synthesis of *arabino*-**103**. However, oxazolidinone thione *lyxo*-**103** was observed to exist as an equilibrating mixture of furanosyl (f-*lyxo*-**103**) and pyranosyl (p-*lyxo*-**103**) isomers. It is thought that this equilibration occurs in the *lyxo*-isomer due to the increased steric bulk associated with locating all the *lyxo*-substituent on one face of the furanosyl ring (Scheme 2.3B).



Scheme 2.3: A) *Proposed mechanism for formation of oxazolidinone thiones from pentose sugars. Shown here for the synthesis of ribo-**44**. Equilibrium between thiocyanic acid **97** and iso-thiocyanate **116** (inset). B) Postulated mechanism for interconversion lyxo furanose oxazolidinone thione (f-lyxo-**103**) and lyxo pyranose oxazolidinone thione (p-lyxo-**103**).*

The reaction of ribose *ribo*-**102** and thiocyanic acid **97**, was monitored, at room temperature by ^{1}H NMR spectroscopy (with single water suppression). The disappearance of sugar resonances and the appearance of oxazolidinone thione *ribo*-**103** was observed to occur over the course of 228 h. However, the transient intermediates (depicted in Scheme 2.3), which are supposed to be accessed *en route* to *ribo*-**103**, were not detected during our time course experiment. It is possible that some of these intermediates may have been detected at lower temperatures, but no further efforts were made to observe these structures.

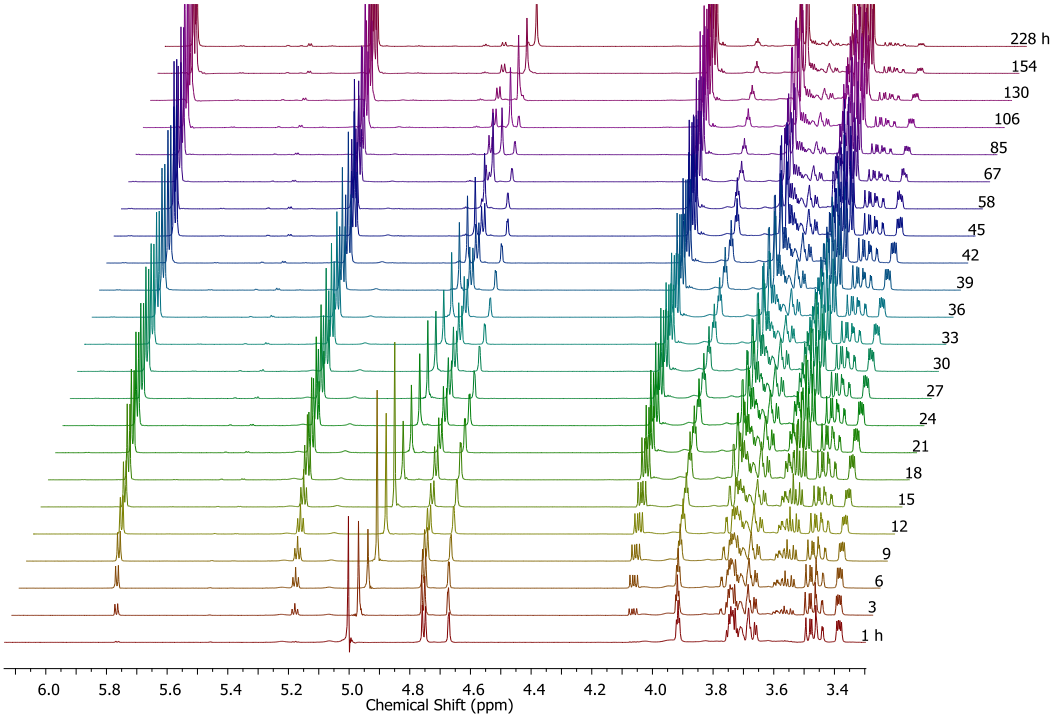


Figure 2.1: *A stack of ^{1}H NMR spectra (600 MHz, D_2O, 3.3 - 6.2 ppm) to show products observed upon incubation of ribose ribo-**102** (1.1 mM) and thiocyanic acid **97** (2.2 mM) at rt and pH -1.5 for 228 h. The peaks for ribo-**102** is decreasing and peaks for oxazolidinone thione ribo-**103** is increasing over time. No intermediate species were detected in ^{1}H NMR spectra with single solvent suppression.*

The formation of oxazolidinone thione *ribo*-**103** was found to be synthesised at an increased initial rate than the other pentose oxazolidinone thiones (*arabino*-***103*** and *xylo*-**103**) from their corresponding pentose sugars (Figure 2.2). The differences in reaction kinetics are attributed to the greater availability of the open-chain aldehydic form of ribose *ribo*-**102** relative to the other pentose *arabino*-**102**, *xylo*-**102** and *lyxo*-**102**.[98]

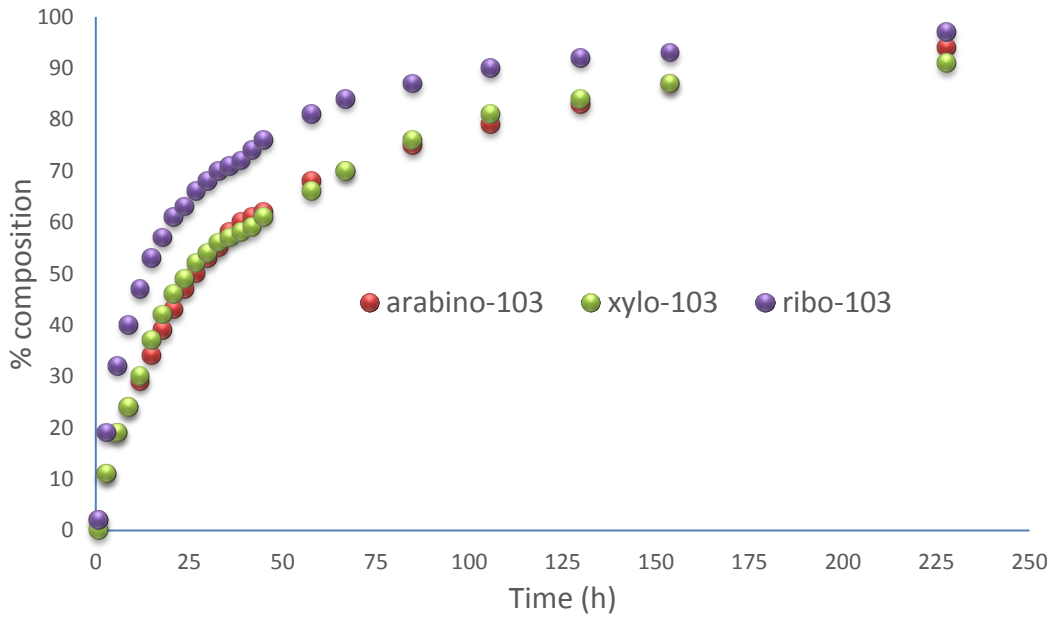


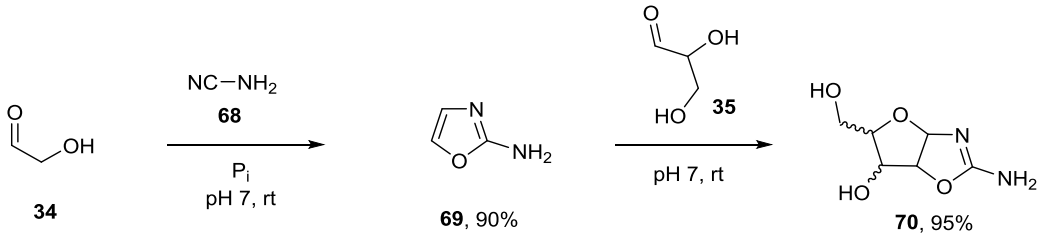
***Figure** 2.2**:** Comparative reaction of ribo-**102**, arabino-**102** and xylo-**102** with potassium thiocyanate **97**, over a period of 228h at rt to yield ribo-**103,** arabino-**103** and xylo-**103** respectively. Each point in the graph is ^{1}H NMR data obtained by integration of H-(C1´) of thione **103** against sugar **102** peaks at 3.39 and 4.75 ppm. Each pentose **102** (110 mM) was incubated with **97** (220 mM) at pH -1.5 and rt for 228h and then data collected for all three oxazolidinone thiones, ribo-**103**, arabino-**103** and xylo-**103** are plotted for comparison*

Despite the furanosyl selectivity of oxazolidinone thione synthesis being impressive and an excellent step toward chemoselective ribonucleotide synthesis, the synthesis and stability of free-pentose sugars **102** are highly problematic under prebiotic conditions.[70, 71, 83] Therefore we next investigated a potential prebiotic route for the synthesis of oxazolidinone thiones **103** that bypasses the pentose sugars **102**. Our strategy was designed to build directly on the observation that the pentose sugar moiety of pyrimidine ribonucleotides can be assembled by the condensation of C_2 + C_3 sugar fragments.[22, 121] The prebiotic synthesis of C_2 and C_3 sugars has been successfully reported by the Sutherland group,[89] furthermore it has also been shown that these simple C_2 and C_3 sugars can be sequestrated via crystallization from complex prebiotic mixtures likely to have been present on early Earth (see Section 1.7.2).[134]

2.2. Prebiotic synthesis of oxazolidinone thiones (by-passing free pentose sugars)

2.2.1. Synthesis of 2-thiooxazole

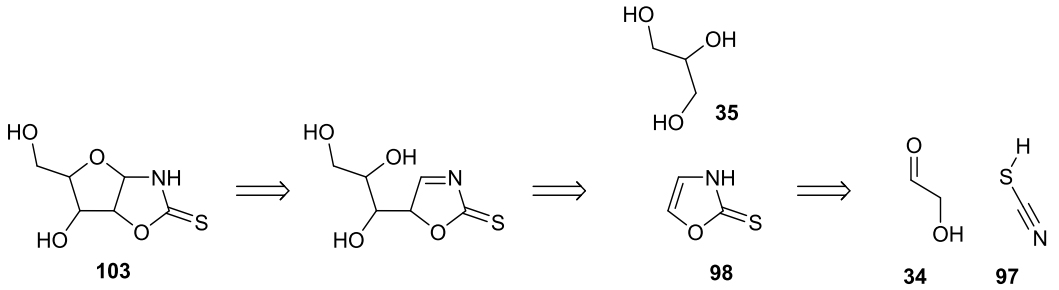
Powner *et al.* have previously reported that pentose aminooxazolines are key prebiotic intermediates for pyrimidine ribonucleotide synthesis.[22, 121] The masked-aldol reaction of 2-aminooxazole **69** and glyceraldehyde **35** to give aminooxazoline **70** was essential to the chemoselective stereocontrolled synthesis of pyrimidine ribonucleotides by-passing free sugars, which are extremely unstable[70, 71]. The heterocyclic molecule, 2-aminooxazole **69** is synthesised (under prebiotically plausible conditions) upon the addition of glycolaldehyde **34** to cyanamide **68** at pH 7 with phosphate acting as a general acid/base catalyst, previously this reaction had only been reported at pH 10 with organic co-solvent (Schemes 2.4 and 1.19).[160]



Scheme 2.4: *Prebiotic synthesis of aminooxazolines **70** from feedstock of prebiotic small molecules.*

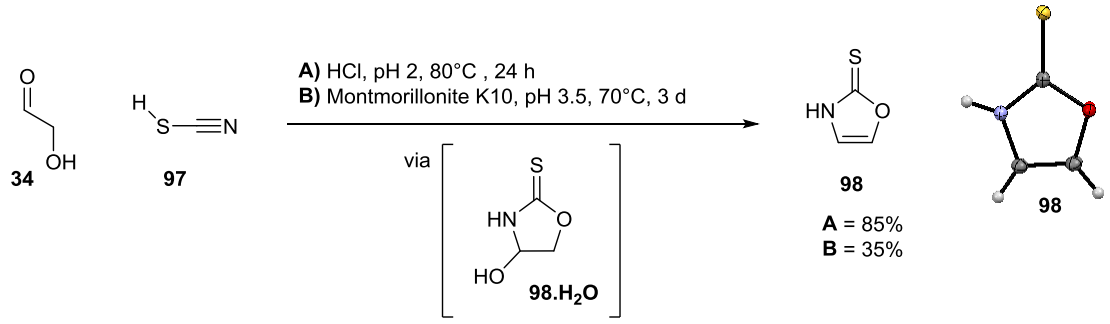
We aimed to build our synthetic strategy by direct modification of the chemistry reported by Powner *et al.* Specifically, we aimed to modify aminooxazoline synthesis to develop oxazolidinone thione synthesis that could be applied to purine ribonucleotide synthesis, but retaining the prebiotic advantage of the previous reported oxazolines synthesis. 2-Aminooxazole **69** is a key ribonucleotide precursor, which derives from glycolaldehyde **34** and cyanamide **68** in the synthesis reported by Powner *et al.* (Scheme 2.4).[22] The C2 carbon atom of **69** is regiospecifically positioned as the C2-pyrimidine carbon atom during abiotic pyrimidine synthesis upon sequential reaction with glyceraldehyde **35** and then cyanoacetylene **30**. Whereas in our proposed

purine synthesis, the C2 carbon atom of oxazole **69** will be activated by introducing a sulfur atom at C2. It was envisaged that this C2 sulfur atom would provide the required chemoselectivity for *in situ* activation towards regiospecific purine nucleobases synthesis upon the thione sugar scaffold. Accordingly, it was hypothesised that synthesis should commence with prebiotically plausible thiocyanic acid **97**,[45, 92, 161] which can be generated quantitatively from the reaction of hydrogen cyanide **14** with sulfur or hydrogen sulfide (H_2S),[162] replacing cyanamide **68** that had been deployed in pyrimidine synthesis (Scheme 2.5).



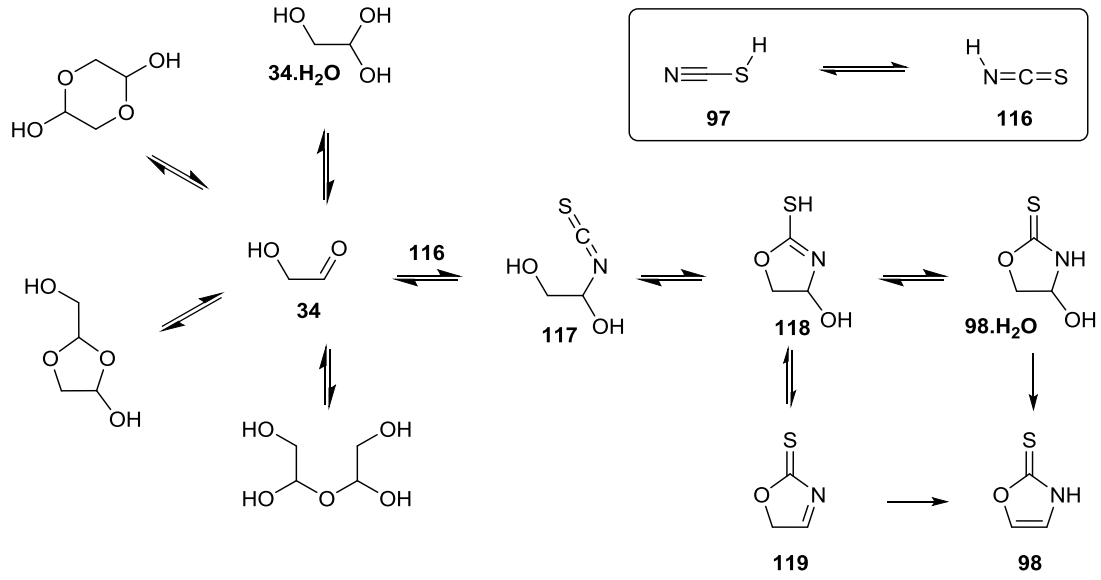
Scheme 2.5: *Retrosynthetic analysis for prebiotic synthesis of oxazolidinone thione* ***103****, by-passing pentose sugars.*

We first synthesised 2-thiooxazole **98** via a literature procedure reported by Leconte *et al.*,[163] and pleasingly this led to 85% isolated yield of **98** (Scheme 2.6). This was then used to investigate the masked-aldol reaction of thiocyanic acid **97** and glycolaldehyde **34**. It is also of note that the literature procedure is a prebiotically plausible reaction. Glycolaldehyde **34** is a corner stone of prebiotic chemistry, which has been observed in non-biological environments (outside our biosphere),[164] and recently its prebiotic synthesised from photoredox chemistry of hydrogen cyanide (H_2S) has been reported (cyanosulfidic protometabolism, Section 1.52).[89] Furthermore, glycolaldehyde **34** has also been used in the prebiotic synthesis of pyrimidine ribonucleotides.[22] Interestingly, Montmorillonite clay, which has been proposed to be an important catalyst in template free oligomerisation of RNA monomers,[165] was found to provide a suitable replacement for mineral acid in the synthesis of **98**, giving 35% yield of thiooxazole **98** upon incubation of glycolaldehyde **34** (416 mM) and potassium thiocyanate (416 mM) at 40°C for 4 days (Scheme 2.6).



Scheme 2.6: *Reaction of thiocyanic acid* ***97*** *with glycolaldehyde* ***34*** *to produce 2-thiooaxazole* ***98****. X-ray structure obtained for* ***98*** *(right).*

The ambident nature of thiocyanate **97** nucleophilicity has been extensively investigated in the chemistry literature. Mayr and co-workers report that, under kinetic control thiocyanate **97** is observed to yield thiocyanates, but under thermodynamic equilibrium **97** lead to isothiocyanates **116**.[166] Accordingly, we speculated that specific acid catalysed thiooxazole synthesis is highly efficient at low pH due to the rapid equilibration of the two tautomeric forms of thiocyanic acid. Such that **97** reacts as an N-nucleophile rather than an S-nucleophile (Scheme 2.7).



Scheme 2.7: *Proposed mechanism for formation of 2-thiooxazole* ***98*** *from the reaction of glycolaldehyde* ***34*** *and thiocyanic acid* ***97****. The two tautomers of* ***97*** *(inset).*

It is thought that the synthesis of 2-thiooxazole **98** proceeded *via* the addition of **116** to the aldehydic form of the glycolaldehyde **34** (~5% in aqueous solution)[167]. The resulting intermediate, **117**, undergoes 5-*exo-dig* cyclisation to yield **118** (Scheme 2.7), followed by tautomerisation to form the initially observed hydrate **98.H$_2$O**,

which subsequently undergoes dehydration to yield the aromatic 2-thiooxazole **98** (Figure 2.3). Alternatively the cyclised product **118** may undergo dehydration to form product **119**, followed by tautomerisation to yield **98.** The driving force of the reaction is thought to be aromatisation. This is different to the formation of 2-amininooxazole **69**,[22] which needs phosphate catalysis and is more electron rich and therefore weakly aromatic/more polarised than 2-thiooxazole **98**. Furthermore, it is also of note that compound **98** can be purified by crystallization directly from water and can also be transported by sublimation. These physical properties provide simple prebiotically plausible mechanisms for the purification, accumulation and transport of 2-thiooxazole **98**.[168, 169]

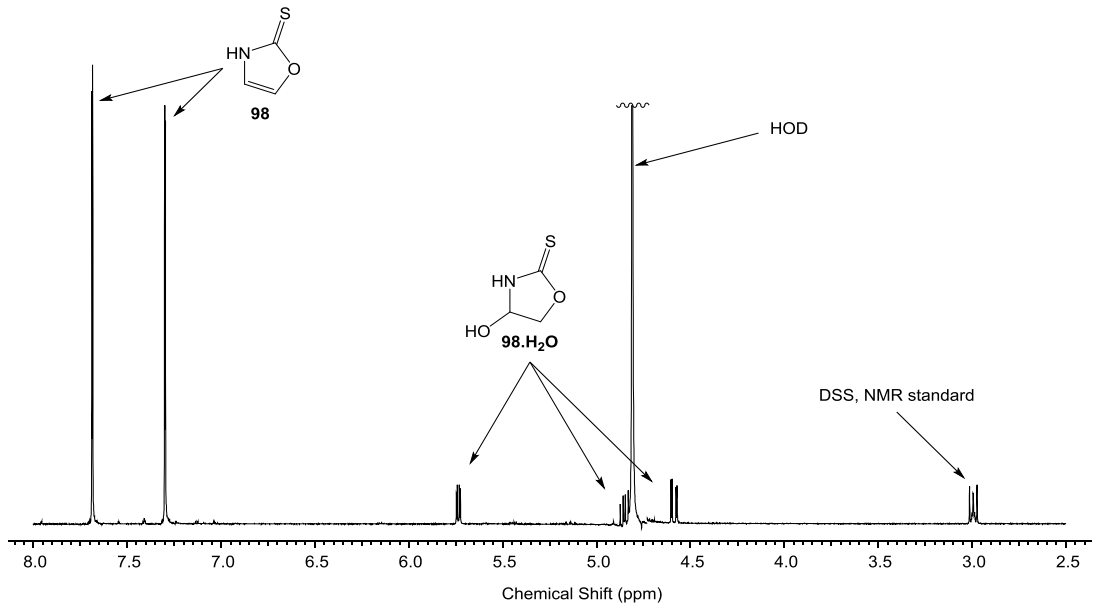
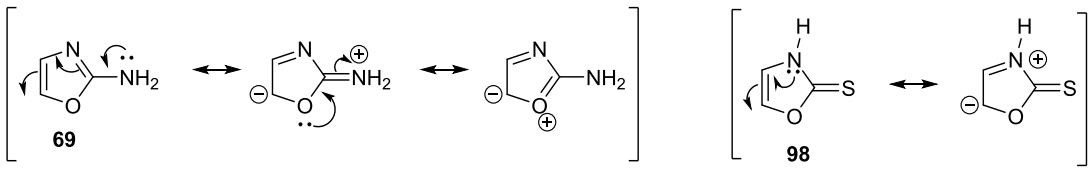


Figure 2.3: *^{1}H NMR (600 MHz, D_2O, 2.5 – 8.0 ppm) to show the reaction of glycolaldehyde **34** (100mM), thiocyanic acid **97** (500mM) and 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, NMR standard) in D_2O (1 mL) after incubated at 80 °C for 24 h. Calibration to internal standard indicates an 85% yield of 2-thiooxazole **98** and a 15% yield of its hydrate, 4-hydroxyoxazolidine-2-thione **98.H_2O**.*

2.2.2. Reaction of 2-thiooxazole with glyceraldehyde

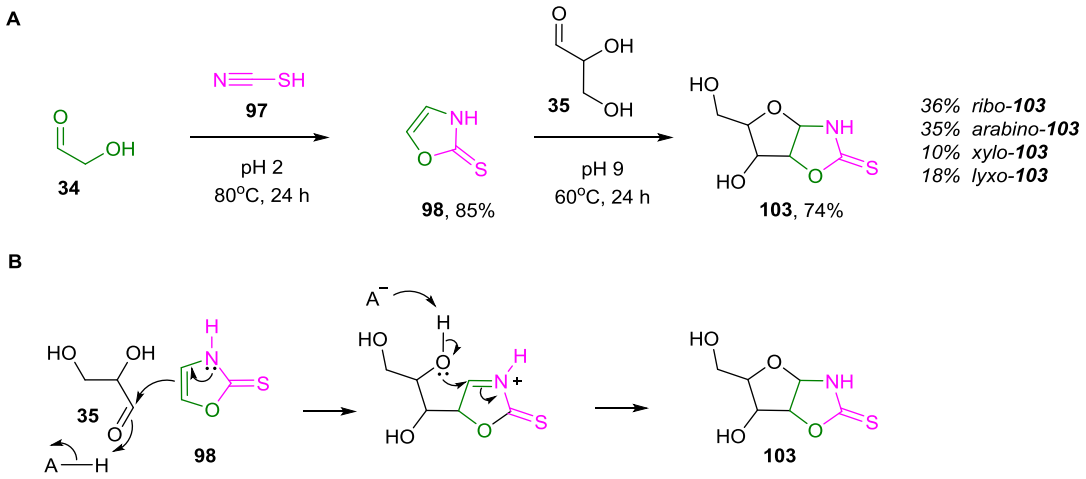
After successfully establishing a prebiotic synthetic route for 2-thiooxazole **98**, we next investigated the reaction of **98** with glyceraldehyde **35** to yield oxazolidinone

thiones **103**. The reaction of 2-aminooxazole **69** with glyceraldehyde **35** has previously been reported to give a high yield of aminooxazolines, *ribo*-**70** and *arabino*-**70** aminooxazolines (44% and 30%, Section 1.7).[22] 2-Aminooxazole **69** is a highly electron rich oxazole, and it has a high degree of enamine character, this renders **69** as an excellent aldol donor in aqueous solution (Scheme 2.8). Although 2-thiooxazole **98** is inherently less electronically activated than 2-aminooxazole **69**, we predicted that the weak nitrogen-sulfur electronic overlap would result in appreciable enamine character in **98**. Nonetheless we anticipated that the condensation of **98** with glyceraldehyde **35** might require more forcing conditions than the reaction of **35** with 2-aminooxazole **69**.



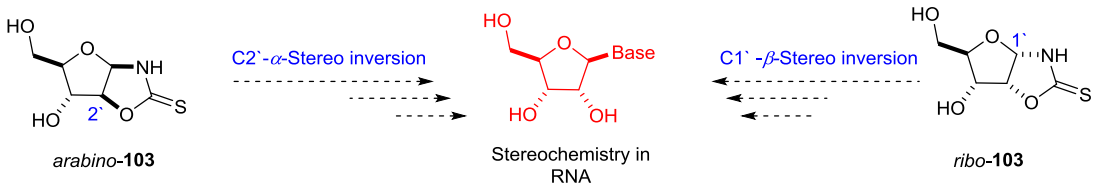
Scheme 2.8. *The nucleophilic character of 2-aminooxazole* ***69*** *and 2-thiooxazole* ***98*** *illustrated by resonance stabilisation.*

Based on our knowledge of the synthesis of aminooxazolines,[22] we reacted 2-thiooxazole **98** with glyceraldehyde **35** to selectively position sulfur at the C2-position of the resulting pentose oxazolidinone thiones **103** (Scheme 2.9). At near neutral pH (pH 4 – 9; Figure 2.4), the reaction of 2-thiooxazole **98** (0.25-0.5 M) with glyceraldehyde **35** (0.5–1 M) in water at 60°C yielded a mixture of the pentose oxazolidinone thiones **103**. The reaction was observed to be slow below pH 6, but more rapid above pH 7, yielding up to 74% thione **103** in 24 h.



Scheme 2.9: *Synthesis of the oxazolidinone thione* ***103*** *from prebiotic feedstock of small molecules.* ***A)*** *Reaction of glycolaldehyde* ***34*** *with thiocyanic acid* ***97*** *give 85% 2-thioooxazole* ***98****, which undergoes masked-aldol reaction with glyceraldehyde* ***35*** *to give 74% oxazolidinone thione* ***103****, arabino and ribo are the major isomers of the reaction.* ***B)*** *Tangible mechanism for the masked aldol reaction of* ***98*** *with* ***35****.*

The reaction proceeds with high *ribo*-/*arabino*- diastereoselectivity (70%, *ribo*-**103**/*arabino*-**103** 1:1; 30% *xylo*-**103**/*lyxo*-**103**), which is of note because both *ribo*-**103** and *arabino*-**103** thiones are only one stereochemical inversion away from the *β*-*ribo*-stereochemistry of RNA (Scheme 2.10).



Scheme 2.10: *Comparison of stereochemistry of oxazolidinone thiones arabino-****103*** *and ribo-****103*** *with the required natural β-ribo-stereochemistry of RNA. Arabino-****103*** *will need to undergo C2´-α-stereochemical inversion and ribo-****103*** *C1´-β-stereochemical inversion to achieve β-ribo-stereochemistry of RNA.*

Furthermore, furanosyl selectivity is equally important *en route* to RNA, and X-ray crystallographic analysis of all the diastereomeric products of the reaction of 2-thiooxazole **98** and glyceraldehyde **35** had demonstrated that only the minor *lyxose* component was furnished as a mixture of furanosyl- and pyranosyl-isomers (p-*lyxo*-**103**/f-*lyxo*-**103**; 2:1; Figure 2.4).

A

				% distribution of oxazolidinone thione products **103**				
98 (M)	**35** (M)	**103** (%)	**98** (%)	*arabino* **103**	*ribo* **103**	*xylo* **103**	*p-lyxo* **103**	*f-lyxo* **103**
0.25	0.5	51	36	35	35	12	11	7
0.50	0.5	35	44	37	36	10	12	5
0.50	1.0	59	25	35	36	12	11	6
0.25	1.0	74	8	36	35	10	12	7

B

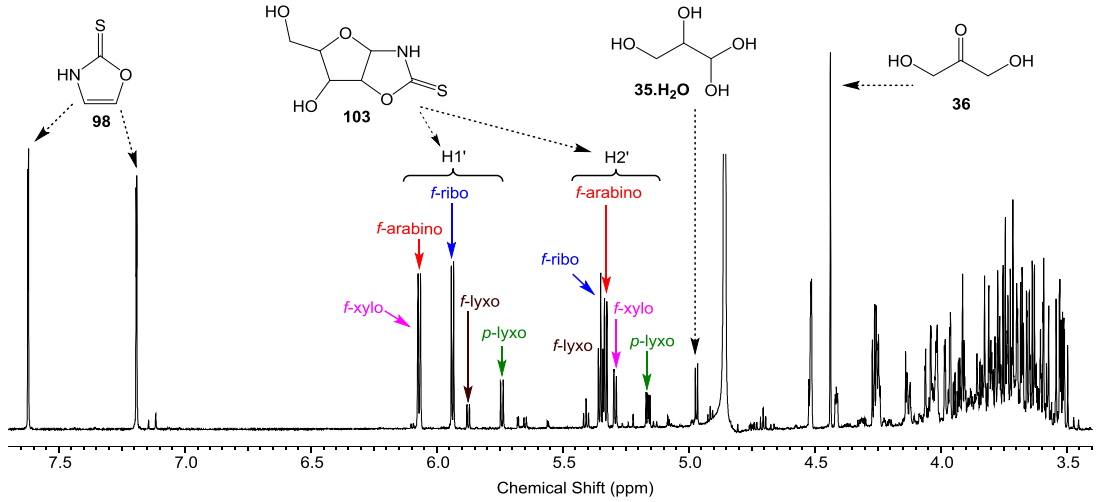


Figure 2.4: A) *Yields of pentose oxazolidinone thiones* ***103*** *and their relative diastereomeric ratio across a variety of concentrations of 2-thiooxazole* ***98*** *and glyceraldehyde* ***35****. All reactions were heated for 24 h at 60°C in pH 9.* ***B)*** *^{1}H NMR spectrum (600 MHz, D_2O, 3.4 – 7.7 ppm) showing reaction of 2-thiooxazole* ***98*** *with glyceraldehyde* ***35****. 2-Thiooxazole* ***98*** *(0.25M) and glyceraldehyde* ***35*** *(1M) at pH 7 were incubated at 60°C for 24 h.*

The identity of the furano and pyrano oxazolidinone thiones **103** were established by spiking with authentic furano and pyrano oxazolidinone thiones (Figure 2.5), which were synthesised using the conventional free pentose sugar synthesis reported in section 2.1. (Scheme 2.2).

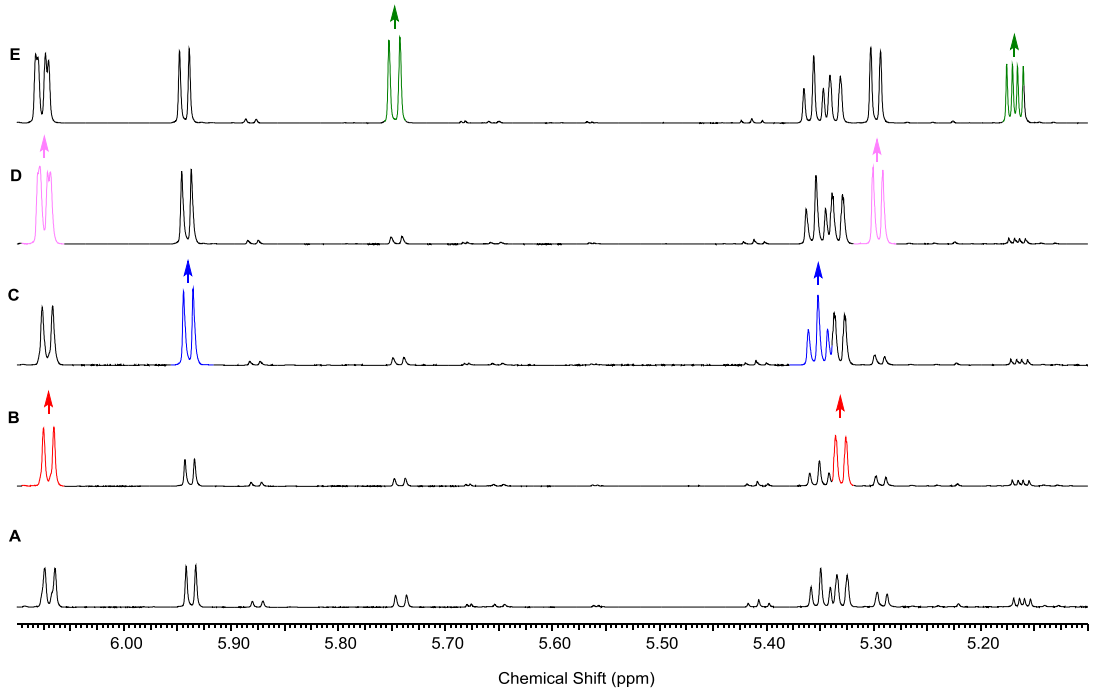
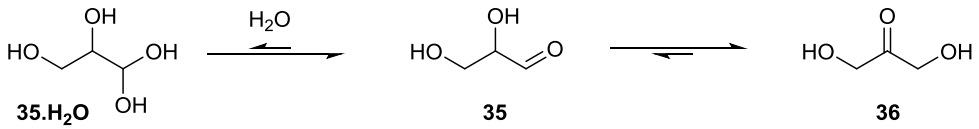


Figure 2.5: *^{1}H NMR spectra (600 MHz, D_2O, 5.1 – 6.1 ppm) showing reaction of 2-thiooxazole **98** with glyceraldehyde **35**. A) **25** (0.25M) and **98** (1M) at pH 7, incubated at 60°C for 24 h. B) Spiked with arabinose oxazolidinone thione arabino-**103**. C) Spiked with ribose oxazolidinone thione ribo-**103**. D) Spiked with xylose oxazolidinone thione xylo-**44**. E) Spiked with lyxose oxazolidinone thione p-lyxo-**44**.*

A major by-product of the reaction of 2-thiooxazole **98** with glyceraldehyde **35** is dihydroxyacetone **36**, which is produced from enol-tautomerisation of glyceraldehyde **35** (Scheme 2.11). Due to the equilibration of **35** with **36**, and thermodynamic preference for ketone **36** over aldehyde **35**, an excess of glyceraldehyde **35** was required in the reaction with **98** to drive consumption of **98** to completion.



Scheme 2.11: *Equilibration of glyceraldehyde **35** into dihydroxyacetone **36** and hydrate **35.H₂O**.*

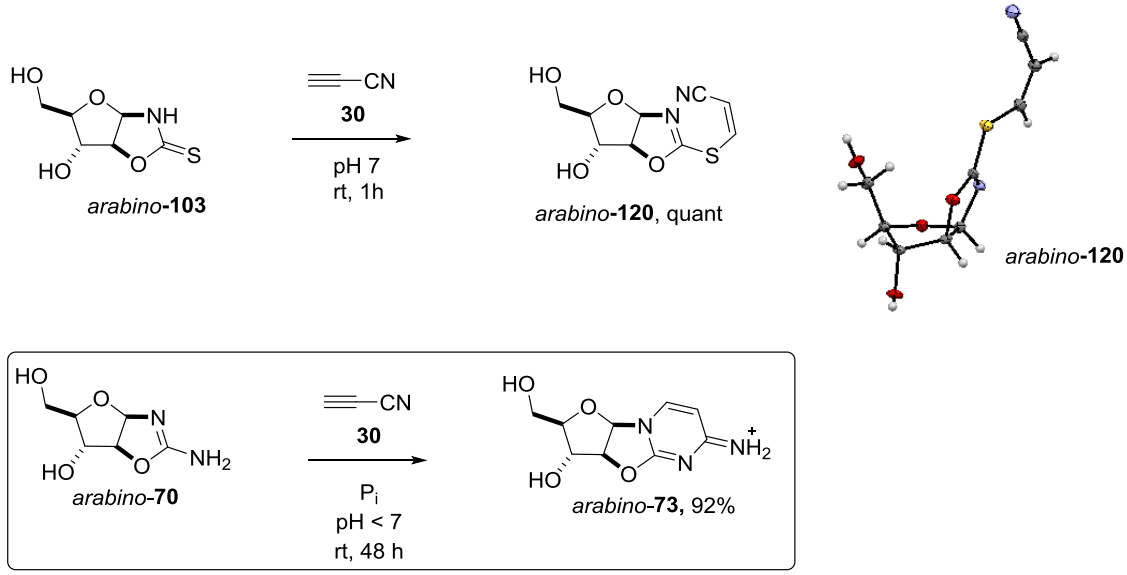
The reaction of **98** with **35** is also, as expected, observed to be slower than with 2-aminooxazole **69** with **35**. The more sluggish nature of the addition of **98** to **35** provides more time for equilibration of glyceraldehyde **35** to its ketose isomer dihydroxy acetone **36**. Moreover, unlike the reaction of the 2-aminooxazole **69**, no phosphate was used in this reaction, as phosphate catalyses formation of **36** from **35**,[133] and therefore was observed to decrease the yield of oxazolidinone thiones **103** at the expense of **36**. The problem of dihydroxyacetone **36** in the synthesis of activated pyrimidine ribonucleotides has recently been resolved by Islam *et al*. by selective sequestration of glyceraldehyde aminals from an equilibrating mixture of triose sugars (Section 1.7.2).[134] The effect of these aminal crystallisations, have not been explored in the context of purine synthesis, however, the observed yield and stoichiometry limitation imposed by triose isomerisation in the purine pathway suggest that further investigation of these aminal chemistries coupled with thione **103** synthesis should be carried out in future investigation. It may be that iterative reactions of glyceraldehyde aminals with thiooxazole **98**, followed by *in situ* recovery of dihydroxyacetone **36** (through aminal crystallisation) could be used to overcome the requirement for excess glyceraldehyde **35** that has been exploited in these current investigations. It is also important to note that unlike **69** (which reacts readily with dihydroxyacetone **36**), the reaction of **98** and **36** was not observed, it may be possible that future experiment can be designed to exploit this selectivity to realise a kinetic resolution of **35** and **36** to synthesis **103** from **36**. However this kinetic resolution was not pursued at this time.

2.3. Site specific thione activation

2.3.1. Cyanovinylation of oxazolidinone thione

With the *ribo*- and *arabino*-furanosyl pentose backbone now installed, we next investigated the selective electrophilic activation of oxazolidinone thiones **103** to facilitate our strategy to construct the core heterocyclic structure of purine nucleotides on the sugar-thione scaffold of **103**. We began our investigation with cyanoacetylene **30**, which plays a fundamental and essential role in pyrimidine synthesis.[22] Cyanoacetylene **30** has been detected in abiotic environments[45] and is observed to be

the second highest abundance product of reduced atmosphere spark discharge experiments.[42, 170] Cyanoacetylene **30** reacts with aminooxazoline *arabino*-**70** to yield 2,2´-anhydrocytidine *arabino*-**73** (<92%), however absolute pH control is required, if the pH of the reaction is unbuffered the reaction pH increases and causes rapid hydrolysis of *arabino*-**73** and then subsequent unwanted reactions of the hydrolysis product with cyanoacetylene **30** (Scheme 1.22).[22, 124]



Scheme 2.12: *Cyanovinylation of oxazolidinone thione arabino-**103** with cyanoacetylene **30** under plausible prebiotic conditions (pH 7, rt, 1h) to give quantitative yield for S-cyanovinylated oxazolidinone thione arabino-**120**. Crystallographic studies also confirmed the C1´-C2´ β-configuration in arabino-**120.** The reported phosphate buffered (pH <7) cyanovinylation of aminooxazoline arabino-**70** to give 2,2´-anhydrocytidine arabino-**73** (inset).*

We envisaged cyanoacetylene **30** was ideally suited to activate oxazolidinone thione **103** in aqueous solution due the slow reaction of cyanoacetylene **30** with water and the excellent electron withdrawing propensity of the cyanovinyl moiety, which will activate the C-2 carbon atom of **103**.[42, 171] Upon treatment of *arabino*-**103** (130 mM) with cyanoacetylene **30** (260 mM) at pH 7, a rapid and completely selective (exocyclic reaction at sulfur rather than endocyclic reaction at *N*1) *S*-cyanovinylation was observed with a click-like efficiency[172] to give a quantitative yield of *S*-cyanovinylated oxazolidinone thione *arabino*-**120** (Scheme 2.12). Due to the higher pKa of anhydronucleoside 2,2´-anhydrocytidine *arabino*-**73** compared to *S*-cyanovinyl thione *arabino*-**120**,[173] and sulfur-prohibited annulation, no increase in pH was observed during cyanovinylation of thione *arabino*-**103** (Figure 2.6). Increasing

pH is a hallmark of the addition of cyanoacetylene **30** to aminooxazoline *arabino*-**70** in water, rendering pH buffered cyanovinylation essential to pyrimidine synthesis,[22] however no buffer was required to control the reaction of *arabino*-**103** with cyanoacetylene **30**. Thus cyanoacetylene provides a superbly controlled and quantitative activation of *arabino*-**103** in water. Although the activated thione *arabino*-**120** was used in the subsequent nucleophilic reactions without any purification (other than removal of excess cyanoacetylene **30** during concentration), *arabino*-**120** can easily be purified by flash column chromatography.

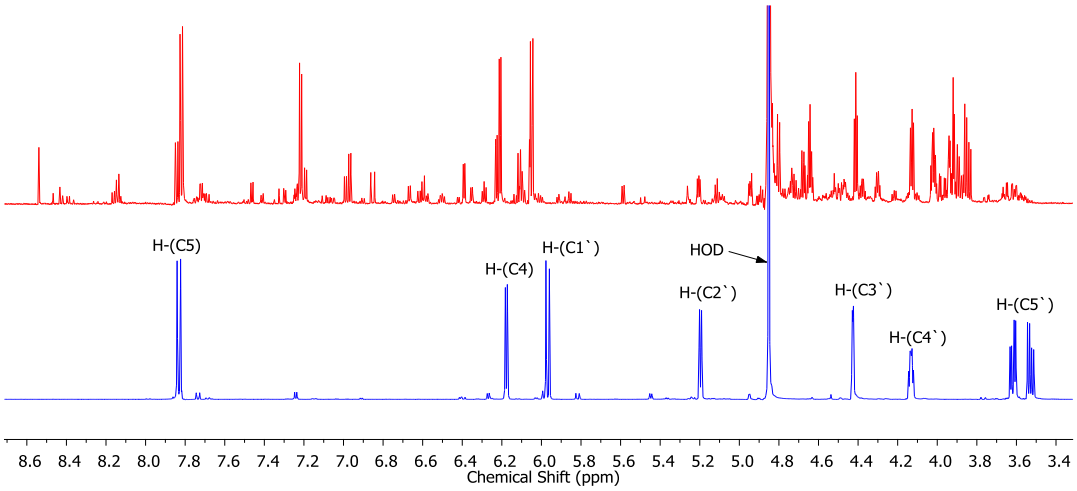


Figure 2.6: *^{1}H NMR spectra (600 MHz, D_2O, 3.4 – 8.6 ppm) comparing crude reaction for the cyanovinylation of aminooxazoline arabino-**70** and oxazolidinone thione arabino-**103**. Unbuffered reaction of arabinose oxazolidinone thione arabiono-**103** with cyanoacetylene **30** at neutral pH after 1 h cleanly furnished S-cyanovinyl oxazolidinone thione arabnino-**120** in quantitative yield (blue). Unbuffered cyanovinylation of aminooxazoline arabino-**70** with **30** after 24 h giving rise to a plethora of hydrolysis and cyanovinylated products (red).*

Cyanovinylation of oxazolidinone thiones *arabino*-**103** at pH 4.0 and pH 10 was slow with the latter also showing complex unidentified cyanovinyl peaks in the ^{1}H NMR spectra, likely due to alkaline hydrolysis of cyanoacetylene **30**. However at pH 7 the reaction was both rapid and clean. The sluggish reactivity of *arabino*-**103** below pH 5, is attributed to substrate protonation (*arabino*-**103**, pKaH = 5.2). Small quantities (~5-10%) of additional cyanovinylation were observed in extended incubation of *arabino*-**103** with cyanoacetylene **30** at neutral pH, these were determined to be due to C3´-OH cyanovinylation.

Powner *et al.* have previously reported cyanovinylation of the 3´OH of aminooxazoline *arabino*-**70** at pH 7.5 (Figure 2.7A). Based on the close structural relationship of *S*-cyanovinyl thione *arabino*-**120** with aminooxazoline *arabino*-**70** we suspected cyanovinylation of the 3´OH in *arabino*-**120** to yield dicyanovinylated product **122.** The structure of compound **122** was determined upon its partial isolation, both COSY and ^{1}H NMR spectroscopic analysis indicated for the formation of **122** (Figure 2.7B).

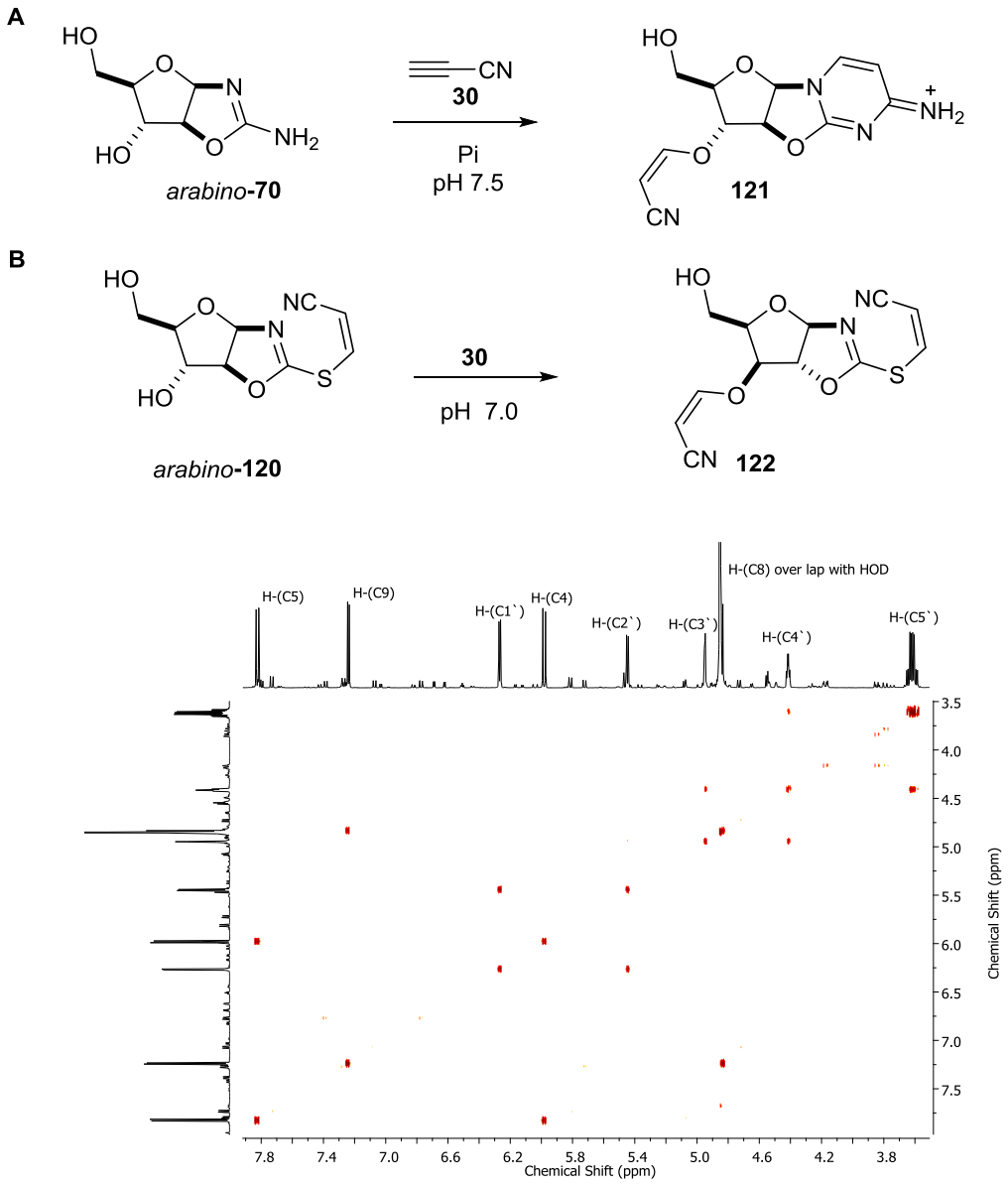
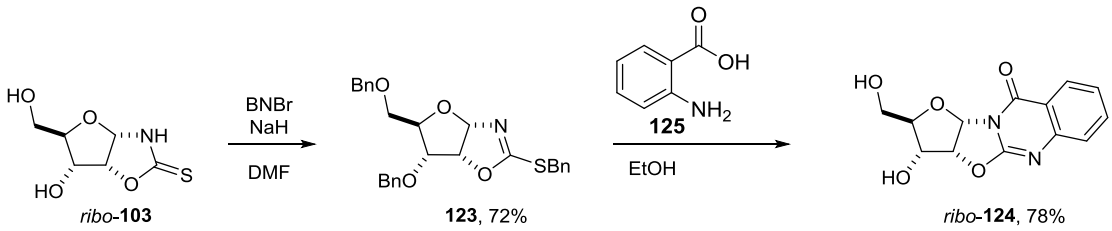


Figure 2.7: A) *3´OH cyanovinylation of aminooxazoline buffered at pH 7.5.* ***B)*** *3´OH cyanovinylation of S-cyanovinyl oxazolidinone thione arabino-**120** at neutral pH and ^{1}H NMR spectra (600 MHz, D_2O, 3.5 – 8.1 ppm) of biscyanovinylated product **122**, the C3´-H is shifted upfield due to the presence of cyanovinyl group on C3´-O.*

Prebiotic protecting group strategies provide very powerful tools for chemical selectivity, however this is a largely unexplored field in prebiotic chemistry. Recently, selective acylation of nucleosides, nucleotides and glycerol-2-phosphocholine in water has been published in the Powner laboratory.[174] We therefore hypothesise that the observed C3´-OH cyanovinylation may also find uses as protecting strategy due to its remarkable selectivity for the secondary C3´-OH (instead of the primary C5´-OH). This selectivity was thought to be due to the partial donation of electron density of C5´-OH into C2 carbon atom through $n \rightarrow \pi^*$ orbital overlap which has been observed in pyrimidine anhydronucleoside *arabino*-**73**.[126] However, upon crystallographic analysis of *S*-cyanovinyl oxazolidinone thione *arabino*-**120** a similar overlap was not observed. It is possible that crystal packing precludes this interaction for *arabino*-**120** in the solid state, however it is also possible that the observed activation is inductively induced and results from the electronic withdrawing imposed upon C2´-oxygen atom through the cyanovinylated thione moiety. No further experiments were undertaken at this time to observe the average solution state bond angle at C5´-OH (for example NOE experiment).

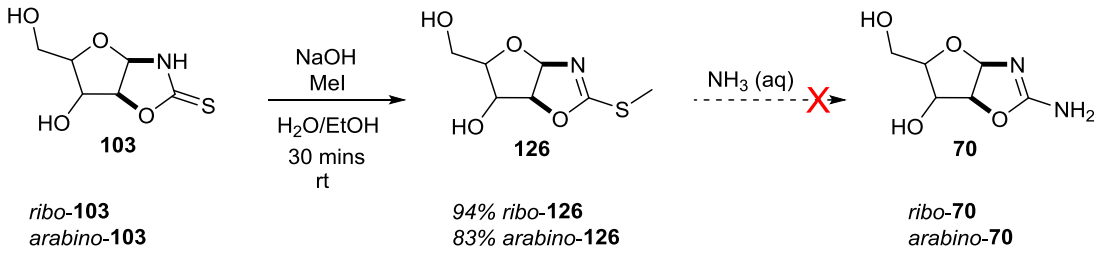
2.3.2. Methylation of oxazolidinone thione

Shaw *et al.* have previously demonstrated both selective methylation of oxazolidinone thiones **103** and reported that *S*-methylated derivatives of thione **103** are "singularly unreactive towards nucleophiles".[175] For example, they did not observe reactivity of *S*-methyl oxazolidinone thione **126** with aqueous or methanolic ammonia. To our understanding, thiolate displacement from *S*-methylated thione **126** to yield aminooxazoline **70** has only been observed upon treatment with formamide at 90°C for 3 h (which was likely to be contaminated with ammonium formate and can slowly release ammonia and formic acid).[176] However, Girniene *et al.* have previously reported the successful substitution of *S*-benzyl thione **123** with anthranilic acid **125** to synthesise2 quinazolinedione *ribo*-**124** (Scheme 2.13).[158] Importantly, their studies demonstrated the concept of ambident reactivity that is essential to our purine synthetic strategy, *albeit* in reports that are limited to anthranilic acid derivatives in organic solvent (ethanol or *t*-butanol).



Scheme 2.13: *Formation of quinazolinone ribo-**124** via benzylthiooxazoline intermediate **123** using ethanol and protection of the free C3´OH and C5´-OH as benzyl ethers.*

We hypothesized that thione protonation and weak amine solvation were both essential to these limited examples in formamide or absolute alcohol solvents. It was unclear to what scope the nucleophilic additions of methylated thiones were investigated by Shaw *et al.*[175] and Davidson *et al.*[176] Therefore, we decided to investigate methylation of oxazolidinone thiones **103** to gain further insight into their utility in nucleophilic addition reactions.



Scheme 2.14: *An adapted synthesis of methylated oxazolidinone thiones **126**. Both ribo-**126** and arabino-**126** did not undergo displacement with aqueous ammonia to form aminooxazoline **70**.*

Following an optimised procedure based on literature reactions,[175] *S*-methylated oxazolidinone thiones, *ribo*-**126** and *arabino*-**126** were synthesised in excellent yields from the reaction of oxazolidinone thiones *ribo*-**103** and *arabino*-**103** respectively with iodomethane under basic conditions (Scheme 2.14). We then investigated the reactivity *S*-methyl thione **126** and in agreement with literature reports, we observed no detectable reaction between *S*-methyl thione **126** (pK_{aH} = 2.4) and ammonia (pK_{aH} = 9.2) in aqueous solution. However, we anticipated that *S*-cyanovinyl oxazolidinone thione *arabino*-**120**, may take part in nucleophilic displacement reactions due to the greatly increased electron-withdrawing effect of the cyanovinyl moiety which is unavailable in *S*-methyl and *S*-benzyl thiones (**126** and **123**). To test displacement of

cyanovinyl group, and to find a plausible prebiotic route to furnish **126**, we investigated sulphide displacement in *arabino*-**120**. It was observed that sequential addition of cyanoacetylene **30** and then methanethiol **127** to thione *arabino*-**103** furnished *arabino*-**126** in up to 50% yield at pH 6, providing a prebiotically plausible route to *S*-methylated oxazolidinone thione **126** (Figure 2.8).

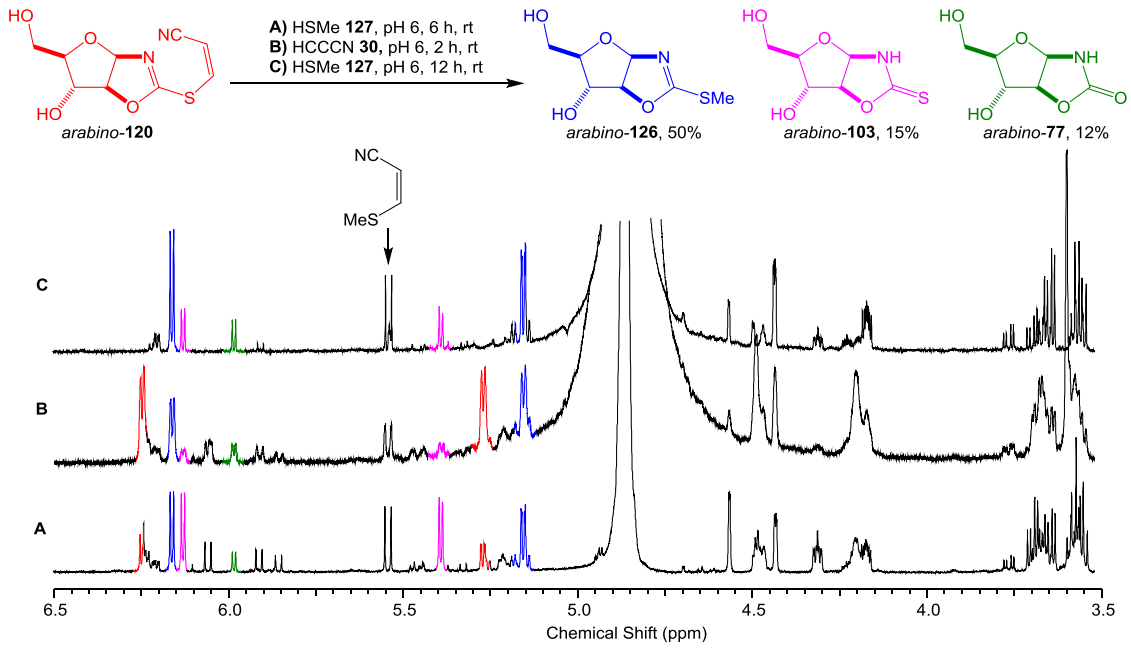


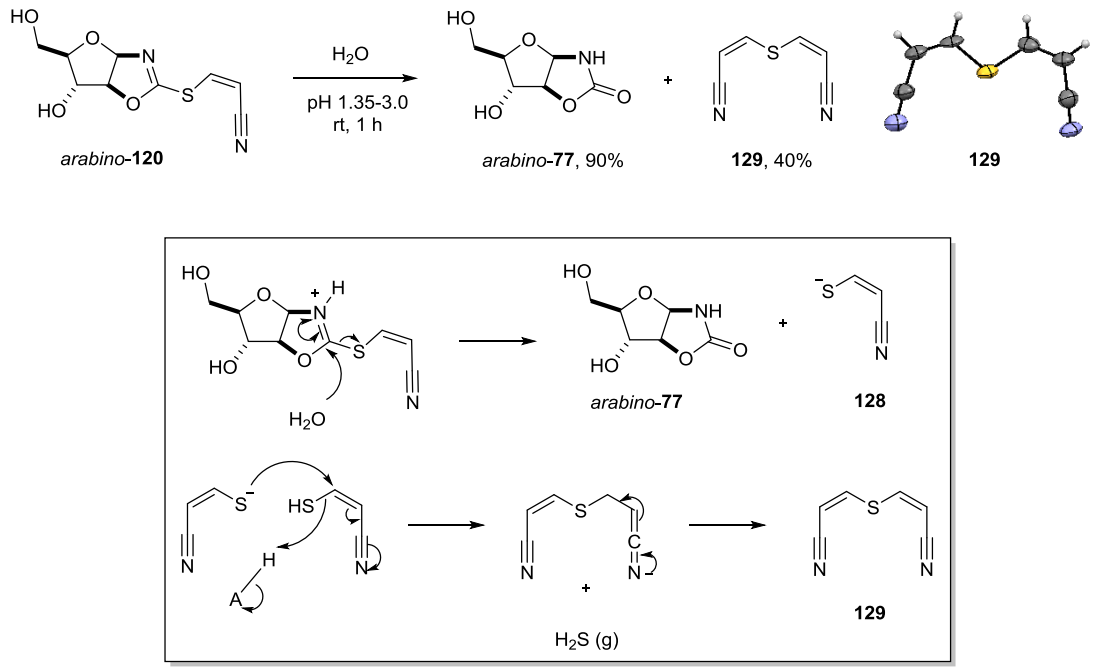
Figure 2.8: *^{1}H NMR spectra (600 MHz, D_2O, 3.5 – 6.5 ppm) showing conversion of S-cyanovinyl oxazolidinone thione arabino-**120** to S-methyl oxazolidinone arabino-**126**. An aqueous solution of arabino-**120** at pH 6 was treated with methanethiol **127** for 6 h **A)** NMR to show conversion at 6 h, mixture contained 30% arabino-**126**, 30% oxazolidinone thione arabino-**103**, 10% oxazolidinone arabino-**77**. The solution was briefly sparged with nitrogen, an aqueous solution of cyanoacetylene **30** was added and the mixture was left to stand for 2 h. **B)** ^{1}H NMR to show near complete conversion of arabino-**120** to arabino-**126**. The solution was again saturated with methanethiol **127** and incubated for 16 h. **C)** ^{1}H NMR to show mixture containing 50% arabino-**126**, 15% arabino-**103** and 12% arabino-**77** after 16 h incubation with methanethiol **127**.*

As preparative synthesis of *S*-methylated oxazolidinone thiones **126** was significantly easier and cheaper to achieve on large scale than *S*-cyanovinylated oxazolidinone thiones **120**, which involves the use of cyanoacetylene **30** (which is not commercial and is unstable). These thiones were used as a model reaction system to develop thione displacement reactions. Prebiotically, it might also be argued that *S*-alkylated thione

126 are structurally simple and therefore, perhaps more readily available than *S*-cyanovinyl thione **120**, indeed alkyl thiols may have been widely accessible.[177] The exchange of cyanovinyl thiol **128** for simple alkyl thiols may present a mechanism for the prolong storage/stabilisation of activated thiones in aqueous solution. Nevertheless, we have demonstrated two (related) pathways (cyanovinylation and cyanovinylation followed by thiol exchange) to access activated oxazolidinone thiones **120** and **126**.

2.3.3. Stability of *S*-cyanovinyl and *S*-methyl thiones

The successful prebiotic synthesis of *S*-cyanovinyl thione *arabino*-**120** and *S*-methyl thione *arabino*-**126** were crucial to our proposed synthetic strategy of purine ribonucleotides, since at this stage we had selectively positioned the sulfur at the C2 carbon atom and these activated thiones were ready to accept nucleophiles. Next, it was deemed necessary to study the stability of *arabino*-**120** and *arabino*-**126** to find suitable nucleophiles, leading toward elaboration of purine nucleobases. We hypothesised that *S*-methyl thione *arabino*-**126** would be more stable than *S*-cyanovinyl thione *arabino*-**120** due to the enhanced electrophilicity of C2 carbon in the latter compound. To test our hypothesis we investigated stability of *arabino*-**120** and *arabino*-**126** in water at varying pHs. When a solution of *arabino*-**120** (50 mM) was incubated at room temperature at low pH (1.3 – 3.0), *arabino*-**120** hydrolysed to form oxazolidinone *arabino*-**77** and the dicyanovinylsulfide **129** as the only major products. It is thought that the protonated *S*-cyanovinyl thione *arabino*-**120** hydrolysed to *arabino*-**77** and *S*-cyanovinyl thiol **128**. The thiol **128** then combined to form the white precipitate of **129** and hydrogen sulfide gas (H_2S) characterised by a strong rotten egg smell (Scheme 2.15). The dicyanovinylsulfide **129** isolated from the reaction was almost completely insoluble in water but dissolved readily in DMSO-d6 for NMR analysis.[142] Finally, an X-ray structure of **129** was obtained to unambiguously confirm the structure of this by-product.



Scheme 2.15: *Stability of S-cyanovinylated oxazolidinone thione arabino-**120** at lower pH (1.35-3.0) and plausible mechanism for acid hydrolysis of arabino-**120** to form oxazolidinone arabino-**77** and dicyanovinylsulfide **129** (inset).*

In order to test the stability of *S*-cyanovinyl thione *arabino*-**120** at elevated pH, *arabino*-**49** (50 mM) in water at pH 7.0 was stirred for 120 h. Interestingly, after 1 week, ^{1}H NMR spectroscopic analysis revealed 37% oxazolidinone thione *arabino*-**103**, 12% dicyanovinylsulfide **129**, 7% oxazolidinone *arabino*-**77** and 33% of *arabino*-**120** (Figure 2.9). When the reaction was repeated at pH 8.0 and 9.0 similar results were obtained, but with a slightly higher ratio of *arabino*-**103** (41%). At pH 11 the half-life ($t_{1/2}$ = 15 h) of the *S*-cyanovinylated oxazolidinone thione *arabino*-**120** was considerably reduced; after 72 h the ^{1}H NMR spectroscopic analysis indicated that all of *arabino*-**120** had been consumed, leading to 31% *arabino*-**103**, 13% *arabino*-**77** and 2% **129**. From these experiments it was clear that at elevated pHs thione *arabino*-**103** was the major product. We suspect that *arabino*-**103** was being formed by Micheal addition of cyanovinyl thiol **128** into *arabino*-**120** (Figure 2.9).

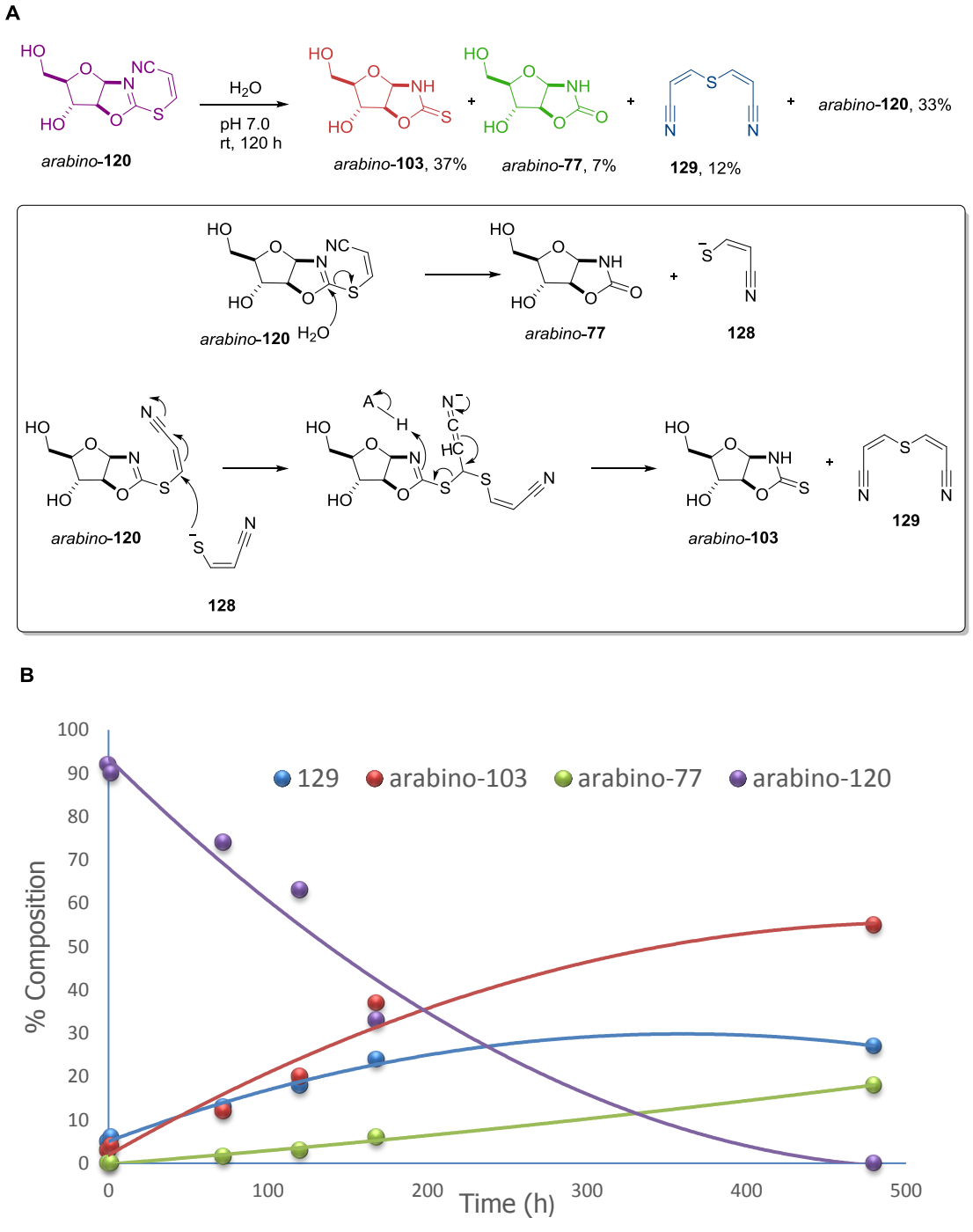
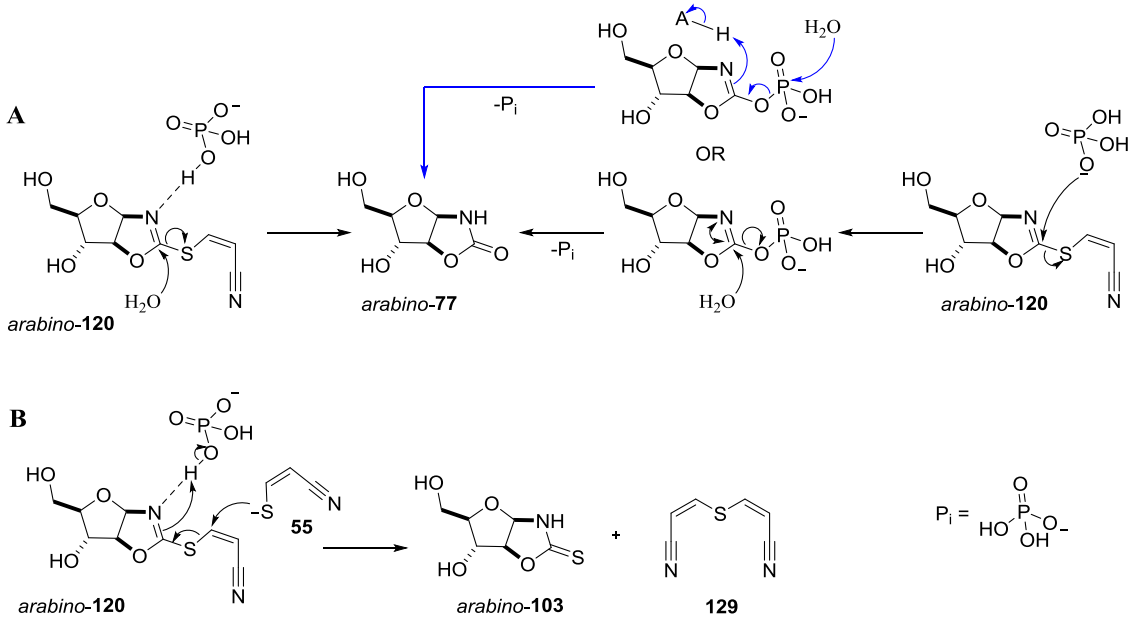


Figure 2.9: A) *Stability of S-cyanovinylated oxazolidinone thione arabino-**120** at higher pH (7.0-11) and plausible mechanism (addition elimination) for base hydrolysis of arabino-**120** to form oxazolidinone arabino-**77**, oxazolidinone thione arabino-**103** and dicyanovinylsulfide **129** (inset).* ***B)*** *Graph for pH stability of arabino-**120** at pH 8.0. Each point in the graph is ^{1}H NMR data obtained by integration of H-(C1´) of arabino-**103**, arabino-**77** and H-(CS) of **129** against H-(C1´) of arabino-**120**.*

Powner *et al.* demonstrated the dual role (acid/base catalyst and pH buffer) of inorganic phosphate in the synthesis of activated pyrimidines.[22] Therefore, it was deemed important to investigate how *S*-cyanovinylated oxazolidinone thione *arabino*-**120** and *S*-methyl thione *arabino*-**126** behaved in the presence of inorganic phosphate as phosphate is eventually required to synthesise ribonucleotides. This would also give us an indication for the potential one pot prebiotic synthesis of pyrimidine and purine ribonucleotides.



Scheme 2.16: *Postulated mechanism for hydrolysis of S-cyanovinylated oxazolidinone thione arabino-**120** in the presence of inorganic phosphate.* ***A)*** *Three possible routes in which phosphate mediated hydrolysis of arabino-**120** leads to formation of oxazolidinone arabino-**77**.* ***B)*** *Mechanism for formation of oxazolidinone thione arabino-**103** and dicyanovinylsulfide **129** in the presence of inorganic phosphate.*

Solutions of *arabino*-**120** (50 mM) and *arabino*-**126** (50 mM) were stirred in a solution of sodium dihydrogen phosphate (100 mM) over a range of pHs (pH = 5, 7 or 9) and at room temperature. ^{1}H NMR spectroscopic analysis showed that the rate of hydrolysis was greatest at pH 5.0, and after 24 h all of *S*-cyanovinyl thione *arabino*-**120** converted to oxazolidinone *arabino*-**77**, while *S*-methyl thione *arabino*-**126** remained unchanged under same conditions. At pH 7 the rate of hydrolysis of *arabino*-

120 decreased but for *arabino*-**126** it increased, indicated by an increase in the production of the oxazolidinone thione *arabino*-**103**. At pH 9 no hydrolysis of the activated thiones (*arabino*-**120** and *arabino*-**126**) was detected after 24 h (Figure 2.10). It was concluded that the presence of phosphate buffer increased hydrolysis in both thiones (*arabino*-**120** and *arabino*-**126**) and any "one-pot" synthesis of purine and pyrimidines will need to consider this issue carefully, since the use of phosphate buffer for pyrimidine ribonucleotide synthesis is vital.[22] (N.B. Further observations in mixed reaction systems have demonstrated that this buffering requirement can be removed during the mixed cyanovinylation of thione **103** and aminooxazoline **70**; vide infra, chapter 5).

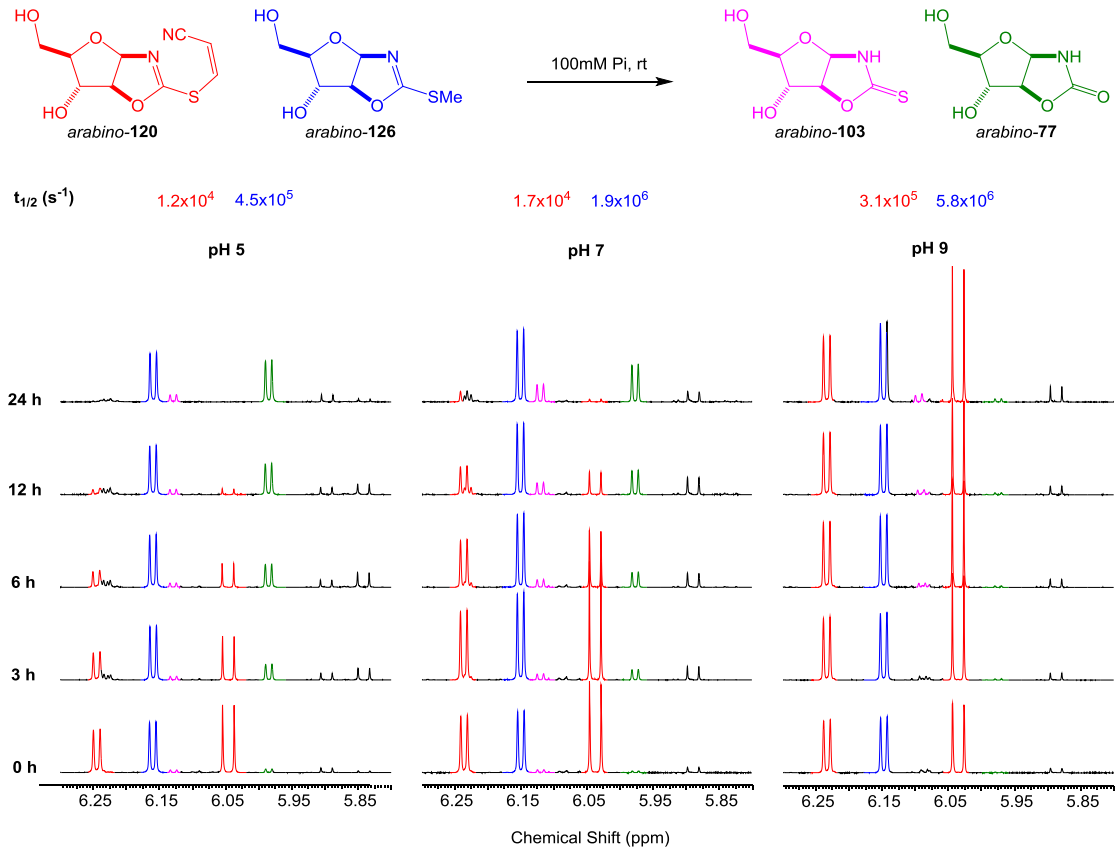


Figure 2.10: *^{1}H NMR (600 MHz, D_2O, 5.8 – 6.3 ppm) showing hydrolysis of S-cyanovinyl oxazolidinone thione arabino-**120** and S-alkyl oxazolinone thione arabino-**126** in the presence of inorganic phosphate across a range of pHs.*

2.3.4. Activation of thione with acrylonitrile to prevent side reactions

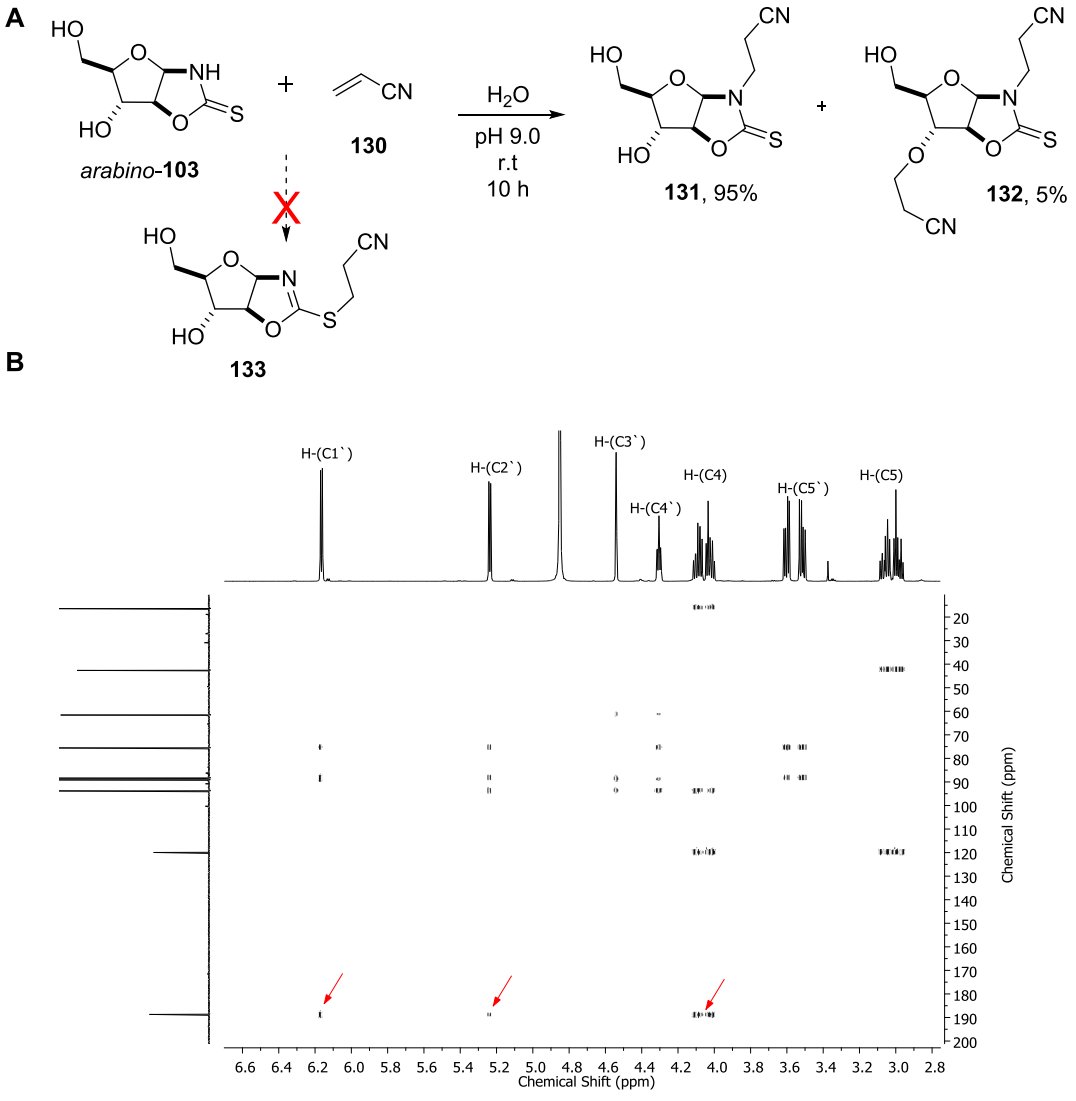


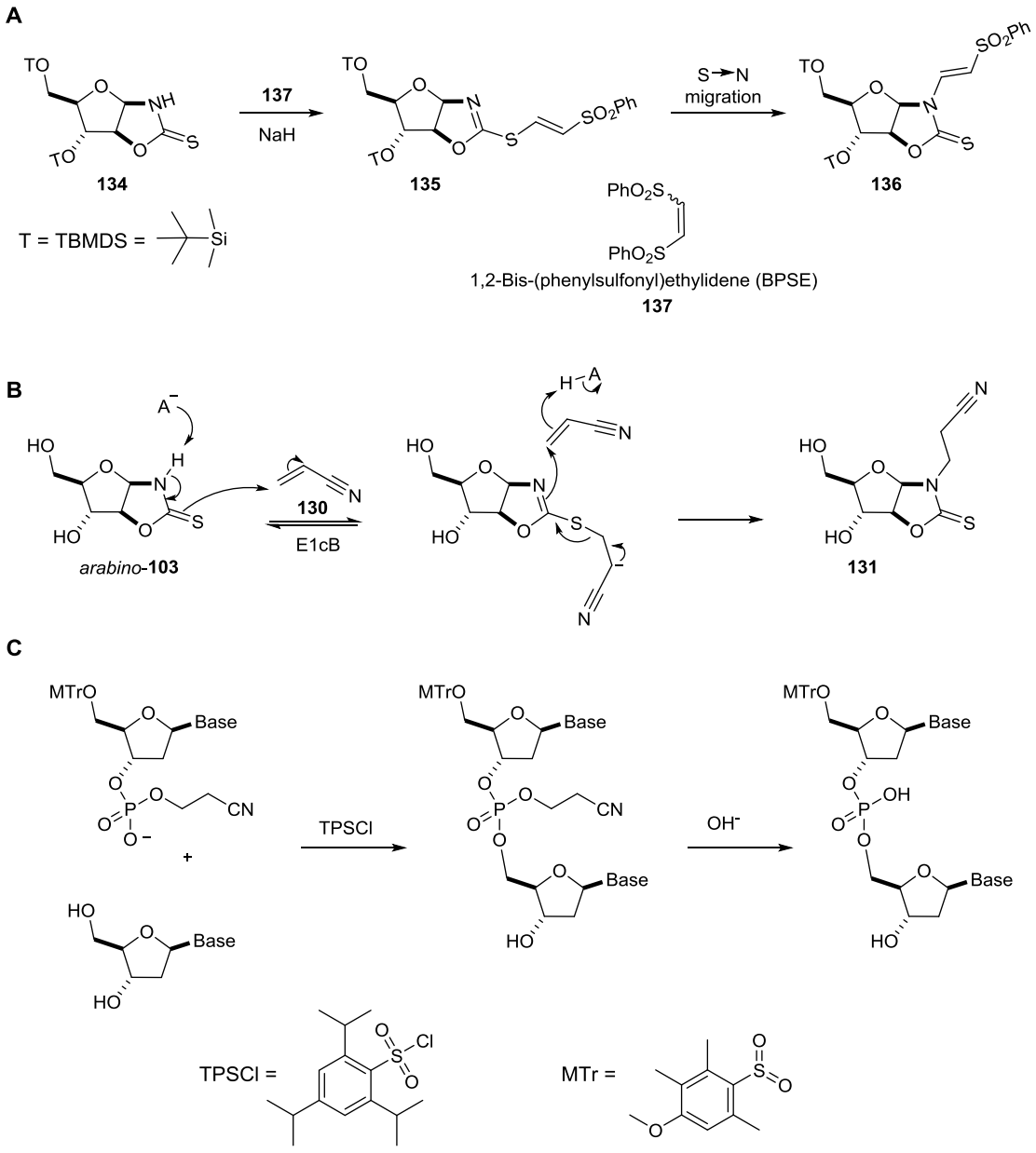
Figure 2.11: A) *Cyanoethylation of the oxazolidinone thione arabino-**103** proceeded via addition to N1 atom instead of the exocyclic sulfur, yielding product **131 (95%)** and **132 (5%)** in favour of the desired product **133**. **B)** HMBC NMR (600 MHz, D_2O, ^{1}H 2.7-6.7 ppm, ^{13}C 10-200 ppm) of **131** showing coupling of C1´-H and C2´-H with C3 carbon of the cyanoethyl group (red arrows), confirming that the cyanoethyl moiety is attached to the endocyclic N1 atom instead of the exocyclic sulfur.*

High yields of oxazolidinone thione *arabino*-**103** were being observed in the displacement reactions of *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (Figure 2.9). This suggested that competing conjugate addition at the α-position of the β-cyanovinyl moiety *arabino*-**120** was occurring. Therefore, we next chose to explore

alternative electrophilic activation to prevent Michael addition to the cyanovinyl group of *arabino*-**120** and improve regioselectivity of nucleophilic addition at the C2 carbon atom. We investigated electrophiles that would result in a saturated thiolate leaving group. Our first electrophile, due to its ostensive similarity to cyanoacetylene **30** and prebiotic plausibility was acrylonitrile **130**.[45] Interestingly, the reaction of oxazolidinone thione *arabino*-**103** and **130** in water at pH 9 rapidly give a single product, which we initially assumed to be **133** (Figure 2.11A). This was then subjected to a range of displacement reactions, and disappointingly was found to be highly stable to all conditions, returning only starting material **133**. At this point, further investigation (and particularly analysis of the HMBC spectrum) of the supposed compound **133** made it obvious to us that we had quite unexpectedly, cleanly furnished the *N*1-cyanoethyl adduct **131** and not *S*-cyanoethyl adduct **133** (Scheme 2.17). We were initially surprised by the apparent change in reactivity; based upon our results with cyanoacetylene **30**, which selectively reacted with exocyclic sulfur of thione *arabino*-**103** to furnish *S*-cyanovinyl thione *arabino*-**120** (Figure 11).

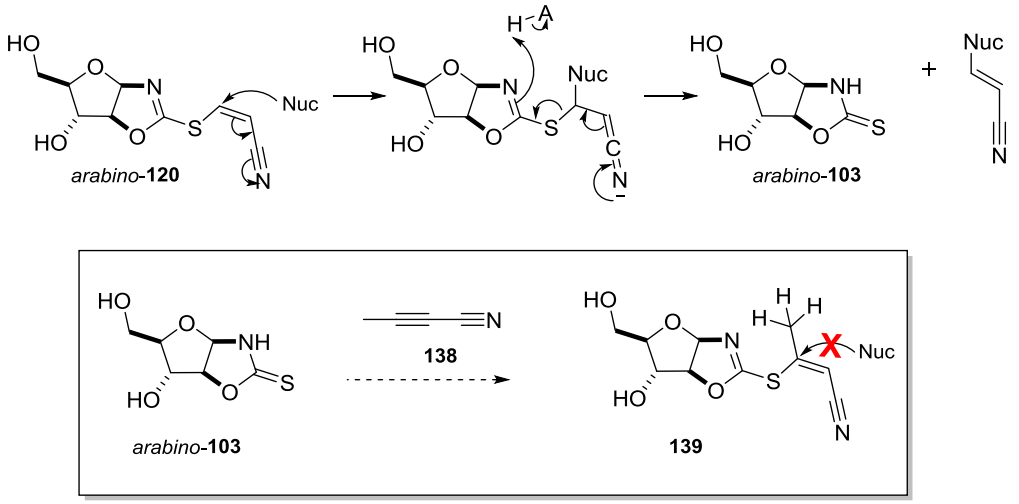
However, Girniene *et al.* have previously observed a mixture of *N*1 and *S*-benzylation upon the reaction of oxazolidinone thione *arabino*-**103** with benzyl bromide and sodium hydride (NaH).[158] The *N*-benzylated derivative of oxazolidinone thione *ribo*-**103** has been isolated (24%), however there is no information on the mechanism and no rational is provided.[158] A tentative mechanism for *N*-vinylation of silylated-thione **134** upon reaction with 1,2-bis-(phenylsulfonyl)ethylidene **137** under basic conditions has been proposed.[157] Girniene *et al.* suggest that the initial reaction occurs, as we anticipated at sulfur giving the *S*-cyanoethyl product **135**, followed by intramolecular rearrangement to give **136** (Scheme 2.17A).[157] However, though Girniene *at al.* describe a plausible pathway for migration of phenylsulfonyl vinyl moiety, with a conjugated sp^2 system, it seems difficult to translate the same migratory process to our observed selectivity in a saturated example. Therefore, we tentatively propose that an addition/elimination mechanism may be occurring under our aqueous conditions (Scheme 2.17B). Elimination would likely be inhibited in *S*-cyanovinyl oxazolidinone thione *arabino*-**120** due to the high energy of the required vinyl anion, but be readily achieved at an sp3 centre adjacent to an anion stabilising cyano group. Furthermore, upon cyanovinylation, the *β*-cyanovinyl moiety of *arabino*-**120** would likely

significantly decrease the electron density at *N*1, and therefore prevent bis-*N*1-cyanovinylation, where alkylation would be expected to be electron donating and favour bis-cyanovinylation. An E1cB mechanism seems highly plausible for the cyanoethyl moiety **131**. Indeed cyanoethyl protection and E1cB deprotection strategies are commonly used in phosphitylation methods for phosphomono- and phosphodiester synthesis (Scheme 2.17C).[178-181]



Scheme 2.17: A) *Tentative mechanism for N1-phenylsulfonyl vinylation of* ***135*** *via S→N migration.* ***B)*** *Proposed mechanism for N1-cyanoethylation of oxazolidinone thione arabino-****103*** *in aqueous solution.* ***C)*** *Use of cyanoethyl group in nucleotide oligomerisation.*

Selective cyanethylation was not observed upon the reaction of oxazolidinone thione *arabino*-**103** with acrylonitrile **130**, therefore a different approach to the problem maybe required. Specifically it may be that steric blockade of the addition of nucleophiles to the cyanovinyl group could circumvent this problem. Therefore, in future it would be worth investigating how more hindered cyanovinylated products such as **139**, which could be synthesised from methylcyanoacetylene **138**, will behave to Michael addition of nucleophiles. (Scheme 2.18).



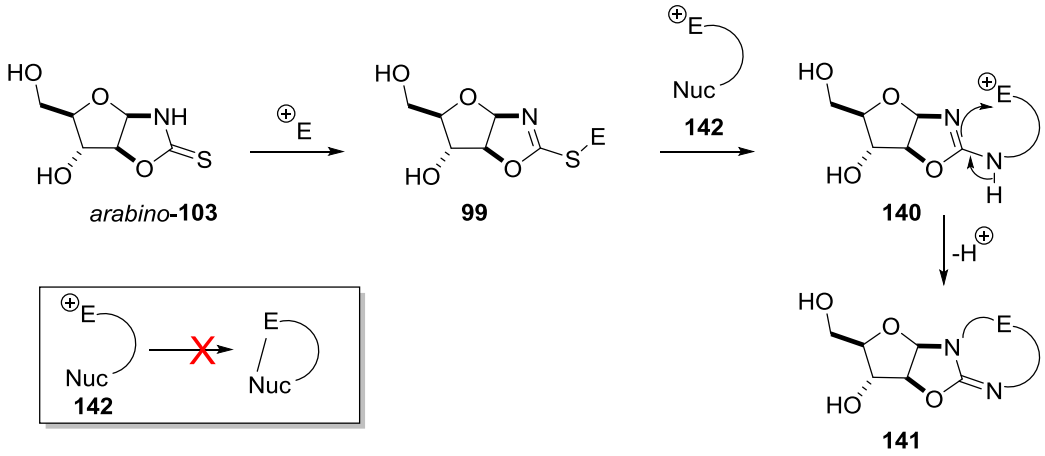
Scheme 2.18: *Plausible mechanism for formation of oxazolidinone thione arabino-**103** during nucleophilic attack on the S-cyanovinylated oxazolidinone thione arabino-**120**. Cyanovinylation of arabino-**103** by methylcyanoacetylene **138** to from **139** which might block Michael addition type reactions, based on steric grounds (inset).*

2.4. Choosing the nucleophile

From the pH stability studies of *S*-cyanovinyl thione *arabino*-**120** and *S*-methyl thione *arabino*-**126**, it was deduced that the nucleophile to be used needed to be meticulously chosen in order to achieve successful displacement of the thiolate group. If the pH is too low, *N*1 of activated thiones (*arabino*-**120** and *arabino*-**126**) is protonated and hydrolysis will take over, forming oxazolidinone *arabino*-**77** and at elevated pH Michael addition of cyanovinyl thiolate **128** leads to the formation of mainly oxazolidinone thione *arabino*-**103** and dicyanovinylsulfide **129**. Therefore, the pH needs to be fine-tuned to get to a pH environment which will maximise the addition of the respective nucleophile and minimise formation of *arabino*-**77** and thione

arabino-**103**. Furthermore, we will specifically be using amine nucleophiles, as we need a nitrogen atom at the C2 carbon atom to give species capable for yielding nucleotides. Alkyl amines normally have a pKa of >9, therefore at neutral pH, the amine will be substantially protonated and ineffective as a nucleophile. Therefore an amine nucleophile with a lower pKa was required so protonation will not impede the amine's reactivity at pHs where thione activation by protonation would occur (*arabino*-**120** pKaH = 5.2, *arabino*-**126** pKaH = 2.4).

To recap, it is remarkable to note the difference in reactivity that has been observed in the reaction of aminooxazolines *arabino*-**70** and oxazolidinone thione *arabino*-**103** with cyanoacetylene **30** (Section 2.3), the former reacting quantitatively on *N*1-nitrogen atom whereas the latter is observed to react near quantitatively on sulfur (Scheme 2.12). As we are exploring sulfur displacement by nitrogen nucleophiles on thiones (*S*-cyanovinyl *arabino*-**120** *and S*-methyl *arabino*-**126***)*. Once the displacement has occurred and we have synthesised an aminooxazoline from a thione. We would predict an interesting switch in reactivity, namely that upon addition the nitrogen adjacent to C2 of the aminooxazoline would become nucleophilic at *N*1. Therefore, if thiolate can be displaced for *arabino*-**120** or *arabino*-**126** with an ambident reagent, a nitrogen nucleophile tethered to an electrophile (such as **142**), we would anticipate intermolecular nucleophilic trapping of the aminooxazoline generated upon nucleophilic addition at *N*1-nitrogen (e.g. intermediate **140** would generate 5 or 6 member ring **141**). This mode of ambident reactivity would then lead to an opportunity to directly build fused tricyclic heterocyclic structures. However, the "electrophilic-amine" (such as **142**) must be wisely chosen and/or judiciously positioned to avoid intramolecular trapping (Scheme 2.19). But, before embarking on our investigation of ambident nitrogen nucleophile, it was decided to choose simple or literature precedented nucleophiles to optimise the nucleophilic addition conditions (pH, concentration, temperature etc.).



Scheme 2.19: *Schematic representation of displacing thiolate with a nitrogen nucleophile tethered to an electrophilic centre, in order to build the heterocyclic core of what is eventually going to become the purine nucleobases. The undesired intramolecular cyclisation of "electrophilic-amine" such as* ***142*** *(inset).*

2.4.1. Displacement of the *S*-methyl thiolate and *S*-cyanovinyl thiolate by a model ambident amine

Having found that oxazolidinone thione *arabino*-**103** can be converted to *S*-cyanovinylated oxazolidinone thione *arabino*-**120** with remarkable efficiency under milder and prebiotically plausible conditions, and also demonstrated the prebiotic synthesis of *S*-methyl oxazolidinone thione *arabino*-**126** via efficient displacement of cyanovinyl thiolate **128** by methanethiol **127** (Section 2.3). Our next goal was to investigate the displacement of thiolate in *arabino*-**120** and *arabino*-**126** with amines.

We first chose to investigate the reactivity of simple *S*-methyl thione **126** which is more stable than *S*-cyanovinyl thione **120** in aqueous conditions. Though Shaw *et al.*[175] and Davidson *et al.*[176] reported **126** to be unreactive with amines, we predicted that the reactivity of **126** could be controlled (switched on/off) through protonation. To test this hypothesize pH switch, we investigated the synthesis of quinazolinedione **124** in water across a broader pH range. Interestingly, we observed near quantitative conversion of **126** via the addition product **143** to *ribo*-**124** and *arabino*-**124** in water between pH 2 and 6, supporting our hypothesis that thione protonation was essential for the activation of *S*-methyl thione. Moreover, the reaction was severely retarded under alkaline conditions, demonstrating that reactivity of **126** is readily modulated

through pH-control (Figure 2.12). We were able to detect the addition product **143** in the first hour of the reaction, identified by the relatively lower chemical shifts of C1´-H (6.2 ppm) and C2´-H (5.4 ppm) in the ^{1}H NMR spectroscopy. X-ray structure of *ribo*-**124** and *arabino*-**124** were obtained to confirm identity of the quinazolinediones.

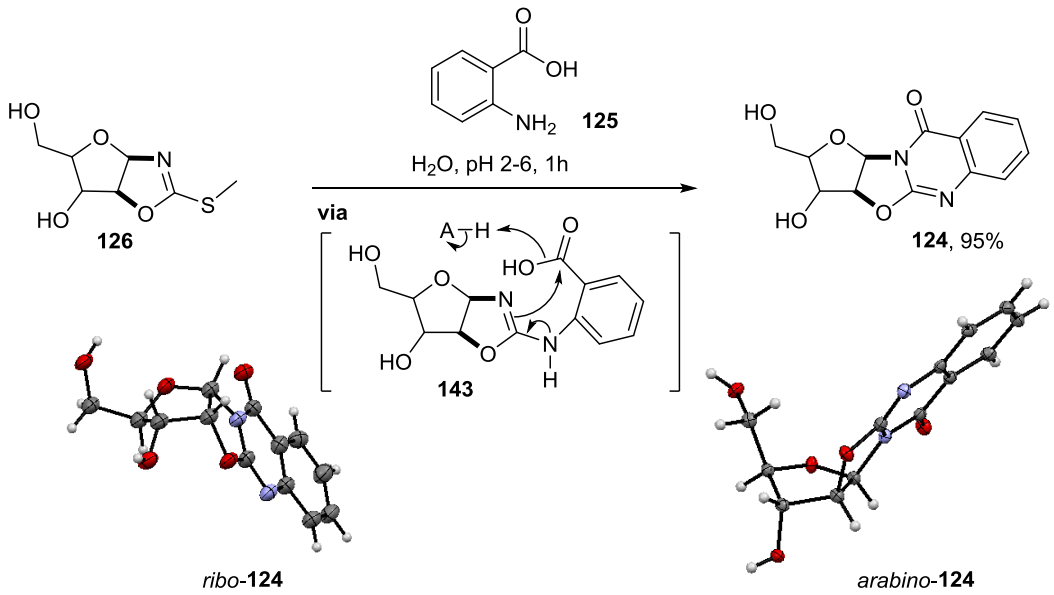
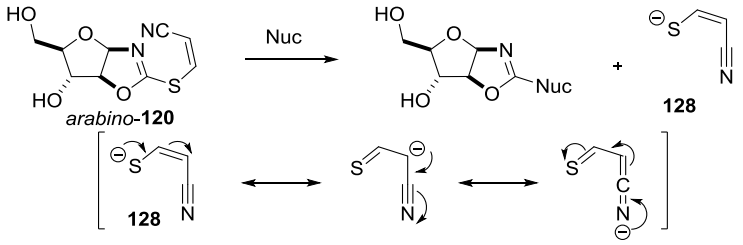


Figure 2.12: *Synthesis of quinazolinedione* ***124*** *by displacement of the S-methyl* ***126*** *(250mM) by anthranilic acid* ***125*** *(500mM) in water at room temperature. The reaction mechanism is thought to proceed via the annulation of the addition product* ***143*** *aided by the carboxylic acid group.*

We suspected *S*-cyanovinyl thiolate **128** to be an excellent nucleofuge, given the conjugation between S and CN groups (Scheme 2.20). Though the pKa of *S*-cyanovinyl thiolate **128** is not known, Ferris *et al.* have reported the pKa of cyanoacetaldehyde **31** to be pKa 8.1 (previously Ferris and Orgel also reported pKa 8.4). Therefore we would expect the mercapto analogue to have a comparable (or indeed perhaps lower) pKa, allowing for the difference in electronic overlap and polarizability of sulfur.



Scheme 2.20: *Conjugation between sulfur and nitrile in S-cyanovinyl thiolate* ***128****.*

Furthermore, **128** has been used recently as nucleofuge during prebiotically plausible oligoribonucleotide ligation by Sutherland and co-workers.[142] To test the reactivity of *S*-cyanovinyl thione *arabino*-**120** with amines, we next investigated the aqueous reaction of anthranilic acid **125** with *arabino*-**120**. Pleasingly, the reaction of *arabino*-**120** (250 mM) with anthranilic acid (500 mM) was found at pH 3.8 to successfully yield 96% (NMR) quinazolinedione *arabino*-**124**. However, intriguingly, we observed the reaction in water at all pH's investigated (3 – 11).

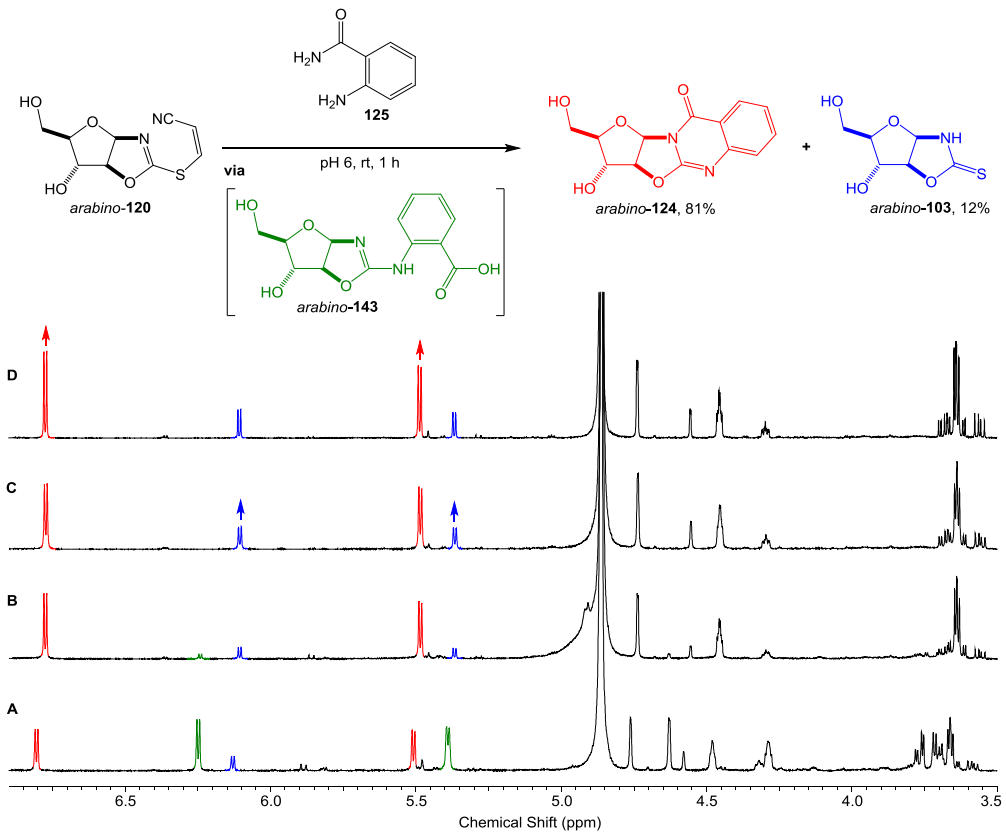
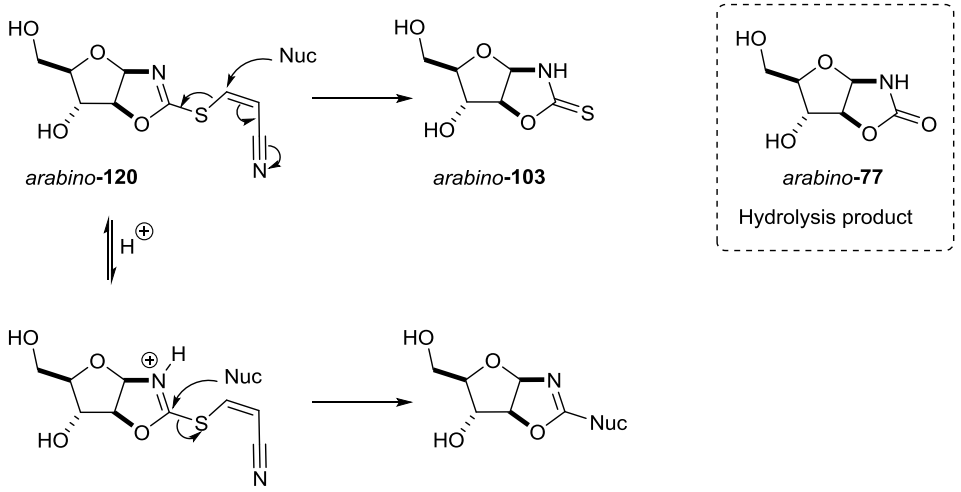


***Figure 2.13**: ^{1}H NMR spectra (600 MHz, D_2O, 3.5 – 6.9 ppm) showing reaction of S-cyanovinyl thione arbino-**120** with anthranilic acid **125**. **A)** Arabino-**120** (250mM) and **125** (500mM) at pH 6, incubated at rt for 1 h. **B)** After 6 h, calibration against internal standard (DSS) gave yields of 81% anthranilic quinazolinone arabino-**124** and 12% oxazolidinone thione arabino-**103**. **C)** Spiked with arabino-**103**. **D)** Spiked with arabino-**124**.*

The pH dependence of quinazolinedione **124** synthesis is thought to be due to the protonation of *N*1 nitrogen, facilitating C2 reactivity instead of conjugate addition to cyanovinyl group (Scheme 2.21). However, it was hypothesised that higher pKa amine nucleophiles would have the potential to add to unprotonated *S*-cyanovinyl

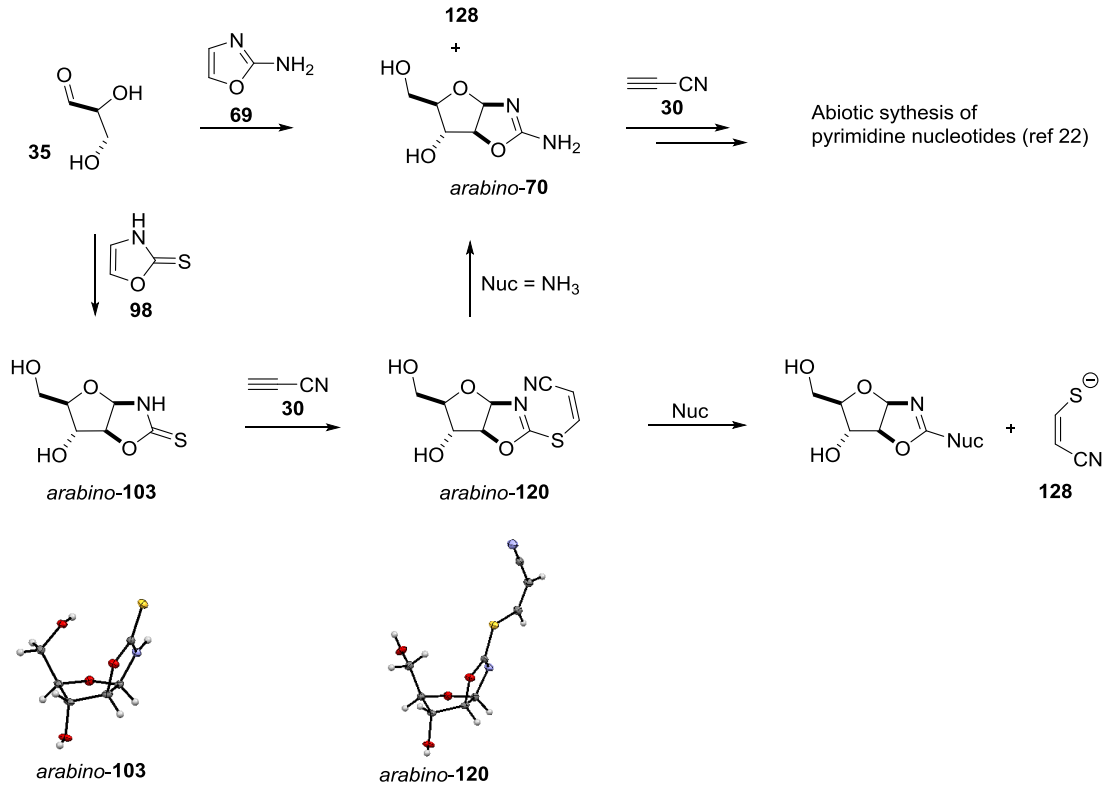
thione *arabino*-**120**, if so a second route to aminooxazoline **70** (a key compound in the prebiotic synthesis of pyrimidine ribonucleotides)[22] may be accessed and to test this hypothesis ammonia (NH_3) was next investigated in the displacement reaction of the thiolate groups.



Scheme 2.21: *Plausible mechanism for the pH dependence of nucleophilic addition to S-cyanovinyl oxazolidinone thione arabino-**120** and formation of oxazolidinone arabino-**77**, formed as the hydrolysis product during the nucleophilic additions to arabino-**120** in almost all reactions with varying rates. Hydrolysis product of nucleophilic additions arabino-**77** (inset).*

2.4.2. Displacement of the thiolate group by ammonia to form aminooxazoline

Ammonia (pKa 9.2),[92] as well as being itself a prebiotically plausible amine,[182] theoretically could displace *S*-cyanovinylated oxazolidinone thione *arabino*-**120** to yield aminooxazoline *arabino*-**70**. This would provide access to a key pyrimidine nucleotide precursor via an alternative route (which would be an exciting discovery on its own), however perhaps more importantly would open the potential for *arabino*-**120** acting as a point of divergence in the synthesis of both pyrimidine and purines (Scheme 2.22). The divergent synthesis of pyrimidine and purine ribonucleotides will be discussed in more detail, later in this thesis (Chapter 5).



Scheme 2.22: *Displacement chemistry of S-cyanovinylated oxazolidinone thione arabino-**120** and its possible central role divergent synthesis of pyrimidine and purine ribonucleotides. Desired stereochemistry of oxazolidinone thione arabino-**103** and arabinoe-**120** shown by X-ray crystallography.*

As mentioned earlier (section 2.3.2) *S*-methylated derivatives of oxazolidinone thione **126** were supposedly "singularly unreactive towards nucleophiles".[175] For example no reactivity with aqueous and methanolic ammonia has been observed. Unlike the case for simple *S*-alkyl thiones such as **126**, aqueous ammonia efficiently displaces thiolate **128** from *S*-cyanovinyl thione *arabino*-**120**. Pleasingly, we found that incubation of *arabino*-**120** or *ribo*-**120** (0.25M) with ammonium chloride (1M, pH 8.5-10.5) returns *arabino*- and *ribo* aminooxazoline **70** (15-21%), from their respective *S*-cyanovinyl thiones **120** (Figure **2.14**). Interestingly the major by-products are the precursor thione **103** and a white crystalline precipitate of dicyanovinyl sulfide **129**;[142] these by-products suggested regeneration of thione **103** results from rapid nucleophilic addition of thiolate **128** to the cyanovinyl moiety of **120**. However, the greatly increased efficacy of cyanovinylation of oxazolidinone thione **103** (Section 2.3) suggested that

S-cyanovinyl thione **120** could be regenerated *in situ* by the reactivation of **103** (in the presence of aminooxazoline **70**) by cyanoacetylene **30**.

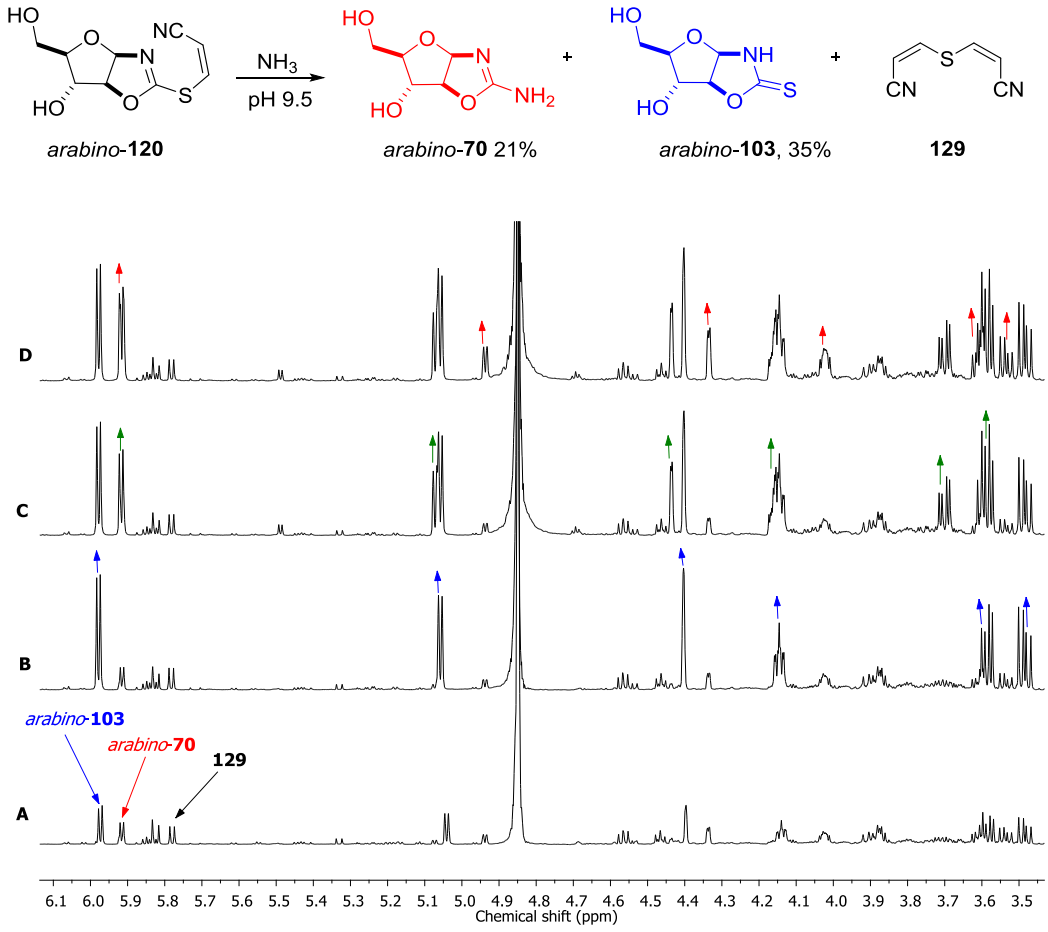


Figure 2.14: *Showing ^{1}H NMR (600 MHz, D_2O, 3.4 – 6.2 ppm) spectroscopic evidence by sequential spiking for formation of aminooxazoline arabino-**70** from the reaction of S-cyanovinyl thione arabino-120 and ammonia at pH 9.5. **A**) Arabino-**120** (250 mM) and ammonia (1M) were incubated at pH 9.5 at rt for 24 h. **B**) Spiked with oxazolidinone thione arabino-**103**. **C**) Spiked with oxazolidinone arabino-**77**. **D**) Spiked with aminooxazoline arabino-**70**.*

2.4.2.1. Regeneration of *S*-cyanovinylated thione in ammonia displacement reaction

The successful displacement of the thiolate group **128** by ammonia shows that pyrimidine nucleotides can (in principle) be synthesised by our strategy, giving us further momentum to investigate purine synthesis. However, the main by-product of ammonia displacement reaction was oxazolidinone thione *arabino*-**103**, which likely

arises from Michael addition of ammonia to the cyanovinyl group. The thiolate displacement reaction also generates the nucleophilic thiolate **128**, which can form *arabino*-**103** via Micheal addition and lead to the formation of the observed (second major by-product) *β*-*β*-dicyanovinyl-thioether **129**,[142] which precipitates form the solution. We hypothesised that addition of cyanoacetylene **30** solution to the reaction of *S*-cyanovinylated oxazolidinone thione *arabino*-**120** with ammonia would trap the thiolate **129** to form **128** and also regenerate the starting material *arabino*-**120** by further reacting with the thione *arabino*-**103**.

Interestingly, we observed that addition of cyanoacetylene **30** (0.25 M) to an aqueous solution of oxazolidinone thione *arabino*-**103** (0.24 M) and aminooxazoline *arabino*-**70** (0.24 M) between pH 7 and 10.5, led to chemospecific cyanovinylation of *arabino*-**103**, which then reacted with ammonia to yield aminooxazoline *arabino*-**70** (60%, with respect to *aranbio*-**103**) (Figure 2.15). Furthermore, we observed that repeated cyanovinylation and incubation of thione *arabino*-**103** in ammonia solution yielded up to 45% aminooxazoline *arabino*-**70** (over two cycles of cyanovinylation at pH 10.5, without need for any intermediate steps of purification) leading to a remarkably pure solution of aminooxazoline *arabino*-**70** (Figure 2.15). The process of adding cyanoacetylene **30** can be continued for further cycles to completely consume all of the oxazolidinone thione *arabino*-**103**. The regeneration experiment was also repeated with *ribo*-**103** to yield 33% *ribo*-**70**, *ribo*-**70** was observed to spontaneously crystallize from the reaction mixture after two cycles of cyanovinylation and ammonolysis.[134] As it has been demonstrated that aminooxazolines *arabino*-**70**[22, 123] and *ribo*-**70**[21, 134] are both key intermediates *en route* to pyrimidine ribonucleotides this additional synthetic strategy further augments the potential of aminooxazolines **70** as a prebiotically plausible precursor of RNA.

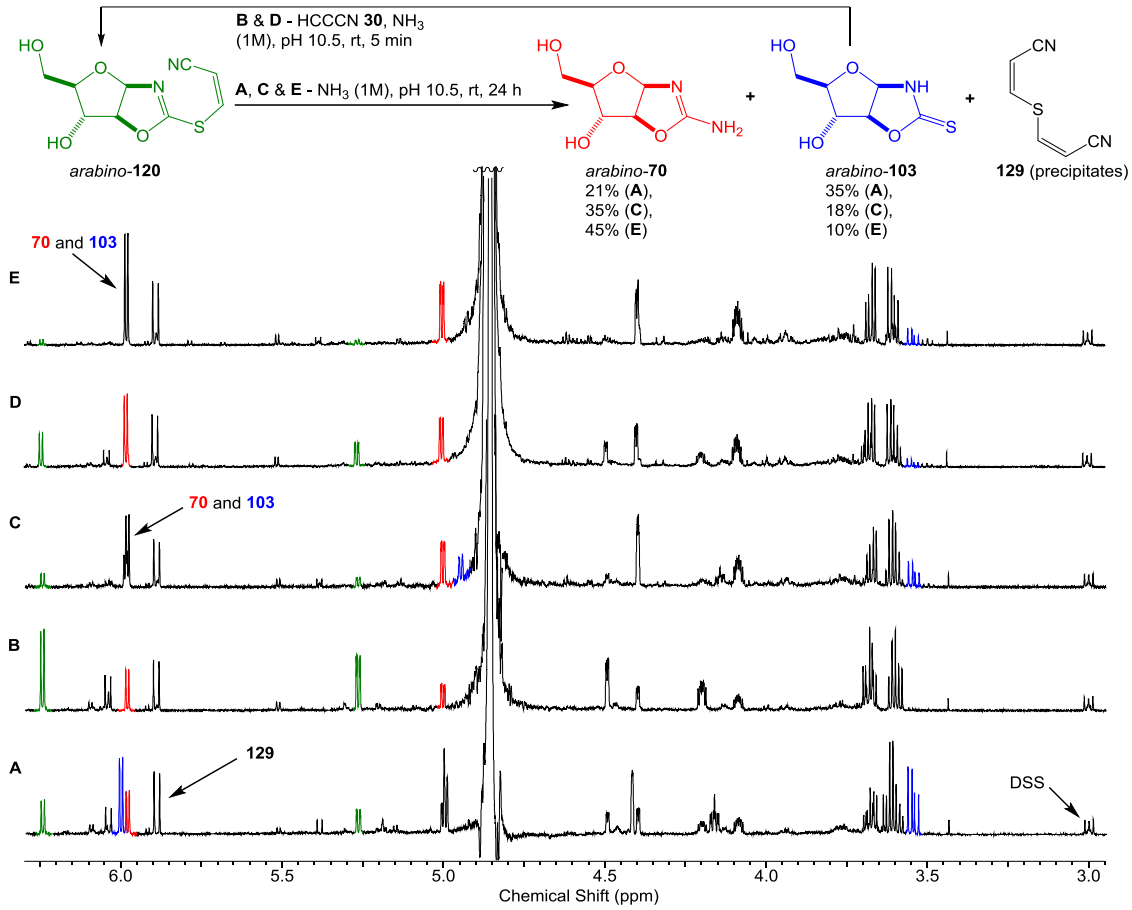


Figure 2.15: *^{1}H NMR spectra (600 MHz, 2.95 – 6.3 ppm) showing reaction of S-cyanovinyl oxazolidinone thione arabino-**120** with ammonia (NH_3). **A)** Arabino-**120** (250 mM) and NH_3 (1M) were incubated at pH 10.5 at rt for 24 h. **B)** Oxazolinone thione arabino-**103** was quantitatively converted back to arabino-**120** by addition of cyanoacetylene **30**. **C)** After a further 24 h at pH 10.5 at rt. **D)** After a 2nd recyanovinylation by addition of **30**. **E)** After another 24 h at pH 10.5 at rt.*

2.4.3. Displacement of the thiolate group by cyanamide

The pH dependence of the reaction suggested that nucleophile's pKa will be highly important in establishing efficient displacement of the thiolate **128**. It would seem that the low pKa of aniline nitrogen (pKa 4.8) is ideally suited to retaining nucleophilicity where thione nucleophilic displacement is specific acid catalysed. To test the hypothesis that the protonation state of the nucleophile is highly important, we next investigated cyanamide **68** (pKaH = 1.1). We observed that cyanamide **68** (1M) reacted with S-cyanovinylated oxazolidinone thione *arabino*-**120** (250mM) most efficiently at pH 6.5 to yield 26% aminooxazoline **144** (Figure 2.16).

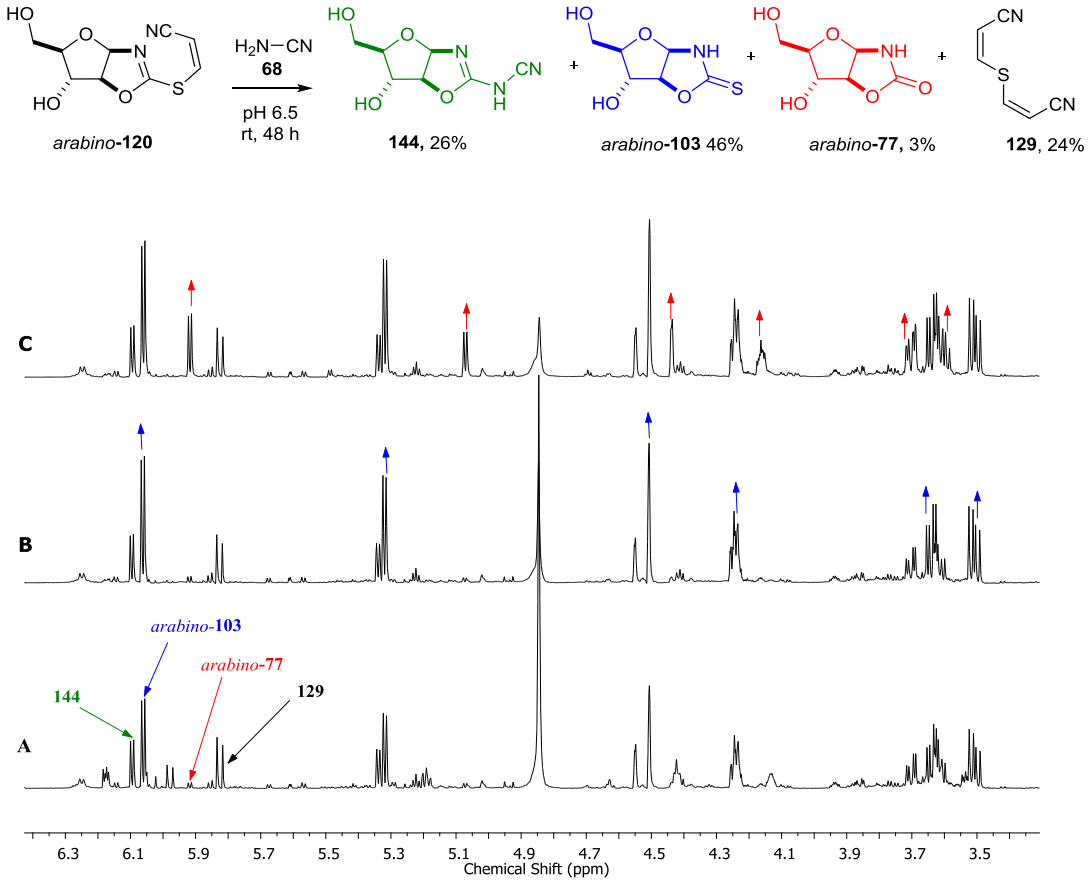


***Figure 2.16**: [1]H NMR spectra (600 MHz, 3.3.3 – 6.37 ppm) showing evidence by sequential spiking for formation of aminooxazoline **144** from the reaction of S-cyanovinylated oxazolidinone thione arabino-**120** and cyanamide **68** at pH 6.5. **A)** The crude reaction of cyanamide **68** (1M) and arabino-**120** (250mM) at pH 6.5 in rt after 48 h. **B)** Spiked with oxazolidinone thione arabino-**103**. **C)** Spiked with oxazolidinone arabino-**77**.*

However, the yield considerably decreased in lower pH. At pH 4.3 we obtained 6% aminooxazoline **144**, 75% oxazolidinone *arabino*-**77** and 19% *S*-cyanovinylated oxazolidinone thione *arabino*-**120**. This suggested that at lower pH hydrolysis outcompeted the addition of cyanamide **68** and as a result mostly oxazolidinone *arabino*-**77** formed by competing hydrolysis of water. Moreover, the lone pair in cyanamide **68** is highly conjugated into the nitrile group, as indicated by its very low pKa, this electron withdrawing effect have inhibited competition with hydrolysis. This further suggested that it would be essential to find the correct pKa of amine, to balance reactivity of the material (*S*-cyanovinylated oxazolidinone thione *arabino*-**120)**, the desired nucleophile and water. If the pH of the reaction is too low hydrolysis takes

over and the major product will be oxazolidinone *arabino*-**77**. If the pH is too high then the thiolate **128** adds to the cyanovinyl group in *arabino*-**120** via Michael addition giving oxazolidinone thione *arabino*-**103** as the major product. Therefore, the interplay of reaction pH and nucleophile pKa appears to be very important. It seems likely that the availability of the amine nucleophile's nitrogen lone pair is essential to achieve efficient displacement, and ideally we require an available nitrogen lone pair at a pH where the electrophilic thione is activated by (partial) protonation. To test this hypothesis we next chose to focus on our primary objective of purine ribonucleotide synthesis and the addition of aminonitrile nucleophiles.

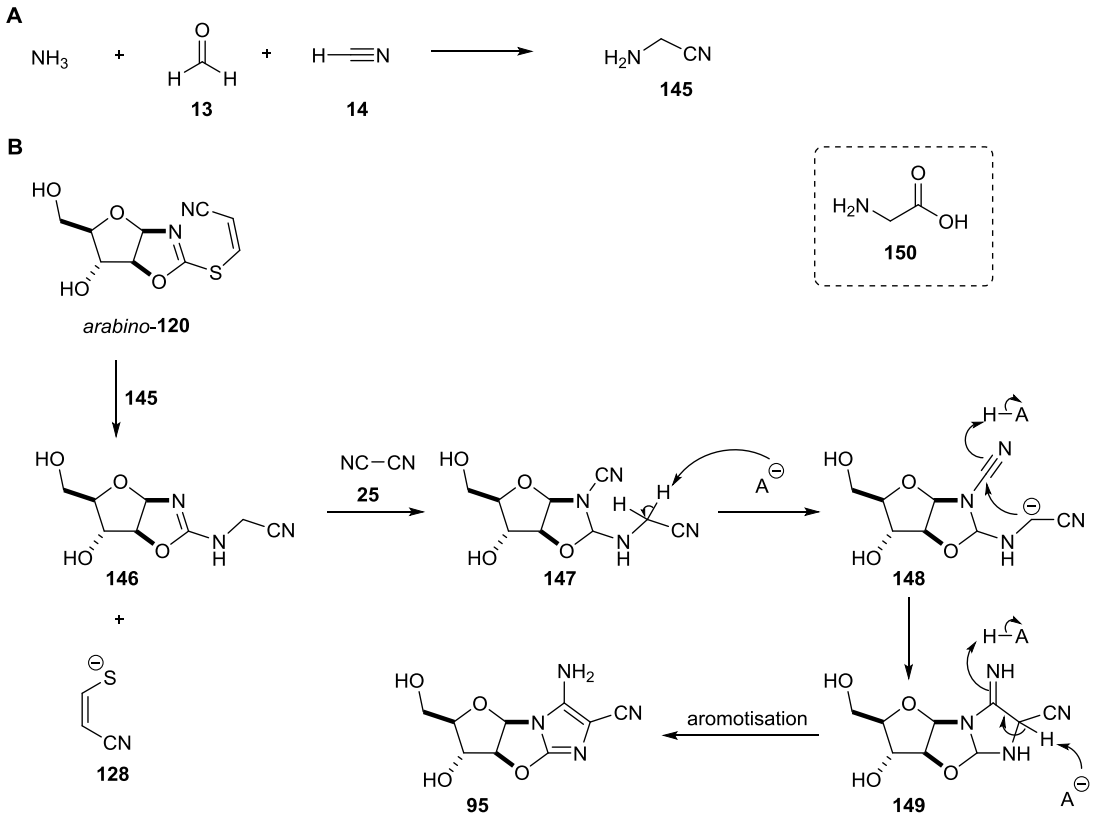
2.4.4. Displacement of the thiolate group by aminonitriles

The observations and experimental analysis that was gained from the reaction of different nucleophiles led us to believe that we were at a stage to tether hydrogen cyanide **14** derivatives. As previously mentioned these are the potential nucleophiles which will help us achieve the abiotic synthesis of purines. HCN **14** can easily be formed from electric discharge experiments[43, 101] and is present in our solar system, comets and atmosphere of other planets.[183-185] Furthermore, the oligomerisation of hydrogen cyanide **14** has widely been proposed as a key route to purine nucleobases,[46, 59, 96, 146, 186] and guided by the recently reported synthesis of purine precursors from aminonitriles,[59] we recognised that hydrogen cyanide derivatives,[51, 52, 156, 187] have the ideal ambident reactivity to displace cyanovinyl thiolate **128** in *S*-cyanovinyl thione *arabino*-**120** and then induce cyclisation in the aminooxazolines to build the imidazole moiety of purine nucleobases upon a sugar scaffold tethered by the 8,2´-anhydro-linker required for phosphorylation and C2´-stereochemical inversion.[22, 188]

2.4.4.1. 2-Aminoacetonitrile (glycine nitrile)

Investigation into the nucleophilicity of aminonitriles was started by studying glycine nitrile **145**, which can easily be synthesised from plausible prebiotic molecules (NH_3, H_2CO **13** and HCN **14**) via a Strecker reaction (Scheme 2.23A). It is also a product of Urey-Miller discharge experiment and Sutherland's cyanosulfidic protometabolism.[38, 89] In 2008 it was also discovered that glycine nitrile **145** was present in the interstellar cloud (Sagittarius B2).[189] Furthermore, Commeyras *et al.* have reported extensive

kinetic and thermodynamic studies of glycine nitrile **145** to get a better understanding of the plausible prebiotic synthesis of the proteinogenic amino acid glycine **150**, the first and (likely) the most abundant amino acid on the primitive Earth.[190]



Scheme 2.23: A) *Abiotic synthesis of glycine nitrile* ***145*** *from Strecker reaction of formaldehyde* ***13****, ammonia (*NH_3*) and hydrogen cyanide* ***14****.* ***B)*** *Plausible mechanism for reaction of S-cyanovinylated oxazolidinone thione arabino-****120*** *with glycine nitrile* ***145*** *to give aminooxazoline* ***146*** *and subsequent N1-cyanation of* ***146*** *with cyanogen* ***25*** *to give compound* ***147****, leading towards the synthesis of potential purine precursor* ***95*** *via intramolecular cyclisation. Structure of proteinogenic amino acid glycine* ***150*** *(inset).*

Given that the *α*-proton of glycine nitrile **145** (~pKa = 20)[191] is relatively acidic, we conjectured that the deprotonation of α-carbon will open up various functionalisation opportunities, such as carboxylation, which may be exploited *en route* to purine elaboration.[192] However, structural similarity of aminooxazoline **146** and aminooxazoline *arabino*-**70**, suggested that electrophiles would preferentially react at *N*1 of **146** in water. We based this hypothesis on the selective *N*1-cyanovinylation of aminooxazoline *arabino*-**70** in the synthesis of pyrimidines.[22] Interestingly, if

cyanogen **25**, by analogy to cyanoacetylene **30** was observed to furnish *N*1 cyanation of **146**, formation of cyanated aminooxazoline **147** would be expected, then intramolecular cyclisation of **147** would yield the AICN tricyclic cyclic structure **95**, a key intermediate in our purine ribonucleotide synthesis (Scheme 2.23B). This would then provide a way to by-pass HCN **14** oligomerisation chemistry, which despite being one of the most cited and referenced processes in chemical origins of life, is a very sluggish reaction that yields a large array of discrete products and polymers.[90, 91 83]

When *S*-cyanovinylated oxazolidinone thione *arabino*-**120** (250mM)) was reacted with glycine nitrile **145** (500mM) at pH 4.2 in water for 24 h, we successfully obtained 52% aminooxazoline **146**, as well as 34% thione *arabino*-**103**, and 14% oxazolidinone *arabino*-**77**. We confirmed the presence of thione *arabino*-**146** and oxazolidinone *arabino*-**77** by spiking the crude ^{1}H NMR with authentic samples which were synthesised from known literature procedures. We already knew that at lower pH, hydrolysis of *arabino*-**120** was more rapid. Therefore, to optimise the addition of glycine nitrile **145** and minimise hydrolysis the reaction was repeated at pH 5.2. Interestingly, at pH 5.2 the reaction furnished 70% aminooxazoline **146**, 13% oxazolidinone thione *arabino*-**103**, and 14% the hydrolysis product *arabino*-**77** (Figure 2.17). At this pH *N*1 nitrogen of *S*-cyanovinyl thione *arabino*-**120** would be partially protonated (pKa of *arabino*-**120** = 5.2) and a significant proportion of glycine nitrile **145** (pKa = 5.6) would not be protonated. To further test if protonation of the *N*1 in *arabino*-**120** or the protonation state of the amine nucleophile is important to displace the *S*-cyanovinyl thiolate **128**, we repeated the reaction of glycine nitrile **145** with *arabino*-**120** at pH 7.0 and 10.0. These displacement reactions furnished 65% (pH 7.0) and 52% (pH 10.0) aminooxazoline **146**. The result confirms that for optimum nucleophilic substitution at the C2 carbon atom of *arabino*-**120**, the pH of the reaction needs to be at a point where *N*1 is protonated, but there should also be a portion of the amine nucleophile unprotonated. However, this pH window of optimum reactivity will vary for other amine nucleophiles as we have seen in ammonia (pKa = 9.2), cyanamide **68** (pKa = 1.1) and anthranilic acid **125** (pKa = 2.14).

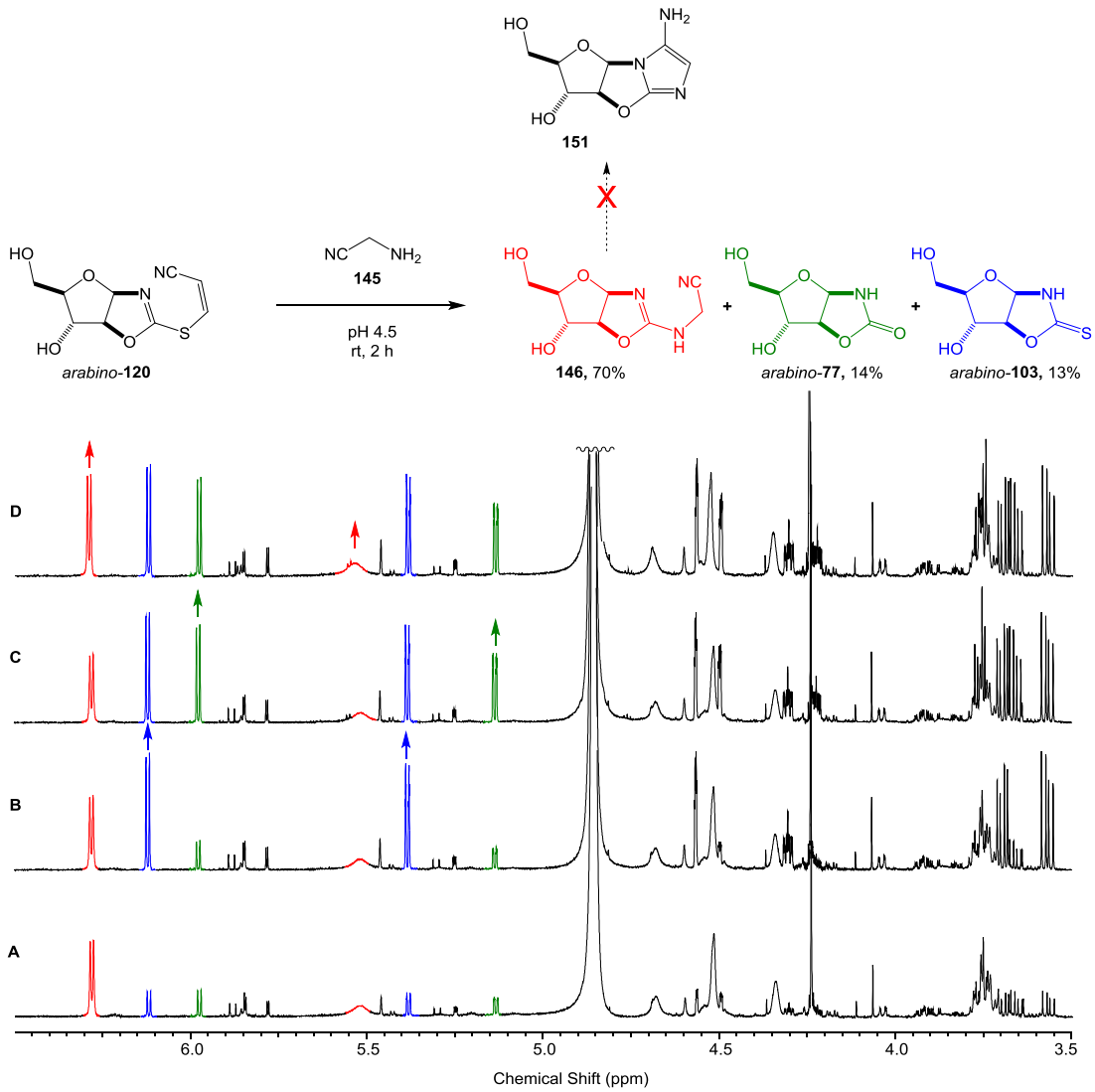


Figure 2.17: *^{1}H NMR spectra (600 MHz, D_2O, 3.5 – 7.0 ppm) for reaction of S-cyanovinyl oxazolidinone thione arabino-**120** with glycine nitrile **145**. **A)** Arabino-**120** (250mM) and **145** (500mM) in D_2O (500 µL) at pH 4.5, incubated at rt for 2 h. Calibration to an internal standard (DSS) gave yields of 70% arabinofuranosyl-N-acetonitrile-aminooxazoline **146**, 13% oxazolidinone thione arabino-**103** and 14% oxazolidinone arabino-**77**. **B)** Spiked with arabino-**103**. C) Spiked with arabino-**77**. **D)** Spiked with **146**.*[a]

After observing the successful reaction of glycine nitrile **145** with *S*-cyanovinyl thione *arabino*-**120**, next the reaction of **145** with *S*-methyl thione *arabino*-**126** was investigated. It was hypothesised that the reaction of **145** with *arabino*-**126** would be more feasible at a lower pH given the stability of *arabino*-**126** over *arabino*-**120**

[a] Pure sample of **146** for NMR spiking was obtained from Dr Shaun Stairs in the Powner laboratory.

(Section 2.3). Indeed incubation of *arabino*-**126** with glycine nitrile **145** at pH 4.5 at room temperature give 76% product **146** and 8% oxazolidinone arabino-**77**.

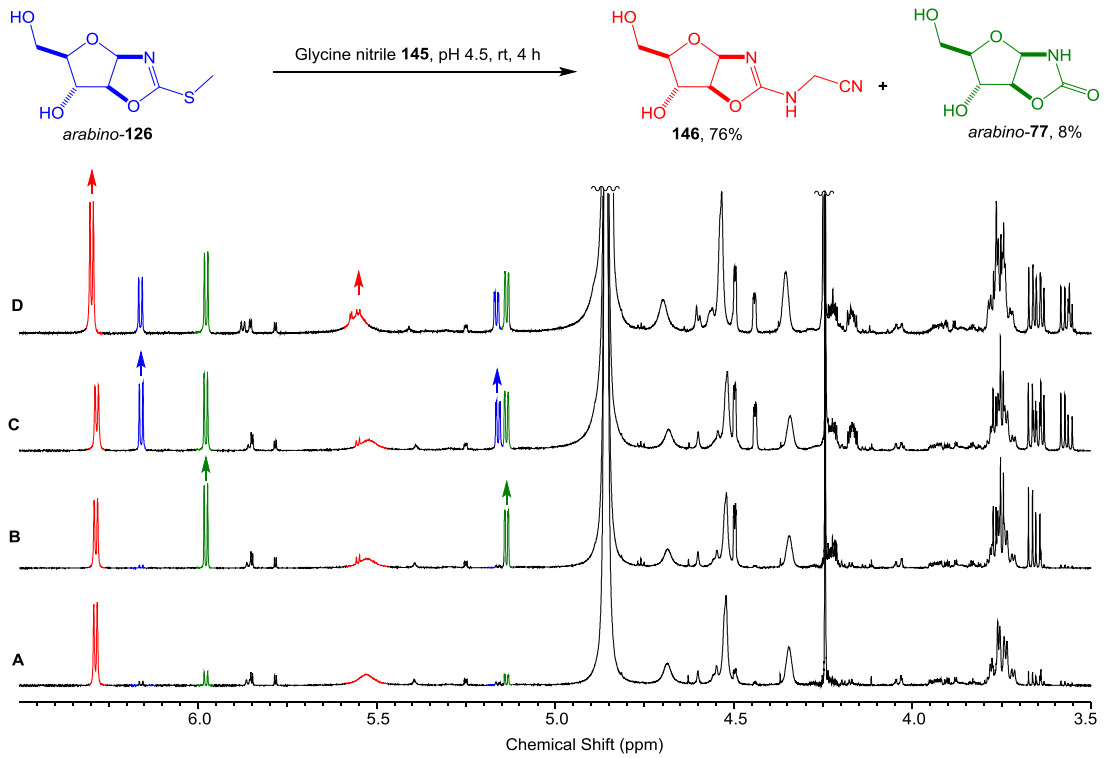
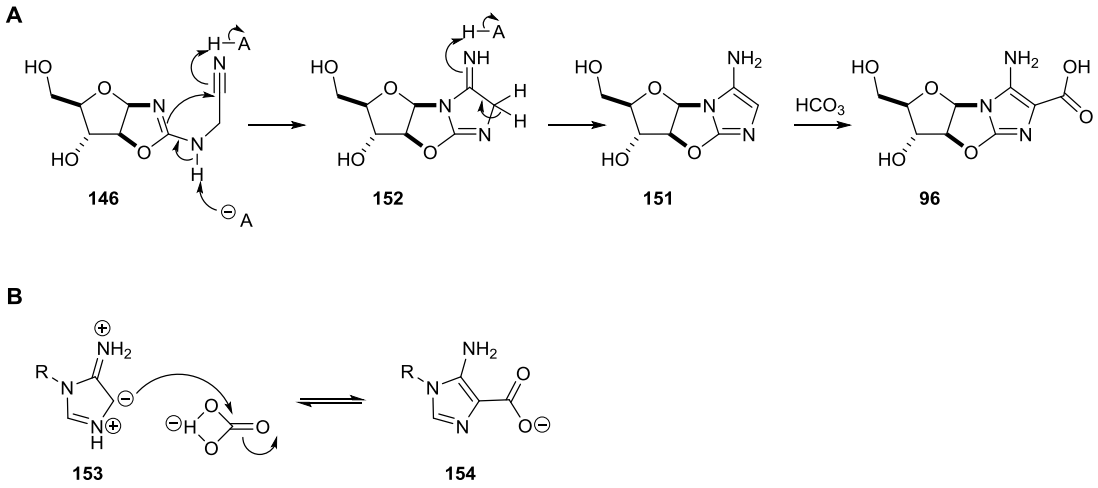


Figure 2.18: *^{1}H NMR spectra (600 MHz, D_2O, 3.5 – 7.0 ppm) showing reaction of S-methylated oxazolidinone thione arabino-**126** with glycine nitrile **145**. **A)** Arabino-**126** (250mM) and **145** (500mM) at pH 4.5, incubated at rt for 4 h. Calibration to an internal standard (DSS) gave yields of 76% aminooxazoline **146** and 8% oxazolidinone arabino-**77**. **B)** Spiked with arabino-**77**. **C)** Spiked with arabino-**126**. **D)** Spiked with **146**.*

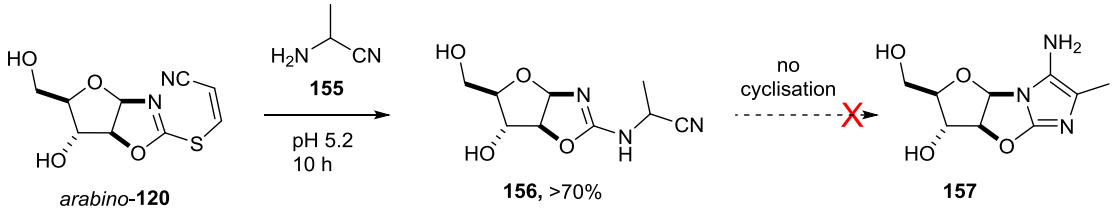
As hypothesised (Scheme 2.23), *N*1-cyanated product of aminooxazoline **146** will furnish an important prebiotic purine precursor AICN tricyclic moiety **95**. However, before taking on the more complex task of investigating cyanation of *N*1, attempts to cyclise aminooxazoline **146** were undertaken. Initially, the intramolecular cyclisation of **146** would be expected to give an imine **152**, which will significantly drop the pKa of the α-protons to produce an enamine **151** also aided by aromatisation. Enamine **151** could be envisaged to be a potential nucleophile (for example Shaw *et al.* reported the reversible carboxylation of amino imidazole **153** to form **154)** and would be expected to undergo carboxylation at C4 carbon atom to furnish the AICA tricyclic moiety **96**, (Scheme 2.24).[192, 193] Therefore, to test how aminooxazoline **146** reacted, we left the reaction mixture at pH 5.2 for 48 h but no visible change was observed. We hypothesised that probably the *N*1 protonation of **146** may be preventing the

intramolecular cyclisation (as the pKa of aminooxazoline = ~ 6.5), therefore the pH of the crude mixture was raised to pH 8.0 and then incubated at room temperature for another 48 h. Surprisingly, once again no evidence of cyclisation was observed from ^{1}H NMR spectroscopic analysis.



Scheme 2.24: A) *Postulated mechanism for cyclisation of aminooxazoline* ***146*** *to furnish the prebiotically plausible purine precursor* ***151*** *followed by prebiotic carboxylation based on Shaw et al. studies to yield AICA tricyclic moiety* ***96****.* ***B)*** *Mechanism for aqueous carboxylation of aminoamidazoles* ***153*** *in potassium hydrogen carbonate by Shaw et al.*[192, 193]

In order to test if Thorpe-Ingold effect[194] contributed to the cyclisation of aminooxazoline **146** reactions of alanine nitrile **155** and *S*-cyanovinylated oxazolidinone thione *arabino*-**120** were explored.[b] Interestingly alanine nitrile **155** (250 mM) successfully tethered with *arabino*-**120** (500mM) to yield aminooxazoline **156**, however repeated attempts at various pH and temperature conditions to cyclise **156** again failed to furnish the tricyclic compound **157**.



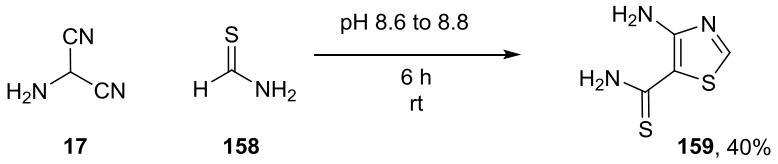
[b]This work was carried out by Dr Shaun Stairs in the Powner laboratory.

Scheme 2.25: *Formation of aminooxazoline* ***156*** *from the reaction of S-cyanovinylated oxazolidinone thione arabino-****120*** *and alanine nitrile* ***155****.* ***156*** *failed to cyclise to give the tricyclic compound* ***157****, suggesting that Thorpe Ingold effect does not contribute to cyclisation in aminooxazolines.*

The reactions of amino nitriles gave us a very positive indication that we were heading in the right direction in our pursuit of finding a lead compound to prebiotically synthesise purine nucleotides. The missing link, however, was cyclisation of the aminonitriles to give us the important tricyclic compounds (**95** and **96**). We next hypothesised that substitution of tetrahedral centre with an electron withdrawing groups might induce desired cyclisation and the synthesis of the 5-membered imidazole ring. Accordingly, the HCN trimer **17** was perceived to be the mostly likely option to marry our desired reactivity with prebiotic plausibility.

2.4.4.2. HCN trimer and tetramer

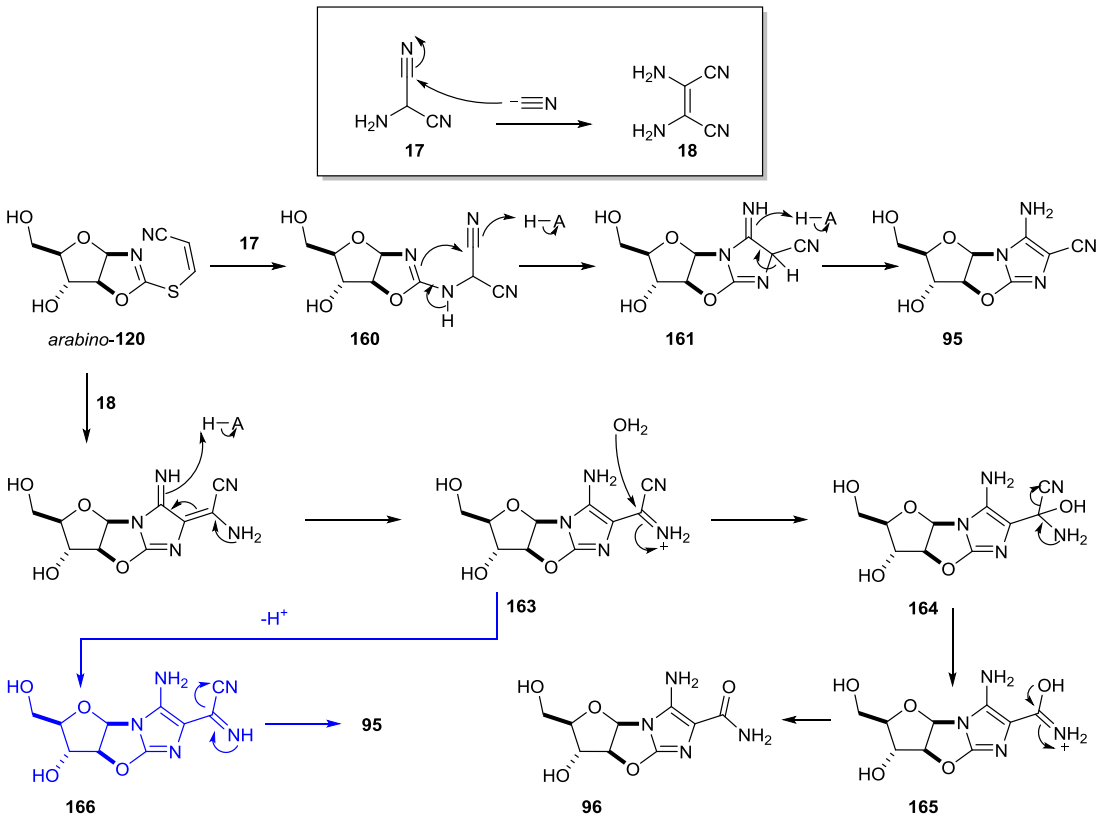
HCN oligomers, amino malononitrile (trimer) **17** and diamino malononitrile (tetramer) **18** have already been used to synthesise purine precursors by Orgel and are observed intermediates in the formation of adenine by cyanide oligomerisation (Scheme 1.7).[46, 52, 59, 96, 186] Orgel *et al.* have also shown that slow addition of thioformamide **158** buffered solution (0.1M, pH 8.75) to HCN trimer **17** (0.01M) yields 40% isochrysean **159** (Scheme 2.26).[52]



Scheme 2.26: *pH controlled reaction of aminomalononitrile* ***17*** *with thioformamide* ***158*** giving 40% isochrysean **159** .[52]

The trimer **17** is constitutionally the simplest nucleophile to tether to C2 carbon atom of *S*-cyanovinylated oxazolidinone thione *arabino*-**120** and would lead towards the core AICA moiety of purine nucleotides. We predicted that the addition product of the trimer aminooxazoline **160** would readily undergo intramolecular cyclization to give the AICN derived product **95**. We not only expected the electrophilicity of the nitrile moieties to be increased in **160** with respect to **146**, but also recognised that

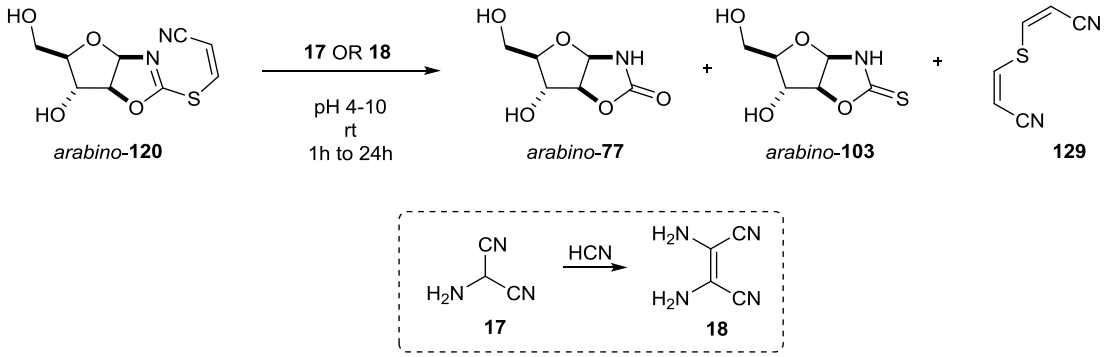
positioning two nitrile moiety on the same carbon atom would statistically increase the reactivity of both nitriles toward cyclisation (Scheme 2.27).



Scheme 2.27: *Tentative mechanism for reaction of S-cyanovinyl oxazolidinone thione arabino-**120** with aminomalononitrile **17** and diaminomalononitrile **18**. Intramolecular cyclisation of aminooxazoline **160** can lead to AICN tricyclic **95**. Addition of HCN tetramer **18** can produce both key purine precursors AICN tricyclic moiety **95** and AICA tricyclic moiety **96** (blue). Oligomerisation of **17** into **18** (inset).*

Using the same conditions that were employed in the reaction of other nucleophiles with *S*-cyanovinyl thione *arabino*-**120,** trimer **17** (500mM) was reacted with *arabino*-**120** (250mM) at pH 4. However, disappointedly, we were unable to see any new product derived from **17**. The reaction was repeated at pH 7 and 10, but again to our disappointment we did not observe any reaction between *arabino*-**120** and HCN trimer **17**, instead only oxazolidinone *arabino*-**77** and oxazolidinone thione *arabino*-**103** were observed. Presence of *arabino*-**77** and *arabino*-**103** was confirmed by spiking with authentic samples. The absence of reactivity of the trimer **17** addition reaction was suspected to be due to its half-life (10^{-3}M, pH 9, rt, 1 h).[52] We suspected that in our reactions the trimer **17** hydrolysed or decomposed prior to its addition to the C2 carbon atom of *arabino*-**120**. It was therefore decided to follow Orgel's method of slow addition of HCN trimer **17** to the *S*-cyanovinylated thione *arabino*-**120** as this

should give enough time for the HCN oligomer **17** to react. But, when aminomalononitrile **17** solution (10mM) was added to *arabino*-**120** (100mM) at pH 4.0 over a 6 h period, however, again we were still unable to detect any displacement by the HCN trimer **17**, the two major products were again oxazolidinone *arabino*-**77** and oxazolidinone thione *arabino*-**103** (Scheme 2.28).



Scheme 2.28: *Reaction of S-cyanovinyl oxazolidinone thione arabino-**120** with HCN trimer **17** and tetramer **18** in various pH and time scales returned oxazolidinone arabino-**77**, oxazolidinone thione arabino-**103** and β-β-dicyanovinyl-thioether **129.** No observable addition reaction of **17** and **18** with arabino-**120** was detected.*

The HCN tetramer **18** is a classic of prebiotic chemistry. Diamino malononitrile **17** is the first stable and isolable intermediate of HCN **14** oligomerisation chemistry. We predicted that once the tetramer **18** is tethered to the C2 carbon atom centre of *S*-cyanovinyl oxazolidinone thione *arabino*-**120**, it will undergo a similar intramolecular cyclisation as the trimer **17**, however, given that the tetramer **18** is relatively more complex, we proposed that both AICN **95** and AICA **96** derived sugar moieties will be synthesised via our proposed mechanism (Scheme 2.27). Therefore, we applied the same reaction conditions that we have been using in the reaction of other nucleophiles. However, we were unable to detect any new product from the reaction of the tetramer **17** at pH 4, 7 or 10, and were only able to confirm the presence of oxazolidinone thione *arabino*-**103**, oxazolidinone *arabino*-**77**, and dicyanovinyl-thioether **129** (Scheme 2.28).

After repeated attempts to displace the (more stable) *S*-cyanovinyl thione *arabino*-**120** with HCN trimer **17** and tetramer **18**, we decided to use the *S*-methyl thione *arabino*-**126** in the displacement reactions (Figure 2.19).

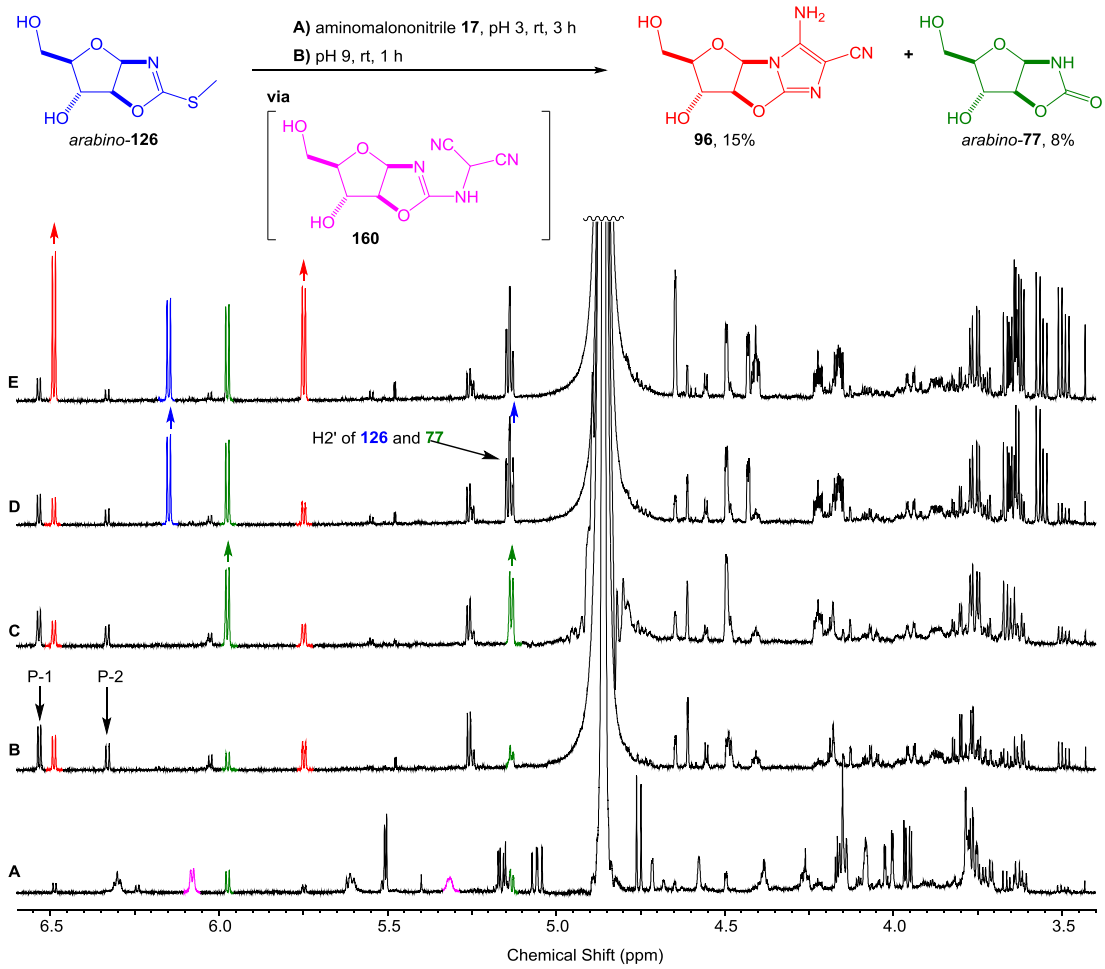


Figure 2.19: *^{1}H NMR spectra (600 MHz, D_2O, 3.4 – 6.6 ppm) showing reaction of S-methylated oxazolidinone thione arabino-**126** with aminomalononitrile **17**. **A)** Arabino-**126** (250mM) and aminomalononitrile p-toluenesulfonate **17** (1M) at pH 3 and rt incubated for 3 h. Aminooxazoline **160** was produced in a 15% yield. **B)** Aliquot (50 µL) from **A** was added to ammonium hydroxide in D_2O (450 µL, 100mM), the solution was adjusted to pH 9 with NaOH (4M) and incubated for 1 h at rt. By-products P-1 and P-2 are derived from arabino-**126** and observed upon incubation of arabino-**126** alone under comparable conditions. P-1 and P-2 are suspected to be the products of (intermolecular) S-N methyl migration. Calibration to an internal standard gave yields of 15% aminoimidazole-4-carbonitrile-β-furanosylarabinoside **95** and 8% oxazolidinone arabino-**77**. **C)** Spiked with arabino-**77**. **D)** Spiked with arabino-**126**. **E)** Spiked with **95**.*

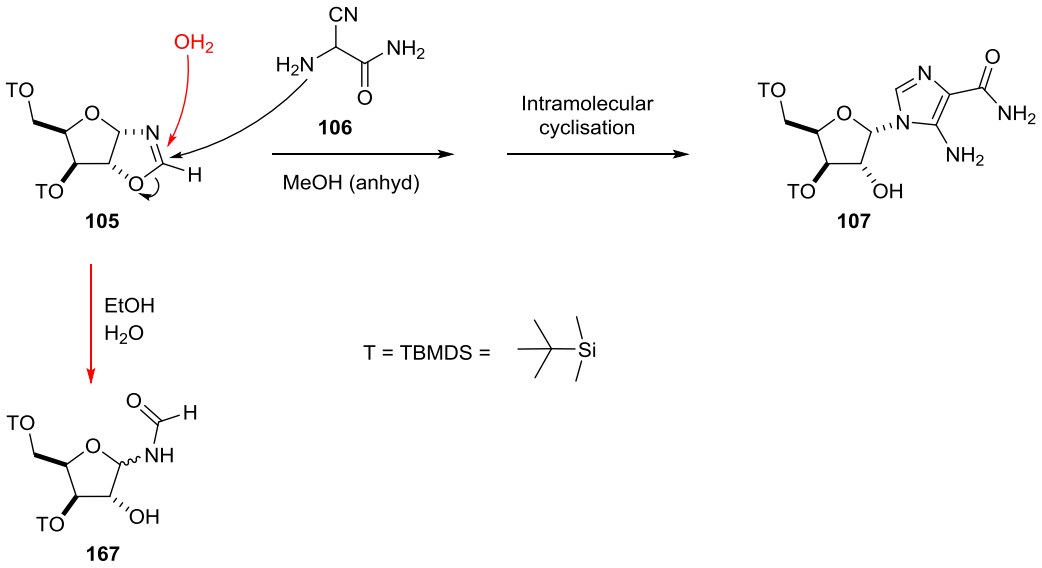
It was hypothesised that the low pH required for the sustainability of the HCN trimer **17** would also be decreasing the half-life of *arabino*-**120**, and knowing that *arabino*-**126** has a longer half-life than *arabino*-**120** in acidic conditions, the reaction of HCN trimer **17** and *S*-alkyl thione *arabino*-**126** was investigated. To our surprise, and delight incubation of thione *arabino*-**126** (250mM) and aminomalononitrile **17** (1M) at pH 3 in room temperature for 3 h give 15% of aminooxazoline **160** which

quantitatively cyclised to the AICN moiety **95** in pH 9 (Figure 2.19). The identity of AICN was confirmed by spiking with authentic sample of AICN tricyclic **95** (prepared by Dr Shaun Stairs through derivatisation of *arabino*-inosine).[195]

The prebiotic synthesis of AICN tricyclic moiety **95** meant that we were heading in the right direction, as **95** was one of our proposed tricyclic structure in our project aims, leading to the synthesis of 8-oxo purine ribonucleotides (Section 1.9). However, the 15% yield of **95** was low and therefore it was decided to investigate the displacement reaction of 2-amino-2-cyanoacetamide **106**, the hydrolysis product of aminomalononitrile **17**.

2.4.4.4. 2-Amino-2-cyanoacetamide

Given our observed stability problem with HCN **14** oligomers in displacement reaction of *S*-cyanovinyl oxazolidinone thione *arabino*-**120**, and the unsuccessful cyclisation attempts of the glycine nitrile product **146**, we next looked to investigating other HCN-oligomer derivatives with a lower propensity to degrade by liberating HCN before reacting with *arabino*-**120**. This immediately suggested 2-amino-2-cyanoacetamide **106**, which is the first hydrolysis product of the HCN trimer **17** and is also significantly more stable than **17**.[196] Furthermore, Imbach *et al.* have used **106** as an ambident nucleophile/electrophile in the conventional organic synthesis of *xylo*furanosyl nucleosides (Scheme 2.29 and 2.1).[156] However, Imbach and co-workers reported that 3´,5'-bis-*O*-(*tert*-butyldimethylsilyl) oxazoline **105** "readily hydrolyses" to give *N*-formyl-*xylo*furanosylamine **167**, and therefore suggested it is essential the reaction is carried out in anhydrous methanol. Irrespective of the fact that this synthesis is not abiotically constrained, and all stereochemistry relationships in the product are incorrect for our purposes, the observed reactions reported by Imbach and co-workers suggests, the role of cyanoacetamide **106** in our own chemistry is worth investigating.



Scheme 2.29: *Utilisation of 2-amino-2-cyanoacetamide* ***106*** *in the conventional systematic synthesis of xylofuranosyl nucleosides .*[156]*Under dry conditions* ***106*** *attacks the C2 carbon atom of* ***105*** *to make* ***107****, however in aqueous conditions hydrolysis takes over the addition reaction leading to formation of* ***167*** *(red).*

It was initially found that no reaction between *S*-cyanovinyl oxazolidinone thione *arabino*-**120** and 2-amino-2-cyanoacetamide **106** was observed at pH 7 and pH 10. However treatment of **106** (500mM)) with *arabino*-**120** (250mM) at low pH furnished the desired aminooxazoline **168**. The yield of **168** product was observed to be 25% and 52% at pH 5 and 4.5, respectfully. Mass spectroscopic analysis also showed the correct mass for the desired aminooxazoline **168** and ^{1}H NMR spiking confirmed that it was indeed new product **168** (Figure 2.20).

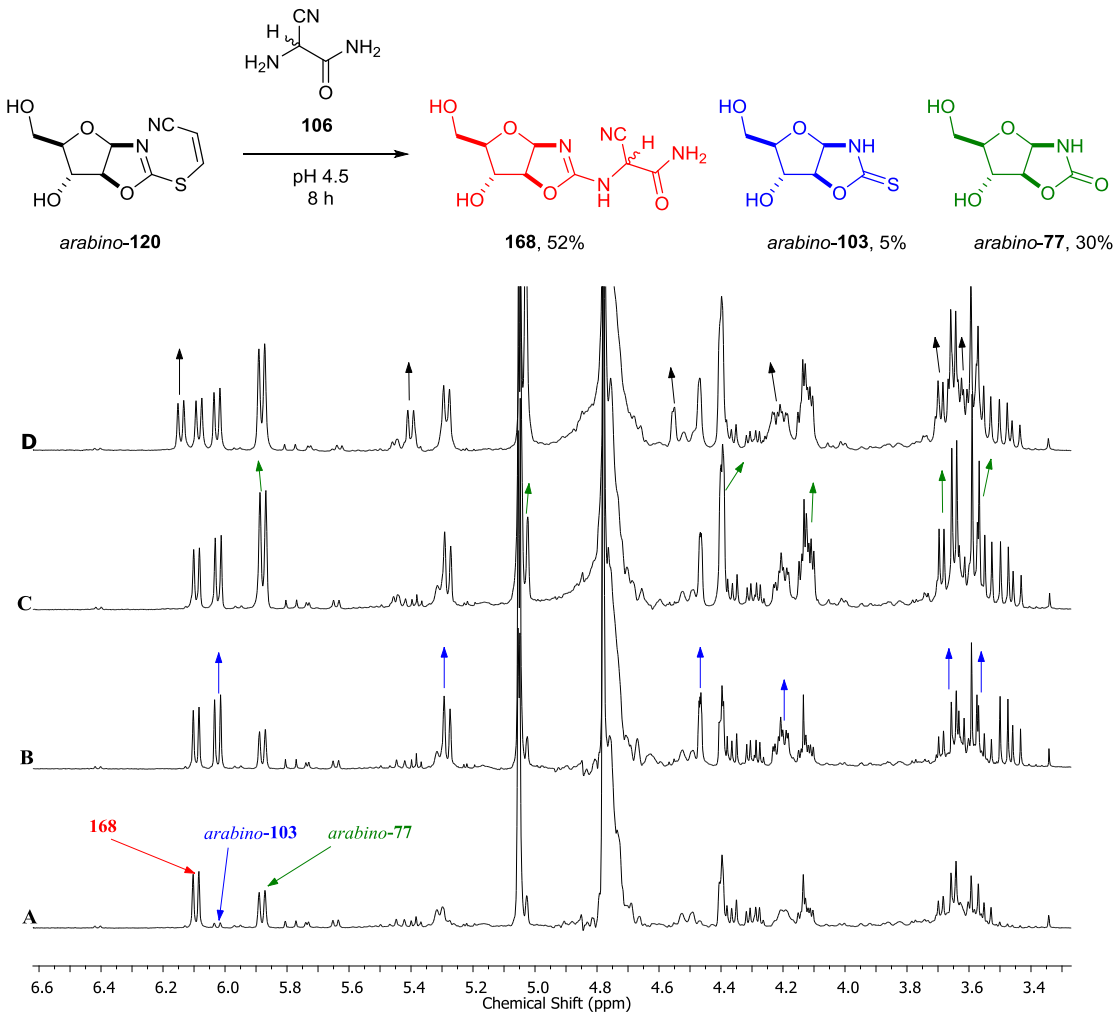


Figure 2.20: *Showing ^{1}H NMR spectroscopic evidence by sequential spiking for formation of aminooxazoline **168** from the reaction of S-cyanovinyl oxazolidinone thione arabino-**120** and 2-amino-2-cyanoacetamide **106** at pH 4.5. **A**) The crude reaction of **106** and arabino-**120** at pH 4.5 after 8 h. **B**) Spiked with oxazolidinone thione arabino-**103**. **C**) Spiked with oxazolidinone arabino-77. **D**) Spiked with aminooxazoline arabino-**70** (precursor for pyrimidine ribonucleotides).*

Aminooxazoline **168** is a set of distereoisomers due to the fact that **106** is racemic. The two diastereoisomers of **168** are evident in the ^{1}H NMR spectra, where C3´-H peaks are distinguishable by two set of peaks and the C5´-H appear as multiples instead of ABX splitting due to overlap of C5´-H in the two diastereoisomers (Figure 2.21A). There is literature precedence that aminooxazolines exist in its open chain form in aqueous solution.[22, 123, 197] We hypothesised that aminooxazoline **168** can also exist as an open chain form and will lead to the deuteration of C2´-H in D_2O. As a

result, the C2´-H peak of **168** was observed to be a broad as singlet, but on continued incubation in D_2O it became fully deuterated (Figure 2.21B).

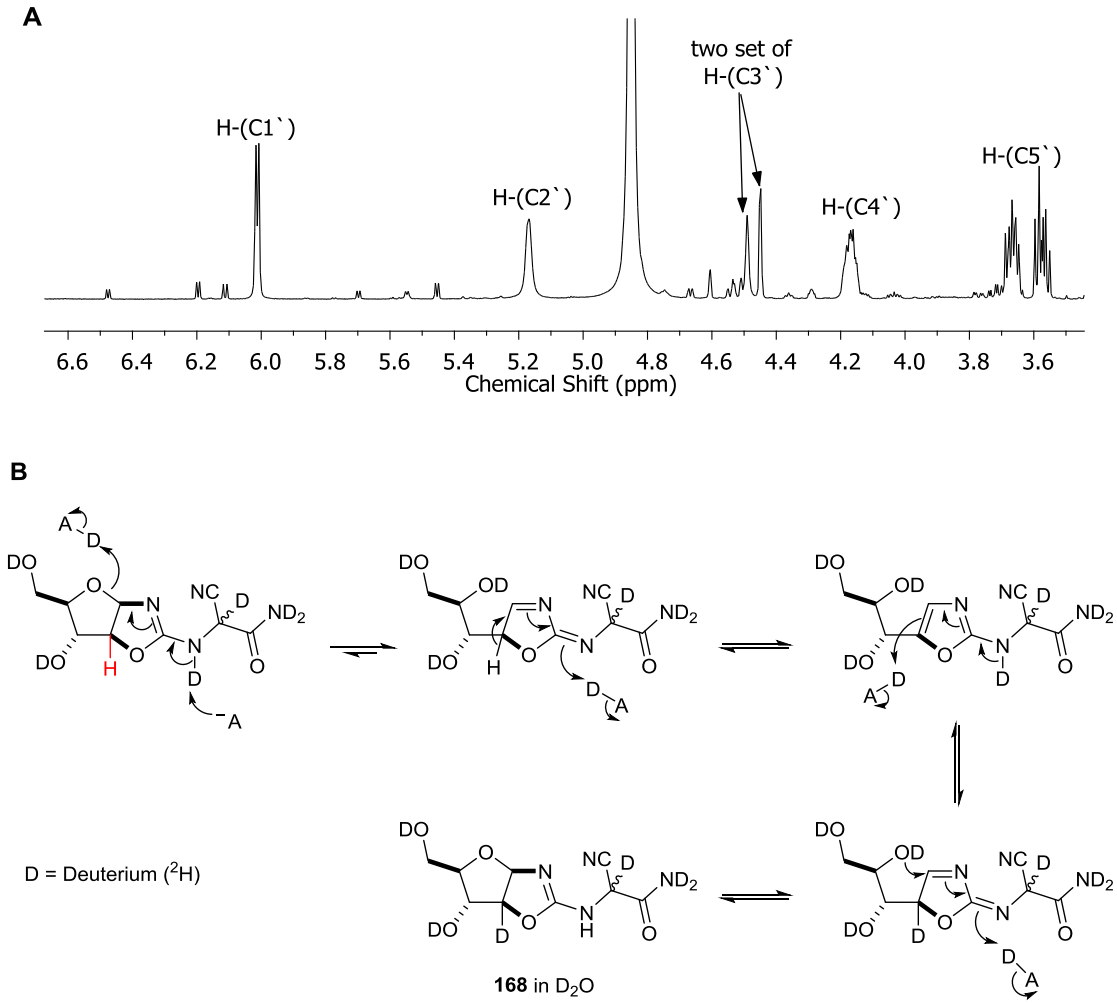


Figure 2.21: A) *^{1}H NMR (600 MHz, D_2O, 3.5 – 6.7 ppm) showing partially purified ^{1}H spectrum of **168** after partially purified by preparative TLC indicating two set of distereoisomers.* ***B)*** *Plausible mechanism for deuteration of C2´-H of aminioxazoline **168** in D_2O.*

Upon close inspection of the crude ^{1}H NMR spectra for the formation of aminooxazoline **168**, two new sets of coupled C1´-H and C2´-H at 6.45 and 5.67 ppm were observed. These are the expected chemical shifts for C1´ and C2´ protons of the anhydro-8-oxo-AICA-arabinoside **96** based on the chemical shift values for the C1´ and C2´ of anhydro-8-oxo-AICN-arabinoside **95** (6.5 and 5.8 ppm, Figure 2.19). It was hypothesised that although the formation of **168** was optimal at low pH (pH 4.5), the subsequent intramolecular cyclisation would require free base of aminooxazoline **168** to produce AICA moiety **96**. The anticipated pKa of **168** would be ~ 6 (based on previously reported aminooxazolines), and this was confirmed upon observed

chemical shift changes in ^{1}H NMR spectra during a pH titration of an NMR sample of 168. Pleasingly, it was found that with pH elevation (pH > 8) the tricyclic moiety **96** was obtained in up to 40% yield at the expense of aminooxazoline **168** (Figure 2.22A). We observed an isolated 13% yield of **96** (2.92 M aminocyanoacetamide **106**, and 4.4M *S*-cyanovinyl thione *arabino*-**120**, 40^{o}C, pH 4.5) upon direct crystallisation from the reaction mixture. Crystals of **96** were then used for NMR and crystallographic studies to confirm its structure (Figure 2.22A and B).

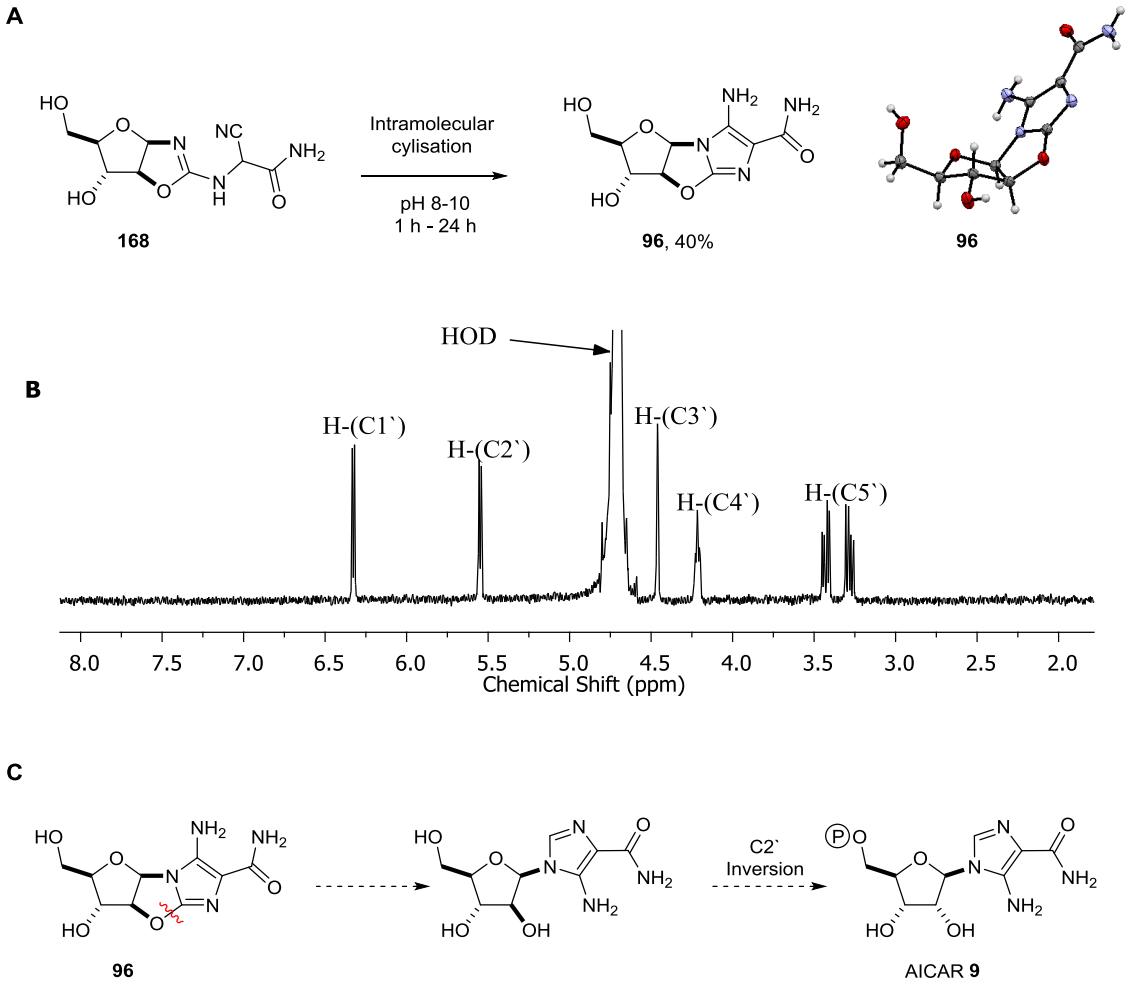


Figure 2.22: A) *Intramolecular cyclisation of aminooxazoline **168** at elevated pH yielding upto 40% anhydro-8-oxo-AICA-arabinoside **96**. X-ray crystallographic structure of **96**, confirming the arabino stereochemistry required for the synthesis of ribonucleotides.* ***B)*** *^{1}H NMR (600 MHz, D2O, 1.5 – 8.2 ppm) spectrum of pure **96** obtained via crystallisation.* ***C)*** *Structural relationship of **96** with aminoimidazole-4-carboxamide ribonucleotide (AICAR **9**).*

It is worthy to note that direct crystallization of aminoimidazole **96** from water provides a prebiotic purification mechanism for accumulation and delivery of **96** under plausible prebiotic conditions.[83, 134] The crystallisation of pure **96** is perceived to offset the low observed yield, because it ensures a clean pathway towards purine ribonucleotides synthesis that can be followed after crystallisation. It is also important to note that anhydro-8-oxo-AICA-arabinoside **96** has a striking similarity with aminoimidazole-4-carboxamide ribonucleotide (AICAR **9)** (Figure 2.22C). AICAR **9** is the first intermediate in the modern de nova biosynthesis pathway of purine ribonucleotides to contain a heterocyclic moiety of the purine base (Section 1.4).[128, 198, 199]

The successful synthesis of AICN tricyclic **96** via the displacement reaction of *S*-cyanovinylated oxazolidinone thione *arabino*-**120** and 2-amino-2-cyanoacetamide **106** has taken us a step closer in our pursuit of building the imidazole moiety of purine nucleobases upon a sugar scaffold tethered by the 8,2'-anhydro-linker required for phosphorylation and C2'-stereochemical inversion. It was hypothesised that the hydrolysed variant of aminomalononitrile **106** will also displace the *S*-methylated oxazolidinone thione *arabino*-**126**. To test this hypothesis, it was decided to react **106** with *arabino*-**126** so that the reaction can be scaled up to make a larger amount of AICA tricyclic **96** and investigate its forward chemistry. As *S*-methylated thione *arabino*-**126** is accessible on a preparative scale relative to the *S*-cyanovinylated thione *arabino*-**120**.

When 2-aminino-2-cyanoacetamide **106** (500mM) and *S*-methyl thione *arabino*-**126** (250mM) were reacted in water at pH 4.5 and heated to 45°C for 3 h, we were able to detect 82% aminooxazoline **168**, 10% *arabino*-**126** and 8% oxazolidinone *arabino*-**77**. Furthermore, **168** cyclised to furnish up to 59% anhydro-8-oxo-AICA-arabinoside **96** under basic conditions (Figure 2.23).

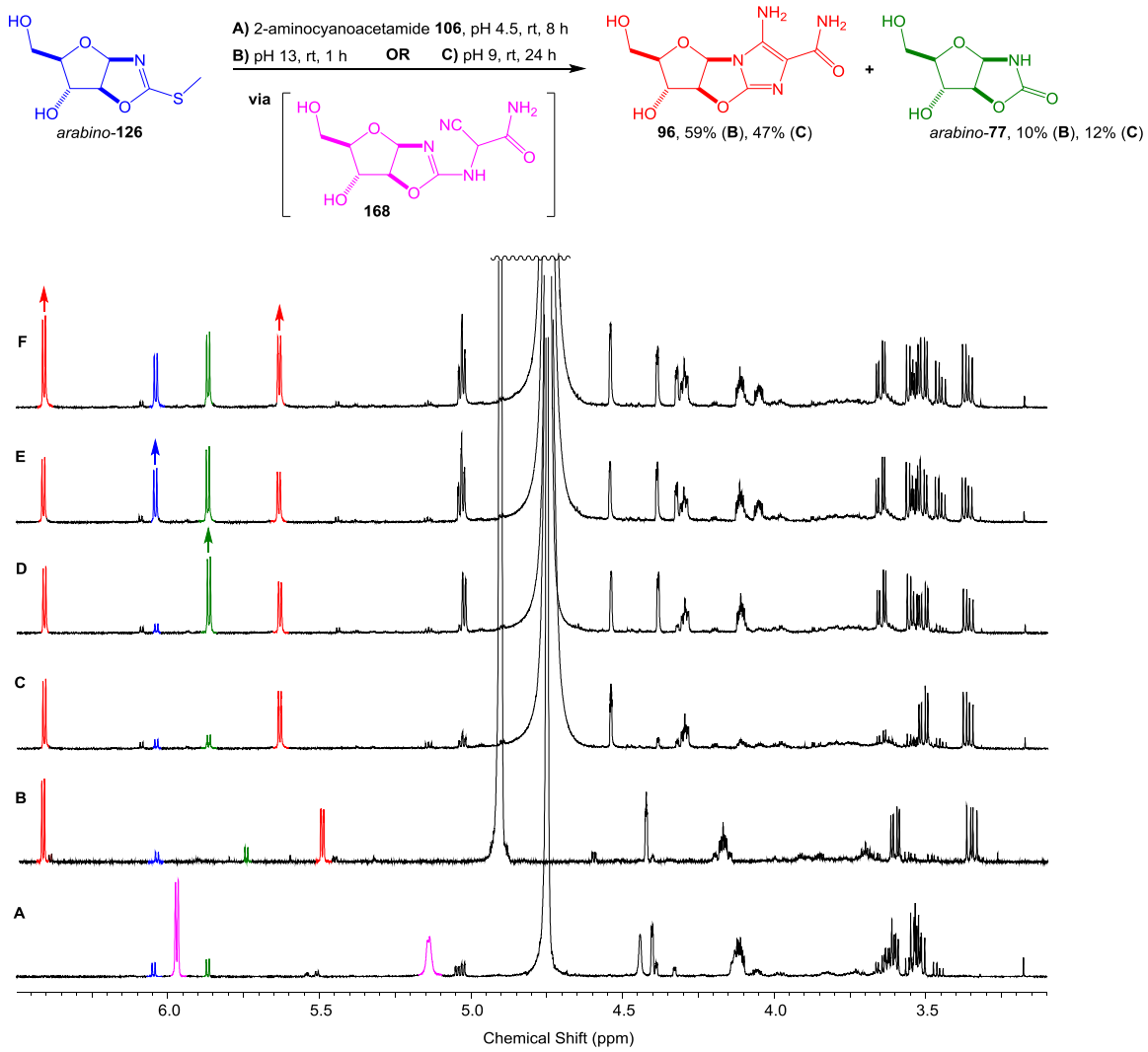
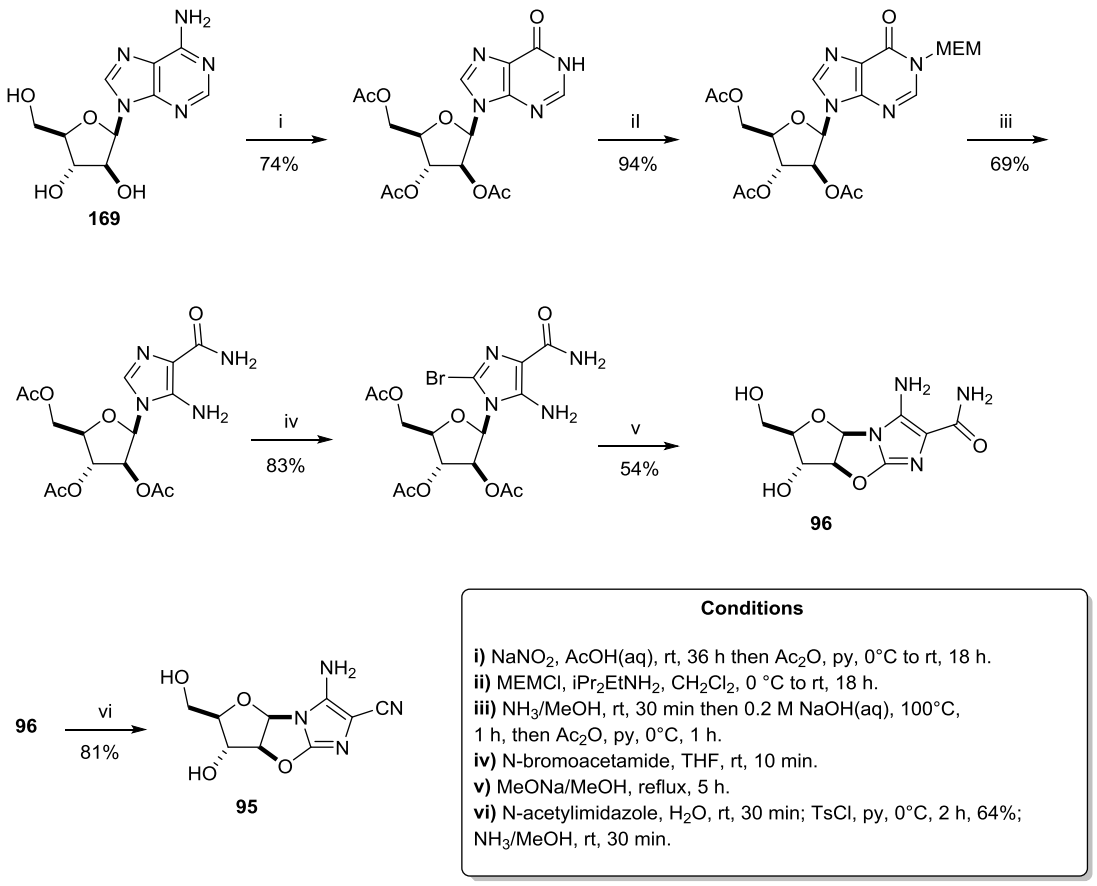


Figure 2.23: *^{1}H NMR spectra (600 MHz, D_2O, 3.1 – 7.0 ppm) for reaction of S-methyl oxazolidinone thione arabino-**126** with 2-amino-2-cyanoacetamide **106**. **A)** Arabino-**126** (250mM) and **106** (500mM) in D_2O (500 µL) at pH 4.5 incubated at rt for 8 h. Calibration to an internal standard gave a yield of 81% (aminooxazoline **168**, 5.96 ppm). **B)** Aliquot (50 µL) of **A** incubated at pH 13 at rt for 1 h giving a yield of 59% **96**. **C)** Aliquot (50 µL) of **A** was added to ammonium hydroxide in D_2O (450 µL, 100 mM), the solution was adjusted to pH 9 with NaOH (4M) and incubated for 24 h at rt. Calibration to an internal standard gave yields of 47% **96**, 12% arabino-**77** with 7% arabino-**126** remaining. **D)** Spiked with arabino-**77**. **E)** Spiked with arabino-**126**. **F)** Spiked with **96**.*

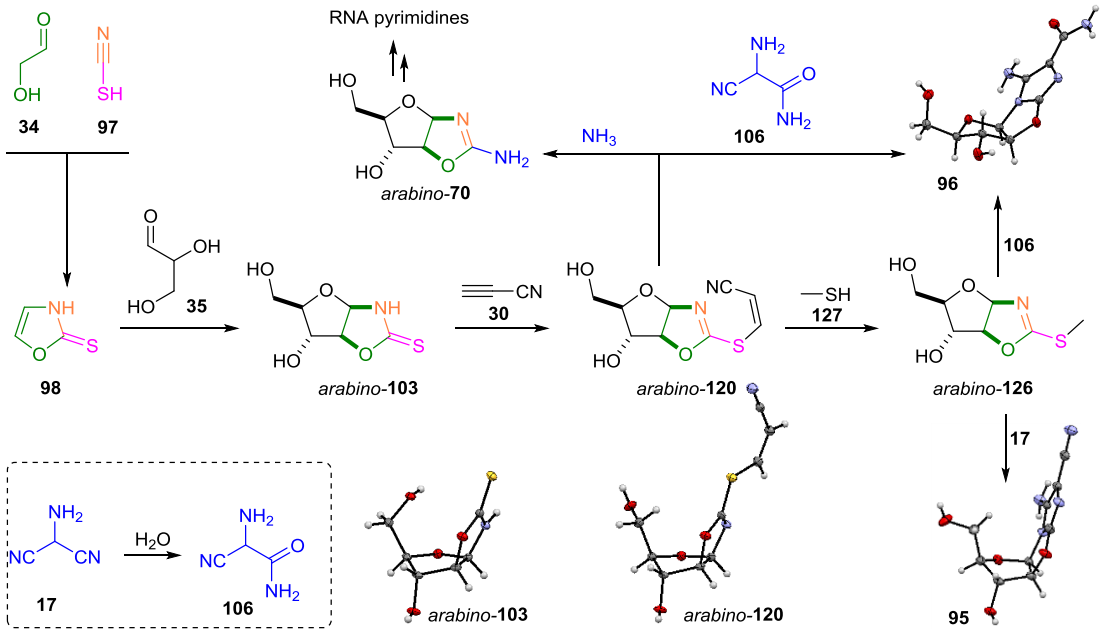
During the course of writing this thesis a second scable synthesise of 2,2´-anhydro-arabinofuranosyl imidazoles was also established in the Powner laboratory (Scheme 2.30).[195]



Scheme 2.30: *Scalable synthesis of 2,2'-anhydro-arabinofuranosyl imidazoles (**96** and **96**) by employing conventional organic methods. Reaction conditions (inset).*

2.5. Summary of nucleophilic additions

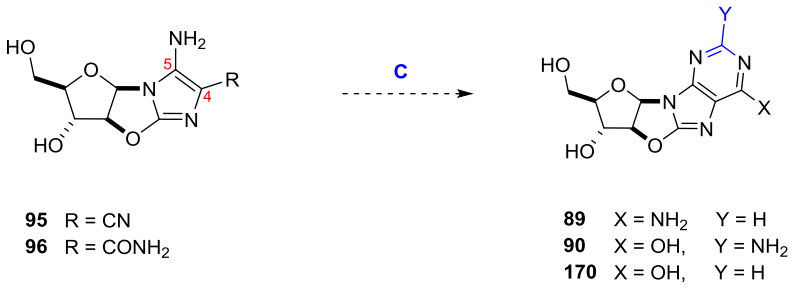
In summary, the two important AICA and AICN tricyclic moieties **96** and **95** were successfully synthesised from plausible prebiotic molecules. And a perceived mutual synthetic relationship between pyrimidine and purine ribonucleotide was demonstrated by successful thiolate exchange in *S*-cyanovinyl thione *arabino*-**120** by ammonia to furnish aminooxazoline *aranino*-**70**. Central to this remarkable concomitant chemistry is *arabino*-**120**, synthesised quantitatively from thione *arabino*-**103** (Scheme 2.31),



Scheme 2.31: *A summary of prebiotic synthesis of AICA* ***96*** *and AICN* ***95*** *tricyclictricyclic compounds and mutal connection of pyrimdine and purines discussed in this chapter. X-ray structure for* ***95*** *was obtained from Dr Shaun Stairs in the Powner laboratory.*

3. Purine cyclonucleosides

With imidazoles **95** and **96** in hand our next goal was to investigate the synthesis of native tetracyclic motif of purine nucleotides **89**, **90** and **170** by direct elaboration of these novel structures. Formally this required the addition of the fifth, and final carbon atom of the nucleotide structure between the C4 carboxyamide/nitrile moiety and C5 amine moiety of the tricyclic iminoimidazoles **95** and **96** to yield 8,2´-anhydro-cyclopurines **89**, **90** and **170** (Scheme 3.1).



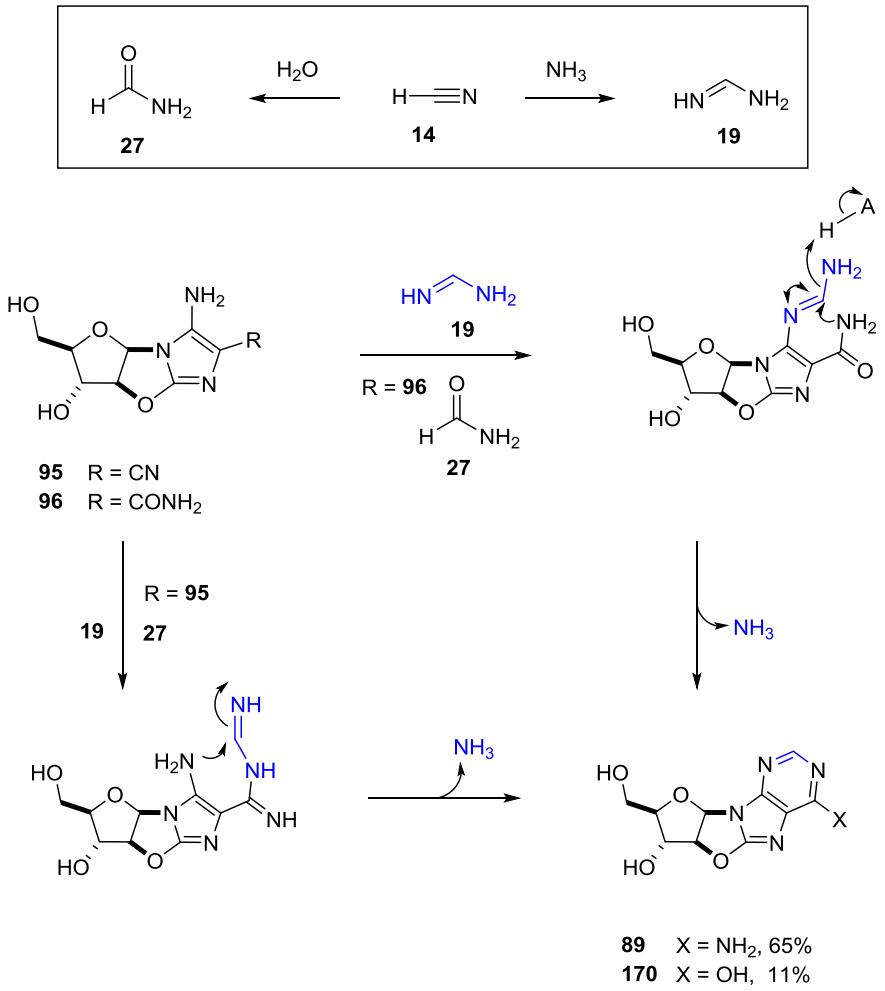
Scheme 3.1: *8,2´-anhydrocyclo- adenosine **89**, guanosine **90** and inosine **170** are the required structures to give natural purine ribonucleotides after phosphorylation. It is thought that these tetracyclic compounds may be synthesised from 2,2'-anhydro-arabinofuranosyl imidazoles **124** and **36**.*

3.2. Plausible prebiotic synthesis of 8,2´-anhydro purine nucleosides

Hydrogen cyanide **14** oligomerisation products aminomalononitrile **17** and 2-amino-2-cyanoacetamide **106** are crucial to the synthesis of our AICA and AICN tricyclic moieties (**95** and **96**). Furthermore, hydrogen cyanide **14** chemistry has previously been widely exploited in the (prebiotic) synthesis of purine nucleobases (Section 1.5.1).[46-49] Therefore, HCN **14** derivatives again appeared to be the ideal prebiotic choice to convert arabinoside **95** and **96** to 8,2´-anhydro adenosine **89** and inosine **170**. To test this hypothesis, **95** and **96** were incubated in formamide **27** (first hydrolysis product of **14**) at 100°C, and over the course of several days we observed the direct conversion of **95** to **89** (10%, 96 h) and **96** to **170** (3%, 72 h).[c] Pleasingly, addition of formamidine **19** (10 equiv.), which is the addition product of ammonia

[c] Prebiotic synthesis of 8,2´-anhydro purine nucleosides was carried out by Dr Shaun Stairs in the Powner laboratory.

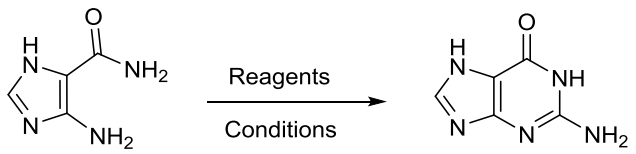
(NH_3) and HCN **14**, significantly improved the yield and rate of 8,2´-anhydro-adenosine **89** synthesis. Following the addition of formamidine **19** we observed both more rapid elaboration of the adenine moiety and a substantially improved yield, such that after only 5 h incubation we observed a 65% conversion to **89**. Conversely, addition of formamidine **19** only marginally improved the yield of inosine **170** (11%, 48 h) and this chemistry remained sluggish (Scheme 3.2).



Scheme 3.2: *Incubation of anhydro-8-oxo-AICN-arabinoside* ***95*** *(40mM) and formamidine.HCl* ***19*** *(400mM) in formamide* ***27*** *after 5 h at 100 °C gave a yield of 65% 2',8-O-anhydro-9-β-arabinofuranosyl-cycloadenosine* ***89.*** *Incubation of anhydro-8-oxo-AICA-arabinoside* ***96*** *(40mM) and* ***19*** *(400mM) in* ***27*** *after 48 h at 100 °C gave a yield of 11% 2',8-O-anhydro-9-β-arabinofuranosyl-cycloinosine* ***170****. Aminolysis and hydrolysis of HCN* ***14*** *yielding formamidine* ***19*** *and formamide* ***27*** *respectively (inset).*

It is of note that canonical nucleobase adenine (A) is efficiently synthesized, but wobble base-pairing inosine (I)[200] is only ineffectually synthesized under these conditions. It is thought that a mechanistic change from electrophilic derivatisation

of the amine moiety of **95** and **96** (which have substantial vinylogous urea **15** character, and are consequently poor nucleophiles) to the initial nucleophilic addition of formamidine **19** to the nitrile moiety of **95** is responsible for the improved reactivity and increased yield of **89** (Scheme 3.2). The difference in adenine and inosine nucleobase syntheses observed suggests that further investigation of the concomitant elaboration of aminoimidazoles **95** and **96** may uncover conditions leading to both anhydro-adenosine **89** and anhydro-guanosine **90**. However, though numerous attempts were made to synthesise anhydro-guanosine **90** from **96**, at this time conversion to **90** has not been observed within the limits of detection. Specifically, we investigated a range of reactions that have been reported to yield guanine **23** from AICA-riboside **21** (Scheme 3.3),[201-203] including cyanogen **25**, cyanate, cyanamide **68** or urea **15**, as well as thiourea, thiocyanate and carbonyl sulfide in formamide at temperatures up to 100 °C, but **90** was not observed during screening. At this time no further attempts were made to synthesise **90**. This is ongoing research within the Powner group. We decided to shift our focus to phosphorylation and chemical inversion of 8,2´-anhydro-cyclopurines **89**, **90** and **170**.



21 **23**

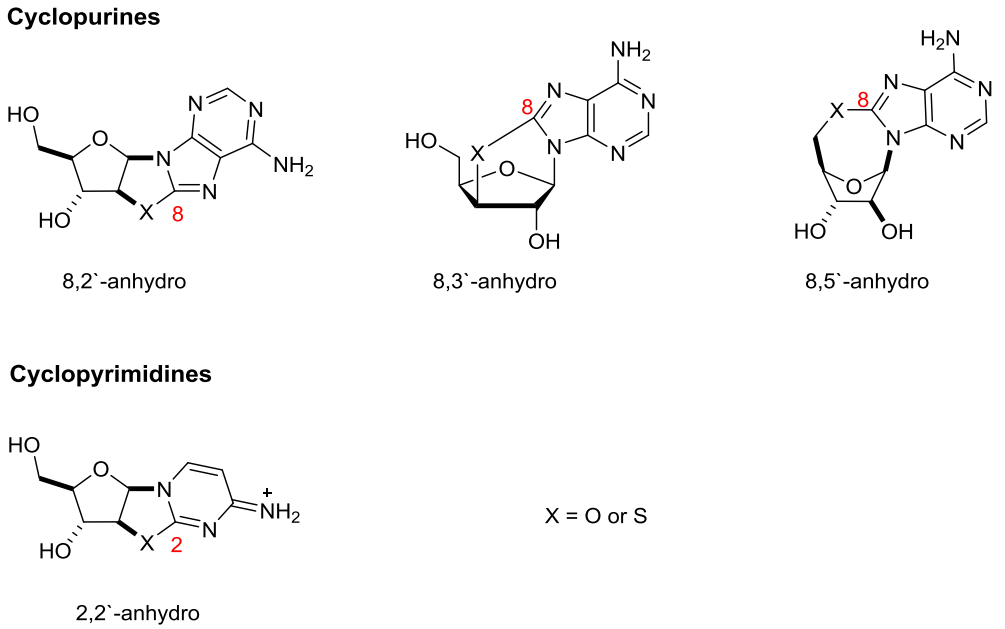
Reagent	Conditions	Yield of **23** (%)
Cyanogen **25**	0.2M, pH (8-10), 100°C, 24 h	43
Cyanate	0.1M, pH (8-10), 100°C, 19 h	20
Cyanamide **68**	0.1M, 100°C, 20 h	1-5
Urea **15**	0.5M, 100°C, 6 d	5-10

Scheme 3.3: *Reported synthesis of guanine nucleobase **23** from 5-aminoimidazole-4-carboxamide (AICA) **21** under varying (prebiotic) conditions. These conditions were attempted with arabinoside **96,** but failed to give 8,2-anhydro guanosine **90** within the limits of detection by ^{1}H-NMR.*

Therefore to facilitate our study of the chemistry of 8,2´-anhydro purines **89**, **90** and **170**, we needed to develop conventional (non-prebiotic) methods to access large quantities of these materials to carry out detailed studies.

3.3. Conventional chemical synthesis of 8,2´-anhydro purine nucleosides

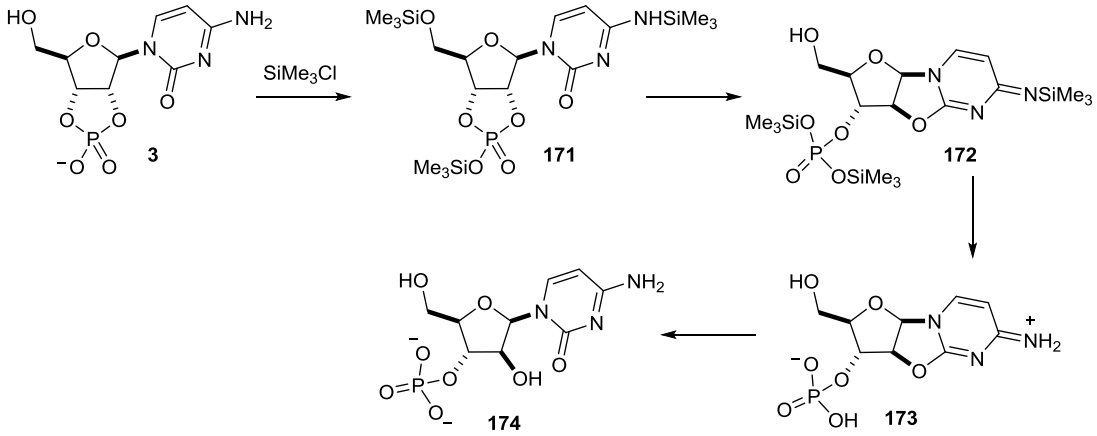
Purine cyclonucleosides differ from canonical purines through a covalent linkage, either directly or via a bridging atom (O or S) between the 2´, 3´, or 5´ carbon of the sugar moiety and C8 carbon of the purine moiety (Scheme 3.4). These nucleotide analogues have been extensively studied for more than half a century, and are found to have a variety of chemical, chemotherapeutic, biochemical and biophysical properties of interest.[204-209]



Scheme 3.4: *Anhydro purine and pyrimidine cyclonucleosides*

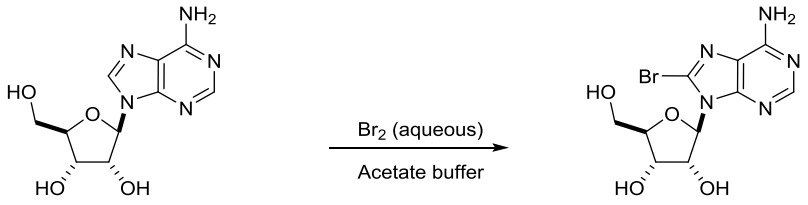
Structural comparison of pyrimidine and purine cyclonucleosides suggests that the C8 position of an 8-oxo purine could react in an analogous manner to the C2 position of a pyrimidine nucleosides (Scheme 3.4). For example, one could draw comparison to the synthesis of anhydro-cytidine-3´-phosphate **174** reported by Nagyvary *et al.*[210] by intramolecular rearrangement of 2´,3´-cyclicphosphate cytidine **171**, which is induced

by silylation of the cyclic phosphate moiety (Scheme 3.5), Ikehara has previously hypothesized that "if a keto or thioketo function could be introduced into the 8 position of a purine nucleoside, nucleophilic attack by either of these groups on a carbon of the carbohydrate moiety bearing an alkyl- or arylsulfonyloxy group should give rise to a cyclonucleoside".[211]



Scheme 3.5: *Synthesis of cytidine-arabino-3´-phosphate* ***174*** *via hydrolysis of 2,2´-anhydro-3´-phosphate-cytidine* ***173****, which is induced by silylation of 2´,3´-cyclicphosphate cytidine* ***3****.*

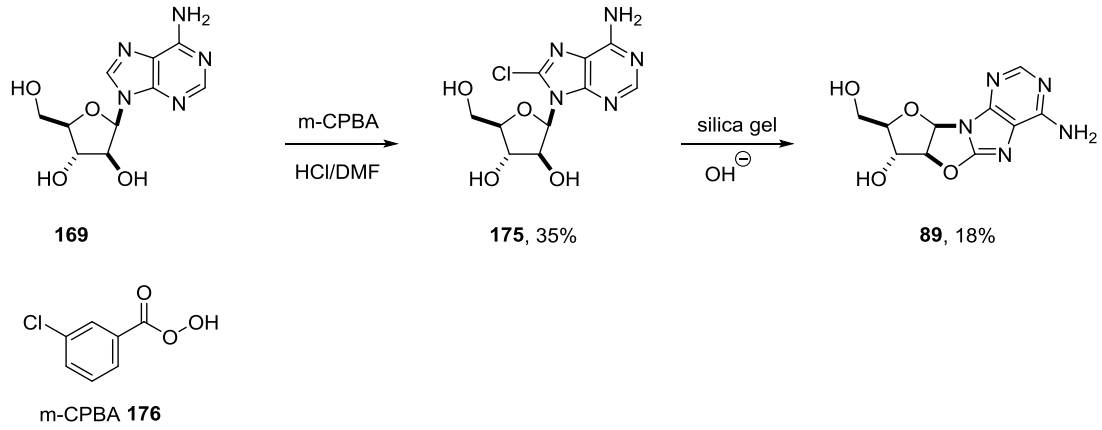
We set out to develop a convenient synthetic route to obtain 8,2´-cyclopurines (**89**, **90** and **170**) so that it can be tested for prebiotic phosphorylations. Therefore, we sought to synthesize 8,2´-O cyclonucleosides by established literature procedures. Chemical modification of purine nucleoside at C8 is readily achieved via oxidation and Ikehara *et al.* introduced an excellent methodology for the direct C8 brominiation of purine nucleosides, via the direct reaction of nucleosides and bromine water with acetate buffer (Scheme 3.6).[211]



Scheme 3.6: *Near quantitative bromination of purine nucleosides by bromine water with acetate buffer reported by Ikehara et al.*[211]

These brominated nucleosides could then be readily converted to purine cyclonucleosides, but requiring multistep protecting group strategies or selective C2´-

OH tosylation catalyzed by dibutyltin oxide.[212, 213] Further development of this chemistry has been reported by Ogilvie *et al.*[207, 208, 214] however, exploiting the commercial available *arabino*-adenosine **169**, given the simplicity of oxidation and ring closure in the *arabino*-series, removes the complication of stereochemical inversion, multistep protecting group chemistry, or the use of toxic tin oxide activation of C2´-OH. We therefore followed the synthetic procedure reported by Ryu *et al.* to make the 8-2´-O anhydro nucleosides by treatment of arabino-adenosine **169** in HCl/DMF and m-CPBA **176** to give 8-chloro adenine **175**, followed by silica gel-induced cyclisation to give 8,2´-*O*-anhydro-arabino-adnine **89** as the sole product.[215] (Scheme 3.7). However, despite repeated attempts with acidic and neutralized silica and both neutral and basic alumina, we were unable to collect any cyclisation product **89**.



Scheme 3.7: *Synthesis of 8,2´-O-anhydro-arabino-adenine **89** via cyclisation of 8-cholro-arabino-adenine **175** reported by Ryu et al.*[215]

Ryu *et al.* also reported that treatment of 8-chloro adenine **175** with 1M NH_4OH or 1M NaOH gave **89** quantitatively (by TLC analysis).[215] Therefore, we next decided to isolate **175** by flash column chromatography (FCC), then incubate **175** under alkaline conditions to induce cyclisation. Following oxidation, chloroadenine **175** was isolated (in 35% yield). Chloroadenine **175** was then treated with 1M NaOH and purified via silica gel to obtain **89** (18%). Although the overall yield of the reaction (6.5%) was low, we were pleased to obtain an authentic sample of 8,2´-*O*-anhydro adenine **89**. The charactersation data for **89** matched that presented in the literature and **89** was also readily recrystallised from hot water to allow for single X-ray crystallographic analysis to further compare the structural relationship between anhydro purine **89** and

anhydro cytidine *arabino*-**73**. The X-ray structure unambiguously showed the anhydro-*O* linkage and *arabino* stereochemistry as expected, but interestingly we also observed a short interaction between the 5´-hydroxyl and the C8 carbon in **89** (Figure 3.1). This interaction compares directly with similar observations in the crystal structure of anhydrocytidine *arabino*-**73**. The n→π* interaction observed in *arabino*-73 has previously been proposed to kinetically favour 3´-phosphorylation under prebiotic phosphorylation conditions.[22, 117] Therefore, the observation of a similar n→π* interactions in anhydro adenine **89** suggested that the 8,2´-*O*-anhydro-cyclopurines such as **89** may be excellent candidates for chemoselective (prebiotic) phosphorylations and have the required activation to forming purine 2´,3´-cyclic phosphate, as previously seen in the pyrimidine series.[22] However, we deemed it was important to optimize the synthesis of 8,2´-anhydro-cyclopurines before we undertook an investigation of phosphorylation reactions.

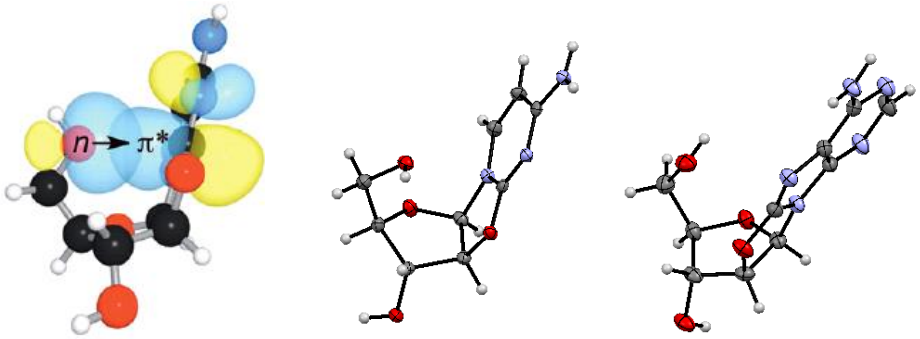
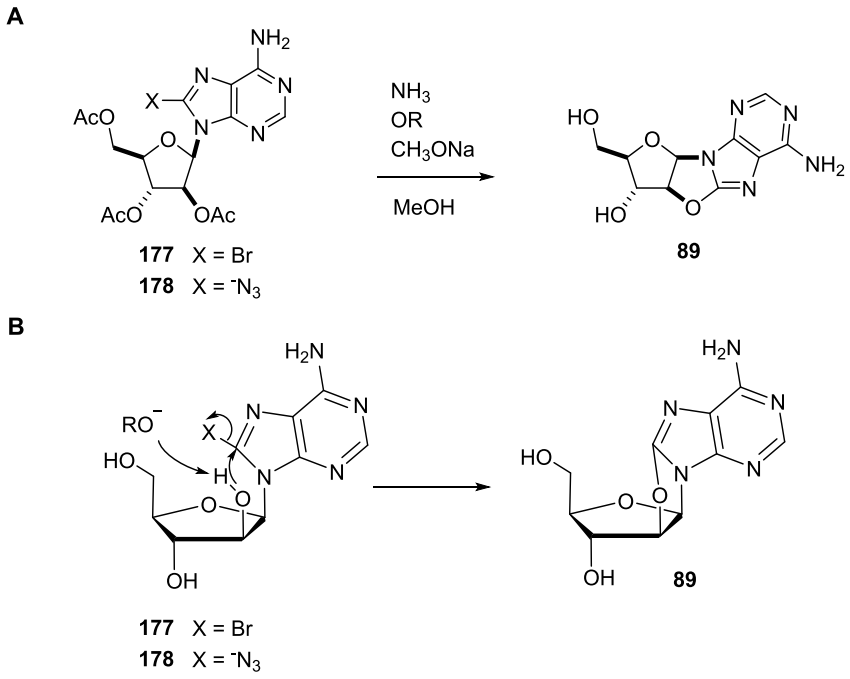


Figure 3.1: *X-ray crystal structures of 8,2´-anhydrocycloc adenine* ***89*** *and 2,2´-anhydrocytidine arabino-****73*** *showing the similarity of the n → π* interaction of the C5´-OH into the C2 carbon atom. X-ray structure for arabino-****73*** *reported by Powner et al.*[22] *and DFT structure for arabino-****73*** *reported by Choudhary et al.*[126]

Reist *et al.*, have reported that treatment of bromo-arabinoside **177** with methanolic ammonia at room temperature led to the formation of **89** in good yield. [216] Anhydropurine **89** was also obtained when **177** was treated with sodium methoxide (Scheme 3.7). These reaction conditions both demonstrate ester cleavage but that once formed the C8 anhydropurine ligation appear relatively stable to nucleophilic attack (by methoxide and ammonia) and **89** is described as the sole product of these reactions. **89** was also the sole product, when 8-azide arabinoside **178** was

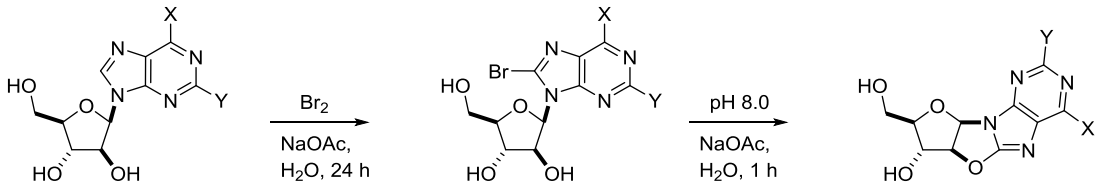
deacetylated under similar conditions.[216] The efficiency of this reaction suggest that significant improvement could be made to the synthesis of **89**, and gave the first indication of a difference in reactivity between anhydrocytidine *arabino*-**73** and anhydroadenine **89**. Whilst adenine **89** appear to be highly stable to nucleophilic attack under basic reaction condition, cytidine *arabino*-**73** is known to be highly labile and readily undergoes hydrolysis at the C2,2'-anhydronucleoside linkage.



Scheme 3.8: A) *The reaction of arabinoside* ***177*** *and* ***178*** *with ammonia or sodium methoxide in methanol to form* 8,2´-anhydrocycloc adenine ***89*** *via the intramolecular* S_N^2 *displacement of the C8 leaving group by the C2´-hydroxyl, rather than by the external nucleophile (*MeO^- *or* NH_3*).* ***B)*** *Plausible mechanism for nucleophilic base-catalysed displacement of the C8 leaving group (bromo, chloro or azide) by the 2´ hydroxyl to make* ***89****.*

Our next attempt to synthesise adenine **89** was undertaken by employing a combination of the methods describe by Ikehara *et al.* for the direct bromination of purine nucleosides,[211] and the alkaline cyclisation reported by Reist *et al.*[216] Commercially available *arabino* adenosine **169** was reacted with saturated bromine water in sodium acetate buffer to cleanly give 8-bromo adenosine **179** in 56% yield as a white powder upon (direct) crystallization from reaction mixture (Scheme 3.9). The bromination of **179** (56%) was promising when compared to the lower yielding

chlorination reported by Ryu *et al.* particularly given that the latter required purification by flash column chromatography,[215] whereas the former could be isolated by direct crystallisation. Interestingly, however, incubation of **179** with sodium acetate at pH 8.0 in room temperature furnished the desired 8,2´-anhydro-*arabino*-cycloadenine **89** in 73% yield. Once again **89** was observed to crystallize directly from the reaction (Scheme 3.9). By telescoping the oxidation and cyclisation, and through exploiting the remarkable crystallinity of **89**, we had in hand an excellent large-scale strategy for the synthesis of **89**. Therefore, we next tested the generality of these condition by investigating the synthesis of 8,2´-anhydro-arabino-guanine **90**. *Arabino* guanosine **180** was again directly brominated under mildly alkaline (pH 8) condition to again furnish 8-bromo *arabino* guanosine **181** in excellent yield (80%). Cyclization was induced in **181** by treatment with sodium acetate buffer (1M, pH 8.0) to give **90** (80%) as a direct crystalline precipitate from the reaction solution (Scheme 3.9). Unfortunately, due to the poor quality of the crystals formed we were unable to get single crystal data for **90** directly, and no attempt to recrystallize **90** were made at this time.



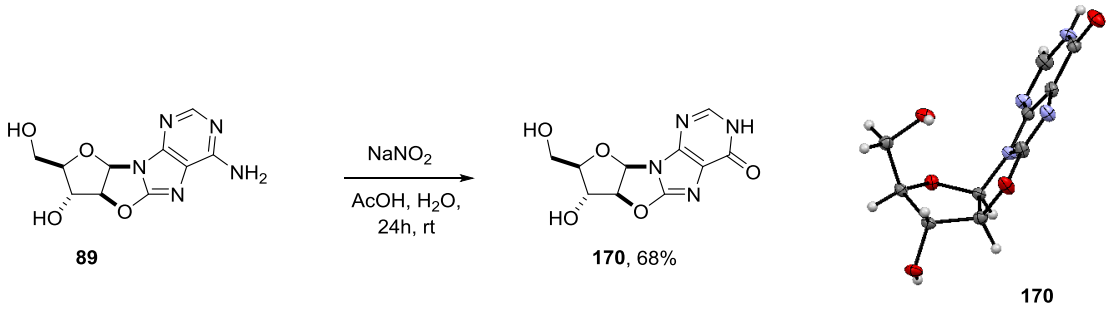
169 X = NH_2, Y = H
180 X = OH, Y = NH_2

179, 56% X = NH_2, Y = H
181, 80% X = OH, Y = NH_2

89, 73% X = NH_2, Y = H
90, 80% X = OH, Y = NH_2

Scheme 3.9: *An optimised large scale synthesis of 8,2´-anhydro-cyclopurines **89** and **90** via the facile intramolecular cyclization of 8-bromo arabino adenosine **179** and 8-bromo arabino guanosine **181** respectively.*

Finally, 8,2´-anhydro-*arabino*-inosine **170** was also synthesised employing a literature procedure from adenine **89** (Scheme 3.10).[217] Pleasingly, 8,2´-anhydro cycloinosine **170** was obtained in 68% yield as pale yellow crystalline platelets (Scheme 3.10). An X-ray crystal structure confirmed the identify of **170**.[22]



Scheme 3.10: *8,2´-anhydro-arabino-cycloinosine* ***170*** *via nitration of 8,2´-anhydro-arabino-cycloinosine* ***89*** *by sodium nitrite.*

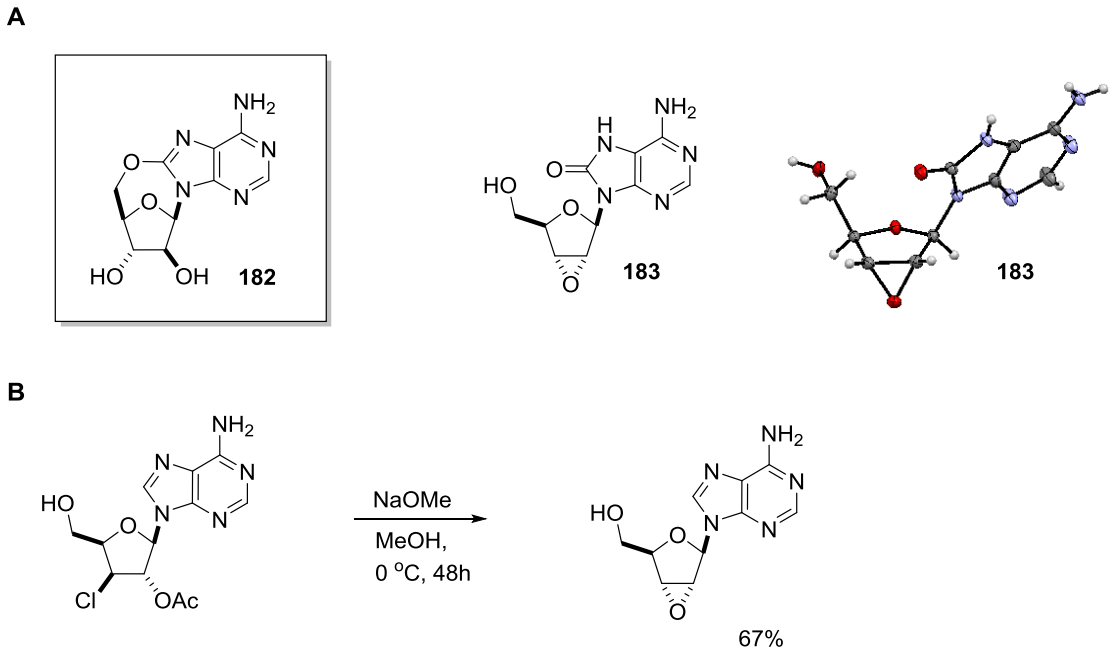
It may be important to note that the synthesis of **170** from **89** is potentially prebiotic plausible given that nitrous acid (HNO_2) can readily be synthesised from nitric oxide and water.[45, 218] However no further investigation of the prebiotic plausibility of this process was undertake here. Rather, with scalable synthesise of 8,2´-anhydro-*arabino*-cyclopurines **89**, **90**, and **170** in hand and an indication that in the solid state **89** and **170** showed n→ π* orbital overlap thought to block the C5´-OH reactivity we proceeded to investigate the desired selective C3´-OH phosphorylation of **89**, **90**, and **170.** [22, 126, 219]

3.4. Isomerisation of anhydro-cyclopurines

During cyclisation studies of 8-bromo-*arabino*-adenosine **179**, a batch was unintentionally treated with 2M sodium hydroxide. To our surprise, this led to the unforeseen discovery that **179** under extremely alkaline condition led to a second set of nucleoside peaks in the crude ^{1}H NMR spectra. The proton resonances were similar to **89**, suggesting that the sugar and the nucleobase moieties were intact and depurination had not occurred. We predicted the new resonance were likely due to 8,5´-anhydro cycloadenine **182**, because the C5´-proton resonances of the new product were more downfield shifted than those of **89**, and the isomerisation of **89** to **182** would be the final outcome of the reaction of 5'-OH with C8 which is foreshadowed by the n→π* interaction observed in the crystal structure of **89**.

We isolated the unknown nucleoside in 34% yield by flash column chromatography and obtained full characterisation data. However, we were unable to assign rational structures data to the predicted 8,5´-nucleoside **182**. In the meantime, an analytically pure sample of the unknown was obtained by recrystallization and submitted for x-ray

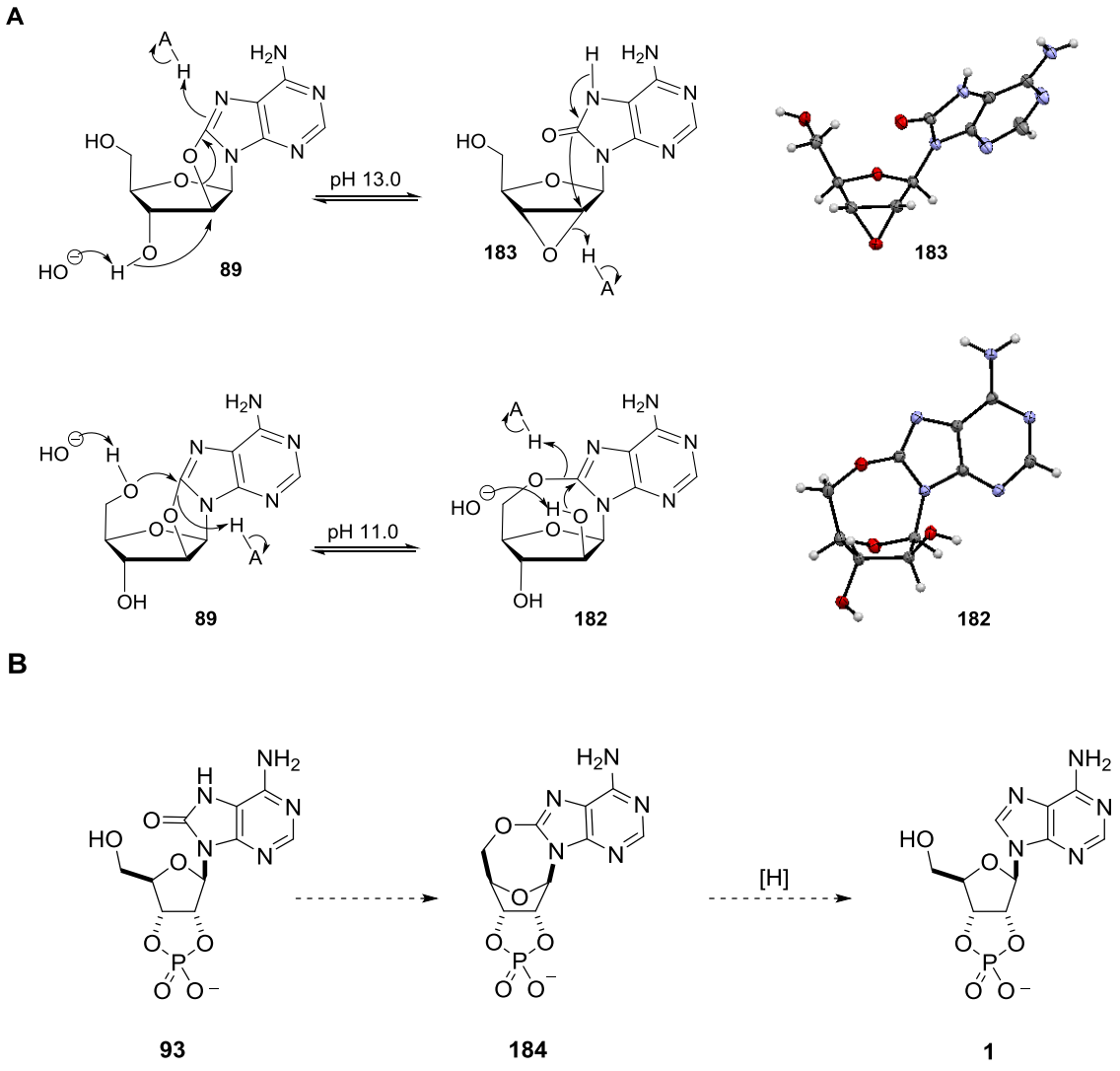
diffraction studies demonstrating that under these highly alkaline conditions intramolecular rearrangement had occurred to furnish the epoxide moiety of 2´,3´-cyclonucleoside **183** (Scheme 3.11A). Indeed these structures had previously been reported (Scheme 3.11B).[220, 221] The confirmation that the 8-oxo-epoxy adenine arabinoside **183** was observed upon isomerisation of purine cyclonucleoside **89**, and remarkable stability of anhydropurine, such as **89** to alkaline hydrolysis prompted us to increase our understanding of isomerisation in purine cyclonucleosides.



Scheme 3.11: A) *X-ray structure of 2´,3´-8-oxy-adenosine* ***183*** *is shown to confirm the identity of the unknown nucleoside formed during incubation of* ***89*** *under extremely alkaline conditions. Structure of 8,5´-anhydro-arabino-cycloadenine* ***182*** *(inset).* ***B)*** *Nucleophilic attack of the C2´-hydroxyl on the C3´-Cl moiety to yield 2´,3´-cycloadenosine (67%) in methanol, at 0°C for 48h reported by Russell and Benitez.*[98,99]

Rees *et al.* have shown that treatment of 8,2´-anhydro-arabino-cycloadenine **89** with 1.25M sodium hydroxide furnished 2´,3´-anhydro-8-oxyadenosine **183** in excellent yield (88%).[222] They have also shown that conversion of **89** into **183** is reversible. Furthermore, Ikehara and Ogiso, have reported conversion of **89** into 8,5´-anhydro-cycloadenosine **182** with 0.01M aqueous sodium hydroxide and also suggest that this reaction is reversible.[223] Enlightened by the reports of Rees and Ikehara, we decided to reinvestigate this interconversion. ^{1}H NMR spectroscopic analysis of the incubation of **89** (20mM) at pH 11 at 40°C for 24 h indicated that 8,5´-anhydro-cycloadenine **182**

(35%), **89** (55%) and **183** (7%) had been obtained. We repeated the above reaction and isolated **182** (11%) as white prismatic crystals, and the structure of **182** was confirmed by X-ray analysis (Scheme 3.12A). Observing the formation of 8,5´-anhydro linkage in the isomerisation of **182** was deemed to be of significant interest suggesting a possible method to get to the native purine nucleotides from 8-oxo purines (via the reductive cleavage of the 8,5´-linker, Scheme 3.12B), a topic that will be discussed later in this thesis.



Scheme 3.12: A) *Plausible mechanism for interconversion of 8,2´-anhydro cycloadenine* ***89*** *between 2´,3´-anhydro-8-oxo cycloadenine* ***183*** *and 8,5´-anhydro cycloadenine* ***182*** *under strong alkaline conditions.* ***B)*** *A postulated reaction of 8-oxy-2´,3´-cyclic phosphate adenosine* ***93*** *to form native 2´,3-cyclicphosphate adenosine* ***1*** *via the reductive cleave of 8,5´-anhydro linkage in* ***184****.*

To test the isomerisation of 8,2´-anhydro-*arabino*-cycloadenine **89** in highly alkaline solutions, we next incubated **89** at pH 13 for 24 h at 40°C. ^{1}H NMR analysis of the

mixture now showed that residual **89** (12%), was accompanied by 8,5´-anhydro cycloadenine **182** (20%) but now predominately 2´,3´-anhdyro-8-oxo cycloadenine **183** (60%). Although Rees was unable to detect any **182** at pH 13, we were able to observe 20% **182** in our reaction. Indeed, we were surprised to observe both isomers **182** and **183** in our reactions, since Ikehara and Rees had both reported the observation of only one isomer **182** or **183** respectively (Figure 3.2).

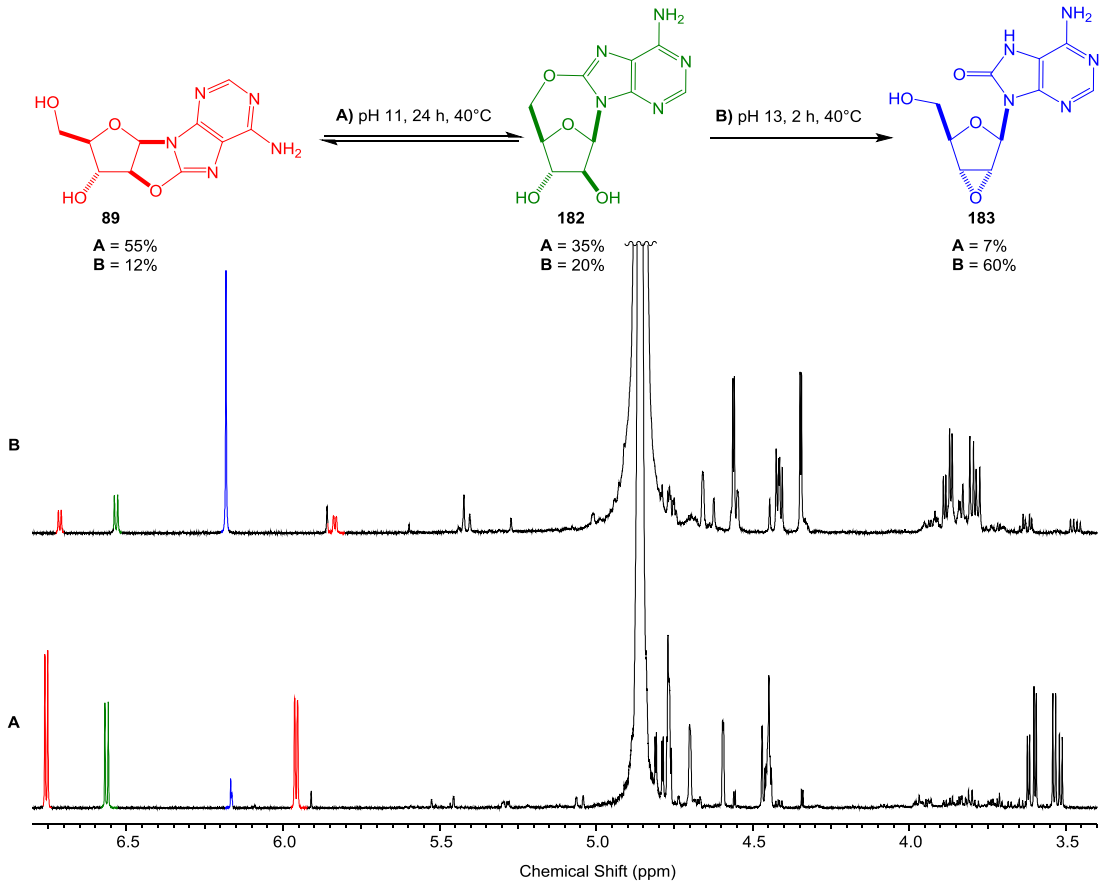
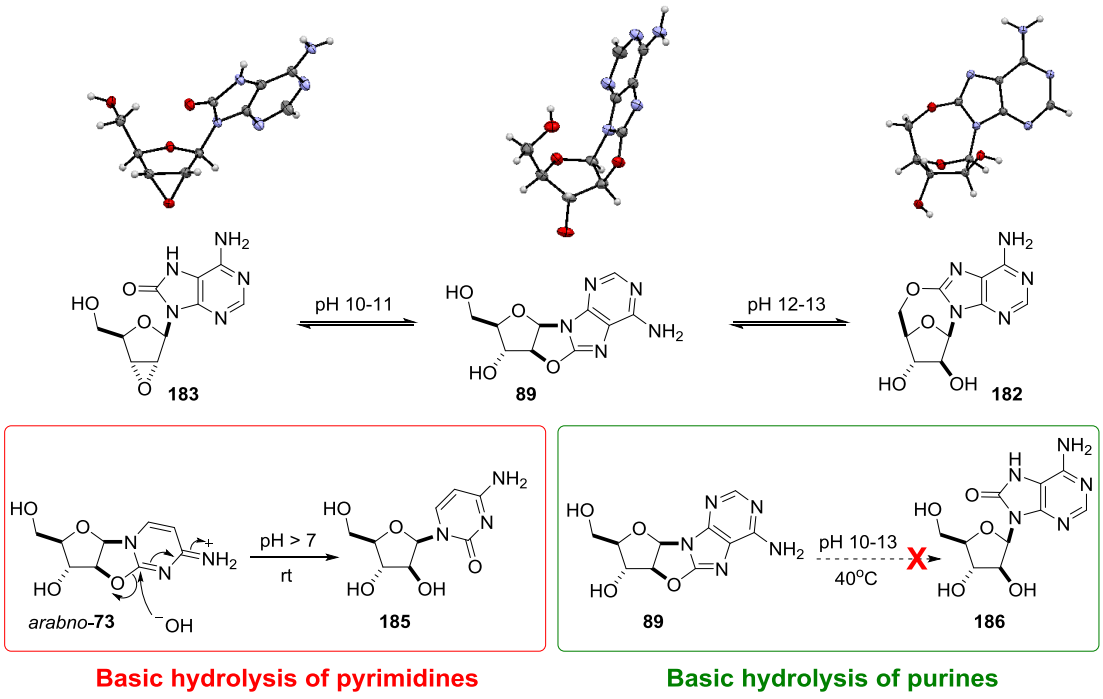


Figure 3.2: *^{1}H NMR (600, MHz, 3.4 – 6.8 ppm) showing isomerisation of anhydropurines. 8,2´-anhydro-arabino-cycloadenosine **89** (20mM) was incubated at pH 11 at 40°C.* ***A)*** *Reaction after 24 h showing a mixture of **89** (55%), 8,5´- anhydro-arabino-cycloadenosine **182** (35%) and 2',3'-epoxy-arabino-8-oxoadenosine **183** (7%). The pH was increased to 13 and the reaction was again incubated at 40°C.* ***B)*** *Reaction after 2 h showing a mixture of **89** (12%), **182** (20%) and **183** (60%).*

The isomerisation of **89** under alkaline condition (> pH 11), *albeit* the conditions being more alkaline than reasonable for prebiotic reactivity, are of note because of the intricate remarkable stability of the anhydropurine structures (**89**, **90** and **170**) to hydrolysis. Conversely, the prebiotic synthesis of activated pyrimidine ribonucleotides requires the careful pH buffering at pH < 7, to avoid the deleterious, and rapid, hydrolysis of 2,2´-anhydrocytidine *arabino*-**73** to *arabino*-cytidine **185** (Scheme 3.13). This switch in stability was deemed to be of significant interest

because hydrolysis of *arabino*-**73** to **185** has also been observed upon (prebiotic) urea-mediated phosphorylation of *arabino*-**73**, and this hydrolysis prevents intramolecular C2-inversion and synthesis of 2′,3′-cyclic phosphate cytidine **3**, and accordingly contribute to a major pathway for loss of prebiotic *ribo*-cytidine nucleotides during their synthesis. However, our studies have confirmed that 8,2′-anhydro-*arabino*-cyclopurines, such as **89**, are remarkably resistant to hydrolysis. Even upon extended incubation at elevated pH, **89** was not observed to not undergo hydrolysis. Instead, **89** undergoes isomerisation to 8,5′-anhydronucleoside **182** (55%, pH 11, 40°C, 24 h) and 2′,3′-anhydronucleoside **183** (60%, pH 13, 40°C, 24 h) (Scheme 3.13 and Figure 3.2). It was therefore anticipated that the isomerisation mechanism, which protected **89** from alkaline hydrolysis, may also protect **89** from hydrolysis under prebiotic phosphorylation conditions.

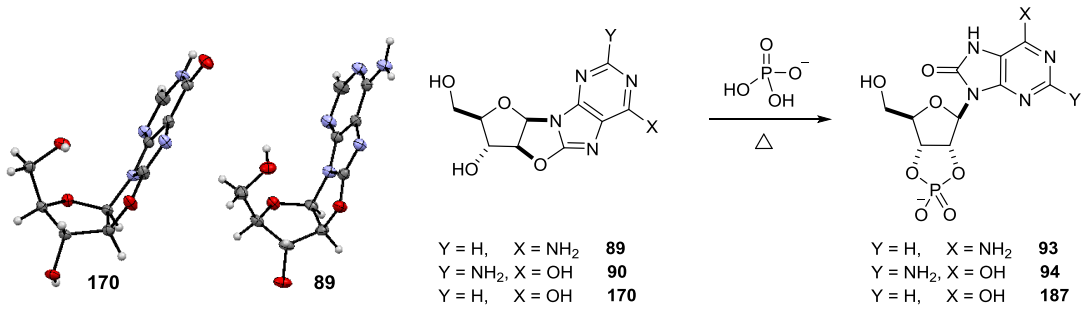


Scheme 3.13: *Comparing the hydrolysis of pyrimidine ribonucleotides at pH 6.5-13. 8,2′-Anhydro-arabino-cycloadenosine* ***89*** *does not hydrolyse under basic conditions (pH 10-13, 40°C) to give* ***186,*** *but instead* ***89*** *isomerises to yield 8,5′-anhydronucleoside* ***182*** *and 2′,3′-anhydronucleoside* ***183*** *(green). 2,2′-Anhydrocytidine arabino-****73*** *readily undergoes hydrolysis above pH 7 at rt to yield arabino-cytidine* ***185*** *(red), hence the prebiotic synthesis of pyrimidine ribonucleotides requires inorganic phosphate to maintain pH at 6.3.*[22]

With an optimised and large scale reproducible synthesis of 8,2′-anhydro cyclopurines (**89**, **90** and **170**) and the knowledge that cyclopurines isomerise under highly alkaline conditions, forming 5′-8-*O* (**182**) or 3′-*O*-2′ (**183**) linkages, we next sought to investigate the phosphorylation and chemical inversion of these anhydro cyclopurines.

4. Phosphorylation and stereochemical inversion of anhydro cyclonucleosides

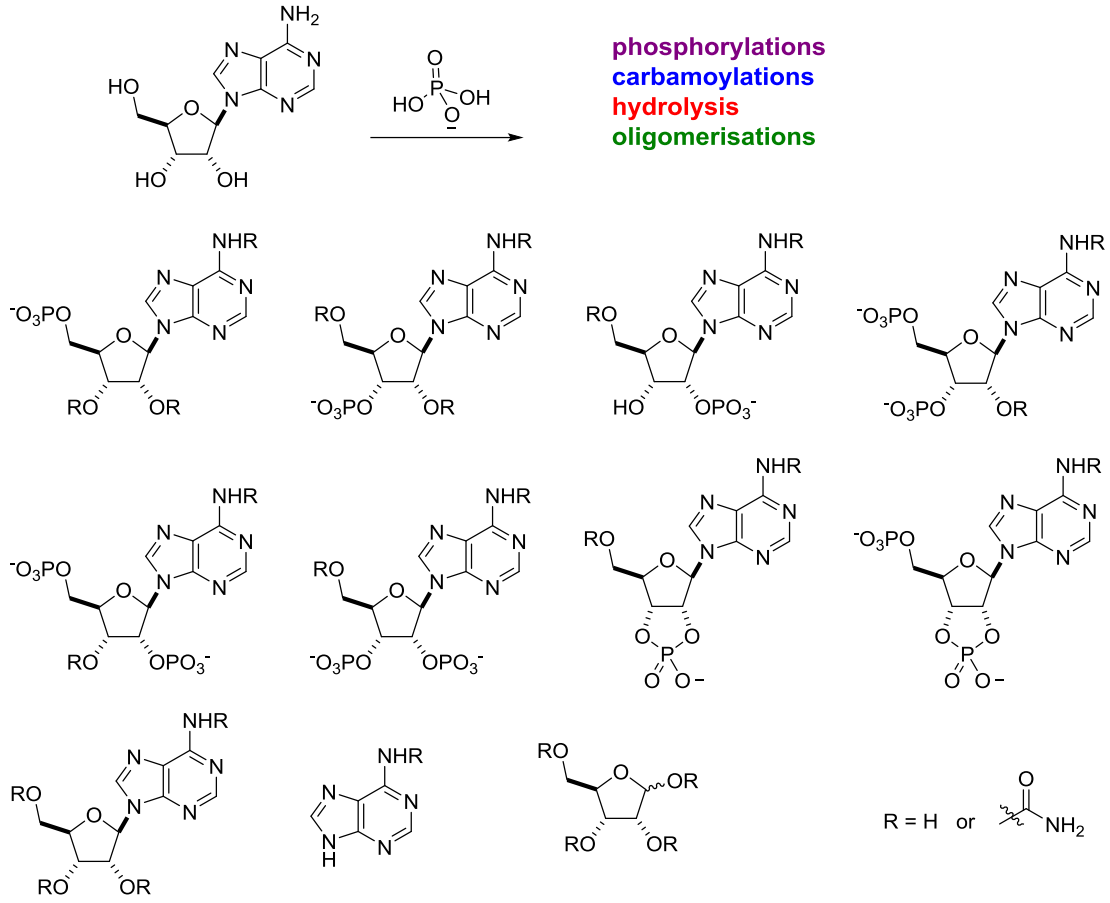
The next vital step of this project was to investigate the prebiotic phosphorylation of 8,2´-anhydro cyclonucleosides **89**, **90**, and **170** to furnish 8-oxo-2´,3-cyclphosphates **93**, **94**, and **187** respectively (Scheme 4.1).



Scheme 4.1: *Proposed prebiotic phosphorylation of anhydronucleosides **89**, **90** and **170** yielding 8-oxo-2´,3-cyclphosphates **93**, **94**, and **187** respectively.*

Water is the most obvious solvent for prebiotic chemistry because it is the most abundant on Earth and has excellent physiochemical properties that enable it to solvate a broad spectrum of polar compound across a very large temperature range.[224] However, phosphorylation is formally a dehydration reaction, which engenders a number of challenges for aqueous phosphorylation.[225, 226] The most widely exploited approach to the deal with the problem is the activation of orthophosphates with condensing agents such as cyanamide or cyanogen,[227] and polyphosphate, such cTMP.[81, 82, 228, 229] However, a more simple approach to the challenge of prebiotic phosphorylation is to avoid or limit the water activity in the system, through evaporation. Dry state phosphorylations, using simple orthophosphates as the source of phosphate, have been used extensively in the prebiotic synthesis of nucleotides.[22, 230-234] Ponnamperuma and Mack demonstrated phosphorylation yields (16%, 2h, 160°C) for a series of nucleotides (including adenosine, guanosine, uridine, cytidine and thymidine).[232] However, experiments at moderate temperatures (65-85°C) require several months to achieve similar results. The results were slightly improved by the use of ammonium phosphates, as evaporation of ammonium phosphates causes loss of ammonia and then provide an acidic environment for relatively faster

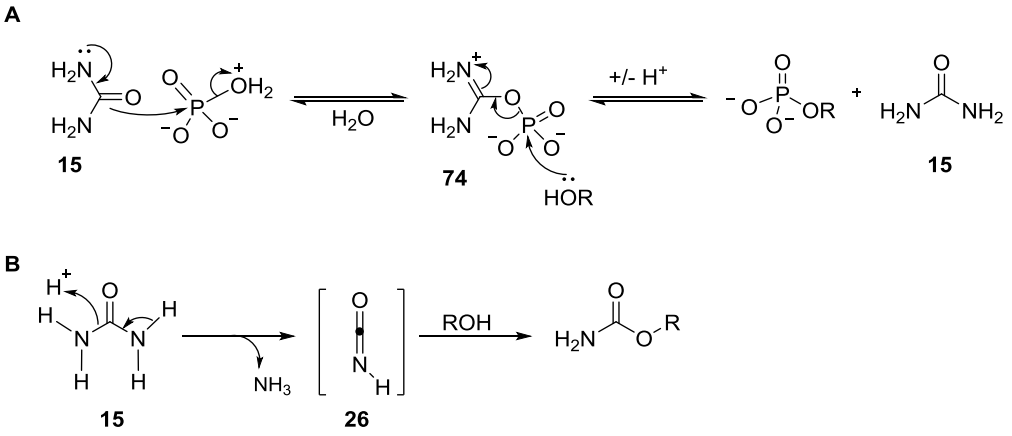
phosphorylations.[234] However, the most important contribution to this area was realised when Lohrmann and Orgel incorporated urea in the phosphorylating mixture.[231]



Scheme 4.2: *Dry state phosphorylation of nucleosides using excess urea and Na_2HPO_4, giving a complex mixture of carbamoylated and phosphorylated nucleotides and nucleosides.*

Urea **15** is a classic component of prebiotic chemistry, and one of the products of Urey-Miller discharge experiments, has been observed in both meteorite and interstellar dust clouds,[235] and is a by-product of pyrimidine ribonucleotide synthesis through the hydrolysis of cyanamide **68** under aqueous phosphate conditions.[22, 146] Importantly urea **15** melts at 140 °C to form a liquid that has excellent solvating properties for polar compounds and salts,[38, 236, 237] and can form a part of a range of deep eutectic solvents and ionic liquids.[230, 238] As well as urea's remarkable solvation properties it is also a nucleophilic catalyst for the displacement of water from phosphate under mildly acidic conditions, such that in the dry state the displacement of water from mono-basic phosphates by urea leads to the formation of ureidophosphates that can propagate the reversible transfer of phosphate moieties

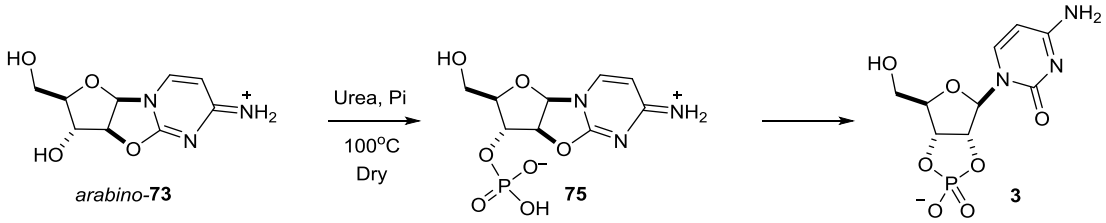
between alcohols (and water). Accordingly, Orgel reported that when water is evaporated from a phosphate/nucleoside/urea solution, leaving the nucleoside, phosphate and urea in close proximity for reaction,[231] incubation (100°C, 24 h) gave excellent yields for phosphate incorporation into nucleotides (> 96% for pyrimidine nucleotides),[231] but as a complex mixture of nucleotides with different phosphorylations patterns (Scheme 4.2). Interestingly, however, a strong preference for the formation of nucleotide-2',3'-cyclic phosphates and nucleotide-2',3'-cyclic-5'-bis-phosphates was observed over extended reaction times and with excess phosphate.[239] The exact mechanism for these phosphorylations is not known and Orgel proposed that the activation of phosphate was due to acid/base catalysis. However, it is not clear what is meant by this statement.[219, 240] It is more likely that urea displaces water from the orthophosphate to generate an activated ureidophosphate intermediate **74**, capable of transferring phosphate between hydroxyl moieties (Scheme 4.3A). This would account for both the reversibility of monophosphate synthesis and irreversible of cyclic phosphate synthesis.[22, 225, 241] Orgel also reported that carbamoylation was a major by-products in these urea mediated phosphorylations, but that addition of ammonium suppressed the carbamoylation of alcohols and the nucleobases. The carbamoylation is most likely due to the release of ammonia under elevated temperatures, to form the reactive intermediate, isocyanate **26**, which readily reacts with hydroxyl and amines (Scheme 4.3B).



Scheme 4.3: *Plausible mechanism for urea mediated phosphorylation and carbamoylation.* ***A)*** *Phosphorylation****:*** *urea eliminates water from the orthophosphate to generate an activated ureidophosphate transient intermediate* ***74****, which transfers phosphate to hydroxyl groups.* ***B)*** *Carbamoylation: Thermal decomposition of urea*

*releases ammonia and form isocyanate **26**, which readily reacts with the nucleophile to form carbamoylated compounds.*

The widely accepted prebiotic synthesis of pyrimidine ribonucleotides utilises Orgel's dry state phosphorylations to furnish 2´,3´-cyclic pyrimidine ribonucleotides, however with greater selectivity than free nucleosides (Scheme 4.4).[22] The key to these reactions was thought to be the use of cyclonucleosides (2,2´-anhydrocytidine *arabino*-**73)** for phosphorylations instead of ribonucleosides.[22] Efficient phosphorylation of *arabino*-**73** is achieved by drying an aqueous solution of *arabino*-**73**, urea and inorganic phosphate. Phosphorylation is then observed to be selective for the 3′-OH, leading into intramolecular rearrangement via **75** to the cytidine-2′,3′-cyclic phosphate **1** with the desired *β*-ribo-stereochemistry. Both kinetic and thermodynamic properties of this system control are thought to control the regioselectivity; $n \rightarrow \pi^*$ donation suppresses nucleophilicity of the C5′-hydroxyl and monophosphate synthesis is reversible, whereas 2′,3′-cyclic phosphates are generated by a different (associative phosphoryl transfer) mechanism and are synthesized irreversibly.[126, 146, 219, 231100, 242]



Scheme 4.4: *Selective urea-mediated 3´-phosphorylation of pyrimidine anhydronucleoside arabino-**73** giving β-2´,3´-cyclic phosphate cytidine **3** via the intramolecular cyclisation of the oxygen atom in the 3´-phosphate, β-2´,3´-cyclic phosphate.*

4.1. Phosphorylation of 8,2´-anhydro-cyclopurines

Based on the prebiotic synthesis of activated pyrimidine ribonucleotides[22] and our x-ray crystallographic data of **89** and **170** (Scheme 4.1), we hypothesised that treatment of 8,2´-anhydro purines (**89**, **90** and **170**) with orthophosphates in urea/formamide will chemoselectively phosphorylate 3´-OH, due to thermodynamic and kinetic effect. Kinetically, phosphorylation on 3´-OH is favoured because of an n→π* stereoelectronic effect between the 5´- oxygen atom and C8 carbon atom of 8,2´-cyclopurines (**89**, **90** and **170**), reducing electron density in 5´-O and reducing its electron density and increasing the steric encumbrance relative to the secondary 3´-OH.[126] Thermodynamically, 3´-phosphorylation will be favoured because the 3´-phosphates can form cyclic phosphates which are stable to these reaction condition, however the 5´-phosphate cannot form a cyclic phosphate and can therefore undergo dissociative phosphate transfer (back to urea to form ureidophosphate intermediate **74**).[243] We also hypothesised that the yields of 2´,3´-cyclic phosphates in purines ribonucleotides will be greater than in pyrimidine series due the suppression of purine C2 hydrolysis (see Section 3.3).

4.2. Synthesis of *β*-8-oxo- 2´,3´-cyclic phosphate purine ribonucleotides

We first investigated the phosphorylation of 8,2´-anhydro adenosine **89**. We started the investigation by repeating the fibre glass disc experiments reported in the phosphorylation of pyrimidine ribonucleotides, in which fibre glass discs are used to stimulate dry surface and maximise surface area to facilitate drying.[22] However when **89**, ammonium dihydrogen phosphate (1 equiv), and urea **15** (10 equiv) were dissolved in water (Method **A** of phosphorylation, see Experimental) and the resultant solutions were dried onto glass fibre disc after evaporation (40°C, 24h) and then heated at 100°C for 24 h, very disappointingly we were unable to detect any observable phosphorylation of **89** by ^{1}H or ^{31}P NMR spectroscopy. As purine ribonucleosides are significantly less soluble than pyrimidine ribonucleosides, and we had observed the readily crystallisation of **89**, **90** and **170** from dilute aqueous solution, it occurred to us that substrate precipitation may be inhibiting this reaction.

Therefore, a prebiotic polar organic solvent, mixed with water, might increase solubility of **89** (during drying) and allow **89** be solvated in the urea/phosphate mixture. Formamide **27**, the simplest naturally occurring amide (and the first hydrolysis product of HCN **14**) has previously been used in prebiotic phosphorylations and has excellent miscibility with water,[22, 242] and was chosen as our co-solvent. Formamide **27** can also be prebiotically produced from ultraviolet irradiation of icy mixtures of HCN **14**, NH_3 and H_2O,[244] or pyrolysis of mixtures of carbon monoxide (CO), NH_3 and H_2O.[245] Furthermore, formamide **14** has been detected in numerous abiotic environments including planets, satellites, meteorites, gas clouds and even dense stars.[246-248] Therefore the phosphorylation was repeated (Method **B** of phosphorylation, see Experimental), however in water/formamide (5:1) solution and pleasingly, ^{1}H NMR analysis of the product mixture showed clearly that **89** had undergone transformation to two major and several minor species (Figure 4.1A). Inspection of the ^{31}P NMR confirmed 45% phosphorus incorporation as 2´,3´-cyclic phosphates by the highly characteristic downfield shifted ^{31}P NMR resonance (18-25 ppm) that is indicative of 2´3´-cyclic phosphates, as well as mono- and bis-phosphates species (Figure 4.1B).

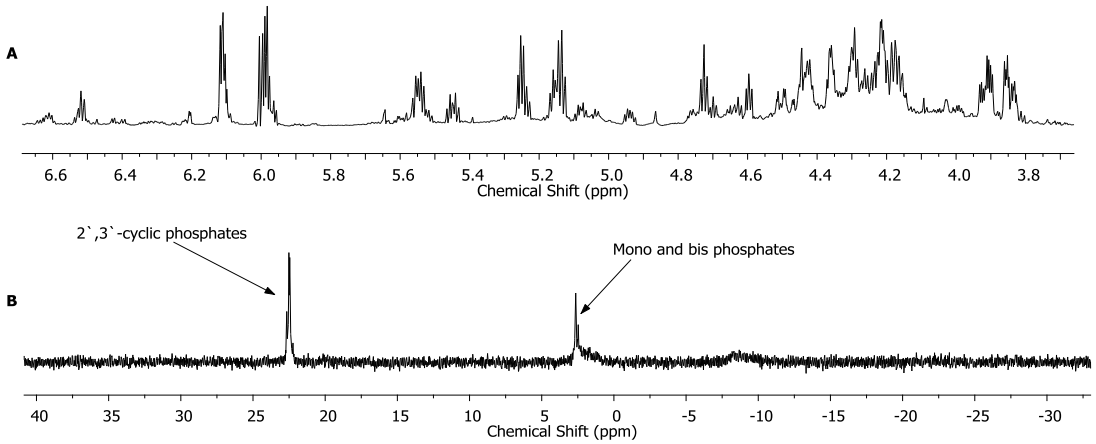


Figure 4.1: *^{1}H and ^{31}P NMR showing formation of complex mixture for the dried down phosphate reaction of 8,2´-anhydro-arabino-cycloadenine **89**, which used water/formamide solution adsorbed onto glass fibre discs (Method **B** of phosphorylation). **A)** 1H NMR (400 MHz, D_2O, 3.9 – 6.7 ppm) showing phosphorylation of **89** giving a mixture of complex species. **89** cannot be detect in the ^{1}H NMR. **B)** ^{31}P NMR (162 MHz, D_2O, -15 – 30 ppm) showing 2´,3´-cyclic phosphate species, mono and bis phosphorylated species.*

The successful phosphorylation of **89** under these conditions (Method **B**, See Experimental), gave us optimism that the 8-*O*-2´ anhydro linkage can be exploited

for chemical inversion and phosphorylation, as seen in the prebiotic synthesis of pyrimidine ribonucleotides.[22] However the ^{1}H and ^{31}P NMR for the mixture was complex and challenging to extract further specific information about the identities of the unknown 2´,3´-cyclic and mono phosphates, therefore a thorough investigation of the mixture was required. It was predicted that initially *β*-8-oxo-2´,3´-cyclic phosphate adenine **93** was formed in the reaction via the cyclisation of 3´-phosphate, **93** then undergoes bis phosphorylation and/or carbamoylation, this giving rise to a complex mixture of carbamoylated and phosphorylated nucleotide 2´3´cylcic phosphates. Although Orgel reported the observation of carbamoylation during dry phosphorylations of nucleosides, no data is available for these compounds, and therefore isolation and purification of the nucleotide mixture was necessary.

The dry state phosphorylation was scaled up using the formamide/water (5:1) solvation system (Method **B** phosphorylation in the experimental section) and purified by HPLC to reveal five major compounds, in agreement with the ^{31}P NMR of the crude mixture, which also showed 5 peaks in 23 ppm region associated with 2´,´3-cyclic phosphates. Pleasingly, all 5 nucleotides were successfully isolated as white crystalline solids (Figure 4.2). ^{1}H and ^{31}P NMR spectroscopy analysis revealed that all 5 compounds were 2´,3´-cyclic phosphates of 8-oxo-adenosine. The presence of 2´,3´-cyclic phosphate was could be readily confirmed by ^{31}P-HMBC NMR spectroscopy, which clearly showed coupling of C2´-H and C3´-H with phosphorus for all 5 compounds. Furthermore, ^{31}P also coupled with ^{13}C, and pleasingly phosphorus-carbon coupling was also observed in ^{13}C NMR spectroscopy of all 5 compounds.[249] To narrow down the identification process, high resolution mass spectroscopy analysis was also used to assign the functional groups and the presence of oxo-moiety at C8 was confirmed by the appearance of a sharp peak at 1650 cm^{-1} in IR spectrum.

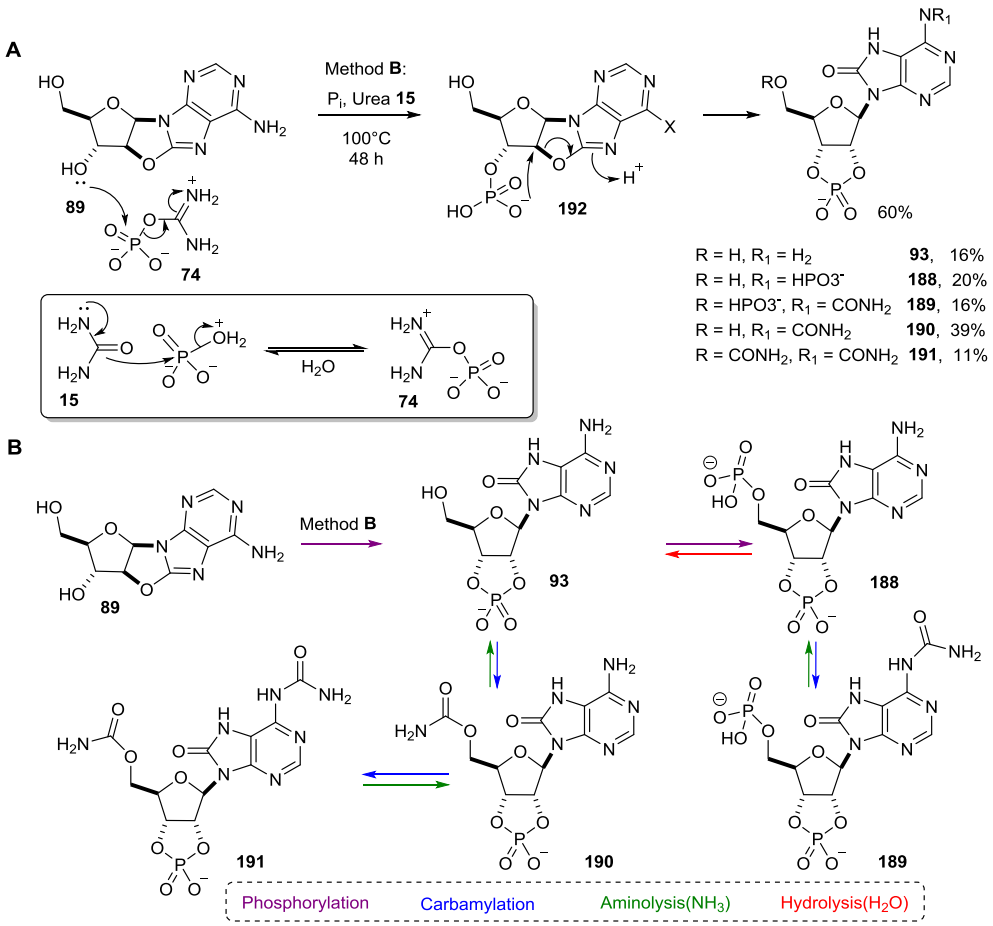


Figure 4.2: A) *Dry state phosphorylation of 8,2´-anhydro-cycloadenine* ***89*** *by inorganic phosphate in the presence of urea (Method* ***B*** *of phosphorylation) furnishing 60% 2´,3´-cyclic phosphate adenines. The individual yields shown are only for 2,´3-cyclic phosphates determined by the HPLC trace. Postulated mechanism for activation of phosphate by urea (inset). Enrichment of the nucleotidic mixture with β-8-oxo-2´,3´-cyclic phosphate adenosine* ***93*** *by hydrolysis and aminolysis of acyclic phosphates (****188*** *and* ***189****) and carbamoylated species (****190*** *and* ***191****) respectively. The colour scheme for the arrows is as follow: phosphorylation, carbamoylation, aminolysis and hydrolysis.*

Interestingly ^{1}H NMR analysis indicated that all 4 compounds, with exception of **93**, had downfield shifted C5´-proton resonances (4.0-4.2 ppm, normally 3.7-3.9 ppm), indicating attachment of an electron withdrawing group at C5 and high resolution mass spectroscopy indicated these to be either carbamoylation ($CONH_2$) as compounds **189** and **191** or phosphorylation (HPO_3^-) in compounds **188** and **189** (Figure 4.3). Further evidence for the structure of **188** and **189** was found in the dt (doublet of triplets) splitting pattern of the C5´-proton resonances in their ^{1}H NMR spectrum indicating a ^{1}H and ^{13}P coupling (which was confirmed by ^{1}H-^{31}P HMBC cross coupling patterns), while C5´-H signal for **190** and **191** showed the characteristic ABX system, inherent in C5´-OH coupling of nucleoside moieties. Finally, the

downfield shift observed of the nucleobase C2-proton in compounds **190** and **191** suggested that *N*6-carbamoylation had likely occurred.

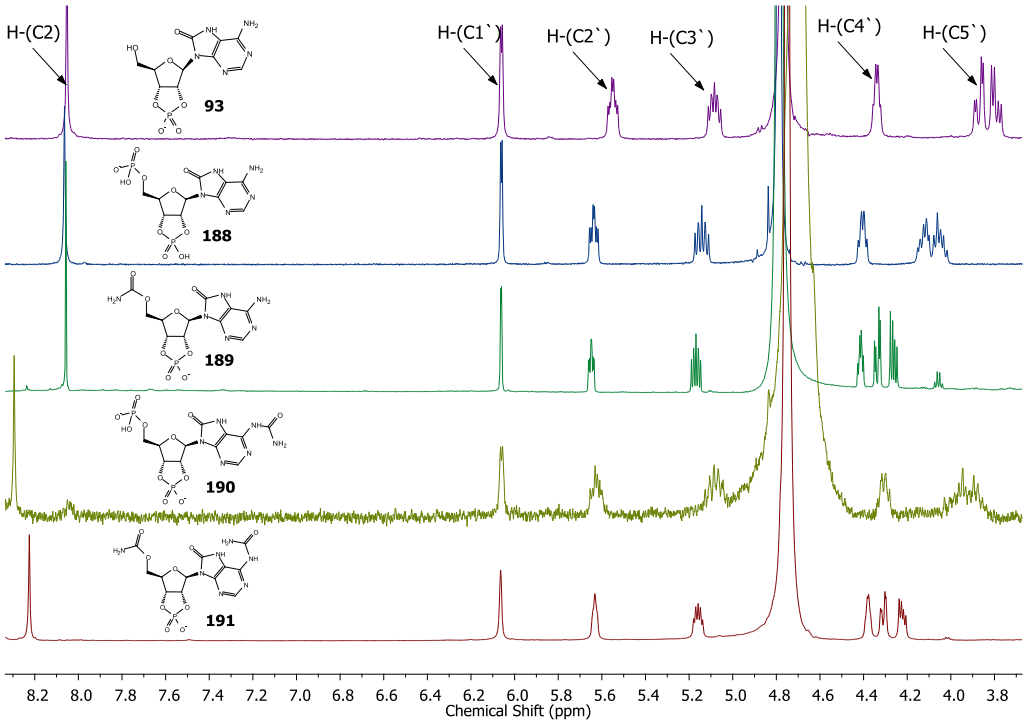


Figure 4.3: *^{1}H NMR (600 MHz, D2O, 3.7-8.3 ppm) for β-8-oxo-2´,3´-cyclic phosphate adenine **93** and its bis phosphorylated and carbamoylated products (**188-191**). The products were isolated from the nucleotidic mixture of the dry state phosphorylation of **89**. The downfield shift in C5´-H is due to presence of HPO_3^- or $CONH_2$ is clearly shown,as well as a visibly downfield shift of C2-H by the carbamoylation of NH_2 in the nucleobases of **190** and **191**.*

Although the dry state phosphorylation (Method **B**) of 8,2´-anhydro-arabino-cycloadenione **89** produced a mixture of nucleotides (**188-191**), all the derivatives produced were 2´,3´-cyclic phosphates and other derivatives are deemed to be reversible to the core nucleotide **93**.[83, 225, 231, 239] Orgel *et al.* reported that use of ammonium salts in dry phosphorylation supressed carbamoylation.[231, 239] Therefore we repeated the dry down phosphorylation experiment (Method C of phosphorylation in experimental) with the addition of ammonium chloride (1 equiv) to the phosphorylating mixture. Now ^{31}P NMR analysis indicated that only two major 2´,3´-cyclic products (presumably **93** and **188**) were formed, alongside two other minor products thought to be acyclic nucleotides of adenine. Importantly, however, a significant decrease in the carbamoylation of **89** was observed.

The dry state phosphorylation of **89** (Method **B**) furnished approximately 30% more 2´,3´-cyclic phosphates than the phosphorylation of adenosine as reported by Orgel.[231] However, in order to further increase the yield and selective of these phosphorylation, as well as reduce the reaction time to expedite further investigation, we decided to directly explore phosphorylation in a urea melt. Urea **15** melts have been employed as a solvent since 1950 and are observed to yield water-like solvation properties.[236] Urea melts have also found use in the prebiotic synthesis of polypeptides.[250] Accordingly we set out to investigate phosphorylation in urea-melts. A rapid screen of urea stoichiometry and melting temperatures indicated that the optimal conditions for phosphorylation required a large excess of urea **15** and incubation at 140°C, under these condition (for example when 8,2´-anhydro cycloadenine **89**, ammonium dihydrogen phosphate (1 equiv.) and urea **15** (27 equiv.) were heated at 140°C (Method **D**)), ^{1}H and ^{31}P NMR spectroscopic analysis revealed that 22% **93** and 33% **188** where synthesised in only 20 minutes (Figure 4.4).

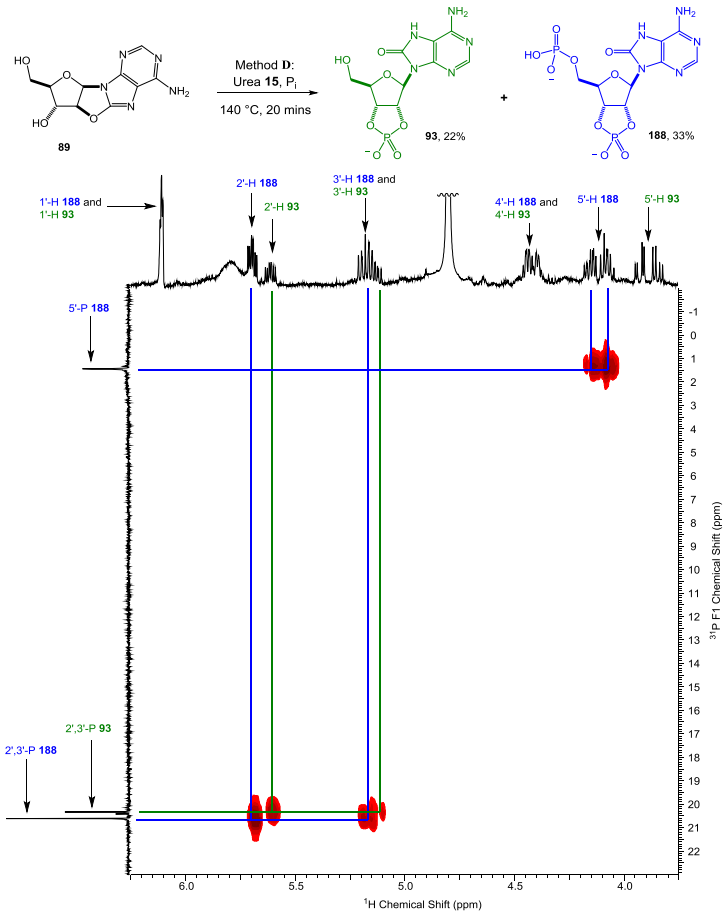
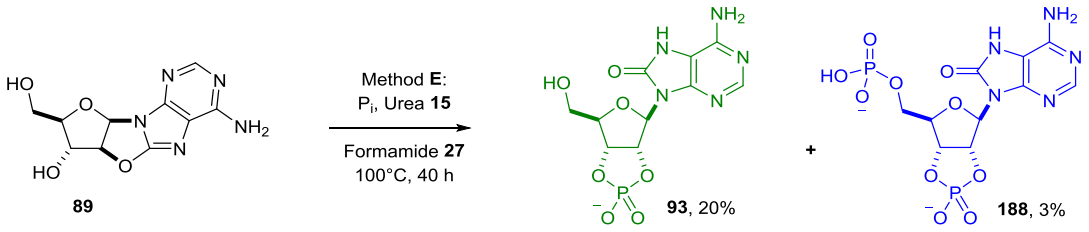


Figure 4.4: *^{1}H - ^{31}P HMBC NMR spectrum (400/162 MHz, D_2O, 3.75 – 6.25; -2.0 – 23.0 ppm) to show phosphorylation of 2',8-anhydro-arabino-cycloadenosine **89** by method **D**.*

To investigate the practicality of formamide **27** solvent system in the prebiotic phosphorylation of 8,2´-anhydro cyclopurines, of 8,2´-anhydro adenosine **89**, ammonium dihydrogen phosphate and urea were dissolved in formamide **27**. The reaction mixture was heated to 100°C for 48 h (Method **E** of phosphorylation in the experimental section). Pleasingly, phosphorylation of 8,2´-anhydro cycloadenosine **89** in formamide was very selective, furnishing 20% **93** and only 3% **188 (**Scheme 4.5).



Scheme 4.5: *Phosphorylation of 2',8-anhydro-arabino-cycloadenosine* ***89*** *by method* ***E***.

Having developed our model urea-mediate phosphorylation methods for cycloadenine **89**, we next investigated the phosphorylation of cycloguanosine **90** and cycloinosine **170**.

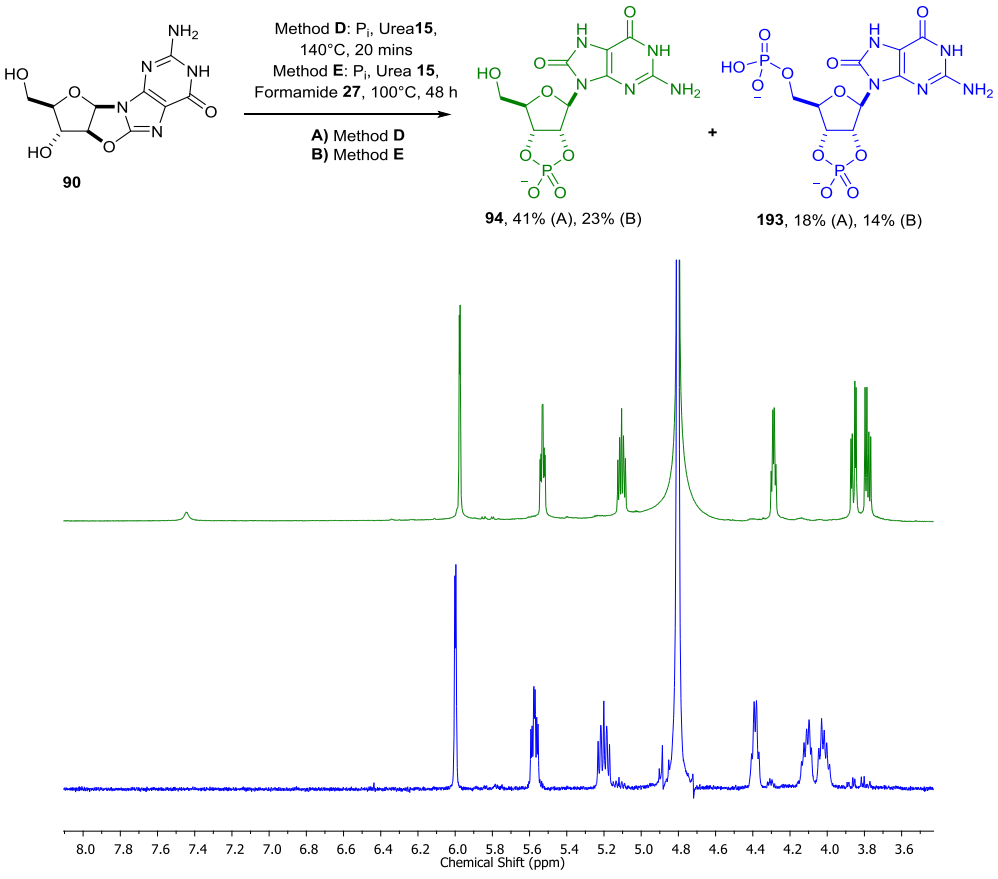


Figure 4.5: *Phosphorylation of 8,2´-anhydro cycloguanosine* ***90*** *using method* ***D*** *and* ***E*** *furnishing β-8-oxo-guanine-2´,3´-cyclic phosphate* ***94*** *(41%, 18% respectively) and*

*β-8-oxo-guanine-2´,3´-cyclic-5´-bisphosphate **193** (23% and 14% respectively) also showing ^{1}H NMR (600 MHz, D_2O, 3.5-8.3 ppm) of HPLC purified **90** and **193**.*

Following method **D**, cycloguanosine **90** was phosphorylated to furnish *β*-8-oxo-guanine-2´,3´-cyclic phosphate **94** and *β*-8-oxo-guanine-2´,3´-cyclic-5´-bisphosphate **193** in 41% and 18% yield respectively. To investigate phosphorylation of 8,2´-anydro-*arabino*-cyclo guanosine **90** in formamide **27** solution, **89** was exposed to phosphorylating conditions of method **E**. Pleasingly, ^{1}H and ^{31}P NMR studies indicated, 23% **94** and 14% **193**, giving an overall phosphorus incorporation of 37%. The cyclic phosphates of guanosine **94** and **193** were purified via HPLC and full characterisation data was obtained. NMR data of **94** and **193** was also compared with *β*-8-oxo-adenine-2´,3´-cyclic-5´-bisphosphate **93** and *β*-8-oxo-adenine-2´,3´-cyclic-5´-bisphosphate **188** respectively to confirm structures of **94** and **193** (Figure 4.5).

Next we concentrated on the phosphorylation of 8,2´-anhydro-*arabino*-cycloinosine **170**. Remarkably, application of method **D** and **E** on **170** give total 2´,3-cyclic phosphates in 70% and 57% yields respectively (Figure 4.6).

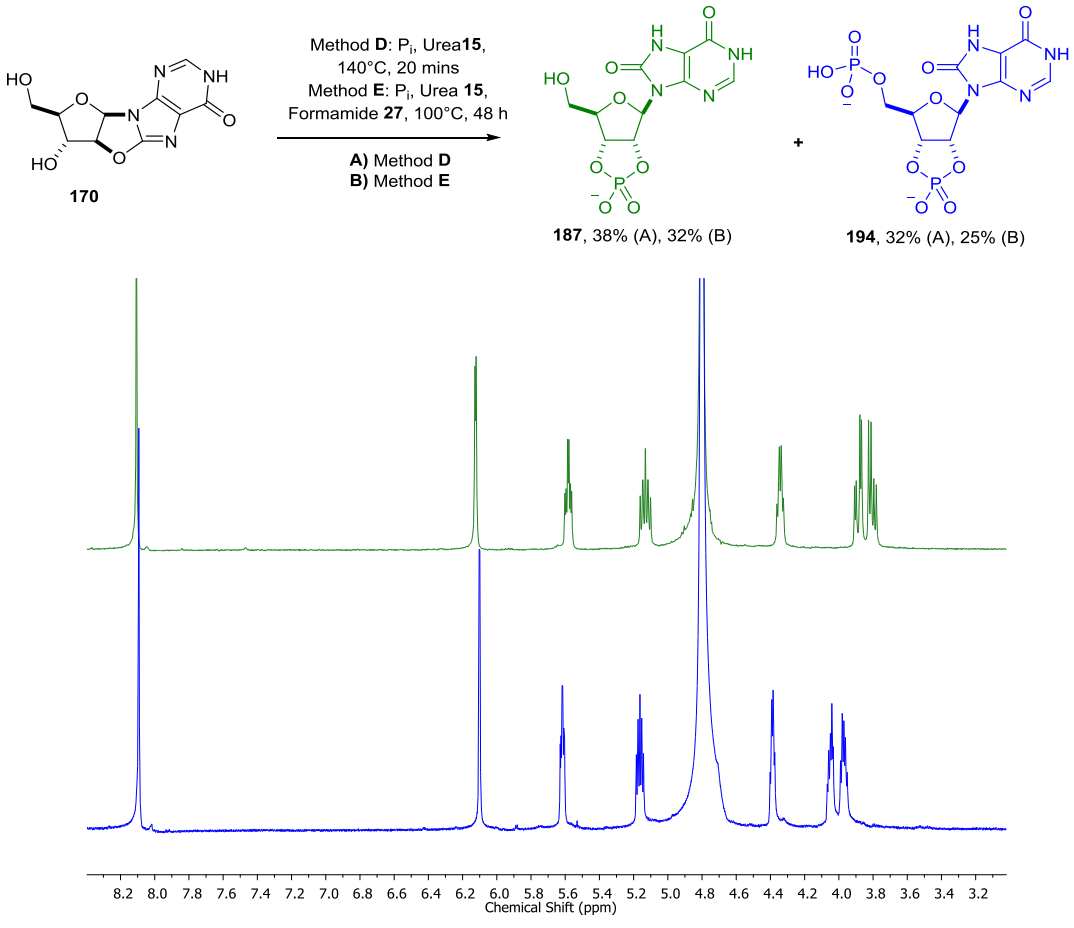
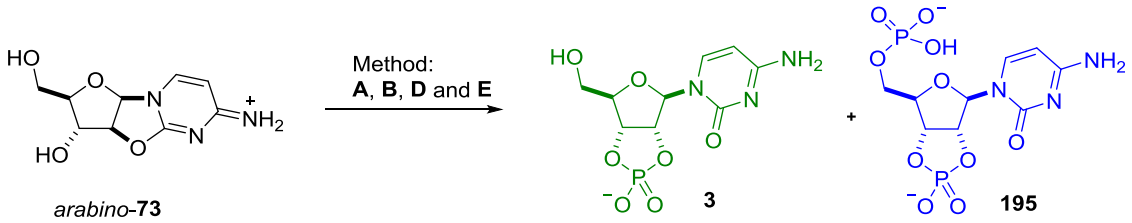


Figure 4.6: *Phosphorylation of 8,2´-anhydro cycloinosine **170** using method **D** and **E** furnishing β-8-oxo-inosine-2´,3´-cyclic phosphate **187** (38%, 32% respectively) and β-8-oxo-inosine-2´,3´-cyclic-5´-bisphosphate **194** (32% and 25% respectively) also showing ^{1}H NMR (600 MHz, D_2O, 3.5-8.3 ppm) of HPLC purified **187** and **194**.*

The incredible 70% phosphorus incorporation in **170** once again show that the obvious prebiotic phosphorylation route must have proceeded via the chemical inversion of 8-*O*-2´ linkage in cyclo anhydro nucleosides and not via the phosphorylation of nucleosides (cytidine, uridine, adenosine, guanosine and inosine) which are unselective and low yielding. Inosine is a wobble base pair.[200] Wobble base pairs play a fundamental role in RNA secondary structures and for the correct translation of the genetic code. Thermodynamically wobble base pair can be compared with Watson-Crick base pairing. Furthermore the incredible similarity (only differing by oxidation states) of inosine to adenosine and guanosine make inosine worthwhile candidate of prebiotic chemistry.

It was deemed necessary to repeat Powner *et al.* phosphorylation of 2,2´-anhydrocytidine *arabino*-**73** under the conditions applied to purine **89**, **90** and **170**. These studies will be important if a one pot prebiotic synthetic route of both purine and pyrimidine ribonucleotides is to be pursued, a topic that will be discussed later on in this thesis. Pleasingly, methods **A**, **B**, **D** and **E** all gave comparable phosphorylation yield to those previously report by Powner *et al.*, giving a total 2´,3-cyclic phosphate incorporation of 30, 46, 65 and 48% respectively (Figure 4.7).

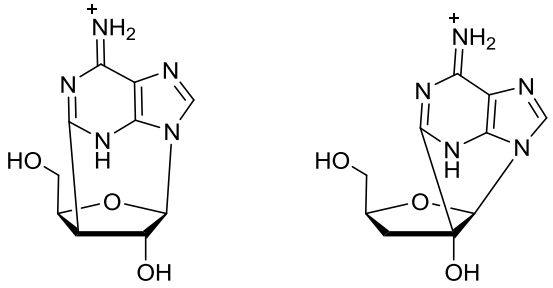


Method	*β*-cytidine-2´,3´-cyclic phosphate 3 (%)	*β*-cytidine-2´,3´-cyclic-5´ bisphosphate 195 (%)	Total 2´,3´ P (%)
A	18	12	30
B	35	11	46
D	47	18	65
E	30	18	48

Figure 4.7: *Incubation of 2,2´-anhydro cytidine arabino-**73** utilising method **A**, **B**, **D** and **E**. Table to show yields of β-cytidine-2´,3´-cyclic phosphate **3** and β-cytidine-2´,3´-cyclic-5´ bisphosphate **195** under **A**, **B**, **D** and **E**.*

4.3. Phosphorylation of *β*-2´,3´-epoxy-8-oxo-adenine

During our studies of cyclopurines we observed that cycloadenine **89** is remarkably resistant under alkaline conditions (Section 3.2). Even upon extended incubation at elevated pH **89** was not observed to undergo hydrolysis, instead, **89** undergoes isomerisation to 8,5´-anhydronucleoside **182** (55%, pH 11, 40°C, 24 h) and 2´,3´-anhydronucleoside **183** (60%, pH 13, 40°C, 24 h) (Figure 3.9A). Given these observed isomerisation (*albeit* only at extreme pHs) we deemed it prudent to investigate phosphorylation of these isomers. It was hypothesised that **182** and **183** would undergo isomerisation with **89** under the condition of phosphorylation to re-establish the C2'-*O*-C8 linkage leading to the same (or similar) product distribution observed for the phosphorylation of **89**. There is also literature precedence for formation of both *N*3-C2´ and *N*3-C3´ bond in cyclopurine ribonucleosides, *albeit* under conventional organic chemistry conditions (Scheme 4.6).[199, 205, 251, 252]



*N*3-C3` *N*3-C2`

Scheme 4.6: *N3-C2´ and N3-C3´ bond in cyclopurine ribonucleosides.*

However, upon C3'-OH and C2'-OH phosphorylation, respectively, the *N*3-anhydronucleotides would also be anticipated to undergo intramolecular rearrangement to furnish 2',3'-cyclic phosphates. To examine our hypothesis, the epoxide **183** was exposed to phosphorylation conditions of method **D**. Pleasingly, ^{1}H and ^{31}P NMR spectroscopic analysis both showed formation of 2´,3´-cyclic phosphate (in 23% yield; as a mixture of adenines **93** and **188**) (Figure 4.8).

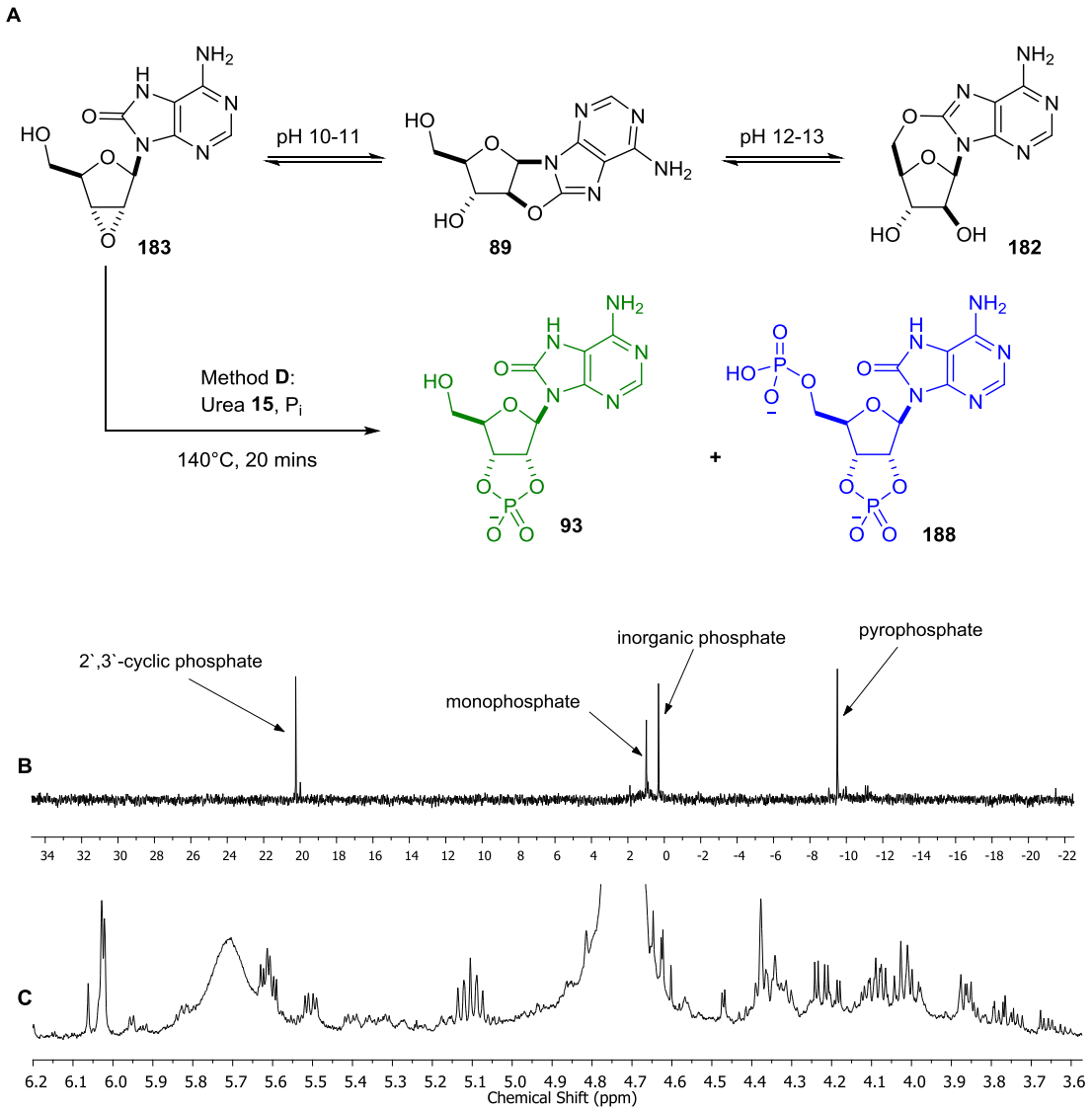
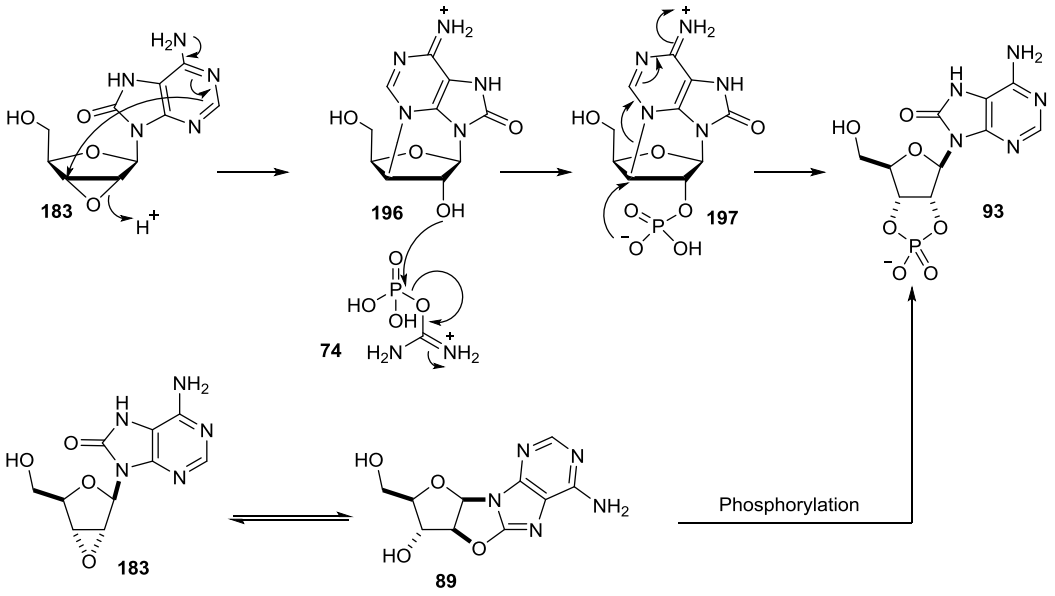


Figure 4.8: *Phosphorylation of isomers of 8,2´-anhydro adenine* ***89****.* ***A)*** *phosphorylation (Method* ***D****) of epoxy 2´,3´-anhydro adenine* ***183*** *to form β-8-oxo-2´,3´-cyclic phosphate adenine* ***93*** *and β-8-oxo-2´,3´-cyclic-5´ bisphosphate adenine* ***188****.* ***B)*** *^{31}P NMR (162 MHz, D_2O, -25 – 40 ppm) showing ~ 23% 2´,3´-cyclic phosphates (assumed to be* ***93*** *and* ***188****.* ***C)*** *^{1}H NMR (400 MHz, D_2O, 3.5 – 6.3 ppm) for phosphorylation of* ***183*** *via method* ***D****, C2´-H and C3´-H coupling with phosphorus can also be observed.*

The phosphorylation mechanism is thought to proceed either with re-equilibration with **89** and subsequent C3'OH phosphorylation or by nucleophilic addition of purine-*N*3-nitrogen atom to C3´ carbon and concomitant epoxide ring opening, to form **196**. The newly formed 2´-OH in **196** can then be phosphorylated and by a second S_N2-type nucleophilic attack of the C2´-phosphate at C3' carbon atom, formation of *β*-8-oxo-adenine-2´,3´-cyclicphosphate **89** would be achieved by a double stereochemical inversion at C3' stereo-centre (Scheme 4.7).



Scheme 4.7: *Plausible mechanism for phosphorylation of 2´,3´-epoxy adenosine* ***183*** *with inorganic phosphate in the presence of molten urea or formamide solvent.*

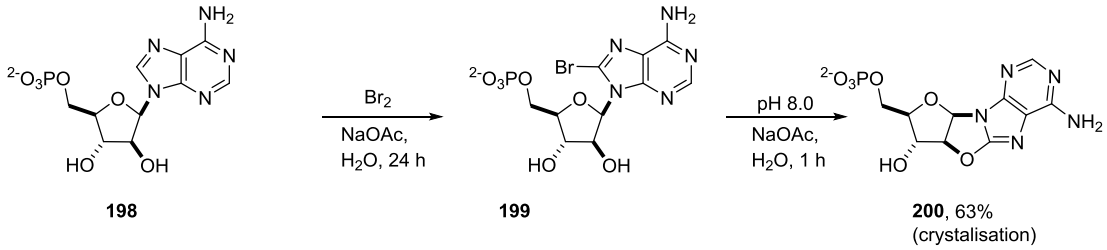
Due to time limitation, this topic is outside the scope of this thesis, and no further investigation of the specific mechanism for phosphorylation of **183** was undertaken. However our results indicate that prebiotic phosphorylation of 2´,3´-epoxy adenosine **183** yield 2´,3´-cyclic phosphates, further demonstrating the predisposed nature of the 2´,3´-cyclic phosphate moiety under urea-mediated phosphorylation conditions.

4.4. 5´→ 3´ phosphorus migration in 5´-phosphate cyclonucleotides

During our phosphorylation studies we were unable to apply method **A** or **B** to cyclopurines due to solubility issues during the drying phase of the reaction process. It therefore occurred to us that 5´-phosphorylation of these cyclopurines may increase their solubility in water due to ionic nature of the phosphate moiety. Moreover, we hoped that phosphoryl transfer across the nucleotide (5´-phosphate to 3´-phosphate) would give superior phosphorus integration. We suspected that the same methodology we had already developed for the synthesis of 8,2´-anhydro purines could be directly used to synthesis its 5'-phosphorylated analogues.

Therefore we next incubated commercially available arabinose adenine 9-*β*-D-arabinofuranoside-5´-monophosphate **198** in sodium acetate buffer at room temperature was treated with saturated bromine water, to oxidatively brominate the

C8 carbon atom. Cyclisation was induced at pH 8.0 upon addition of NaOH to produce 8,2´-anhydro-arabino-5´-phosphate adenine **200**. Surprisingly, crystals of **200** (63%) again spontaneously formed in the crude mixture as colourless sheets (Scheme 4.8).



198 **199** **200**, 63% (crystalisation)

Scheme 4.8: *Direct bromination of arabino-5´-phosphate adenosine **198** in bromine water to yield 8-bromo-arabino-5´-phospahte adenosine **199** which cyclizes under mild alkaline conditions to furnish 8,2´-anhydro-arabino-5´-monophosphate adenine **200** (63%) via the elimination of bromine. **200** spontaneously crystalizes from the crude solution.*

Spontaneous crystallisation is an excellent prebiotic tool for purification.[134] However we were surprised to observe **200** crystallise directly from aqueous solution, phosphorylated compounds usually being highly soluble as their sodium salts in water and often difficult to crystallise. However, **200** rapidly precipitated from the crude mixture, indicating a potentially plausible prebiotic purification strategy for 5´-phosphorylated 8,2´-anhydro purines. As expected the C5´-proton resonances for **200** were shifted downfield and the correct multiplet splitting due to ^{1}H-^{31}P coupling was observed as well as ^{31}P NMR and ^{1}H-^{31}P HMBC spectra confirming the presence of the C5´-phosphate in **200** (Figure 4.9).

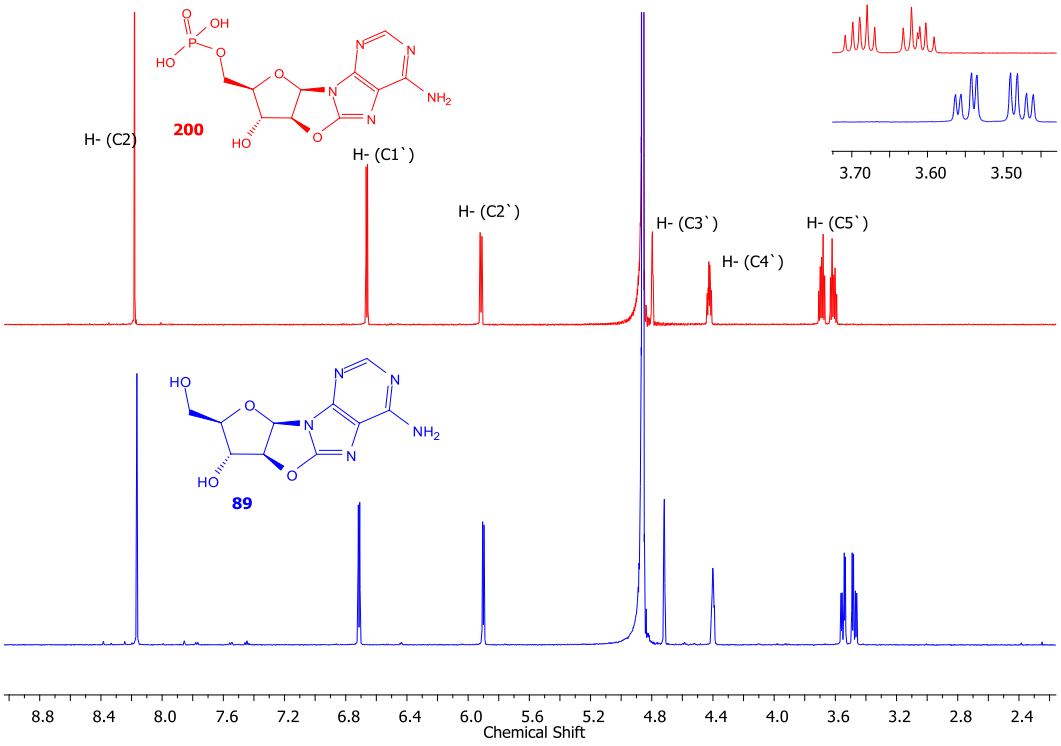
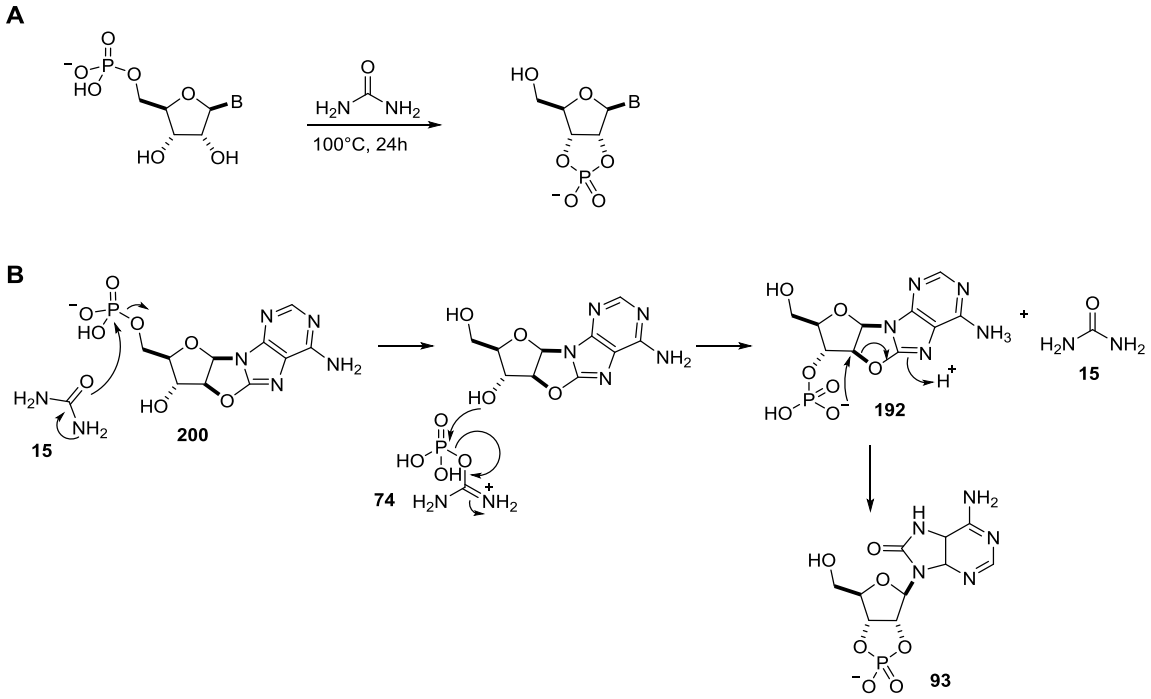


Figure 4.9: *Comparison ^{1}H NMR (600 MHz, D_2O, 2.2 – 9.0 ppm) of 8,2´-anhydro-cycloadenine **89** with **200**. Expanded NMR indicating ^{1}H-^{31}P coupling in **200** has shifted the C5´-H to the left (down field) and given rise to multiple splitting.*

Orgel had previously demonstrated that when 5´-monophosphate nucleotides are heated in urea **15** a mixture of nucleotide product are observed, and 2´,3´-cyclic phosphates are the major product. Orgel's results suggest that 5´→3´ phosphate migration occurs under these condition (Scheme 4.9A).[231, 239] However, to our knowledge, there are no literature reports of 5´→3´ phosphorus migration in cyclonucleotides. We hypothesized that incubation of **200** in urea would yield 2´,3´-cyclic phosphates via the intermediacy of ureidophosphate complex **74** (Scheme 4.9B). We were particularly interested by this transformation, because installing the phosphate on cyclonucleoside prior to dry heating would allow the accurate control of nucleoside/phosphate stoichiometry and likely assist in the co-localisation of phosphate at the site of required phosphorylation.



Scheme 4.9: A) *Orgel's direct heating of 5´-phosphate nucleotides in urea* ***15*** *to furnish 2´,3´-cycli phosphates.* ***B)*** *Proposed mechanism for formation of 2´,3´-cyclic phosphates via 5´→3´ phosphoryl transfer involving ureidophosphate complex* ***74*** *when 8,2´-anhydro-arabino-5´-monophosphate adenine* ***200*** *is heated in urea* ***15*** *in the absence of stoichiometric inorganic phosphate.*

Therefore, to test 5´→3´ phosphorus migration in cyclonucleosides, **200** was put through phosphorylating conditions of method **D** (without inorganic phosphate). Interestingly, ^{31}P NMR spectrum clearly indicated the formation of 2´, 3´-cyclic phosphate even in the absence of inorganic phosphate. The ^{1}H NMR spectroscopic analysis indicated at least 2 cyclic phosphate species, presumed to be 8-oxo-2´,3-cyclic phosphate adenosine **93** and its 5´-bisphosphorylated variant **188** (Figure 4.10).

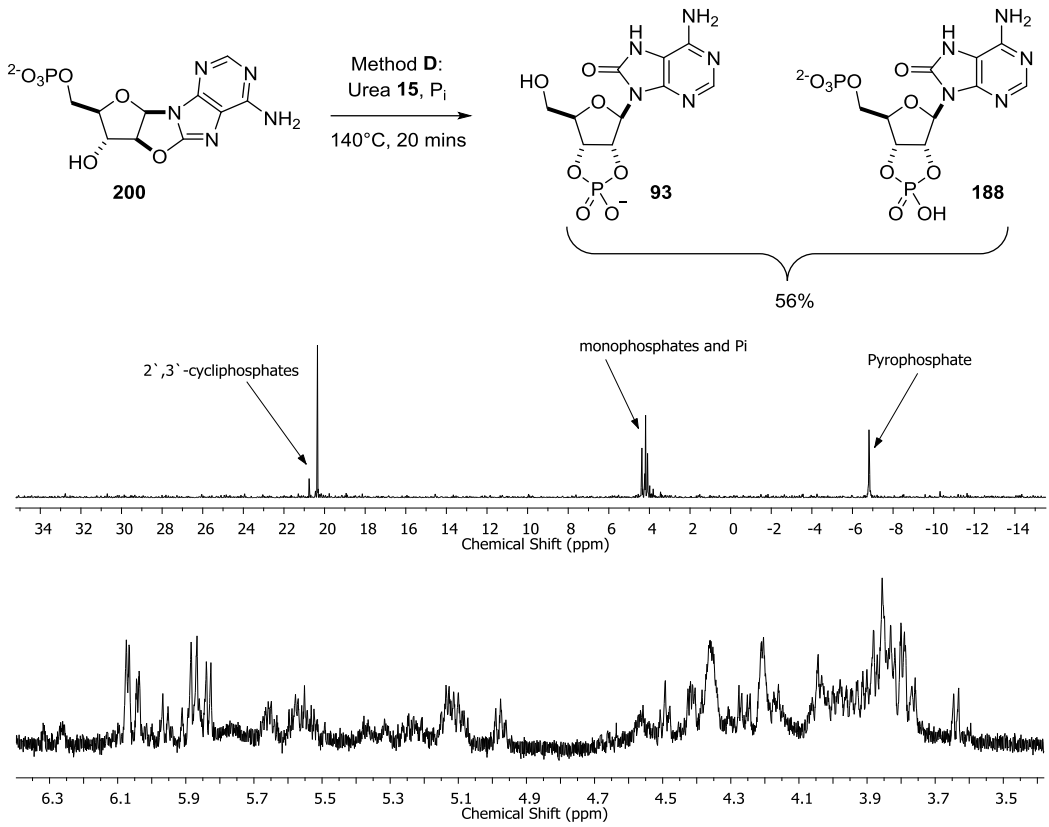
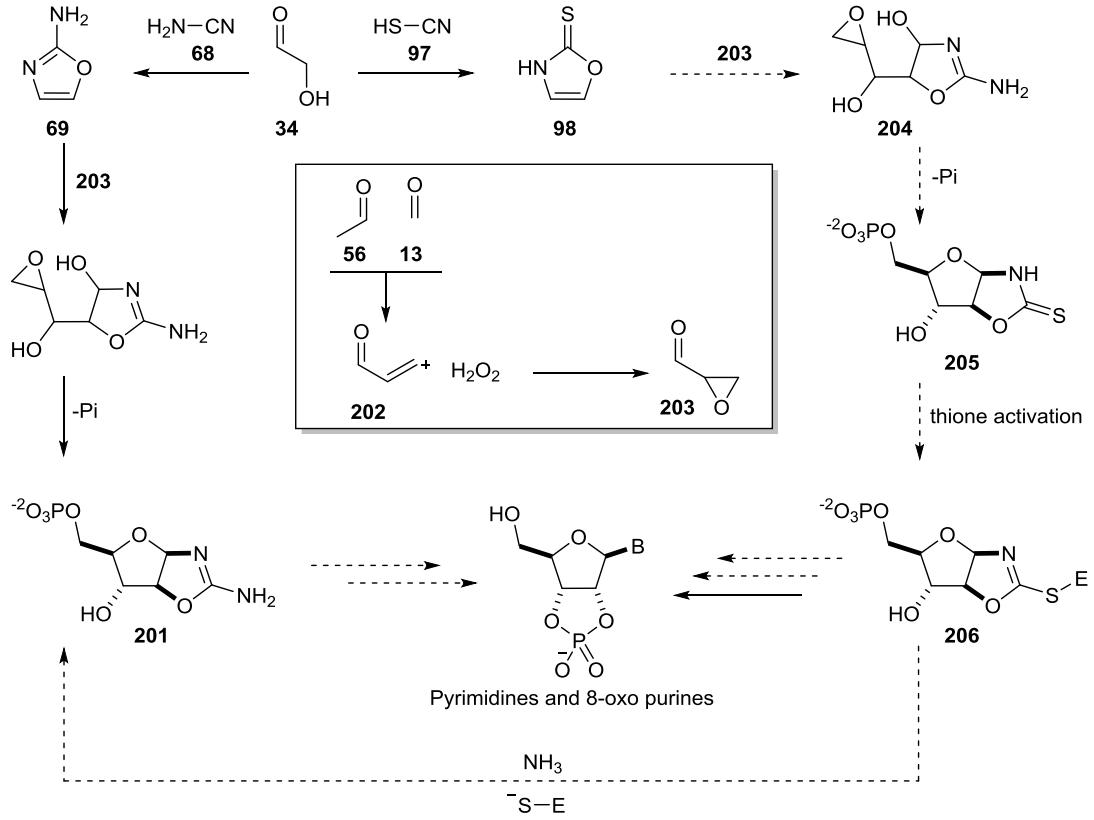


Figure 4.10: *^{1}H NMR (400 MHz, D_2O, 3.4 – 6.4 ppm) and ^{31}P NMR (162 MHz, D_2O, -14–36 ppm) showing incubation of 8,2´-anhydro-arabino-5´-monophospahte adenine **200** in urea at 140°C for 20 min (Method **D**). Both NMRs are clearly showing formation of 2´,3´-cyclic phosphates (resumed to be **93** and **188**). Pyrophosphates also detected in the phosphorus NMR, indicating some dephosphorylation of **200**.*

From these initial experiments it has been demonstrated that 5´-phosphates such as **200** can furnish 2´,3´-cyclic phosphates (Method **D**). This potentially opens up a stoichiometrically controlled way of introducing phosphorus into pyrimidine and purine nucleotides, however for this to happen a prebiotic synthesis of 5´-phosphate nucleotides is required. Prebiotic synthesis of aminooxazoline-5´-phosphates **201** has already been established in the Powner laboratory by oxidative phosphorylation.[174] Crucial to this synthesis is acrolein **202**, which can be prebiotically synthesized from acetaldehyde **31** and formaldehyde **13**. Acrolien **202** is also implicated in the prebiotic synthesis of several amino acids.[253, 254] It seem possible that further development of this oxidative phosphorylation strategy may lead to synthesis of 5´-phosphorylated anhydro purines as well as the pyrimidine that have been investigated already (Scheme 3.22).

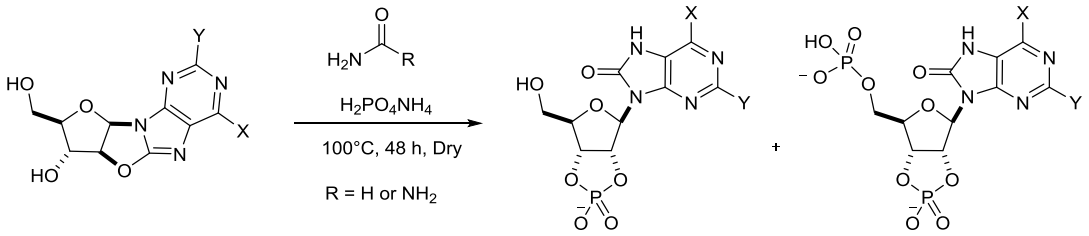


Scheme 4.10: *A plausible one pot synthesis of pyrimidines and 8-oxo purines by application of 5´-phosphorylation chemistry. Dashed arrows, proposed chemistry and solid arrows, reported chemistry. Arabino-5´aminooxazoline* ***201*** *could also be synthesised via the reaction of activated thione* ***206*** *with ammonia,* ***206*** *will be acting as the key precursor both for pyrimidine and 8-oxo purine synthetic pathways, leading to a (potentially) plausible one pot (prebiotic) synthesis of pyrimidine and purine ribonucleotides via 5´-phosphorylated intermediates.*

4.6. Phosphorylation summary

From these phosphorylation studies we have successfully established that 8,2´-cyclopurines can be phosphorylated to furnish 8-oxo-2´,3´-cyclic phosphates in good yields (32-70%, Figure 4.11). We have also demonstrated that these reactions can be scaled up and the products isolated so that further chemistry of these valuable compounds can be elucidated. We have also established that 8,2´-anhydro purines are not prone to alkaline hydrolysis (pH 11 - 13, 40°C) instead they isomerise to furnish 8,5´-anhydro and 2´,3´-epoxy purines. We have successfully shown that the epoxide

isomer also forms 2´,3'-cyclic phosphates under urea mediated phosphorylation conditions.



89 Y = H, X = NH_2
90 Y = NH_2, X = OH
170 Y = H, X = OH

93 Y = H, X = NH_2
94 Y = NH_2, X = OH
187 Y = H, X = OH

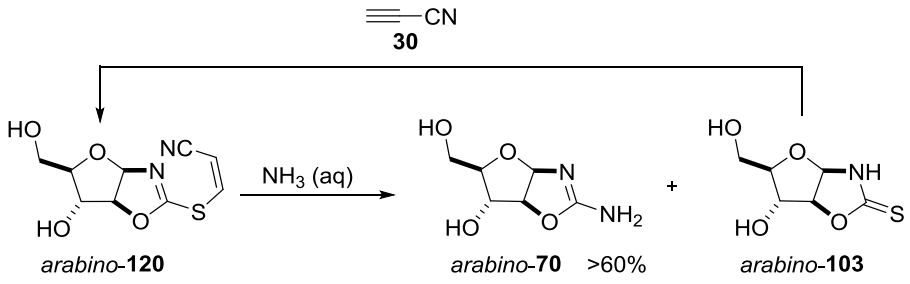
188 Y = H, X = NH_2
193 Y = NH_2, X = OH
194 Y = H, X = OH

Anhydro-nucleoside	Method	2′,3′-cyclic phosphate (%)	2′,3′-cyclic-5′-bisphosphate (%)	Total (%)
2′,8-*O*-anhydro cycloadenosine **89**	**D**	*22*	*33*	*55*
	E	*24*	*8*	*32*
2′,8-*O*- anhydro cycloguanosine **90**	**D**	*41*	*18*	*59*
	E	*23*	*14*	*37*
2′,8-*O*-anhydro cycloinosine **170**	**D**	*38*	*32*	*70*
	E	*32*	*11*	*43*
Ancitabine *arabino*-**73**	**D**	*47*	*18*	*65*
	E	*44*	*19*	*63*

***Figure 4.11**: Summary table of yields of 2´,3-cyclic phosphate formation from of cyclonucleoside **89**, **90** and **170** under general methods **D** and **E**. Yields were calculated based on comparison with an internal standard (DSS).*

5. Divergent prebiotic synthesis of activated pyrimidines and 8-oxo purines

Given the importance of the biochemical interplay between purines and pyrimidines, we next sought to investigate the concomitant synthesis of both classes of nucleotides together. Previous attempts to find a one pot synthesis of purine and pyrimidine nucleotides has been unsuccessful (Section 1.7.1).[128, 129, 146] We have already shown in this thesis that displacement of *S*-cyanovinyl oxazolidinone thione *arabino*-**120** by aqueous ammonia gave aminooxazoline *arabino*-**70** (section 2.5.1). Furthermore, it was shown that the by-product oxazolidinone thione *arabino*-**103,** formed during displacement reactions, can be activated with cyanoacetylene **30** to regenerate *arabino*-**120** (even in the presence of *arabino*-**70** due to increase nucleophilicity of thione *arabino*-**103** with respect to aminooxazoline *arabino*-**70**). Subsequent to re-cyanovinylation of thione *arabino*-**103** and displacement with aqueous ammonia, the yield can be up to 60% yields of aminooxazoline *arabino*-**70**, the key pyrimidine ribonucleotide precursor (Scheme 5.1).[22]



Scheme 5.1: *Reaction of S-cyanovinyl oxazolidinone thione arabino-**120** with aqueous ammonia in the presence of cyanoacetylene **30** to furnish 60% aminooxazoline arabino-**70**. The thione arabino*-**103** produced during displacement reaction**s** can be re-cyanovinylated to regenerate *arabino*-**120.**

For a concomitant reaction pathway to furnish both pyrimidine and purine ribonucleotide it has be to established that the different components of pyrimidine and

purines synthesis can be synthesised together under the same conditions, and this has never been achieved. For example, *S*-cyanovinyl thione *arabino*-**120** can be chemoselectively formed in the presence of aminooxazoline *arabino*-**70**, and upon addition of cyanoacetylene **30** (250mM) to an aqueous solution of oxazolidinone thione *arabino*-**103** (240mM) and *arabino*-**70** (240mM) at pH 7 after 1 h at room temperature a quantitative conversion of *arabino*-**103** to *S*-cyanovinyl thione *arabino*-**120** was observed whilst leaving *arabino*-**70** untouched (Figure 5.1).

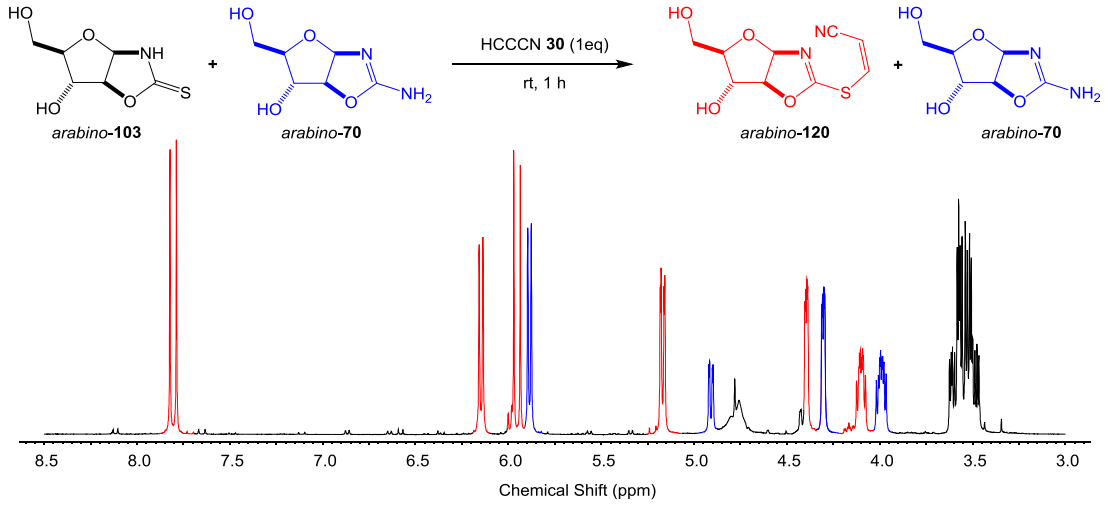


Figure 5.1: *^{1}H NMR spectrum (600 MHz, H_2O/D_2O, 3.0 – 8.0 ppm) of 1:1 mixture of oxazolidinone thione arabino-**103** and aminooxazoline arabino-**70** incubated with cyanoacetylene **30** showing chemoselective cyanovinylation of arabino-**103** to give S-cyanovinyl oxazolidinone thione arabino-**120** in the presence of arabino-**70**.*

Cyanovinylation of aminooxazoline *arabino*-**70** to furnish 2,2`-anhydrocytidine *arabino*-**73** is, however, a vital intermediate in pyrimidine nucleotide synthesis. Moreover, this is a delicate step during the synthesis of activated pyrimidines, requiring careful pH control (pH < 7). Therefore, any cohesive synthesis of purine and pyrimidine ribonucleotides has to also guarantee that *arabino*-**70** can be cyanovinylated near neutral conditions, whilst avoiding the hydrolysis of *arabino*-**73** (into the undesired *arabino*-cytidine **185**, Scheme 3.12**)**. Therefore, the cyanovinylation of *arabino*-**70** in the presence of oxazolidinone thione *arabino*-**103** was then tested. An aqueous solution of *arabino*-**103** (40 mM), *arabino*-**70** (40 mM) and cyanoacetylene **30** (420 mM) was incubated at pH 7 for 24 h. Interestingly, co-

cyanovinylation of both *arabino*-**103** and *arabino*-**70** to yield *S*-cyanovinyl thione *arabino*-**120** and 2,2`-anhydrocytidine *arabino*-**73** respectively in excellent yield was observed (Figure 5.2). Moreover, importantly, no hydrolysis of *arabino*-**73** or *arabino*-**120** was detected, even in the absence of phosphate, suggesting coexistence of *arabino*-**73** and *arabino*-**120** *en route* to the synthesis of pyrimidines and purines, respectively.

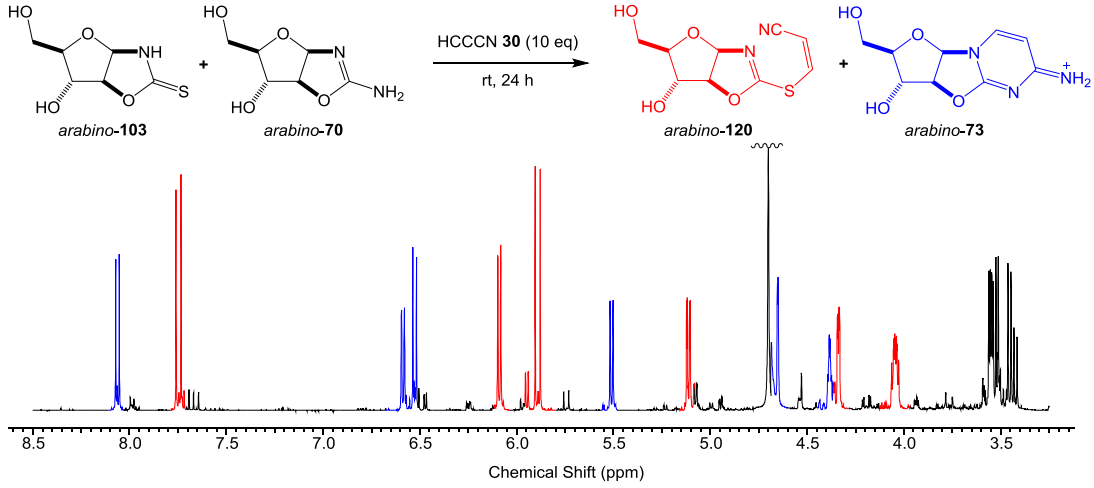


***Figure 5.2**: [1]H NMR spectrum (600 MHz, H_2O/D_2O, 3.25 – 8.5 ppm) to show the products of incubating a 1:1 mixture of oxazolidinone thione arabino-**103** and* aminooxazoline *arabino-**70** with cyanoacetylene **30** (10 eq.) at pH 7 in room temperature leading to cyanovinylation of S-cyanovinyl thione arabino-**120** and the formation of ancitabine arabino-**73**.*

Having explored the prebiotic synthesis of *arabino*-**73** in the presence of purine precursors (thione *arabino*-**103** and *arabino*-**120**), we were intrigued to investigate the elaboration of the purine nucleobases by displacement of thiolate group **128** in *S*-cyanovinyl thione *arabino*-**120** by 2-amino-2-cyanoacetamide **106** in the presence of pyrimidine precursors (aminooxazoline *arabino*-**70** and ancitabine *arabino*-**73**). Therefore, next an aqueous solution of *arabino*-**120** (250mM), *arabino*-**70** (250mM) and **106** (500mM) was incubated at pH 4.5 and room temperature for 8 h. Pleasingly, we observed successful addition of **106** to *arabino*-**120** whilst aminooxazoline *arabino*-**70** remained completely unaffected under these conditions. We then test cyclisation in the presence of aminooxazoline *arabino*-**70**. To an aliquot from the above mixture was added ammonium hydroxide, the solution was adjusted to pH 9 and incubated for 24 h at room temperature. Remarkably, we observed a 28% yield of AICA tricyclic moiety **96** (alongside 29% oxazolidinone thione *arabino*-**103** and 26% oxazolidinone *arabino*-**77**), but 94% of the aminooxazoline *arabino*-**70** remained

(Figure 5.3) demonstrating that our key purine precursor **96** can be synthesised in the presence of the key pyrimidine precursor *arabino*-**70**.

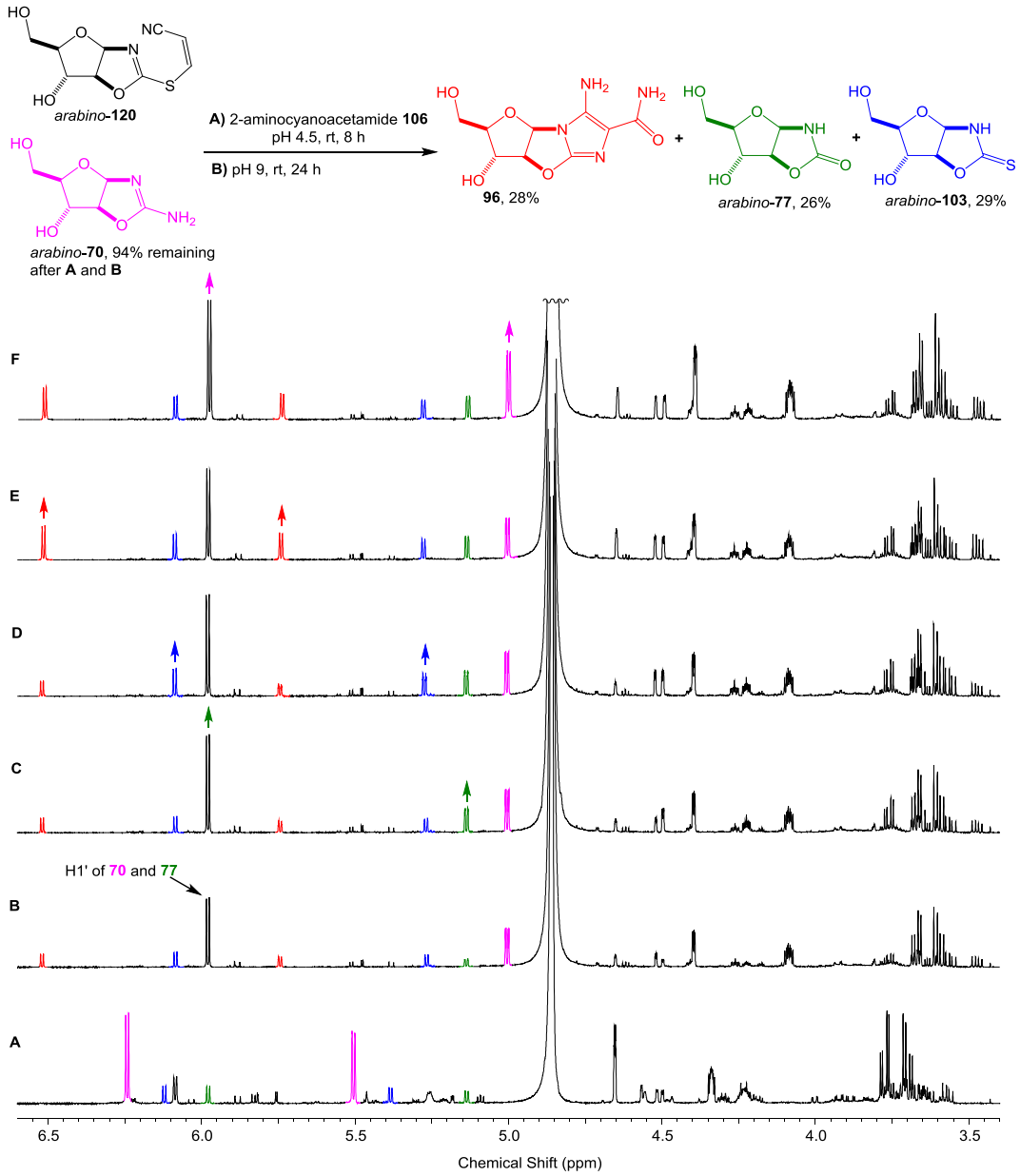


Figure 5.3: *^{1}H NMR spectra (600 MHz, D_2O, 3.4 – 6.6 ppm) showing reaction of S-cyanovinyl oxazolidinone thione arabino-**120** with 2-aminocyanoacetamide **106** in the presence of aminooxazoline arabino-**70**. **A)** arabino-**120** (250mM), arabino-**70** (250mM) and **106** (500mM) at pH 4.5 incubated at rt for 8 h. **B)** An aliquot (50 µL) was added to ammonium hydroxide, the solution was adjusted to pH 9 with NaOH and incubated for 24 h at rt. Calibration to an internal standard gave yields of 28% 2,2′-anhydro-AICA-arabinoside **96**, 29% oxazolinone thione arabino-**103**, 26% oxazolidinone arabino-**77**, additionally 94% arabino-**70** remained. **C)** Spiked with arabino-**77**. **D)** Spiked with arabino-**103**. **E)** Spiked with **95**. **F)** Spiked with arabino-**70**.*

From these experiments we were able to successfully demonstrate that the presence of aminooxazoline *arabino*-**70** in the reaction mixture has no adverse effect in the synthesis of AICA **96**. Next, we sought to investigate the formation of AICA **96** in the presence of 2`.2-anhydrocytidine *arabino*-**73**.

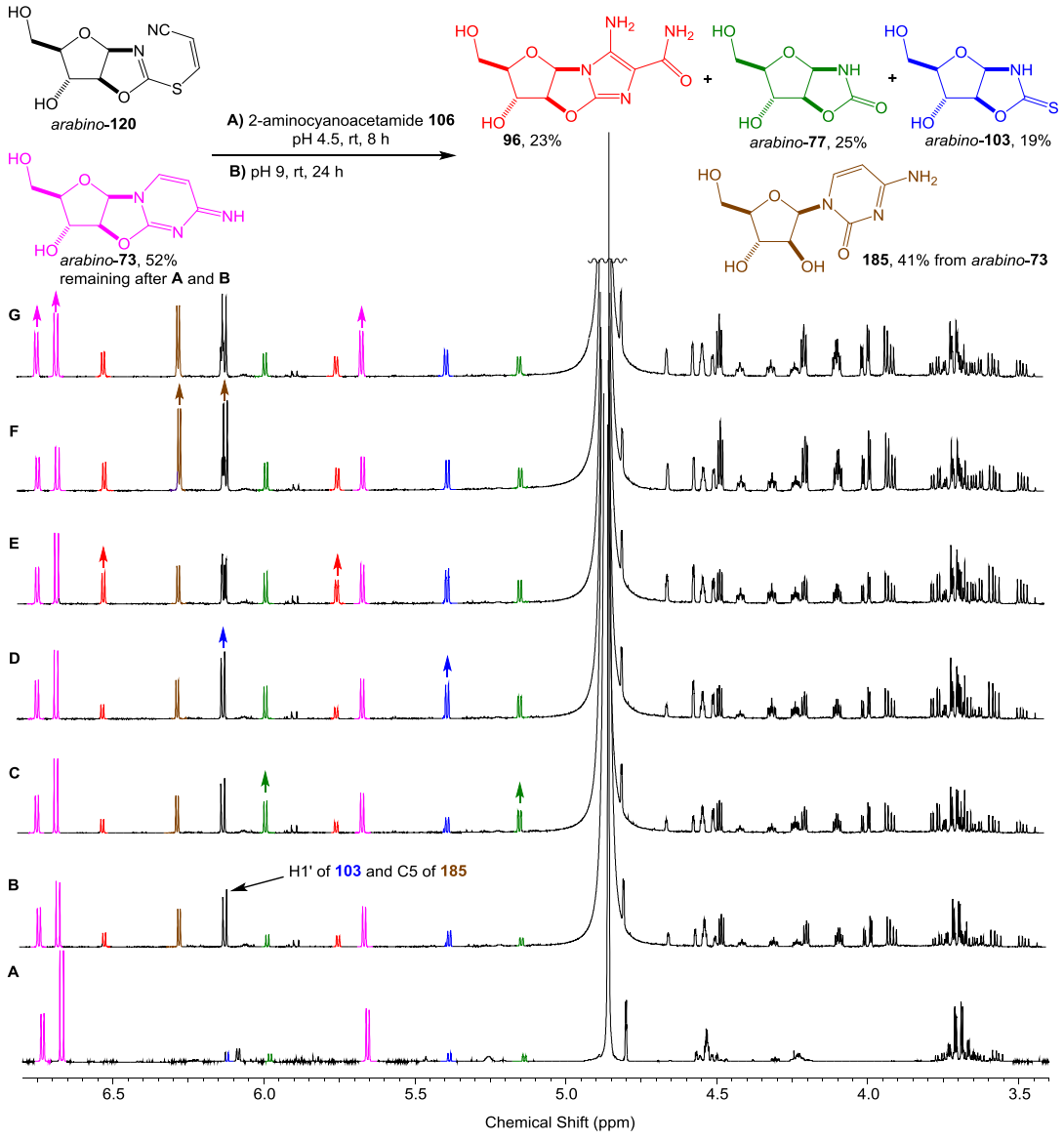


Figure 5.4: *^{1}H NMR spectra (600 MHz, D_2O, 3.4 – 6.8 ppm) showing reaction of S-cyanovinyl oxazolidinone thione arabino-**120** with 2-amino-2-cyanoacetamide **106** in the presence of ancitabine arabino-**73**. **A)** Arabino-**120** (250mM), arabino-**73** (250mM) and **106** (500mM) at pH 4.5 incubated at rt for 8 h. **B)** An aliquot (50 µL) was added to ammonium hydroxide, the solution was adjusted to pH 9 with NaOH and incubated for 24 h at rt. Calibration to an internal standard gave yields of 23% 2,2'-anhydro-5-AICA-arabinoside **106**, 19% oxazolinone thione arabino-**103**, 25% oxazolidinone arabino-**77**, additionally hydrolysis of arabino-**73** gave 41% yield of arabino-cytidine **185** with 52% arabino-**73** remaining. **C)** Spiked with arabino-**77**. **D)** Spiked with arabino-**103**. **E)** Spiked with **106**. **F)** Spiked with arabino-cytidine **185**. **G)** Spiked with arabino-**73**.*

It is important to note that *arabino*-***73*** undergoes facile hydrolysis (pH>6.5, room temperature) to form *arabino*-cytidine **185**.[21, 22] *S*-Cyanovinyl thione *arabino*-**120** (250mM), *arabino*-**73** (250mM) and **106** (500mM) were incubated at pH 4.5 and room temperature for 8 h, after which we observed the successful nucleophilic addition of **106** into *arabino*-**120**, whilst 30% cytidine **185** has formed by to hydrolysis of *arabino*-**73** (70% *arabino*-***73*** remained unchanged). An aliquot of the crude reaction mixture was added to ammonium hydroxide and the solution to adjusted the solution to pH 9. This aliquot was further incubated for 24 h at room temperature. Remarkably, analysis of the crude ^{1}H NMR spectrum and through spiking with authentic samples, we were able to observe a 23% yield of AICA **96** (alongside 19% thione *arabino*-**103** and 25% oxazolidinone *arabino*-**77**) additionally though hydrolysis of *arabino*-**73** gave 41% yield of *arabino*-cytidine **185**, 52% of the original *arabino*-**73** was still remaining in the mixture (Figure 5.4).

Finally, a one pot phosphorylation of purine and pyrimidine anhydronucleosides would be required to complete the divergent synthesis of activated purine and pyrimidine nucleotides. In the previous section (Section 4.4) we have already established optimised phosphorylation conditions. Therefore, to investigate simultaneous phosphorylation of 8,2`-cyclopurines and 2,2`-anhydropyrimidines, adenine **89** and 2,2`-anhydrocytidine *arabino*-**73** were phosphorylated together (Method **F**). ^{1}H and ^{31}P NMR spectroscopic analysis of the crude reaction mixture revealed phosphorus incorporation of 34% and 44% yields for **89** and *arabino*-**73** respectively (Figure 5.5).

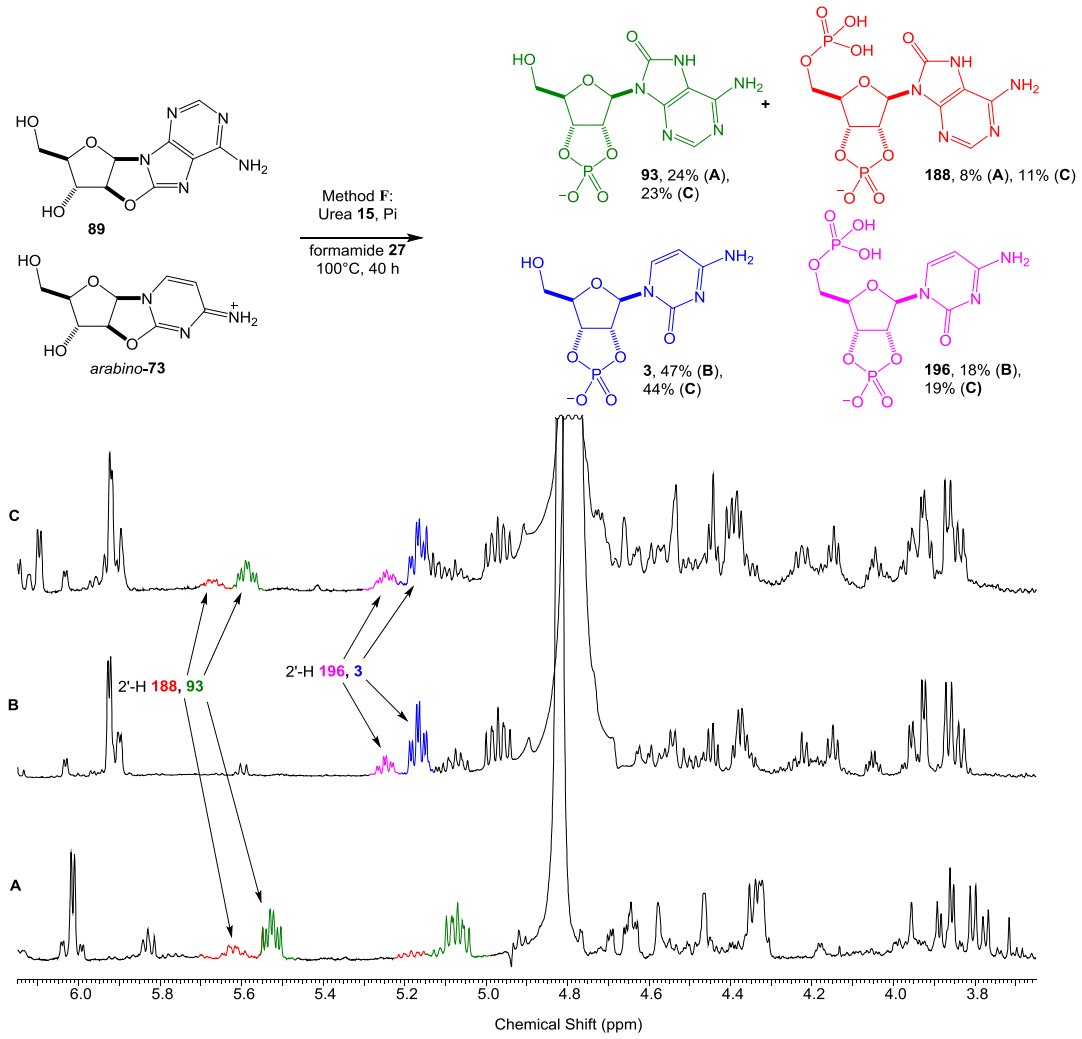


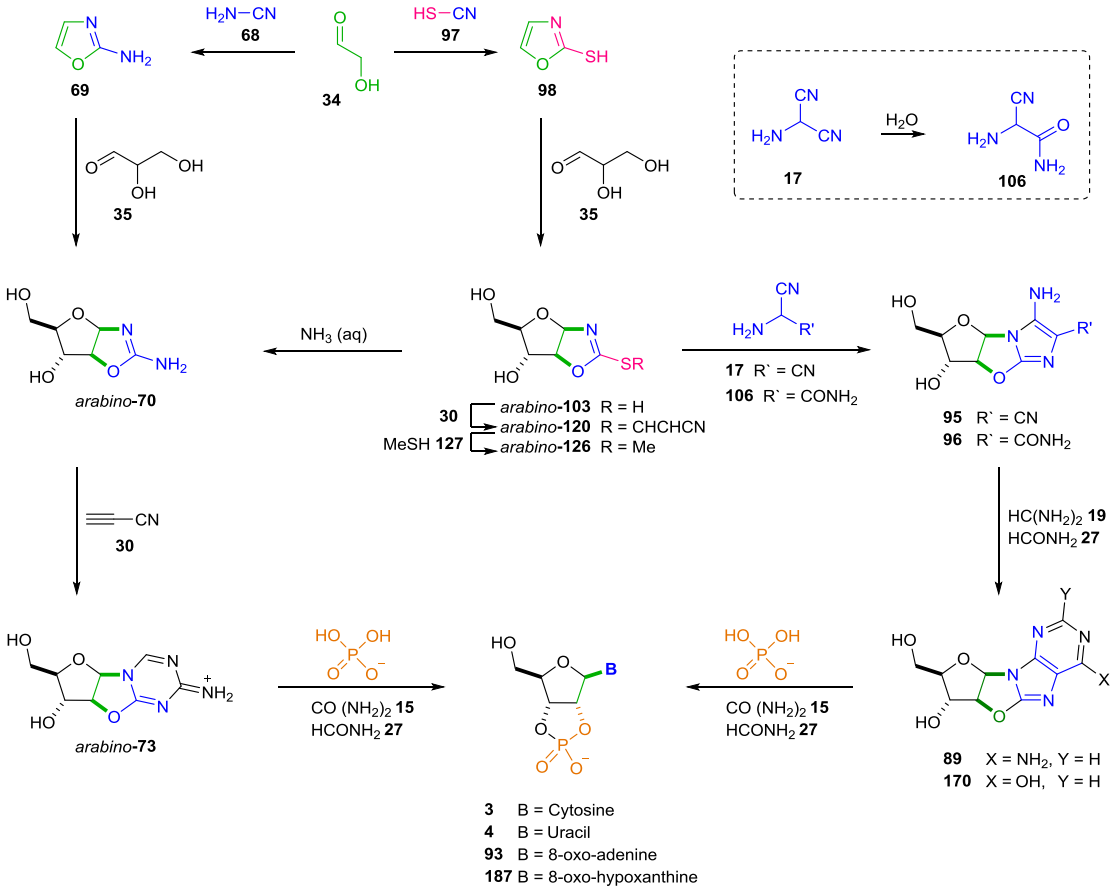
Figure 5.5: *^{1}H NMR spectrum (400 MHz, D_2O, 3.65 – 6.15 ppm) showing co-phosphorylation of ancitabine arabino-**73** and 2',8-O-anhydro-cycloadenosine **89**. **A)** phosphorylation of **89** by method **E** . **B)** Phosphorylation of arabino-**73** by method **E**. **C)** Co-phosphorylation of **89** and arabino-**70** (1:1) by method **F**.*

6. Conclusion

Throughout this thesis we have pursued the goal of developing a divergent synthesis of purine and pyrimidine nucleotides. We have been able to show that the reaction of glycolaldehyde **34** with thiocyanic acid **97** furnishes 2-thiooxazole **98** in excellent yield and its addition to glyceraldehyde **35** establishes the furanosyl ring of ribonucleotide in excellent yield and with excellent selectivity for the *ribo*/*arabino* stereochemistries in the corresponding thione (**103**) products. The *arabino* thione *arabino*-**103** was then be exploited in the first divergent prebiotic synthesis of ribonucleotides as a precursor of both pyrimidine nucleotides (**3** and **4**) and 8-oxo-purine nucleotides (**93** and **187**). After oxazole **98** undergoes reaction with glyceraldehyde **34** to yield thione *arabino*-**103** with complete pentose selectivity, thione *arabino*-**103** is activated by cyanoacetylene **30** to furnish *S*-cyanovinyl oxazolidinone thione *arabino*-**120**. A second point of divergence (following the derivatisation of glycolaldehyde **34** with cyanamide **68** or thiocyanic acid **97**) is observed by the nucleophilic displacement reactions of *arabino*-**103**, which yields pyrimidine precursor *arabino*-**70** upon reaction with aqueous ammonia or purine precursors **95** or **96** upon reaction with hydrogen cyanide oligomers (**17** or **106**, respectively). Cyanovinylation of *arabino*-**70** yields *arabino*-**73** and formylation of **95** and **96** yields **89** and **170**, respectively. The anhydro cyclonucleosides (*arabino*-**73**, **89** and **170)** following subsequent urea-mediated phosphorylation leads to the congruent synthesis of pyrimidine nucleotides (**3** and **4**) and 8-oxo-purine nucleotides (**93** and **187**).

In conclusion, we have demonstrated a synthetic relationship between pyrimidine and purine nucleotides by constructing both heterocycles on the same sugar scaffold. It is also of note that glycine nitrile adduct **146**, which is not at the correct oxidation level to yield an aminoimidazole, was not observed to undergo cyclization; we only observed cyclization in hydrogen cyanide trimers **17** and **106** adducts forming the core imidazole motif of the purines (**95** and **96**). These highly selective cyclizations point towards a renewed investigation of prebiotic synthesis of aminonitrile **17**,[92, 255] which is a pitfall of HCN **14** oligomerisation. The recent reports of the "cyanosulfidic protometabolism" system chemistry by Sutherland and colleagues which avoids the

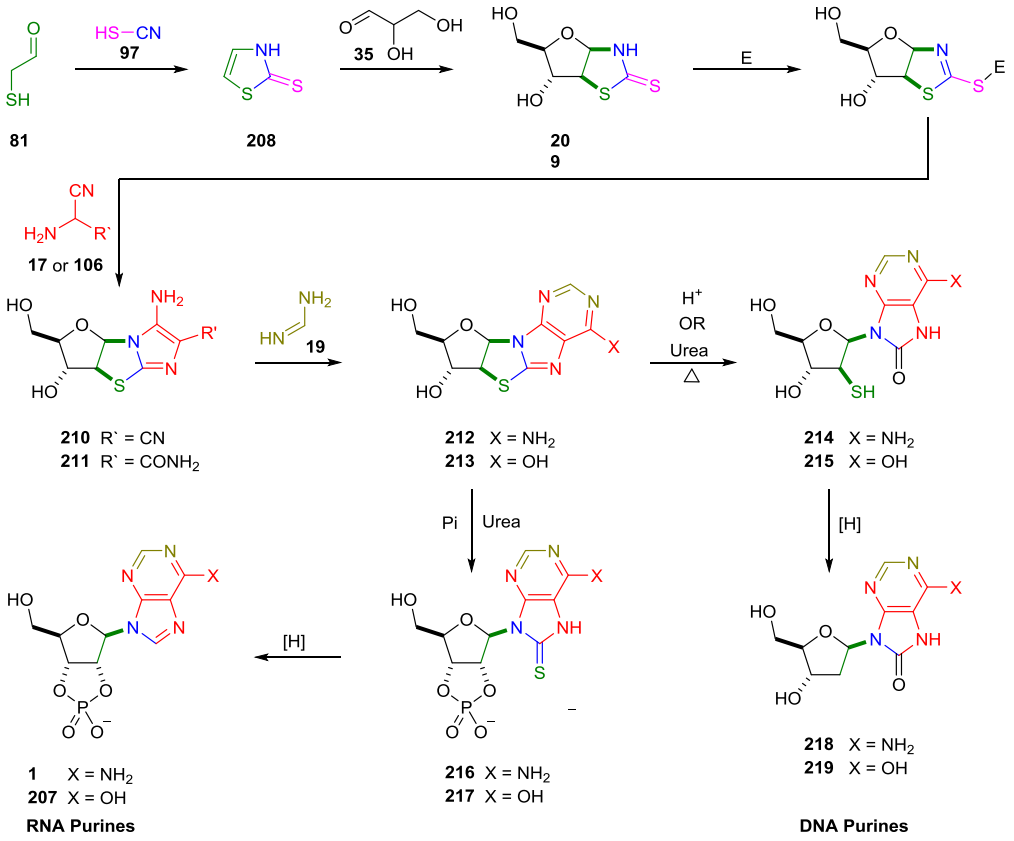
traditional uncontrolled high-pH formaldehyde oligomerisation during prebiotic C_2 and C_3 sugar synthesis,[92] proposes that similar redox coupling network reactions may be used to synthesise HCN oligomers such as **17** and **106.** The chemoselective (prebiotic) reduction of the 8-oxo-purine ribonucleotide-2′,3′-cyclic phosphates (**93**, **94** and **187**) to the canonical ribonucleotides **1**, **2** and **207** is currently being investigated in the Powner laboratory (Scheme 6.1).



Scheme 6.1: *A summary of divergent prebiotic synthesis of pyrimidine and 8-oxo purine ribonucleotides. Plausible reductive cleavage of 8-oxo-2′,3′-cyclic phosphate purines* ***93***, ***94*** *and* ***187*** *furnishing native purine ribonucleotides* ***1***, ***2*** *and* ***207*** *is to be explored for future work (inset).*

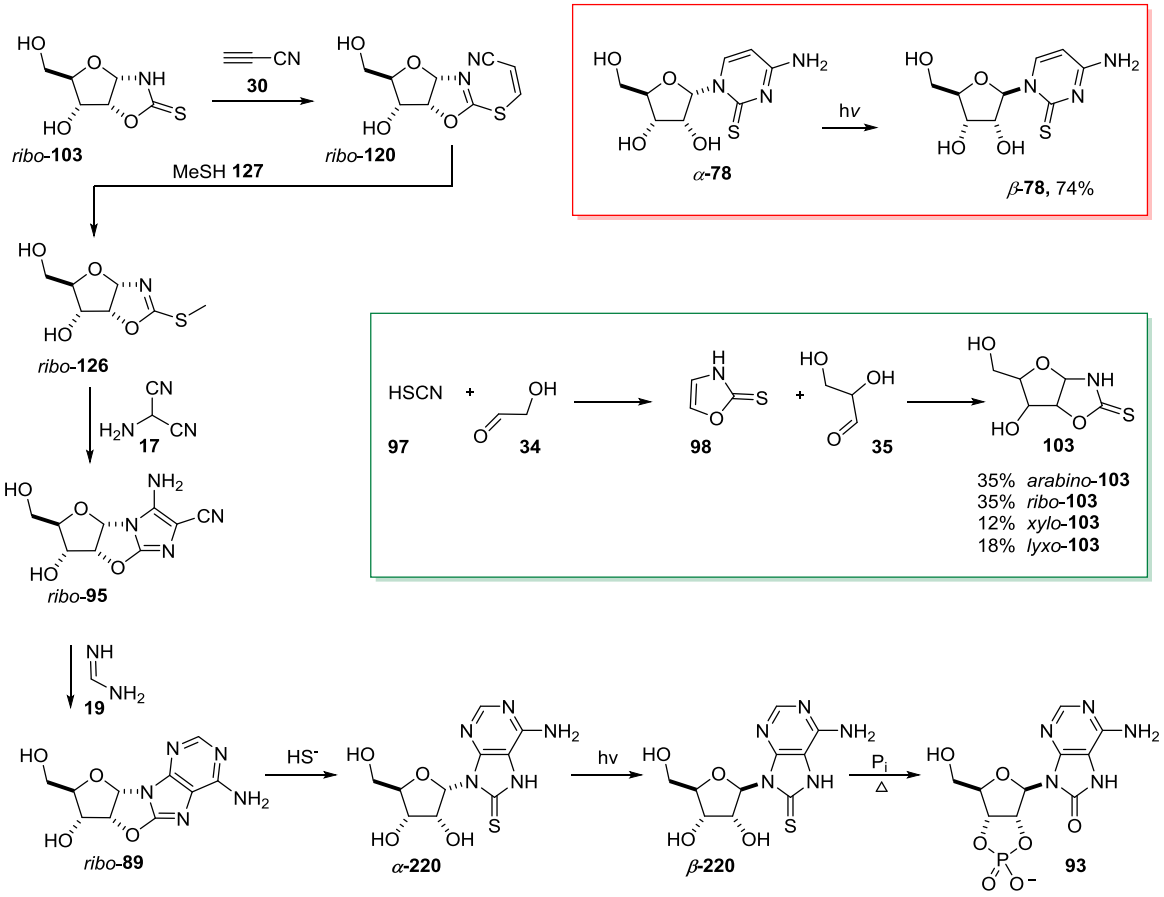
7. Further Future Work

It is hypothesised that 8,2`-anhydro-*S*-purines may provide a possible solution for conversion of 8-oxo purines to natural purine ribonucleotides. Interestingly, the chemistry anticipated to yield 8,2`-anhydro-*S*-purines (**212** and **213**) may follow a very similar strategy to the chemistry that has been already described in this thesis, differing only by the introduction of sulfur at C2`. There is some literature precedence for the synthesis of 8,2`-*S*-anhydro nucleosides using conventional organic synthesis,[208, 256] and these reactions would be a good starting point to investigate the phosphorylation and desulfuration of 8,2`-S-anhydro nucleosides, before embarking on the elucidation of a prebiotic synthesis of **212** and **213**. However, tentatively it could be proposed that the thiozolidines **209** and 2-thiothiazole **208** that have also been reported in the literature,[257, 258] may be chemical precursors of **212** and **213** (Scheme 7.1).



Scheme 7.1: *Plausible prebiotic route for one pot synthesis of RNA and DNA purine ribonucleotides. This route provides possible solution to the reductive cleavage of 8-oxo purines. Most of the chemistry is similar to what has already been established in this thesis, apart from the final reductive, desulfurization steps.*

The reaction of 2-thiooxazole **98** with glyceraldehyde **35** produces 35% oxazolidinone thione *ribo*-**103** in addition to other pentose thiones (*arabino*, *xylo* and *lyxo*). *Ribo*-**103** has not been utilised or investigated in the synthesis reported here. However, based on a new pyrimidine ribonucleotide synthesis reported by Sutherland *et al.* exploiting the photochemical anomerisation of *α*-ribonucleotide to *β*-ribonucleotides,[21] an analogous synthesis for the purine series may also be explored using *ribo*-**103** (Scheme 7.2).



Scheme 7.2: *Proposed synthesis of purines ribonucleotides via C1´ photoanomersation. The efficient (74%) photoanomersation of α-thiocytidine α-**78** to β-thiocytidine β-**78** reported by Sutherland et al. (red inset). Reaction of thiocyanic acid **97** with 2-thiooxazole **98** furnishing pentose oxazolidinone thiones **103** including 35% ribo-**103** (green inset).*

Synthesis of *ribo*-**89** from *ribo*-**103**, followed by thiolysis is expected to produce *α*-8-mercapto adenosine *α*-**220.** Given that *α*-thiocytidine *α*-**78** can be readily be converted to *β*-thiocytidine *β*-**78** (74%) via photoanomerization of C1` carbon, and the structural relationship between *α*-**78** and *α*-**220** (and the position of the

thiocarbonyl group and C1` hydrogen 5-atoms separate in each molecule) it is possible that α-**220**, like α-**78**, can also be converted to β-**220** by anomerisation under photochemical irradiation. These similarities warrant further investigation of purine synthesis under photochemical conditions to complement those presented in this thesis.

Furthermore, polymerisation reactions of pyrimidine and purine 2`,3`-cyclic phosphates would increase our understanding of RNA polymers in the initial steps of chemical to biological transitions.

In a broader prospect and in the hope of a more concerted approach to the origins of life research, aqueous sulfur activation of amine nucleophiles in peptide ligation could be investigated so that a coinciding protein and RNA world can be envisioned. Finally, a more challenging and extremely important aspect of the origins of life research would be to carry out the divergent synthesis of purines and pyrimidines in protocells, so that the "first cell" capable of mimicking minimal life can be synthesised without any biological interference.

8. Experimental

8.1. General Experimental Techniques

Reaction conditions

Moisture sensitive reactions were carried out under atmosphere of nitrogen using oven (80°C) or flame dried glassware. All transfers were performed using plastic or glass syringes using a positive pressure of nitrogen or argon gas. Degassed solutions were prepared by rapidly bubbling nitrogen gas through the required solvent for approximately 10-15 min prior to use. Stirring was achieved by an internal magnetic follower.

Solvents

Petroleum ether or petrol refers to the fraction of petroleum ether with a boiling point between 40°C and 60°C, unless otherwise stated. Anhydrous diethyl ether, dichloromethane, toluene and DMF were purified using a MBRAUN MB SPS-800 solvent purification system. Acetone was distilled over magnesium sulphate. All other solvents were purified by standard procedures or used as supplied from commercial sources. Solvent removal under reduced pressure was carried out using a Büchi Rotavapor equipped with a water or dry ice condenser as appropriate. For heat or concentration sensitive produces water was removed by lyophilisation.

Reagents

Reagents and solvents were obtained and used without further purification, unless specified, from the following commercial sources: Sigma Aldrich, Alfa Aesar, Fluorochem, Acros Organics, Merck, Fisher Scientific, VWR International, Carbosynth, Manchester Organics, BDH, Lancaster, Apollo Scientific, Molekula, TCI and Santa Cruz Biotechnology. Dowex® 50W × 8 ion-exchange resin (200-400 mesh) ion exchange resin was purchased from Acros Organics and was washed with methanol and sodium hydroxide solution before being regenerated with hydrochloric

acid solution. Deionized water was obtained from an Elga Option 3 purification system. The commercially available compound from Acros, 2-amino-2-cyanoacetimide was purified by recrystallisation from hot ethanol.

Chromatography

Flash chromatography was carried out using Fluka silica gel 60 (220-240 mesh) (Brockmann 2-3); samples were applied as a concentrated solution in an appropriate solvent. Thin Layer Chromatography (TLC) was performed on pre-coated aluminum plates with either Merck Kieselgel 60 F_{254} or Merck Aluminum Oxide 60 F_{254}. Visualization was either by ultraviolet light (λ = 254 nm) or by staining with aqueous potassium permanganate ($KMnO_4$) solution followed by heating with a heat gun.

Instrumentation

^{1}H, ^{13}C and ^{31}P NMR spectra were recorded on *Bruker* NMR spectrometers *AVANCE III* 600, *AVANCE III* 400 and *AVANCE* 300 equipped with a *Bruker* 5 mm cryoprobe (600 MHz) and a gradient probe (400 and 300 MHz). All chemical shifts (δ) are reported in parts per million (ppm) relative to residual solvent peaks, and ^{1}H and ^{13}C chemical shifts relative to TMS were calibrated using the residual solvent peak. When a mixed H_2O/D_2O solvent system was used a solvent suppression pulse sequence (noesygppr1d, *Bruker*) was used to obtain ^{1}H NMR spectra. Coupling constants are reported in Hertz (Hz). Spin multiplicities are indicated by symbols: s (singlet); d (doublet); t (triplet); q (quartet); qn (quintet); spt (septet); oct (octet), m (multiplet); obs. (obscured/coincidental signals), or a combination of these. Diastereotopic geminal (AB) spin systems coupled to one or two additional nuclei are reported as ABX and ABXY, respectively. NMR data are reported as follows: chemical shift (number of protons, multiplicity, coupling constants (*J*), nuclear assignment). Spectra were recorded at 298 K. Melting points were determined using an *Electrothermal* standard digital apparatus for all solids and are quoted to the nearest ºC and are uncorrected. Infrared spectra (IR) were recorded on a *Shimadzu IR Tracer 100* FT-IR spectrometer. Absorption maxima are reported in wavenumber (cm^{-1}). Ultraviolet spectra (UV) were recorded on a *Shimzadu UV-2700* spectrometer at 298K using a 1 mL quartz cuvette. Absorption maxima are reported in nm. Mass spectra and accurate mass measurements were recorded on a *VG70-SE*, *Waters LCT Premier XE* or *Thermo Finnigan MAT 900XP* instrument at the Department of Chemistry, University College

London. Flash Column Chromatography (FCC) was carried out on a *Biotage Isolera One* purification platform using either *Biotage SNAP* or *Kinesis TELOS* cartridges. Solution pH values were measured using a *Mettler Toledo Seven Compact* or a *Corning* 430 pH meter equipped with either a *Mettler Toledo InLab* semi-micro pH probe or a *Fisherbrand* FB68801 semi-micro pH probe. pD values were corrected according to *Covington et al.*[1] D_2-formamide was prepared by repeatedly (5×) lyophilising formamide (50 mL) with D_2O (100 mL) until the amide proton signals were no longer visible by ^1H NMR. Unless otherwise stated reactions conducted in formamide were repeatedly (3x) lyophilised with D_2O (10x volume of formamide) before analysis by ^1H NMR.

HPLC was carried out on an *Agilent 1260 Infinity LC* system using an *Agilent Polaris* 5 *C18-A* 150 × 10.0 mm column.

HPLC method used to purify nucleotide cyclic phosphates:

Time (min)	A	B	Flowrate (mL/min)
0.0	100	0	0
0.5	100	0	5
4.0	100	0	5
13.0	80	20	5
13.5	10	90	5
16.5	10	90	5
17.0	100	0	5
20.0	100	0	5
20.5	100	0	0

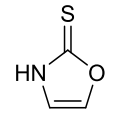
A = 100 mM triethylammonium formate at pH 5.0

B = Acetonitrile

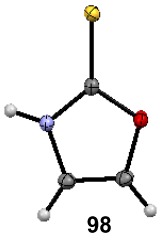
8.2. Prebiotic synthesis of 2-thiooxazole 37

8.2.1. Method A

2-Thiooxazole **98**



C_3H_3NOS
101.12



98

Following a literature procedure,[163] glycolaldehyde **34** (1.00 g, 16.7 mmol) and potassium thiocyanate **97** (3.24 g, 33.3 mmol) were dissolved in minimum volume of water (3 mL). The mixture was cooled to -5°C and HCl (37%, 2.10 mL) was added drop wise. The reaction mixture was initially left for 2 h at rt and then left 24 h at 80°C, completion of reaction was monitored by ^{1}H NMR spectroscopy. The mixture was allowed to cool and then the organics were extracted with ethyl acetate (3 × 50 mL). The organic layers were combined, washed with brine solution (3 × 30 mL), and then dried over $MgSO_4$. The $MgSO_4$ was removed by filtration and solvent was removed under reduced pressure. The resulting solids were crystallized from CH_2Cl_2 to yield 2-thiooxazole **98** (1.43 g, 14.2 mmol, 85%) as a yellow crystalline solid: M.p 140-144ºC (Lit.[168] 147ºC). IR (Solid, cm^{-1}) 3117 (NH), 1587 (C=C), 1478 (C=S). ^{1}H NMR (600 MHz, D_2O) 7.33 (1H, d, *J* = 4.5, H-(CN)), 7.05 (1H, d, *J* = 4.5, H-(CO)). ^{13}C NMR (151 MHz, D_2O) 188.1 (C=S), 130.0 (C-O), 117.2 (C-NH). HRMS (m/z) calculated for C_3H_3NOS $[M]^+$, 100.9930; found, 100.9930.

Crystal structure obtained; see appendix for crystallographic data.

8.2.2. Method B

Montmorillonite KSF clay (1 g) was added to water (0.75 mL) and the pH of the mixture was adjusted to pH 4.0 with aqueous HCl and NaOH as required. To the resultant slurry was added glycolaldehyde **34** (125 mg, 2.08 mmol) and potassium thiocyanate **97** (404 mg, 4.16 mmol) dissolved in water (0.25 mL) and pre-adjusted

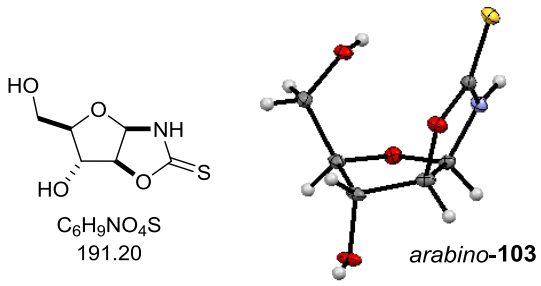
with to pH 4 with aqueous HCl. The reaction was stirred for 3 d at 70°C and then the clay was removed by filtration. The filtrate and washings were combined and lyophilised. The lyophilisate was dissolved in D_2O (2 mL) and an accurately weighed sample of potassium hydrogen phthalate (5.0 mg) was added. The yield of **98** was calculated with reference to the phthalate ^{1}H NMR standard by ^{1}H NMR (600 MHz) signal integration. The calculated yield of 2-thiooxazole **98** was observed to be 35% and not starting material **34** observed due to loss on the clay.

8.3. Synthesis of pentose oxazolidinone thiones

8.3.1. Conventional synthetic protocol: From aldopentose sugars

General procedure for synthesis of pentose oxazolidinone thiones (**103**) from free aldopentose sugars (**102**), adapted from Girniene *et al.*[159] Pentose (1.00 g, 6.67 mmol) and potassium thiocyanate **97** (1.30 g, 13.4 mmol) were dissolved in water (60 mL). HCl (37%, 1.15 mL) was added drop wise at 0°C. The pink solution was kept at rt for 2 h then incubated to 60°C. NMR spectra were periodically acquired until the reaction was complete. Silica gel (10 mL, dry) was added and the mixture was then evaporated to dryness to give a free flowing powder. The powder was purified by silica gel flash coloumn chromoatography, eluting with EtOAc/MeOH (0-10% MeOH) resulting product was recrystallized from $CHCl_3$/EtOH to yield the desired oxazolidinone thione. Analytical samples were recrystallized from hot water.

Arabinofuranosyl oxazolidinone thione *arabino*-**103**

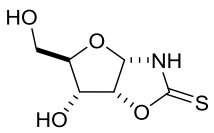


(8.28 g, 43.3 mmol, 65%) from arabinose *arabino*-**102** (10.0 g, 66.6 mmol) after 6 d at 60°C. M.P. 137-139°C (Lit.[259] 132-133°C[259]). IR (solid, cm^{-1}) 3394 (NH), 3299

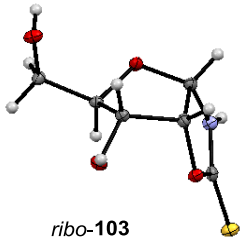
(OH), 2991, 2943, 2931, 2871 (CH), 1482 (C=S).^{1}H NMR (600 MHz, D_2O) 6.07 (1H, d, *J* = 5.8 Hz, H-(C1`)), 5.33 (1H, d, *J* = 5.8 Hz, H-(C2`)), 4.51 (1H, d, *J* = 1.4 Hz, H-(C3`)), 4.25 (1H, ddd, *J* = 7.2, 5.1, 1.4 Hz, H-(C4`)), 3.64 (1H, ABX, *J* = 12.3, 5.1 Hz, H-(C5`)), 3.51 (1H, ABX, *J* = 12.3, 7.2 Hz, H-(C5``)).^{13}C NMR (151 MHz, D_2O) 190.0 (C2), 92.5 (C1`), 90.5 (C4`), 87.8 (C2`), 75.2 (C3`), 61.6 (C5`). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -4.08. HRMS (*m/z*) calculated for $C_6H_{10}NO_4S$ $[M+H^+]^+$, 192.0331; found, 192.0332.

Crystal structure obtained; see appendix for crystallographic data.

Ribofuranosyl oxazolidinone thione *ribo*-**103**



$C_6H_9NO_4S$
191.20

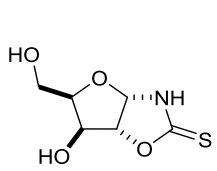


ribo-**103**

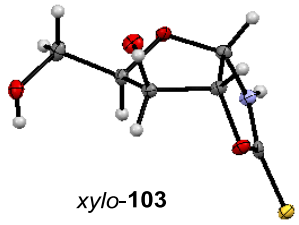
(4.40 g, 23.0 mmol, 69 %) from ribose *ribo*-**102** (5.00 g, 33.4 mmol) after 4 d at 60°C. M.P. 172-175°C decomp. (Lit.[259] 167-168°C). IR (Solid, cm^{-1}) 3438 (NH), 3329 (OH), 2993, 2920, 2902 (CH), 1521 (C=S). ^{1}H NMR (600 MHz, D_2O) 5.80 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.13 (1H, t, *J* = 5.4 Hz, H-(C2`)), 4.18 (1H, dd, *J* = 9.5, 5.4 Hz, H-(C3′)), 3.96 (1H, ABX, *J* = 12.7, 2.3 Hz, H-(C5′)), 3.88 (1H, ddd, *J* = 9.5, 4.9, 2.3 Hz, H-(C4′)), 3.75 (1H, ABX, *J* = 12.7, 4.9 Hz H-(C5′)). ^{13}C NMR (151 MHz, D_2O) 161.1 (C2), 86.0 (C1′), 80.4 (C4′), 78.6 (C2′), 70.4 (C3′), 60.0 (C5′). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) +8.73. HRMS (m/z) calculated for $C_6H_{10}NO_4S$ $[M+H^+]^+$, 192.0331; found, 192.0340.

Crystal structure obtained; see appendix for crystallographic data.

Xylofuranosyl oxazolidinone thione *xylo*-**103**



$C_6H_9NO_4S$
191.20

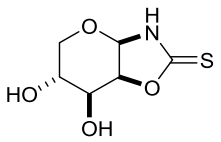


xylo-**103**

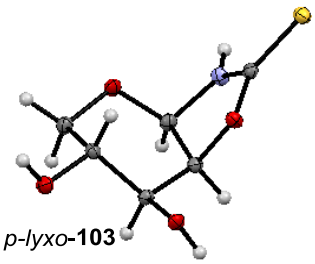
(1.46 g, 7.66 mmol, 46%) from xylose *xylo*-**102** (2.50 g, 16.6 mmol) after 14 d at 60°C. M.P. 133-135°C (Lit.[259] 129-130°C). IR (Solid, cm^{-1}) 3251 (OH), 2930 (CH), 1488(C=S). ^{1}H NMR (600 MHz, D_2O) 6.07 (1H, d, *J* = 5.5 Hz, H-(C1′)), 5.29 (1H, d, *J* = 5.5 Hz, H-(C2′)), 4.52 (1H, d, *J* = 2.8 Hz, H-(C3′)), 4.04 (1H, ddd, *J* = 7.4, 4.3, 2.8 Hz, H-(C4′)), 3.96 (1H, ABX, *J* = 12.0, 4.3 Hz, H-(C5′)), 3.84 (1H, ABX, *J* = 12.0, 7.4 Hz, H-(C5′)). ^{13}C NMR (151 MHz, D_2O) 190.3 (C2), 91.1 (H2′), 89.5 (H1′), 80.8 (H4′), 73.5 (H3′), 59.5 (H5′). HRMS (*m/z*) calculated for $C_6H_{10}NO_4S$ $[M+H^+]^+$, 192.0331; found, 192.0332.

Crystal structure obtained; see appendix for crystallographic data.

Lyxofuranosyl oxazolidinone thione *p-lyxo*-**103**



$C_6H_9NO_4S$
191.20



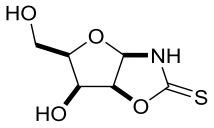
p-lyxo-**103**

(0.25 g, 1.31 mmol, 20%) from lyxose *lyxo*-**102** (1.00 g, 6.66 mmol) after 7 d at 60°C. M.P. 185-187°C (Lit.[259] 178°C). IR (Solid, cm^{-1}) 3342 (NH), 3237, 3130 (OH), 2975 (CH), 1496(C=S). ^{1}H NMR (600 MHz, D_2O) 5.74 (1H, d, *J* = 6.1 Hz, H-(C1′)), 5.16 (1H, dd, *J* = 6.1, 3.2 Hz, H-(C2′)), 4.02 (1H, dd, *J* = 7.5, 3.2 Hz, H-(C3′)), 3.95-3.89 (2H, m, H-(C4′) and H-(C5′)), 3.69 (1H, m, H-(C5′)). ^{13}C NMR (151 MHz, D_2O) 191.4 (C2), 84.0 (C2′), 83.8 (C1′), 70.3 (C3′), 68.9 (C4′), 66.2 (C5′). HMBC was

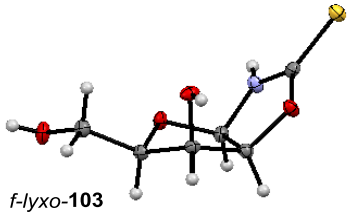
observed between H-(C1′) and C5′ indicating pyranose structure. HRMS (*m/z*) calculated for $C_6H_{10}NO_4S$ $[M+H^+]^+$, 192.0331; found, 192.0328.

Crystal structure obtained; see appendix for crystallographic data.

Lyxopyranosyl oxazolidinone thione *f-lyxo*-**103**[260]



$C_6H_9NO_4S$
191.20



f-lyxo-**103**

^{1}H NMR (600 MHz, D_2O) 5.87 (1H, d, *J* = 5.8 Hz, H-(C1′)), 5.40 (1H, t, *J* = 5.8 Hz, H-(C2′)), 4.70 (1H, dd, *J* = 6.4, 5.8 Hz, H-(C3′)), 4.31 (1H, ddd, *J* = 9.0, 6.4, 3.9 Hz, H-(C4′)), 3.81 (1H, ABX, *J* = 12.4, 3.9 Hz, H-(C5′)), 3.56 (1H, ABX, *J* = 12.4, 9.0 Hz, H-(C5′)).

Crystal structure obtained; see appendix for crystallographic data.

Method large scale synthesis of oxazolidinone thiones purified without chromatography:[261]

Arabinose *arabino*-**102** (120 g, 0.80 mol) and potassium thiocyanate **97** (156 g, 1.60 mol) were dissolved in water (120 mL) and then HCl (37%, 148 mL) was added drop wise at 0°C. The pink solution was kept at rt for 2 h then heated to 60°C for 6 d. After 6 d, the crude mixture was cooled, filtered to the remove yellow amorphous solids. The filtrate was then concentrated and air-dried over night to give glassy solids which were triturated with ethanol (3 × 500 mL). The ethanolic fractions were combined and concentrated to give yellow solids. The solids were dissolved in boiling ethanol (500 mL) and mixed with activated charcoal (5.0 g). The organic mixture was filtered to remove all solids and then concentrated to dryness. The solids obtained were dissolved in boiling H_2O (70 mL) and then left at 5°C overnight to give pale yellow crystals, which were further recrystallised from hot water (70 mL) to give of arabinofuranosyl oxazolidinone thione *arabino*-**103** (50.0 g, 0.26 mol, 33%) as a crystalline product.

General procedure for time course reactions of aldopentose sugars with potassium thiocyanate.

Aldopentose sugar **102** (100 mg, 0.66 mmol) and potassium thiocyanate **97** (130 mg, 1.34 mmol) were dissolved in H_2O (0.50 mL) and D_2O (0.1 mL) and then HCl (37%, 0.11 mL) was added drop wise at 0°C. The resultant solution was then transferred to an NMR tube and data was collected at the required intervals at rt to generate an NMR time course for the synthesis of different oxazolidinone thiones (*arabino*-**103**, *ribo*-**103** or *xylo*-**103**). The graphs were analysed to investigate any detectable intermediates during the reaction of the aldopentose sugars **102** with potassium thiocyanate **97**.

Ribofuranosyl oxazolidinone thione *ribo*-**103**

Following the above method, ribose *ribo*-**102** and potassium thiocyanate **97** to furnish *ribo*-**103**, the pH of the reaction was monitored and ^{1}H NMR data was collected at different time intervals.

Arabinofuranosyl oxazolidinone thione *arabino*-**103**

Following the above method, arabinose *arabino*-**102** and potassium thiocyanate **97** to furnish *arabino*-**103**, the pH of the reaction was monitored and ^{1}H NMR data was collected at different time intervals.

Xylofuranosyl oxazolidinone thione *xylo*-**103**

Following the above method, xylose *xylo*-**102** and potassium thiocyanate **97** to furnish *xylo*-**103**, the pH of the reaction was monitored and ^{1}H NMR data was collected at different time intervals.

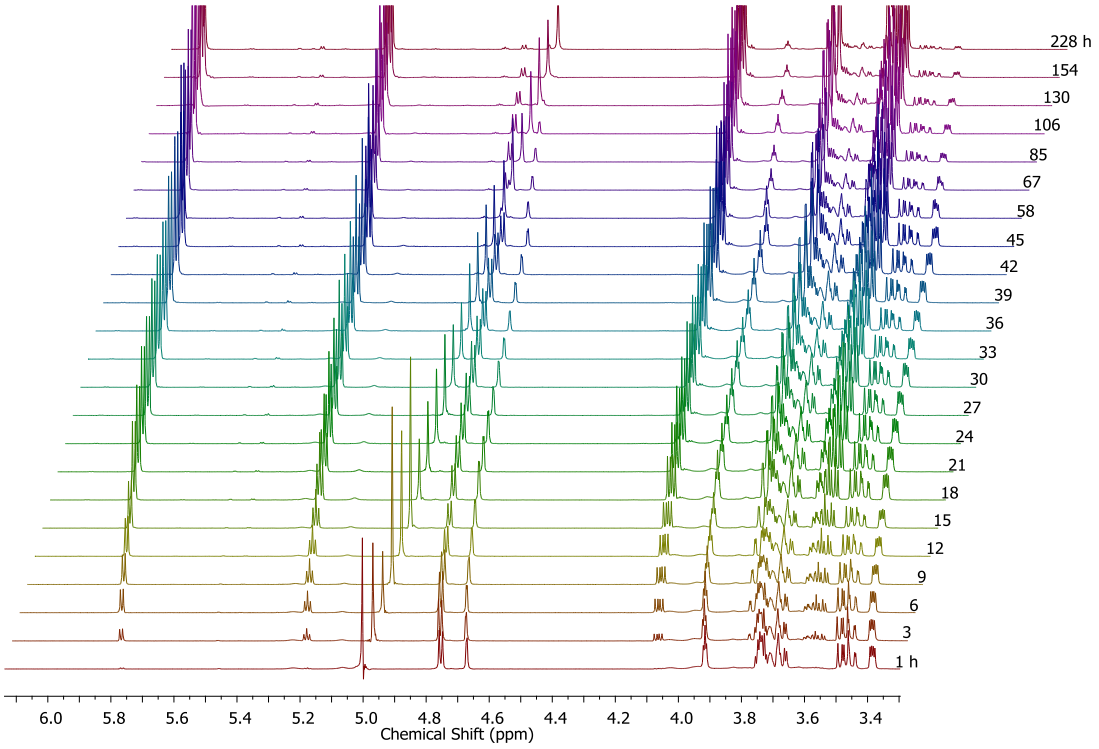
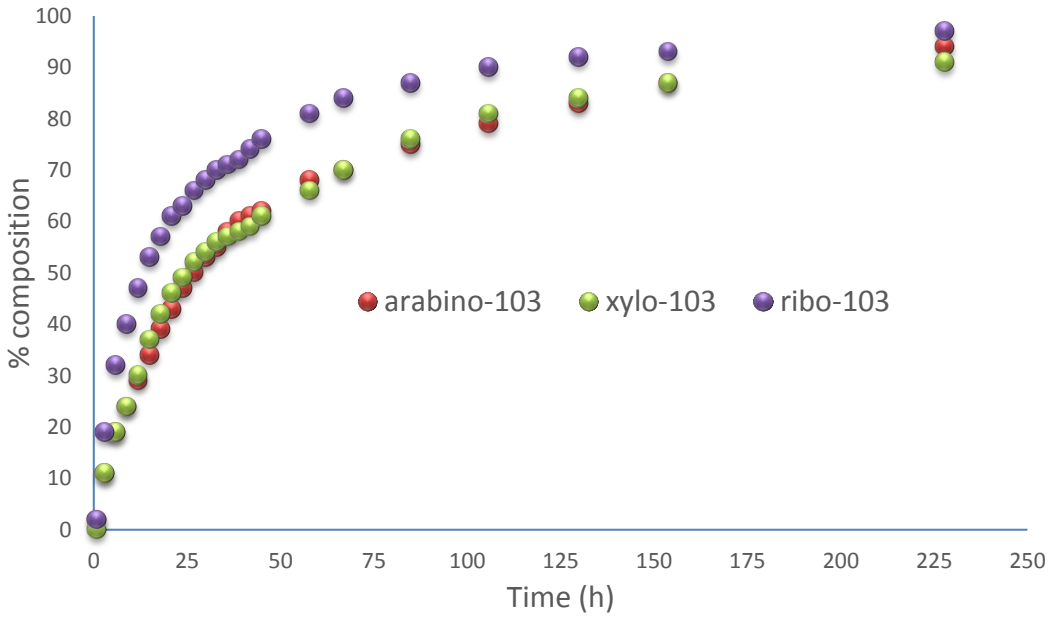


Figure 8.1: *A stack of ^{1}H NMR spectra (600 MHz, D_2O, 3.3 - 6.2 ppm) for incubation of ribose ribo-**102** (1.1 mM) and potassium thiocyanate **97** (2.2 mM) at rt, pH -1.5 for 228h. The peaks for ribose ribo-**102** are decreasing and peaks for oxazolidinone thione ribo-**103** is increasing over time. No intermediate species were detected in the ^{1}H NMR with single solvent suppression.*



Time (h)	*ribo*-**103**	*arabino*-**103**	*xylo*-**103**
1	2	2	0
3	19	11	11
6	32	19	19
9	40	24	24
12	47	29	30
15	53	34	37
18	57	39	42
21	61	43	46
24	63	47	49
27	66	50	52
30	68	53	54
33	70	55	56
36	71	58	57
39	72	60	58
42	74	61	59
45	76	62	61
58	81	68	66
67	84	70	70
85	87	75	76
106	90	79	81
130	92	83	84
154	93	87	87
228	97	94	91

Figure 8.2: *Comparative reaction of ribo-**102**, arabino-**102** and xylo-**102** with potassium thiocyanate **97**, over a period of 228h at rt to yield ribo-**103,** arabino-**103** and xylo-**103** respectively. Each point in the graph is ^{1}H NMR data obtained by integration of H-(C1`) of thione **103** against sugar **102** peaks at 3.39 and 4.75 ppm. Each pentose **102** (110 mM) was incubated with **97** (220 mM) at pH -1.5 and rt for 228h and then data collected for all three oxazolidinone thiones, ribo-**103**, arabino-**103** and xylo-**103** are plotted for comparison.*

8.3.2. Prebiotic protocol: Bypassing free aldopentose sugars

2-Thiooxazole **98** (75.6 mg, 0.75 mmol) and 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS; 20 mg; internal standard) were dissolved in D_2O (1.5 mL) to give a 500mM solution of **98**. The solution was adjusted to pH 7 with NaOH (4M) and 250 or 500 µL was added to glyceraldehyde **35** (500mM or 1M). Solutions were made up to 500 µL with D_2O where necessary. The solutions were stirred at 60°C for 24 h and the

solution was readjusted to pH 7 every 6 h (if necessary). Aliquots (50 µL) were taken after 24 h, diluted with D_2O (450 µL) and analysed by NMR spectroscopy. The presence of oxazolinone thiones **103** was confirmed by spiking with authentic samples, prepared using conventional organic synthesis as reported above. Yields were calculated by measuring the ratio of product **103**:DSS at 24 h.

				% distribution of oxazolidinone thione products **103**				
98 (M)	**35** (M)	**103** (%)	**98** (%)	*arabino* **103**	*ribo* **103**	*xylo* **103**	*p-lyxo* **103**	*f-lyxo* **103**
0.25	0.5	51	36	35	35	12	11	7
0.50	0.5	35	44	37	36	10	12	5
0.50	1.0	59	25	35	36	12	11	6
0.25	1.0	74	8	36	35	10	12	7

Figure 8.3: *Yields of pentose oxazolidinone thiones* ***103*** *and their relative diastereomeric ratio across a variety of concentrations of 2-thiooxazole* ***98*** *and glyceraldehyde* ***35****. All reactions were heated for 24 h at 60°C in pH 9.*

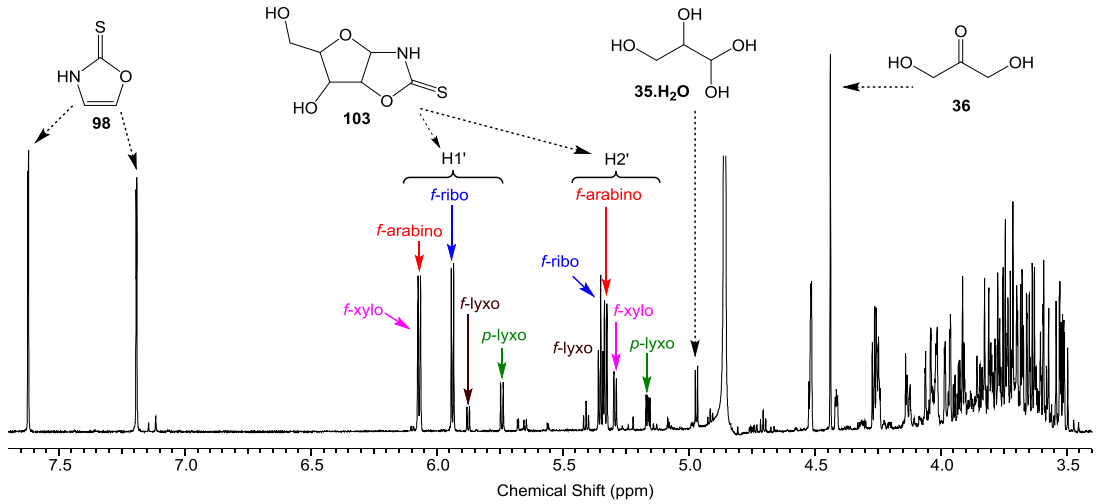


Figure 8.4: *^{1}H NMR spectrum (600 MHz, D_2O, 3.4 – 7.7 ppm) showing reaction of 2-thiooxazole* ***98*** *with glyceraldehyde* ***35****. 2-Thiooxazole* ***98*** *(0.25M) and glyceraldehyde* ***35*** *(1M) at pH 7 were incubated at 60 °C for 24 h.*

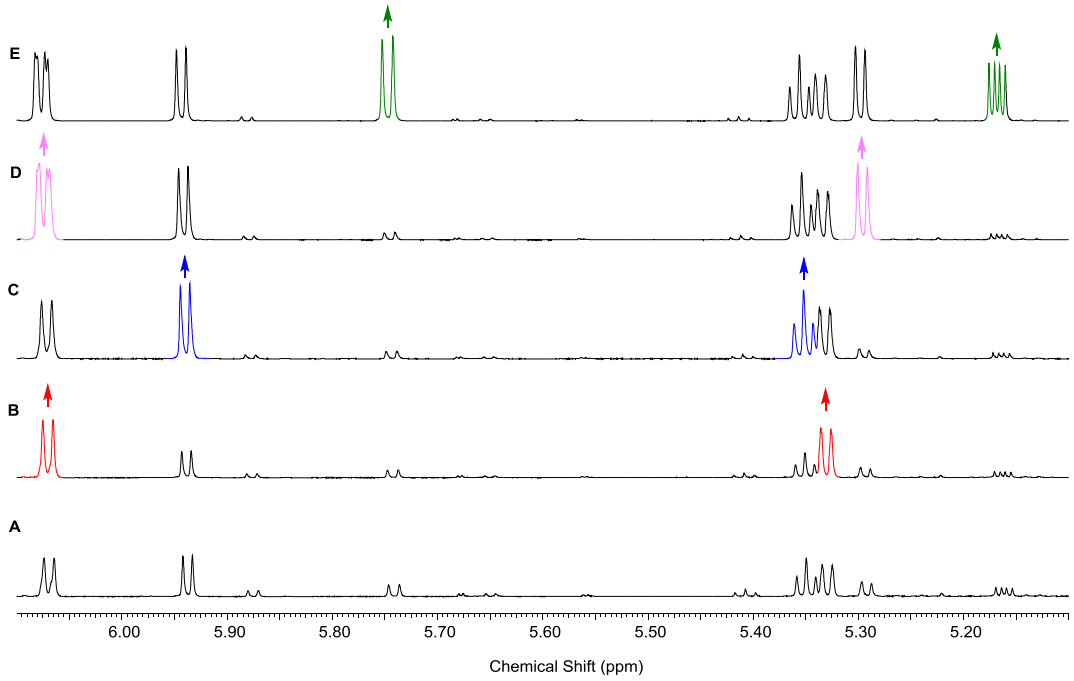
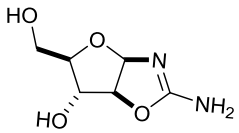


Figure 8.5: *^{1}H NMR spectra (600 MHz, D_2O, 5.1 – 6.1 ppm) showing reaction of 2-thiooxazole **98** with glyceraldehyde **35**. A) **25** (0.25M) and **98** (1M) at pH 7, incubated at 60°C for 24 h. B) Spiked with arabinose oxazolidinone thione arabino-**103**. C) Spiked with ribose oxazolidinone thione ribo-**103**. D) Spiked with xylose oxazolidinone thione xylo-**44**. E) Spiked with lyxose oxazolidinone thione p-lyxo-**44**.*

8.4. Synthesis of pentose aminooxazolidinones

Pentose aminooxazolines were synthesised using a protocol modified from Sanchez and Orgel.[124]

Arabinofuranosyl aminooxazoline *arabino*-**70**

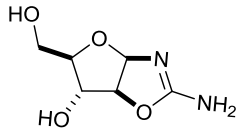


$C_6H_{10}N_2O_4$
174.16

To a syrup of arabinose *arabino*-**102** (5.00 g, 33.3 mmol) in aqueous ammonia solution (3.5%, 50.0 mL) was added cyanamide **68** (2.28 g, 54.2 mmol). The resultant solution was heated at 60°C for 1 h. After this time the reaction was cooled to rt and MeOH (100 mL) was added to the solution to promote crystallisation. After 16 h at 4°C the crystals were collected by filtration, washed with ice-cold MeOH (80 mL)

and dried under vaccum to yield of arabinofuranosyl aminooxazoline *arabino*-**70** (3.20 g, 18.4 mmol, 55%) as a fine white powder: M.p 195-197°C (Lit.[124] 175°C). IR (solid, cm^{-1}) 3407 (NH_2), 3140 (OH), 2919 (CH), 1660 (C=N). ^1H NMR (600 MHz, D_2O) δ 5.91 (1H, d, *J* = 5.5 Hz, H-(C1`)), 4.94 (1H, d, *J* = 5.5 Hz, H-(C2`)), 4.33 (1H, d, *J* = 3.6 Hz, H-(C3`)), 4.02 (1H, ddd, *J* = 6.9, 5.4, 3.6 Hz, H-(C4`)), 3.61 (1H, ABX, *J* = 12.2, 5.4 Hz, H-(C5`)), 3.54 (ABX, *J* = 12.2, 6.9 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 165.6 (C2), 99.29 (C1`), 89.3 (C4`), 85.0 (C2`), 75.9 (C3`), 61.8 (C5`). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -8.31. HRMS (*m/z*) calculated for $C_6H_{10}N_2O_4$ $[M+H^+]^+$, 175.0719; found, 175.0719.

Ribofuranosyl aminooxazoline *ribo*-**70**



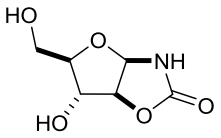
$C_6H_{10}N_2O_4$
174.16

To a syrup of ribose *ribo*-**102** (2.00 g, 13.3 mmol) in aqueous ammonia solution (3.5%, 20.0 mL) was added cyanamide **68** (1.12 g, 26.6 mmol). The resultant solution was heated at 60°C for 1 h. After this time the reaction was cooled to rt and MeOH (40 mL) was added to the solution to promote crystallisation. After 16 h at 4°C the crystals were collected by filtration, washed with ice-cold MeOH (30 mL) and dried under vaccum to yield of ribofuranosyl aminooxazoline *ribo*-**70** (1.63 g 9.36 mmol, 70%) as a fine white powder: M.p 182-184°C (Lit.[124] 195°C). IR (solid, cm^{-1}) 3426 (NH_2), 3319 (OH), 3147 (CH), 1664 (C=N). ^1H NMR (600 MHz, D_2O) 5.80 (1H, d, *J* = 5.3 Hz, H-(C1`)), 4.98 (1H, t, *J* = 5.3 Hz, H-(C2`)), 4.13 (1H, dd, *J* = 9.6, 5.3 Hz, H-(C3`)), 3.93 (1H, ABX, *J* = 12.7, 2.3 Hz, H-(C5`)), 3.73 (1H, ABX, *J* = 12.7, 4.7 Hz, H-(C5`)), 3.62 (1H, ddd, *J* = 9.6, 4.7, 2.3 Hz, H-(C4``)). ^{13}C NMR (151 MHz, D_2O) 166.7 (C1), 97.7 (C1`), 82.6 (C4`), 77.9 (C2`), 71.0 (C3`), 60.2 (C5`). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) +3.94. *m/z* (CI+) 175 (78%, M+). HRMS (*m/z*) calculated for $C_6H_{10}N_2O_4$ $[M+H^+]^+$, 175.0719; found, 175.0721.

8.5. Synthesis of pentose oxazolidinones

Pentose oxazolidinones were synthesised according to a procedure by Kovács *et al.*[262]

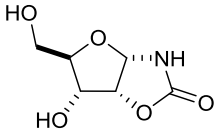
Arabinofuranosyl oxazolidinone *arabino*-**77**



$C_6H_9NO_5$
175.14

To arabinose *arabino*-**102** (3.00 g, 20.0 mmol), in H_2O (20 mL) was added NH_4Cl (1.60 g, 30.0 mmol) and potassium cyanate **26** (2.40 g, 30.0 mmol). The resultant mixture was heated at 60°C for 8 h, allowed to cool to rt and lyophilised to give a yellow oil. The oil was then purified by silica gel flash column chromatography eluting with (EtOAc:MeOH 9:1). After evaporation the residue was recrystallized from EtOH/H_2O to give arabinofuranosyl oxazolidinone *arabino*-**77** (1.60 g, 9.14 mmol, 46%) as a white solid: M.p 164-168°C. IR (solid, cm^{-1}) 3398 (NH), 3233 (OH), 2956, 2935, 2897 (CH), 1747 (C=O). ^{1}H NMR (600 MHz, D_2O) 5.92 (1H, d, *J* = 5.7 Hz, H-(C1`)), 5.08 (1H, d, *J* = 5.7 Hz, H-(C2`)), 4.44 (1H, s, H-(C3`)), 4.16 (1H, dd, *J* = 6.9, 4.8 Hz, H-(C4`)), 3.70 (1H, ABX, *J* = 12.3, 4.8 Hz, H-(C5`)), 3.60 (1H, ABX, *J* = 12.3, 6.9 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 160.1 (C1), 87.6 (C1`), 87.2 (C4`), 86.9 (C2`), 75.5 (C3`), 61.9 (C5`). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -5.91. HRMS (*m/z*) calculated for $C_6H_9NO_5$ $[M+H^+]^+$, 176.0559; found, 176.0551.

Ribofuranosyl oxazolidinone *ribo*-**77**

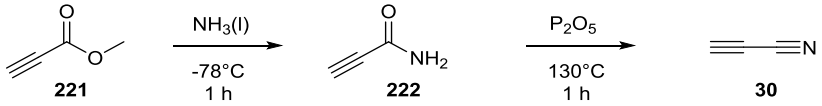


$C_6H_9NO_5$
175.14

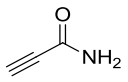
To ribose *ribo*-**102** (3.00 g, 20.0 mmol), in H_2O (20 mL) was added NH_4Cl (1.60 g, 30.0 mmol) and potassium cyanate **26** (2.40 g, 30.0 mmol). The resultant mixture was

heated at 60°C for 8 h, allowed to cool to rt and lyophilised to give a yellow oil. The oil was then purified by silica gel flash column chromatography eluting with (EtOAc:MeOH 9:1). After evaporation the residue was recrystallized from EtOH/H_2O to give ribofuranosyl oxazolidinone *ribo*-**77** (2.20 g, 12.56 mmol, 63%) as a yellow solid: M.p 125-127°C (Lit.[262] 161-164°C). IR (solid, cm^{-1}) 3341 (NH), 3217 (OH), 3005,2934, 2879 (CH), 1715 (C=O). ^{1}H NMR (600 MHz, D_2O) 5.80 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.13 (1H, t, *J* = 5.4 Hz, H-(C2`)), 4.18 (1H, dd, *J* = 9.5, 5.4 Hz, H-(C3`)), 3.96 (1H, ABX, *J* = 12.7, 2.3 Hz, H-(C5`)), 3.88 (1H, ddd, *J* = 9.5, 4.8, 2.3 Hz, H-(C4`)), 3.75 (1H, ABX, *J* =12.7, 4.8 Hz H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 161.1 (C2), 86.0 (C1`), 80.4 (C4`), 78.6 (C2`), 70.4 (C2`), 60.0 (C5`). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) +6.76. HRMS (*m/z*) calculated for $C_6H_9NO_5$ $[M+H^+]^+$, 176.0559; found, 176.0555.

8.6. Synthesis of cyanoacetylene



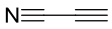
Propiolamide **222**



C_3H_3NO
69.02

Methyl propiolate **221** (9.45g, 10.0 mL, 112 mmol) was added to liquid ammonia (100 mL) at -78°C and stirred. After 1 h excess ammonia was evaporated by warming the solution to rt, followed by co-evaporation with dry chloroform (100 mL), to yield solids. The solids were dissolved in hot anhydrous CH_2Cl_2 (30 mL) and left at -18°C to yield propiolamide **222** (5.74 g, 83.0 mmol, 74%) as long colourless needle-like crystals which were isolated by filtration: M.p 57-59°C (Lit.[262] 58-61°C). ^{1}H NMR (600 MHz, D_2O) 3.53 (1H, s, H-(C3). m/z (CI+) 70 (100%, M+). Spectra matched lietrature data.[263]

Cyanoacetylene **30**

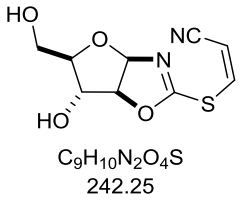


C_3HN
51.05

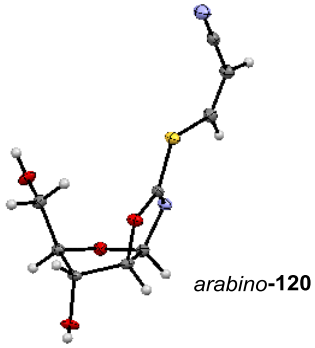
Propiolamide **222** (5.00 g, 72.5 mmol) and oven dried sand (37.5 g) were thoroughly grounded together with a pestle and mortat. P_2O_5 (15.5 g, 109 mmol) was added and the mixture was rapidly transferred to a dry flask under an argon atmosphere. Cyanoacetylene **30** was then distilled from the brown sticky mixture at 130°C and 100 torr over 1 h. During this time a white solid (2.98 g, 58.4 mmol, 80%) was collected at -78°C. The solid was immediately dissolved in water (47 mL) to give a 1M aqueous solution of cyanoacetylene **30**: ^{1}H NMR (600 MHz, D_2O) 3.42 (1H, s, H-(C3)). ^{13}C NMR (151 MHz, D_2O) 105.5 (C1); 77.1 (C3); 56.7 (C2).

8.7. Cyanovinylation of pentose oxazolidinone thiones

(*S*-*Z*-Cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120**



$C_9H_{10}N_2O_4S$
242.25



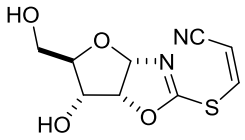
arabino-**120**

Arabinofuranosyl oxazolidinone thione *arabino*-**103** (2.50 g, 13.1 mmol) was added to an aqueous solution of cyanoacetylene **30** (20 mL, 1M, 20.0 mmol). After stirring the reaction mixture for 1 h at rt, it was immediately lyophilised to yield (*S*-*Z*-cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (3.26 g, 13.4 mmol, quant) as a white solid which was used without further purification: M.p 197-203°C. IR (solid, cm^{-1}) 3372 (OH), 3068, 3056, 2956 (CH), 2212 (C≡N), 1601 (C=N), 1575 (C=C). ^{1}H NMR (500 MHz, D_2O) δ 7.75 (1H, d, *J* = 10.5 Hz, H-(C4)), 6.10 (1H, d, *J* = 6.0 Hz, H-(C1`)), 5.90 (1H, d, *J* = 10.5 Hz, H-(C5)), 5.12 (1H, dd, *J* = 6.0, 1.4 Hz,

H-(C2`)), 4.35 (1H, ddd, *J* = 6.0, 3.4, 1.4, Hz, H-(C3`)), 4.06 (1H, ddd, *J* = 6.4, 4.7, 3.4 Hz, H-(C4`)), 3.54 (1H, ABX, *J* = 12.4, 4.7 Hz, H-(C5`)), 3.45 (1H, ABX, *J* = 12.4, 6.4 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 167.3 (C2), 143.5 (C5), 115.9 (C6), 100.3 (C4 and C1`), 91.5 (C2`), 86.4 (C4`), 76.0 (C3`), 61.5 (C5`). Analysis (% calcd, % found for $C_9H_{10}N_2O_4S$): C (44.62, 43.06), H (4.16, 4.03), N (11.56, 10.87). UV/Vis: λ_{max} 265 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -1.83. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_9H_{10}N_2O_4S$ calcd 243.0440, found 243.0447.

Crystal structure obtained see appendix for data.

(*S*-*Z*-Cyanovinyl)-ribofuranosyl oxazolidinone thione *ribo*-**120**



$C_9H_{10}N_2O_4S$
242.25

Ribofuranosyl oxazolidinone thione *ribo*-**103** (331 mg, 1.73 mmol) was added to an aqueous solution of cyanoacetylene **30** (5.0 mL, 1M, 4.90 mmol). After stirring the reaction mixture for 1 h at rt, it was immediately lyophilised to yield (*S*-*Z*-cyanovinyl)-ribofuranosyl oxazolidinone thione *ribo*-**120** (400 mg, 1.65 mmol, 95%) as a white solid which was used without further purification: M.p 193-198°C. IR (solid, cm^{-1}) 3344 (OH), 3055, 2926, 2876 (CH), 2218 (C≡N), 1653 (C=N), 1601 (C=C). ^{1}H NMR (600 MHz, D_2O) δ 7.88 (1H, d, *J* = 10.8 Hz, H-(C4)), 6.10 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.99 (1H, d, *J* = 10.8 Hz, H-(C5)), 5.22 (1H, t, *J* = 5.4 Hz, H-(C2`)), 4.27 (1H, dd, *J* = 9.3, 5.4 Hz, H-(C3`)), 3.95 (1H, dd, *J* = 13.2, 2.4 Hz, H-(C5`)), 3.76 (1H, dd, *J* = 13.2, 4.6 Hz, H-(C5``)), 3.60 (1H, ddd, *J* = 9.3, 4.6, 2.4 Hz, H-(C4`)). $^{13}$C NMR (151 MHz, $D_2O$) δ 168.4 (C2), 143.7 (C4), 116.0 (CN), 100.3 (C5), 98.9(C1`), 84.7 (C3`), 78.9 (C4`), 70.9 (C3`), 59.9 (C5`). UV/Vis: λ_{max} 258 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) +9.01. $[M\text{-}H^+]^-$ $C_9H_{10}N_2O_4S$ calcd 243.0493, found 243.0495.

8.7.1. One pot cyanovinylation of *ribo*- and *arabino* oxazolidinone thione

To a solution of arabinofuranosyl oxazolidinone thione *arabin*o-**103** (47.0 mg, 0.24 mmol) and ribfuranosyl oxazolidinone thione *ribo*-**103** (47.0 mg, 0.24 mmol) in H_2O (0.5 mL) at pH 7.0 was added aqueous cyanoacetylene **30** (1M, 0.25 mL, 0.24 mmol). The reaction mixture was stirred at rt for 1 h. After this time the sample was submitted for ^{1}H NMR spectroscopy analysis using single solvent suppression, which showed that there was no observable difference in the rate of cyanovinylation of *arabino*-**103** and *ribo*-**103** in the above reaction conditions.

8.7.2. Cyanovinylation of arabino aminooxazoline 70

In order to compare cyanovinylation of aminooxazoline *arabino*-**70** with with oxazolinone thione *arabino*-**103**, arabinofuranosyl aminooxazoline *arabino*-**70** (25.0 mg, 0.14 mmol) and cyanoacetylene **30** (1M, 1.5 mL) was stirred over night at rt without altering the contact pH. After 24 h the reaction mixture was submitted for ^{1}H NMR spectroscopic analysis using single solvent suppression indicating a range of cyanovinylated products including ancitabine *arabino*-**73** and arabino cytidine **185**.[22]

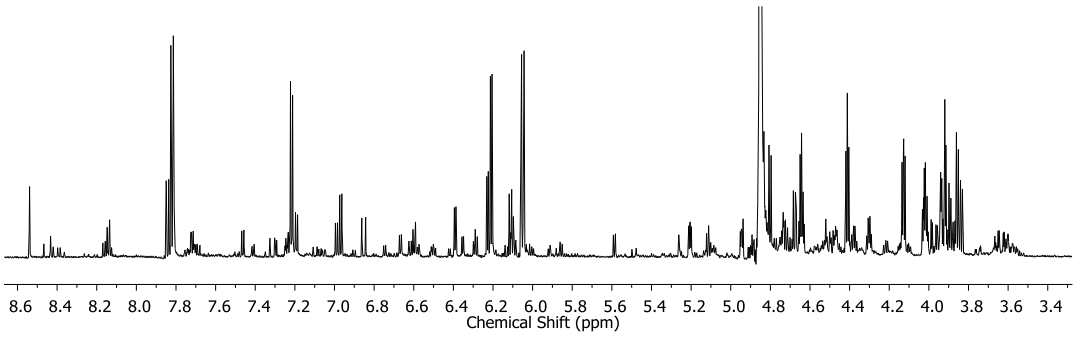
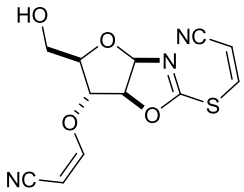


Figure 8.6: *Unbuffered cyanovinylation of arabinose aminooxazoline arabino-**70** with cyanoacetylene **30** after 24 h giving rise to a plethora of hydrolysis and cyanovinylated products.*

8.7.3. Bis-cyanovinylation of S-cyanovinyl thione

*S-Z-O*3`-biscyanovinyl-arabinofuranosyl oxazolidinone thione **122**



$C_{12}H_{11}N_3O_4S$
293.30

(*S-Z*-Cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (50 mg, 206 mmol) was added to an aqueous solution of cyanoacetylene **30** (1M, 10 mL). The resultant solution was incubated at pH 7.0 and rt. After 72 h, the mixture was lyophilised to give a powder, which was purified by silica gel flash column chromatography eluting with CH_2Cl_2/MeOH (0-5% MeOH). After evaporation of the fractions containing product **122**, *S-Z-O*3`-biscyanovinyl-arabinofuranosyl oxazolidinone thione **122** (12 mg, 41 mmol, 20%) was obtained as yellow powder. Unfortunately, the product obtained was not completely pure and contained residual peaks from the crude mixture, however both ^{1}H and COZY NMR spectra of the solids obtained indicated formation of **122** (Figure 8.7)

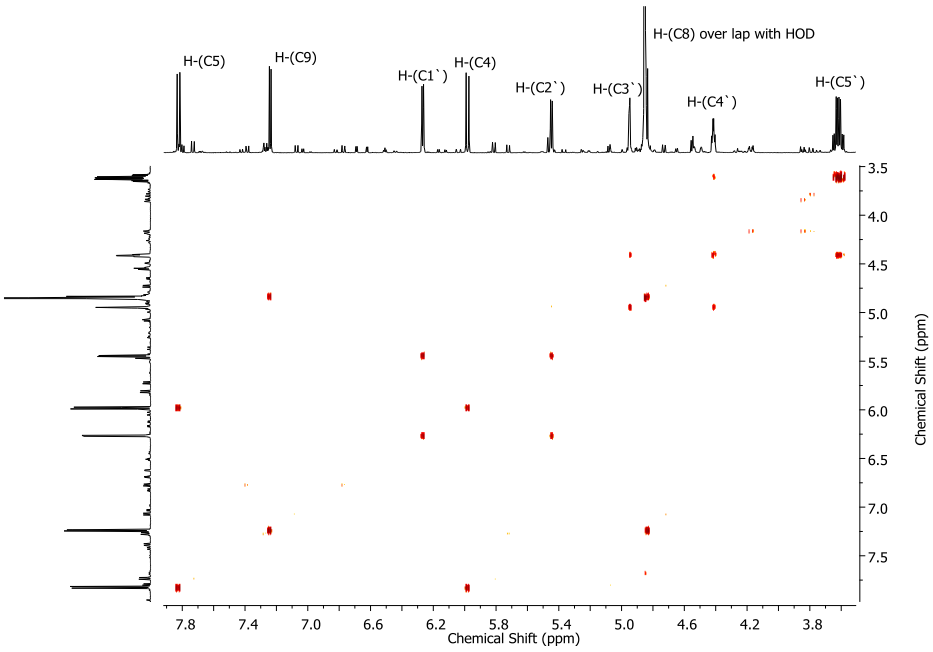


Figure 8.7: *^{1}H NMR spectra (600 MHz, D_2O, 3.5 – 8.1 ppm) for partially purified biscyanovinylated product **122**, the C3`-H is shifted upfield due to the presence of cyanovinyl group on C3`-O.*

8.7.4. Cyanovinylation of oxazolidionone thiones in the presence of aminooxazoline

To a solution of oxazolidinone thione *arabino*-**103** (47.0 mg, 0.24 mmol) and aminooxazoline *arabino*-**70** (42.0 mg 0.24 mmol) in H_2O (0.5 mL) at pH 7.0 and rt was added cyanoacetylene **30** (0.50 mL, 0.55 M 0.24 mmol). The resultant solution was left to react and the pH was monitored every 30 min. After 2 h when an aliquot of the reaction mixture was submitted for ^{1}H NMR spectroscopic analysis, it was observed that oxazolidinone thione *arabino*-**103** was selectively cyanovinylated to form *S*-cyanovinylated oxazolidinone thione *arabino*-**120** in the presence of aminooxazoline *arabino*-**70**. The ratio of *arabino*-**70** to *arabino*-**120** was 1:1, which was the same as the ratio of oxazolidinone thione *arabino*-**44** and aminooxazoline *arabino*-**81** in the beginning of the reaction (Figure 8.8).

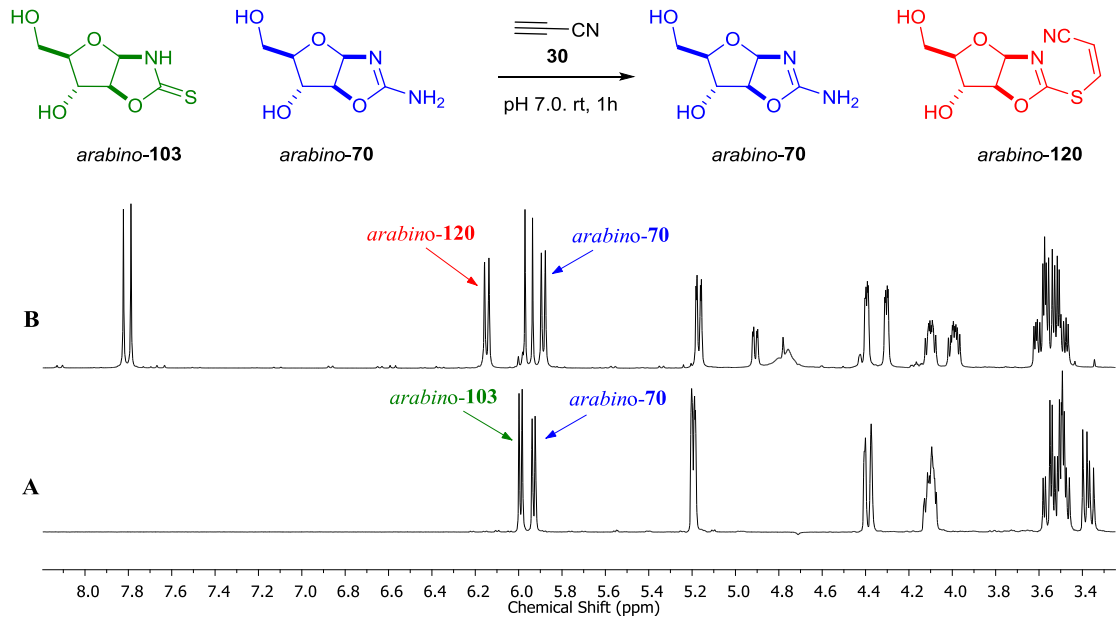
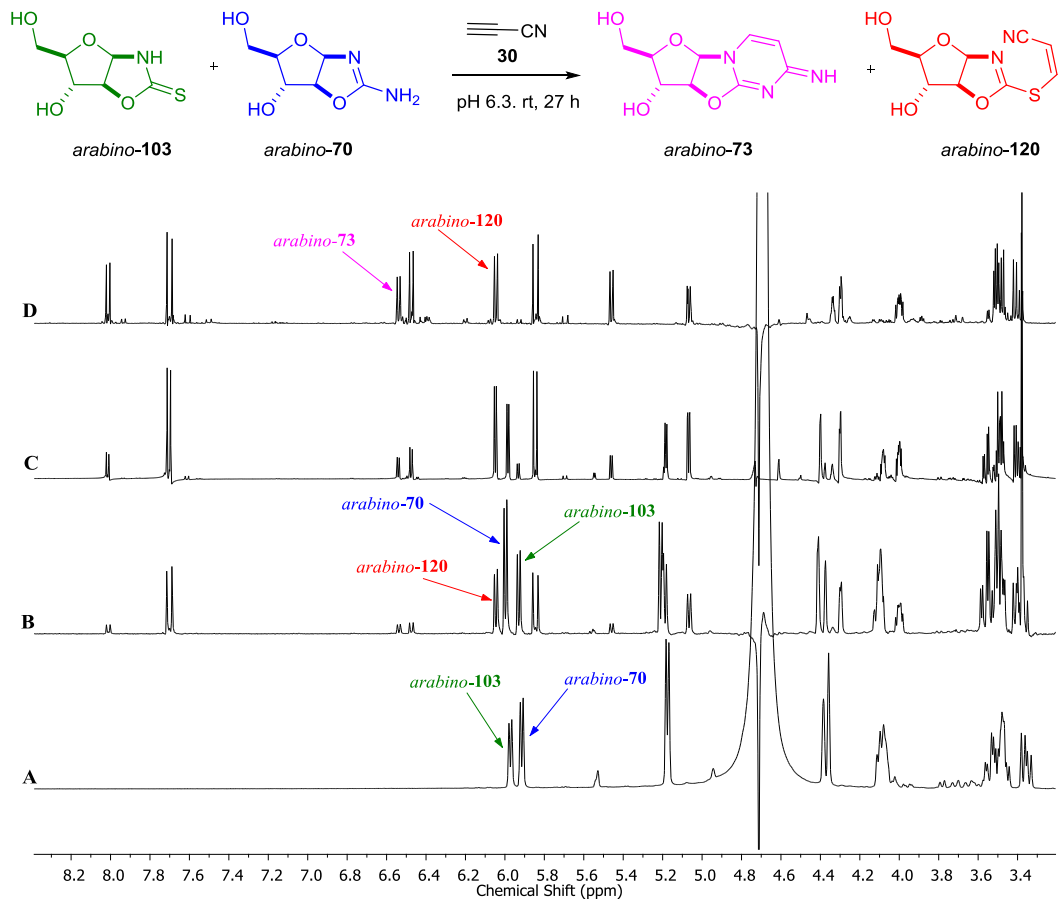


Figure 8.8: *^{1}H NMR spectra (600 MHz, D_2O, 3.3 – 8.2 ppm) for selective cyanovinylation of oxazolidionone thione arabino-**103** to furnish S-cyanovinyl oxazolidinone thione arabino-**120** in the presence of aminooxazoline arabino-**70**. **A)** Reaction at 0 h before addition of cyanoacetylene **30**. **B)** 1 h after addition of cyanoacetylene **30**.*

8.7.5. Cyanovinylation of aminooxazoline in the presence of oxazolidionone thiones

To as solution of aminooxazoline *arabino*-**70** (0.06 mmol) and oxazolidinone thione *arabino*-**103** (11.50 mg, 0.06 mmol) in H_2O (0.25mL) at pH 6.3 and rt was added cyanoacetylene **30** (1.25 mL, 0.50 M, 0.61 mmol).The resultant solution was then left to stand without stirring. The pH of the reaction was monitored and did not require any adjustments. The reaction was then analysed by ^{1}H NMR spectroscopy at 0, 1, 6, and 27 h (Figure 8.9).



Time (h)	*arabino*-**103** (%)	*arabino*-**70** (%)	*arabino*-**73** (%)	*arabino*-**120** (%)
0	55	45	0	0
1	30	45	0	25
6	6	37	12	45
27	0	0	45	55

Figure 8.9: *^{1}H NMR spectra (600 MHz, D_2O, 3.3 – 8.2 ppm) co-cyanovinylation of oxazolidinone thione arabino-**103** and aminooxazoline arabino-**70** to furnish S-cyanovinyl oxazolidinone thione arabino-**120** and ancitabine arabino-**73** respectively. **A)** Reaction at 0 h before addition of cyanoacetylene **30**. **B)** 1 h after addition of cyanoacetylene **30**. **C)** 6 h after addition of cyanoacetylene **30**. **D)** 27 h after addition of cyanoacetylene **30**.*

8.7.6. Attempted cyanovinylation of aminooxazoline by *S*-cyanovinyl oxazolidinone thione

Aminooxazoline *arabino*-**70** (21.0 mg, 0.12 mmol), *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (29.0 mg, 0.12 mmol) and sodium hydrogen phosphate (15 uL, 1 M) were dissolved in H_2O (1.5 mL) at pH 6.3, the resultant solution was then left to stand at rt for 96 h. After this time an aliquot of the reaction mixture was submitted for ^1H NMR spectroscopic analysis. The ^1H NMR showed approximately 1:1 oxazolidinone *arabino*-**77** and aminooxazoline *arabino*-**70**. ancitabine *arabino*-**73** and *arabino*-**120** were not observed (Figure 8.10).

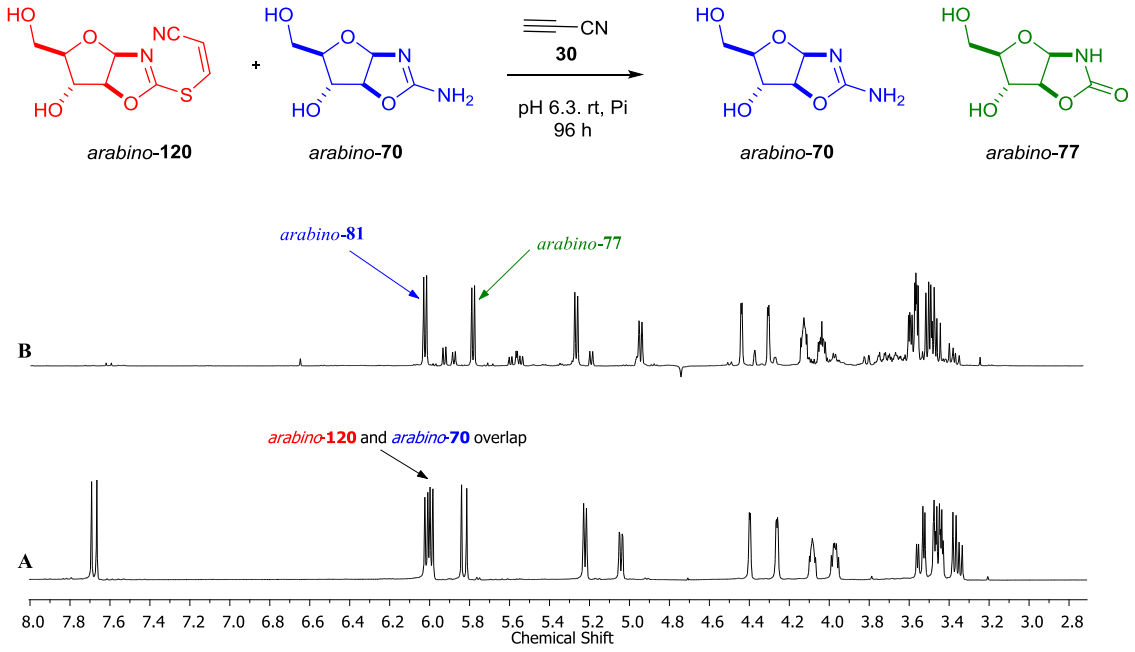


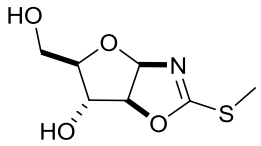
Figure 8.10: *^1H NMR spectra for cyanovinylation of aminooxazoline arabino-**70** with S-cyanovinyl oxazolidinone thione arabino-**120** to furnish ancitabine arabino-**73**. **A)** 0 h. **B)** 96 h.*

8.8. Synthesis of S-methyl oxazolidionone thiones

8.8.1. Conventional synthesis of S-methyl oxazolidionone thiones

S-Methylated oxazolidinone thiones were synthesised using an adapted protocol from Davidson *et al.*[176]

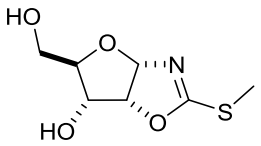
Arabinofuranosyl-(2-thiomethyl)-oxazoline *arabino*-**126**



$C_7H_{11}NO_4S$
205.23

Arabinofuranosyl oxazolidinone thione *arabino*-**103** (7.00 g, 36.4 mmol), iodomethane (5.84 g, 41.2 mmol) and NaOH (1.65 g, 41.2 mmol) were vigorously stirred in ethanol (280 mL) and H_2O (140 mL) at rt for 30 min. The solution was then concentrated and purified withby flash column chromatography eluting by $CHCl_3$/MeOH (9:1) to yield arabinofuranosyl-(2-thiomethyl)-oxazoline *arabino*-**92** (6.2 g, 30.2 mmol, 84%) as a white solid. M.p. 84-78°C: IR (solid, cm^{-1}) 3289 (OH), 2899, 2843 (CH), 1585 (C=N). ^{1}H NMR (600 MHz, D_2O) 6.10 (1H, d, *J* = 5.9 Hz, H-(C1`)), 5.09 (1H, d, *J* = 5.9, 1.8 Hz, H-(C2`)), 4.38 (1H, dd, *J* = 3.2, *J* = 1.8 Hz, H-(C3`)), 4.11 (1H, ddd, *J* = 6.9, 5.3, 3.2 Hz, H-(C4`)), 3.59 (1H, ABX, *J* = 12.2, 5.3 Hz, H-(C5`)), 3.51 (1H, ABX, *J* = 12.2, 6.9 Hz, H-(C5``)), 2.51 (3H, s, CH_3). ^{13}C NMR (151 MHz, D_2O) 174.5 (C2), 100.1 (C1`), 90.9 (C2`), 86.1 (C4`), 76.0 (C3`), 61.6 (C5`), 14.2 (CH_3). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -23.52. *m/z* (EI+) 205 (20%, M+). Data matches literature data.[176]

Ribofuranosyl-(2-thiomethyl)-oxazoline *ribo*-**126**



$C_7H_{11}NO_4S$
205.23

Ribofuranosyl oxazolidinone thione *ribo*-**103** (2.00 g, 10.5 mmol), iodomethane (1.67 g, 11.8 mmol) and NaOH (0.47 g, 11.8 mmol) were vigorously stirred in ethanol (80 mL) and H_20 (40 mL) at rt for 30 min. The solution was then concentrated and purified by flash column chromatography eluting with $CHCl_3$/MeOH (9:1) to yield ribofuranosyl-(2-thiomethyl)-oxazoline *ribo*-**126** (2.00 g, 9.74 mmol, 92%) as a white solid: M.p. 73-77°C. IR (solid, cm^{-1}) 3272 (OH), 2903, 2836, 2693 (CH), 1578 (C=N). ^{1}H NMR (600 MHz, D_2O) 6.00 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.10 (1H, t, *J* = 5.4 Hz, H-(C2`)), 4.21 (1H, dd, *J* = 9.4, 5.4 Hz, H-(C3`)), 3.94 (1H, ABX, *J* = 12.8, 2.1 Hz, H-(C5`)), 3.74 (1H, ABX, *J* = 12.8, 4.6 Hz, H-(C5``)), 3.56 (1H, ddd, *J* = 9.4, 4.6, 2.1 Hz H-(C4`)), 2.55 (3H, s, $CH_3$). $^{13}$C NMR (151 MHz, $D_2O$) 175.6 (C2), 98.7 (C1`), 84.0 (C2`), 78.6 (C4`), 71.0 (C3`), 60.0 (C5`), 14.4 (CH_3). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) +15.77. *m/z* (EI+) 205 (20%, M+). Data matches literature data.[176]

8.8.2. Prebiotic synthesis of S-methyl oxazolidionone thiones

(*S*-*Z*-Cyanovinyl)-arabinofuranosyl oxazolinone thione *arabino*-**120** (24.2 mg, 0.1 mmol) and 4,4-dimehtyl-4-silapentane-1-sulfonic acid (DSS, NMR standard, 15 mg) were dissolved in D2O (500 µL). Sodium acetate buffer (500 µL, 2M, pH 6, D_2O) was added and the solution adjusted to pH 6 with 1M NaOH/HCl. Methanethiol **127*** was bubbled through the solution for 10 min. Further methanethiol gas was added every 2 h. After 6 h the solution was briefly sparged with nitrogen, then cyanoacetylene **30** (200 µL, 1M) was added and the mixture was left to stand for 2 h. The solution was again saturated with methanethiol **127** gas for 10 min and incubated for 16 h before NMR spectra were acquired. NMR analysis indicated the presence of arabinofuranosyl-(2-thiomethyl)-oxazolinone *arabino*-**126** (50%), arabinofuranosyl

oxazolidone thione *arabino*-**103** (15%) and arabinofuranosyl oxazolidinone *arabino*-**77** (12%). (See Section 2.3.3 for NMR spectra)

*Methanethiol **127** gas was generated by dropping sodium thiomethoxide solution (21% v/v) onto anhydrous monobasic sodium phosphate at rt.

8.9. Stability of *S*-cyanovinyl oxazolidinone thione

An aqueous solution of *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (50 mM) at the required pH was stirred at rt and ^{1}H NMR spectra were periodically taken over 448 h. Aliquots of the reaction mixtures, which contained 4-dimehtyl-4-silapentane-1-sulfonic acid (DSS) from the beginning of the experiment, were analysed by ^{1}H NMR spectroscopy. Spectra were then carefully analysed for the presence of oxazolidinone thione *arabino*-**103**, oxazolidinone *arabino*-**77**, and *β*-*β*-dicyanovinyl-thioether **129**, each compound was confirmed by spiking with authentic samples. (Figure 8.11 and 8.12).

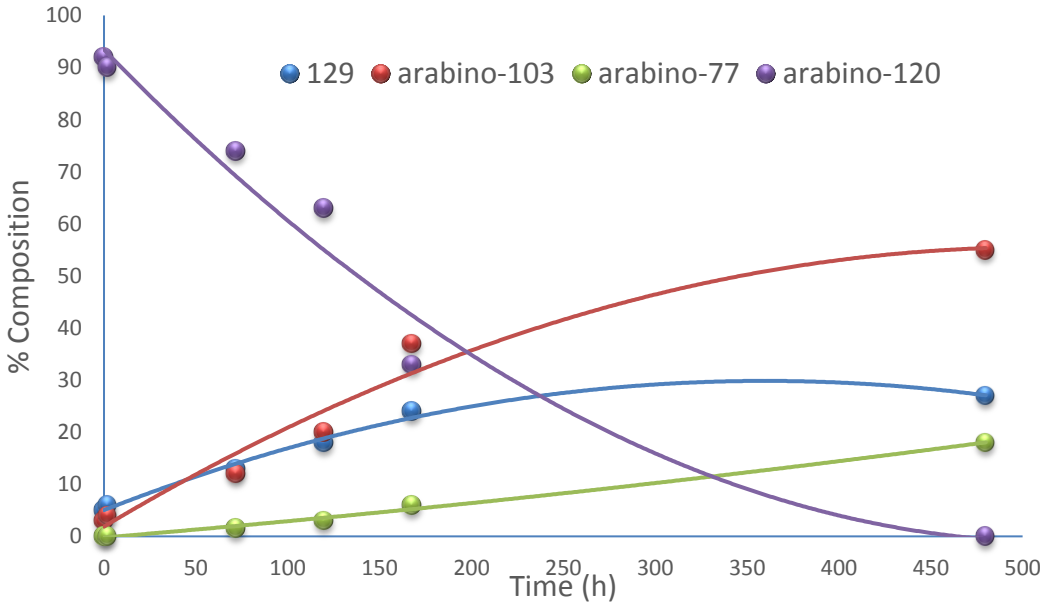


Figure 8.11: *Graph for pH stability of S-cyanovinylated oxazolidinone thione arabino-**120** at pH 8.0.*

pH 7.0				
Time (h)	*arabino*-**120**	**129**	*arabino*-**103**	*arabino*-**77**
0	92	5	3	0
2	90	6	4	0
72	74	13	12	1.5
120	63	18	20	3
168	33	24	37	6
480	0	27	55	18

pH 8.0				
Time (h)	*arabino*-**120**	**129**	*arabino*-**103**	*arabino*-**77**
0	92	5	3	0
2	90	6	4	0
72	73	13	12	2
120	56	18	21	6
168	33	24	36	7
480	0	27	56	17

pH 11.0				
Time (h)	*arabino*-**120**	**129**	*arabino*-**103**	*arabino*-**77**
0	92	5	3	0
2	90	6	4	0
72	0	13	31	2
120	-	-	-	-
168	-	-	-	-
480	-	-	-	-

Figure 8.12: *Tabulated data showing pH stability of S-cyanovinyl oxazolidinone thione arabino-**120** at pH 7.0, 8.0, and 11.0.*

Solution of sodium dihydrogen phosphate (50 mM) and *S*-cyanovinylated oxazolidinone thione *arabino*-**120** (50 mM) at pH 6.2 were incubated at rt for 3 d. After this time ^{1}H NMR spectra were acquired indicating: 75% oxazolidinone

arabino-**77**, 19% oxazolidinone thione *arabino*-**103** 6% *β*-*β*-dicyanovinyl-thioether **129** were obsereved (Figure 8.13).

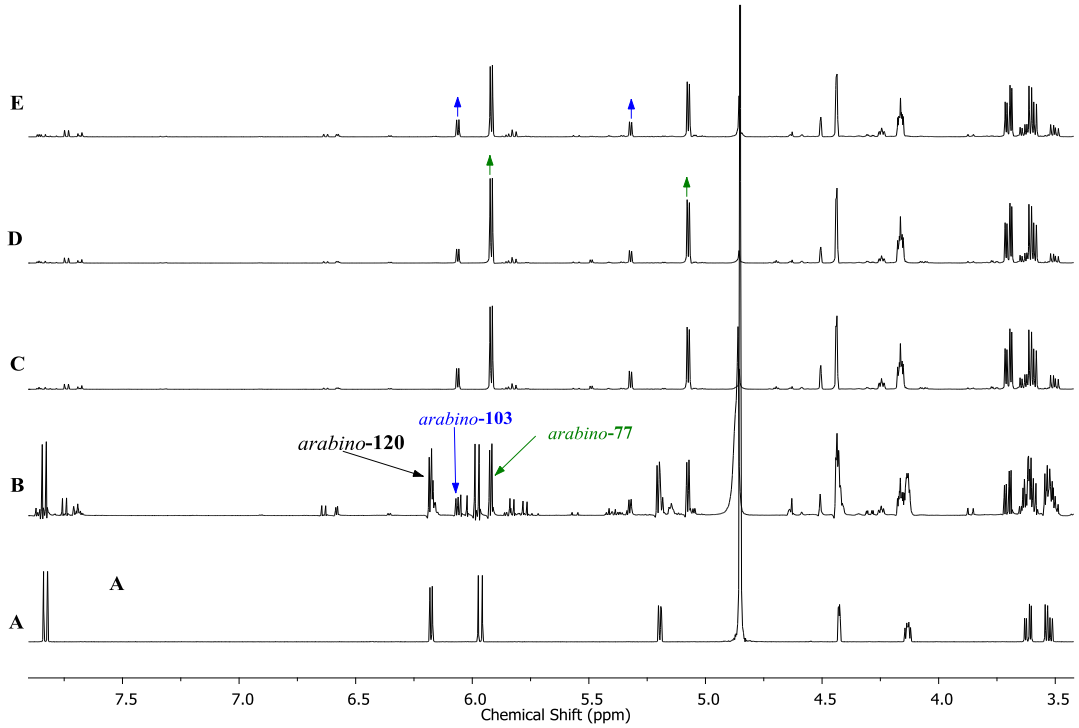
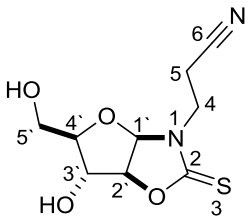


Figure 8.13: *^{1}H NMR spectra (600 MHz, D2O, 3.4 – 8.1 ppm) for pH stability of S-cyanovinyl oxazolidinone thione arabino-**120** in the presence of inorganic phosphate.* ***A)*** *0 h.* ***B)*** *10 h.* ***C)*** *96 h.* ***D)*** *Spiked with oxazolidinone arabino-**77**.* ***E)*** *Spiked with oxazolidinone thione arabino-**103**.*

8.10. Activation of oxazolidinone thione by acrylonitrile

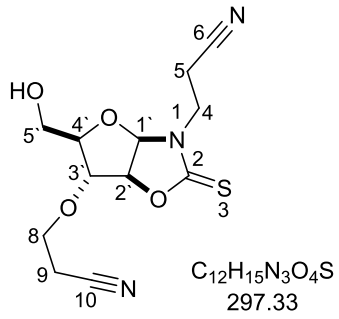
Arabinofuranosyl-(*N*1-cyanoethyl)-oxazolidinone thione **131**



$C_9H_{12}N_2O_4S$
244.05

Arabinofuranosyl oxazolidinone thione *arabino*-**103** (500 mg, 2.60 mmol), acrylonitrile (2.68 mL, 26.0 mmol) and H_2NaPO_4 (2.40g, 20.0 mmol) dissolved in water (28.0 mL) were stirred at pH 9 for 10 h. After this time the solution was lyophilised and purified with flash column chromatography eluting with EtAOc/MeOH (9:1) to yield the unexpected product arabinofuranosyl-(*N*1-cyanoethyl)-oxazolidinone thione **131** (606 mg, 2.48 mmol, 95%) as white solid: M.p 190-194°C. IR (solid, cm^{-1}) 3286 (OH), 3000, 2975, 2951, 2922 (CH), 2253 (C≡N), 1495 (C=S). ^1H (600 MHz, D_2O) 6.16 (1H, d, *J* = 5.9 Hz, H-(C1`)), 5.24 (1H, d, *J* = 5.9 Hz, H-(C2`)), 4.54 (1H, s, H-(C3`)), 4.33 – 4.27 (1H, m, H-(C4`)), 4.09 (2H, ABXY, *J* = 14.4, 7.5, 6.4 Hz, H-(C5), 4.02 (2H, ABXY, *J* = 14.4, 6.4, 6.4 Hz, H-(C4)), 3.60 (1H, ABX, *J* = 12.2, 5.3 Hz, H-(C5`)), 3.51 (1H, ABX, *J* = 12.2, 7.1 Hz, H-(C5``)), 3.06 (1H, ABXY, *J* = 17.3, 7.5, 6.4 Hz, (C4)), 2.99 (1H, ABXY, *J* = 17.3, 6.4, 6.4 Hz, (C4)). ^{13}C NMR (151 MHz, D_2O) 188.7 (C2), 120.0 (C6), 93.8 (C1`), 89.1 (C4`), 88.4 (23`), 75.6 (C3`), 61.6 (C5`), 42.7 (C4), 16.4 (C5). UV/Vis: λ_{max} 233 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -0.56. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_9H_{12}N_2O_4S$ calcd 245.0596, found 245.0587.

<u>Arabinofuranosyl-(*N*1, *O*3`-dicyanoethyl)-oxazolidinone thione **132**</u>



$C_{12}H_{15}N_3O_4S$
297.33

The above procedure also yielded arabinofuranosyl-(*N*1, *O*3`-dicyanoethyl)-oxazolidinone thione **132** (32.0 mg, 0.11 mol, 5%) as white solid: $^1H$ (600 MHz, $CDCl_3$) 5.91 (1H, d, *J* = 6.1 Hz, H-(C1`)), 5.14 (1H, d, *J*= 6.1 Hz, H-(C2`)), 4.28 (1H, d, *J* = 4.8 Hz, H-(C3`)), 4.22 (1H, dd, *J* = 8.7, 4.8 Hz, H-(C4`)), 4.06 (2H, ABXY, *J* = 14.0, 7.0, 7.0 Hz, H-(C5)), 3.83 (3H, m, H-(C5), H-(C5`) and H-(C5``)), 3.67 (2H, qd, *J* = 12.0, 4.9, 4.9 Hz, H-(C4)), 2.94 (2H, t, *J* = 6.6 Hz, H-(C8)), 2.67 (2H, t, *J* = 6.6 Hz, H-(C9)). ^{13}C NMR (151 MHz, $CDCl_3$) 187.3 (C2), 118.1 (C6), 117.5 (C9), 93.1 (C1`), 86.6 (C4`), 84.8 (C2`), 83.6 (C3`), 65.2 (C5`), 61.5 (C4), 42.6 (C5), 19.2

(C7), 16.3 (C8). HRMS (*m/z*): $[M\text{-}Na^+]^-$ $C_{12}H_{15}N_3O_4S$ calcd 320.0681, found 320.0678.

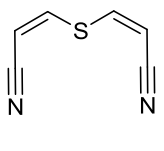
8.11. Nucleophilic additions

8.11.1. General procedure

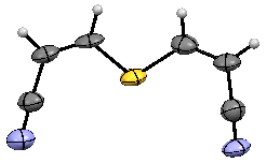
The nucleophile (0.25 mmol) was dissolved in H_2O (400 µL) at the desired pH. The activated thione (*S*-cyanovinyl *arabino*-**120** or *S*-metheyl *arabino*-**126**) and 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, 50 µL, 244 mM; internal standard) were then added. The solution was readjusted to the desired pH and the pH was maintained over the course of the reaction by addition of 1M NaOH/HCl as required. The reaction was monitored by periodic acquisition of ^1H NMR spectra with single solvent suppression technique. Each spectrum was acquired upon dilution of an aliquot (20 µL) of the reaction mixture adjusted to pH 9 by addition of ammonium hydroxide (450 µL, in 100mM D_2O).

Nucleophile additions to *S*-cyanovinyl oxazolidinone thione *arabino*-**120** resulted in the precipitation of a white solid that was isolated, characterised and determinded to be *β*-*β*-dicyanovinyl-thioether **129**.

β-*β*-dicyanovinyl-thioether **129**



$C_6H_4N_2S$
136.17



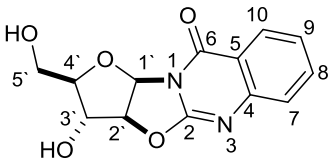
129

M.p 140-144°C (Lit.[264] 142.4-143.2°C). IR (Solid, cm^{-1}) 3073, 3061, 3046 (=C-H), 2214 (CN), 1578 (C=C), 1547 (C-C). ^1H NMR (600 MHz, DMSO) δ 8.05 (2H, d, *J* = 10.3 Hz, H-(CCN)), 6.12 (2H, d, *J* = 10.3 Hz, H-(CS)). ^{13}C NMR (151 MHz, DMSO) δ 147.2 (C-S), 115.4 (CN), 97.2 (C-CN). UV/Vis: λ_{max} 290 nm. *m/z* (CI+) 137 (100%, M+).

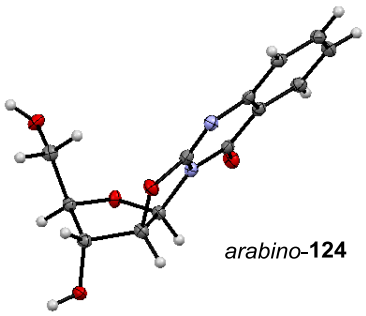
Crystal structure obtained; see appendix for crystallographic data.

8.11.2. Anthranilic acid

Arabinofuranosyl anthranilic quinazolinone *arabino*-**124**



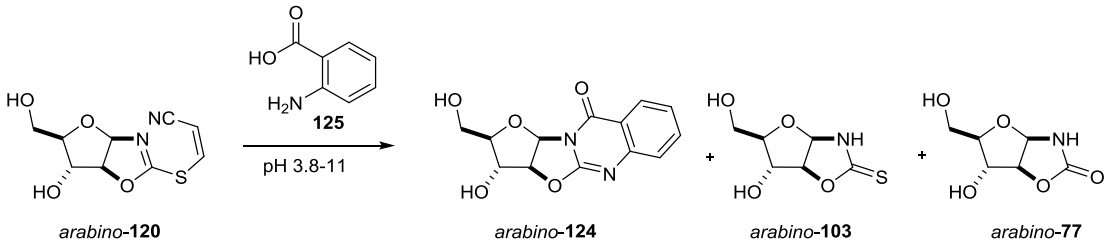
$C_{13}H_{12}N_2O_5$
276.25



(*S*-Z-Cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (100 mg, 0.041 mmol) was added to a solution of anthranilic acid **125** (113 mg, 0.082 mmol) in water (10.0 mL) at pH 3.8. The reaction was stirred for 8 h at rt, after this time the solution was lyophilised and purified by flash column chromatography eluting with EtAOc/MeOH (9:1) to yield arabinofuranosyl anthranilic quinazolinone *arabino*-**124** (40 mg, 0.14 mmol, 36%) as a white solid: M.p. 234-238°C. IR (solid, cm^{-1}) 3422 (OH), 3067, 2959, 2929, 2875 (CH), 1699 (C=O), 1644 (C=N), 1608 (C-N), 1562 (C=C). ^{1}H NMR (600 MHz, DMSO) δ 8.06 (1H, dd, *J* = 7.9, 1.3 Hz, H-(C10)), 7.82 – 7.65 (1H, m, 1H, H-(C9)), 7.44 (1H, dd, *J* = 8.1, 0.5 Hz, (C7)), 7.38 (ddd, *J* = 8.1, 7.2, 1.1 Hz, 1H, (C8)), 6.48 (1H, d, *J* = 5.6 Hz, H-(C1`)), 5.89 (1H, d, *J* = 4.3 Hz, H-(3O`)), 5.18 (1H, d, *J* = 5.6 Hz, H-(C2`)), 4.93 (1H, t, *J* = 5.3 Hz, H-(5O`)), 4.40 (1H, dd, *J* = 4.3, 2.0 Hz, H-(C3`)), 4.07 (1H, td, *J* = 5.3, 2.0 Hz, H-(C4`)), 3.30 (1H, ABXY, *J* = 12.0, 5.3, 5.3 Hz, H-(C5`)), 3.27 (1H, ABXY, *J* = 12.0, 5.3, 5.3 Hz, H-(C5``)). ^{13}C NMR (151 MHz, DMSO) δ 159.5 (C2), 155.0 (C10), 149.0 (C9), 135.1 (C8), 126.5 (C7), 125.9, 124.6, 118.6, 89.0 (C4`), 88.3 (C2`), 87.0 (C1`), 74.4 (C3`), 60.8 (C5`). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -32.5. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{13}H_{12}N_2O_5$ calcd 276.0746, found 276.0739.

Time course and pH experiments of anthranilic quinazolinone *arabino*-124

Following the general procedure for nucleophilic additions, (*S*-Z-cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (250 mM) and anthranilic acid **125** (500 mM) were incubated at varying pH conditions for 8 h at rt, and ^{1}H NMR spectroscopic data was collected (Figure 8.14).



arabino-**120** *arabino*-**124** *arabino*-**103** *arabino*-**77**

Contact pH	*arabino*-**124** (%)	*arabino*-**103** (%)	*arabino*-**77** (%)
3.8	96	3	1
5.0	88	8	2
6.0	81	12	3
7.0	70	29	1
8.5	70	30	0
11.0	65	33	2

Figure 8.14: *Tabulated data illustrating, how pH increase hinders the formation of arabinofuranosyl anthranilic quinazolinone arabino-**124** and promotes the formation of arabinofuranosyl oxazolidinone thione arabino-**103**.*

Ribofuranosyl anthranilic quinazolinone *ribo*-**124**

Following general procedure for nucleophilic additions, (*S*-Z-cyanovinyl)-ribofuranosyl oxazolidinone thione *ribo*-**120** (250mM) and anthraniclic acid **125** (500mM) in H_2O (500 µL) at pH 6, incubated at rt for 1 h to furnish the addition product *ribo*-**143** (Figure 8.15A). The reaction was then left to stand for 6 h and the cylisation of *ribo*-**143** was moitored by ^{1}H NMR spectroscopy. After 6 h, calibration to an internal standard (DSS) demonstrated that 70% ribofuranosyl anthranilic quinazolinone *ribo*-**124** and 13% ribofuranosyl oxazolinone thione *ribo*-**103** had been obtained (Figure 8.14B).

Crystal structure obtained; see appendix for crystallographic data.

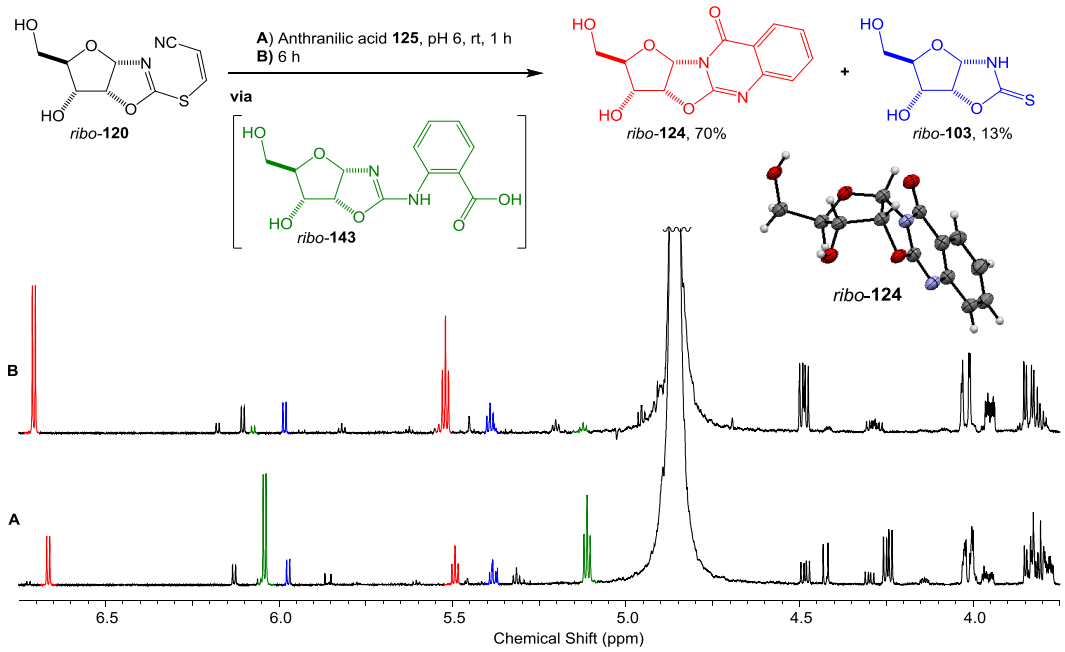


Figure 8.15: *^{1}H NMR spectra (600 MHz, D_2O, 3.75 – 6.75 ppm) for reaction of (S-Z-cyanovinyl)-ribofuranosyl oxazolinone thione ribo-**120** with anthranilic acid **125**. **A)** 1h. **B)** 6 h.*

8.11.3. Ammonia (NH_3)

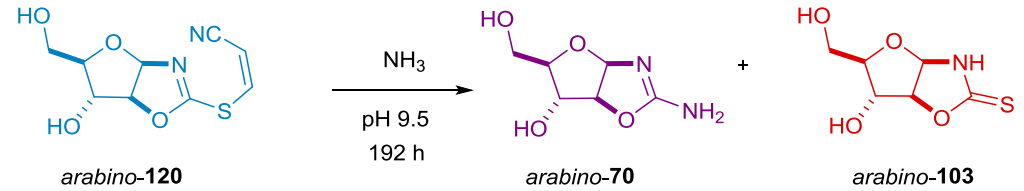
Reaction with *S*-cyanovinyl oxazolidinone thione

Following the general procedure for nucleophilic additions, (*S*-Z-cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (250 mM) and ammonia (1M) were incubated at pH 9.5 and rt for 48 h. ^{1}H NMR spectroscopic analysis indicated that 21% arabinofuranosyl aminooxazoline *arabino*-**70**, 35% arabinofuranosyl oxazolidinone thione *arabino*-**103** and *β*-*β*-dicyanovinyl-thioether **129** precipitate were obtained. (See Section 2.5.2 for ^{1}H NMR spectra).

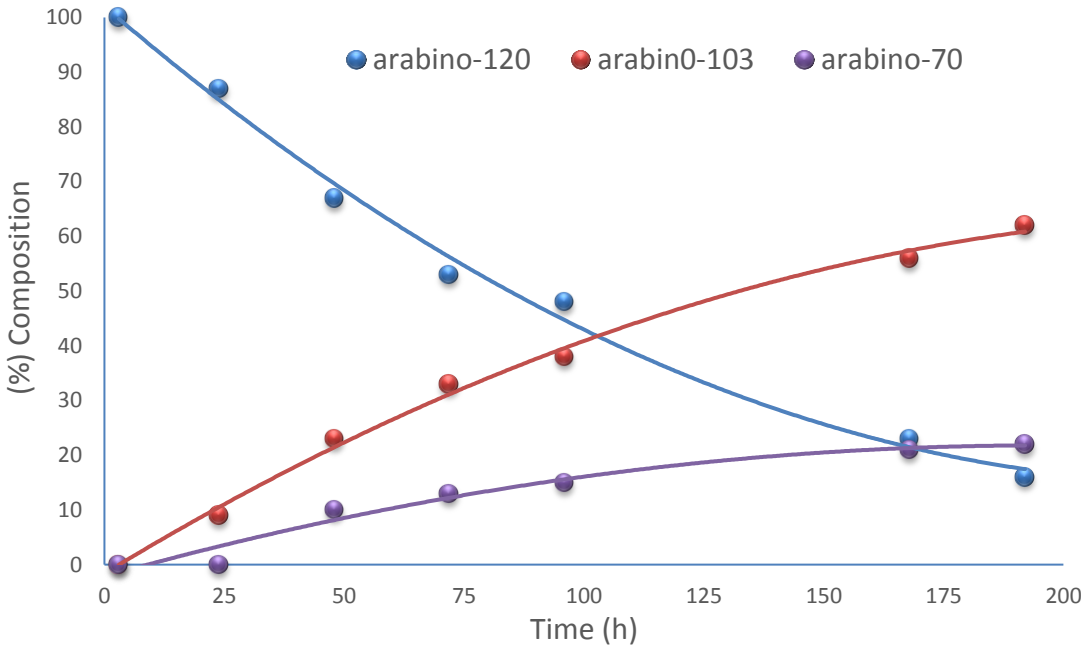
Time course and pH experiments of aminooxazoline *arabino*-70

Following the general procedure for nucleophilic additions, (*S*-Z-cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (250 mM) and ammonia (1M)

were incubated at pH 9.5 and rt for 48 h. ^{1}H NMR Data was collected periodically at time intervals of 0, 1, 18, 72, 96, 120, and 192 h (Figure 8.16)



arabino-**120** *arabino*-**70** *arabino*-**103**



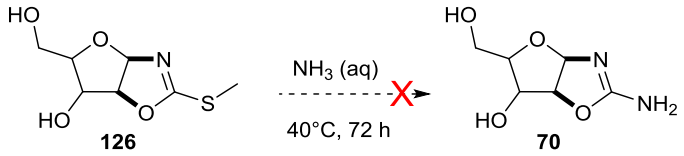
Time (h)	*arabino*-**120**	*arabino*-**103**	*arabino*-***70***
0	100	0	0
24	87	9	0
48	67	23	10
72	53	33	13
96	48	38	15
168	23	56	21
192	16	62	22

Figure 8.16: *Time course graph and tabulated data for the reaction of S-cyanovinylated oxazolidinone thione arabino-**120** with ammonia to furnish aminooxazoline arabino-**70**. Graph was plotted by taking ^{1}H NMR time points at 0, 1, 72, 96, 120 and 192 h of the crude reaction.*

Reaction with *S*-methyl oxazolidinone thione

To confirm Davidson *et al.* and Shaw *et al.* observations that *S*-methyl thione arabino-**126** was unreactive towards ammonia,[175, 176] ammonia was reacted with *arabino*-**126**.

Following the general procedure for nucleophilic additions, *S*-methyl oxazoline **126** (250 mM) and ammonia (1M) were incubated at pH 9.5 and rt for 72 h. ^{1}H NMR spectroscopic analysis did not indicate formation of aminooxazoline **70**. The reaction was repeated at elevated temperature (40°C), but again no reaction was obsereved (Scheme 8.1).



ribo-**126**
arabino-**126**

ribo-**70**
arabino-**70**

Scheme 8.1: *Reaction of S-methyl oxazolidine thione* ***126*** *and ammonia (NH_3).*

8.11.3.1. Regeneration of *S*-cyanovinylated oxazolidinone thione in ammonia displacement reaction

(*S*-Z-cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (250 mM) and ammonia (1M) were incubated at pH 10.5 and rt for 24 h. After this time cyanoacetylene **30** (0.25 mL, 0.5 M) was added to the mixture. ^{1}H NMR spectroscopic analysis of the solution indicated quantitative conversion of oxazolidinone thione *arabino*-**103** to *arabino*-**120** within 5 min. The pH of the reaction was at this point pH 10.5 and did not require any adjustments. Therefore the reactions was left to stand for further 24 h. After this time cyanoacetylene **30** (0.25 mL, 0.5 M) was added to the mixture. The pH of the reaction was again 10.5 and did not require any adjustments. Therefore the reaction was incubated for further 24 h at rt. Following this second recyanovinylation, ^{1}H NMR analysis revealed that 45% aminooxazoline *arabino*-**70** and 10% oxazolidinone thione *arabino*-**103** were present in the reaction mixture (See Section 2.5.2.1 for ^{1}H NMR spectra).

The above method for regerarting *S*-cyanovinyl *arabino*-**120** from oxazolidinone thione *arabino*-**103**, was also repeated for the *ribo* series. (*S*-Z-cyanovinyl)-

ribofuranosyl oxazolidinone thione *ribo*-**120** (250 mM) and ammonia (1M) were incubated at pH 10.5 at rt for 24 h and the mixture was then treated with cyanoacetylene **30** at pH 10.5 and rt for 5 min, which quantitatively converted oxazoidione thione *ribo*-**103** to *S*-cyanovinyl thione *ribo*-**120**. This cycle was repeated and after the second treatment with cyanoacetylene **30** the reaction mixture was incubated for a further 24 h at pH 10.5 and rt. After 24 h, ^{1}H NMR spectroscopic analysis calibrated to an internal DSS standard indicated that 35% ribofuranosyl aminooxazoline *ribo*-**70** and 11% ribofurnosyl oxazolidinone thione *ribo*-**103** had been formed (Figure 8.17). Upon standing it was obsereved that *ribo*-**70** (15 mg, 30%) directly crystalised from the mixture.

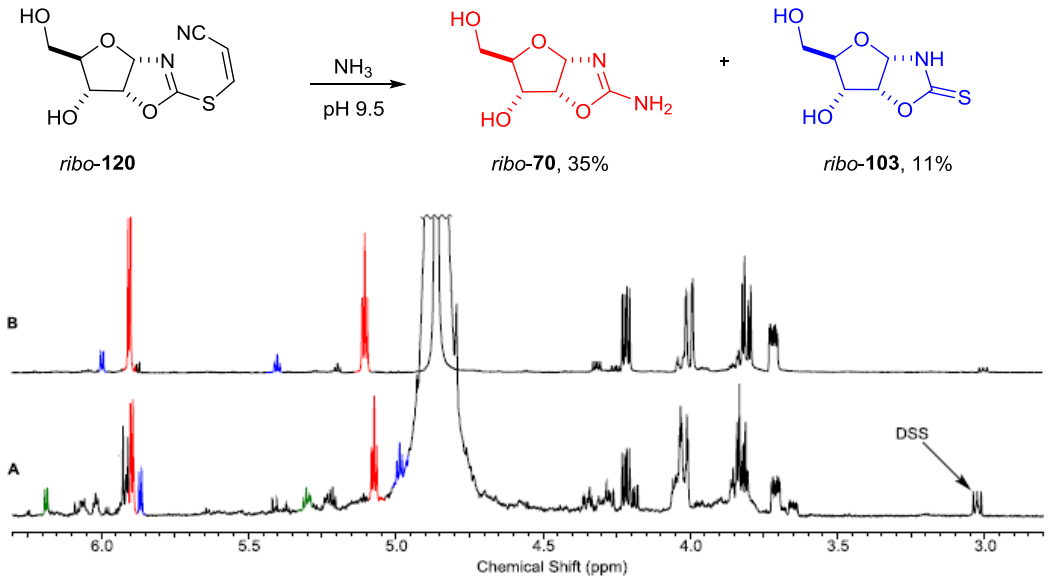
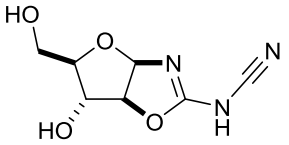


Figure 8.17: *^{1}H NMR spectra (600 MHz, 2.95 – 6.3 ppm) for the reaction of (S-Z-cyanovinyl)-ribofuranosyl oxazolidinone thione ribo-**120** with ammonia (NH_3).* ***A)*** *ribo-**120** (250 mM) and ammonia (1M) were incubated at pH 10.5 at rt for 24 h and the mixture was then treated with cyanoacetylene **30** at pH 10.5 and rt for 5 min. This cycle was repeated and after the 2nd treatment with **30** the reaction mixture was incubated for a further 24 h at pH 10.5 at rt.* ***B)*** *Crystals of ribofuranosyl aminooxazoline ribo-**70** obtained from the crude reaction mixture (crystals were not washed before analysis and were contaminated with ribo-**103**, it is not clear if ribo-**103** had co-crystalised or was supernant contamination).*

8.11.4. Cyanamide

(3-Acetonitrile)-arabinofuranosyl oxazolidinone thione **144**



$C_7H_9N_3O_4$
199.17

(*S*-Z-Cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (192 mg, 0.79 mmol) was added to a solution of cyanamide **68** (134 mg, 3.17 mol) in water (20.0 mL) at pH 6.5. The reaction was then stirred for 48 h at rt, and after this time the solution was lyophilised and purified by flash column chromatography eluting with EtAOc/MeOH (9:1) to yield (3-acetonitrile)-arabinofuranosyl oxazolidinone thione **144** (60.0 mg, 0.21 mmol, 26%) as white solid: M.p 192-195°C. IR (Solid, cm^{-1}) 3331 (NH), 3217 (OH), 2193 (CN), 1630 (C=C), 1551 (C=N). ^{1}H NMR (600 MHz, D_2O) 6.10 (1H, d, *J* = 5.7 Hz, H-(C1`)), 5.35 (1H, d, *J* = 5.7 Hz, H-(C2`)), 4.56 (1H, d, *J* = 1.8 Hz, H-(C3`)), 4.24 (1H, ddd, *J* = 6.8, 4.3, 1.8 Hz, H-(C4`)), 3.70 (1H, ABX, *J* = 12.6, 4.3 Hz, H-(C5`)), 3.61 (1H, ABX, *J* = 12.6, 6.8 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 166.6 (C2), 117.1 (CN), 91.7 (C1`), 89.6 (C4`), 88.1 (C2`), 75.3 (C3`), 61.7 (C5`). UV/Vis: λ_{max} 213 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -8.17. HRMS (*m/z*): $[M-H^+]^-$ $C_7H_9N_3O_4$ calcd 200.0671, found 200.0680.

8.11.5. 2-Aminoacetonitrile (glycine nitrile)

Reaction with *S*-cyanovinyl oxazolidinone thione

Following the general procedure for nucleophilic additions, a solution of glycine nitrile.HCl **145** (23.1 mg, 0.25 mmol) in H_2O (0.50 mL) was adjusted to pH 5.2 with $NaOH_{aq}$ (1M) then (*S*-Z-cyanovinyl)-arabinofuranosyl oxazolidinone thione *arabino*-**120** (30.3 mg, 0.12 mmol) was added in a single portion and the solution was readjusted to pH 5.2. The reaction was stirred for 6 h at rt. After 6 h an aliquot (50 µL) was submitted to NMR studies. ^{1}H NMR spectroscopic analysis revealed that

70% aminooxazoline **146** and 14% oxazolidinone thione *arabino*-**103** had been formed (Figure 8.18). The identities of each compound in the crude mixture was confirmed by spiking with authentic samples.

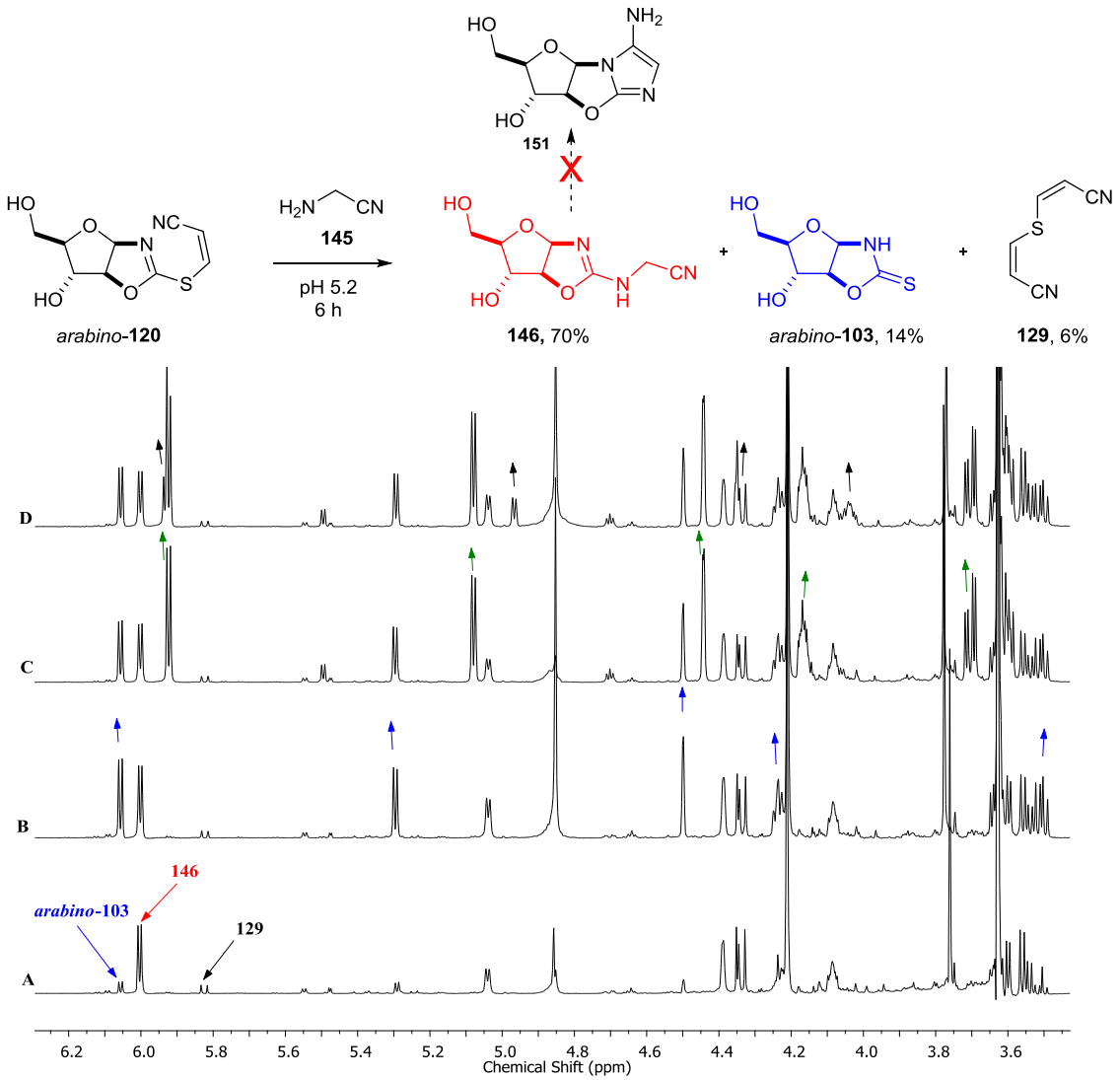


Figure 8.18: *^{1}H NMR spectra (600 MHz, D_2O, 3.5 – 6.3 ppm) showing spectroscopic evidence by sequential spiking for formation of aminooxazoline **146** from the reaction of S-cyanovinylated oxazolidinone thione arabino-**120** and glycine nitrile **145** **A)** The crude reaction of glycine nitrile **145** (500mM) and arabino-**120** (250mM) at pH 5.2 after 10 h. **B)** Spiked with oxazolidinone thione arabino-**103**. **C)** Spiked with oxazolidinone arabino-**77**. **D)** Spiked with aminooxazoline arabino-**70**.*

To investigate the cyclisation in aminooxazoline **146**, the above reaction was repeated then an aliquot (100 µL) was diluted with NaOH solution to 400 µL, and the

specificed pH 8–10. After adjusting the reaction to specified pH (between pH 8–10), the reaction was monitor by NMR spectroscopy periodically for 24 h. ^{1}H NMR spectroscopic analysis revealed that no cyclisation to **151** had occurred, only residual aminooxazoline **146** was observed at all pHs investigated.

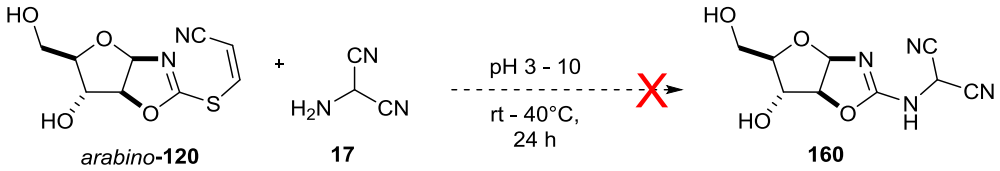
Reaction with *S*-methyl oxazolidinone thione

Following the general procedure for nucleophilic additions, solution of glycine nitrile.HCl **145** (23.1 mg, 0.25 mmol) in H_2O (0.5 mL) was adjusted to pH 4.5 with $NaOH_{aq}$ (1M) then *S*-methyl oxazolidinone thione *arabino*-**126** (24.0 mg, 0.12 mmol) was added in a single portion and the solution was readjusted to pH 4.5. The reaction was stirred for 2 h at rt, and after 2 h an aliquot (50 µL) was submitted for NMR studies. ^{1}H NMR spectroscopic analysis revealed that 76% aminooxazoline **146** and 8% oxazolidinone *arabino*-**77** were present in the solution. The identities of each compound in the crude mixture was confirmed by spiking with authentic samples. (See section 2.5.4.1 for NMR spectra)

8.11.6. 2-Aminomalononitrile (HCN trimer)

Reaction with S-cyanovinyl oxazoidione thione

Following the general procedure for nucleophilic additions, solution of aminomalononitrile *p*-toluenesulfonate **17** (254 mg, 1.00 mol) in H_2O (1 mL) was adjusted to required pH (3–10) with $NaOH_{aq}$ (1M) then *S*-cyanovinyl oxazolidinone thione *arabino*-**49** (60.6 mg, 0.25 mmol) was added in a single portion and the solution was readjusted to the required pH (3–10). The reaction was stirred for 24 h at rt, and aliquots (50 µL) were submitted periodically (every 2 h) for NMR studies. ^{1}H NMR spectroscopic analysis did not show the formation of aminooxazoline **160** and only oxazolidinone thione *arabino*-**103** and oxazolidionone *arabino*-**77** were obsereved (Scheme 8.2). The identity of each compound in the crude mixture was confirmed by spiking with authentic samples.



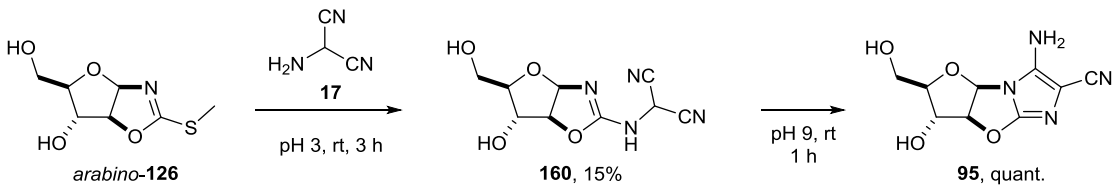
Scheme 8.2: *Reaction of 2-aminomalononitrile* ***17*** *with S-cyanovinyl oxazolinone thione arabino-****120****.*

Reaction with *S*-cyanovinylated oxazolidinone thione (Slow addition method)[52]

Following the adapted literature procedure of Orgel *et al.*[52], for the synthesis of HCN tetramer **18**, to an aqueous solution of *S*-cyanovinylated oxazolidinone thione *arabino*-**120** (10 mL, 0.20 mM), adjusted to pH 3.5 with dilute HCl, and was added aminomalononitrile *p*-toluenesulfonate **17** (10 mL, 0.02 mM) over a 6 h period. The pH of the reaction was maintained between 3.5-4.0 by slow addition of dilute HCl. At 1 h intervals, aliquots (0.2 mL) of the crude mixture were diluted with D_2O (0.3 mL) and then submitted to ^1H NMR spectroscopic analysis. ^1H NMR analysis indicated that no addition of HCN trimer **17** had occurred, only oxazolidinone thione *arabino*-**103** and oxazolidinone *arabino*-**77** was detected.

Reaction with S-methyl oxazolidinone thione

Following the general procedure for nucleophilic additions, solution of aminomalononitrile *p*-toluenesulfonate **17** (127 mg, 0.25 mol) in H_2O (0.5 mL) was adjusted to pH 3.0 with $NaOH_{aq}$ (1M) then *S*-methyl oxazolidinone thione *arabino*-**126** (24.0 mg, 0.12 mmol) was added in a single portion and the solution was readjusted to pH 3. The reaction was stirred for 3 h at rt, and after 3 h an aliquot (50 µL) was submitted to NMR studies. ^1H NMR spectroscopic analysis revealed that 10% aminooxazoline **160** was present in the reaction mixture. To induce cylisation in **160**, an aliquot (50 µL) from the above reaction was added to 100 mM ammonium hydroxide in D_2O (450 µL). The solution was adjusted to pH 9 with NaOH (4M) and incubated for 1 h at rt. ^1H NMR spectroscopic analysis and calibration to an internal standard demonstrated that aminoimidazole-4-carbonitrile-*β*-furanosylarabinoside **95** and oxazolidinone *arabino*-**77** had been furnished in 15 and 8% yield, respectively (Scheme 8.3). The identity of each compound in the crude mixture was confirmed by spiking with authentic samples. (See section 2.5.4.2 for NMR spectra).

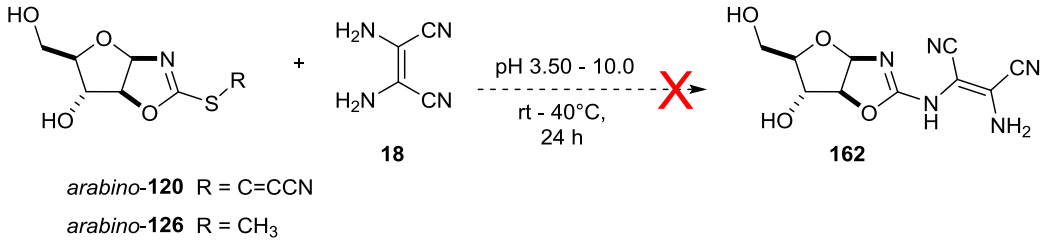


Scheme 8.3: *Reaction of 2-aminomalononitrile **17** with S-methyl oxazolinone thione arabino-**126**.*

8.11.7. Diaminomalononitrile (HCN tetramer)

Reaction with *S*-cyanovinyl and *S*-methyl oxazoidoinone thione

Following the general procedure for nucleophilic additions, diaminomalononitrile **18** (54 mg, 0.5 mmol) in H_2O (1 mL) was adjusted to the required pH (pH 3–10) with $NaOH_{aq}$ (1M) then *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (60.6 mg, 0.25 mmol) or *S*-methyl oxazolidinone thione *arabino*-**126** (48.0 mg, 0.25 mmol) was added in a single portion. The resultant solution was readjusted to the required pH (3–10). The reaction was stirred for 24 h at rt, and aliquots (50 µL) were submitted periodically (every 2 h) to NMR studies. ^{1}H NMR spectroscopic analysis did not show formation of aminooxazoline **162**, only oxazolidinone thione *arabino*-**103** and oxazolidionone *arabino*-**77** were obsereved (Scheme 8.4). The identity of each compound in the crude NMR spectrum was confirmed by spiking with authentic samples.



Scheme 8.4: *Reaction of S-cyanovinyl oxazolidinone thione arabino-**120** and S-methyl oxazolidinone thione arabino-**126** with diaminomalononitrile **18**.*

8.11.8. 2-Amino-2-cyanoacetamide

Reaction with S-cyanovinyl oxazoidione thione

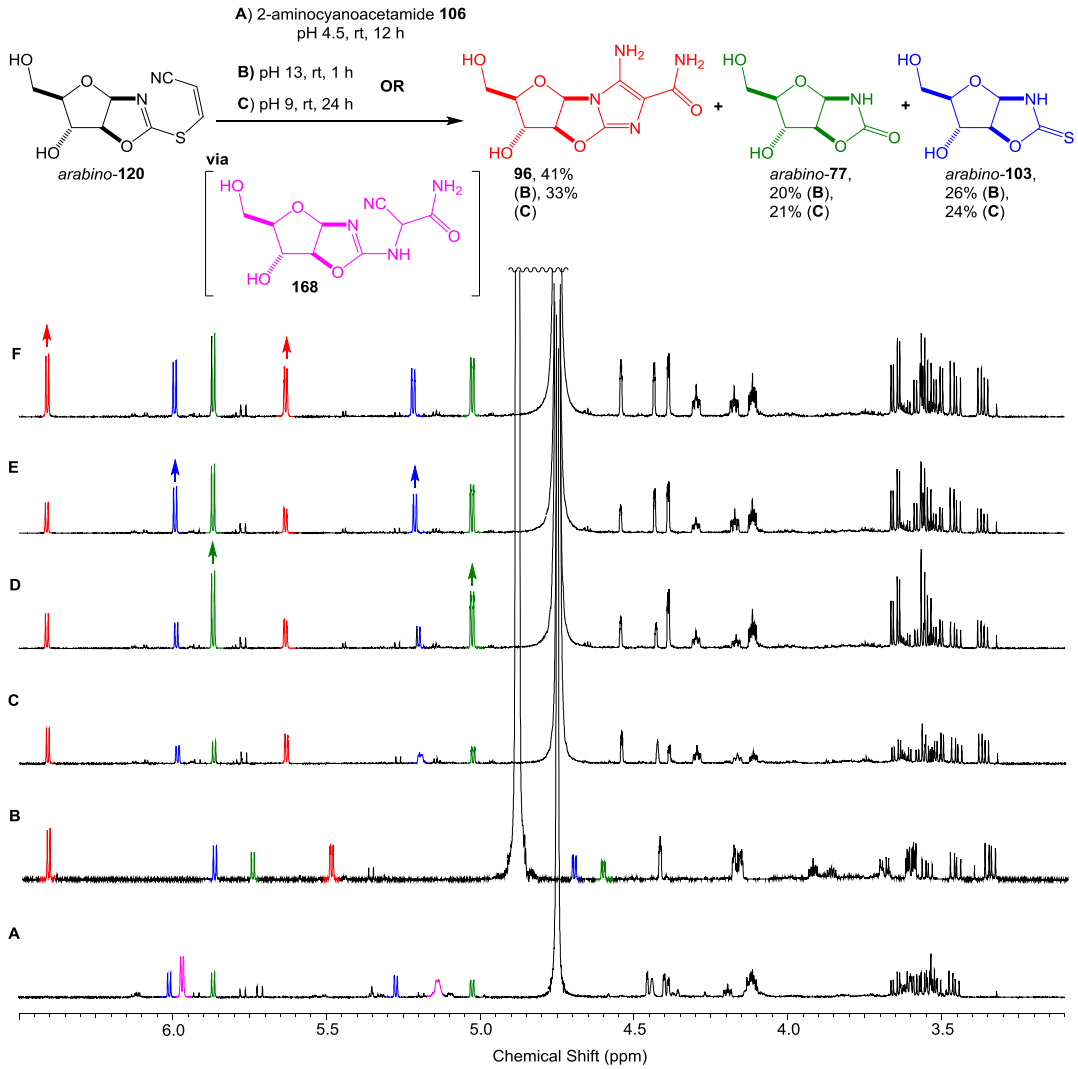


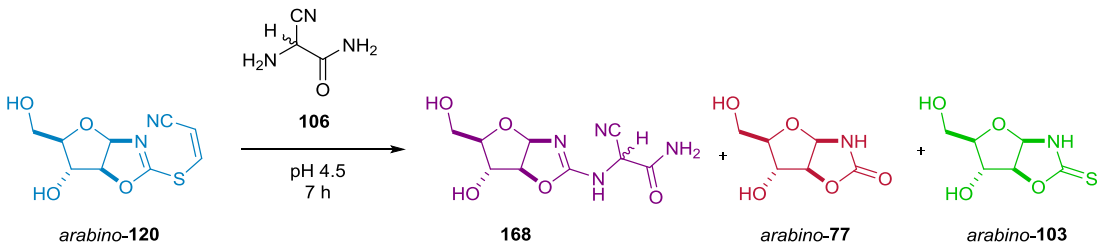
Figure 8.19: *^{1}H NMR spectra (600 MHz, D_2O, 3.1 – 7.0 ppm) for reaction of S-cyanovinyl oxazolinone thione arabino-**120** with 2-aminocyanoacetamide **106**. **A)** arabino-**120** (250mM) and **106** (500mM) in H_2O (500 µL) at pH 4.5 incubated at rt for 12 h. **B)** Aliquot of **A** incubated at pH 13 and rt for 1 h giving a yield of 41% 2,2'-anhydro-5-aminoimidazole-4-carboxamide-β-furanosylarabinoside **96**. **C)** An aliquot (50 µL) of **A** was added to ammonium hydroxide in H_2O (450 µL, 100mM), the solution was adjusted to pH 9 with NaOH (4M) and incubated for 24 h at rt. Calibration to an internal standard gave yields of 33% **96**, 24% oxazolinone thione arabino-**103**, 21% oxazolidinone arabino-**77**. **D)** Spiked with arabino-**77**. **E)** Spiked with arabino-**103**. **F)** Spiked with **96**.*

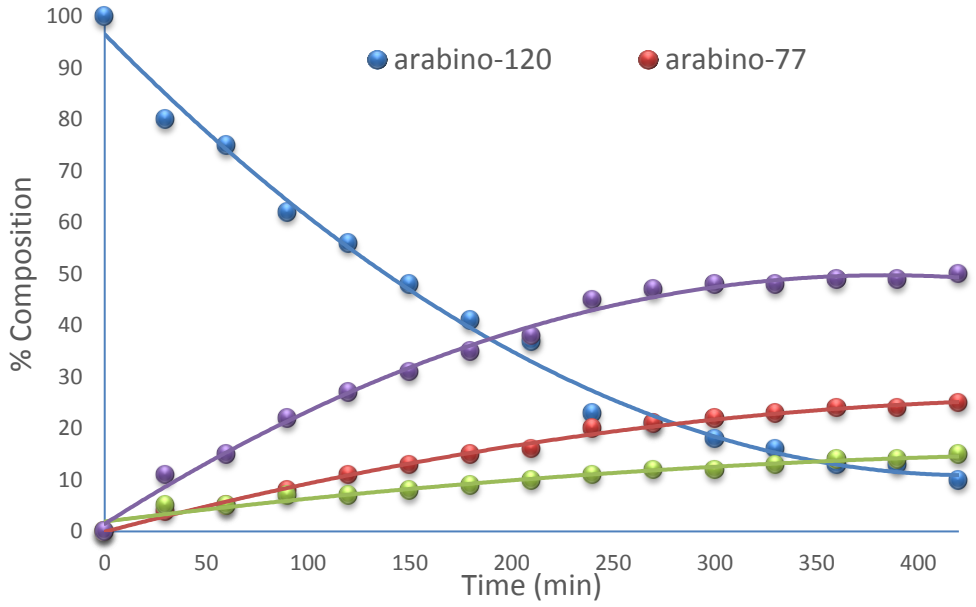
Following the general procedure for nucleophilic additions, a solution of 2-amino-2-cyanoacetamide **106** (24.8 mg, 0.25 mol) in H_2O (0.5 mL) was adjusted to required pH 4.5 with $NaOH_{aq}$ (1M), then *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (30.3 mg, 0.12 mmol) was added in a single portion. The resultant solution was readjusted to pH 4.5 and the reaction was stirred for 12 h at rt. After 12 h an aliquot (50 µL) of the reaction mixture was submitted for NMR studies. ^{1}H NMR spectroscopic analysis revealed that aminooxazoline **168** had been furnished in 50% yield. (Figure 8.19A).

To induce cylisation in **168**, an aliquot (50 µL) from the above reaction was added to 100 mM ammonium hydroxide in D_2O (450 µL), the solution was then adjusted to pH 13 with NaOH (4M) and incubated for 1 h at rt. ^{1}H NMR spectroscopic analysis and calibration to an internal standard demonstrate that aminoimidazole-4-carbonitrile-*β*-furanosylarabinoside **96**, oxazolidinone *arabino*-**77**, and oxazolidinone thione *arabino*-**103** have been synthesised in 41, 20 and 26% yield, respectively (Figure 8.19B). Cyclisation of **168** was also investigated under milder alkaline conditions by incubation of a second aliquot of the above reaction at pH 9 and rt for 24 h. ^{1}H NMR spectra of this reaction revealed that **96**, *arabino*-**77**, and *arabino*-**103** had been furnished in 33, 21 and 24% yield, respectively (Figure 8.19C). The identity of each compound in the crude NMR spectrum was confirmed by spiking with authentic samples (Figure 8.19D-F).

Time course and pH experiments of aminooxazoline 168

Following the general procedure for nucleophilic additions, solution of 2-amino-2-cyanoacetamide **106** (50.4 mg, 0.50 mol) in H_2O (1 mL) was adjusted to required pH 4.5 with $NaOH_{aq}$ (1M) then *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (60.6 mg, 0.25 mmol) was added in a single portion and the solution was readjusted to pH 4.5. The reaction was stirred for 7 h at rt, and aliquots (50 µL) which had 4-dimehtyl-4-silapentane-1-sulfonic acid (DSS, NMR standard) from the beginning of the experiment, were submitted for ^{1}H NMR studies periodically (at 1 h intervals). ^{1}H NMR data was then carefully analysed and presence of aminooxazoline **168**, oxazolidinone thione *arabino*-**103**, and oxazolidinone *arabino*-**77** were confirmed by spiking with authentic samples (Figure 8.20).



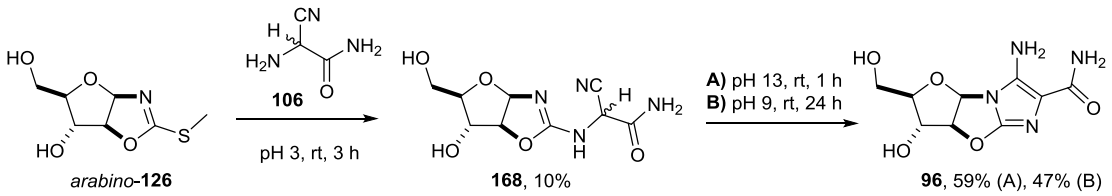


Time (min)	*arabino*-**120**	*arabino*-**77**	*arabino*-**103**	**168**
0	100	0	0	0
30	80	4	5	11
60	75	5	5	15
90	62	8	7	22
120	56	11	7	27
150	48	13	8	31
180	41	15	9	35
210	37	16	10	38
240	23	20	11	45
270	21	21	12	47
300	18	22	12	48
330	16	23	13	48
360	13	24	14	49
390	13	24	14	49
420	10	25	15	50

Figure 8.20: *Time course reaction for formation of aminooxazoline* ***168*** *from S-cyanovinyl oxazolidinone thione arabino-****120****.*

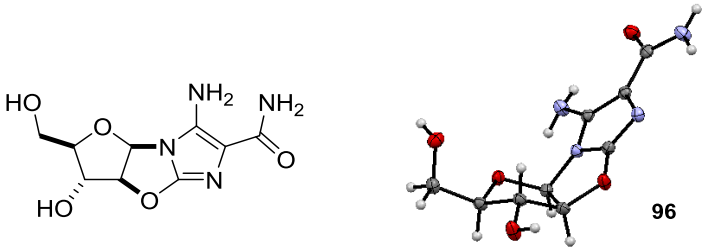
Reaction with *S*-methyl oxazolidinone thione

Following the general procedure for nucleophilic additions, solution of of 2-amino-2-cyanoacetamide **106** (24.8 mg, 0.25 mol) in H_2O (0.5 mL) was adjusted to pH 4.5 with $NaOH_{aq}$ (1M) then *S*-methyl oxazolidinone thione *arabino*-**126** (24.0 mg, 0.12 mmol) was added in a single portion and the solution was readjusted to pH 4.5. The reaction was stirred for 6 h at rt, after 6 h an aliquot (50 µL) was submitted for NMR studies. ^{1}H NMR spectroscopic analysis revealed that aminooxazoline **168** had been furnished in 81% yield. To induce cylisation in **168**, an aliquot (50 µL), from the above reaction was added to 100 mM ammonium hydroxide in D_2O (450 µL), the solution was adjusted to pH 13 with NaOH (4M) and incubated for 1 h at rt. ^{1}H NMR spectroscopic analysis, and calibration to an internal standard, demonstrated that aminoimidazole-4-carbonitrile-*β*-furanosylarabinoside **96** had been synthesised in 59% yield. Cyclisation of **168** was also investigated under mildly alkaine conditions; an aliquot (50 µL) from the above reaction was incubated at pH 9 and rt for 24 h to furnished a 47% yield of **96** (Scheme 8.5). The identity of each compounds in the crude NMR spectrum was confirmed by spiking with authentic samples. (See section 2.5.4.3 for NMR spectra)



Scheme 8.5: *Reaction of 2-amino-2-cyanoacetamide **106** with S-methyl oxazolinone thione arabino-**126**.*

2,2′-anhydro-5-aminoimidazole-4-carboxamide-*β*-furanosylarabinoside **96**

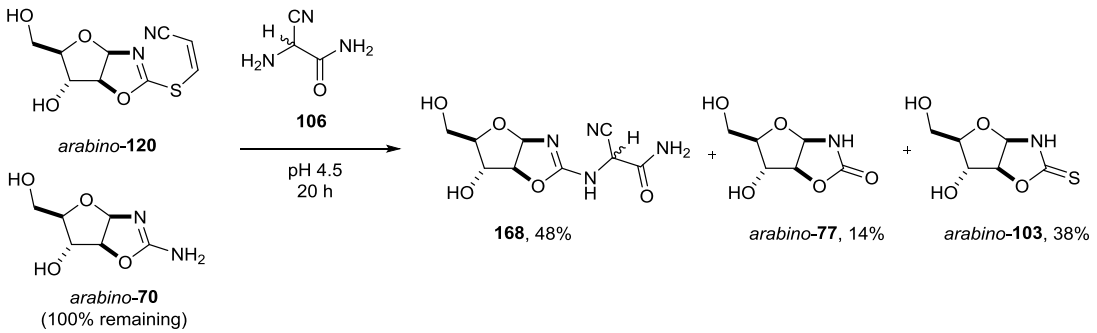


To a solution of 2-amino-2-cyanoacetamide **106** (108 mg, 1.09 mmol) in water (1 mL) at pH 4.5 was added *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (354 mg, 1.46 mmol). The resultant mixture was stirred for 3 h at 45°C and the solution was kept between pH 4.3–4.7 during the course of the reaction. After 3 h the pH of the reaction was raised to pH 8.1 and the reaction was left to stir overnight at 45°C. The mixture was then allowed to cool to rt and then placed in a fridge (7°C). After 3 h in the fridge white solids had precipitated from the crude reaction mixture. These solid were isolated by filteration, washed with cold methanol and dried to yield 2,2′-anhydro-5-aminoimidazole-4-carboxamide-*β*-furanosylarabinoside **96** (37.2 mg, 0.15 mmol, 13%) as white solid: M.p 234-238°C. IR (solid, cm^{-1}) 3493 (NH_2-C=O), 3359 (NH_2-C), 3210 (OH), 1638 (C=O), 1581 (C=N), 1533 (C=C). ^{1}H NMR (600 MHz, D_2O) 6.46 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.68 (1H, d, *J* = 5.4 Hz, H-(C2`)), 4.60 (1H, br s, H-(C3`)), 4.38 (1H, ddd, *J* = 6.8, 4.9, 1.9 Hz, H-(C4`)), 3.56 (1H, ABX, *J* = 12.4, 4.9 Hz, H-(C5`)), 3.42 (1H, ABX, *J* = 12.4, 6.8 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 169.1 (C6),152.6 (C2), 140.0 (C5), 109.9 (C4), 97.9 (C2`), 89.3 (C4`), 86.5 (C1`), 75.3 (C3`), 61.5 (C5`). UV/Vis: λ_{max} 240 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -14.92. HRMS (*m/z*): $[M\text{-}H^+]^-C_9H_{12}N_4O_5$ calcd 257.0886, found 257.0909.

Crystal structure obtained; see appendix for crystallographic data.

8.11.8.1. Displacement of *S*-cyanovinyl thiolate group by 2-amino-2-cyanoacetamide in the presence of aminooxazoline *arabino*-70

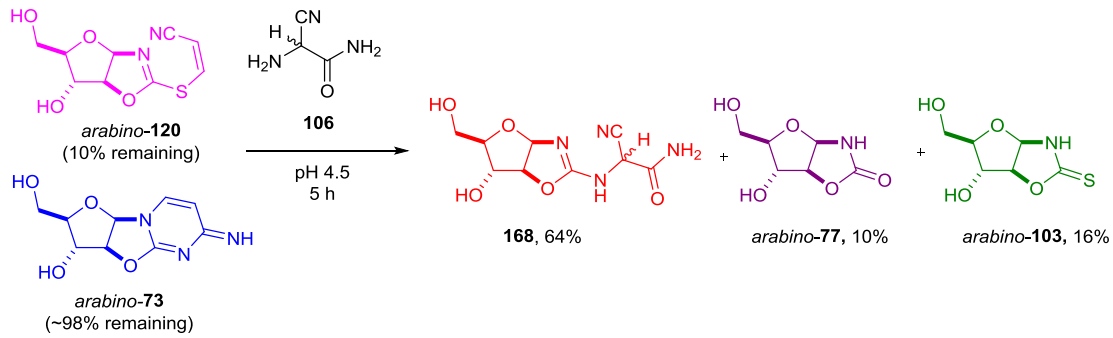
To a solution of aminooxazoline *arabino*-**70** (0.04 mmol, 7.50 mg) and 2-amino-2-cyanoacetamide **106** (0.12 mmol, 13 mg) in H_2O (0.50 mL) and D_2O (0.10 mL) at pH 4.5, was added *S*-cyanovinylated oxazolidinone thione *arabino*-**120** (0.04 mmol, 10.0 mg). The resultant solution was left to stand at rt for 20 h, and the pH was verified every 30 minutes for the first 8 h and then after 20 h. After 20 h the sample was submitted to NMR spectroscopy. The acquired ^{1}H NMR spectrum indicated that 48% aminooxazoline **168** had been obtained alongside the formation of 38% oxazolidionoe thione *arabino*-**103** and 14% oxazoidione *arabino*-**77**, whist aminooxazoline *arabino* **-70** remained unchanged (Scheme 8.6).

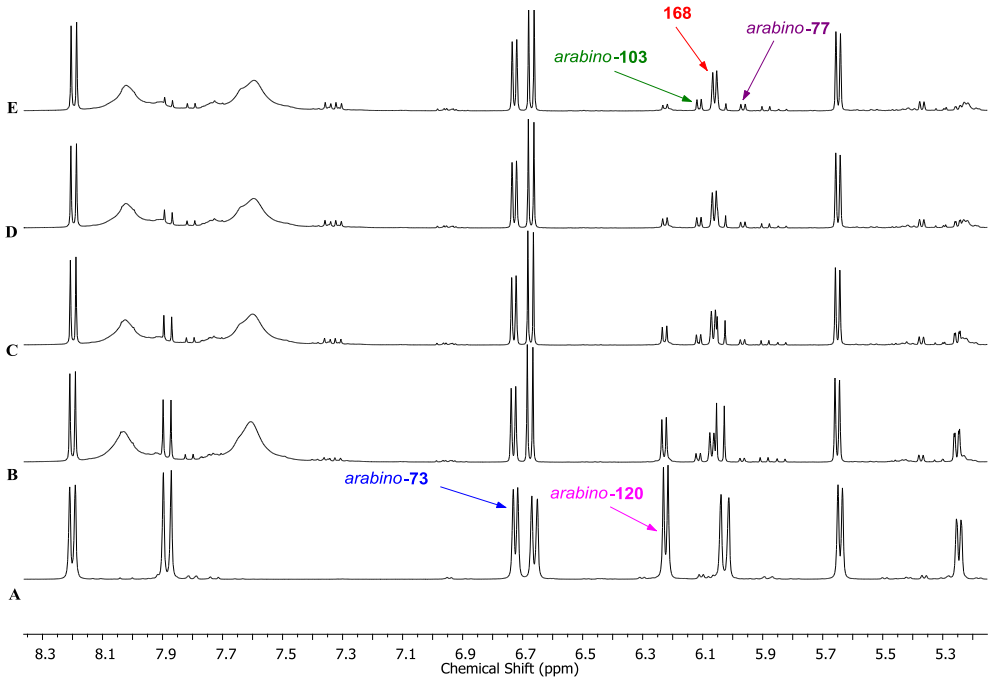


Scheme 8.6: *Incubation of S-cyanovinyl thione arabino-**120** with 2-amino-2-cyanoacetamide **106** in the presence of aminooxazoline arabino-**70** at pH 4.5 and rt for 20 h.*

8.11.8.2. Displacement of *S*-cyanovinyl thiolate group by 2-amino-2-cyanoacetamide in the presence of ancitabine

To a solution of 2-amino-2-cyanoacetamide **106** (0.83 mmol, 110 mg) in H_2O (0.60 mL) at pH 4.5 was added *S*-cyanovinyl oxazolidinone thione *arabino*-**120** (0.08 mmol, 20.0 mg) and ancitabine *arabino*-**73** (0.08 mmol, 20.0 mg) in D_2O (0.20 mL). The resultant solution was left to stand at rt for 5 h, and the pH was checked every hour. The sample was submitted periodically for ^{1}H NMR spectroscopy after 0, 1, 3, 4 and 5 h. Calibration to internal standard indicated that after 5 h 64% aminooxazoline **168**, 16% oxazolidinone thione *arabino*-**103**, 10% oxazolidinone *arabino*-**77** had formed and 10% *arabino*-**120** remaining, whilst ancitabine *arabino*-**73** remained unchangaed (Figure 8.21).





Time (h)	*arabino*-**120** (%)	**168** (%)	*arabino*-**103** (%)	*arabino*-**77** (%)
1	54	36	12	4
3	26	54	14	8
4	16	60	16	10
5	10	64	16	10

Figure 8.21: *Expanded ^{1}H NMR spectra (600 MHz, D_2O, 5.4 – 8.3 ppm) for the time course reaction of 2-amino-2-cyanoacetamide* ***106*** *with S-cyanovinyl oxazolidinone thione arabino-****120*** *in the presence of ancitabine arabino-****73****.* ***A)*** *0 h,* ***B)*** *1 h.* ***C)*** *3 h.* ***D)*** *4 h.* ***E)*** *5 h.*

In order to induce cyclisation in aminooxazoline **168**, to furnish 2,2′-anhydro-5-aminoimidazole-4-carboxamide-*β*-furanosylarabinoside **96** in the presence of ancitabine *arabino*-**73**, the pH of the above reaction mixture was slowly raised to pH 7.0 with NaOH (0.5M). The sample was submitted for NMR spectroscopy using single

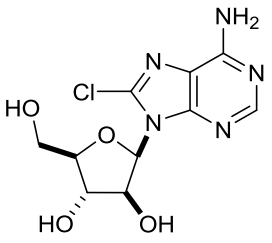
solvent suppression. Calibration to internal standard was used to assess reaction yields at 30 mins, 2 h and then every 6 h for 96 h (Figure 8.22). After 96 h the sample was spiked with authentic compounds to confirm the identity of each species present.

Time (h)	*arabino*-**73**	**185**	**168**	**96**	*arabino*-**103**	*arabino*-**77**	Unknown
0.5	96	0	72	0	20	10	2
2	96	0	72	0	20	10	2
6	96	0	66	4	20	10	4
12	92	4	62	6	20	10	4
18	90	8	54	8	22	10	8
24	90	8	48	10	20	12	12
36	88	12	38	14	20	12	12
48	80	14	36	16	22	12	18
60	68	18	34	20	20	12	24
72	64	22	30	20	20	14	30
84	58	26	28	20	22	14	32
96	50	28	26	22	20	14	36

Figure 8.22: *Tabulated data for cyclisation of aminooxazoline **168** to AICA tricyclic moiety **96** in the presence of ancitabine arabino-**73**. Hydrolysis of arabino-**73** to arabino cytidine **185** was reasonably slow to accommodate cyclisation of **168** to **96** at pH 7. The NMR data obtained was calibrated with inter NMR standard DSS.*

8.12. Conventional synthesis of 8,2`-anhydro-arabino-cycloadenine

9-*β*-Arabinofuranoside-8-chloroadenine **175**

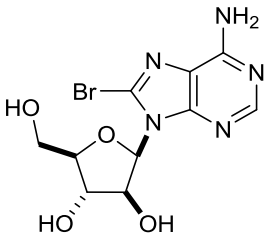


$C_{10}H_{12}ClN_5O_4$
301.69

Following the literature procedure described by Ryu *et al.*,[215] to a solution of arabinose adenine 9-*β*-arabinofuranoside **169** (50.0 mg, 0.19 mmol) in 0.55 M HCl in DMF

(0.58 mL) was added mCPBA **176** (57.0 mg, 0.19 mmol) in DMF (0.20 mL). After 30 min a second batch of mCPBA **176** (27.0 mg, 0.09 mmol) in DMF (0.1 mL) was added. The yellow/orange solution was stirred for 2 h at rt. After 2 h the mixture was concentrated by removing DMF in *vacuo*. Water (3 mL) was added and a white precipitate formed. The precipitate was isolated by filteration and washed with water (5 mL). The filtrate was then dried to yield a gummy residue which was purified by flash column chromatography (MeOH/$CHCl_3$, 5:95). All fractions containing the required product were combined and evaporated to give 9-*β*-arabinofuranoside-8-chloroadenine **175** (20 mg, mmol 0.07, 35%) as a white solid. No characterisation data available in literature. M.p. 225°C (decomp.). IR (Solid cm^{-1}) 3316 (NH_2), 3188 (OH), 2940, 2866 (CH), 1668 (C=C), 1606, 1578 (C=N). ^{1}H NMR (600 MHz, DMSO) δ 8.10 (1H, s, H-(C2)), 7.45 (2H, s, H-(NH_2)), 6.28 (1H, d, *J* = 7.0 Hz, H-(C1`)), 5.65 (1H, d, *J* = 5.6 Hz, H-(2`OH), 5.49 (1H, d, *J* = 5.6 Hz, H-(3`-OH)), 5.27 (1H, dd, *J* = 6.2, 4.4 Hz, H-(5`OH)), 4.37 (1H, td, *J* = 7.0, 5.6 Hz, H-(C2`)), 4.33 (1H, td, *J* = 7.0, 5.6 Hz, H-(C3`)), 3.73 (3H, m, H-(C4`) and H-(C5`)). ^{13}C NMR (151 MHz, DMSO) δ 155.0 (C6), 152.30 (C2), 150.2 (C4), 136.7 (C8), 117.5 (C5), 85.0 (C1`), 82.8 (C4`), 76.4 (C2`), 74.2 (C3`), 60.9 (C5`). UV/Vis: λ_{max} 256 nm. $[\alpha]_D^{20.0}$ (c = 1.00, DMSO) +8.87. HRMS (*m/z*): $[M\text{-}H^+]^-C_{10}H_{12}ClN_5O_4$ calcd 302.0656, found 302.0642.

9-*β*-Arabinofuranoside-8-bromoadenine **179**

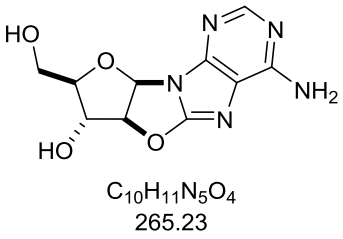


$C_{10}H_{12}BrN_5O_4$
346.14

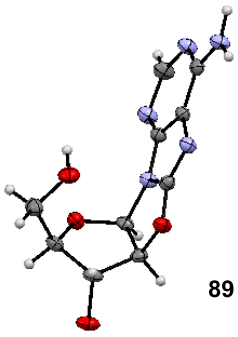
Following the literature procedure described by Ikehara *et al*.,[211] saturated bromine water (5.8 mL in 100 mL H_2O) was added to a solution of arabinose adenine 9-*β*-arabinofuranoside **169** (10.0 g, 34.4 mmol) in sodium acetate solution (1.5 L, 1 M, pH 4.0) at rt. The resultant solution was stirred vigorously for 3 h, during which time the reaction progress was followed by ^{1}H NMR specrroscopy. After 3 h the solution was

decolourised with 5 M sodium bisulfite, and then adjusted to pH 7.0 with 5 M NaOH. The decolourised solution was then concentrated overnight by air-blowing over the solution, after 24 h a precipitate was formed. The resulting crystals were collected by filtration, and then washed with water (3 × 50.0 mL) and acetone (100 mL). The desired product 9-*β*-arabinofuranoside-8-bromoadenine **179** (6.72 g, 19.4 mmol, 56 %) was isolated as white powder. An analytically pure sample was obtained by recrystallization of the product from EtOH-H_2O. M.p. 203ºC (decomp.) (Lit.[265] 202-204°C, decomp). IR (Solid, cm^{-1}) 3375 (NH_2), 3177 (OH), 2866 (CH), 1736 (C=C), 1603 (C=N). ^{1}H NMR (600 MHz, DMSO) 8.07 (1H, s, H-(C2)), 7.46 (2H, br s, H-NH_2), 6.23 (1H, d, *J* = 7.1 Hz, H-(C1`)), 5.64 (1H, d, *J* = 5.8 Hz, H-(2`OH)), 5.49 (1H, d, *J* = 5.7 Hz, H-(3`OH)), 5.35 (1H, dd, *J* = 6.5, 4.4 Hz, H-(5`OH)), 4.45 (1H, td, *J* = 7.1, 5.7 Hz, H-(C3`)), 4.34 (1H, td, *J* = 7.1, 5.8 Hz, H-(C2`)), 3.83 – 3.68 (3H, m, H-(C4`), H-(C5`), H-(C5``)). ^{13}C NMR (151 MHz, DMSO) 155.0 (C6), 152.0 (C2), 150.4 (C4), 126.3 (C8), 119.3 (C5), 86.2 (C1`), 82.8 (C4`), 76.5 (C2`), 74.1 (C3`), 60.9 (C5`). UV/Vis: λ_{max} 267 nm. $[\alpha]_D^{20.0}$ (c = 1.00, DMSO) +6.90. HRMS (*m/z*): [M-H^+]$^-$ $C_{10}H_{12}BrN_5O_4$ calcd 346.0151, found 346.0133.

8,2`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloadenosine **89**



$C_{10}H_{11}N_5O_4$
265.23



89

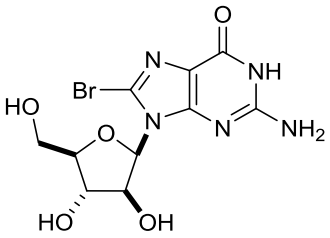
9-*β*-Arabinofuranoside-8-bromoadenine **169** (3.28g, 9.48 mmol) was dissolved in 1M sodium acetate (20 mL). The solution was adjusted to pH 8.0 and stirred for 1 h. The reaction was followed by TLC and single solvent suppressed ^{1}H NMR. A precipitate was observed to form during the course of the reaction and after 1h the resulting crystals were collected by filtration. The precipitate was washed with water (3 × 20 mL) and acetone (20 mL), to give the desired product 8,2`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloadenosine **89** (2.0 g, 7.5 mmol, 80%) as white powder. An analytical sample was obtained by recrystallisation of **89** from hot water. M.p. 205°C

(decomp.). IR (Solid, cm^{-1}) 3297 (NH_2), 3142 (OH), 2962, 2883 (CH), 1667 (C=C), 1623 (C=N). ^{1}H NMR (600 MHz, D_2O) δ 8.16 (1H, s, H-(C2)), 6.71 (1H, d, *J* = 5.5 Hz, H-(C1`)), 5.90 (1H, d, *J* = 5.5 Hz, H-(C2`)), 4.72 (1H, s, H-(C3`)), 4.40 (1H, ddd, *J* = 6.5, 5.4, 4.3 Hz, H-(C4`)), 3.55 (1H, ABX, *J* = 12.7, 4.3 Hz, H-(C5`)), 3.48 (1H, ABX, *J* = 12.7, 5.4 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) δ 161.2 (C8), 154.4 (C5), 151.6 (C2), 146.2 (C4), 121.1 (C6), 99.5 (C2`), 89.6 (C4`), 86.5 (C1`), 75.6 (C3`), 61.4 (C5`). Analysis (% calcd, % found for $C_{10}H_{11}N_5O_4$): C (45.12, 42.48), H (4.18, 4.45), N (26.41, 24.53), O (24.13, 28.14). UV/Vis: λ_{max} 257 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -18.03. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{11}N_5O_4$ calcd 266.0861, found 266.0866.

Crystal structure obtained; see appendix for crystallographic data.

8.13. Conventional synthesis of 8,2`-anhydro-arabino-cycloguanine

9-*β*-Arabinofuranoside-8-bromoguanine **181**

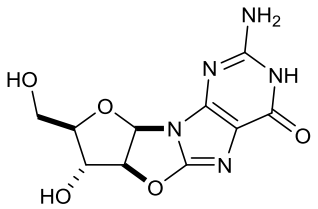


$C_{10}H_{12}BrN_5O_5$
362.14

This method was adapted from a reported literature procedure by Ikehara *et al*.[211] To a solution of arabinose guanine 9-*β*-arabinofuranoside **180** (250.0 mg, 0.88 mmol) dissolved in sodium acetate solution (37.5 mL, 1M, pH 4.0) at rt was added saturated bromine water (145 µL in 2.5 mL H_2O). The resultant solution was vigorously stirred at r.t for 24 h. The progress of the reaction was followed by single solvent suppressed ^{1}H NMR and TLC. The solution was decolourised with 5 M sodium bisulfite and then adjusted to pH 7.0 with 5 M NaOH. The decolourised solution was then concentrated overnight by blowing air over the solution and after 16 h a precipitate had formed.

The resulting crystals were collected by filtration, washed with water (3 × 5.0 mL) and acetone (20 mL), to give the desired product 9-*β*-arabinofuranoside-8-bromoguanine **181** (265 mg, 0.73 mmol, 73%) as white powder. An analytical sample was obtained by recrystallisation of **181** from DMSO-H_2O. M.p. 205°C (decomp.). IR (Solid, cm^{-1}) 3397 (NH_2), 3331 (NH), 3152 (OH), 2913, 2889 (CH), 1672 (C=O), 1651 (C=C), 1599 (C=N). ^{1}H NMR (600 MHz, DMSO) δ 10.77 (1H, s, H-(N1)), 6.45 (2H, s, H-(NH_2)), 6.06 (1H, d, *J* = 6.5 Hz, H-(C1`)), 5.55 (1H, d, *J* = 6.0 Hz, H-(2`OH)), 5.35 (1H, d, *J* = 5.6 Hz, H-(3`OH)), 4.85 (1H, dd, *J* = 5.7, 4.2 Hz, H-(5`OH)), 4.30 (1H, td, *J* = 6.5, 5.6 Hz, H-(C3`)), 4.22 (1H, td, *J* = 6.5.0, 6.0 Hz, H-(C2`)), 3.79 – 3.73 (1H, m, H-(C4`)), 3.72 – 3.64 (2H, m, H-(C5`)). ^{13}C NMR (151 MHz, DMSO) δ 155.5 (C6), 153.2 (C2), 152.2(C4), 120.6 (C8), 117.0 (C5), 85.0 (C1`), 83.0 (C4`), 76.7 (C2`), 75.2 (C3`), 61.7 (C5`). UV/Vis: λ_{max} 250 nm. $[\alpha]_D^{20.0}$ (c = 1.00, DMSO) -4.22. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{10}N_5O_4$ calcd 362.0260, found 362.0223.

8,2`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloguanine **90**



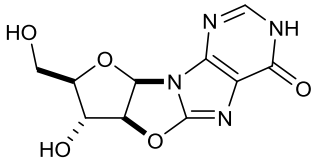
$C_{10}H_{11}N_5O_5$
281.23

9-*β*-Arabinofuranoside-8-bromoguanine **181** (100.0 mg, 0.28 mmol) was dissolved in 1 M sodium acetate (5 mL). The solution was adjusted to pH 8.0 and stirred for 4 h. The progress of reaction was followed by TLC and ^{1}H NMR. After 4 h a white precipitate was observed to form, and the resulting crystals were collected by filtration, washed with water (3 × 5 mL) and acetone (5 mL), to give the desired product 8,2`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloguanine **90** (63.0 mg, 0.22 mmol, 80%) as white powder. An analytical sample was obtained by recrystallisation of **90** from DMSO-H_2O. M.p 285°C (decmp.). IR (Solid, cm^{-1}) 3440 (NH), 3352 (OH), 2914, 2776 (CH), 1659 (C=O), 1605 (C=C), 1570 (C=N). ^{1}H NMR (600 MHz, DMSO) δ 10.62 (1H, s, H-(N1), 6.45 (2H, s, H-(NH_2)), 6.32 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.86 (1H, d, *J* = 4.2 Hz, H-(3`OH)), 5.54 (1H, d, *J* = 5.4 Hz, H-(C2`)), 4.95 (1H, t, *J* = 5.5 Hz, H-(5`OH)), 4.38 (1H, d, *J* = 4.2 Hz H-(C3`)), 4.04 – 3.97 (1H, m,

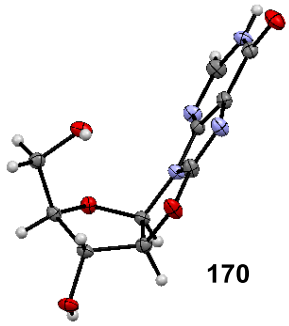
H-(C4`)), 3.20 (1H, ABXY, *J* = 11.4, 5.5, 5.5 Hz, H-(C5`)), 3.10 (1H, ABXY, *J* = 11.4, 6.6, 5.5 Hz, H-(C5``)).^{13}C NMR (151 MHz, DMSO) δ 156.1 (C2), 155.9 (C9), 153.4 (C4), 147.4 (C5), 116.7 (C7), 97.8 (C2`), 88.4 (C4`), 84.9 (C1`), 74.2 (C3`), 60.6 (C5`)**.** Analysis (% calcd, % found for $C_{10}H_{11}N_5O_5$): C (42.71, 40.60), H (3.94, 4.16), N (24.9, 23.50), O (28.46, 31.74). UV/Vis: λ_{max} 255 nm. $[\alpha]_D^{20.0}$ (c = 1.00, DMSO) -14.65. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{11}N_5O_5$ calcd 282.0838, found 282.0839.

8.14. Synthesis of 8,2`-anhydro-arabino-cycloinosine

8,2`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloinosine **170**



$C_{10}H_{10}N_4O_5$
266.21



170

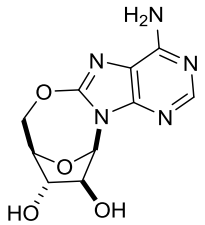
Following a known literature procedure reported by Ikehara *et al.*,[217] 9-*β*-arabinofuranosyl-cycloadenosine **89** (265 mg, 1.0 mmol) was dissolved in 2M AcOH (20 mL), and then sodium nitrite (345 mg, 5 mmol) dissolved in H_2O (5 mL) was added. The resultant mixture was stirred at rt for 24 h, whlist the reaction progress was monitored by TLC. After 24 h the solvent was evaporated, and then repeatedly co-evaporated with EtOH, until no odour of AcOH remained. The residue was recrystallised from H_2O to give desired product 8,2`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloinosine **170** (180 mg, 0.68 mmol, 68%) as pale yellow plates. M.p. 255°C (decmp). IR (Solid, cm^{-1}) 3216 (OH) or (NH), 3076, 2006, 2953 (CH), 1673 (C=O), 1598 (C=C), 1572 (C=N). ^{1}H NMR (600 MHz, D_2O) δ 8.17 (H, s, H-(C7)), 6.75 (1H, d, *J* = 5.5 Hz, H-(C1`)), 5.91 (1H, d, *J* = 5.5 Hz, H-(C2`)), 4.72 (1H, s, H-(C3`)), 4.42 (1H, ddd, *J* =5.5, 5.5, 4.1 Hz, H-(C4`)), 3.56 (1H, ABX, *J* = 12.7, 4.1 Hz, H-(C5`)), 3.49 (1H, ABX, *J* = 12.7, 5.5 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) δ 160.7 (C2), 158.4 (C9), 145.8 (C4), 145.6 (C7), 126.0 (C5), 99.5 (C2`), 89.9 (C4`), 86.9 (C1`), 75.6 (C3`), 61.4 (C5`). UV/Vis: λ_{max} 250 nm. Analysis (% calcd, % found for $C_{10}H_{10}N_4O_5$): C (45.12, 44.12), H (3.79, 3.70), N (21.05, 20.56), O (30.05,

31.62). $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -16.30. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{10}N_5O_4$ calcd 267.0729, found 267.0723.

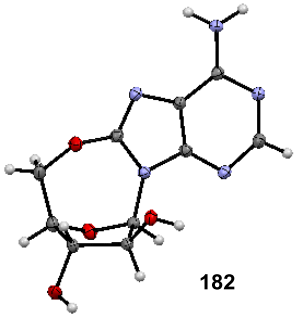
Crystal structure obtained; see appendix for crystallographic data.

8.16. Isomerisation of anhydronucleoside

8,5`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloadenine **182**



$C_{10}H_{11}N_5O_4$
265.23

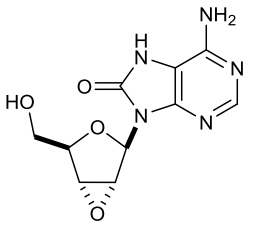


182

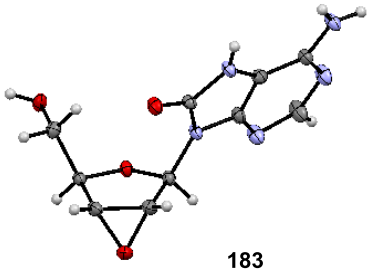
Following a known literature procedure reported by Ikehara and Ogiso,[266] 8,2`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloadenosine **89** (265 mg, 1.0 mmol) was heated in 0.01 M NaOH (30 mL) at 60°C for 3 h then neutralised with 0.1 M HCl and evaporated in vacuo to approximately 5 mL, precipitating crystalline material. The crystals were collected by filtration and washed with MeOH, which dissolved the starting material. The residual crystals were recrystallised from water to give the desired product 8,5`-*O*-anhydro-9-*β*-arabinofuranosyl-cycloadenine **182** (29 mg, 0.11 mmol, 11%) as white prisms. M.p. 236°C (decomp.). IR (Solid, $cm^{-1}$) 3563 (NH), 3420 (OH), 2961, 2916 (CH), 1620 (C=C), 1576 (C=N). $^1H$ NMR (600 MHz, D6-DMSO) 8.06 (1H, s, H-(C2)), 6.30 (1H, d, *J* = 6.4 Hz, H-(C1`)), 4.53 (1H, ABX, *J* = 13.0, 2.6 Hz, H-(C5`)), 4.38 (1H, d, *J* = 6.4 Hz, H-(C2`)), 4.35 (1H, d, *J* = 2.6 Hz, H-(C4`)), 4.18 (1H, s, H-(C3`)), 4.12 (1H, ABX, *J* = 13.0 Hz, H-(C5``)). ^{13}C NMR (151 MHz, DMSO) 155.0 (C8), 153.9 (C6), 151.4 (C2), 148.4 (C4), 114.2 (C5), 86.6 (C4`), 83.8 (C2`), 79.8 (C1`), 77.4 (C3`), 74.7 (C5`). UV/Vis: λ_{max} 260 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -9.52. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{11}N_5O_4$ calcd 266.0889, found 266.0898.

Crystal structure obtained; see appendix for crystallographic data.

2`,3`-epoxy-9-*β*-arabinofuranoside-8-oxoadenine **183**



$C_{10}H_{11}N_5O_4$
265.23



183

This method was adapted from a reported literature procedure by Rees *et al*.,[222] 9-*β*-arabinofuranoside-8-bromoadenine **179** (1.0 g, 2.89 mmol) was dissolved in NaOH (1 M, 10 mL). The resultant mixture was then stirred for 2 h at r.t, whilst the reaction progress was monitored by TLC and ^{1}H NMR. After 2 h neutralised silica (6.0 g in 1.50 mL aqueous ammonia) was added and the slurry was concentrated in *vacuo* to yield a fine free flowing powder. This silica was the applied to the top of a silica gel column. The column was then eluted with ethyl acetate and methanol (95:5) to yield desired product 2`,3`-epoxy-9-*β*-arabinofuranoside-8-oxoadenine **183** (260 mg, 0.98 mmol, 34%) as white solids. An analytical sample of **183** was obtained by recrystallisation from hot water. M.p 202°C (decomp.). IR (Solid, cm^{-1}) 3323 (NH_2), 3194 (OH), 1715 (C=O), 1651 (C=C), 1595 (C=N). ^{1}H NMR (600 MHz, D2O) 8.13 (1H, s, H-(C2)), 6.17 (1H, s, H-(C1`)), 4.56 (1H, d, *J* = 2.7 Hz, H-(C2`)), 4.40 (1H, dd, *J* = 7.7, 5.0 Hz, H-(C4`)), 4.27 (1H, d, *J* = 2.7 Hz, H-(C3`)), 3.82 (1H, ABX, *J* = 12.1, 5.0 Hz, H-(C5`)), 3.72 (1H, ABX, *J* = 12.1, 7.7 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D2O) 153.5 (C8), 151.8 (C2), 148.2 (C6), 147.1 (C4), 105.3 (C5), 81.5 (C4`), 81.3 (C1`), 61.4 (C5`), 59.9 (C3`), 58.5 (C2`). UV/Vis: λ_{max} 301 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -7.89. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{11}N_5O_4$ calcd 266.0895, found 266.0866.

Crystal structure obtained; see appendix for crystallographic data.

8.17. Prebiotic phosphorylation of nucleosides

8.17.1. General methods for prebiotic phosphorylation

Method **A:**

The nucleoside (0.06 mmol), ammonium dihydrogen phosphate (0.06 mmol) and urea (0.6 mmol) were dissolved in H_2O (0.6 mL) by warming. The resultant suspension was evenly spread on glass fibre disc (2 × 25mm × 0.68 mm). The discs were dried at 40°C for 48 h and then heated at 100°C for 24 h. The discs were then washed with H_2O (10 mL) and D_2O (10 mL), the washings were then lyophilised to yield a brown solid. The brown solids were dissolved in D_2O (1 mL) and then the solids were removed by centrifuge. The resultant solution was left overnight and some precipitate formed which was further removed by centrifuge. The solution obtained was then submitted for ^{1}H and ^{31}P NMR analysis. Yields were calculated based on comparison with an internal standard (DSS) and nucleotide phosphates were purified using HPLC.

Method **B:**

The nucleoside (0.06 mmol), ammonium dihydrogen phosphate (0.06 mmol) and urea (0.6) were dissolved in H_2O (0.5 mL) and formamide (0.1 mL) by warming. The resultant suspension was evenly spread on glass fibre disc (2 × 25 mm × 0.68 mm). The discs were dried at 40°C for 48 h and then heated at 100°C for 24 h. The discs were then washed with H_2O (10 mL) and D_2O (10 mL), the washings were then lyophilised to yield a brown solid. The brown solids were dissolved in D_2O (1 mL) and then the solids were removed by centrifuge. The resultant solution was left overnight and some precipitate formed which was further removed by centrifuge. The solution obtained was then submitted for ^{1}H and ^{31}P NMR analysis. Yields were calculated based on comparison with an internal standard (DSS) and nucleotide phosphates were purified using HPLC.

Method **C:**

The nucleoside (0.06 mmol), and a mixture of salts containing dihydrogen phosphate/ammonium chloride/ammonium bicarbonate/urea (1:5:10:10 equiv.) were dissolved in H_2O (0.6 mL). The resultant solution was evenly applied to glass fibre

discs (2 × 25 mm × 0.68 mm). The discs were dried at 40°C for 48 h and then heated at 100°C for 24 h. The disc was then washed with H_2O (10 mL) and D_2O (10 mL), the washings were then lyophilised to yield a brown solid. The brown solid was dissolved in D_2O (1 mL) and then the solids were removed by centrifuge. The resultant solution was left overnight and some precipitate formed which was further removed by centrifuge. The solution obtained was then submitted for ^{1}H and ^{31}P NMR analysis. Yields were calculated based on comparison with an internal standard (DSS) and nucleotide phosphates were purified using HPLC.

Method **D**:

Nucleoside (0.06 mmol), ammonium dihydrogen phosphate (0.06 mmol) and urea (1.6 mmol) were thoroughly mixed and heat at 140°C for 20 min. The reaction mixture was then cooled to rt, dissolved in D_2O (3 mL) and lyophilised. The residue was thrice dissolved in D_2O (3 mL) and lyophilised. The final lyophilite was then dissolved in D_2O (0.5 mL) and NMR spectra were acquired. Yields were calculated based on comparison with an internal standard (DSS) and nucleotide phosphates were purified using HPLC.

Method **E**:

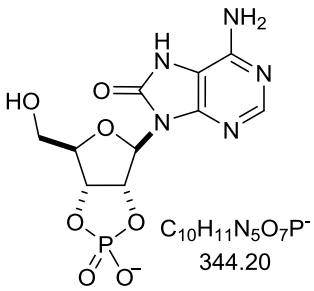
Nucleoside (0.06 mmol), ammonium dihydrogen phosphate (0.06 mmol) and urea (0.6 mmol) were suspended in formamide (0.6 mL). The reaction mixture was heated at 100°C for 72 h. The reaction was diluted with D_2O (3 mL) and lyophilised for 3 d (to remove formamide), diluted with D_2O (3 mL) and further lyophilised. The lyophilite was dissolved in D_2O (0.5 mL) and NMR spectra were acquired. Yields were calculated based on comparison with an internal standard (DSS).

Method **F**:

Nucleoside (0.03 mmol), ancitabine **11** (0.03 mmol), ammonium dihydrogen phosphate (0.06 mmol) and urea (0.6 mmol) were suspended in formamide (0.6 mL). The reaction mixture was heated at 100°C for 72 h. The reaction was diluted with D_2O (3 mL) and lyophilised for 3 d (to remove formamide), diluted with D_2O (3 mL) and further lyophilised. The lyophilite was dissolved in D_2O (0.5 mL) and NMR spectra were acquired. Yields were calculated based on comparison with an internal standard (DSS).

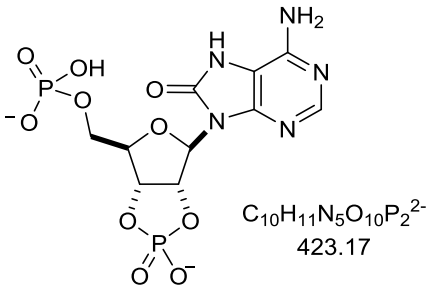
8.17.2. Synthesis of *β*-8-oxo- 2`,3`-cyclic phosphate purine ribonucleotides from anhydro purines.

β-Ribofuranosyl-8-oxo-adenine-2`,3`-cyclic phosphate **93**



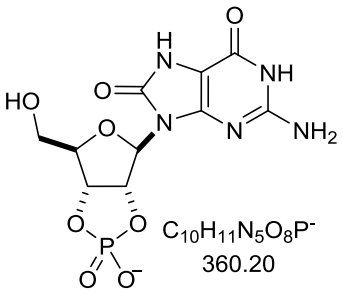
Following method **E** of phosphorylation: 8,2`-anhydro-adenosine **89** (32.0 mg, 0.12 mmol), ammonium dihydrogen phosphate (13.5 mg, 0.12 mmol) and urea (72.0 mg, 1.2 mmol) were suspended in formamide (1.2 mL). The reaction mixture was heated at 100°C for 48 h. The pale yellow solution was allowed to cool, diluted with H_2O (1.8 mL) to make a 3 mL solution, which was then purified by HPLC with a gradient of 100 mM triethylammonium formate at pH 5 and acetonitrile from 80:20 at a flow rate of 5 ml/min. The product was detected at λ = 250 nm. The fractions containing the desired product were combined, and lyophilised to yield *β*-ribofuranosyl-8-oxo-adenine-2`,3`-cyclic phosphate **93** (8.2 mg, 0.024 mmol, 20%) as white solids. M.p. 220°C (decomp.). IR (Solid, cm^{-1}) 1711 (N-C=N), 1649 (N-C=O), 1060 (PO_4). ^{1}H NMR (600 MHz, D_2O) 8.10 (1H, s, H(C2)), 6.11 (1H, d, *J* = 3.4 Hz, H-(C1`)), 5.61 (1H, ddd, *J* = 7.8, 7.0, 3.4 Hz, H-(C2`)), 5.14 (1H, ddd, *J* = 10.9, 7.0, 4.9 Hz, H-(C3`)), 4.40 (1H, ddd, *J* = 5.5, 4.9, 3.3 Hz, H-(C4`)), 3.93 (1H, ABX, *J* = 12.3, 3.6 Hz, H-(C5`)), 3.85 (1H, ABX, *J* = 12.3, 5.5 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 153.3 (C8), 151.8 (C2), 148.3 (C6), 146.7 (C4), 105.2 (C5), 87.0 (d, *J* = 5.5 Hz, (C1`)), 85.6 (d, *J* = 2.2 Hz, (C4`)), 79.8 (d, *J* = 2.2 Hz, (C2`)), 78.3 (C3`), 62.0 (C5`). ^{31}P NMR (162 MHz, D_2O) δ 20.37 (dd, *J* = 10.5, 7.5 Hz). UV/Vis: λ_{max} 273 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -7.32. HRMS (*m/z*): $[M-H^+]^-$ $C_{10}H_{11}N_5O_7P^-$ calcd 344.0396, found 344.0390.

β-Ribofuranosyl-8-oxo-adenine-2`,3`- cyclic-bis-5`-phosphate **188**



The above procedure also yielded *β*-ribofuranosyl-8-oxo-adenine-2`,3`- cyclic-bis-5`-phosphate **188** (4.1 mg, 0.010 mol, 8%) as white solid. M.p. 235°C (decomp.). IR (Solid, $cm^{-1}$) 1707 (N-C=N), 1655 (N-C=O), 1038 ($PO_4$). $^{1}H$ NMR (600 MHz, $D_2O$) 8.12 (1H, s, H-(C2)), 6.11 (1H, d, *J* = 2.9 Hz, H-(C1`)), 5.69 (1H, td, *J* = 6.3, 2.9, H-(C2`)), 5.19 (1H, ddd, *J* = 12.0, 6.3, 5.8, H-(C3`)), 4.44 (1H, ddd, *J* = 6.2, 5.8, 5.0 Hz, H-(C4`)), 4.14 (1H, ABXY, *J* = 11.4, 5.0, 5.0 Hz, H-(C5`)), 4.06 (1H, ABXY, *J* = 11.4, 6.2, 6.2 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 153.2 (C8), 151.9 (C2), 148.2 (C6), 146.8 (C4), 105.1 (C5), 86.7 (d, *J* = 6.6 Hz, (C1`)), 84.4 (d, *J* = 7.2 Hz, (C4`)), 79.4 (d, *J* = 2.8 Hz, (C2`)), 77.9 (C3`)), 64.5 (d, *J* = 3.9 Hz, (C5`)).^{31}P {decoupled} NMR (162 MHz, D_2O) δ 20.57, 0.85 (broad). ^{31}P NMR (162 MHz, D_2O) δ 20.58 (dd, *J* = 12.3, 5.5 Hz), 0.85 (s, broad). UV/Vis: λ_{max} 265 nm. HRMS (*m/z*): $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) - 8.17. $[M\text{-}H^+]^-$ $C_{10}H_{11}N_5O_{10}P_2^{2-}$ calcd 424.0060, found 424.0056.

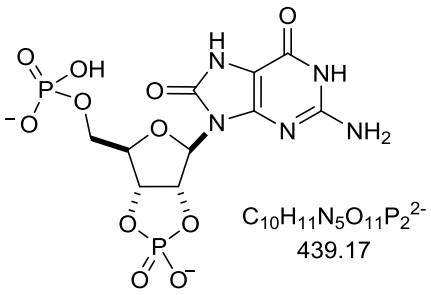
β-Ribofuranosyl-8-oxo-guanine-2`,3`-cyclic phosphate **94**



Following method **E** of phosphorylation: 8,2`-anhydro-guanosine **90** (16.0 mg, 0.056 mmol), ammonium dihydrogen phosphate (6.5 mg, 0.056 mmol) and urea (90.0 mg, 1.50 mmol) were thoroughly mixed together in a small test tube. The mixture was then pushed down the test making sure that there were no solids on the walls of the container. The test tube was then lowered in an oil bath (140°C) and left standing for

20 min. After this time the reaction mixture was allowed to cool, dissolved in H_2O (2 mL) and the pale yellow solution obtained was purified via HPLC with a gradient of 100 mM triethylammonium formate at pH 5 and acetonitrile from 80:20 at a flow rate of 5 ml/min. The desired product was detected at λ = 250 nm. The fractions containing the desired product were combined, and lyophilised to yield *β*-ribofuranosyl-8-oxo-guanine-2`,3`-cyclic phosphate **94** (6.15 mg, 0.017 mmol, 30%) as white solids. M.p. 295°C (decomp.). IR (Solid, cm^{-1}) 3557 (NH), 3312 (OH), 1720 (C=O), 1638 (C=C), 1609 (C=N), 1053 (PO_4). ^{1}H NMR (600 MHz, D_2O) 6.03 (1H, d, *J* = 2.9 Hz, H-(C1`)), 5.59 (1H, ddd, *J* = 7.0, 6.1, 2.9 Hz, H-(C2`)), 5.16 (1H, ddd, *J* = 12.0, 6.1, 4.8 Hz, H-(C3`)), 4.34 (1H, ddd, *J* = 6.0, 4.8, 4.0 Hz, H-(C4`)), 3.91 (1H, ABX, *J* = 12.4, 4.0 Hz, H-(C5`)), 3.84 (1H, ABX, *J* = 12.4, 6.0 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) δ 154.2 (C6), 153.8 (C5), 153.2 (C8), 148.7 (C4), 100.3 (C2), 86.7 (d, *J* = 7.1 Hz (C1`)), 85.6 (C4`), 80.2 (d, *J* = 2.9 Hz (C2`)), 78.3 (C3`), 61.9 (C5`). ^{31}P {decoupled} NMR (162 MHz, D_2O) δ 20.44. ^{31}P NMR (162 MHz, D_2O) δ 20.4 (m). UV/Vis: λ_{max} 300 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -7.46. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{11}N_5O_8P^-$ calcd 362.0496, found 362.0501.

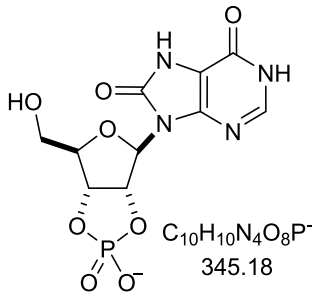
<u>*β*-Ribofuranosyl-8-oxo-guanine-2`,3`- cyclic-bis-5`-phosphate **193**</u>



The above procedure also yielded *β*-ribofuranosyl-8-oxo-guanine-2`,3`- cyclic-bis-5`-phosphate **193** (2.5 mg, 0.006 mol, 10%) as white solid. M.p. 260°C (decomp.). IR (Solid, $cm^{-1}$) 3337 (NH), 3242 (OH), 2918, 2851 (CH), 1659 (C=O), 1603 (C=C), 1553 (C=N), 1047 ($PO_4$). $^1$H NMR (400 MHz, $D_2O$) δ 6.00 1H NMR (400 MHz, D2O) 6.00 (1H, d, *J* = 2.5 Hz, H-(C1`)), 5.57 (1H, td, *J* = 6.3, 2.5 Hz, H-(C2`)), 5.26 (1H, ddd, *J* = 12.0, 6.3, 5.8 Hz, H-(C3`)), 4.38 (1H, ddd, *J* = 6.0, 5.8, 5.3 Hz, H-(C4`)), 4.11 (1H, ABXY, *J* = 11.3, 5.3, 5.3 Hz, H-(C5`)), 4.01 (1H, ABXY, *J* = 11.3, 6.0, 6.0 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) δ 154.3 (C6), 153.9 (C5), 153.3 (C8), 148.9 (C4), 100.3 (C2), 86.7 (d, *J* = 6.8 Hz (C1`)), 84.6 (d, *J* = 8.1 Hz (C4`)), 80.1 (d, *J* =

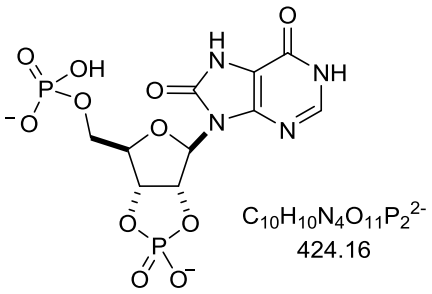
3.0 Hz, (C2`)), 78.3 (C3`), 64.6 (d, *J* = 4.5 Hz, (C5`)). ^{31}P {decoupled}NMR (162 MHz, D_2O) δ 20.51, 1.78 (broad). ^{31}P NMR (162 MHz, D_2O) δ 20.51 (m), 1.78 (m, broad). UV/Vis: λ_{max} 290 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -6.20. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{11}N_5O_{11}P_2^{2-}$ calcd 442.0160, found 442.0159.

β-Ribofuranosyl-8-oxo-inosine-2`,3`-cyclic phosphate **187**



Following method **D** of phosphorylation: 8,2`-anhydro-inosine **170** (32.0 mg, 0.12 mmol), ammonium dihydrogen phosphate (13.5 mg, 0.12 mmol) and urea (72.0 mg, 1.2 mmol) were suspended in formamide (1.2 mL). The reaction mixture was heated at 100°C for 24 h. The colourless solution was allowed to cool, then diluted with H_2O (1.8 mL) to make a 3 mL solution, which was purified via HPLC with a gradient of 100 mM triethylammonium formate at pH 5 and acetonitrile from 100:0 to 80:20 at a flow rate of 5 ml/min. The desired product was detected at λ = 250 nm. The fractions containing the desired product were combined, and lyophilised to yield *β*-ribofuranosyl-8-oxo-inosine-2`,3`-cyclic phosphate **187** (16.6 mg, 0.048 mmol, 40%) as white solids. M.p. 310°C (decomp.). IR (Solid, cm^{-1}) 3404 (NH), 3183 (OH), 2967 (CH), 1724 (C=O), 1674 (C=C), 1661 (C=N), 1063 (PO_4). ^{1}H NMR (600 MHz, D_2O) 8.14 (1H, s, H-(C2)), 6.16 (1H, d, *J* = 2.9 Hz, H-(C1`)), 5.62 (1H, td, *J* = 6.6, 2.9 Hz, H-(C2`)), 5.17 (1H, ddd, *J* = 12.1, 6.6, 5.4 Hz, H-(C3`)), 4.38 (1H, ddd, *J* = 5.9, 5.4, 3.9 Hz, H-(C4`)), 3.92 (1H, ABX, *J* = 12.4, 3.9 Hz, H-(C5`)), 3.85 (1H, ABX, *J* = 12.4, 5.9 Hz, H-(5``)). ^{13}C NMR (151 MHz, D_2O) 153.3 (C6), 153.2 (C8), 145.9 (C2), 145.7 (C4), 109.8 (C5), 87.0 (d, *J* = 6.6 Hz, C1`), 85.7 (d, *J* = 1.7 Hz, C4`), 80.2 (d, *J* = 2.2 Hz, C2`), 78.2 (C3`), 61.8 (C5`). ^{31}P {decoupled} NMR (162 MHz, D_2O) δ 20.44. ^{31}P NMR (162 MHz, D_2O) δ 20.44 (m). UV/Vis: λ_{max} 261 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -6.48. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{10}N_4O_8P^-$ calcd 346.0315, found 346.0318.

β-Ribofuranosyl-8-oxo-inosine-2`,3`- cyclic-bis-5`-phosphate **194**



The above procedure also yielded *β*-ribofuranosyl-8-oxo-inosine-2`,3`- cyclic-bis-5`-phosphate **194** (5.9 mg, 0.014 mol, 12%) as white solid. M.P. 275°C (decomp.). IR (Solid $cm^{-1}$) 3397 (NH), 3225 (OH), 2924, 2853 (CH), 1680 (C=O), 1059 ($PO_4$). $^1H$ NMR (600 MHz, $D_2O$) 8.15 (1H, s, H-(C2)), 6.16 (1H, d, *J* = 2.8 Hz, H-(C1`)), 5.67 (1H, td, *J* = 6.4, 2.8 Hz, H-(C2`)), 5.22 (1H, ddd, *J* = 12.6, 6.4, 5.7, H-(C3`)), 4.45 (1H, ddd, *J* = 6.1, 5.7, 4.9 Hz, H-(C4`)), 4.10 (1H, ABXY, *J* = 11.3, 4.9, 4.9 Hz, H-(C5`)), 4.03 (1H, ABXY, *J* = 11.3, 6.1, 6.1 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) 153.5 (C6), 153.2 (C8), 146.1 (C2), 145.8 (C4), 109.8 (C5), 86.8 (d, *J* = 6.6 Hz, C1`), 84.7 (d, *J* = 8.3 Hz, C4`), 79.9 (d, *J* = 2.8 Hz, C2`), 78.2 (C3`), 64.4 (d, *J* = 4.4 Hz, C5`). ^{31}P {decoupled} NMR (162 MHz, D_2O) δ 20.58, 2.72 (broad). ^{31}P NMR (162 MHz, D_2O) δ 20.58 (dd, *J* = 12.1, 4.1 Hz), 2.72 (m, broad). UV/Vis: λ_{max} 256 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -4.79. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{10}N_4O_{11}P_2^{2-}$ calcd 424.9900, found 424.9904.

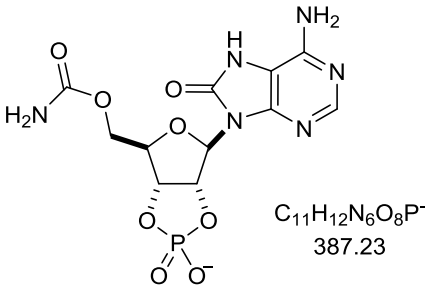
8.17.3. Phosphorylation of 8,2`-anhydro adenosine using method B

8,2`-Anhydro-adenosine **89** (15.0 mg, 0.06 mmol), ammonium dihydrogen phosphate (6.6 mg, 0.06 mmol) and urea (34.0 mg, 0.6) were dissolved in H_2O (0.5 mL) and formamide (0.1 mL) by warming. The resultant suspension was evenly spread on glass fibre disc (2 × 25 mm × 0.68 mm). The discs were dried at 40°C for 48 h and then heated at 100°C for 48 h. The discs were then washed with H_2O (10 mL) and D_2O (10 mL), the washings were then lyophilised to yield a brown solid. The brown solids were dissolved in D_2O (1 mL) and then the residual solids were removed by centrifugation. The resultant solution was left overnight and some precipitate formed which was further removed by centrifuge. The orange solution obtained was diluted with H_2O (1 mL) to make 2 mL solution, which was purified via HPLC with a gradient

of 100 mM triethylammonium formate at pH 5 and acetonitrile from 100:0 to 80:20 at a flow rate of 5 ml/min. The desired products was detected at λ = 250 nm, to yield the following 2`,3`-cyclic phosphate compounds.

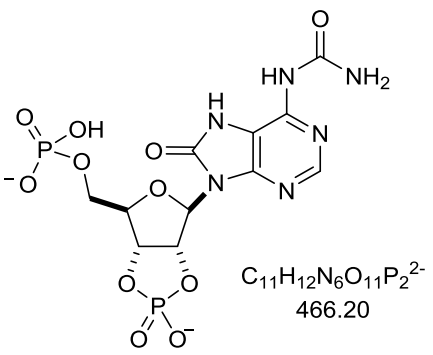
2`, 3`-cyclic compound	HPLC yield (%)
Adenine-2`, 3`-cyclic phosphate **93**	12
Adenine-2`, 3`-cyclic-5`-bisphosphate **188**	20
Adenine-5`-carbamoyl-2`, 3`-cyclic phosphate **189**	16
Adenine-*N*9-carbamoyl-2`, 3`-cyclic-5`-bisphosphate **190**	39
Adenine-5`,*N*9-dicarbamoyl-2`, 3`-cyclic phosphate **191**	11

β- Ribofuranosyl adenine-5`-carbamoyl-2`, 3`-cyclic phosphate **189**



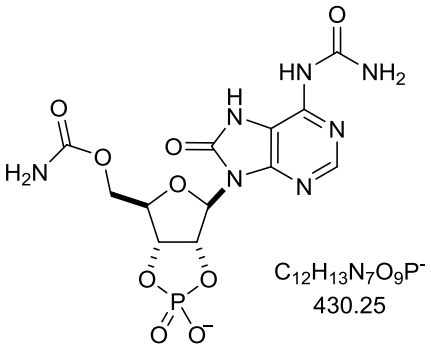
^{1}H NMR (600 MHz, D_2O) 8.13 (s, 1H, H-(C2)), 6.13 (1H, d, J = 2.4 Hz, H-(C1`)), 5.72 (1H, ddd, $J$ = 7.0, 6.6, 2.4 Hz, H-(C2`)), 5.24 (1H, ddd, J = 12.3, 6.6, 5.9 Hz, H-(C3`)), 4.48 (1H, ddd, $J$ = 6.4, 5.9, 3.6 Hz, H-(C4`)), 4.41 (1H, ABX, J = 12.0, 3.6 Hz, H-(C5`)), 4.33 (1H, ABX, J = 12.0, 6.4 Hz, H-(C5``)). ^{31}P {decoupled} NMR (162 MHz, D_2O) δ 22.52. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{11}H_{12}N_6O_8P^-$ calcd 387.1036, found 387.1031.

β-Ribofuranosyl adenine-*N*9-carbamoyl-2`, 3`-cyclic-5`-bisphosphate **190**



^{1}H NMR (600 MHz, D_2O) δ 8.40 (1H, s, (H-(C2)), 6.17 (1H, d, *J* = 2.8 Hz, H-(C1`)), 5.73 (1H, td, *J* = 6.6, 2.8 Hz, H-(C2`)), 5.19 (1H, ddd, *J* = 12.0, 6.6, 5.8 Hz, H-(C3`)), 4.42 (1H, dt, *J* = 5.8, 5.0 Hz, H-(C4`)), 4.07 (1H, ABXY, *J* = 11.5, 5.0, 5.0 Hz, H-(C5`)), 3.99 (1H, ABXY, *J* = 11.5, 5.8, 5.8 Hz, H-(C5``)). ^{31}P {decoupled} NMR (162 MHz, D_2O) δ 22.89, 6.20 (broad). HRMS (*m/z*): $[M-H^+]^-$ $C_{11}H_{12}N_6O_{11}P_2^{2-}$ calcd 466.8659, found 466.8668.

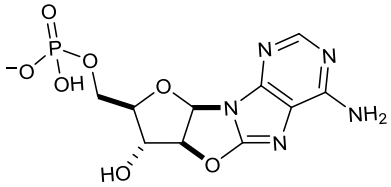
β-ribofuranosyl adenine-5`,9-dicarbamoyl-2`, 3`-cyclic phosphate



^{1}H NMR (600 MHz, D_2O) δ 8.33 (1H, s, H-(C2)), 6.17 (1H, d, *J* = 2.9 Hz, (C1`)), 5.74 (1H, td, *J* = 6.1, 2.9 Hz, H-(C2`)), 5.27 (1H, ddd, *J* = 12.5, 6.1, 4.9 Hz, H-(C3`), 4.49 (1H, ddd, *J* = 6.3, 4.9, 3.3 Hz, H-(C4`)), 4.42 (ABX, 1H, *J* = 12.0, 3.3 Hz, H-(C5`)), 4.33 (ABX, 1H, *J* = 12.0, 6.3 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) δ 159.5 ($CONH_2NH$), 158.0 ($CONH_2$), 150.3 (C8), 149.1 (C5), 141.5 (C4), 86.9 (d, *J* = 6.7 Hz (C1`), 83.3 (C4`), 79.8 (C2`), 78.0 (C3`), 64.2 (C5`). ^{31}P {decoupled} NMR (162 MHz, D_2O) δ 22.62. HRMS (*m/z*): $[M-H^+]^-$ $C_{12}H_{13}N_7O_9P^-$ calcd 430.0011, found 430.0009.

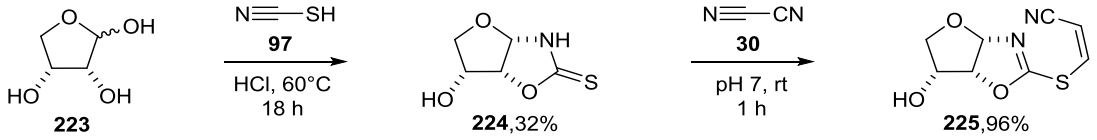
8.18. Synthesis of 8,2`-anhydro adenosine 5`-monophosphate

8,2`-*O*-Anhydro-9-*β*-arabinofuranosyl-cycloadenosine-5`-monophosphate **200**

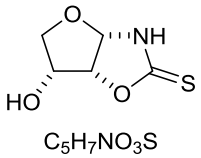


This method was adapted from a known literature procedure reported by Ikehara *et al.*[211] To a solution of 9-*β*-arabinofuranoside-adenosine-5`-monophosphate **198** (573 mg, 1.67 mmol) dissolved in sodium acetate solution (100 mL, 1 M, pH 4.0) at rt was added saturated bromine water (170 µL in 10 mL H_2O). The resultant solution was vigorously stirred at rt for 24 h. The progress of the reaction was followed by ^1H NMR spectroscopy and TLC. After 24 h the solution was decolourised with 5M sodium bisulfite, and then adjusted to pH 8.0 with 5M NaOH. The decolourised solution was then concentrated overnight by blowing air over the solution. After 16 h a precipitate had formed. The resulting crystals were collected by filtration, washed with water (3 × 5.0 mL) and then thrice recrystallized from hot H_2O, to give the desired product 8,2`-*O*-Anhydro-9-*β*-arabinofuranosyl-cycloadenosine-5`-monophosphate **200** (365 mg, 1.1 mmol, 63%) as colourless glass sheets. M.p. 285°C (decomp.). ^1H NMR (600 MHz, D_2O) δ 8.18 (1H, s, H-(C2)), 6.66 (1H, d, *J* = 5.5 Hz, H-(C1`)), 5.92 (1H, dd, *J* = 5.5, 1.5 Hz, H-(C2`)), 4.80 (1H, dd, *J* = 3.6, 1.5 Hz, H(C3`)), 4.42 (1H, td, *J* = 6.0, 3.6 Hz, H-(C4`)), 3.69 (1H, ABXY, *J* = 12.0, 6.0, 6.0 Hz, H-(C5`)), 3.61 (1H, ABXY, *J* = 12.0, 6.0, 6.0 Hz, H-(C5``)). ^{13}C NMR (151 MHz, D_2O) δ 160.9 (C8), 154.6 (C6), 151.7 (C2), 146.3 (C4), 121.2 (C5), 99.39 (C2`), 86.7 (d, *J* = 7.8 Hz, (C4`)), 85.73 (C1`), 75.2 (C3`), 63.27 (d, *J* = 4.4 Hz, (C5`)). ^{31}P NMR (162 MHz, D_2O) δ 3.15 (t, *J* = 6.2 Hz). ^{31}P {decoupled} NMR (162 MHz, D_2O) δ 3.15. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) -5.30. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_{10}H_{12}N_5O_7P$+H calcd 346.0547, found 346.0545.

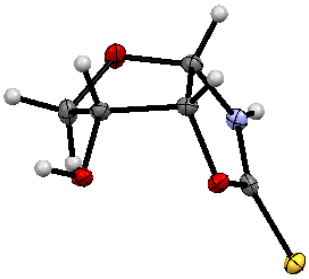
8.19. Synthesis of *S*-cyanovinylated erythrose furanosyl oxazolidinone thione



Erythrosefuranosyl oxazolidinone thione **224**



$C_5H_7NO_3S$
161.18

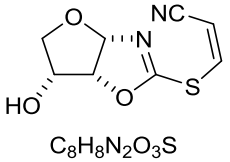


224

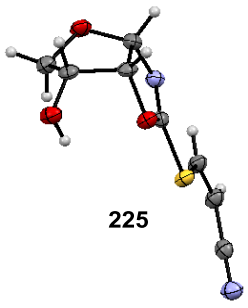
Erythrose **223** (150.0 mg, 1.30 mmol) and potassium thiocyanate **97** (248.0 mg, 2.6 mmol) were dissolved in water (2 mL) and then HCl (37%, 200 µL) was added dropwise at 0°C. The pink solution was kept at rt for 2 h and then heated to 60°C for 18 h. After this time the solution was cooled to rt and mixed with silica gel (1.0 g) and then the solvent was evaporated to obtain a fine free flowing powder. The powder was then loaded on a silica gel column (10.0 g) and eluted with ethyl acetate and chloroform (1:1). Fractions containing the product were concentrated to give erythrosefuranosyl oxazolidinone thione **224** (67 mg, 0.41 mmol, 32%) as crystalline white solid. M.p. 160-163°C. IR (Solid, cm^{-1}) 3298 (NH), 3167 (OH), 2984, 2864 (CH), 1519 (C=S). ^1H (600 MHz, D_2O) 5.94 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.30 (1H, t, *J* = 5.4 Hz, H-(C2`)), 4.51 (1H, dt, *J* = 9.5, 5.4 Hz, H-(C3`)), 4.15 (1H, dd, *J* = 9.5, 5.4 $D_2O$ Hz, H-C4`)), 3.47 (1H, t, *J* = 9.5 Hz, H-(C4``)). ^{13}C NMR (151 MHz, D_2O) δ 191.1 (C2), 90.1 (C1`), 85.3 (C2`), 71.0 (C3`), 67.4 (C4`). UV/Vis: λ_{max} 242 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) +12.68. HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_5H_7NO_3S$ calcd 162.0225, found 162.0226.

Crystal structure obtained; see appendix for crystallographic data.

<u>α-(2-Cyanovinyl)-erythrosefuranosyl oxazolidinone thione **225**</u>



$C_8H_8N_2O_3S$
212.22



225

Erythrose-furanosyl oxazolidinone thione **224** (30.0 mg, 0.186 mmol) was added to a solution of cyanoacetylene **30** (91.0 mg, 0.372 mmol) in water (1.0 mL). After stirring

the reaction mixture for 1 h at rt. The reaction was then lyophilised to yield α-(2-cyanovinyl)-erythrosefuranosyl oxazolidinone thione **225** (38.0 mg, 0.179 mmol, 96%) as white solid which was used without further purification: M.p. 136-138°C. IR (Solid, cm^{-1}) 3298 (NH), 3167 (OH), 2984, 2864 (CH), 1519 (C=S). ^{1}H NMR (600 MHz, D_2O) δ 7.87 (1H, d, *J* = 10.5 Hz, H-(C4)), 6.11 (1H, d, *J* = 5.4 Hz, H-(C1`)), 5.99 (1H, d, *J* = 10.5 Hz, H-(C5)), 5.18 (1H, t, *J* = 5.4 Hz, H-(C2`)), 4.51 (1H, dt, *J* = 9.7, 5.4 Hz, H-(C3`)), 4.11 (1H, dd, *J* = 9.7, 5.4 Hz, H-(C4`)), 3.29 (1H, t, *J* = 9.7 D_2O Hz, H-(C4``)). ^{13}C NMR (151 MHz, D_2O) δ 168.2 (C2), 143.7 (C4), 115.9 (CN), 100.3 (C1`), 100.1 (C5), 84.2 (C2`), 71.1 (C3`), 67.0 (C4`). UV/Vis: λ_{max} 267 nm. $[\alpha]_D^{20.0}$ (c = 1.00, H_2O) +0.84.HRMS (*m/z*): $[M\text{-}H^+]^-$ $C_8H_8N_2O_3S$ calcd 213.0335, found 213.0334.

Crystal structure obtained; see appendix for crystallographic data.

References

1. C. Mora, D. P. Tittensor, S. Adl, A. G. B. Simpson and B. Worm, *PLoS Biol.*, 2011, **9**, e1001127.
2. L. A. Hug, B. J. Baker, K. Anantharaman, C. T. Brown, A. J. Probst, C. J. Castelle, C. N. Butterfield, A. W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D. A. Relman, K. M. Finstad, R. Amundson, B. C. Thomas and J. F. Banfield, *Nat. Microbiol.*, 2016, **1**, 16048.
3. R. N. Glud, F. Wenzhofer, M. Middelboe, K. Oguri, R. Turnewitsch, D. E. Canfield and H. Kitazato, *Nat. Geosci.*, 2013, **6**, 284-288.
4. D. Wacey, M. R. Kilburn, M. Saunders, J. Cliff and M. D. Brasier, *Nat. Geosci.*, 2011, **4**, 698-702.
5. A. P. Nutman, V. C. Bennett, C. R. L. Friend, M. J. Van Kranendonk and A. R. Chivas, *Nature*, 2016, **537**, 535-538.
6. S. J. Mojzsis, G. Arrhenius, K. D. McKeegan, T. M. Harrison, A. P. Nutman and C. R. L. Friend, *Nature*, 1996, **384**, 55-59.
7. S. A. Wilde, J. W. Valley, W. H. Peck and C. M. Graham, *Nature*, 2001, **409**, 175-178.
8. https://www.nasa.gov/press-release/nasa-releases-kepler-survey-catalog-with-hundreds-of-new-planet-candidates
9. S. A. Benner, *Astrobiology*, 2010, **10**, 1021-1030.
10. http://www.merriam-webster.com/dictionary/life
11. G. Joyce, *Origins of Life: The Central Concepts*, 1994.
12. P. Luisi, *Origins Life Evol. Biosphere*, 1998, **28**, 613-622.
13. A. Eschenmoser and M. Volkan Kisakürek, *Helv. Chim. Acta*, 1996, **79**, 1249-1259.
14. J. W. Szostak, *J. Biomol. Struct. Dyn.*, 2012, **29**, 599-600.
15. P. Ball, *Sci. Amer.*, 2011, **305**, 48-53.
16. Clancy, S., Nat. Ed., 2008, 7(1), 60
17. E. A. Harwood, P. B. Hopkins and S. T. Sigurdsson, *J. Org. Chem*, 2000, **65**, 2959-2964.
18. E. R. Kandimalla, D. Yu, Q. Zhao and S. Agrawal, *Biorg. Med. Chem.*, 2001, **9**, 807-813.
19. Y. Ambroise, C. Mioskowski, G. Leblanc and B. Rousseau, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 1125-1127.
20. F. Crick, *Nature*, 1970, **227**, 561-563.
21. J. Xu, M. Tsanakopoulou, C. J. Magnani, R. Szabla, J. E. Sponer, J. Sponer, R. W. Góra and J. D. Sutherland, *Nat. Chem.*, 2017, **9**, 303-309.
22. M. W. Powner, B. Gerland and J. D. Sutherland, *Nature*, 2009, **459**, 239-242.
23. R. H. Symons, *Annu. Rev. Biochem*, 1992, **61**, 641-671.
24. T. R. Cech, *Angew. Chem. Int. Ed.*, 1990, **29**, 759-768.
25. S. Altman, *Angew. Chem. Int. Ed.*, 1990, **29**, 749-758.
26. T. Inoue, G. F. Joyce, K. Grzeskowiak, L. E. Orgel, J. M. Brown and C. B. Reese, *J. Mol. Biol.*, 1984, **178**, 669-676.
27. A. Zaug and T. Cech, *Science*, 1986, **231**, 470-475.
28. A. J. Zaug, M. D. Been and T. R. Cech, *Nature*, 1986, **324**, 429-433.
29. J. Piccirilli, T. McConnell, A. Zaug, H. Noller and T. Cech, *Science*, 1992, **256**, 1420-1424.
30. J. A. Doudna and J. W. Szostak, *Nature*, 1989, **339**, 519-522.
31. P. Khaitovich, A. S. Mankin, R. Green, L. Lancaster and H. F. Noller, *Proc. Natl. Acad. Sci.*, 1999, **96**, 85-90.
32. H. White, III, *J. Mol. Evol.*, 1976, **7**, 101-104.
33. S. An, R. Kumar, E. D. Sheets and S. J. Benkovic, *Science*, 2008, **320**, 103-106.

34. K. Leu, B. Obermayer, S. Rajamani, U. Gerland and I. A. Chen, *Nucleic Acids Res.*, 2011, **39**, 8135-8147.
35. G. F. Joyce, *Nature*, 2002, **418**, 214-221.
36. J. F. Kasting, *The Chemistry of Life's Origins*, 1993, 149-176.
37. H. Cleaves, *Life*, 2013, **3**, 331.
38. S. L. Miller, *Science*, 1953, **117**, 528-529.
39. F. Wöhler, *Annalen der Physik*, 1828, **88**, 253-256.
40. J. L. Bada, *Chem. Soc. Rev.*, 2013, **42**, 2186-2196.
41. P. H. Abelson, *Proc. Natl. Acad. Sci.*, 1966, **55**, 1365-1372.
42. R. A. Sanchez, J. P. Ferris and L. E. Orgel, *Science*, 1966, **154**, 784-785.
43. J. P. Ferris and W. J. Hagan Jr, *Tetrahedron*, 1984, **40**, 1093-1120.
44. J. R. Cronin, *Adv Space Res:The Official Journal of the Committee on Space Res*, 1989, **9**, 59-64.
45. P. Thaddeus, *Philos. Trans. R. Soc. London, Ser. B*, 2006, **361**, 1681-1687.
46. J. Oró, *Biochem. Biophys. Res. Commun.*, 1960, **2**, 407-412.
47. J. Oro, *Nature*, 1961, **191**, 1193-1194.
48. J. Oró and A. P. Kimball, *Arch. Biochem. Biophys.*, 1961, **94**, 217-227.
49. J. Oró and A. P. Kimball, *Arch. Biochem. Biophys.*, 1962, **96**, 293-313.
50. J. P. Ferris and L. E. Orgel, *J. Am. Chem. Soc.*, 1966, **88**, 1074-1074.
51. J. P. Ferris and L. E. Orgel, *J. Am. Chem. Soc.*, 1966, **88**, 3829-3831.
52. R. A. Sanchez, J. P. Ferris and L. E. Orgel, *J. Mol. Biol.*, 1967, **30**, 223-253.
53. L. E. Orgel and R. Lohrmann, *Acc. Chem. Res.*, 1974, **7**, 368-377.
54. O. Leslie E., *Origins Life Evol. Biosphere*, 2004, **34**, 361-369.
55. S. Miyakawa, H. James Cleaves and S. L. Miller, *Origins Life Evol. Biosphere*, 2002, **32**, 195-208.
56. S. Miyakawa, H. J. Cleaves and S. L. Miller, *Origins Life Evol. Biosphere*, 2002, **32**, 209-218.
57. H. Yamada and T. Okamoto, *Chem. Pharm. Bull*, 1972, **20**, 623-624.
58. W. Traube, *Ber. Dtsch. Chem. Ges*, 1900, **33**, 1371-1383.
59. S. Becker, I. Thoma, A. Deutsch, T. Gehrke, P. Mayer, H. Zipse and T. Carell, *Science*, 2016, **352**, 833-836.
60. Z. Wang, in *Comprehensive Organic Name Reactions and Reagents*, John Wiley & Sons, Inc., 2010, pp. 2789-2792.
61. S. L. Miller and H. J. Cleaves, *Syst. Biol.*, 2006, **1**, 3-56.
62. M. P. Robertson and S. L. Miller, *Nature*, 1995, **375**, 772-774.
63. R. Shapiro, *Origins Life Evol. Biosphere*, 1995, **25**, 83-98.
64. R. Shapiro, *Origins Life Evol. Biosphere*, 2002, **32**, 275-278.
65. R. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 4396-4401.
66. P. Decker, H. Schweer and R. Pohlamnn, *J. Chromatogr. A*, 1982, **244**, 281-291.
67. A. Butlerow, *Justus Liebigs Ann. Chem.*, 1861, **120**, 295-298.
68. H. J. Cleaves, *Encyclopedia of Astrobiology*, 2014, 1-8.
69. R. F. Socha, A. H. Weiss and M. M. Sakharov, *J. Catal.*, 1981, **67**, 207-217.
70. R. Shapiro, *Origins Life Evol. Biosphere*, 1988, **18**, 71-85.
71. R. Larralde, M. P. Robertson and S. L. Miller, *Proc. Natl. Acad. Sci.*, 1995, **92**, 8158-8160.
72. H.-J. Kim, A. Ricardo, H. I. Illangkoon, M. J. Kim, M. A. Carrigan, F. Frye and S. A. Benner, *J. Am. Chem. Soc.*, 2011, **133**, 9457-9468.
73. R. Shapiro, *Origins of Life*, 1984, **14**, 565-570.
74. A. W. Schwartz, *Origins Life Evol. Biosphere*, 1993, **23**, 185-194.
75. A. W. Schwartz and R. M. de Graaf, *J. Mol. Evol.*, 1993, **36**, 101-106.
76. S. Chapelle and J.-F. Verchere, *Tetrahedron*, 1988, **44**, 4469-4482.
77. A. Ricardo, M. A. Carrigan, A. N. Olcott and S. A. Benner, *Science*, 2004, **303**, 196.
78. B. E. Prieur, *Comptes Rendus de l'Académie des Sciences - Series IIC - Chemistry*, 2001, **4**, 667-670.

79. Y. Furukawa, M. Horiuchi and T. Kakegawa, *Origins Life Evol. Biosphere*, 2013, **43**, 353-361.
80. D. Müller, S. Pitsch, A. Kittaka, E. Wagner, C. E. Wintner and A. Eschenmoser, *Helv. Chim. Acta*, 1990, **73**, 1410-1468.
81. R. Krishnamurthy, G. Arrhenius and A. Eschenmoser, *Origins Life Evol. Biosphere*, 1999, **29**, 333-354.
82. R. Krishnamurthy, S. Guntha and A. Eschenmoser, *Angew. Chem. Int. Ed.*, 2000, **39**, 2281-2285.
83. S. Islam and M. W. Powner, *Chem*, **2**, 470-501.
84. R. Breslow and Z.-L. Cheng, *Proc. Natl. Acad. Sci.*, 2009, **106**, 9144-9146.
85. A. Eschenmoser, *Angew. Chem. Int. Ed.*, 2011, **50**, 12412-12472.
86. E. Fischer, *Ber. Dtsch. Chem. Ges*, 1889, **22**, 2204-2205.
87. A. S. Serianni, E. L. Clark and R. Barker, *Carbohydr. Res.*, 1979, **72**, 79-91.
88. D. Ritson and J. D. Sutherland, *Nat. Chem.*, 2012, **4**, 895-899.
89. B. H. Patel, C. Percivalle, D. J. Ritson, C. D. Duffy and J. D. Sutherland, *Nat. Chem.*, 2015, **7**, 301–307
90. G. Schlesinger and S. L. Miller, *J. Am. Chem. Soc.*, 1973, **95**, 3729-3735.
91. T. Arrhenius, G. Arrhenius and W. Paplawsky, *Origins Life Evol. Biosphere*, 1994, **24**, 1-17.
92. D. J. Ritson and J. D. Sutherland, *Angew. Chem. Int. Ed.*, 2013, **52**, 5845-5847.
93. G. Zubay, *Origins Life Evol. Biosphere*, 1998, **28**, 13-26.
94. C. Meinert, I. Myrgorodska, P. de Marcellus, T. Buhse, L. Nahon, S. V. Hoffmann, L. L. S. d'Hendecourt and U. J. Meierhenrich, *Science*, 2016, **352**, 208-212.
95. R. Krishnamurthy, *Acc. Chem. Res.*, 2017, **50**, 455-459.
96. W. D. Fuller, R. A. Sanchez and L. E. Orgel, *J. Mol. Biol.*, 1972, **67**, 25-33.
97. W. D. Fuller, R. A. Sanchez and L. E. Orgel, *J. Mol. Evol.*, 1972, **1**, 249-257.
98. K. N. Drew, J. Zajicek, G. Bondo, B. Bose and A. S. Serianni, *Carbohydr. Res.*, 1998, **307**, 199-209.
99. C. Fonseca Guerra, F. M. Bickelhaupt, S. Saha and F. Wang, *J. Phys. Chem. A*, 2006, **110**, 4012-4020.
100. O. Leslie E., *Crit. Rev. Biochem. Mol. Biol.*, 2004, **39**, 99-123.
101. A. Eschenmoser and E. Loewenthal, *Chem. Soc. Rev.*, 1992, **21**, 1-16.
102. A. Eschenmoser, *Chem. & Biod.*, 2007, **4**, 554-573.
103. S. Pitsch, S. Wendeborn, R. Krishnamurthy, A. Holzner, M. Minton, M. Bolli, C. Miculca, N. Windhab, R. Micura, M. Stanek, B. Jaun and A. Eschenmoser, *Helv. Chim. Acta*, 2003, **86**, 4270-4363.
104. A. Eschenmoser and M. Dobler, *Helv. Chim. Acta*, 1992, **75**, 218-259.
105. M. Beier, F. Reck, T. Wagner, R. Krishnamurthy and A. Eschenmoser, *Science*, 1999, **283**, 699-703.
106. A. A. Koshkin, P. Nielsen, M. Meldgaard, V. K. Rajwanshi, S. K. Singh and J. Wengel, *J. Am. Chem. Soc.*, 1998, **120**, 13252-13253.
107. K.-U. Schöning, P. Scholz, X. Wu, S. Guntha, G. Delgado, R. Krishnamurthy and A. Eschenmoser, *Helv. Chim. Acta*, 2002, **85**, 4111-4153.
108. S. Islam, J. A. Aguilar, M. W. Powner, M. Nilsson, G. A. Morris and J. D. Sutherland, *Chem. Eur. J.*, 2013, **19**, 4586-4595.
109. X. Wu, G. Delgado, R. Krishnamurthy and A. Eschenmoser, *Org. Lett.*, 2002, **4**, 1283-1286.
110. P. Nielsen, *Origins Life Evol. Biosphere*, 1993, **23**, 323-327.
111. H. Knudsen and P. E. Nielsen, *Nucleic Acids Res.*, 1996, **24**, 494-500.
112. J. G. Schmidt, P. E. Nielsen and L. E. Orgel, *Nucleic Acids Res.*, 1997, **25**, 4797-4802.
113. L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174-4175.
114. C.-H. Tsai, J. Chen and J. W. Szostak, *Proc. Natl. Acad. Sci.*, 2007, **104**, 14598-14603.
115. M. P. Robertson and G. F. Joyce, *Cold Spring Harb Perspect Biol.*, 2012, **4**.

116. J. D. Sutherland, *Nat. Rev. Chem.*, 2017, **1**, 1-7.
117. J. D. Sutherland, *Angew. Chem. Int. Ed.*, 2015, **55**, 104-121.
118. M. W. Powner, *Nat. Chem.*, 2013, **5**, 355-357.
119. J. W. Szostak, *Nature*, 2009, **459**, 171-172.
120. G. Springsteen and G. F. Joyce, *J. Am. Chem. Soc.*, 2004, **126**, 9578-9583.
121. C. Anastasi, M. A. Crowe, M. W. Powner and J. D. Sutherland, *Angew. Chem. Int. Ed.*, 2006, **45**, 6176-6179.
122. J. E. Hein and D. G. Blackmond, *Acc. Chem. Res.*, 2012, **45**, 2045-2054.
123. M. W. Powner and J. D. Sutherland, *Angew. Chem.*, 2010, **122**, 4745-4747.
124. R. A. Sanchez and L. E. Orgel, *J. Mol. Biol.*, 1970, **47**, 531-543.
125. M. W. Powner, C. Anastasi, M. A. Crowe, A. L. Parkes, J. Raftery and J. D. Sutherland, *ChemBioChem*, 2007, **8**, 1170-1179.
126. A. Choudhary, K. J. Kamer, M. W. Powner, J. D. Sutherland and R. T. Raines, *ACS Chem. Biol.*, 2010, **5**, 655-657.
127. M. W. Powner and J. D. Sutherland, *ChemBioChem*, 2008, **9**, 2386-2387.
128. M. W. Powner, J. D. Sutherland and J. W. Szostak, *J. Am. Chem. Soc.*, 2010, **132**, 16677-16688.
129. M. W. Powner, S.-L. Zheng and J. W. Szostak, *J. Am. Chem. Soc.*, 2012, **134**, 13889-13895.
130. S. A. Benner, H.-J. Kim and Z. Yang, *Cold Spring Harb Perspect Biol.*, 2012, **4**, a003541.
131. Nicholas V. Hud, Brian J. Cafferty, R. Krishnamurthy and Loren D. Williams, *Chem. & Biol.*, 2013, **20**, 466-474.
132. A. W. Schwartz, *Astrobiology*, 2013, **13**, 784-789.
133. R. W. Nagorski and J. P. Richard, *J. Am. Chem. Soc.*, 2001, **123**, 794-802.
134. S. Islam, D.-K. Bučar and M. W. Powner, *Nat. Chem.*, 2017, **9**, 584–589
135. A. D. Keefe, G. L. Newton and S. L. Miller, *Nature*, 1995, **373**, 683-685.
136. M. C. L. Rosaly and A. W. David, *Rep. Prog. Phys.*, 2005, **68**, 303.
137. R. V. Eck and M. O. Dayhoff, *Science*, 1966, **152**, 363-366.
138. J. E. Goldford, H. Hartman, T. F. Smith and D. Segrè, *Cell*, 2017, **168**, 1126-1134.
139. C. Bonfio, L. Valer, S. Scintilla, S. Shah, D. J. Evans, L. Jin, J. W. Szostak, D. D. Sasselov, J. D. Sutherland and S. S. Mansy, *Nat. Chem.*, 2017, DOI: 10.1038/nchem.2817.
140. P. Dawson, T. Muir, I. Clark-Lewis and S. Kent, *Science*, 1994, **266**, 776-779.
141. C.-F. Liu, C. Rao and J. P. Tam, *Tetrahedron Lett.*, 1996, **37**, 933-936.
142. F. R. Bowler, C. K. W. Chan, C. D. Duffy, B. Gerland, S. Islam, M. W. Powner, J. D. Sutherland and J. Xu, *Nat. Chem.*, 2013, **5**, 383-389.
143. H. H. Zepik, S. Rajamani, M. C. Maurel and D. Deamer, *Origins Life Evol. Biosphere*, 2007, **37**, 495-505.
144. M.-C. Maurel and L. E. Orgel, *Origins Life Evol. Biosphere*, 2000, **30**, 423-430.
145. L. Leman, L. Orgel and M. R. Ghadiri, *Science*, 2004, **306**, 283-286.
146. M. W. Powner, J. D. Sutherland and J. W. Szostak, *Synlett*, 2011, **2011**, 1956-1964.
147. S. Kanvah, J. Joseph, G. B. Schuster, R. N. Barnett, C. L. Cleveland and U. Landman, *Acc. Chem. Res.*, 2010, **43**, 280-287.
148. K. C. Cheng, D. S. Cahill, H. Kasai, S. Nishimura and L. A. Loeb, *J. Biol. Chem.*, 1992, **267**, 166-172.
149. K. V. Nguyen and C. J. Burrows, *J. Am. Chem. Soc.*, 2011, **133**, 14586-14589.
150. L. Haracska, S.-L. Yu, R. E. Johnson, L. Prakash and S. Prakash, *Nat. Genet.*, 2000, **25**, 458-461.
151. K. D. Carlson and M. T. Washington, *Mol. Cell. Biochem.*, 2005, **25**, 2169-2176.
152. L. A. Lipscomb, M. E. Peek, M. L. Morningstar, S. M. Verghis, E. M. Miller, A. Rich, J. M. Essigmann and L. D. Williams, *Proc. Natl. Acad. Sci.*, 1995, **92**, 719-723.
153. Y. Oda, S. Uesugi, M. Ikehara, S. Nishimura, Y. Kawase, H. Ishikawa, H. Inoue and E. Ohtsuka, *Nucleic Acids Res.*, 1991, **19**, 1407-1412.

154. S. K. Kim, S. H. Lee, O. S. Kwon and B. J. Moon, *J. Biochem. Mol. Bio.*, 2004, **37**, 657-662.
155. T. Hofer, A. Y. Seo, M. Prudencio and C. Leeuwenburgh, *Biol. Chem.*, 2006, **387**, 103-111.
156. G. Gosselin, M. C. Bergogne, J. De Rudder, E. De Clercq and J. L. Imbach, *J. Med. Chem.*, 1986, **29**, 203-213.
157. J. Girniene, S. Tardy, A. Tatibouët, A. Sačkus and P. Rollin, *Tetrahedron Lett.*, 2004, **45**, 6443-6446.
158. J. Girniene, G. Apremont, A. Tatibouët, A. Sackus and P. Rollin, *Tetrahedron*, 2004, **60**, 2609-2619.
159. J. Girniene, D. Gueyrard, A. Tatibouët, A. Sackus and P. Rollin, *Tetrahedron Lett.*, 2001, **42**, 2977-2980.
160. A. F. Cockerill, A. Deacon, R. G. Harrison, D. J. Osborne, D. M. Prime, W. J. Ross, A. Todd and J. P. Verge, *Synthesis*, 1976, **1976**, 591-593.
161. C. Chyba and C. Sagan, *Nature*, 1992, **355**, 125-132.
162. J. K. Bartlett and D. A. Skoog, *Anal. Chem.*, 1954, **26**, 1008-1011.
163. N. Leconte, S. Silva, A. Tatibouët, A. P. Rauter and P. Rollin, *Synlett*, 2006, **2006**, 301-305.
164. M. T. Beltrán, C. Codella, S. Viti, R. Neri and R. Cesaroni, *Astrophys. J. Lett.*, 2009, **690**, L93.
165. J. P. Ferris, *Philos. Trans. R. Soc. London, Ser. B*, 2006, **361**, 1777-1786.
166. R. Loos, S. Kobayashi and H. Mayr, *J. Am. Chem. Soc.*, 2003, **125**, 14126-14132.
167. Y.-B. Xiang, S. Drenkard, K. Baumann, D. Hickey and A. Eschenmoser, *Helv. Chim. Acta*, 1994, **77**, 2209-2250.
168. G. Lacasse and J. M. Muchowski, *Can. J. Chem.*, 1972, **50**, 3082-3083.
169. W. A. Jr. and J. S. Chickos, *J. Phys. Chem. Ref. Data*, 2010, **39**, 043101.
170. L. Orgel, *Origins Life Evol. Biosphere*, 2002, **32**, 279-281.
171. H. Bader, L. C. Cross, I. Heilbron and E. R. H. Jones, *J. Chem. Soc*, 1949, DOI: 10.1039/jr9490000619, 619-623.
172. H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed. Engl.*, 2001, **40**, 2004-2021.
173. E. R. Walwich, Roberts, W. K. & Dekker, C. A., *Proc. Chem. Soc*, 1959, DOI: 10.1039/ps9590000073, 73-108.
174. C. Fernandez-Garcia, N. M. Grefenstette and M. W. Powner, *Chem. Commun.*, 2017, DOI: 10.1039/c7cc02183f.
175. H. J. Brown, G. Shaw and D. Wright, *J. Chem. Soc., Perkin Trans. 1*, 1981, DOI: 10.1039/p19810000657, 657-660.
176. R. M. Davidson, G. D. Byrd, E. White, V. Samm, A. Margolis and B. Coxon, *Magn. Reson. Chem.*, 1986, **24**, 929-937.
177. C. E. Hand and J. F. Honek, *J. Nat. Prod.*, 2005, **68**, 293-308.
178. G. M. Tener, *J. Am. Chem. Soc.*, 1961, **83**, 159-168.
179. R. L. Letsinger and K. K. Ogilvie, *J. Am. Chem. Soc.*, 1969, **91**, 3350-3355.
180. V. A. Efimov, A. A. Buryakova, S. V. Reverdatto, O. G. Chakhmakhcheva and Y. A. Ovchinnikov, *Nucleic Acids Res.*, 1983, **11**, 8369-8387.
181. C. B. Reese, *Tetrahedron*, 1978, **34**, 3143-3179.
182. S. Pizzarello, L. B. Williams, J. Lehman, G. P. Holland and J. L. Yarger, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 4303-4306.
183. B. Donn, *J. Mol. Evol.*, 1982, **18**, 157-160.
184. A. T. Tokunaga, S. C. Beck, T. R. Geballe, J. H. Lacy and E. Serabyn, *Icarus*, 1981, **48**, 283-289.
185. R. Hanel, B. Conrath, F. M. Flasar, V. Kunde, W. Maguire, J. Pearl, J. Pirraglia, R. Samuelson, L. Herath, M. Aallison, C. D., D. Gautier, P. Gierasch, L. Horn, R. Koppany and C. Ponnamperuma, *Science*, 1981, **212**, 192-200.
186. C. U. Lowe, M. W. Rees and R. Markham, *Nature*, 1963, **199**, 219-222.
187. F. Freeman, *Synthesis*, 1981, **12**, 925-954.

188. M. Ikehara, T. Nagura and E. Ohtsuka, *Chem. Pharm. Bull*, 1974, **22**, 2578-2586.
189. A. Belloche, K. M. Menten, C. Comito, H. S. P. Müller, P. Schilke, J. Ott, S. Thorwirth and C. Hieret, *Astron. Astrophys.*, 2008, **482**, 179-196.
190. G. Moutou, J. Taillades, S. Bénefice-Malouet, A. Commeyras, G. Messina and R. Mansani, *J. Phys. Org. Chem.*, 1995, **8**, 721-730.
191. G. W. Stevenson and D. Williamson, *J. Am. Chem. Soc.*, 1958, **80**, 5943-5947.
192. N. J. Cusack, G. Shaw and G. J. Litchfield, *J. Chem. Soc. C*, 1971, DOI: 10.1039/j39710001501, 1501-1507.
193. G. J. Litchfield and G. Shaw, *Chem. Commun.*, 1965, DOI: 10.1039/c19650000563, 563-565.
194. R. M. Beesley, C. K. Ingold and J. F. Thorpe, *J. Chem. Soc, Trans.*, 1915, **107**, 1080-1106.
195. S. Stairs and M. W. Powner, *Synlett*, DOI: 10.1055/s-0036-1590968.
196. K. Koch, W. B. Schweizer and A. Eschenmoser, *Chem. & Biod.*, 2007, **4**, 541-553.
197. J. D. Sutherland, *Cold Spring Harb Perspect Biol.*, 2010, **2**.
198. H. S. Bernhardt and R. K. Sandwick, *J. Mol. Evol.*, 2014, **79**, 91-104.
199. K. Ace and J. D. Sutherland, *Chem. & Biod.*, 2004, **1**, 1678-1693.
200. I. Alseth, B. Dalhus and M. Bjørås, *Current Opinion in Genet. & Dev.*, 2014, **26**, 116-123.
201. M. Levy, S. L. Miller and J. Oró, *J. Mol. Evol.*, 1999, **49**, 165-168.
202. B. Basile, A. Lazcano and J. Oró, *Adv. Space Res.*, 1984, **4**, 125-131.
203. R. A. Sanchez, J. P. Ferris and L. E. Orgel, *J. Mol. Biol.*, 1968, **38**, 121-128.
204. R. J. Capon and N. S. Trotter, *J. Nat. Prod.*, 2005, **68**, 1689-1691.
205. A. P. Martinez, W. W. Lee and L. Goodman, *J. Org. Chem*, 1966, **31**, 3263-3267.
206. V. M. Clark, A. R. Todd and J. Zussman, *J. Chem. Soc*, 1951, DOI: 10.1039/jr9510002952, 2952-2958.
207. K. K. Ogilvie and L. A. Slotin, *Can. J. Chem.*, 1973, **51**, 2397-2405.
208. K. K. Ogilvie, L. A. Slotin, J. H. Westmore and D. C. K. Lin, *J. Heterocycl. Chem.*, 1972, **9**, 1179-1180.
209. M. Ikehara, *Acc. Chem. Res.*, 1969, **2**, 47-53.
210. J. Nagyvary, *J. Am. Chem. Soc.*, 1969, **91**, 5409-5410.
211. M. Ikehara, S. Uesugi and M. Kaneko, *Chem. Commun.*, 1967, DOI: 10.1039/c19670000017, 17-18.
212. M. Ikehara and T. Maruyama, *Tetrahedron*, 1975, **31**, 1369-1372.
213. M. Ikehara, S. Uesugi and J. Yano, *J. Am. Chem. Soc.*, 1974, **96**, 4966-4972.
214. K. K. Ogilvie, L. Slotin, J. B. Westmore and D. Lin, *Can. J. Chem.*, 1972, **50**, 2249-2253.
215. E. K. Ryu and M. MacCoss, *J. Org. Chem*, 1981, **46**, 2819-2823.
216. E. J. Reist, D. F. Calkins, L. V. Fisher and L. Goodman, *J. Org. Chem*, 1968, **33**, 1600-1603.
217. M. Ikehara and M. Muraoka, *Chem. Pharm. Bull*, 1976, **24**, 672-682.
218. R. Shapiro and S. H. Pohl, *Biochem*, 1968, **7**, 448-455.
219. M. W. Powner and J. D. Sutherland, *Philos. Trans. R. Soc. London, Ser. B*, 2011, **366**, 2870-2877.
220. A. F. Russell, S. Greenberg and J. G. Moffatt, *J. Am. Chem. Soc.*, 1973, **95**, 4025-4030.
221. A. Benitez, O. P. Crews, L. Goodman and B. R. Baker, *J. Org. Chem*, 1960, **25**, 1946-1950.
222. J. B. Chattopadhyaya and C. B. Reese, *J. Chem. Soc., Chem. Commun.*, 1976, DOI: 10.1039/c39760000860, 860-862.
223. M. Ikehara and Y. Ogiso, *Tetrahedron*, 1972, **28**, 3695-3704.
224. V. P. Sokhan, A. P. Jones, F. S. Cipcigan, J. Crain and G. J. Martyna, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112**, 6341-6346.
225. C. Fernández-García, A. J. Coggins and M. W. Powner, *Life*, 2017, **7**, 31.
226. M. Gull, *Challenges*, 2014, **5**, 193-221.

227. M. Halmann, R. A. Sanchez and L. E. Orgel, *J. Org. Chem*, 1969, **34**, 3702-3703.
228. M. Karki, C. Gibard, S. Bhowmik and R. Krishnamurthy, *Life*, 2017, **7**, 32.
229. O. T. Quimby and T. J. Flautt, *Z. Anorg. Allg. Chem.*, 1958, **296**, 220-228.
230. B. Burcar, M. Pasek, M. Gull, B. J. Cafferty, F. Velasco, N. V. Hud and C. Menor-Salván, *Angew. Chem. Int. Ed.*, 2016, **55**, 13249-13253.
231. R. Lohrmann and L. E. Orgel, *Science*, 1971, **171**, 490-494.
232. C. Ponnamperuma and R. Mack, *Science*, 1965, **148**, 1221-1223.
233. J. Škoda and J. Morávek, *Tetrahedron Lett.*, 1966, **7**, 4167-4172.
234. A. Beck, R. Lohrmann and L. E. Orgel, *Science*, 1967, **157**, 952-952.
235. G. W. Cooper and J. R. Cronin, *Geochim. Cosmochim. Acta*, 1995, **59**, 1003-1015.
236. R. E. D. Clark, *Nature*, 1951, **168**, 876-876.
237. G. Imperato, S. Hoger, D. Lenoir and B. Konig, *Green Chem.*, 2006, **8**, 1051-1055.
238. M. Gull, M. Zhou, F. M. Fernández and M. A. Pasek, *J. Mol. Evol.*, 2014, **78**, 109-117.
239. R. Österberg, L. E. Orgel and R. Lohrmann, *J. Mol. Evol.*, 1973, **2**, 231-234.
240. R. Kluger, P. P. Davis and P. D. Adawadkar, *J. Am. Chem. Soc.*, 1979, **101**, 5995-6000.
241. R. Osterberg and L. E. Orgel, *J. Mol. Evol.*, 1972, **1**, 241-248.
242. A. M. Schoffstall, *Origins of Life*, 1976, **7**, 399-412.
243. C. M. Tapiero and J. Nagyvary, *Nature*, 1971, **231**, 42-43.
244. P. A. Gerakines, M. H. Moore and R. L. Hudson, *Icarus*, 2004, **170**, 202-213.
245. M. P. Bernstein, L. J. Allamandola and S. A. Sandford, *Adv. Space Res.*, 1997, **19**, 991-998.
246. A. Segura and R. Navarro-González, *Geophy. Res. Lett.*, 2005, **32**, 1-4.
247. R. L. Hudson and M. H. Moore, *Icarus*, 2004, **172**, 466-478.
248. G. Yu, L. C. Chris, M. S. Philip and A. V. B. Paul, *Astrophys. J. Lett.*, 2007, **660**, L93.
249. G. A. Krudy and R. S. Macomber, *J. Chem. Educ.*, 1979, **56**, 109.
250. H. Mita, S. Nomoto, M. Terasaki, A. Shimoyama and Y. Yamamoto, *Int. J. Astrobiol.*, 2005, **4**, 145-154.
251. R. Mengel and W. Muhs, *Chem. Ber.*, 1979, **112**, 625-639.
252. G. S. G. De Carvalho, J.-L. Fourrey, R. H. Dodd and A. D. Da Silva, *Tetrahedron Lett.*, 2009, **50**, 463-466.
253. E. T. Parker, H. J. Cleaves, M. P. Callahan, J. P. Dworkin, D. P. Glavin, A. Lazcano and J. L. Bada, *Origins Life Evol. Biosphere*, 2011, **41**, 201-212.
254. H. J. Cleaves Ii, *Monatshefte für Chemie / Chemical Monthly*, 2003, **134**, 585-593.
255. M. Fiore and P. Strazewski, *Angew. Chem. Int. Ed.*, 2016, **55**, 13930-13933.
256. K. Tekenaka, T. Tsuji and M. Muraoka, *Nucleosides and Nucleotides*, 1998, **17**, 869-874.
257. V. Skaric and J. Matulic-Adamic, *J. Chem. Soc., Perkin Trans. 1*, 1985, DOI: 10.1039/p19850000779, 779-783.
258. R. A. Mathes and A. J. Beber, *J. Am. Chem. Soc.*, 1948, **70**, 1451-1452.
259. J. C. Jochims, A. Seeliger and G. Taigel, *Chem. Ber.*, 1967, **100**, 845-854.
260. J. D. Street, O. S. Mills and J. A. Joule, *Acta Crystallographica Section C*, 1991, **47**, 1523-1525.
261. W. H. Bromund and R. M. Herbst, *J. Org. Chem*, 1945, **10**, 267-276.
262. J. Kovács, I. Pintér, U. Lendering and P. Köll, *Carbohydr. Res.*, 1991, **210**, 155-166.
263. EU Pat., EP1007509, 2000.
264. F. Scotti and E. J. Frazza, *J. Org. Chem*, 1964, **29**, 1800-1808.
265. W. R. Mancini, M. S. Williams and T. S. Lin, *Biochem*, 1988, **27**, 8832-8839.
266. M. Ikehara and Y. Ogiso, *Tetrahedron*, 1972, **28**, 3695-3704.

Appendix

Crystallographic and refinement parameters

Compound	*xylo*-**103**	*arabino*-**120**	**98**	*arabino*-**124**
chemical formula	$C_6H_9NO_4S$	$C_9H_{10}N_2O_4S$	$C_3H_3NO_S$	$C_{13}H_{12}N_2O_5$
Mr/gmol-1	191.20	242.25	101.12	276.25
crystal system	orthorhombic	triclinic	monoclinic	monoclinic
space group	*P212121*	*P1*	*P21/c*	*C2*
a/Å	4.7529(2)	5.6835(2)	3.985(3)	21.1572(7)
b/Å	7.0665(2)	5.6922(2)	10.002(3)	8.1067(3)
c/Å	22.8605(7)	17.1686(5)	10.602(3)	6.8795(2)
α/°	90	94.661(2)	90	90
β/°	90	95.058(2)	95.398(4)	107.2490(10)
γ/°	90	109.588(2)	90	90
V/Å3	767.80(5)	517.60(3)	420.7(3)	1126.87(6)
Z	4	2	4	4
Dc/gcm−3	1.654	1.554	1.596	1.628
F(000)	400	252	208	576
λ/Å	0.71073	0.71073	0.71073	1.54178
μ/mm−1	0.394	0.313	0.590	1.080
T/K	100(2)	100(2)	100(2)	100(2)
crystal size/mm	0.28 × 0.14 × 0.12	0.34 × 0.20 × 0.16	0.22 × 0.18 × 0.16	0.24 × 0.16 × 0.12
index range	-6 → 6 -9 → 9 -30 → 30	-7 → 7 -7 → 7 -23 → 23	-5 → 5 -13 → 13 -14 → 14	-24 → 24 -9 → 9 -8 → 8
collected reflections	11456	14693	7757	13377
unique reflections	2008	5322	1099	1835
Rint	0.0453	0.0305	0.0275	0.0330
reflections with I > 2σ(I)	1911	5030	1030	1831
no. parameters	121	308	59	190
R(F), F > 2σ(F)	0.0272	0.0306	0.0233	0.0217
wR(F2), F > 2σ(F)	0.0291	0.0332	0.0248	0.0217
R(F), all data	0.0671	0.0691	0.0618	0.0556
wR(F2), all data	0.0679	0.0704	0.0629	0.0556
Δr (min., max.)/e Å−3	-0.181, 0.279	-0.190, 0.281	-0.304, 0.308	-0.200, 0.226

Compound	*ribo*-**124**	**129**	**96**	**D-96**
chemical formula	$C_{13}H_{14}N_2O_6$	$C_6H_4N_2S$	$C_9H_{10}N_4O_4$	$C_9H_{12}N_4O_5$
Mr/gmol-1	294.26	136.17	238.21	256.23
crystal system	orthorhombic	monoclinic	orthorhombic	orthorhombic
space group	*P212121*	*Pc*	*P212121*	*P212121*
a/Å	6.6843(4)	3.8091(2)	6.36040(10)	5.16040(10)
b/Å	8.0162(4)	9.4498(5)	9.83730(10)	13.4735(2)
c/Å	24.7739(14)	9.2941(4)	16.07840(10)	14.8581(2)
α/°	90	90	90	90
β/°	90	100.940(4)	90	90
γ/°	90	90	90	90
V/Å3	1327.45(13)	328.46(3)	1006.01(2)	1033.06(3)
Z	4	2	4	4
Dc/gcm−3	1.472	1.377	1.573	1.647
F(000)	616	140	496	536
λ/Å	1.54178	1.54178	1.54184	1.54184
μ/mm−1	1.009	3.567	1.081	1.172
T/K	100(2)	100(2)	150.0(1)	152(5)
crystal size/mm	0.28 × 0.16 × 0.10	0.20 × 0.18 × 0.08	0.23 × 0.14 × 0.11	0.26 × 0.08 × 0.03
index range	-6 → 7 -9 → 9 -29 → 29	-4 → 4 -11 → 11 -11 → 10	-7 → 7 -12 → 11 -20 → 20	-6 → 6 -16 → 16 -17 → 17
collected reflections	30166	5548	34327	33377
unique reflections	2274	1006	2022	1821
Rint	0.0771	0.0667	0.0340	0.0804
reflections with I > 2σ(I)	2168	921	2010	1771
no. parameters	206	82	164	177
R(F), F > 2σ(F)	0.0546	0.0761	0.0279	0.0289
wR(F2), F > 2σ(F)	0.0558	0.0813	0.0280	0.0289
R(F), all data	0.1495	0.1814	0.0730	0.0747
wR(F2), all data	0.1505	0.1872	0.0731	0.0754
Δr (min., max.)/e Å−3	-0.250, 0.413	-0.293, 0.924	-0.215, 0.250	-0.195, 0.180

Compound	**L-96**	**89**	**170**	**182**
chemical formula	$C_9H_{12}N_4O_5$	$C_{10}H_{17}N_5O_7$	$C_{10}H_{10}N_4O_5$	$C_{10}H_{11}N_5O_4$
Mr/gmol-1	256.23	319.28	266.22	265.24
crystal system	orthorhombic	orthorhombic	monoclinic	monoclinic
space group	*P212121*	*P212121*	*P21*	*P21*
a/Å	5.16330(10)	5.9577(2)	8.59170(10)	8.3830(3)
b/Å	13.4533(2)	8.6625(3)	6.88950(10)	6.73040(10)
c/Å	14.8581(2)	27.0257(8)	9.4985(2)	10.1823(3)
α/°	90	90	90	90
β/°	90	90	109.654(2)	112.672(4)
γ/°	90	90	90	90
V/Å3	1032.09(3)	1394.76(8)	529.484(16)	530.10(3)
Z	4	4	2	2
Dc/gcm−3	1.649	1.521	1.670	1.662
F(000)	536	672	276	276
λ/Å	1.5418	1.54184	1.54184	1.54184
μ/mm−1	1.174	1.119	1.178	1.127
T/K	100(1)	200(1)	150(1)	150(1)
crystal size/mm	0.38 × 0.06 × 0.04	0.37 × 0.04 × 0.02	0.40 × 0.23 × 0.13	0.47 × 0.08 × 0.07
index range	-6 → 6 -16 → 16 -17 → 17	-5 → 7 -10 → 10 -32 → 28	-10 → 10 -8 → 8 -10 → 11	-9 → 9 -8 → 5 -11 → 12
collected reflections	14385	3241	6422	2041
unique reflections	1820	2132	1851	1393
Rint	0.0603	0.0186	0.0131	0.0181
reflections with I > 2σ(I)	1772	2022	1838	1374
no. parameters	177	225	178	182
R(F), F > 2σ(F)	0.0320	0.0281	0.0245	0.0239
wR(F2), F > 2σ(F)	0.0330	0.0307	0.0248	0.0243
R(*F*), all data	0.0835	0.0686	0.0645	0.0613
wR(F2), all data	0.0844	0.0703	0.0646	0.0619
Δr (min., max.)/e Å−3	-0.204, 0.284	-0.182, 0.156	-0.183, 0.178	-0.176, 0.192

Compound	183
chemical formula	$C_{10}H_{11}N_5O_4$
Mr/gmol-1	265.24
crystal system	orthorhombic
space group	*P212121*
a/Å	6.9590(2)
b/Å	7.4942(2)
c/Å	20.5342(5)
α/°	90
β/°	90
γ/°	90
V/Å3	1070.90(5)
Z	4
Dc/gcm−3	1.645
F(000)	552
λ/Å	1.54184
μ/mm−1	1.116
T/K	150(1)
crystal size/mm	0.13 × 0.12 × 0.09
index range	-8 → 5 -7 → 8 -19 → 24
collected reflections	2611
unique reflections	1707
Rint	0.0190
reflections with I > 2σ(I)	1631
no. parameters	182
R(F), F > 2σ(F)	0.0292
wR(F2), F > 2σ(F)	0.0311
R(F), all data	0.0712
wR(F2), all data	0.0724
Δr (min., max.)/e Å−3	-0.215, 0.176