Fakultät für Chemie und Pharmazie
Ludwig-Maximilians-Universität München

Dissertation zur Erlangung des Doktorgrades

Die Selektivität der Einfachheit
-
Präbiotische Betrachtungen zur Synthese von Desoxyribonukleosiden

Florian Michael Kruse

2021

Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

Die Selektivität der Einfachheit

-

Präbiotische Betrachtungen zur Synthese von Desoxyribonukleosiden

The Selectivity of Simplicity - Prebiotic Investigations on Deoxyribonucleosides

Florian Michael Kruse
aus
München, Deutschland

2021

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Oliver Trapp betreut.

Eidestattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 17.05.2021

.....
Florian Michael Kruse

Dissertation eingereicht am:	20.05.2021
1. Gutachter:	Prof. Dr. Oliver TRAPP
2. Gutachter:	Prof. Dr. Thomas CARELL
Mündliche Prüfung am:	28.06.2021

Zusammenfassung

Diese Dissertation beschreibt die präbiotische Synthese von Ribo- (RNS) und Desoxyribonukleosiden (DNS). Es wurde ein neuer retrosynthetischer Ansatz gewählt, um Nukleoside präbiotisch plausibel aufzubauen. Die neuartige Ribosylierung verbessert damit die bisher gängige Praxis. Es sind alle kanonische Desoxyribonukleoside zugänglich. Die Reaktionsbedingungen sind mild und präbiotisch äußerst plausibel. Es sind keine sequenziellen Schritte nötig und der Bedarf an Ribose/Desoxyribose wurde vermieden.

Anders als bisher gezeigt, erzielt diese DNS-Nukleosid eine exzellente β-Stereoselektivität, sowie eine exklusive furano-Selektivität. Aufgrund dieser hohen Selektivitäten wird davon ausgegangen, dass die dargestellte präbiotische Synthese der Weg zu DNS-Nukleosiden ist. Des Weiteren wurde ein möglicher direkter Vorfahre der DNS-Nukleoside identifiziert. Dieser ist durch Variation der verwendeten Zucker erreichbar. Es wird angenommen, dass die Desoxyapio-Nukleoside (DApiNA) im Verlauf einer molekularen Evolution ersetzt wurden. Die Totalsynthese der DApiNA-Nukleoside, sowie deren präbiotischer Nachweis wird beschrieben. Ein weiterer Zugang zu Nukleosidanaloga ist möglicherweise auch außerhalb der präbiotischen Chemie von Bedeutung.

Durch eine weitere Änderung der genutzten Zucker ist ein äquivalenter Weg zu Ribonukleosiden möglich. Im Gegensatz zu bisherigen Synthesen kommt dieser ohne drastische Änderung der Reaktionsbedingungen aus. Des Weiteren ist eine gleichzeitige Synthese von DNS- und RNS-Nukleosiden unter den gleichen Bedingungen möglich. Es wird daher davon ausgegangen, dass beide Nukleoside am Ursprung des Lebens eine Rolle gespielt haben.

Darüber hinaus wird eine präbiotische Cytosin Synthese beschrieben. Die Synthese wurde so gewählt, dass möglichst ähnliche Ausgangsmaterialien wie in der DNS-Nukleosidsynthese verwendet werden. Ferner wurde untersucht, ob die neuartige Ribosylierung auf weitere Heterozyklen angewendet werden kann. Die Variation der Heterozyklen eröffnet eine mögliche präbiotische Synthese des Kofaktors Riboflavin.

Die beschriebene Ribosylierung von Heterozyklen ist einfach, effektiv und läuft unter milden Bedingungen. Die Entdeckung dieser eröffnet einen Weg zu DApiNA-Nukleosiden, Riboflavin, RNS-Nukleosiden und DNS-Nukleosiden. Es wird daher davon ausgegangen, dass dieser Weg präbiotisch äußerst relevant ist.

Abstract

This doctoral thesis presents the prebiotic synthesis of ribonucleosides (RNA) and deoxyribonucleosides. In search of prebiotic access to deoxyribonucleosides (DNA), a new approach to disconnect ribonucleosides was discovered. This new sugar-forming reaction improves the so far applied practice leading to ribonucleosides. Utilising this novel pathway, synthetic access to all canonical deoxyribonucleosides was achieved in one pot. The reactions are mild and highly prebiotically plausible. The implementation of multiple steps was avoided as well as the need for pure deoxyribose.

By this approach, exclusive β-stereo- and furano-regioselectivity was achieved. Given its outstanding selectivity, the presented pathway to deoxyribonucleosides is considered the most plausible. By changing the sugar-precursors, nucleoside predecessors were discovered. A possible direct progenitor of DNA-nucleosides named deoxyapionucleosides (DApiNA) was discovered. For verification, a conventional synthetic method to DApiNA-nucleosides is presented, which might be relevant in fields apart from prebiotic interest.

Another variation of the DNA-yielding pathway lead to a rational pathway to ribonucleosides, lacking impracticable and drastic variations of the synthetic conditions as previously applied. A concomitant DNA- and RNA-nucleoside synthesis was performed under the same conditions. Other than previously postulated, it is concluded that RNA- and DNA-nucleosides developed at the same stage of molecular evolution.

Employing a cut-set of starting materials facilitated a simple, one-step formation of cytosine and adenine. Further, the application of this novel sugar-forming reaction to other heterocycles provided a rationale for the key-step in prebiotic riboflavin synthesis.

The simple, but effective design of our novel sugar-forming reaction lead to the discovery of DNA- and RNA-nucleosides, its DApiNA-progenitor and synthetic access to riboflavin and is thus considered as highly prebiotically plausible.

Scientific contributions

Publications

J. S. Teichert, F. M. Kruse, O. Trapp, *Angew. Chem. Int. Ed.* **2019**, *58(29)*, 9944–9947.

F. M. Kruse, J. S. Teichert, O. Trapp, *Chem. Eur. J.* **2020**, *26(65)*, 14776–14790.

Poster presentations

J. S. Teichert, F. M. Kruse, O. Trapp, Science of Early Life 2019, Kloster Seeon, Germany;
Direct prebiotic pathway to the DNA world

J. S. Teichert, F. M. Kruse, O. Trapp, SFB 749 Meeting 2019, San Servolo, Italy;
Direct prebiotic pathway to the DNA world

F. M. Kruse, O. Trapp, Molecular Origins of Life, CAS Conference 2018, Munich, Germany;
Time resolved investigation of complex prebiotic mixtures by LC-MS/MS analysis

Contents

Abbreviations

acetyl-CoA	Acetyl Coenzyme A
AICA	4-Aminoimidazole-5-carboxamide
AICA	4-Aminoimidazole-5-carboxamide
AICIA	5-amino-1*H*-imidazole-4-carboximidamide
AICN	4-Aminoimidazole-5-carbonitrile
ATCTA	4-aminothiazole-5-carbothioamide
ADP	Adenosine Diphosphate
AMP	Adenosine Monophosphate
ATP	Adenosine Triphosphate
DApiNA	Deoxyapionucleic Acid
DNA	Deoxyribonucleic Acid
DIBAL-H	Diisobutyl Aluminum Hydride
EIC	Extracted Ion Chromatogram
FAD	Flavine Adenine Dinucleotide
GTP	Guanosine Triphosphate
HILIC	Hydrophilic Interaction Chromatography
HPLC	High-Performance Liquid Chromatography
HRMS	High-Resolution Mass Spectrometry
NADH	Nicotinamide Adenine Dinucleotide
NMR	Nuclear Magnetic Resonance
mtDNA	Mitochondrial DNA
QTOF	Quadrupole Time of Flight
RNA	Ribonucleic Acid
RP	Reversed-Phase
TNA	Threose Nucleic Acid
tRNA	Transfer Ribonucleic Acid
T	Temperature
UDP	Uridine Diphosphate
UPLC	Ultra-high Performance Liquid Chromatography
UV/VIS	Ultraviolet-Visible Light

1 Introduction & Aim

> 'Life cannot have had a random beginning. The trouble is that there are about 2000 enzymes, and the chance of obtaining them all in a random trial is only one part in 10^{40000}, an outrageously small probability that could not be faced even if the whole universe consisted of organic soup.'
>
> *Fred Hoyle (Evolution from Space)*

Once in a while, I was asked the following question at casual events: *What is the topic of your Ph.D. thesis?* Answering that I am researching on the origin of Life, is always responded with stunning wide eyes. A facial expression of utter incomprehension results from the precise definition that I am investigating the prebiotic organic synthesis on bio-informational polymers.
Let's perform a mind experiment to get a conception about this rather complex topic. Imagine the early Earth roughly 4 billion years ago. What circumstances did rule Earth during that time? Popularly speaking, the Earth was a devastating place. The atmosphere consisted of nitrogen, water, carbon dioxide and hydrogen and their reaction products. Spark discharges by lightning, volcanic activity and bombardment by extraterrestrial meteorites were the order of the day. Luridly put, we must fill the gap and answer the question– *How did we get from 'dead' matter, 4 billion years ago, to us complex creatures now sitting at the bar?*

Examining the formation of complex creatures from single cells is already located a

few step too late in evolution. Usually, that is the transition for which it is asked for. We need to simplify this further, by going beyond the evolution of cells and examine processes at a molecular level. *How did photosynthesis come into play?*, *How did cell membranes develop?*, *How did this beautiful cellular machinery develop and start interacting?*–are questions to answer. Indispensable for Life is its hard drive, our genetic code. Oversimplifying, the admirable helical structure once evolved from gas–and it is the task of us researches to find out how.

The basic concept of Miller-Urey's *prebiotic soup* is familiar from school. However, in a first approximation, direct access *via* this approach to result in the polymeric species of DNA is still a level too complex. First, a rationale for the construction of the assembling lower-level building blocks needs to be found. The coding polymers of Life, RNA and DNA, are composed of small building blocks, called nucleosides commonly known by the letters A, T, G and C, for RNA it is U instead of T. Consequently, 4 billion years ago there must have been a pathway from these nucleosides to the coding polymers. A preceding pathway to construct these nucleosides from simple organic molecules must have existed. Nucleosides can be separated into two parts, a nucleobase and a sugar, which were so far synthesised in two distinct environments. As we now almost reached the bottom of the molecular level that is exactly the research topic.
How can we assemble nucleosides from small building blocks under conditions present on early Earth?

Relating to this challenge, modern Life is peculiar. All sugars can be present in two different configurations. The left-handed, and the right-handed configuration. In all living organisms, however, there is just one, the right-handed D-sugar, present. The same principle applies for amino acids, where just the left-handed L-species is present in modern biology. *What incidence led to the homochirality of our world?*–is the questions to answer.

In a nutshell, answering these fundamental questions first, we need to outline possible scenarios from the molecular level to more complex structures, explaining the evolution of Life. Having synthetic access to nucleosides, these need to be phosphory-

lated to oligomerise, forming DNA and RNA. Circular–communicating, feedbacking, promoting– reaction cascades need to start, creating new materials, by consuming energy and starting material. To function properly, these cascades need to be set in a protected environment, resembling proto-cells. Rudimentary catalysts evolve, enhancing these circular reaction cascades, finally resulting in enzymes. Steadily evolving, a last universal common ancestor (LUCA) is achieved, inheriting key features of all three kingdoms of Life. Ever evolving and ever adapting to new environments we slowly get to modern biochemistry and organisms. This whole complex process is not a product of coincidence, but governed by the laws of Nature. Finding initial conditions and applying chemical rules, will lead to the reinvention of Life.[1]
At a certain point the examination of the origin of Life can be quite a philosophical task. However, at a molecular level it is, exploring the correct conditions to force a certain transformation, and that's chemistry.

This thesis aims to find chemical conditions on a molecular level, from the view of prebiotic chemistry, leading to the formation of deoxyribonucleosides (DNA), ribonucleosides (RNA) and cofactors. Generally, a holistic approach is pursued, finding not just single answers to single problems, but putting all answers to a bigger picture, by revealing the underlying chemical concepts. Finding comprehensive answers, should spark the light of perception to one of the last dark areas of knowledge of humanity – itself.

2 Related Work

Since the discovery of nucleic acids 150 years ago,[2] major achievements have been made in addressing the challenging scientific questions about the genetic code.[3–14] However, the most tempting circumstance about the origin and the evolution of the genetic code remains unsolved. Especially, the key question on how nature managed to build up such intriguing complex molecules, able to encode structure and function from simple building blocks, is easily asked yet not answered. Over the past decades, promising synthetic concepts were proposed providing clarity in the field of prebiotic nucleic acid research.

In the following, the discovery of nucleosides is illustrated, followed by a brief delineation of possible prebiotic conditions and geochemical scenarios. Second, the conventional approach to nucleosides starts the demonstration of synthetic approaches. An introduction to prebiotic nucleobase synthesis is given, subsequently, all proposed pathways towards ribonucleosides are envisaged. Accompanying, research on flavin-based cofactors is illustrated. Last, recent advances in accessing deoxyribonucleosides are discussed, completed by synthetic access to apiose-nucleosides. The following illustration was earlier published in a similar form.[15]

2.1 The Discovery of Nuclein and the Resulting Investigations

In 1871, a small sentence in the essay 'Über die chemische Zusammensetzung der Eiterzellen' by Miescher indicated the discovery of *nuclein* from white blood cells and sparked the nucleic acid research.[2] The *nuclein* was discovered to be a substance containing nitrogen and being very rich in phosphorous. Elucidation of the molecular structure of the *nuclein* was the main research interest, whereas major contributions

were made by Levene at the end of the 19*th* century.[16]

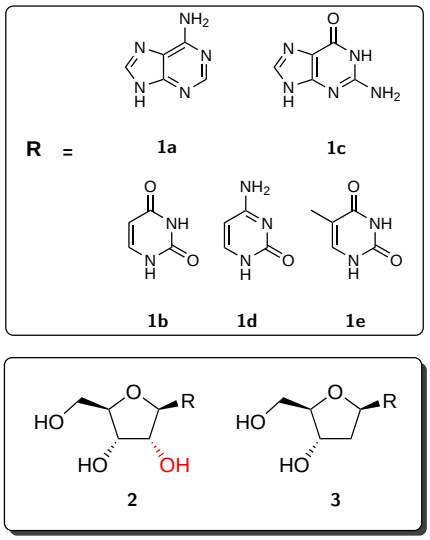


Figure 2.1: Comparison of RNA-nucleosides **2** with DNA-nucleosides **3**. In RNA mostly the bases **1a**–**d** are implemented, whereas in DNA **1a**,**c**–**e** are the only occurring nucleobases.

In vivo extractions resolved the several subunits of *nuclein*. It consists of a heterocyclic unit, linked to a sugar, which he called the nucleoside (see Figure 2.1). With a phosphorylated sugar implemented, he called it a nucleotide. Regarding the state of the art at that time, strikingly precise structures of the nucleosides and nucleotides were postulated. The sugar unit was identified as a pentose structure in 1909; however, it was not clear if it is a ribose **4a** or an arabinose **4b** unit, yet lyxose **4c** could be excluded, see Figure 2.2.[17]

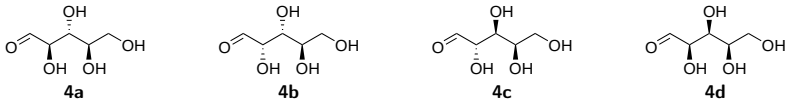


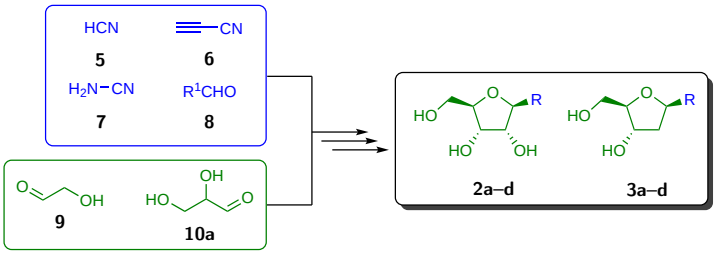
Figure 2.2: Structures of D-ribose **4a**, D-arabinose **4b**, D-lyxose **4c** and D-xylose **4d**.

At the beginning of nucleic acid research, the structure of the RNA and DNA nucleosides were elucidated,[4] while since the middle of the last century their prebiotic synthesis, in the context of the origin of Life, has been intensively studied. It was aimed to achieve the formation of the canonical ribonucleosides under prebiotic conditions.[18,19]

2.2 Definition of Plausible Chemical Prebiotic Conditions

The reaction conditions on the early Earth, leading to the emergence of Life, are far from strictly definable. The different constituents of the atmosphere and their partial pressure, the reducing state of the atmosphere and Earth's mantle, the water content and its aggregate state are under intense discussion. Consequently, the hypotheses vary to a large extend. The current, cross-disciplinary opinion is that early Earth's atmosphere was lacking oxygen.[20–23] Different geological scenarios are exerted, accordingly, the constituents of the atmosphere vary. A state described by geologists is the term *oxygen fugacity* or the ability to oxidise atoms. A fugacity similar to the present Earth is needed, when minerals are used, which are delineated in the pathways, presented in the following.[24] The oxygen fugacity is mainly influenced by the gravitational collection of siderophilic element in Earth's mantle. This is proposed to have happened during the moon forming cosmic impact.[25] Afterwards, the redox state of the atmosphere is set by fumigating constituents of Earth's mantle.[26] Most agree that the main constituents of the early atmosphere of Earth were N_2, H_2O and CO_2.[27,28] According to the reduction state, little H_2 and CO was present.[26,27,29,30] Photochemical studies suggest that NH_3 and CH_4 are not contributing to the atmosphere, as they decompose under UV light.[31] A high amount of H_2O, in its liquid form, was supposed to be present on early Earth.[32–34] This is a small introduction to the geological environment presumably present on early Earth, a detailed view on the conditions and possible scenarios is provided by Kitadai *et al.*[35] Liquid H_2O suggests a temperature between 0–100 °C, which is comparatively warm to the fact that the suns activity must have been 25% lower than today. Insulating greenhouse gases, like N_2O and CO_2, could have prevented Earth from cooling down significantly below 0 °C.[36] One explanation to describe this phenomenon might be NO_2, formed by the influx of solar wind, accompanied by hydrogen cyanide **5** and substituted acetylenes **6**.[37,38]

Both are prominent starting materials towards heterocyclic compounds and in combination with other compounds possible precursors of nucleoside formation. (Scheme 2.1, blue box) Some researchers do have different views on the term of *prebiotic conditions*, which are explicitly emphasised in the pathways presented below.



Scheme 2.1: An exemplary pathway towards ribonucleosides (R=**1a**–**e**), from simple starting materials derived from gas-phase reactions (blue) and sugar-forming reactions (green).

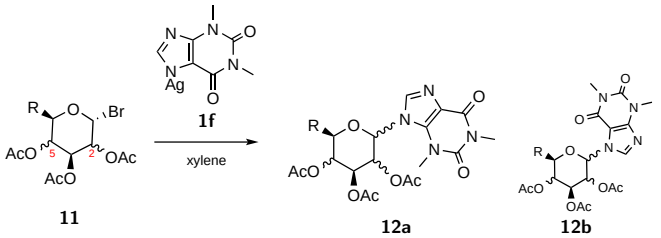
On a global scale, different geochemical scenarios are postulated in which Life could have emerged. In the following, these scenarios are briefly discussed, concerning their restrictions in the synthesis of prebiotic compounds. One could imagine that life emerged in: (i) the *free ocean*, (ii) a *dry lagoon*,[39]

i This model is commonly known as the *prebiotic soup*. The model is described by low concentrations and therefore low reaction rates. In consequence, high destruction/hydrolysis rates make it an implausible scenario for a global scale synthesis of prebiotic precursors. Reactions in such a scenario are driven by thermodynamic rules, leaving no space for far from equilibrium conditions. One backdoor is the freezing of the solution. These advantages are depleted by a vicious cycle of even lower reaction rates.

ii High concentrations are possible in a dry lagoon model or *warm, little pond*, leading to high reaction rates. When aiming for effective phosphorylation and condensation procedures, such a scenario is prominently resorted to. The concentrations need to be achieved by physical enrichment. Unfortunately, most of the starting materials, proposed in the synthesis of heterocyclic compounds, are far too reactive or volatile to reach concentrations at this level. Confinement of space permits the synthesis of some precursor molecules. However, the liberation to the free ocean, finally mixing two distinct environments, depletes the advantages. The construction proceeds at a micro-scale, whereas the destruction happens at a global scale. In consequence, a sufficient number of such geological scenarios needs to be connected in series to achieve the amounts needed on a global scale. Disadvantages are the over-constructive sequential arrangement of such micro-scale reactors.[40]

2.3 Conventional Ribonucleoside Synthesis

At first, conventional synthesis enabled structural elucidation of the nucleosides. Contemplation on nucleosides in the light of Life's origin was developed later. Ribose **4a** and deoxyribose **d4a** were soon discovered to be structural linking moieties in RNA[41] and DNA.[42] Recognising the biological relevance of the structural integrity, synthetic procedures accessing the nucleosides synthetically, were searched for. Still commonly taught in undergraduate study programs and experimentally executed, the disconnection to synthesise nucleosides is set at the glucosidic bond (see Scheme 2.2). This retro-synthetic approach was provided by Fischer synthesising several non-canonical nucleosides under non-prebiotic conditions.[43,44]



11a : R = CH_2OAc, **11b** : R = CH_2OAc (5*R*), **11c** : R = H (2*S*,5*S*)

Scheme 2.2: Conventional synthetic experiments leading to nucleosides,[45] derived from acetobromoglucose **11a**, acetobromogalactose **11b** or acetobromorhamnose **11c**.

Simplifying the system to a minimum, the xanthine derivative theophylline **1f** was studied. **1f** was converted to its silver salt and reacted with the corresponding haloaceto sugar **11** to give the desired nucleoside **12**. The transformation was executed in hot xylene. In reaction with the heterocycle, a glucoside, a rhamnoside and a galactoside were achieved, starting from acetobromoglucose **11a**, acetobromogalactose **11b** and acetobromorhamnose **11c**.[45] Although the sites of the nucleophilic attack were already restricted, a prediction of whether the substitution takes place at *N*7 or *N*9 of the purine ring, could not be made. Further, the connectivity at the anomeric centre could also not be addressed.[45] As **1f** was converted to caffein in silver assisted methylation experiments, it was assumed that *N*7 is connected by the glucosidic bond. The procedure was successful for purine analogues, but not for the corresponding pyrimidine nucleosides. Following this procedure, significant progress was made according to the state of the art, however, access to the canonical nucle-

osides **2** could not be provided. The synthetic procedure for all purine derivatives was provided by Traube.[44,46] Refinement of the presented reaction conditions and more advanced technical equipment resulted in the full structural elucidation of all canonical ribonucleosides.[3,47]

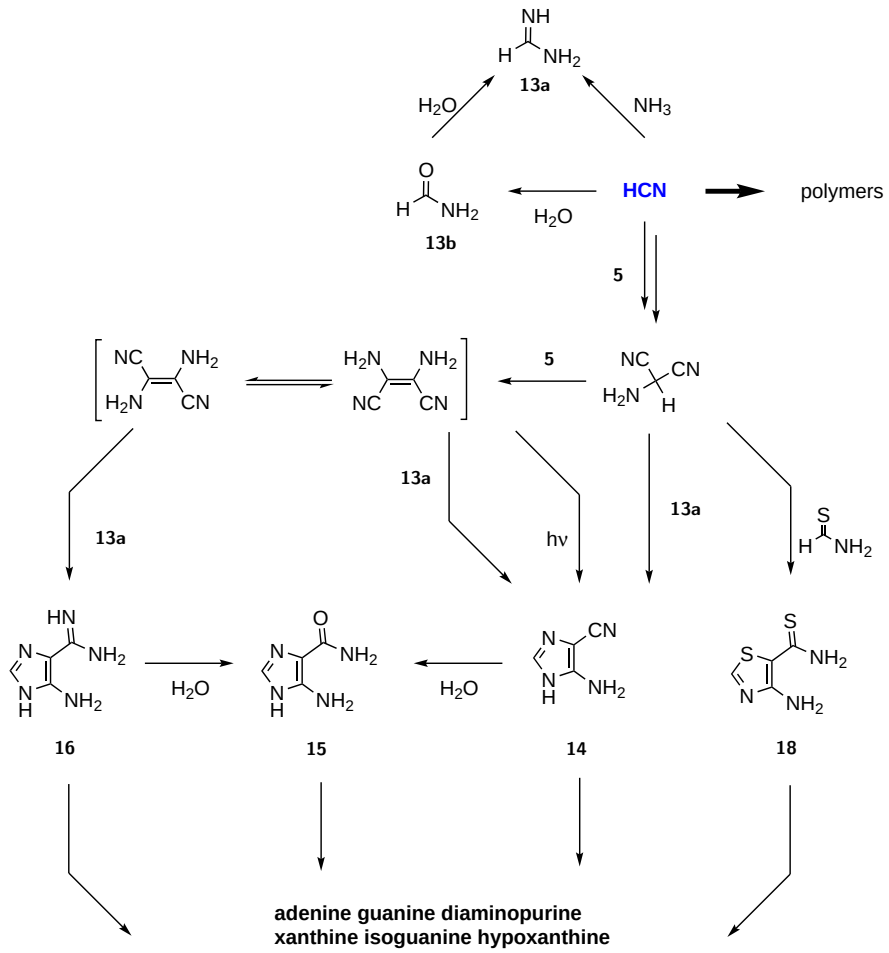
In contrast, purine-*N*9 turned out to be connected to the anomeric centre of **4a**. Moreover, it was determined that all canonical purine and pyrimidine nucleosides are β-configured.[48] The break-through procedure was achieved by implementing 2,8-dichloroadenine and α-acetohalogenoxylopyranose alongside the development of the reaction conditions.[49,50] Thus, access to adenine- and guanine-derived nucleosides were given.[48,50] Succeeding in a procedure for **2d**,[51] a conventional synthetic pathway towards all canonical nucleosides **2a**–**d** could be outlined.

2.4 Prebiotic Nucleobase Synthesis

In the meantime, other scientists pursued the fascinating question of how biological molecules could have formed under the conditions on early Earth.[52,53] Great effort was made delineating prebiotically plausible pathways towards the canonical ribonucleosides **2** and deoxyribonucleosides **3**. First, considering the assembly of the nucleobases from simple organic molecules seemed promising. The most prominent pathways are presented in the following.

Among many interesting insights provided by spark-discharge experiments,[54,55] free nucleobases were found to be key compounds in the gained reaction mixtures. For example, the purine nucleobases adenine **1a** and guanine **1c** were detected in small quantities. Unfortunately, the production of cytosine **1d** under these conditions could not be verified.[56] This leads to some controversial discussion about the presence of **1d**, which is discussed at a later stage of this thesis. Confirming the relationship of **1** and **2**, the hydrolysis of **2** in hot, alkaline solution gave **1a** as decomposition product, which indicates how sensitive potential prebiotic systems are.[53] Spark-discharge experiments provided a good start on the formation of purine nucleosides, however not satisfactory.

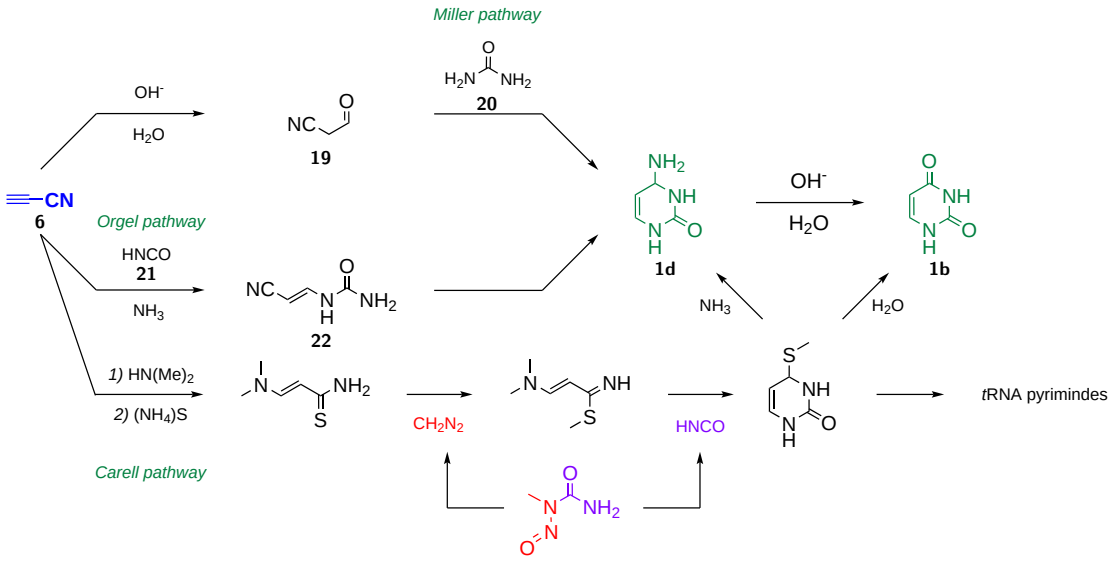
Studying the solution chemistry of **5** in detail, see Scheme 2.3, revealed a promiscu-



Scheme 2.3: Prebiotically plausible pathways, starting from **5** towards the purine nucleobases[57–60]

ous network of **5**-oligomers and hydrolysis products **13**, leading to purine nucleobases *via* several heterocyclic intermediates, namely AICN **14**, AICA **15**, AICIA **16** and ATCTA **17**.[57]

In contrast to the purine nucleobase synthesis, access to the canonical pyrimidine nucleobases was not achieved as easily. The synthesis (*cf.* Scheme 2.4) is mostly proposed to start from cyanoacetylene **6**, which is a possible product of the discharge reaction between CH_4 and N_2. In this spark discharge reaction, **6** is reported to be the second most abundant product after HCN **5**.[58,61] The syntheses towards the canonical pyrimidine nucleobases are mainly based on the two starting molecules, **6** and its hydrolysis product cyanoacetaldehyde **19**.



Scheme 2.4: Cyanoacetylene **6** as a starting point for pyrimidine nucleobase synthesis[39,58,62]

In an early approach to pyrimidine nucleobases, **6** is reacted with cyanic acid **21** to form cyanovinylurea **22**, (*cf.* Scheme 2.4, middle pathway) which subsequently cyclises intramolecularly to **1d**.[58] Another route proceeds via **19**, reacting with urea **20** to form cytosine.[39]

Major doubt against these courses of action is the high susceptibility of **6** to nucleophilic destruction, present in the early Earths atmosphere.[58,63–68] The hydrolysis rate of **6** to **19** has a half-life of 11 d (30 °C, pH = 9.[69]) Under the same conditions hydrolysis of cyanogen **23** proceeds with a half-life of 30 s.[70] This leaves a small time window to react with the nucleophile of choice.

Directly starting from a molecule of lower reactivity, the conversion of **6** with **20** is reported to yield **1d**.[58] In this case, a lower reactivity of **20** needs to be compensated by a higher concentration. Yet **20** is easier available and replenishable than **21**.[71] **6** and **21** on the contrary, are not likely to have accumulated at high concentrations.

In this context, cyanoacetaldehyde **19** has a slightly lower reactivity ($t_{1/2}$ = 31 yr[69]).[39] It was reported that **1d** could be gained in yields of 30–50% from **19** in a highly concentrated solution of urea **20**. Yet facile, the synthesis needs to be stopped after five hours,[39] as **1d** rapidly hydrolyses to **1b**.[72–75] In the laboratory, such highly confined scenarios are easily applicable, however, the translation to an early-earth scenario

reveals the improbability of the chosen setting.

As discussed above, the highly electrophilic properties of **6** are commonly questioned. A recent synthesis of **1d** uses the high electrophilic reactivity of **6**,[62] trapping it as a α,β-unsaturated thioamide (see Scheme 2.4, lower pathway). In an aqueous solution, cyanoacetylene **6** quickly reacts with dimethylamine to form the corresponding acrylonitrile. Subsequent reaction with ammonium sulphide furnishes the α,β-unsaturated thioamide. Despite the presence of multiple nucleophiles, the reaction is reported to proceed in high yields, utilising a slow addition of **6**. Fortunately, the α,β-unsaturated thioamide crystallises upon cooling the reaction. Methylation and carbamoylation of the α,β-unsaturated thioamide with *N*-methyl-*N*-nitrosourea gives *S*-methylpyrimidinone. The latter can be transformed to **1d** (*aq.* NH_3, $FeSO_4$) or **1b** (H_2O, $FeSO_4$) offering different nucleophiles. Non-canonical modified nucleobases present in modern tRNA are accessible as well. It is reported that a one-pot formation of the *S*-methylpyrimidinones is possible, starting from **6**. However, sequential addition of each of the reagents is necessary. Although elegantly outlined, this pyrimidine pathway also suffers from confined laboratory conditions, not transferrable to a global prebiotic scenario.

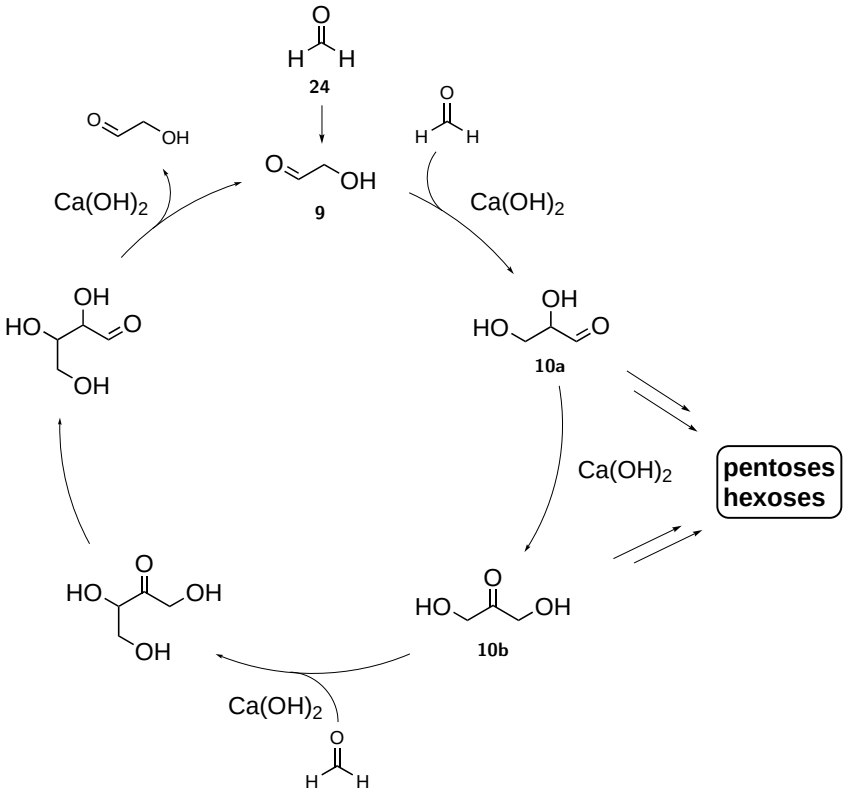
Apart from the complicated construction, any **1d** synthesis suffers from facile deamination to give **1b**. The rate of spontaneous deamination thereby exceeds the rate of the **1d** production. At every pH, the deamination occurs fast.[76] The lowest conversion takes place in the pH range of 6–9. The rates for deamination even are equivalent for **1d**, **2d** and **2d** in single-stranded DNA.[74,77,78] Only double-stranded DNA is less prone to deamination. In Nature, the deamination is balanced by the repair enzyme *uracil-DNA glycosylase*.[79] Additionally, **1d** is readily destroyed to its photohydrate under UV light. In conclusion, a possible synthesis is more feasible in the dark.[80]

The above-presented syntheses imply doubts about the prebiotic availability of **1d** on early Earth. Further evidence is given by the fact that **1d** was not found in discharge experiments[55] and on meteorites. **1d** formation is both a problem of synthesis and stability. There is no robust system that produces **1d** at concentrations nowadays present on earth. Major drawbacks of the syntheses are the high reactivity of volatile

species, which are needed in high concentrations. Additionally, an experimental setup is needed, which is strongly dependent on confined laboratory settings. The above-presented geochemical scenarios, see Section 2.2, are not suited for a comprehensive formation of **1d** either. In a *global ocean* scenario, the deamination rate exceeds the production rate by far. In a *dry lagoon* scenario, **1d** formation could have happened, however, release into a dilute basin would deplete the positive effect of the dry down. The last described scenario can not be ruled out, but distribution on a global scale is not inevitably given.

Based on the above scenarios, it is concluded that the formation of **1d** and its existence at the advent of Life is highly unlikely. However, a tiny loophole-scenario is delineated,[81] in which the transition from the starting materials to **1d** is readily intercepted by a reaction forming nucleotides or oligonucleotides. These assemble instantly to stable RNA or DNA-double strands. This would significantly lower the rate of deamination. The transition should proceed in the same environment, in combination with water, at a temperature of 25 °C. No pH and temperature change should be needed and the production should be facile from atmospheric components under the assistance of minerals.

Explaining the origin of organic heterocycles is far easier than describing a pathway to the sugar moiety. A commonly accepted theory is that sugars originated from the formose reaction.[82,83] Starting from formaldehyde **24** and glycolaldehyde **9**, the formation of complex linear and branched sugars could be illustrated, see Scheme 2.5. However, the base-catalysed, autocatalytic reaction network produces a complex mixture, rather than selectively sugars of choice. Important products of the formose reaction are glyceraldehyde **10a** and its isomer dihydroxyacetone **10b**. Although desperately needed in most of the nucleoside pathways, the high yielding formation of ribose **4a** could not be achieved so far. Experiments to stabilise **4a**, prolonging its lifetime, are based on the *cis*-stabilisation with borate minerals.[84]



Scheme 2.5: The supposed catalytic cycle of the Formose Reaction.[82,83]

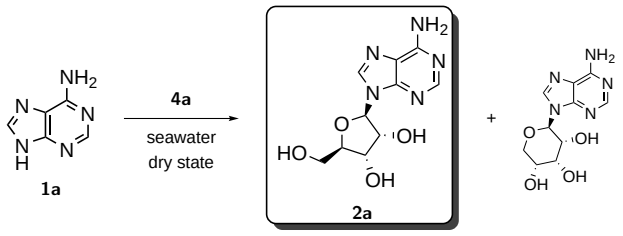
2.5 Prebiotic Nucleoside Synthesis

The discovery of plausible pathways towards **1** under prebiotic conditions led to the investigation of various pathways towards pyrimidine and purine nucleosides. The performed retrosynthetic strategy was almost exclusively based on the disconnection at the anomeric bond, introduced by Fischer in 1909.[45]

2.5.1 Early Nucleoside Syntheses

The syntheses of purine nucleosides succeeded in low yields under UV radiation of a solution of **1a** and **4a**.[85] Repeating these experiments[86] revealed a low reproducibility. The improved synthesis under hot, dry conditions yielded the formation of **2** from **1a**, **1c** and **4a**.

Mimicking the dry down of a sea-water pool, the reaction was studied in synthetic and natural sea-water and the solid-state. Both were assisted by inorganic salts,



Scheme 2.6: Dry state synthesis of ribonucleosides supported by sea water salts.[86,87]

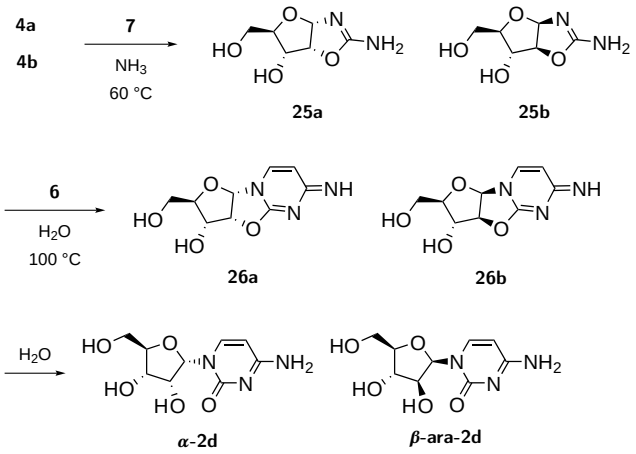
see Scheme 2.6. Drying down the reaction, conditions utilising $MgCl_2$ and $(NH_4)HBO_3$ performed the best. Further, the ribosylation of the primary amino group is reported, which can be converted to the canonical riboside by hydrolysis in neutral to alkaline aqueous media.[88] Direct access to ***β*-2a**,**c** (10% yield) is achieved from free **4a** and **1a**,**c** under the assistance of magnesium-ions or polyphosphates. It is reported that the reaction does not proceed under the assistance of montmorillonite and other clays.[86] Depurination of the formed ribonucleoside and incompatibility of the synthesis to **1b**,**d**,**e** turned out to be the major complication of the procedure. As pyrimidine nucleosides do not hydrolyse under acid catalysis, the corresponding glycosidation is concluded to not take place under these conditions.[86]

The presented syntheses of **2** inherit two major problems. First, coupling **1** to the pentose sugar **4a** is feasible, but not very efficient. This is caused by the instability of **4a** itself. It is the least stable sugar, as it decomposes roughly four times faster than its related pentoses and 16 times faster compared to the hexoses. Expressed in kinetic data, the half-life is three hours at a pD of 10.2 and 55 °C ($k = 7.0 \cdot 10^{-5} s^{-1}$).[89] These depict the conditions of the formose reaction, from which **4a** is supposed to originate.

Second, the intrinsic kinetic barrier of the bond-forming reaction between **4a** and **1** is too high. Accordingly, prebiotic access to **2** in the described manner is not feasible and a different disconnection rationale needs to be found.

Concluding, the critical disconnection of the anomeric bond should take place within a series of reactions or at the very beginning of a pathway, not in the last step. Therefore, heterocyclic starting materials or their precursors with a higher nucleophilicity are needed. A brief overview of the accomplishments in this field is given, using **4a** and cyanamide **7** as the initial compounds of interest, see Scheme 2.7.

Favouring the C-N connection at the beginning, a new pathway produces ***α*-2d** from



Scheme 2.7: Prebiotic pathway towards pyrimidine ribonucleosides, giving a mixture of the non-canonical stereoisomers **α-2d** and **β-ara-2d**, via the newly discovered class of sugar aminooxazolines **25**.[90]

4a or **β-ara-2d** from arabinose **4b**.[91] An important intermediate structure is the newly discovered class of aminooxazolines.[91] The reaction preferentially proceeds with **4a**, however other sugars form their corresponding aminooxazolines as well. The ribo-aminooxazoline **25a** crystallises from a cold aqueous solution and is stable therein, while all possible other sugar aminooxazolines remain in solution. This novel molecular class provides a storage form of **4a** *via* a chemical enrichment step. **25a** thereby decomposes 70 times slower than free **4a**. Advantageously, cyanamide **7** reacts more than 200 times faster ($1.5 \cdot 10^{-2} s^{-1}$) with **4a**, than its decomposition ($7.0 \cdot 10^{-5} s^{-1}$).[89] A subsequent reaction[92] with cyanoacetylene **6** leads to the riboanhydronucleoside **26a** and the arabinoanhydronucleoside **26b**. The terminal step of the pathway is the hydrolysis of **26**, giving **α-2d** and **β-ara-2d**. In the case of **α-2d**, further hydrolysis leads to **α-2b**

Although providing a first prebiotically plausible pathway to ribonucleosides, these are not obtained in the correct stereochemical configuration as the main product is **α-2d**. The canonical nucleoside **β-2d** can be gained in a low yield of 5% after treatment with UV light ($\lambda = 253\,nm$).

In a later examination, the synthesis of the α- and β-anomers of **2b** and **2d** was comprehensively developed.[91] The main objective was a more productive formation of **26**, accompanied by improvement of the reaction conditions. After purification

of **25a**, the reaction with **6** was continued in *N*,*N*-dimethylacetamide, leading to the formation of different acrylonitrile byproducts. The addition of water or an aqueous solution of ammonia yielded **β-ara-2d** and **α-2d**.[91] Later, this possibility of approaching pyrimidine nucleosides[90] was further developed and enhanced,[93] while still relying on the basic concepts.

The synthetic pathways towards the purine nucleosides were not examined for a long period. This is in contrast to their natural abundance, for example in prominent purine cofactors and RNA/DNA polymers. The synthesis leading to purine nucleosides was mostly executed by reacting free **1a** as its hydrochloride salt and free **4a** in a dry molten state.[88] This approach, however, produces complex reaction mixtures and very poor yields with respective to **2a**.[52,53] In consequence, a logical approach to the purine nucleosides was urgently required.

2.5.2 The Aetiology of the Nucleic Acid Structure

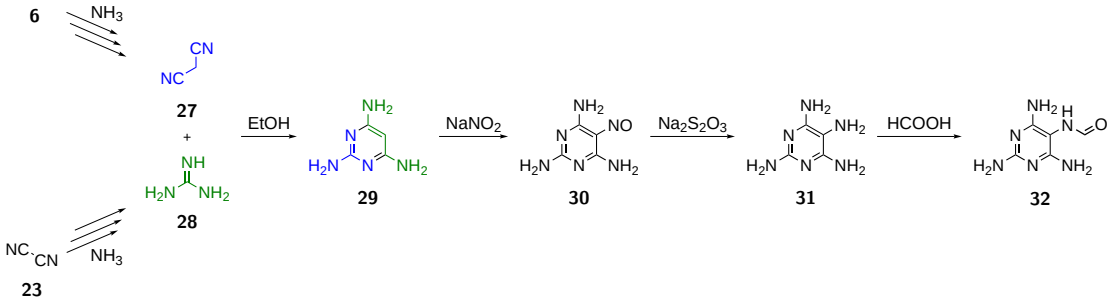
The work described in the following chapter is not directly linked to the prebiotic synthesis of ribonucleosides, but it can not be left unmentioned. Pioneering thoughts and leading impulses on the field of prebiotic chemistry was provided by the research group of Eschenmoser. Besides many fields of interest, one main question was the specific implementation of pentose sugar into the coding biopolymers of Life. Subsumed by the question –Why pentose- and not hexoseribonucleosides?– a model system of homo-DNA was postulated and investigated. By varying the sugar backbone, answers on the Watson-Crick[94] pairing principles and the variation of bonding interactions could be provided. Thereby, a reason for the chemical selection of ribofuranose over all other possible sugars was illustrated.[95–100]

Further research topics within the field of the evolution of RNA and DNA are the postulation of TNA,[101] research on the chemical evolution of cofactors and the chemical aetiology of nucleic acids.[102–104]

Reflecting on the evolution of the nucleic acid structure, an evolution of RNA was proposed. Preceding RNA resembling structures are outlined. These differ from the canonical polymer by variation of the heterocyclic moieties, sugars and linking groups. During evolution, it is proposed that these were ruled out due to chemical selection.[96]

The above-presented approaches serve a holistic comprehension of the molecular evolution of coding polymers.

Besides, a plausible formation of basic heterocyclic structures from simple gas-phase constituents is delineated and linked to the prebiotic synthesis of **2**-nucleosides.



Scheme 2.8: Synthesis of tetraaminopyrimidines starting from **5**-derived precursors.[104]

Ubiquitous pathways from simple organic molecules to complex heterocycles were examined applying strict rules. These rules are: *1) No molecular oxygen, 2) No water, 3) Derivatives of cyanogen, cyanoacetylene and ammonia as precursors, 4) Heat, 5) Monomolecular reactions or bimolecular reactions with one ubiquitous reaction partner;*[104]

Utilising solely HCN **5**-, cyanoacetylene **6** and cyanogene **23**-derived precursors, plausible intermediates in the formation of amino acids, nucleobases and cofactors could be pictured. An important junction molecule towards both nucleobases and cofactors is the class of triaminopyrimidines **29**.

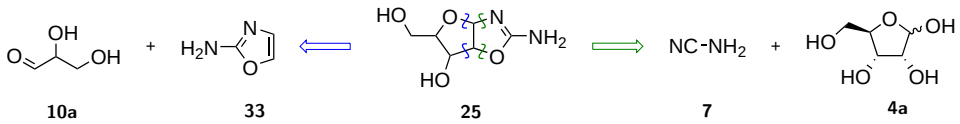
The **23**-descendant guanidine **28**, and the **6**-descendant malononitrile **27**, can be condensed to gain **29**. Subsequent acidic nitrosylation gives nitrosopyrimidines **30**, see Scheme 2.8. The reduction of these with sodium thiosulfate leads to tetraaminopyrimidine **31**. Further, pteridine derivatives are accessible by the facile oxidation of free **31**.

Formamidopyrimidnes **32** are key molecules in prebiotic synthesis (*cf.* Section 2.8.3), which are obtained after heating **31** in formic acid. Upon melting **32**, a plethora of canonical and non-canonical nucleobases is accessible.[104] Refinement of these structural ideas by the Carell group[105] sketched a consistent pathway towards purine ribonucleosides **2a**,**c**.

The so far applied retrosynthetic analysis posing the disconnection at the glyco-

sidic bond in the last possible step is described by Eschenmoser as *the notorious nucleosidation problem*,[1] which needs to be solved.

2.5.3 Sutherland Approach



10a **33** **25** **7** **4a**

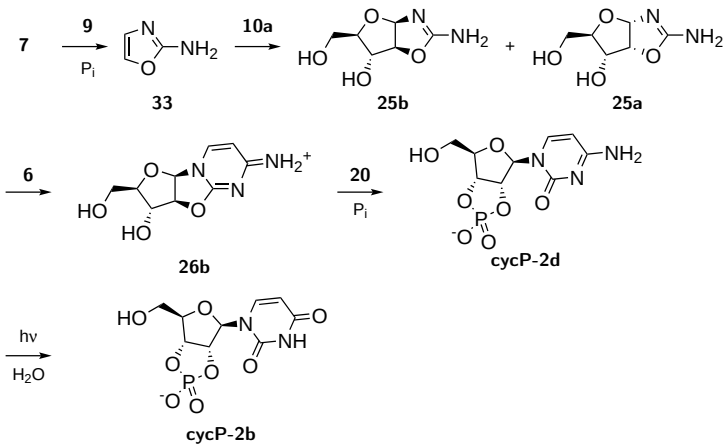
Scheme 2.9: A different disconnection approach of sugaraminoxazolines **25a** to overcome the intrinsic problem of nucleosidation.

The pathway earlier presented by Sanchez and Orgel[90] did provide a new way of approaching the nucleosidation problem, see Section 2.5.1. Sutherland[18] did subtle embark on this task, modifying the pathway of Orgel by introducing phosphate as a general acid-base catalyst. This significantly increased the yield of **25**.

In contrast to the previously presented pathway, the critical use of **4a** is avoided. Achieving a comprehensive approach, **25** is cut differently to gain **10a** as a C_3-synthon and 2-aminooxazole **33** as a fused heterocycle, see Scheme 2.9. Preceding, **33** is achieved by the condensation of glycolaldehyde **9** and **7**. In comparison to **4a**, **9** and **10a** are produced in a higher quantity under formose-reaction conditions and are more stable.[82,83]

The prebiotic origin of **7**, however, is far from being resolved. Speculations on its source are found in interstellar nebulae and ices.[106–109] Delivery to the primordial Earth might have occurred via meteorites, asteroids and comets.

The experimental procedure is outlined as described in the following. A quantitative formation of **33** is achieved by fusing **9** and **7**. The subsequent reaction of **33** with **10a** produces a mixture of ribo-**25a** (25%), arabino-**25b** (15%), lyxo-**25c** (6%) and xylo-**25d** (4%), see Scheme 2.10. Additionally, the hydrolysis products of **25** are reported.[110] Coincidentally, **25a** crystallises with an enantiomeric excess of 60%, after cooling the reaction mixture to 4 °C. This physically enriches one stereoisomer, yet **25a** is in the opposite stereochemical configuration to proceed to the nucleotide in the canonical β-configuration. To achieve that, a procedural method to the canonical configuration is needed, circumventing the presented problems.



Scheme 2.10: A pathway towards the pyrimidine nucleotides **cycP-2d** and **cycP-2b** via stereochemically enriched arabinose-aminooxazoline **25b**, originating from **7**, glycolaldehyde **9**, glyceraldehyde **10a** and **6**.[18]

The easiest access is provided by a phosphate-buffered equilibrium of **25a** and **25b**, involving a furanose ring opening. The equilibrium, however, is on the side of the insoluble **25a**.[110]

In analogy to the previously presented pathway, see Section 2.5.1, the corresponding anhydronucleoside **26b** is obtained after the phosphate-buffered reaction with **6**. Though, one can scrutinise the described mechanism of the **6**-addition.

As the alkyne moiety of cyanoacetyene **6** is highly electrophilic, a direct nucleophilic attack of the free amine moiety at the triple bond is supposed to be more likely than a conjugate addition of the inner cyclic nitrogen. This is confirmed by the fact that nucleophilicity of the free amine moiety (N~13)[111] should be comparatively higher than the inner cyclic nitrogen(N~9).[112] Assuming a direct nucleophilic attack, a different anhydronucleoside is achieved. Downstream, a nucleotide might be formed which can't be distinguished from the canonical one by NMR spectroscopical methods.

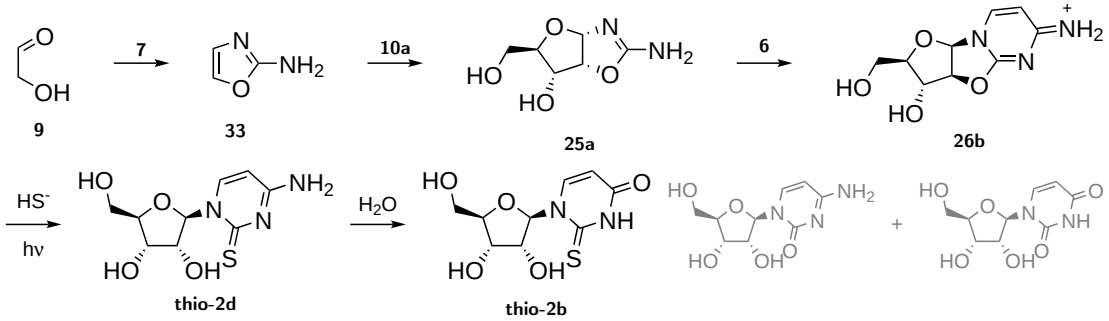
Following the course of the action, the unbuffered reaction of **25b** yields the non-canonical arabinose derivative of **2d**. Therefore, a distinct phosphorylation step was implemented to achieve the correct nucleotide.

Evaporating the reaction mixture containing **26**, inorganic phosphate and urea **20** leads to the selective phosphorylation of the 3'-hydroxy group.[113] The formation of a cyclic phosphate via an intramolecular attack changes the crucial stereochemistry of

the nascent 2'-OH. The pyrimidine ring is released in the correct β-stereochemistry, giving the cyclic phosphates of cytidine **cycP-2d** and uridine **cycP-2b**. In a partial hydrolysis step, the phosphate is cleaved, receiving the desired nucleoside **2b**.[18]

Another feature of the reported reaction series is the destruction of all unwanted side products by radiation with UV light (λ = 248 nm), leaving **2d** and **2b** in equal proportions. This photoanomerisation is just applicable to the 2',3'-cyclic nucleotides.

Although elegant, the pathway relies on the water-soluble, minor component **25b**, which is not as easily accumulable as **25a**. However, directly starting from **25a** is not feasible, as the photoanomerisation of the corresponding ***α*-2d** results in its destruction. Further, the limited depth of penetration of UV light is another issue. These obstacles led to a new route, based on the thiolysis products of **25a** and **25b**.[114] This pathway is described in the following, see Scheme 2.11.



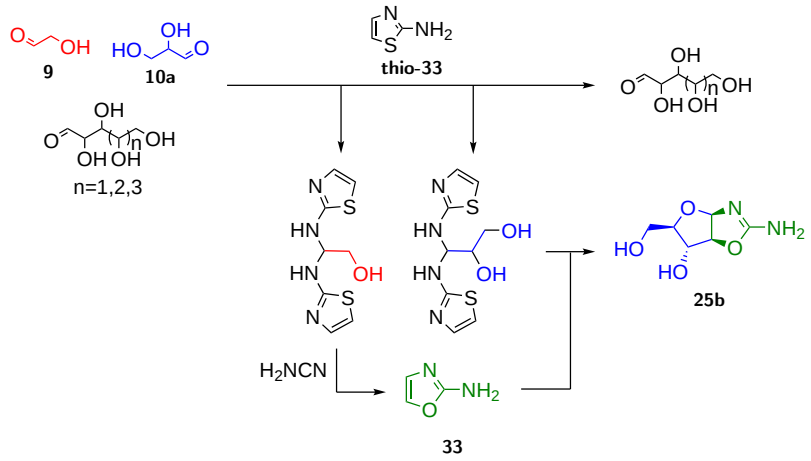
Scheme 2.11: An improved pathway towards **2d**,**b** via ***α*-thio-2d** utilising **25a** as an enrichable compound.[114]

The thiolysis of **26b** in aqueous formamide yields ***α*-thio-2d**. Instead of its destruction, photoanomerisation leads to ***β*-thio-2d** in 76% yield.

From either ***α*-thio-2d** or **thio-2d** hydrolysis leads to ***α*-thio-2b** and **thio-2b**. Finally, the hydrolysis in phosphate buffer (pH = 7) of **thio-2d** gives the canonical ribofuranoside **2d**, whereas a slight decrease in pH gives a significantly higher amount of **thio-2b**.

These syntheses towards pyrimidine nucleotides and nucleosides are elegant solutions to the glycosidation problem. In contrast, the outcome of the pathway is greatly influenced by the order of the sequential addition of reagents and thus dependent

on external interventions. In Nature, however, this scenario is difficult to control. To overcome the inaccuracy provided by sequential addition, a novel chemical scenario is proposed, based on the crystallisation of intermediate products. The crystals should guarantee inertness to avoid accompanying side reactions.[115]

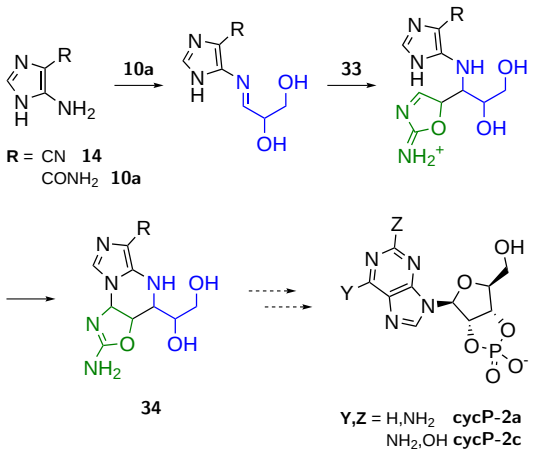


Scheme 2.12: Selective reaction and precipitation of thiazol aminals favouring **9** and **10a** over C_{4-6}-sugars on the way to pyrimidine nucleotides .[115]

The key concept of this procedure is the formation of stable aminals derived from 2-aminothiazole **thio-33**, see Scheme 2.12. In reaction with a sugar mixture, it was found that the aminals of **9** and **10a** crystallise selectively, resolving them as enantiopure compounds. Providing a separation mechanism, the aminals of the tetroses, pentoses or hexoses do not precipitate. From the **9**-derived aminal and the **10a**-derived aminal, the reaction cascade can proceed in an above-described manner. This reaction sequence depicts a possible one-pot pathway to solve the legitimate criticism of the sequential addition of reagents.[115]

So far, the pathways explaining the emergence of **2** considered the purine derivatives separately from the pyrimidine nucleosides. A uniform system leading to both classes of nucleosides would be desirable. Two distinct chemical environments producing the nucleosides separately, but resulting in one informational biopolymer, is statistically hard to argue. On the other hand, analysis of the modern biosynthetic pathway shows exactly two distinct procedural methods.[117]

Investigation on the sole and concomitant synthesis of purine nucleosides is moti-



Scheme 2.13: Proposal of a concomitant synthesis of purine and pyrimidine nucleotides derived from 2-aminooxazole **33** as a common precursor.[116]

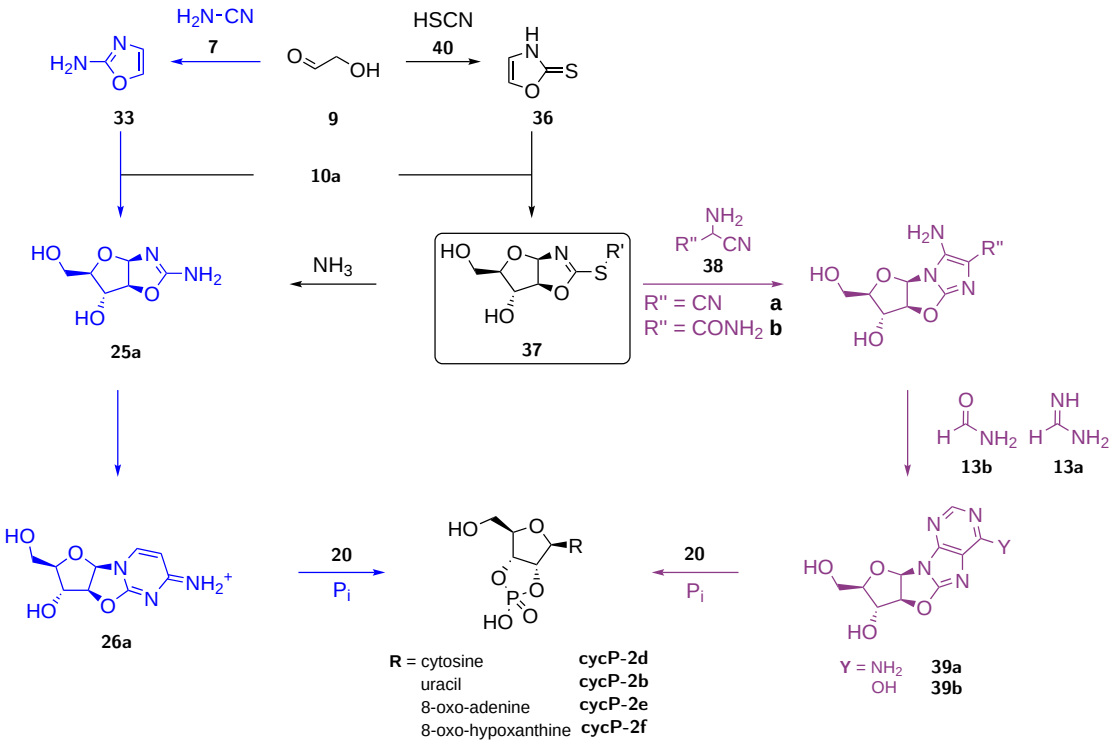
vated by the gained insights in the synthesis of pyrimidine nucleosides. One attempt to find a uniform pathway combines **33** and the intermediates of the **1a**-synthesis, AICN **14** or AICA **15**,[118] as central molecules in reaction with sugars.

A multi-component reaction is proposed, see Scheme 2.13, in which a fused heterocycle **34** is obtained from several aldehydes (**9**, **10a**, formaldehyde **24** and acetaldehyde **35**) via iminium ion formation with **14** or **15** and subsequent cyclisation. Several TNA or RNA precursors are reachable by variation of the aldehydes. The fused heterocycle **34** is reported to crystallise from an aqueous solution. A cyanovinylation-step and subsequent UV-induced loss of pyrimidine leads to the **15**-riboside. It is speculated that a ring closure furnishes the anhydro purine nucleoside, which can be converted to the purine nucleoside by urea-mediated phosphorylation.

However, none of the speculated steps is reported and the mere number of steps proves to be difficult to control from the point of chemical selectivity and prebiotic plausibility.[116]

A unified pathway leading to both pyrimidine and purine nucleosides is proposed.[18] It is based on the variation of the heterocycles of the above-described pathways, furnishing the 8-oxo-purine nucleotides **cycP-2e** and **cycP-2f**.

In contrast, glycolaldehyde **9** is reacted with thiocyanic acid **40**, instead of cyanamide **7**, to give an oxazolthione **36**.[119] The heterocycle creates oxazolidinone thione **37**, after



Scheme 2.14: A unified pathway towards pyrimidine nucleotides **cycP-2d** and **cycP-2b** and 8-oxo-purine nucleotides **cycP-2e** and **cycP-2f** from arabinose-2-aminooxazoline derivatives **37** as common precursor.[119]

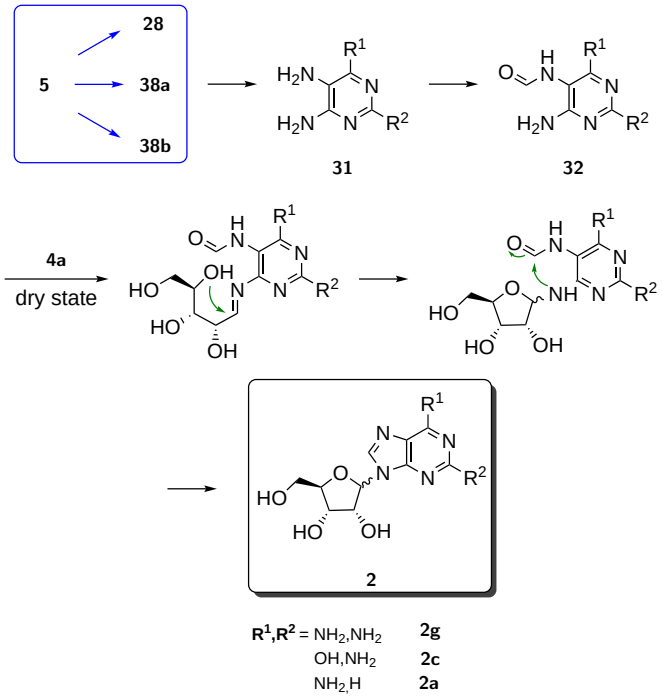
the reaction of **36** with **10a**. **37** serves as a key precursor leading to either the pyrimidine or purine nucleotides. Ammonolysis of **37** yields **25a**, proceeding to **cycP-2d** and **cycP-2b** in the described manner, see Scheme 2.14, left pathway. Whereas the reaction with 2-aminomalononitrile **38a** and 2-amino-2-cyanoacetamide **38b** (right side) furnishes the cyclic nucleotides of 8-oxoadenosine and 8-oxohypoxanthine **cycP-2e** and **cycP-2f**.

In detail, **37** is reacted with 2-aminomalononitrile **38a** and 2-amino-2-cyanoacetamide **38b**. Second, formamide **13b** and formamidine **13c** are implemented to give the anhydropurinnucleoside **39**. Changing the stereochemistry of the nascent 2'-OH in a **20**-mediated phosphorylation step, leads to the 8-oxo-purine ribonucleotides **cycP-2e** and **cycP-2f**.

The synthetic pathway suffers from sequential addition of all reagents, due to the high reactivity/instability of the reagents **6**, **7** and **40**. Further external change of

pH is crucial for a successful outcome, as it guarantees the stability of the formed compounds.

2.5.4 Carell Approach



Scheme 2.15: Prebiotically plausible route towards purine nucleosides: Dry state reaction of formamidopyrimidines **32** and D-ribose **4a** supported in alkaline borate media.[105]

A long searched prebiotic access to the canonical β-purine ribonucleoside is elaborated, relying upon and refining the previously illustrated pathway towards formamidopyrimidines **32**. **32**, which are ancestors of the **5**-regime, are hypothetically easily accessible from guanidine **28**, **38a** or **38b**.[102–104] Triaminopyrimidines **29** were intensively studied in terms of their formation, reactivity and characterisation.[104] Possibly inspired by Nature's salvage pathways, the reactivity of triaminopyrimidine **29** with formic acid and **13a** were studied, resulting in **32** as purifiable intermediate.[103] The latter is found many times in the biological breakdown of nucleosides, caused by oxidative stress.[120]

Recognising a convenient starting point for the bottom-up synthesis of the purine ribonucleosides, **32** was reacted with **4a**, in a dry state reaction, see Scheme 2.15.[105]

It was found that the proposed formylation[104] undergoes a high position selectivity towards the amine at C_4 or C_5. This results from the protonation of the inner-cyclic nitrogen, consequently leaving the exocyclic amino group *N*5 as the sole nucleophilic position.

Purification of **32** by precipitation in the cold, leaves a clean starting material. The subsequent dry-state reaction, furnishes a **4a**-imine, with one of the amines adjacent to the formamidine-moiety. Both amino groups of **32** are mirror-symmetrical to each other, reducing selectivity issues. The canonical ribofuranose configuration is gained after an intramolecular attack of the 4'-hydroxy group at the imine.

In a final step, the fused imidazole ring is closed by nucleophilic attack of the amine at the formamide moiety. Sodium borate is implemented as it is reported to assist the formation of **4a** and its *cis*-stabilisation.[121]

Performing the ribosylation in an alkaline environment poses the main disadvantages of the pathway as two incompatible environments need to be placed consecutively. Further, the synthesis lacks stereoselective control as the ribonucleosides are formed in a furanose/pyranose mixture both containing the α- and β-anomer.

Even more sophisticated, a multi-step wet-dry cycle procedure is deployed, concomitantly forming canonical and non-canonical ribonucleosides.[122] The wet-dry cycles guarantee precursor enrichment and thus increase the overall conversion to **2a**,**c**. Introducing a different starting point, resulting in another purifiable intermediate, the pathway (Scheme 2.17, b) is slightly changed.

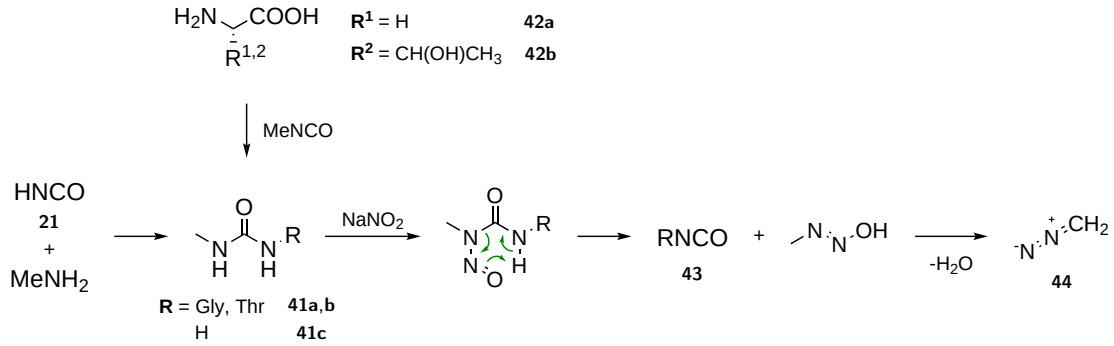
The procedure starts from malononitrile **27**, derived from **6**, which is nitrosylated, giving (hydroxyimino)malononitrile. Besides, wet-dry cycles and day-night cycles are described. A first night cycle, by cooling the reaction mixture, crystallises the desired compound as its guanidinium salt. During the day, upon heating, annealing leads to the corresponding nitrosopyrimidines **30**.[46] As these are not soluble in aqueous media, several derivatisation steps are possible to furnish miscellaneous derivatives of **30**. In further reaction, these lead to non-canonical ribonucleosides. Still present

in modern RNA, they are seen as ancient relics,[123] confirming the validity of the pathway.

Reduction in a formic acid environment, utilising a native iron, nickel or zinc system[124] results in **32**. The formation of hydrogen from formic acid, oxidising iron and nickel to their water-soluble salts, reduces **30**. Later substituted by Zn^{2+} in the unified pathway. Purification is achieved by precipitation of the metal salts under alkaline conditions, while **32** remains soluble and is washed away. In a daylight procedure, the supernatant reaction mixture is dried, crystallising **32**. From there, **2a**,**c** are gained as previously described.

The pathway is characterised by the complex formation of ribonucleosides within the boundaries of simple changes of physical parameters such as temperature, pH and concentration. Yet, these *simple changes* in the environment are the crux of the matter.

A second, different route towards non-canonical ribonucleosides is proposed.[125] Simple derivatisation steps of the canonical ribonucleosides include carbamoylation and methylation reactions, see Scheme 2.16.



Scheme 2.16: Prebiotic formation of diazomethane **44** and a carbamoylation reagent from amino acids to achieve derivatisation reagents for canonical RNA nucleosides.[125]

The high reactivity of isocyanate **21** is preserved in methyl urea **41** via the nucleophilic addition of methylamine. Nitrosylation of **41** under borate-alkaline conditions, rearrangement,[126] and the elimination of water furnishes diazomethane **44**. This household reagent is a potential candidate to methylate **2**. Complementary, the carbamoylation of the nucleosides is achieved, starting from the amino acids glycine **42a** and threonine **42b** as precursors.[125] In analogy, these react with methyl isocyanate

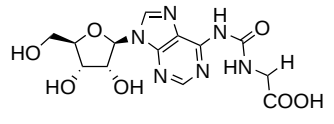
to lead to the glycine- and threonine derivatives **41a**,**b**.

After nitrosylation[128] of the isocyanate-derivatives **41**, carbamoylation of **2** by **43** at the exocyclic amino groups can be achieved. RNA- derivatives like **45a** and **45b** might be a result of this procedure.[127] (*cf.* Figure 2.3) As described earlier, in Section 2.4, methylation/carbamoylation via *N*-methyl-*N*-nitrosourea is a key step in a proposed synthesis towards pyrimidine nucleobases.[62]

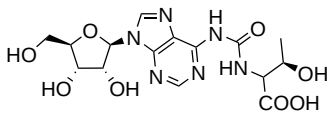
Addressing the corporate formation of pyrimidine and purine nucleosides, a unified pathway was reported, see Scheme 2.17. Increasing the prebiotic plausibility of this pathway, it starts from one common pool of feedstock molecules, derived from **6**. A promiscuous procedure of different environments and removal/addition of reagents to the same flask is described, which is referred to as a one-pot sequence. The synthetic procedure towards **2a**,**c** remains as reported, but an interacting procedure towards the pyrimidine ribonucleosides is designed. It starts by generating 3-aminoisoxazole **47** from **6** and hydroxylamine **46** under basic conditions. In close analogy to the Raschig process, **46** is supposed to originate from N_2 and CO_2 *via* partial reduction of NO_2^-. However, the prebiotic relevance of **46** is questionable.

Mimicking a day-night cycle, dry down of the reaction gives **47**. A clean reaction with **20** gives *N*-isoxazolyl-urea **48**, with only **47** as the remaining impurity. The reaction is catalysed by Zn^{2+} or Co^{2+}-salts. Precipitation of the metal cations as their carbonates leaves all other compounds in solution. A *one-pot* procedure from **6**, **20** and **46** to **48** is proposed.

A pure solution of **4a** is needed, incidentally flooding the **48**-pond to furnish the

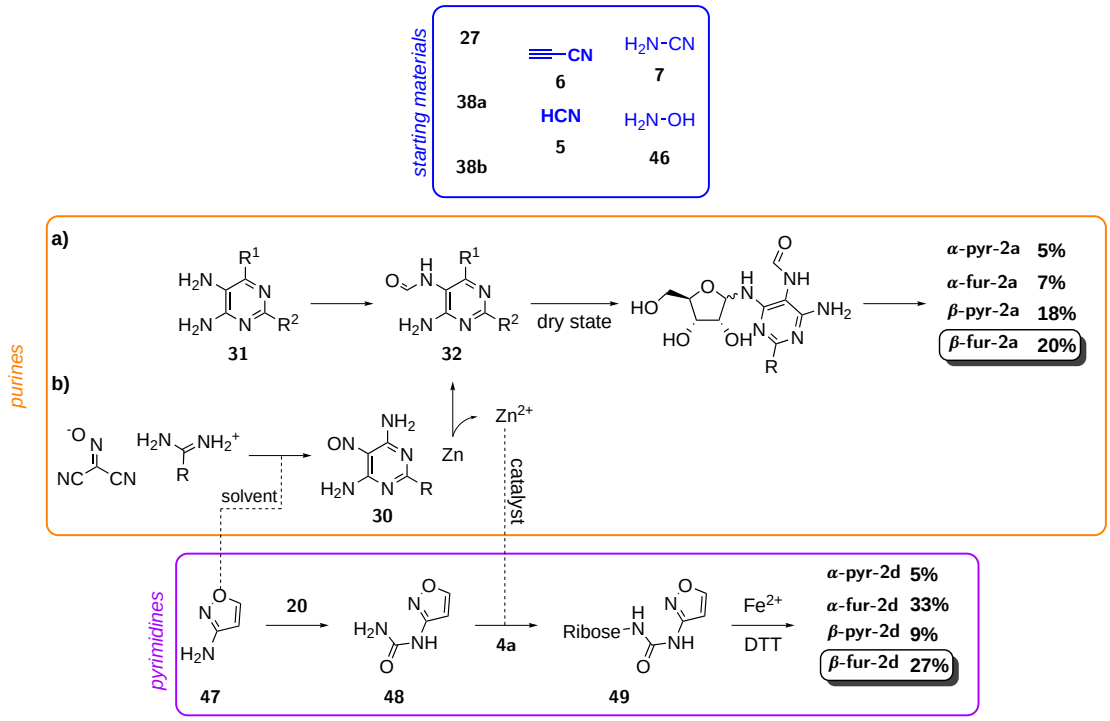


45a



45b

Figure 2.3: g^6A **45a** and t^6A **45b** as examples for derivatives of RNA-nucleosides.[127]



Scheme 2.17: A unified pathway towards ribonucleodsides starting from formamidopyrimides **32** and 3-aminoisoxazole **47**.[19,105,122]

ribosylation. Upon draught, **4a** anneals to give ribo aminoisoxazoleurea **49** in a solid-state reaction. The reaction is catalysed by boric acid or other borate containing minerals, such as lüneburgite[121] or borax.

In a first step, the incorrect conformation of the sugar is achieved, yielding the α-,β-pyranosides of **49**. Heating the mixture under basic conditions corrects the discomfort, giving α-,β-furanosides, see Scheme 2.17. This correction is at the price of losing **49** to hydrolysis. The most complicated step is a consecutive FeS-mediated reduction. The isoxazole-ring of **49** is opened reductively, rearranges and furnishes **2b**,**d**. Unfortunately, the ribonucleosides are obtained in an α-/β-mixture, possessing a furanose/pyranose ratio of 17:1.

Acknowledging a unified ancestor for both purine and pyrimidine nucleosides **2** in a chemically beautiful manner, the pathway inherits severe inconsistencies doubting its prebiotic relevance. A downside is the ambiguous mixing of distinct environments and the incompatibility of certain reactions.

47 is reported to mediate the formation of **30** as a solvent. On the contrary, **32** does not form in this solvent. Therefore, a process is necessary which removes **47** (bp. 228°C) and later adds it to the reaction mixture under prebiotic plausible conditions.

Reduction and formylation of **30** take place after coincidentally mixing pure **30** with a stream of water containing zinc and formic acid. Synergistic effects are described, as the formed zinc salts could also catalyse the reaction of **47** with **20**. This synergistic effect, however, cannot take place as **47** needs to be removed to enable the formation of **32**.

Even though distinct environments are constantly needed, the ribosylation of **32** or **48** proceed under the same conditions and might be accomplished in one pot. These are minor disturbances regarding the geological setting of the scenario. Nevertheless, a major chemical problem is a prebiotically plausible need for **46**, which is not easily explainable. In the last environment **2b**,**d** and **2a**,**c** are formed, giving **2a** and **2d** with the highest yields. The yield of **2a** is decreased by double ribosylation, occurring at the exocyclic amine group.

From a chemical point of view, the described pathway includes well-designed reactions, still, the logical arrangement and the progression of the reaction cascade and its compulsory succession is rather disputable. Further, a strong explanation for the origin of **46** needs to be provided. Again, the outcome of the pathway is decisively dependent on the experimenter and the assumption implies that the formation of Life on Earth is a mere product of coincidence.

The above described pathways[18,19,105,114,117,119,122] are attempts to synthesise **2** from common organic feedstock molecules. Different prebiotic scenarios are outlined, assuming that RNA is the coding polymer at the advent of Life. This theory is highlighted by the facts that RNA exhibits catalytic function, the ability to store information and the possibility to undergo template-based polymerisation, further described in Section 3.1. Resulting from prebiotic inaccessibility of deoxyribose **d4a**, DNA, as the coding polymer of modern Life, is said to be developed at a much later stage. The transition from RNA to DNA is hypothesised to be assisted by enzymes, which cleave the decisive 2'-hydroxyl group.

Recent research results,[129] however, indicate a chimeric formation of DNA-RNA double strands, shifting the focus from RNA. These chimaeras could have played a role in template-directed, autocatalytic and non-enzymatic replication of bioinformational polymers.

Contrary to the instability of the free deoxyribose **d4a**, the corresponding DNA strand is more stable than RNA, except for the acid-mediated depurination of the double-strand.[130] This poses the question, why DNA nucleosides, nucleotides or even the DNA polymer should not have had a role in the origin of Life context.

One could argue that maybe DNA is too stable to participate in the chemical evolution. However, the theory is appealing that at the beginning of Life, a flexible, catalytic polymer (RNA) and a stable, informational polymer (DNA) were working hand-to-hand. Consequently, a prebiotic pathway towards deoxy ribo nucleosides **3** is needed.

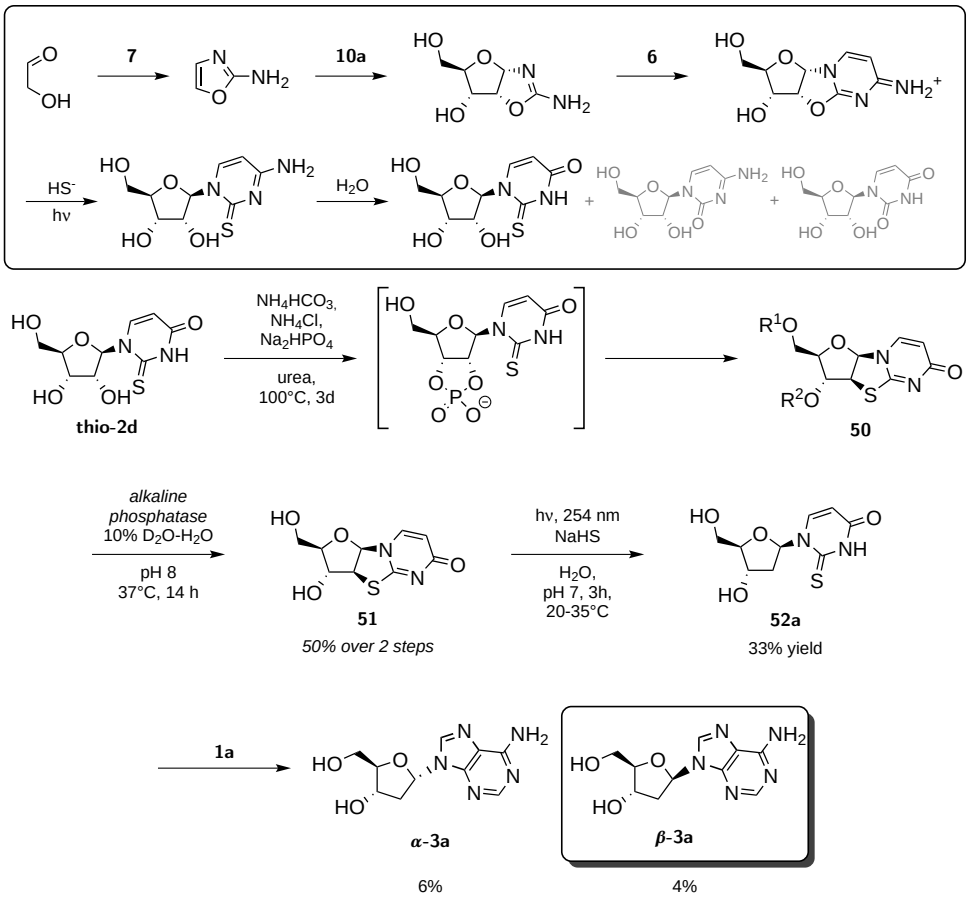
2.6 Chicken or Egg? – Deoxyribonucleoside Synthesis

As the view shifts from the sole presence of RNA ruling the emergence of Life, recent advances to the synthesis of DNA **3** are presented. The presented pathway is proposed by the Sutherland group[131] and is in analogy to the dithioreduction of the enzyme *ribonucleotide reductase*.[132]

In the first contribution to **3**, a variation of the approach to **2** via **26a**,[131] is presented. In a preceding publication,[114] a pathway was found utilising easily enrichable **25a** as a starting material for **2**-synthesis. Reconstructing this known pathway (see box, Scheme 2.18) gave access to **3**.

Studying, the phosphorylation of thiouridine **thio-2b** in semi-molten urea, it was expected to find the 2',3'-cyclic phosphates, as in analogy to the reaction with DAP.[133] Instead, the reversed reaction was observed, leading to the formation of the phosphorylated thioanhydronucleoside **50**, see Scheme 2.18.

50 is the first example of a C-S connection at the 2'-carbon of the furanose ring. The C-S bond of **50** might be prone to reduction, in similarity to the key step of the *ribonucleotide reductase*, a stoichiometric radical dithioreduction. In consequence, it might open up a pathway to **3**.



Scheme 2.18: Phosphorylation of **thio-2b** enables the formation of the corresponding thioanhydronucleoside **51** and subsequent formation of deoxythiopyrimidine ribonucleosides.[131] (R^1, R^2 =$HOPO_2{}^-$,H; R^1, R^2 =H,$HOPO_2{}^-$; R^1, R^2 =$HOPO_2{}^-$)

Reduction under UV-irradiation (254 nm) of **51** in aqueous H_2S-solution resulted in a mixture of 2-thiocytosine **thio-1d**, thiouracil **thio-1b** and 2'-deoxy-2-thiouridine **52a**. Unfortunately, the reduction does not proceed directly from the phosphorylated species **50**, but just from **51**. Lacking a prebiotic pathway for dephosphorylation, **51** was obtained from **50** after treatment with *alkaline phosphatase*. Further, just the thioreduction of uridine-anhydronucleoside **50a** and not the reduction of cytidine-anhydronucleoside **50b** yielded the 2'-deoxyribonucleoside **52a**.

Verification of the products and conditions was provided by control experiments and analysis of the reaction mixture by NMR spectroscopy using pentaerythritol as

an internal standard. The reaction in the dark under the influence of H_2S as well as the irradiation with UV light without H_2S, did not lead to the formation of **52a**. It is concluded that the reduction is light-induced, cleaving the C-S bond after the generation of solvated electrons from HS^-.

The pathway elegantly mimics the active site reaction of the *ribonucleotide reductase*. Increasing its prebiotic plausibility calls for a prebiotic dephosphorylation procedure replacing the *alkaline phosphatase*. Further, a pathway towards **52b** and its transition to **3d** needs to be found, as dU **3e** is not a constituent of modern DNA. Until a procedure reasoning the formation of **3d** and **3b** is found, this multistep pathway is highly scrutinised.

Aware of these constraints, it was tested if a transglycosylation procedure starting from **1a** and **52a** gives access to the canonical purine nucleosides. However, solely dry state conditions at 100 °C successfully showed the low-yielding formation of **3a**. The transglycosylation is not stereoselective and results in both the α- and the β-anomer of **3a**.

Although an elegant pathway is delineated, which avoids the use of pure **4a**, a classical Fischer-disconnection[45] at the anomeric carbon closes this route to **3**. As expected, a vast reaction mixture results, applying the reaction conditions of Orgel.[86] This once again indicates that another disconnection rationale than the one applied for the last 90 years is in great demand.

In total, synthetic access to deoxy uridine **3e** in an overall low selectivity is given. However, routes towards all four canonical nucleosides are still lacking, accompanied by minor stereochemical selectivity caused by a radical process. Adding all downsides of this process results in a poor prebiotic relevance of this scenario.

2.7 Relevance of Apiosyl Nucleosides

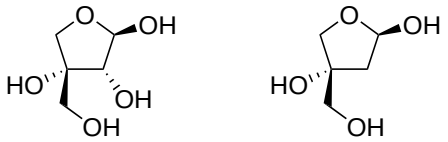


Figure 2.4: Structure of D-apiose **4e** and 2'-deoxy-D-apiose **d4e**.

Apiose **4e** is an example of a 3'-branched sugar, see Figure 2.4. It was first isolated from parsley in 1901.[134] Although its ostensive unusualness, it is commonly found in pectins of all higher plants. Its two main locations are in the polysaccharides rhamnogalacturonan II and apiogalacturonan.[135] However, free **4e** has not been found so far. In rhamnogalacturonan II, (RG-II, *cf.* Figure 2.5) apiose is one of 12 different sugar subunits, building up one of the Earth's most sophisticated carbohydrate structure.[136] The **4e**-subunit cross-links with borate to interconnect two RG-II units. This dimeric structure functions as a stabiliser in plant cell walls. Lack of borate (an essential micro-nutrient for plants) or **4e** hinders normal plant growth and development. The deficit causes phenotypical disorder of the agonised plant.[137] The second **4e**-containing polysaccharide is apiogalacturonan. Apart from highly ordered polysaccharides, glycosides of **4e** with glucose **53a**, galactose **53b**, rhamnose **53c**, xylose **4d** and arabinose **4b** could be found.[138] The most prominent disaccharide is β-D-apiofuranosyl-(16)-β-D-glucopyranose, which is called *acuminose*. Despite its essential function in plants, the origin of **4e** has not been elucidated so far. Especially the question of why energy is invested to create such a peculiar sugar could not be answered.

Biosynthetically, **4e** is achieved from glucuronic acid.[140] Later, UDP-apiose, an activated form was discovered, posing the relevance as secondary metabolites. These play important roles in signal transduction as well as in defending pathogens or attracting symbionts. Synthetic conversion to apiose-UDP is mainly achieved by the two enzymes *UDP-apiose/UDP-xylose synthase* and *UDP-apiose:7-O-(-D-glucosyl)-flavone apiosyltransferase*.[141]

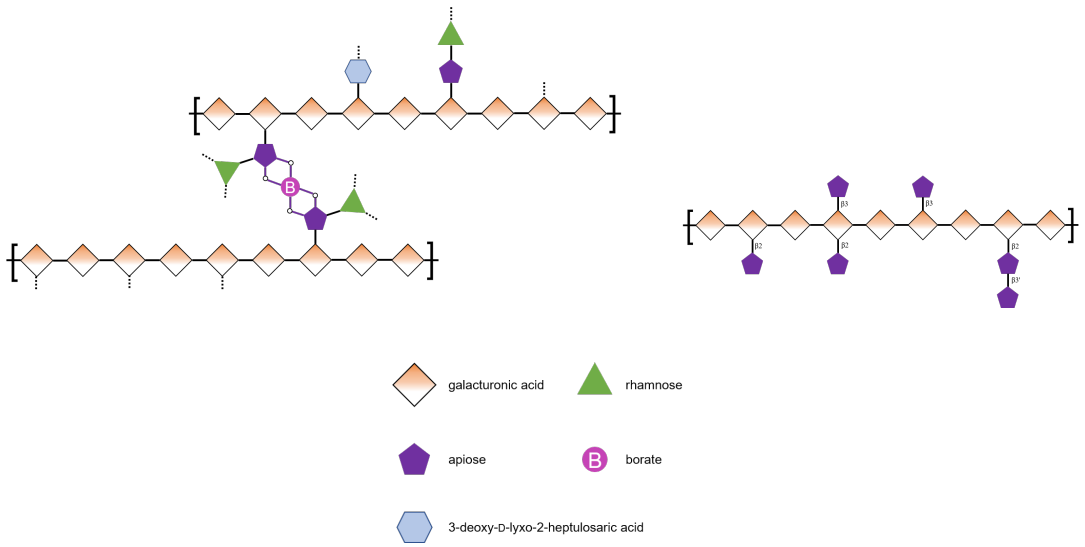
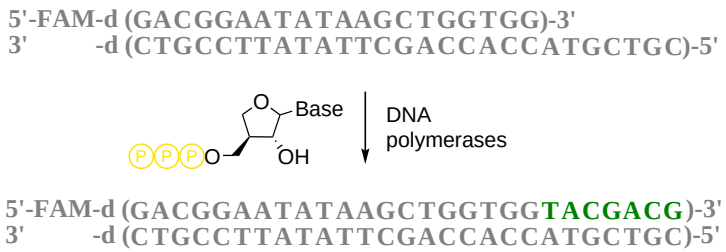


Figure 2.5: Highly simplified structure of rhamnogalacturonan-II (RG-II,left) and apiogalacturonan (right).[139]

Apart from the elucidation of the biosynthetic relevance, the investigation of **4e**-derivatives in a wider prebiotic context and as antiviral agents were conducted. Both are briefly illustrated in the following.

In vitro, one approach correlates apiosyl-nucleosides to the concept of the *aetiology of the nucleic acid structure*, see Section 2.5.2. The polymerisation of threose nucleosides (TNA) could be conducted with DNA polymerase. However, the polymerisation stopped at a certain point.[142] Examining these results closely,[143] it was hypothesised that the number of bonds in the sugar backbone, from phosphodiester to phosphodiester, might correlate with the truncation of the polymerisation. In TNA, the number of bonds from phosphodiester to phosphodiester is five, whereas in DNA/RNA it is six. Therefore, a 3'-deoxyapiosyl nucleotide was designed to have the same connection mode as in TNA, but the same number of bonds as in DNA. Synthetic access under stereo- and regioselective control was achieved from galactose, see Scheme 2.19. Subsequent 3'-phosphorylation furnished the desired 3'-deoxy-apiosyl nucleotide. Finally, primer extension with DNA polymerase incorporated these nucleotides into the DNA template structure.[143]

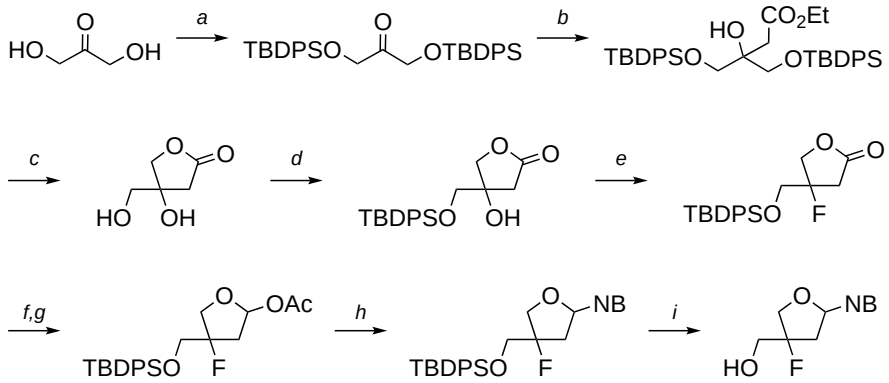
Apart from incorporation into polymeric structures, another field of interest of apiose nucleosides is antiviral drugs. Synthetic access to potent antiviral agents is



Scheme 2.19: Successful polymerisation using 3'-deoxyapionucleosides and DNA polymerases.[143]

a key part of nucleoside chemistry. This field is motivated by natural antiviral and antibacterial nucleoside derivatives. In its resemblance to canonical nucleosides, apiose nucleosides are a promising suspect for potent antiviral and antibacterial drugs.

Inspired by the discovery of nucleocidin, a natural antibiotic containing fluorine, two synthetic procedures are outlined accessing fluoro-apiosyl nucleosides. Replacement of hydrogen atoms by fluorine atoms leads to a similar enzyme recognition, but a significantly different bonding interaction. During therapy, crucial enzymes can be inhibited.[144]



Scheme 2.20: Synthesis of fluoro-apionucleosides. Conditions: a)TBDPSCl, pyridine; b) $BrCH_2CO_2Et$, activated Zn, benzene– toluene (2 : 1), reflux; c) 12% HCl in MeOH, r.t.; d) TBDPSCl, DMAP, Et_3N, DCM, r.t.; e) DAST, DCM, r.t.; f) DIBAL-H, DCM, –78 °C; g) Ac_2O, Et_3N, DMAP, DCM; h) bis-silylated base, TMSOTf, DCE, 0 °C; i) TBAF, THF, r.t.; NH_3, methanol, r.t., then TBAF, THF,r.t.;

One successful synthesis starts from lactose, giving 2'-deoxyapiose- L-furanosyl nucleosides. However, assays with HIV could not fulfil the expectations of antiviral activity.[145] A second procedure, see Scheme 2.20, gives access to another 3'-fluoro

apiosyl nucleoside.[146,147] The potency of antiviral activity is speculated, yet no studies are reported.

In Nature, solely the incorporation of apiose nucleoside and its corresponding nucleotides were found, the integration of deoxyapionucleosides or -nucleotides is not known.

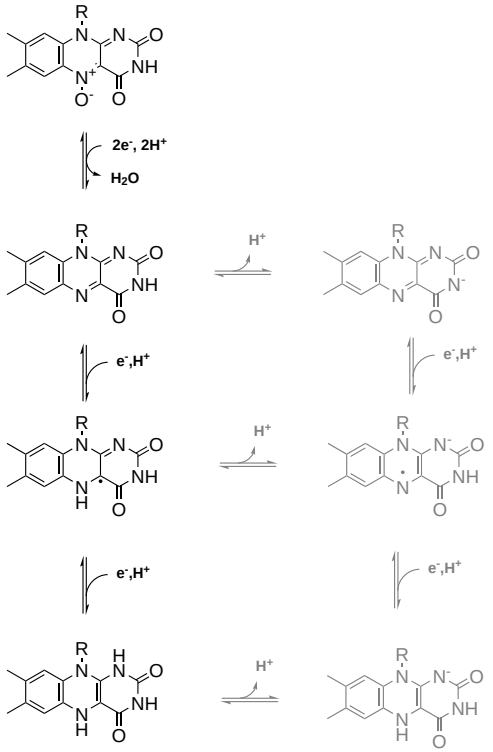
2.8 Research on Cofactors

Cofactors are ubiquitous parts of modern biological life. They can be separated into two different domains: inorganic cofactors and organic cofactors. The latter can be bound to an enzyme either covalently or non-covalently. In elucidating the origin of Life via organic-synthetic methods, the examination of routes towards cofactors definitely needs to be taken into account. Amongst many, the most prominent representatives of cofactors are haem, biotin, flavin and pyridoxal phosphate.

The haem core unit is the basis for haemoglobin.[148] Further, the heterocycle is included in the photosynthetic system. Biotin functions as a cofactor in epigenetic regulation. Pyridoxal phosphate catalyses transamination, decarboxylation and dehydration. Flavin is active in the cellular respiration enzymes as a proton/electron donor and acceptor. Flavins can catalyse reactions, transferring either one or two electrons. In contrast to NAD^+, it is strongly bound to the enzyme. After the reaction, it needs to be regenerated oxidatively, in an energy-consuming step.

Underlining its versatility in redox reactions, it exists in four different states, see Scheme 2.21. In the highest oxidation state, it exists as the flavin-*N*-oxide. The uptake of two electrons leads to the state of the quinone which resembles the oxidation state of FAD. Adding another electron brings the system to the semiquinone-state. Finally, receiving another electron, the system reaches the highest reduction potential in the state of the hydroquinone or $FADH_2$. Every structure is in a protonation equilibrium and has its distinct function in the respiratory cycle. Most reactions catalysed by flavin cofactors implement a one-electron transfer.

In recent investigations on the prebiotic formation of essential molecules cofactors did play a minor role. It was likely that riboflavin was present at an early stage of Life as it is a versatile assistant in many enzymatic reactions. (*cf.* Section 3.1) Therefore, a closer examination of the biosynthesis and the design of prebiotic syntheses might



Scheme 2.21: Different redoxstates of the flavin-cofactor (R=ribityl-residue)

turn out valuable.

2.8.1 Biosynthetic Pathway

Major contribution on the mechanistic elucidation of the riboflavin biosynthesis was provided by the group of Bacher.[149] In 1952, it was found that the biosynthesis is greatly dependent on the concentration of purine nucleosides in the growing medium, concluding that the synthesis of riboflavin is linked to purine nucleosides.[150] Subsequent studies found GTP to be the starting point.[150] In brief, the enzymatic (*GTP cyclohydrolase-II*)[151] breakdown of GTP forms 6-ribityl-5,6-diaminouracil **54**, which is the core structure of the flavin (Figure 2.6).

The pyrimidine ring, the nitrogen atoms of the pyrazin ring and the ribityl-residue are all descendants of this intermediate. Crucial steps in the break down are the loss of formate and the reduction of the **4a**-residue to the ribityl-chain. Salvage ends with

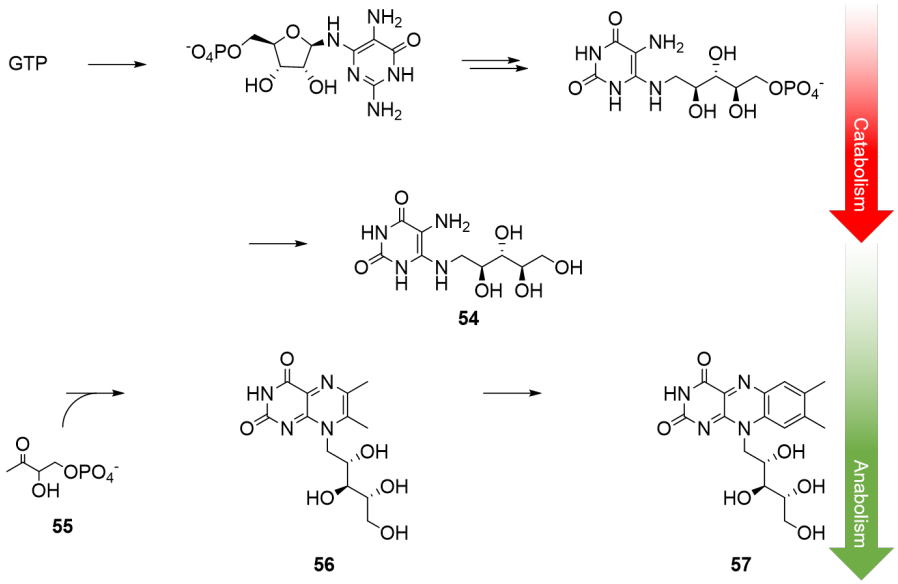


Figure 2.6: Biosynthesis of riboflavin **55**

54 and in a consecutive bottom-up synthesis 6,7-dimethyl-8-D-ribityl-lumazine **56** is formed. Construction of the pyrazin-ring is achieved from 3,4-dihydroxy-2-butanone 4-phosphate **57**, derived from ribulose-5-phosphate **58**-**P**. **58**-**P** itself is a product of the pentose phosphate pathway.[152] During this step, the 5'-phosphate of **54** cannot undergo the pyrazine formation, so the phosphate needs to be cleaved first. In a subsequent dismutation reaction, **55** is gained from **56**. The dismutation only proceeds from **56** and not from the lumazine heterocycle **59**, lacking the ribityl residue.[153]

2.8.2 Early Pteridine Synthesis

Early procedures synthesising the pteridine sub-unit and its biosynthetic predecessors were illustrated.[44] In these cases, the urge to find easily accessible syntheses towards simple, ubiquitous heterocycles was in focus, rather than the prebiotic route towards **55**. In search for a scalable, elementary route towards xanthin **1g** and guanine **1c**, 4,5-diamino uracil **60** was found to be an intermediate structure. This compound is build from cyanoacetic acid **61** and urea **20**. Condensation of these, followed by cyclisation gives 6-amino uracil **62**. The nitrosylation and reduction of **62** then led to **60**. The composition of the pteridine heterocycle was achieved in similarity to the

biosynthesis of **59**. Reacting **60** with diacetyl **63** gives 6,7-dimethyl lumazine **59**.[154] The connection to **55** was mentioned, however, this path was not investigated more closely.

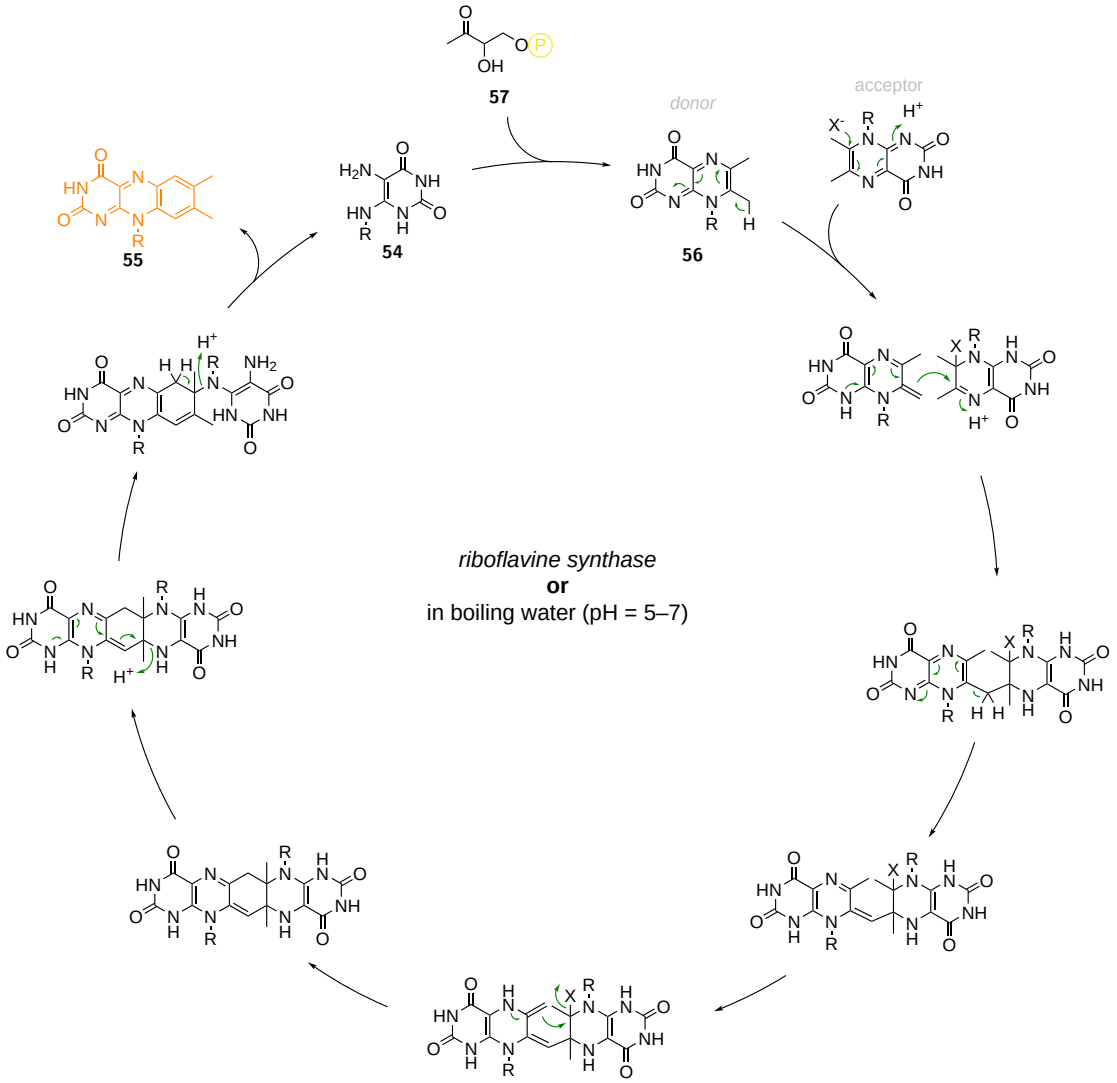
2.8.3 Elucidation of the Biosynthetic Pathway

In the following, a deeper mechanistically insight is given, which led to the elucidation of the **55**-biosynthetic pathway. A careful comprehension of the biosynthetic ideas might lead to derivable prebiotic concepts. In this context, especially two questions are crucial: (i) How is the pyrazin subunit built up and (ii) how does the lumazine dismutation proceed mechanistically?

Concerning the pyrazine ring closure, many C_4-molecules, including acetoin,[155,156] diacetyl **63**[157] and pyruvate[158] were discussed. Because of their lack of regioselectivity, these compounds were ruled out as potential precursors. Conducting *in vivo* studies with ^{13}C-labeled species showed a connection between the substances of the pentose phosphate pool and the incorporated carbon material in the pyrazine ring.[159] The compound closing the pyrazin ring was verified to be 3,4-dihydroxy-2-butanone 4-phosphate **57**, which results from ribulose-5-phosphate **58**-**P**.[152]

Evidence for **57** was found in the yeast *candida guilliermondii*, with an enzyme catalysing the transition from **58**-**P** to **57**.[160]

Examining the lumazine formation, it was discovered that the synthesis proceeds without the need for the enzyme *lumazine synthase*. Studies on the pH- and temperature dependence revealed the optimum at 65 °C and a pH of 7.75.[161] Further, the earlier discussed diacetyl **63** can also be established from **57** by elimination at elevated temperatures. The former incorporates into the pyrazin ring, yet with low regioselectivity because of the symmetry of the molecule.



Scheme 2.22: Proposed mechanism of the lumazine dismutation, posing a autocatalytic cycle, catalysed by *riboflavin synthase*.[162]

Concerning the second question, the final catalysed step to riboflavin **55** is a peculiar dismutation reaction, in which two molecules of ribityl lumazine **56** result in one molecule of **55** and one molecule of the precursor **59**, see Scheme 2.22. The presented reaction thus poses a perfect autocatalytic cycle, needed for the off-equilibrium development of Life. During the course of the reaction, the large electronic flexibility, provided by the aromatic system, is neatly utilised. The mechanism leading to **55** was intensively studied by deuterium exchange experiments and monitoring by NMR

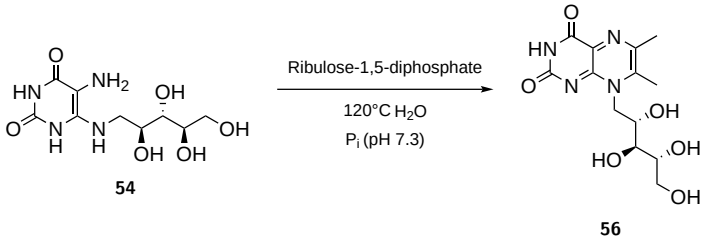
spectroscopy.[153,163] Astonishingly, it was found that the reaction proceeds without the help of an enzyme. The mechanism is yet not fully clarified. The current postulated mechanism[164] depicts one lumazine molecule as the acceptor and the other as the donor molecule.

Closer examination of the biosynthetic pathway reveals the simplicity of the reaction cascade. Most of the reactions are simple conversions, which even proceed without the need for a catalyst. The assumption may be made that the reaction cascade at the beginning of Life was closely related to the one present in modern biology. Further optimisation and enhancement of the reaction sequence by enzyme catalysis, was implemented at a later stage of evolution.

2.8.4 Prebiotic Investigations on Riboflavin

In the following, the major attempts to synthesise riboflavin **55** under prebiotic conditions are presented. The theoretical considerations and experiments were conducted in the group of Eschenmoser. In one examination, the formation of the lumazine ring was investigated,[165] whereas the reactions building the ribityl residue at the heterocyclic unit were examined in a different investigation.[103]

The first one represents a direct formation of **56** in 9.4% yield from ribulose-5-phosphate **58**-**P** and **54**.[165] The reaction proceeds in a phosphate buffer (pH 7.2) under the exclusion of air and light at a temperature of 120 °C, see Scheme 2.23.



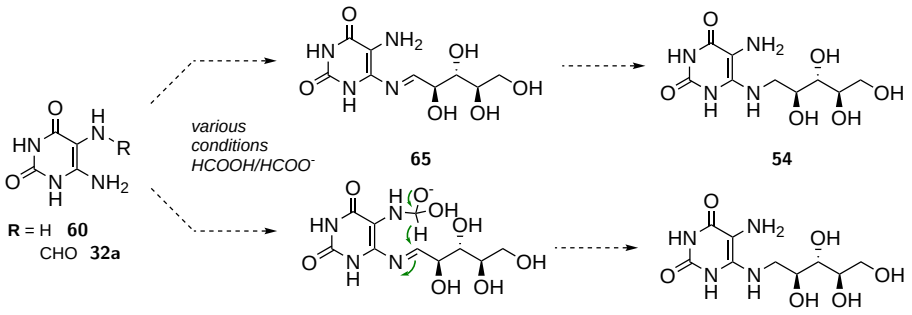
Scheme 2.23: Formation of **56** from **54** as performed bei *Strupp*.[165]

This pathway was examined based on the fact that **58**-**P** does degrade in basic solution to form 3,4-dihydroxy-2-butanone 4-phosphate **57**. Exceptional care was paid to the strict exclusion of water and oxygen to avoid concomitant oxidation products. It

is stated that the unique energy potential gradient provided by hydrolysis should be utilised at a later stage. Further byproducts of the described reaction are 6-methyl- and 7-methyl-8-D-ribityllumazine. To a great extend 6-acetyl-7-methyl-8-D-ribityllumazin was detected as the main byproduct. On the other hand, reacting ribulose **58** with 5,6-diamino uracil **60** does not lead to the desired product.

As no direct prebiotic pathway towards **58**-**P** is known, direct implementation of the C_4-precursors, *e.g.* 2,4,6-pentanetrione **64**, was utilised. The conversion of **64** and **54** leads to **56**, however in low yield and accompanied by a vast mixture of side-products.

Simulating its gradual development from **57**, the slow addition of **64** to the reaction mixture gave significantly higher yields of **56**. Studies with isotopically labelled material indicated that the corresponding mechanism does not proceed via the formation of 7-acetyl-6-methyl-D-ribityllumazine. Concluding, a conversion of **54** to form **56** could be demonstrated. However, the preceding step, constructing the ribityl-residue at the heterocyclic compound, is still unclear.



Scheme 2.24: Reaction of **60** with **4a** under various conditions mostly implementing either formic acid or formiate.[103]

The reductive **4a** coupling was seized in a concurrent work.[103] Whenever the constitution of **4a** at a heterocycle is involved, the main challenge is to solve the *notorious nucleosidation problem.*[1] Towards **55**, it is coupled with a reduction.

The condensation of **4a** and **60** was described in previous experiments, giving a mixture of four isomers.[166,167] For the reduction of the presumably formed **4a**-imine, a formic acid system was proposed. The formic acid system could reduce the ribose residue in a Leuckhart-Wallach-reaction. Further, it is assumed that **60**, in its non-

oxidised state, could perform the reduction of the ribose side-chain.

Repeating the experiments[166,167] revealed their ambiguity as the desired ribityl-derivatives were not found. It is supposed that this fact is due to low regioselectivity as **4a** preferentially reacts with the C5-amine and not with the C6-amine.

Enhancing the selective formation, a system is proposed in which the preferred 6-ribosylamino-5-amino uracil **65** is withdrawn from the equilibrium by reduction to **54**. The 5-ribosylamino- and the 6-ribosylamino-species should be in equilibrium, however, the oxidation potential of the latter should be higher to enable a facile reduction.[103] Analysis of the reaction mixture via TLC in combination with a ninhydrin stain and NMR spectroscopy did not show evidence for **54**. Screening possible conditions, plenty of reactions were tested with no success. However, in resemblance to previous works,[104] the facile formation of 5-formamido-5,6-diaminouracil **32a** was observed, see Scheme 2.24.

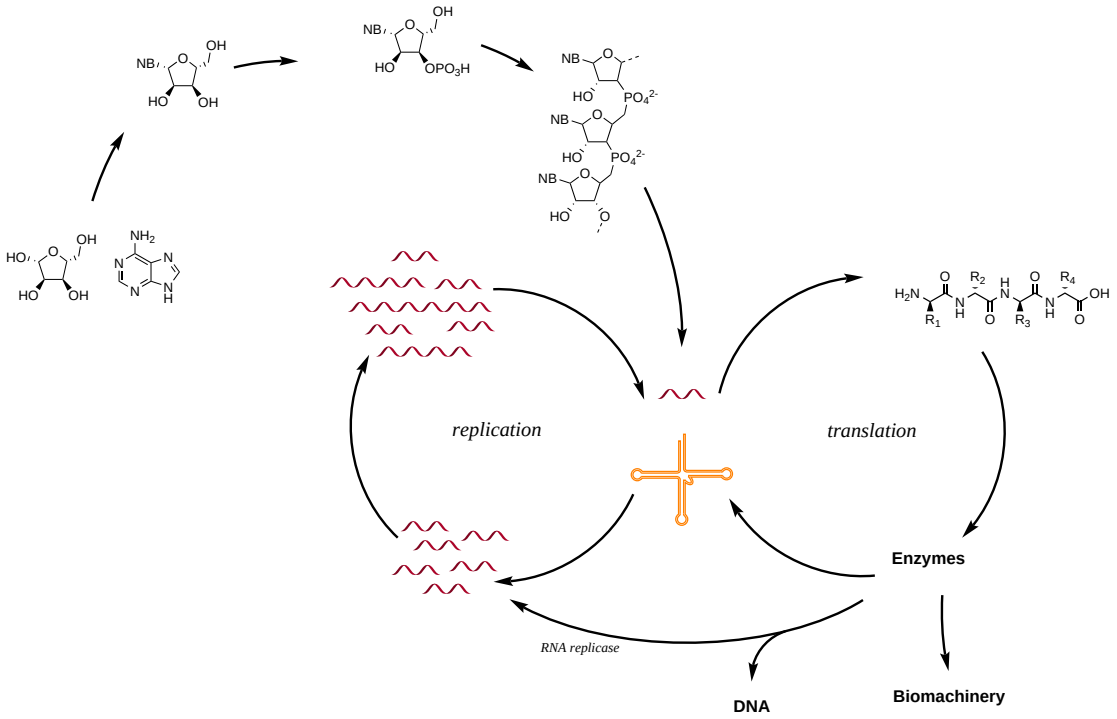
This sparked the idea to intramolecularly reduce a formed **4a**-imine via the formamide. The proximity of the oxidant and the reductant is advantageous. Such a reduction would require the formamide in its hydrated hemiaminal form, which might be present under alkaline conditions. A careful investigation of different reaction conditions indicated the problem that **4a** does not form an imine with **32a**. Consequently, the highly elegant reduction via the formamido-hemiaminal yielding **54** could not be studied. Combining the above-presented biosynthetic pathway[103,165] with ideas resulting from previously conducted prebiotic experiments could open up new pathways to the essential cofactor.

3 Theoretical Framework

In this chapter, the theoretical concepts employed in the *Origins of Life* research are briefly delineated and explained. The description starts with the *RNA-world* hypothesis,[168–171] extended by a hypothesis of the *aetiology of the nucleic acid structure*.[172] Second, the *mechanism-first* hypothesis[173–175] is explained, accompanied by a short note about a *physical approach* to the formation of Life. Last, a short introduction to *systems chemistry*[176,177] and its contribution to the research of the origin of Life is given. Subsuming all theories, a critical evaluation is given. Concluding, the condensed theory forming the basis of this thesis is provided.

3.1 *RNA World* Hypothesis

The *RNA-world* hypothesis is currently the most accepted theory to explain the formation of Life on Earth and the subsequent evolution to modern biology, see Scheme 3.1. Ideas of RNA acting as a *godfather* molecule at the emergence of Life were discussed early.[169] It is believed that this essential biomolecule developed directly under primordial conditions from prebiotic feedstock material. The discovery of the *ribozyme*[178–180] sparked the prebiotic RNA research further by showing the capability of RNA to inherit catalytic function and informational storage.[123,181,182] Catalysis is believed to act on surrounding reactions, however, the key feature of RNA is to catalyse its replication.[183] Resulting from the development of intron and exon splicing, it is supposed that RNA gradually gained flexibility and adapted by mutations. Validation for this principle is seen by the fact that it is still observed in modern tRNA.[184] The RNA oligomers were capable of catalysing transitions in their environment, creating a chemical network based on RNA. The catalysis rates are not expected to be high, but sufficiently higher than without assistance. It is hypothesised[185] that cofactors were



Scheme 3.1: Illustrative development of the modern biomachinery according to the *RNA-world hypothesis.*

present early and increased the rate further. This led to the emergence of protein-like structures, developed from RNA via RNA adapter molecules, leading to amino acid oligomers and short proteins. Over time, it gave rise to an enzyme, equivalent to the *ribonucleotide reductase*, converting RNA into DNA.[186] DNA substituted RNA by providing a more stable medium for bio-informational storage. In catalytic terms, the RNA got paced out by far faster and more efficient enzymes, repelling RNA to its background role in modern biology. The major evidence for this theory is seen in cofactors, which are largely based on RNA nucleotides or derived from heterocyclic ribosides. Further, the intron/exon-splicing mechanism generating the catalytically active ribozyme from tRNA is thought to be another relict from the ancient bio machinery of Life.[169]

3.2 The Evolution of RNA

Opposing, it is believed that RNA emerged via an evolutionary process. The complex structure of RNA is too adapted to its function to allow the assumption that it is a product of a single event. Multiple conditions need to be precisely fulfilled to gain RNA in the canonical state as it acts in modern biology. The synthesis must implement the following requirements: *(i) aldose*, *(ii) pentose*, *(iii) ribose*, *(iv) furanosyl*, *(v)* β *not* α, *(vi)* D *not* L, *(vii) regiospecific glycosylation*, *(viii) regiospecific phosphorylation* and *(ix) activation towards regiospecific oligomerisation*.[187] There is, to date, no plausible explanation of *in situ* nucleoside formation, from prebiotic mixture, fulfilling all of these criteria.

The emergence of RNA can be divided into three hypotheses:
First, a classical approach,[19,86] in which RNA is erected from simple feed-stock molecules to form nucleobases and nucleosides. These are transformed into nucleotides, which are polymerised to form RNA. The second hypothesis sets the formation of the sugar moiety at the centre of the synthesis.[18] The third is a so-called *polymer fusion model*, see Figure 3.1, setting at least one pre-RNA polymer as a predecessor for RNA.[182]

The last theory includes that RNA did not emerge from a *prebiotic soup*, but an informational biopolymer or several evolutionary stages of biopolymers preceded RNA. One could imagine, that these biopolymers already served catalytic tasks and stored information at a lower level. Gradually, these hypothesised biopolymers were subject to selection pressure, adapting to the environment. During evolution, all major constituents of modern RNA, the sugar moiety, the nucleobase and the phosphodiester bond were exchanged, finally forming RNA. It is assumed that there have been simpler base-pairing mechanisms before Watson-Crick base-pairing.[188] Theories include a reduction of possibilities by using just two recognition units or one class of recognition units (*e.g.* purines or pyrimidines).[189,190] Instead of base pairs, information could have been stored using base-tetramers or hexamers. Precursor heterocycles like aminopyrimidine **66**, melamine or barbituric acid might have been possible.[191]

The sugar unit is believed to not have been originated from a *frozen accident* sce-

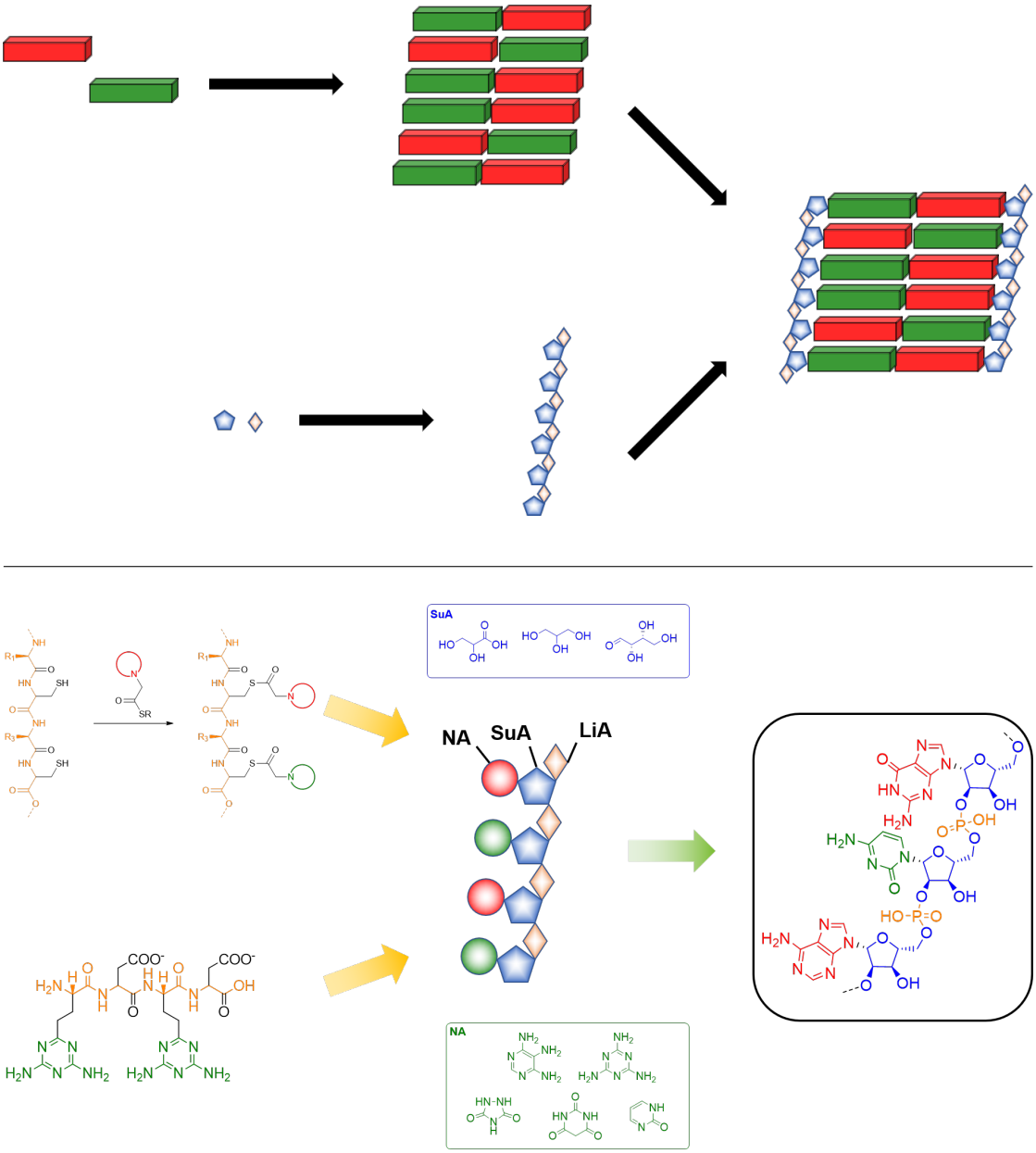


Figure 3.1: Schematic evolution of the RNA according to the polymer fusion model (top) and the evolution of the single building blocks (bottom, NA=nucleobase analogue, SuA= sugar analogue, LiA= linker analogue)[182]

nario as it is too perfectly adapted to its requirements.[1] A *ribose-later* scenario is proposed in which threose, glycerol or glyceric acid preceded ribose **4a** in connecting the recognition units.[101,192,193] In another variation, peptide nucleic acids from aspartate and glutamate are considered as plausible backbone monomers,[194] examples of such structures are depicted in Figure 3.1 (bottom).

A possible unit predating the phosphodiester bond could have been glyoxylate,[195] which is rather similar in size to phosphate. In a peptide nucleic acid, the need for linking units is obsolete. Efforts to find plausible substituent for phosphate are motivated by the assumption that free phosphate was not available to that extent needed on the early Earth.[196] However, there are very plausible reasons for phosphate to be the linker moiety.[197]
The *notorios nucleosdiation problem*[1] could have been attenuated by implementing different recognition units like urazole and 2-pyrimidinone.[198] By substituting the canonical nucleobases, the glycosidic bond is even more prone to hydrolysis.[199] Thus, the resulting pre-nucleoside is easier to form and more flexible.

Following consecutive thinking, it is more logical to first align information-coding subunits to a supramolecular assembly and then interconnecting them to form polymers. The driving force for polymerisation is the constraint of the informational system. As the name suggests, the *polymer fusion model* proposes the separate formation of a polymer *e.g.* a polysaccharide, a polypeptide or a polyester and a supramolecular assembly, see Figure 3.1 (top). These two react with each other, building a pre-RNA structure, whereas the spacing of the polymer fits the supramolecular assembly. The supramolecular assembly might have been present in dimeric, tetrameric or hexameric structures. Following this hypothesis, pre-selection occurred, as not every converging polymer strand fitted the spacing of the assembled monomers.

One example for such pre-RNA polymers might be a thioester peptide nucleic acid (tPNA),[181] reversibly self-assembling from a polypeptide-strand and thioester-derivatives of nucleobases. Adapting a dynamic sequence modification, the gained tPNA aligns with natural RNA and DNA oligomers developing Watson-Crick base pairs. Via such an assembly, template-directed synthesis of successor polymers like RNA or DNA could have been possible.

The role of water under primordial conditions is controversially discussed, as it hinders some necessary reactions. Especially the creation of reversible chemical

bonds, prone to hydrolysis, is a difficult task. Examples are the formation of the glycosidic bond, phosphorylation of the nucleoside and the condensation reaction under cleavage of water to form RNA oligomers. The most prominent answer to this problem is the *drying-lagoon* model.[200] Upon draught, condensation reactions can take place. However, the rates of the reactions are significantly decreased, due to limited molecular motion in the solid-state. Others locate prebiotic reactions to formamide[201] or even liquid ammonia.[202] Deep eutectic solvents, like formamide, can provide polar water-free conditions, in which the nucleoside- and the nucleotide formation[203] and the subsequent polymerisation could take place.
Each of the sub-units of modern RNA has possible progenitors, which are simpler in synthesis and therefore exchanged under selection pressure. Concluding, this theory states the evolution of RNA and not the emergence of RNA as a universal biopolymer, exhibiting all its functions from the beginning. Every change in the transition from pre-RNA to RNA came with a change of function of the related sub-unit, optimising the generated polymer to a new niche of function.

3.3 *Mechanism first* Hypothesis

The so-called *mechanism first* hypothesis sees acetyl-CoA[204] as a relic of an ancient chemistry, which occurred in deep-sea hydrothermal vents. In principle, the hypothesis proposes the reaction of dissolved gases, catalysed by transition metals. These are located at a deep-sea vent, resulting in a chemical network. For a long time, it was believed that *black smokers* provided the circumstances for the formation of life. They are located at the seafloor spreading, at which seawater and magma get in contact. As a consequence, chimneys precipitate, ejecting 400 °C hot water, rich in dissolved transition metals. However, nowadays it is believed that the high temperatures of the *black smokers* did provide unsuitable conditions for the formation of Life.[175]

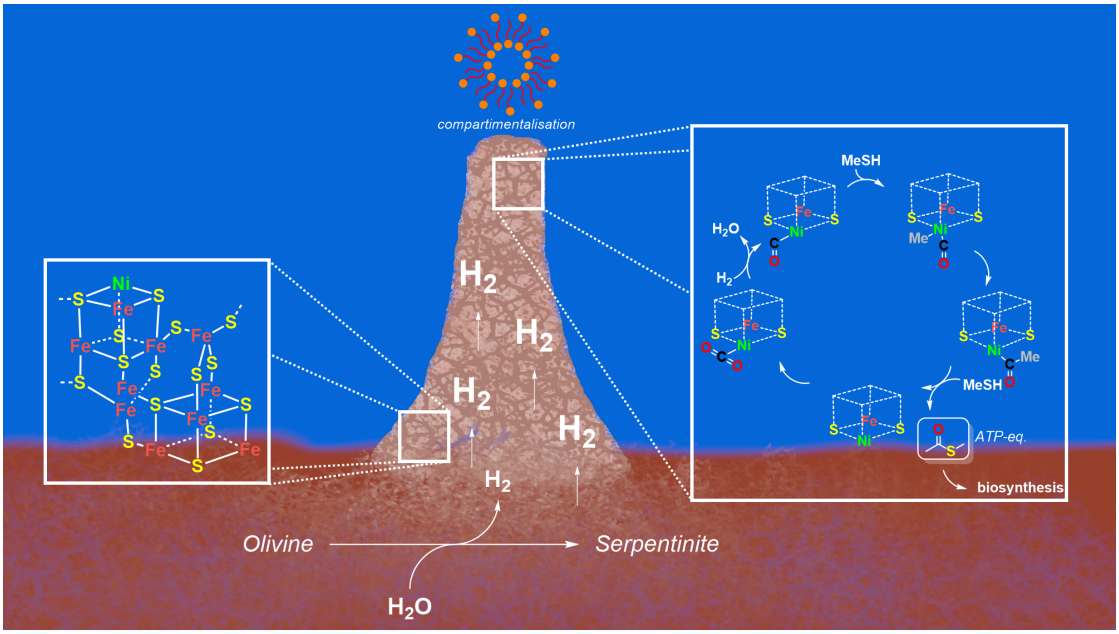


Figure 3.2: Schematic illustration showing the transition from inanimate matter to animate matter according to the *Mechanism first* theory.

The discovery of another class of hydrothermal vents,[205] called *Lost city* lighted the debate. *Lost city*, formed via the serpentinisation[206] of olivine. This iron and nickel catalysed process releases solvated hydrogen and methane.[207] The serpentinisation is exothermic, providing water temperatures of 40 –100 °C under alkaline conditions, see Figure 3.2. The chimneys formed via this process are rich in transition-metal-sulfides and exhibit small cavities, providing the scene for compartmentalisation. At

this interface, it is believed that the transition-metal-sulfides catalysed the conversion of methane and hydrogen to species of higher molecular weight.[175,208] The result might be acetyl thioesters. From these, the transition to more complex chemical processes starts. Molecules, able to act in compartmentalisation were produced, separating the networks from the dilute surrounding. At some stage, RNA evolved, followed by the ribosome, producing enzymes. In consequence, DNA could evolve after the reduction of RNA.

Evidence for such a scenario is seen in archaea bacteria, which fix CO_2 using nickel. Further, the modern enzyme *acetyl-CoA synthase* is based on nickel,[209] giving another point of evidence. In certain microbes, acetyl-CoA is an energy currency directly involved in the synthesis of ATP and in modern autotrophic organisms all synthetic pathways, except for the Calvin cycle, start from acetyl-CoA. Last, the chemiosmotic coupling of energy conservation to an ion gradient is also believed to originate from scenes like *Lost city*, as proton gradients established at the membrane-like interfaces.[175]

Approaching the problem of the emergence of Life chemically is one possibility, the other is a physical approach. Its main objective is to find physical scenarios, like temperature-gradients,[210,211] gas-water interfaces,[212] pH-gradients,[213] flow procedures and capillary effects[213] to explain the formation of certain biomolecules. However, this is not a solitary procedure to explain Life's origin but needs robust chemical reactions or reaction networks underlining the predictions. In principle, the focus in explaining the emergence of the biosphere is on: *Which* reactions needed to proceed to achieve a living organism and not *where* did they proceed.

3.4 Systems Chemistry

How could life evolve, continuously getting more complex, although the second law of thermodynamics predicts the complete opposite? Why does Darwinian evolution seems to disobey the law, taught at the very beginning of every scientific academic study? Is there another physical law that is not yet discovered?

Motivated by the assumption that there was not a sudden transition from inanimate matter to animate,[214] there must be another rationale to explain the out-of-equilibrium state of Life. The *systems chemistry* approach tries to provide a rational explanation for this transition. Possibly many intermediate stages on the convergence to *aliveness* were present. These stages must be temporarily stable enough to have the chance to change towards higher complexity. This requirement excludes a population by the second law of thermodynamics and the Boltzmann distribution.[215] Temporary stability can be achieved by a kinetic energy barrier. The constant dissipation of energy, the exponential re-growth of the stages constituents and their slow decay needs to be fulfilled to ensure a longer lifetime. Following this model, not thermodynamics is driving the complication of matter, but kinetics in a *dynamic kinetic stability* model.[215]
Within this model, see Figure 3.3, there are two mathematical descriptions to define stability. Either by reduction of energy and increase of entropy (second law) or by persistence over time in a metastable state. Stability in the latter case does not mean the inability to react, as in the first case.[216] Life can be described by the latter description, utilising the kinetic power of exponential growth, driven by energy influx to keep the system in a non-equilibrium steady state. One possibility to achieve non-equilibrium states are non-linear, autocatalytic cycles consuming energy.

Two autocatalytic cycles consuming the same source will lead to the extinction of the slower by the faster and more efficient. This principle inherits Darwinian evolution at a molecular level. However, combined autocatalytic cycles are proposed to perform better as selfish ones.[217] The complexity of the autocatalytic system can be increased indefinitely by introducing feedback loops, inhibition cycles and promotion cycles. Subsuming the presented, the emergence of life needs nonlinearity, replication or autocatalysis and the dissipation of free energy to occur. Another key requirement is the irreversibility of the process at the cost of dissipative energy. Based on this theory

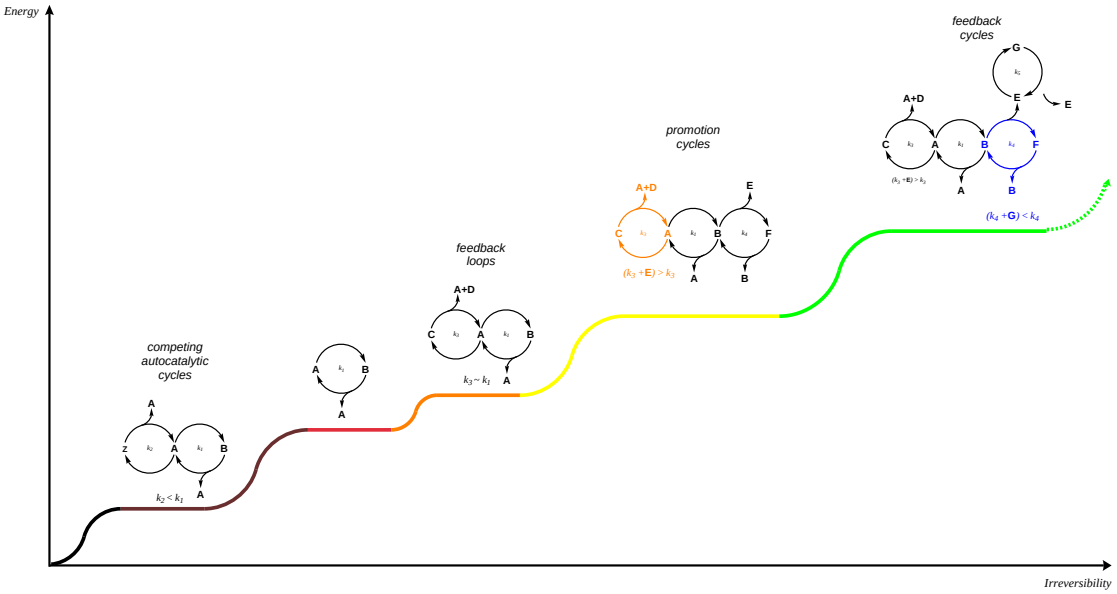


Figure 3.3: Schematical conversion from inanimate to animate matter *via* dissipation of energy in autocatalytic cycles, including molecular evolution, promotion and inhibitions cycles.

the emergence of life via a *mechanism first* model is not explainable, as the reduction of CO_2 by hydrogen is not sufficient to pay the price for irreversibility, however higher carbon analogues like HCN **5** would be possible.[218,219]

The *dynamic kinetic stability* model still obeys the requirements of the second law, however being persistent by an ever-improving, multiplying and replicating system, which populates higher states not according to the Boltzmann distribution. In this context, the survival of the fittest can be translated to the persistence of the most stable form.[220] This concept provides a rationale for the emergence of Life on Earth, while the model does not predict the result of the process. Furthermore, the question must be answered how general such replicative systems are and if there are any, apart from the nucleic acid-protein combination.

Systems chemistry[176] is a relatively young field trying to gain a deeper understanding of the chemical character of the replicative growth and find answers about the generality of this chemistry. Inspired by Nature, it tries to understand and recreate the complex machinery of modern biology and its supramolecular assemblies. Constant

dissipation of energy is needed to keep the system far from equilibrium, the main criteria of Life. In a first step, simple self-assembling systems are searched which behave like natural systems would.[221,222] Deeper understanding of the generality of the chemistry can provide an answer to the question, if Life is restricted to Earth or if there are other species in the universe.

3.5 Critical Discussion

In the following, a brief discussion about the presented theories is given. The discussion is given to the best of the author's knowledge and does not claim absoluteness.

Exaggeratedly said, in a primordial world governed by RNA, the bio-polymer has the role of an omnipotent *godfather*. Every problem of such a complex process is projected to the replicating and catalysing capability of RNA, further enhanced by its ability of information storage. Such an approach is by far too short-sighted, complemented by the fact that no prebiotic synthetic pathway to RNA polymers was found. This is under the major criticism that RNA is too complex to develop from scratch out of a prebiotic mixture.[182] Putting all expectations of primordial chemistry to one shoulder is too simplistic and the fact that this incident happened out of thin air suggest that Life was a random incident.

This evaluation does not exclude RNA as part of a primordial world, however, a sole, omnipotent role is questioned. Within a primordial world, the development of cofactors for catalysis and the development of intron/exon splicing in combination with RNA are logical possibilities. A simpler, easier achievable pathway towards RNA needs to be found, which are not necessarily high yielding but continuously produce the building blocks. Further, the transition to RNA oligomers needs to be outlined.

To overcome the problem of a single incident producing RNA, the emergence of an evolutionary procedure was proposed. Generally, the hypothesis of a slow evolution of RNA sounds appealing as it substitutes the randomisation by an evolutionary process, a natural principle. From a logical point of view, it seems more reasonable to find a direct pathway to nucleosides, without the indirection of evolution. In this case, human

time-efficiency is implied, which not necessarily counts for Nature. A simultaneous formation or a slightly delayed formation of both the precursor polymers and RNA nucleosides can be envisioned. In this scenario, the concomitant synthesis of tPNA-like structures and RNA nucleosides could result in the assistance of oligomerisation towards RNA.

Amongst the *mechanism first* supporters, the omnipotent role of RNA and its complex synthesis is criticised. An omnipotent mechanism is envisioned, producing Life's critical molecules from H_2 and CO_2, catalysed by transition-metal-sulfides. This is an equally incomplete approach to another direction. The proposed potent molecules are thioesters, later found as acetyl-CoA in modern biology. However, no distinct transition from these simple thioesters to more complex structures is given. For the production of heterocyclic compounds, feedstock gases need to include a nitrogen source. The remote location of deep-sea vents prevents access to atmospheric nitrogen, probably its most abundant source. Additionally, the energy needed for N_2 activation is not provided by the conditions found at deep-sea hydrothermal vents. Recent ideas, thus, set the hydrothermal vents to the surface.[223] This addresses not a chemical problem, but a geological one.
Its main statement is that the chemiosmotic coupling in modern cells originates from the circumstances at Hadean deep-sea vents. Evidence for this claim is seen in archaea bacteria gaining energy by metabolising hydrogen and methane. This focuses on a very specific problem and its solution, regardless of whether all accompanying problems are equally solved by this approach. The theory seems far too simplistic as it implies many assumptions. Many boundary conditions are taken as given *e.g.* that cofactors are needed in the first place to harness energy from the reduction of CO_2 by H_2, but no conception is given of how these heterocycles can form under these conditions. To date, no empirical evidence for such a scenario is provided. Above all, this approach implies a logical discontinuity in the transition from mechanism-based Life to Life centred around replication.[224]
In series with all the above-presented hypotheses of the emergence of Life, *systems chemistry* gives the most plausible hypothesis. It combines aspects from every previous hypothesis and thus provides a holistic approach to the field. Nature does not *think* in single steps, relying on single molecules, in single events explaining every-

thing. Nature is a finely balanced interplay of reactions in dependency to each other (promotion, inhibition, feedback). Therefore, the main objective of researchers trying to understand the origin of Life is to find auto-catalytic cycles and networks which resemble Nature. Dissipative, supramolecular structures need to be found, located far from equilibrium. This field is in its infancy and very difficult to verify, yet the most promising. The research on single reactions is necessary, however, they always need to be set to a big picture of reactions communicating with each other, ever-evolving, ever adapting and thus slowly transforming to Life.

3.6 Setting the Scene

The hypothesis underlying this thesis takes a critical view on the ideas of the *RNA world* and the *mechanism first*-hypotheses and considers them as too short-sighted.

This work proposes a simultaneous emergence of RNA and DNA at the advent of Life, set to a holistic approach. It combines the advantages of a flexible, adaptable, catalytically active biopolymer with a chemically stable, informational biopolymer, working hand in hand. Therefore, there is no need for a late emergence of an enzyme corresponding to the *ribonucleotide reductase*. The starting point of synthesis is not set at the disconnection of the anomeric bond. Increasing the plausibility of nucleoside pathways requires simple, chemically robust reactions, implying a small set of precursor molecules. Reaction conditions are required, which omit the need for drastic changes. To the best of our knowledge, the focus on high yielding pathways is obsolete. From a *systems chemical* point of view, high-yielding reactions are thermodynamically driven. However, the creation of Life requires kinetically driven reactions, which thus can deflect in any direction. Prebiotic conditions are proposed to imply the following requirements:

1. Aqueous media[33]
2. Simple starting materials, gained from high-energy gas-phase reaction[225]
3. No serious changes to the initial conditions

4. No sequential additions of reagents
5. Moderate temperatures ($0\,°C > T < 100\,°C$)[226,227]
6. Dissolved-metal salts and porous surfaces as additives or catalysts[228,229]

Maintaining similar conditions for a consecutive synthesis of nucleosides calls for a different approach of pyrimidine and purine synthesis. The nucleobase formation can be intercepted by nucleoside formation and thus overcome nucleobase stability issues. The same principles should apply to cofactor synthesis, as they are ubiquitous parts of modern bio machinery. The investigations executed in this thesis are situated on a molecular level, however, the results must be put to a universal entity.

4 Methodology

This thesis focuses on the discovery of genuine and comprehensible prebiotically plausible one-pot scenarios towards nucleosides, deoxynucleosides, non-canonical nucleosides, nucleobases and cofactors. In this context, new synthetic pathways towards DNA **3** and RNA **2** are searched for. Further, an upstream preceeding pyrimidine synthesis, resorting to a similar cut-set of starting molecules is searched for. The theoretical knowledge gained on prebiotic sugar synthesis at heterocycles, is applied to a logical approach for the synthesis of riboflavin cofactors. In the following, the methodical considerations to achieve these objectives are presented, see Figure 4.1 It is outlined, how the evaluation of the accuracy of the data and the collection of the data, was executed. At last, the tools to analyse the experimental outcome are presented alongside the reasons for choosing this data collection method.

As presented in Section 2.3, the prevailing prebiotic syntheses of ribonucleosides and deoxyribonucleosides set the critical disconnection at the nucleosidic bond. Therefore, the synthesis of DNA nucleosides **3** was not available so far. In this thesis, a prebiotically plausible pathway towards **3** is designed by alteration of the critical disconnection, in analogy to aminocatalysis.[230,231]

Consequently, a new logical approach is pursued to gain control over the connectivity and stereochemistry of the reaction as well as to increase the simplicity of the system. The same logical approach is transferred to gain accessibility to RNA nucleosides **2** under the same conditions. So far, no convincing, simultaneous pathway to **3** and **2** is published, which does not rely on multiple step procedures or on external influences.

Approaching the emergence of Life holistically, *cf.* Section 3.6, a concomitant synthesis of **3** and **2** would be beneficial. The ideal conception is to find such a simple, natural

pathway, which gives access to **3** and **2** with the intrinsic regulation of stereochemistry. Upon investigating the prebiotic formation of **3** and **2**, possible evolutionary non-canonical pre-products, like deoxy apio-nucleosides **67** were identified.

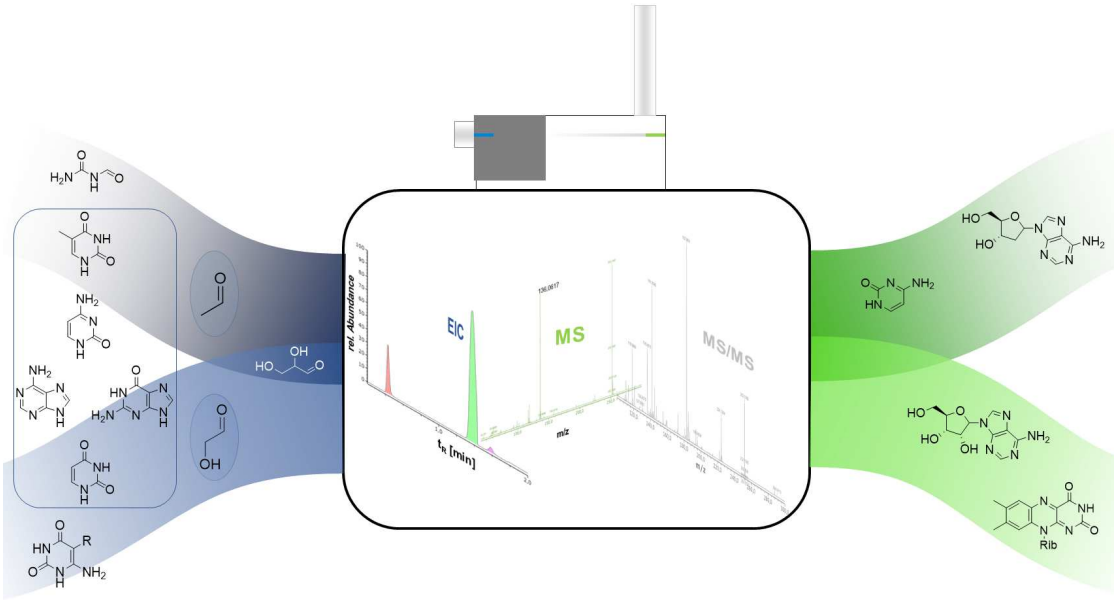


Figure 4.1: Schematic methodology applied in this thesis.

A cytosine **1d** or pyrimidine synthesis, preceding the nucleoside synthesis, is sought after to overcome the problem of instability of **1d**.[81] Similar to the synthesis of **3** the composition of **1d** is based on acetaldehyde **35**. A superordinate aim, not addressed in this thesis, is to reveal concepts enabling an in-line synthesis from basic starting materials to DNA polymers via **3**. Furthermore, it is one possible answer to prevent **1d** from hydrolysis and make it more plausible for the origin of Life.

Riboflavin **55** is one of the most versatile cofactors in modern biology, yet no complete answer to its prebiotic formation was published. The crucial point in the prebiotic synthesis of **55**, is the linkage of a ribityl-residue to a heterocycle. Picking up previous ideas[103,104,165] and extending it with our concept of sugar-formation on heterocycles, illustrates a possible pathway to **55**.

The results collected are interpreted by empirical methods and proven to be valuable by different, independent methods. The evidence for the results is evaluated to

the best of our knowledge by independent methods.

Data collection proving the formation of **3**, **2**, **67**, **1d** and **55** and their preceding intermediates was accomplished by UPLC-analysis. Reversed phase liquid chromatography was implemented. The compounds were either monitored by downstream UV/Vis-detection or downstream HRESI-QTOF-MS. Further, high resolution mass spectrometric data was collected with a HRESI-OrbiTrap-MS. Proof for the reliability of the results was accomplished by comparing their retention times with in-house synthesised or purchased reference samples. Fragmentation patterns and coelution scores for the pure samples were collected and compared to the data resulting from the prebiotic mixture. Further, spiking the original reaction mixture with authentic references, verified the reliability of the results. HRMS spectra were collected as all ions-, targeted-MSMS- or auto-MSMS-spectra. Additionally, analysis of blank samples confirmed the exclusive formation in the reaction mixture and eliminated possible contamination.

The above described analytical methods were chosen in order to obtain independent, orthogonal data about the constituents of the complex mixtures. Distinct statements about the isolated species can be given, as overlapping with interfering species is excluded. The data is collected at a high resolution, which enables definitive assignment of the gained structures and leaves little space for speculative scope.

The program *Agilent MassHunter* was used to acquire the data using the HRESI-QTOF-MS spectrometer. The data was analysed with *Agilent MassHunter Qualitative Analysis*. Targeted analysis was performed by extracting the ion chromatogram of the desired species with an error of 5 ppm. Calibration curves, standard addition procedures and batch processing of data was accomplished with *Agilent MassHunter Quantitative Analysis*. Molecular feature extraction, the summary of connectable fragments, was executed with *Agilent ProFinder*. The blank subtraction and preparation for autoMSMS experiments was achieved with *Agilent MassProfiler*. A database of all reference samples, consisting of the name, its molecular structure, the molecular formula, the ionisation preferences, the exact mass and its fragmentation pattern at 10 V, 20 V, 40 V was set up in *Agilent PCDL database*. The entries of the database

were generated by acquisition of targetedMSMS spectra of pure, authentic samples. Comparison of the database with the experiments was performed executing the *find by formula* tool implemented in *Agilent MassHunter Qualitative Analysis*.

5 Results & Discussion

In this chapter, the experimental results of this thesis are described and discussed. First, motivating experiments, regarding the sugar selectivity of earlier described FaPy **32** and 2-aminoxazole **33** are presented. Second, a new rationale for the synthetic access of DNA nucleosides from all canonical nucleobases **1** and available aldehydes is described. The stereo- and regioselectivity of the reaction are examined and delineated. Alteration of the aldehyde reactants lead to the discovery of DApiNA-nucleosides **67**. Their conventional synthesis is presented. Further, the synthetic access to pyrimidine nucleobases is investigated. Another variation of the aldehydes led to a novel, simple pathway towards RNA-nucleosides **2**, accompanied by the illustration of its stereo- and regiocontrol. Last, rational prebiotic access to the riboflavin **55** cofactor is illustrated. In general, a novel disconnection of the canonical ribonucleosides is developed and examined for its validity. In consequence, the concomitant synthesis of **2** and **3** can be achieved under prebiotically plausible conditions, in a very simple and elegant manner.

5.1 Motivational Work

In this section, time-resolved experiments are described, which examine the sugar selectivity of the previously discussed heterocycles FaPy **32** and 2- aminooxazole **33**. In these contributions, it is stated that the depicted heterocycles lead to the corresponding ribonucleosides in high yields, see Chapter 2. **32** is reacted with pure **4a** and **33** is reacted with pure **10a**. The same reaction outcome is claimed when starting from a sugar mixture, like the formose reaction. Initiated by the kinetic description of the formose network by K. Kohler in our group, the reactivity of the depicted compounds was examined in the reaction with various sugars.

Curious about the sugar-selectivity of the described reactions, **32** and **33** were

reacted with glycolaldehyde **9**, glyceraldehyde **10a** and ribose **4a**. The depicted aldehydes possess a similar carbonyl electrophilicity, however increasing reciprocal with the size of the sugar.[232] From a chemist point of view, the reaction kinetics of **9** and **10a** with the depicted heterocycles should be consequently faster, than with ribose. The possible products are depicted in Figure 5.1. In the case of **33**, both depicted species are thinkable (middle and right), as the nucleophilicity of the free amine might be higher. These products however would not lead to the desired ribonucleotides, *via* the described pathways.

The availability of pure **4a** under prebiotic conditions is intensively doubted.[233] Further, the prebiotic formation under formose-like conditions is also discussable. However, small sugars like **9** and **10a** are abundant in these mixtures. Hence, the reactivity of **9**, **10a** and **4a** in reaction with **32** and **33** are compared separately, to predict their reactivity in a formose-mixture.

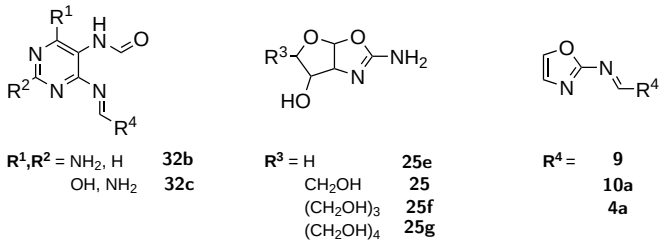


Figure 5.1: Possible **32**- or **33**-sugar adducts

The formation of the corresponding **9**-, **10a**- and **4a**-adducts was monitored by UPLC-QTOF-MS. Aiming for comparability of the reactions, these were performed under standardised conditions, not resembling the reported conditions.[105] The reactions were examined in phosphate buffer (pH=7) at 50 °C. Both heterocycles **32** and **33** were initially stirred in the buffer to ensure adaption to the reaction conditions and their resolution. A blank sample was analysed to prove the absence of any sugar-adduct species before analysis. The corresponding sugar was added to the equilibrated **32**/**33**-solution and instantly a sample was taken (t=0 min) and analysed by UPLC-QTOF-MS. The progression of the reaction was monitored every 10 minutes *via* UPLC-QTOF-MS. The experimental data is depicted in Figure 5.2.

The results strongly indicate a behaviour contrary to the reported. The reaction of **32b**,**c** with **4a** is the slowest in this reaction series, first showing a significant detection

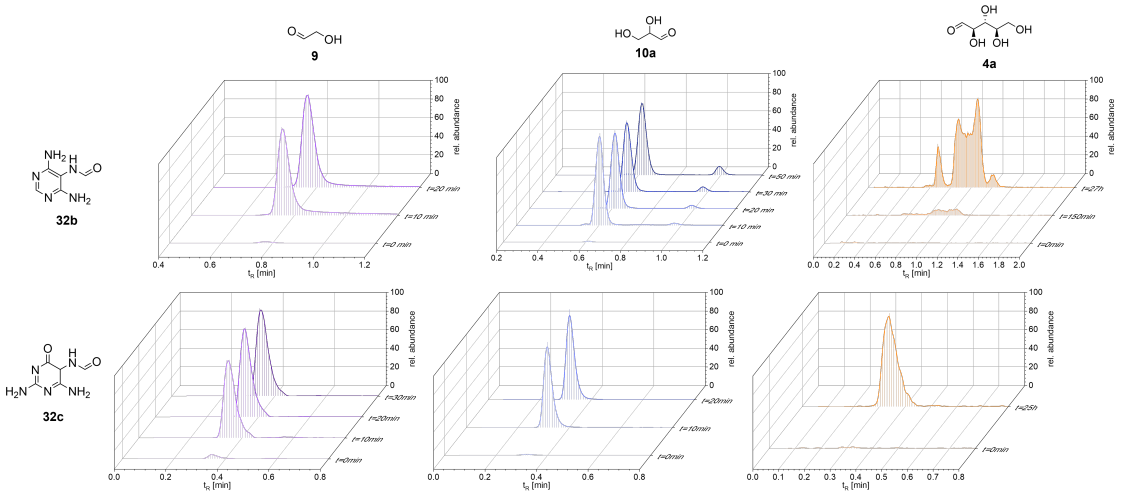


Figure 5.2: Depiction of the time-resolved reaction of **32b** and **32c** with **9** (violet), **10a** (blue) and **4a** (orange). An aliquot of 5 µL was taken in the depicted time intervals and analysed by UPLC-QTOF-MS. QTOF-MS was performed with the following applied mass filters (*m/z* 196.0819–196.0839, 226.0923–226.0946, 286.1132–286.1160, 212.0768–212.0788, 252.0872–252.0896, 302.1080–302.1110;)

within a day. On the opposite the reaction of **32** with **9** and **10a** proceeds within 10 min. In the reaction between **32b** with **10a** a second species develops (t_R = 1.0 min), seemingly at the expense of the first.

In conclusion, offering **32** to a formose sugar mixture would lead to a vast distribution of products, based on the fast reaction with **9** and **10a**. In the case of present **4a** in the formose reaction mixture, it would be outruled by the smaller sugars. In the conducted experiments, possible regioisomers were not distinguished. The correctness of all formed species was positively confirmed by targeted MS/MS experiments, showing a characteristic fragmentation to give the corresponding heterocycle. These results clearly show the lack of selectivity, arising by reacting **4a** with heterocycles like **32** on the road to ribonucleosides. In terms of availability and reactivity, the deployment of smaller C_2- and C_3-sugars would be more favourable.

A different approach was envisaged, circumventing the usage of **4a** in the context of prebiotic ribonucleoside formation.[18] **33** is here reacted with **10a** to give the pyrimidine nucleosides *via* the intermediate formation of **25**, see Section 2.5.3. Offering **33**

to a formose-like mixture of sugars, the sole reaction of **33** with **10a** is discussable. A plethora of **33**-sugar adducts should be the result. To test this hypothesis, the kinetic formation of **33**-adducts with **9**, **10a** and **4a** was studied. The analysis was performed under the same circumstances described for the **32**-species.

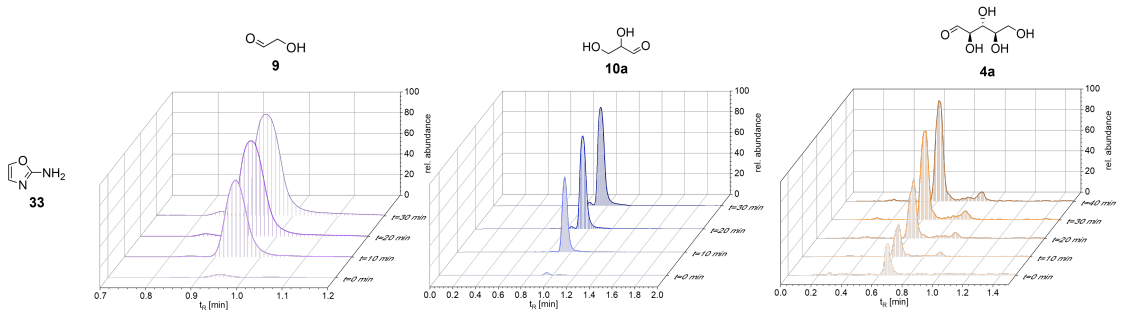


Figure 5.3: Depiction of the time-resolved reaction of **33** with **9** (violet), **10a** (blue) and **4a** (orange). An aliquot of 5 was taken in the depicted time intervals and analysed by UPLC-QTOF-MS. QTOF-MS was performed with the following applied mass filters (*m/z* 228.095–228.099, 145.0601–145.0615, 175.0705–175.0721, 235.0915–235.0935;)

33 reacts fast and unselective with all three species. It is noteworthy that already upon mixing a solution of **33** with either **9**, **10a** or **4a**, the corresponding sugar-adduct can be detected. The formation of the **9**- and **10a**-adduct proceed in the same order of magnitude, whereas **4a** reacts slightly slower. The structural relationship was confirmed by targeted MS/MS experiments. The question of whether the sugar aminooxazoline or the sugar imine is formed could not be addressed in these experiments.

The fast reactions leave no space for the preferred formation of **25** over its C_2-counterpart. Even the reaction with **4a** would perform fast. Expecting a selective formation of one species by offering a formose reaction mixture is futile. For **33** reacting with **10a** four species with different retention times could be identified, after a reaction time of 1 h. These results are following the reported mixture of **25a–d**.[18] Significantly faster development of one species was observed. The correct assignment of the fast- developing species was not provided.

Such fast reaction kinetics leave minor space far an unambiguous selection. This would mean ending up in one thermodynamic equilibrium after another, which is not the criterion for developing Life.

There must be prebiotic access to nucleosides, circumventing the presented difficulties. Instead of **4a**, these pathways should be based on short-chain sugars, which are

preferably formed in the formose reaction. It should overcome the difficulties related to the selection of the sugar. This new approach should implement finely balanced reactions instead of high-yielding reactions leading to the thermodynamical equilibrium. Last, it should open up a pathway to deoxyribonucleosides. The research on a pathway implementing the listed boundary conditions is presented in the following.

5.2 Prebiotic DNA Synthesis

In the past, the synthetic focus was almost exclusively set on the prebiotic access to ribonucleosides. Assuming a *RNA world* at the emergence of Life, the question of how 2'-deoxy ribonucleosides **3** can be achieved, was generally disregarded. Only recently,[129,131] the topic gets into focus, doubting its development from **2** catalysed by *ribonucleotide reductase*. A structure so essential to Life, like DNA, must have developed much earlier than previously assumed. Learning from the presented sugar-selectivities in reaction with heterocycles, a synthetic procedure relying on small starting materials is searched for. In this section the theoretical and experimental work leading to the prebiotic synthesis of **3** is described and discussed. Parts of this section were investigated in collaboration with J. S. Teichert and previously published.[234]

5.2.1 Theoretical Considerations on 3

All presented pathways towards **2** contribute beautifully to the topic of prebiotic accessibility and have their right to exist. However, at some point they suffer from one of the presented drawbacks, see Chapter 2. The presented limitations are:

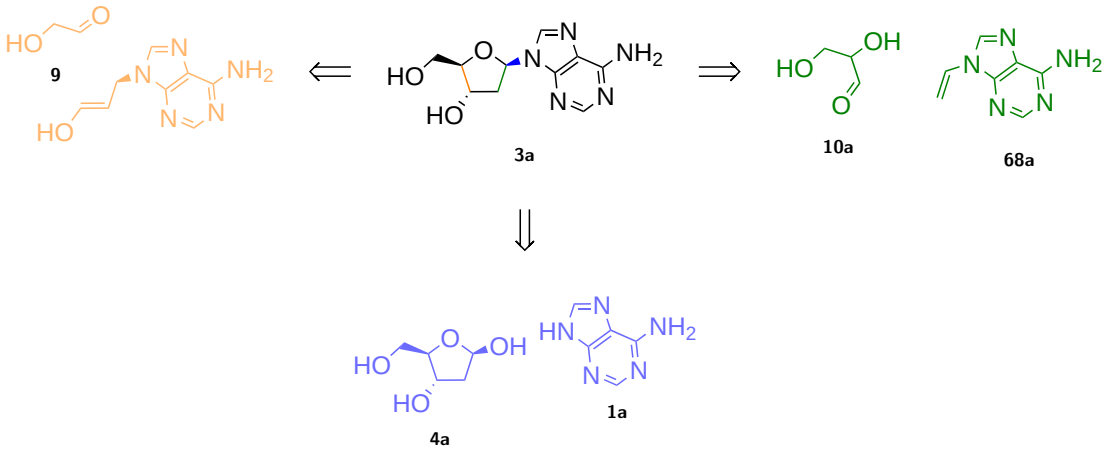
- the difficult presence of **4a** and **d4a** under prebiotic conditions
- multistep procedures with a plethora of different reaction conditions
- issues on regio- and stereoselectivity
- yet no convincing pathway towards **3**

In addition, following these pathways no direct access to deoxy ribonucleosides can be accomplished. Assembling **3** as **2**, would call for free **d4a**, which cannot be synthesised under prebiotic condiditions, nor it is stable under these conditions. Aiming for

the prebiotic access to **3** requests a different synthetic rationale.

Reflection on the so far published routes to ribonucleosides opens up some key questions: 1) *Is there a way to disconnect nucleosides differently?*, 2) *Does a new pathway cause simpler reaction steps/conditions?*, 3) *How would a change in disconnection open up new pathways towards nucleosides?* and 4) *Can we mimic Nature and implement catalytic cycles for nucleoside production?*

In the following, approaches are presented to give answers to parts of these questions and thus creating access to **3**.



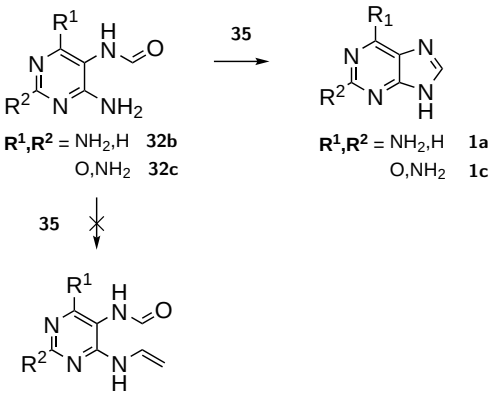
Scheme 5.1: Possible disconnections of DNA nucleosides the scission of **3a** is depicted for clarity, the presented disconnection applies to all **3**, green: novel disconnection, blue: Fischer disconnection, orange: a third possible disconnection with **9** as starting material)

For a century, ribonucleosides were disconnected at the anomeric bond (blue), however, theoretically there are two other scissions possible, see Scheme 5.1 The first thinkable disconnection (orange) gives a malondialdehyde derived nucleobase and **9**. The **4a** problem is clearly circumvented. Implementing **1** as starting materials, this pathway would possibly suffer from the multiple reactivity of the malondialdehyde including the reaction of **1c** to form M_1dG.[235] This fused heterocycle of **1c** is a commonly known exocyclic mtDNA adduct, caused by oxidative stress. Requiring simpler and selective reaction conditions, this dual reactivity is not beneficial.

Another disconnection (green) is set at the 2'-3' carbon bond and offers a vinylated

nucleobase **68** and **10a** as possible starting material for the formation of **3**. Cutting **3** as presented omits the use of **d4a**. The pathway convinces by simple, readily available starting materials and high, directed and thus controllable reactivity. Additionally, it provides uniform access to both the pyrimidine and purine nucleosides. **68** can be achieved by reacting **1** with a C_2 species *e.g.* acetylene or its hydrolysis product **35**. In rethinking the disconnection of nucleosides, a possible pathway is delineated, starting from the simple and readily available reactants **1**, acetaldehyde **35** and **10a**.

In a very first attempt and inspired by the approach of Eschenmoser and Carell,[104,105] **35** was reacted with several **32** derivatives. The goal was to achieve an equal reactivity of **32** with acetaldehyde as with several sugars, see Section 5.1. However, the desired reaction did not proceed and an acetaldehyde adduct of **32** was not detectable. Surprisingly, a significant quantity of the **32**-corresponding nucleobase **1** was found in all experiments, see Scheme 5.2. It was therefore concluded to directly start from **1**, circumventing possible troubles with an additional reaction step.

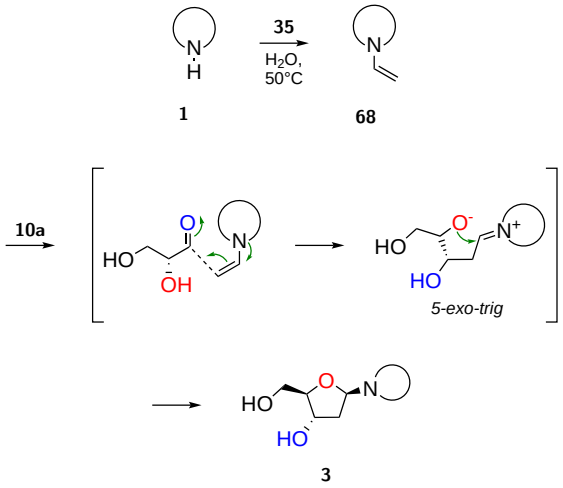


Scheme 5.2: Reaction of acetaldehyde **35** with **32**

Starting from **1** applies the nucleophilicity of *N*9 in purines **1a**,**c** and *N*1 in the pyrimidines **1d**,**e**. The nucleophilicity of these specific sites was quantified by Mayr *et al.*[236] and correlates with a slight acidity of the moieties. (**1a** (pK_a 9.80), **1c** (pK_a 9.21), **1e** (pK_a 9.94) and **1d** (pK_a 12.2)) In consequence, **1** reacts with **35** to give the vinyl nucleobase **68**, an enamine when speaking in terms of aminocatalysis.[230,231] This analogy inherits the catalytic potential and thus implements an answer catalytic cycles for the production of nucleosides.

Following this disconnection, **10a** is revealed as the reaction partner. This scission is beneficial, as it evades *the notorious nucleosidation problem* of free **d4a**. In addition, it was shown in the group, by K. Kohler that C_3-sugars as **10a** and **10b** are readily available from a prebiotic formose reaction mixture.

The nucleophilic attack of **68** at the carbonyl of **10a** from the Si-face takes place in a stereo controlled manner,[237] see Scheme 5.3. The subsequent *5-exo-trig* ring closure[238] is kinetically favoured and exclusively furnishes the furanose ring, leading to the β-deoxyribonucleoside **3**.



Scheme 5.3: Proposed, highly stereo- and regiocontrolled mechanism towards **3** (0.10 mmol **1**, 1.00 mmol **35**, 0.15 mmol **10a** in 1 mL H_2O for 4 days at 50 °C).

5.2.2 Experimental Work

Several reaction conditions were screened including variation of the educts, the pH, the implementation of several additives, variation of the temperature and solvents. Different concentrations of all reaction partners were also tested. The tested conditions are listed in the table below, see Table 5.1. As expected, the reaction only proceeds in neutral to alkaline media to provide the nucleophilicity by partly deprotonating the nucleobase at nitrogen *N*9 and *N*1. Additives promoting neutral to alkaline conditions are successful in enhancing the reaction, however severe alkaline media prevents the nucleosides to form. The optimum is roughly around pH 8. No conversion was observed in an acidic environment. The conditions are considered as mild as just

benign additives or no additives favour the reaction. The conversion takes place at different temperatures in the range of 20–70 °C. At room temperature, the conversion rates are significantly lower and temperatures above 70 °C are too harsh. Several polar solvents, including water, formamide and methanol enable the reaction.

Table 5.1: Tested additives in the reaction towards **3**, performed in methanol:water (99:1) at 50 °C. OrbiTrap-MS was performed with the following applied mass filters (*m/z* 228.095–228.099, 243.095–243.099, 252.107–252.110, 268.102–268.106)

additive	***m/z* ($[M+H]^+$) of 3**
-	found
NaOAc	found
Na_2HPO_4	found
phosphate buffer pH 9	found
imidazole	found
$Ca(OH)_2$	found
$CuCl_2$	traces
$MgCl_2$	traces
phosphate buffer pH 5	-
$AlCl_3$	-
$B(OH)_3$	-
$Mg(OH)_2$	-
NaOet	-
NaOH	-
LiOH	-
HCl	-
AcOH	-
HCOOH	-

The final conditions to yield **3** are very mild and highly prebiotically plausible. The reaction proceeds best in water or methanol:water (99:1) at a temperature of 50 °C under the support of sodium acetate. Performing the reactions in either fused ampoule or in vials with screw-caps did not significantly change the outcome.

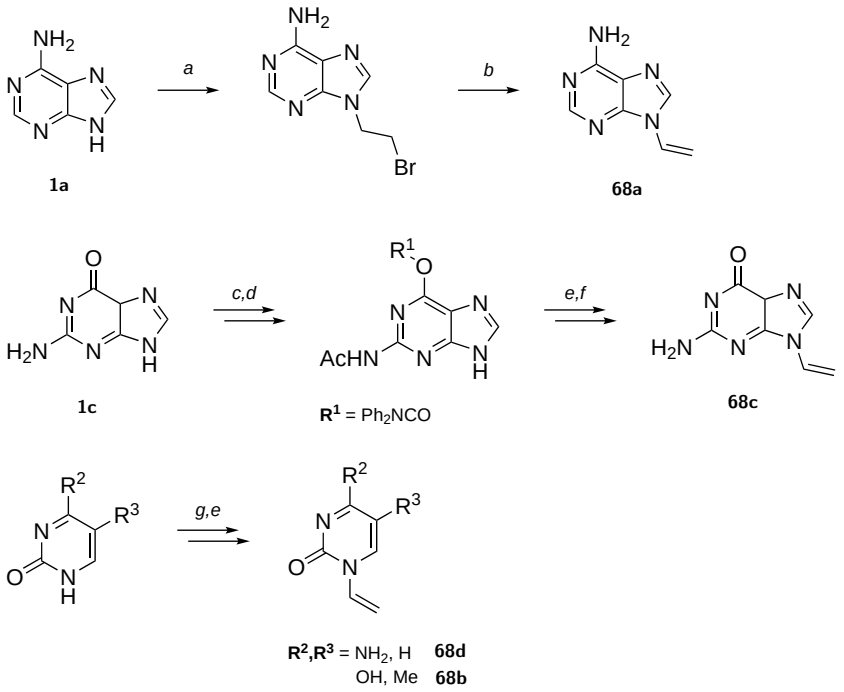
The gained nucleosides **3** were chromatographically separated *via* HPLC or UPLC

and detected in a downstream QTOF-ESI-MS or upon direct injection to OrbiTrap-MS. An isocratic elution did not suffice to separate the reaction mixture. Therefore, a published separation procedure for nucleosides was applied.[239] The chromatography includes a gradient from 100% water to 100% methanol (*cf.* experimental section). Methanol can be substituted by acetonitrile, which does not significantly change the performance of the separation, nor the retention times. In general, ionisation is increased. The described method is a powerful tool, in which a complete separation of all nucleobases **1**, vinylnucleobases **68**, nucleosides **2**, deoxynucleosides **3**, nucleotides **2**-**P** and deoxynucleotides **3**-**P** is achieved.

The formation of **68** was proven *via* coinjection with authentic samples, see Figure 5.4. These were synthesised following the procedures depicted in Scheme 5.4. Initially the procedure applied for **68a** was planned for all four species, however, it turned out that it is not feasible for **68b**–**d**. Instead, a mercury(II)acetate and TMSOTf assisted exchange of acetate by **1** in vinylacetate gave the correct products.[240]

During analysis, it was observed that the retention time of **68** varied and showed a distinct discrepancy. However, the analysed *m/z* remained constant. As an explanation, the concept of prebiotic protection groups was proposed. **35** might react with more than the desired nucleophilic site. As a consequence, intermediate protected structures are formed. These are observed *via* UPLC-separation on a C18 phase by longer retention times. In source however, the corresponding species are cleaved solely showing the correct mass of the pure **68**. This concept can also be applied to the prebiotic mixture itself. Explaining the exclusive formation of the canonical nucleosides, despite the presence of multiple nucleophilic sites. All sites are reactive, however, as soon as **68** is formed at the site for canonical nucleosidation, the reaction proceeds to form **3**. All further vinyl or sugar species act as protection groups. These are formed at all non-canonical nucleophilic sites, and are cleaved due to their instability. This concept provides a rational for the selective formation of canonical **3** and might even open up pathways to selective sugar formation, catalysed by the synthesis of **3**.

Coinjection with the reference samples showed the sole formation of **68** and **3**, see



Scheme 5.4: Synthesis of **68**. a) CH_2Br_2, K_2CO_3, DMF, r.t.; b) NaOEt, EtOH, r.t.; c) Ac_2O, pyridine, reflux; d)Ph_2NCOCl, pyridine 0°C to r.t., e) CH_2CHOAc, $Hg(OAc)_2$, TMS-OTf, reflux; f) NH_3, MeOH, r.t.; g) HMDS, TMSCl, NH_4SO_4;

Figure 5.4. The distinct assignment was verified by targeted MS/MS, OrbiTrap-MS UPLC-QTOF-MS and UV/Vis detection upon UPLC separation.

Due to ionisation abilities in ESI-MS the verification of the **1a**- and **1d**-species turned out to be much more feasible than its counterparts derived from **1e** and **1c**. Larger injection volumes of higher concentrated samples made a unambiguous identification possible. Noteworthy is the significant increase of retention on a C18 solid phase caused by the introduction of a methyl group.

5.2.3 Stereoselectivity of the DNA Synthesis

As theoretically outlined above, the presented reaction pathway should only accumulate the β-deoxyribonucleosides **3**. This is a consequence of selective attack of the nucleophile from the Si-face. An additional degree of regioselective freedom is provided by the choice between the furano- and the pyrano-species. According to theory, both are equally possible by either a *5-exo-trig* or a *6-exo-trig* cyclisation. In a

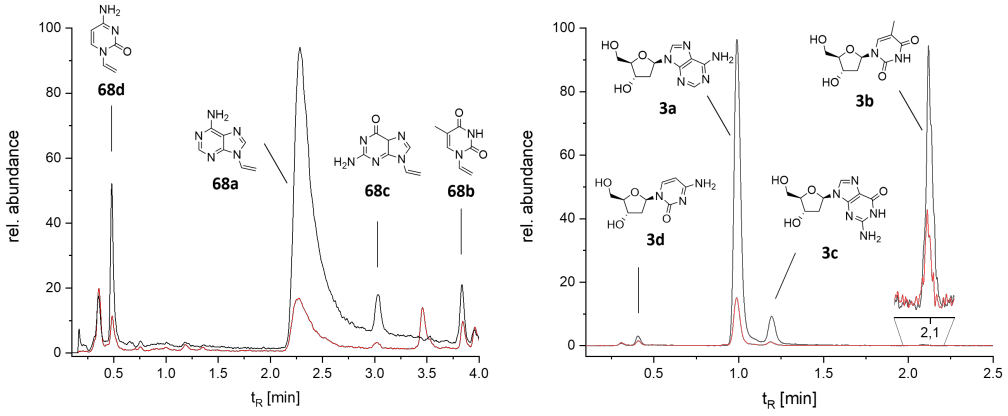


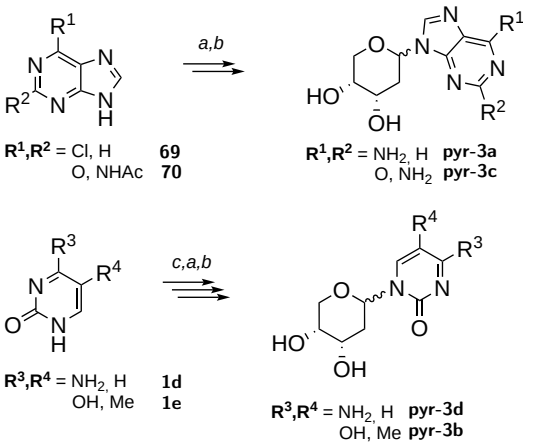
Figure 5.4: Coinjections (left side) with synthetic samples of **68a**–**d** showed the sole formation of the vinyl nucleobases **68**, as analysed by UPLC-QTOF-MS (mass-filter m/z 138.064–138.068, 153.063–153.067, 162.075–162.079, 178.070–178.074). Coinjections (right side) with synthetic samples of **3a**–**d** showed the sole formation of the deoxy ribonucleosides **3**, as analysed by UPLC-QTOF-MS (mass-filter m/z228.095–228.099,243.095–243.099, 252.107–252.110, 268.102–268.106). See 7 for experimental details.

kinetically driven reaction the 5-membered ring is favoured over the thermodynamic six-membered product. Beneficial for the plausibility of the prebiotic reaction would be a sole formation of the β-deoxyribonucleosides **3** in canonical furano selectivity. In comparison to all previously described pathways towards nucleosides, such a exclusive regio- and stereoselectivity for all pyrimidine and purine deoxynucleosides would be a novelty.

To verify the sole formation of ***β*-3**, the α-isomers were purchased for direct coinjection. The corresponding α/β-pyranosides of **3** were gained synthetically by applying Vorbrüggen nucleosiditation methods, see Scheme 5.5. The synthesis of the reference pyranosides is described in the following.

The reaction series leading to the α/β-pyranosides started with the acetate protection of **d4a**, giving a 1:1 mixture of the furanose-triacetate and the pyranose-triacetate. The pyranose-form crystallised at 0 °C and could be separated from the oily mixture *via* trituration with diisopropylether. Applying well-established Lewis acid assisted silyl-Hilbert-Johnson nucleosidation conditions[241] the pyrano-nucleosides could

be obtained after deprotection and purification *via* preparative HPLC. In case of the pyrimidines **1d**,**e**, the nucleobases were silylated prior to the nucleosidation step, whereas for the purines **1a**,**c** the nucleobases were silylated *in situ* during the nucleosidation step. Several nucleosidation protocols were tested, including the use of BSTFA/TMSOTf, BSTFA/$SnCl_4$, $SnCl_4$. Applying $SnCl_4$ in combination with HMDS and TMSCl gave the best results in all nucleosidation steps. It was further tested if other secondary *N*-heterocyclic nucleosides could be achieved *via* theses procedures. It turned out that these nucleosidation procedures just apply for aromatic *N*-heterocycles. Obtaining deoxyribonucleoside derivatives in the described fashion from these *N*-heterocycles did not succeed either. The structural integrity of the formed pyranose nucleosides was confirmed by 2D-NMR, crystal structures and HR-ESIMS, see the experimental section for further details.



Scheme 5.5: Synthesis of **pyr-3**. a) **71a**, HMDS, TMSCl, $SnCl_4$, acetronitrile, 50 °C; b) NH_3, MeOH, 60 °C; c) HMDS, TMSCl, NH_4SO_4;

Coinjection with the synthesised and purchased reference samples confirmed the exclusive β-furano selectivity for **3a**–**c**, see Figure 5.5. The chromatogram showing the correct configuration for **3b** was normalised to ensure better readability. The identification of **3d** in its canonical configuration remains. This deoxynucleoside will be verified in the future.

This novel design of the sugar-formation not only omits the need for pure **4a**, but accesses **3** exclusively as its canonical β-furano isomer. The reaction conditions are mild and conversion from simple starting materials are achieved in one-pot, without

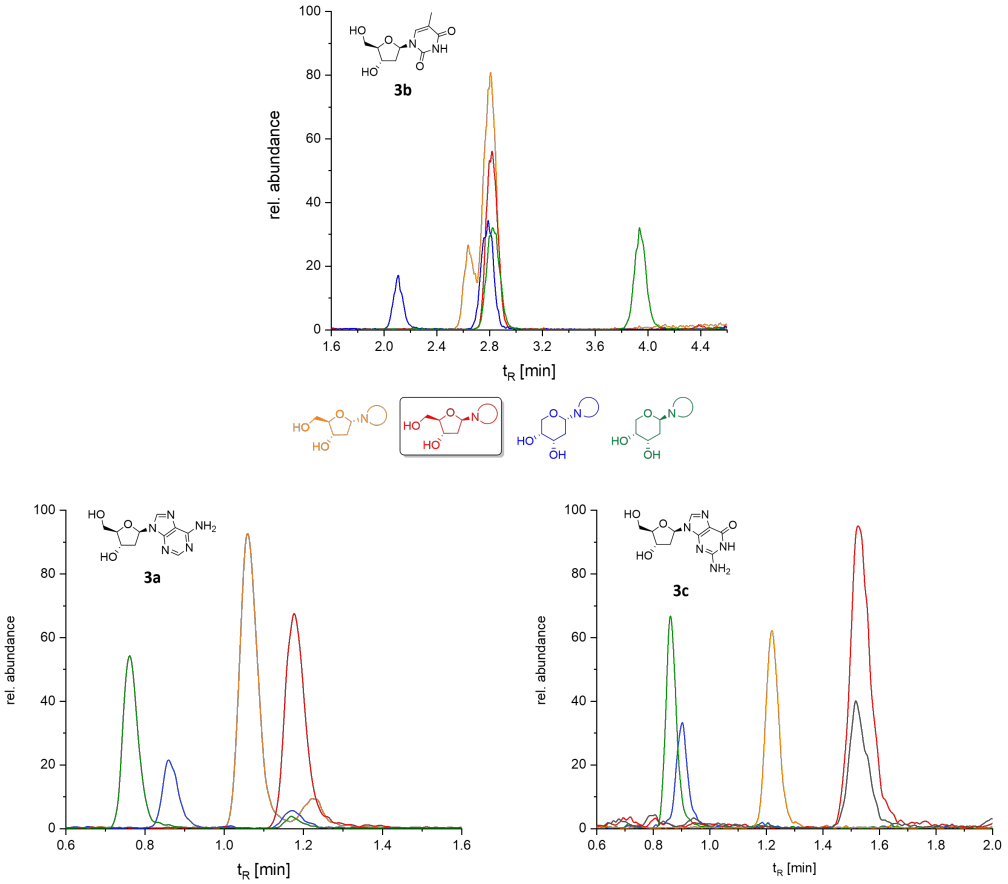


Figure 5.5: Coinjections with ***α*-3**, ***β*-3**, ***α*-pyr-3** and ***β*-pyr-3** shows the sole formation of the canonical β-deoxyribonucleosides **3**, as analysed by UPLC-QTOF-MS (mass-filter m/z 228.095–228.099,243.095–243.099, 252.107–252.110, 268.102–268.106). See experimental section for details.

the need for a change in reaction conditions. These results strongly indicate that the described prebiotic pathway towards **3** can be considered as the most plausible.

5.3 Elucidation of the DApiNA Species

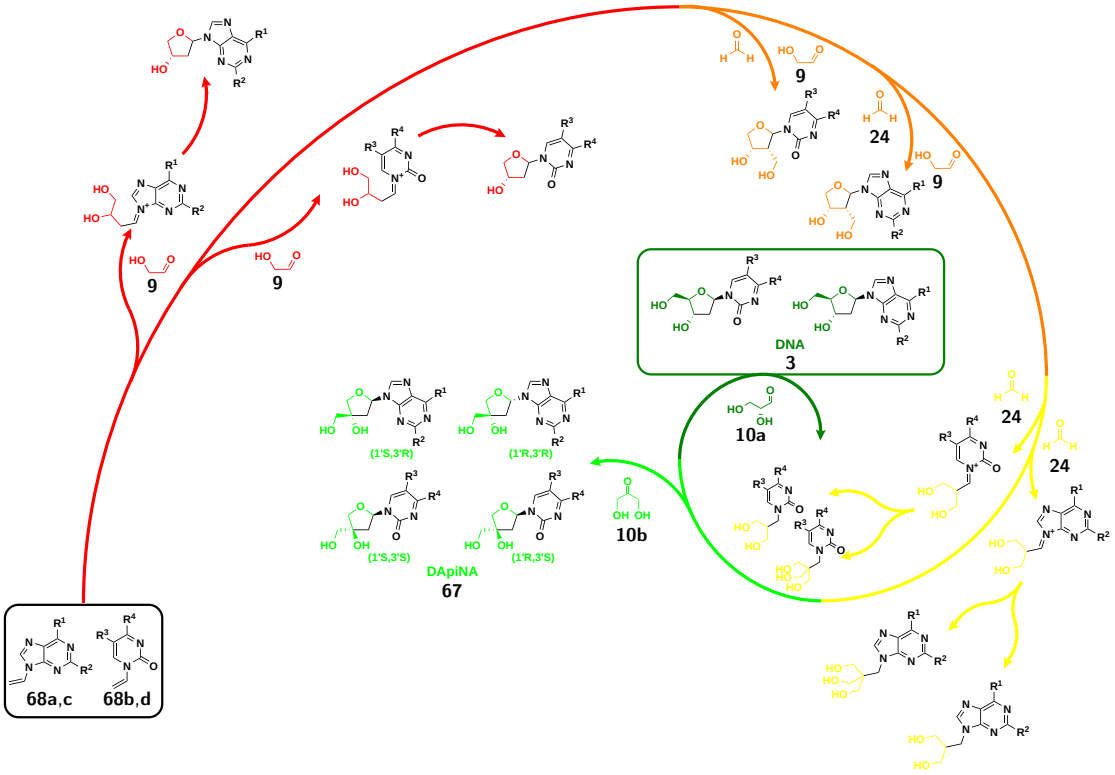


Figure 5.6: Possible and verified compounds derived from **9**, **10a**, **24** and **68**. $\boldsymbol{R^1, R^2}$= NH_2, H (**1a**); $\boldsymbol{R^1, R^2}$= OH, NH_2 (**1c**); $\boldsymbol{R^3, R^4}$= Me, OH (**1e**); $\boldsymbol{R^1, R^2}$= H, NH_2 (**1d**);

Testing the proposed reaction for a general application, other precursors than **35** and **10a** were implemented. The used compounds are starting materials or products of the formose reaction and are all aldehydes. Particular focus was set on not exceeding the maximum number of three carbon atoms in these species. The precursors used were glycolaldehyde **9**, dihydroxyacetone **10b** and formaldehyde **24**. Analysing the reaction mixtures *via* high-resolution OrbiTrap-MS, a plethora of different sugar-nucleobase adducts were found, see Figure 5.6. Examining the resulting products, the idea came to mind that this reaction network spanned by the four simple compounds **1**, **9**, **10a** and **24** could resemble the molecular evolution of deoxyribonulceosides.

In the following, the hypothesis for the selection of **3** out of such a mixture is described. During this evolution, the class of compounds solely derived from **9** were

ruled out because of their inability to polymerise (red). When implementing **9** and **24**, the resulting nucleosides (orange) are able to polymerise and have a more confined structure. Nevertheless they were abolished as well, as they would not lead to a well-defined helical structure. Exclusively starting from **24** gives a class of molecules derived from pentaerythrite (yellow). However, their chemical flexibility, in terms of the molecular degree of freedom and their stereo- and regioflexibility, is too liberate to participate in a confined coding polymer. As a consequence, these were ruled out as well.

When changing the sugar from **10a** to **10b**, an apiose-derived nucleoside (DApinNA) is gained (light green). Leaving just four epimers (*cf.* Figure 5.7), the nucleoside is already confined enough to might have served as a possible DNA progenitor. Yet its stereochmical freedom was not confined enough to be present in modern coding polymers. **3** (green) on the favoured side is just existing in two enantiomers. The properties of maximal stereochemical constraint, its ability to polymerise and to build elegant, stable structures, finally led to its assertation in the DNA-polymer.

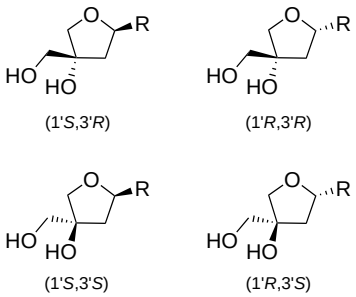
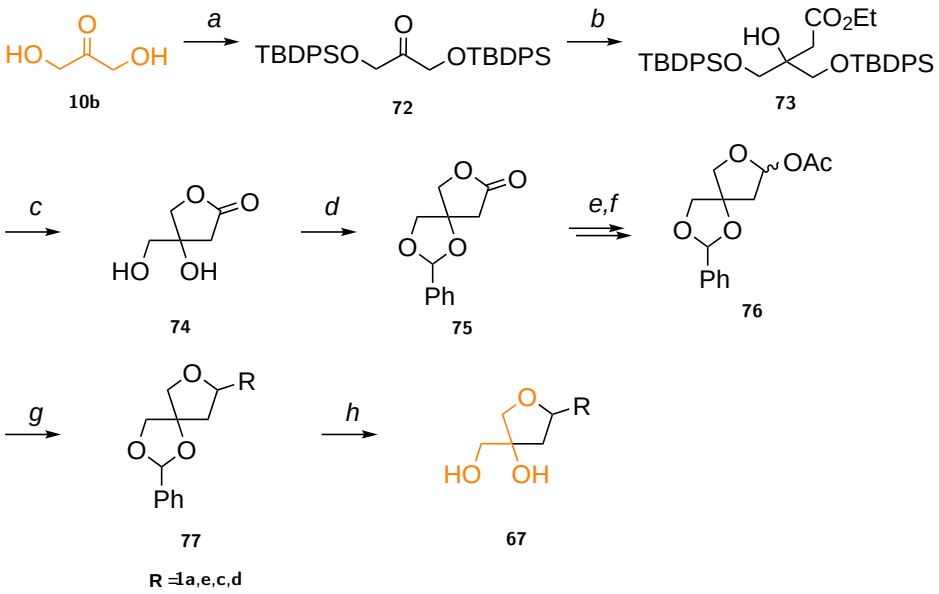


Figure 5.7: The stereoisomers of **67**

The prebiotic synthesis of DApiNA-nucleosides **67** is executed in the same way as for DNA, however, the C_3 building block **10a** is changed to **10b**. The reaction was analysed by UPLC-QTOF-MS and the retention times of the found species varied significantly from those of **3**. The motivation to assemble authentic samples of **67**-nucleosides came from two directions. First, the unambiguous assignment of the found species and second, the production of nucleoside analogues which might inherit bioactive potential.

The synthesis reclines to a previously published synthesis of fluoro-apionucleosides, see Section 2.7.[146] The proposed route is an eight step synthesis with overall good conversion rates, see Scheme 5.6. The procedure starts from **10b** in resemblance of its prebiotic counterpart. In the first step, **10b** is protected, forming the silylether-derivative **72**.



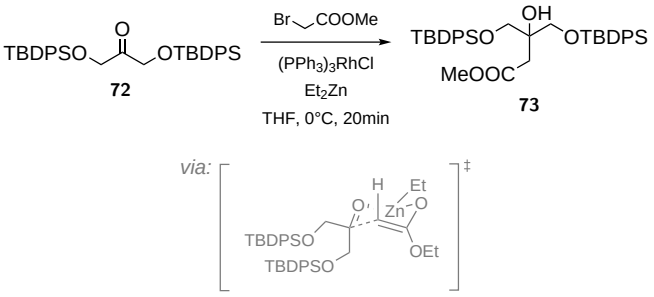
Scheme 5.6: Synthesis of **67**. Conditions: a) TBDPSCl, pyridine; b) $BrCH_2CO_2Et$, $(PPh_3)_3RhCl$, $ZnEt_2$, THF, 0 °C to r.t; c) HF·pyridine, 0 °C to r.t, then TMSOMe; d) PhCHO, $ZnCl_2$, Na_2SO_4, 60 °C; e) DIBAL-H, DCM, −78 °C; f) Ac_2O, DMAP, pyridine, 0 °C to r.t; g) HMDS, TMSCl, **1a**,**c**–**e**, $SnCl_4$, 50 °C; h) NH_3, MeOH, r.t. then HCl in MeOH, r.t.;

In a second step, two more carbon atoms are introduced to furnish the protected skeletal structure of apiose. This is achieved by a Reformatzky-type reaction. The conventional Reformatzky reaction,[242] implementing elementary zinc for the halogene-insertion, was not feasible in our case and did not lead to the desired product. In this case, the elementary zinc might be too inert. Activating agents like iodine or dibromoethane did not help either.

Increasing the reactivity of the zinc, by an in-situ formation of zinc nanoparticles from $ZnCl_2$ in a birch reduction (sodium and naphtalene) was considered next. Also in this case it did not lead to a conversion to the desired product. Possibly the remaining dissolved sodium prevents the reaction.

Careful investigation of the literature brought up a procedure,[243] implementing

diethylzinc and Wilkinson's catalyst. This procedure implies fantastic conversion rates to the desired product **73** and can be executed quite efficiently. On top, the product isolation turned out to be quite facile. The proposed mechanism of this reaction inherits a double insertion, see Scheme 5.7. First, the halogene-rhodium exchange is followed by the rhodium-zinc insertion, creating the zinc-enolate as the active species to attack the carbonyl group of the protected **10b**.



Scheme 5.7: Performed Reformatzky-type reaction and the depicted reactive intermediate.[243]

The subsequent deprotection of the silyl ether with a fluorine species and simultaneous acid-catalysed *5-exo-trig* cyclisation furnishes the apiose derived lactone **74**. The deprotection proceeded quite well for both cases utilising *tert*-butyl ammonium-fluoride (TBAF) and Olah's reagent (HF·pyridine). On the contrary, the purification of the TBAF-procedure turned out to be not feasible, as the excess amount of ammonium salts hindered the extraction of the water soluble product **74**. Thus, the HF·pyridine complex was used. The attention is again on the workup as only quenching with TMS-OMe is feasible, leaving no interfering excess amounts of salts. The resulting trimethylsilylfluoride and pyridine can be removed in the high vacuum, leading to a very mild and gentle purfication procedure.

Protection of the remaining free hydroxyl groups is needed to solely enable the reactions at the anomeric carbon. Careful choice of the protection groups needs to be made in order to guarantee a good preservation over all further steps. On the contrary, the deprotection should be accomplished under mild conditions in order to not disassemble the gained nucleoside in the last step.

Several household protection groups were tested. The formed utilised TBDPS-

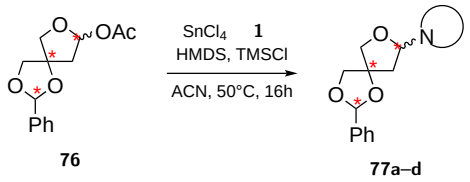
silyl ether is ruled out because of its size. Here, it results in the sole protection of the primary hydroxyl group. Implementing less-hindered silylethers did succeed, however, the purification turned out to be quite challenging. The gained product was not UV active nor could be stained by common and elaborate staining agents.

The choice fell on benzaldehyd, forming the benzylidene-acetal. Being aware of the fact that it might cause some complications during the course of the sequence. However, its advantages are the mild application and deprotection conditions and the conversion can be monitored throughout the sequence by UV activity.

Implementing benzaldehyde, the spirolacton **75** is formed. No attention was turned to the stereoselective control in the following steps. The reaction was performed in benzaldehyde, zinc(II)chlorid was used as acidic catalyst and the resulting water was trapped by sodium sulfate. Optimisation of the reaction conditions with glycol as a model compound showed that a ratio of 3:2 between $ZnCl_2$ and Na_2SO_4 gave the best yields. A further increase of the yield was achieved by raising the reaction temperature from room temperature to 50 °C. Substituting the desiccating Na_2SO_4 by molar sieve (4Å) did not lead to an equivalent conversion, nor did the reaction in a Dean-Stark apparatus implementing toluene at 120 °C.

In the following, **75** was reduced with DIBAL-H, leading to **78**. This step turned out to be the bottle-neck of the reaction sequence as the yield decreased significantly. It is assumed that overreduction at the lactone occurs, accompanied by the reduction of the benzylidenacetal, leading to an overall decrease in yield. Reducing the amount of DIBAL-H from 2 eq. to 1.25 eq. and cooling of both the reaction mixture and the reagent to –78 °C did result in a increase of the yield from 30% to 48%. Further optimisation of the reaction conditions or the choice of a different protection group for the diol would be beneficial.

The acetalisation performed well, when applied after purification of the compound. Although reported for some cases, the one-pot reduction and acetalisation did not give the product **76**. After 6 steps, the nucleosidation precursor **76** was gained as a diasteromeric mixture. The three chiral centers were not controlled throughout the sequence.



76 **77a–d**

Scheme 5.8: Performed silyl-Hilbert-Johnson nucleosidation with **76** as starting material.

In past transformations, the silyl-Hilbert-Johnson-nucleosidation performed quite well, see Scheme 5.8. However, in this case, it turned out to be another critical step as the used Lewis acid might be to harsh for this particular sugar precursor, resulting in low yields. Consequently, in case of **67b** debenzylidation was performed after nucleosidation at the stage of **77b**, leading to a one-pot nucleosidation /deprotection step. This turned out to be quite elegant, however on the expense of yield. For the remaining three nucleosides **67a**,**c**,**d**, additional deprotection is needed to cleave the benzylidene acetal. The purine nucleosidation is just feasible when 6-chloro purine as a precursor for **67a** and *N*-acetyl guanine as the precursor for **67c** are used. The use of the latter increases the solubility in the employed solvent.

In both cases, stirring in methanolic ammonia is needed to gain the correct nucleosides through either aromatic substitution or deprotection.

The deprotection of the benzylidene-acetal was performed by either dilute hydrochloric acid in methanol or by dilute acetic acid in methanol. It is crucial to consider the reaction times and the concentration of the acid to not cleave the gained nucleoside. Deprotection in methanolic hydrochloric acid (1 M) performs fast, but with a high destruction of the desired nucleoside. Deprotection in dilute acetic acid in methanol proceeds very slowly (around two days), thus, the conservation of the nucleosides is higher. Further optimisation of the acid catalyst is needed to loose no product to hydrolysation.

After deprotection, all nucleosides were purified *via* preparative HPLC to give two species. Theoretically, four species are possible, however, a chiral separation to discriminate the enantiomeric pairs was not conducted. The diasteromeric pairs are referred to as *cis*-**67** and *trans*-**67**.

67d was obtained in its diastereomeric pair and was fully characterised by NMR. The isolation of the *trans*-isomer of **67b** was achieved. Isolation of the *cis*-isomer was accomplished, but not in an amount sufficient for NMR-analysis. In case of **67a** and

67c, a full characterisation was not possible, as little amount of both species made NMR analysis impossible. Scale-up of the reaction is needed to obtain a significant amounts suitable for NMR analysis. The prebiotic mixtures were compared with the corresponding crude sample *prior* to preparative HPLC.

To ensure a definite verification of the nucleosides **67** in the prebiotic mixtures, these need to be spiked with the synthesised samples. As a coevolution of **3** and **67** is aimed to be demonstrated, the same reaction conditions are applied.

The prebiotic formation of the **67**-species turned out to not be as robust as indicated by previous results. However, the comparison of the retention times of the reference samples with the prebiotic reaction mixture indicates the formation of the correct structure, see Figure 5.8. Depicted on the left and in the middle are the EIC of **67d** and **67c**. This data was already published earlier.[234] In this context, the data was assigned misleadingly. Several signals of strong ionisation were identified at late retention times. These were erroneously perceived as the corresponding **67**-signals. The correct species eluted much quicker and had a much weaker ionisation. The correct assignment is depicted in Figure 5.8.

When injecting **3d** to a prebiotic sample supposedly containing **67d** (left), a distinct peak of a different retention time adds to the chromatogram. This indicates the formation of different species, when starting from **1d**, **35** and **10b**. The same applies for **67c** and **67a**. The ionisation properties of **67b** are even weaker as observed for **3b**, thus a verification and depiction is missing.

Comparison of the retention times and analysis of the distinct peaks for in-source fragmentation provides a strong indication for the correctness of the data. In analogy to the deoxynucleosides, **67b** poses a weak behaviour of ionisation. It was thus not detectable in the prebiotic mixture and is not listed in the following table. Further analysis is needed to verify its presence unambiguously. In Table 5.2, the retention times of the found **67**-species in the prebiotic mixtures are compared to those of the synthesised samples. Coinjections are indispensable as matrix effects of the prebiotic mixture cause different retention times as for the pure samples. Regardless, the samples are compared to cast a prediction on the correctness of the formed species.

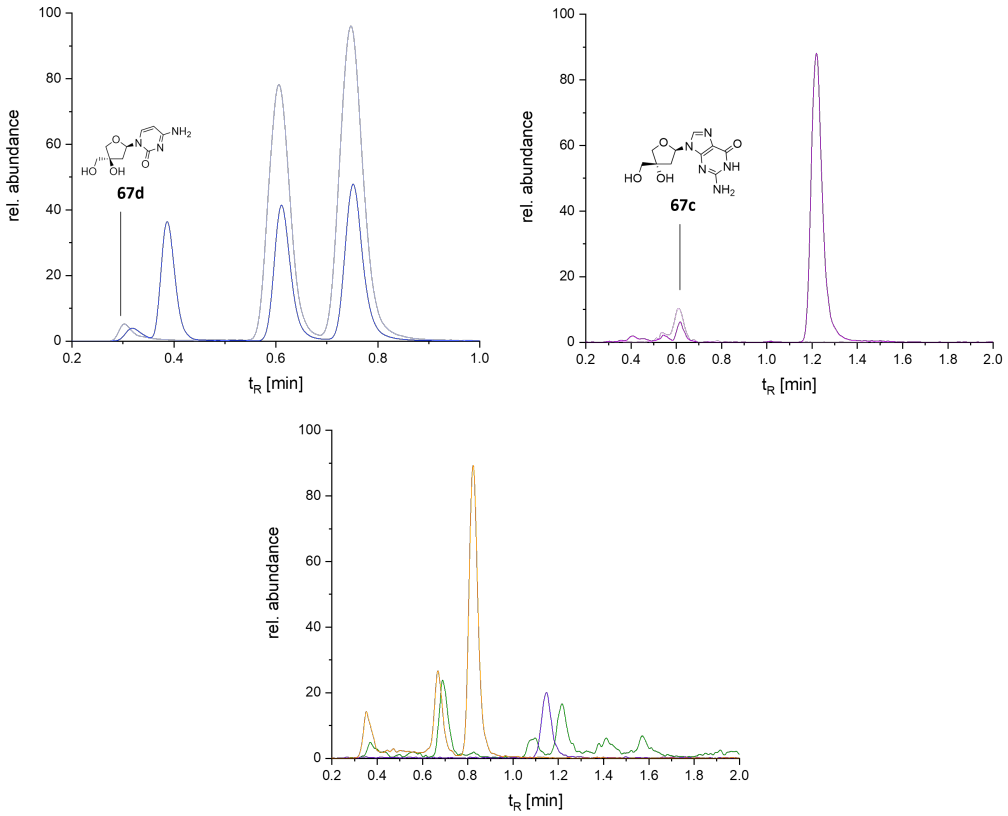


Figure 5.8: Performed prebiotic reactions leading to **67**. Depicted are the EIC of **67d** (top left) and **67c** (top right).[234] (0.1 mmol **1**, 1.00 mmol **35**, 0.15 mmol **10b** in 1 mL1 H_2O for 4 days at 50 °C) The prebiotic reaction mixture is superimposed in a light colour. Depiction (bottom) of **67a** (green), **67c** (violet) and **67d** (orange). The formation of **67b** is not depicted due to detection issues. (mass-filter m/z 228.095–228.099, 252.107–252.110, 268.102–268.106)

In case of **67a** two distinct peaks were found in the region in which nucleosides of **1a** are typically located. Analysis of the mass spectrum of both peaks shows the characteristic molecular peak of **1a** correlating with the molecular peak of supposed **67a**. The spectra are obtained in all-ions mode, during which the analysed sample is constantly, unselectively fragmented at 10 V and 20 V. In case of unambiguity, at least a statement for the assignment can be proposed. In addition, analysis of the crude **67a**-sample *prior* to preparative HPLC suggest that the signal at $t_R = 0.69$ min, found in the prebiotic mixture might correlate with the synthesised species. Unfortunately, purification of the synthesised **67a** was not achieved in quantities, necessary for full

characterisation. For a explicit confirmation of the correct species, co-injections with the reference samples are needed.

Table 5.2: Comparison of the retention times in the prebiotic mixtures of the found **67a**,**c**,**d** with the corresponding reference substances (mass filter *m/z* 228.095–228.099, 252.107–252.110, 268.102–268.106)

Species	$t_{Rpreb.}$ **[min]**	$t_{Rref.}$ **[min]**
67a	0.69	0.52/0.38
trans-**67b**	-	3.37
67c	1.15	-
trans-**67d**	0.35	0.34
cis-**67d**	0.67	0.49

One distinct peak was found in the experiments aiming for the verification of **67c**. The peak is located in the typical region of **1c**-deoxy nucleosides. Further, its mass spectrum contains the characteristic molecular peak of **1c**. The corresponding peak of the reference sample has a similar retention time. In consequence, the found peak points to a formation of **67c**, however a distinct verification by coinjection is needed. The coinjection would also answer the question if solely one isomer of **67c** is formed *via* this reaction setup.

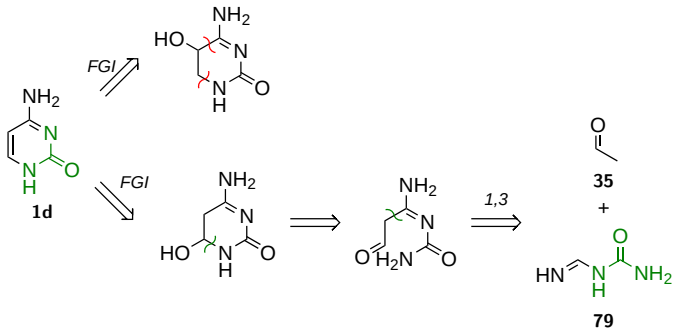
67d is the only species, which is fully characterised by NMR and retention times. Additionally, the found signals in the prebiotic mixture closely relate to the synthetic samples. The first species (t_R = 0.34 min) is in good accordance with the pure sample of *cis*-**67d** and the second signal at t_R = 0.67 min might correlate with the pure sample of *trans*-**67d**. Merely the confirmation by coinjection is still pending.

An efficient route towards **67** was delineated and executed. Comparison of the gained nucleosides with the prebiotic reaction mixture implies a strong evidence for the formation of **67** under prebiotic conditions. This outcome fortifies the existence of **67** as a possible progenitor to **3**. Unambiguous verification of the found species by co-injection is indispensable, to provide clarity about the remaining problems.

5.4 Prebiotic Pyrimidine Synthesis

In various recent studies, the synthetic access to purine nucleobases **1a**,**c** from simple prebiotic molecules could be shown. The oligomerisation of **5** simply gives the desired compounds. However, reaching the pyrimidine nucleobases **1b**,**d**,**e**, turns out to be far more complex, see Scheme 2.4. In general, the syntheses suffer from constructed reaction cascades, which are difficult to control. This work aims to introduce a novel synthesis for **1d**, which is compliant with the applied nucleoside synthesis.

Advantageous for such a task is to use a cut-set of starting materials. Used starting materials in the syntheses of nucleosides are available candidates *e.g.* **9**, **35**, **10a**. Performing a retro-synthetic analysis shows that both sugars cannot function as precursors *e.g.* cytosine build up. An eye-catching motive is the urea moiety, already functionalised twice in the formation of pyrimidine nucleobases. As **20** is easily accessible and reasonably stable, the analysis was built around this core function. Multiple possibilities of scissoring **1d** are conceivable, see Scheme 5.9.

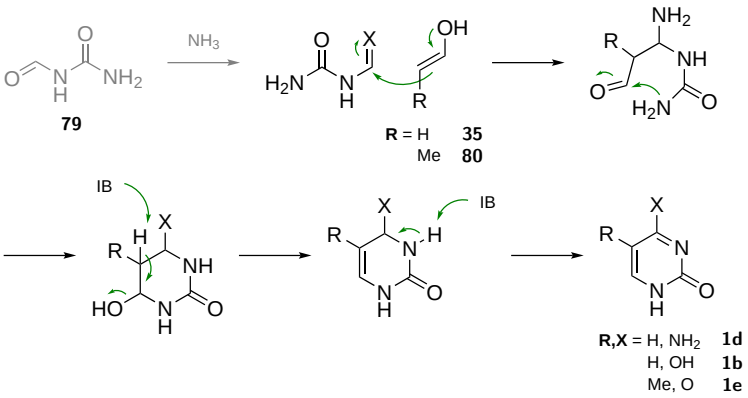


Scheme 5.9: Retrosynthetic analysis of **1d**

The prominent C-C double bond might originate from the elimination of water. The functional group interconversion reveals a possible nucleophilic attack at a carbonyl group, followed by subsequent elimination. The carbonyl can either contribute the 6- or the 5-position of the heterocycle. A carbonyl group at C5 would implicate an incompatible combination of nucleophilic and electrophilic centres, so it was placed at C6. A considerate 1,3-relationship is revealed, derived from the nucleophilic attack of the **35**-enolate, at a carbonyl-functionality. The second molecule turns out to be formylurea in its imine form. A perfect cut-set is achieved, introducing the new

molecule formylurea **79** and relying on the most reactive member **35** of the nucleoside reaction series. **79** is easily accessible via the formylation of urea by formic acid.

In the forward direction, the mechanism is proposed to perform in this manner, see Scheme 5.10. This delineated pathway seems to open up multiple possibilities to access **1d** by choosing the conditions in favour of a formimidamide, to **1b** by directly reacting the formamide. Even the construction of **1e** seems possible by variation of the aldehyde.



Scheme 5.10: Proposed mechanism of the pyrimidine formation.

Resulting from the mechanism and in correspondence with a continuous approach, the reaction conditions need to be alkaline, to favour the enolate formation. Further, the formimidamide needs to be introduced. The reaction was performed in water and several different alkaline additives were tested toward their ability to form the desired **1d**. The applied conditions are depicted in Table 5.3.

Confirmingly, just two of the tested conditions responded positively, both involving ammonia. These are killing two birds with one stone by the formation of the formimidamide and adequate basicity, enabling deprotonation of **35**, the E1cb elimination and the final deprotonation establishing aromaticity. Several experiments were conducted in the following, scrambling ammonia with the further depicted additives. Solely ammonia or ammonia/ guanidine **28** performed best under this task. Hence, the synthesis was executed by diluting **79** and **35** in aqueous ammonia and stirring the reaction mixture at 50 °C for 5 days.

Coinjection of the formed species with commercial **1d** confirmed the formation of the pyrimidine nucleobase, see Figure 5.9.

Table 5.3: Tested additives in the reaction towards **1d**. UPLC-QTOF-MS was performed with the following applied mass filters (*m/z* 112.050–112.051)

additive	***m/z* ([M+H]$^+$) of 1d**
-	-
K_2CO_3	-
NaH_2PO_4	-
Na_2HPO_4	-
NH_3	found
NH_3/**28**	found
NaOH	-
NaOAc	-
$(NH_4)PO_4$	-

Aiming for the most benign conditions, the conversion was examined in different concentrations of ammonia. Tracking the relative yield by UPLC-QTOF-MS, the concentration of ammonia was examined in a range from 0.01–5 M. The analysis indicated a less efficient conversion for the lower ammonia concentration and no conversion for the strong concentration. This indicates a tendency to milder reaction conditions, which is in line with the so far postulated conditions. In this context, it should be noted, that the pH value of the pyrimidine synthesis is significantly above the pH of nucleoside synthesis. Yet, the implementation of **28**, as an additional strong base catalyst, was tested if it enhances the conversion further.

28 is just available as its hydrochloride salt, as it is unstable in its free form. The pH of the ammonia solution (1 M) should suffice enough to partly deprotonate **28**.[244] The introduction of **28** did not enhance the conversion rate considerably, see Figure 5.9.

Performing untargeted analysis of the reaction mixture revealed several other interesting heterocyclic compounds, which might be of relevance in future investigations. The analysis was performed as a preliminary study, as advanced computational and statistical methods would be in demand to enable a precise statement. This would extend the scope of this thesis. The complex matrices were subsumed by their molecular features and the resulting molecular formulas were briefly checked for corresponding

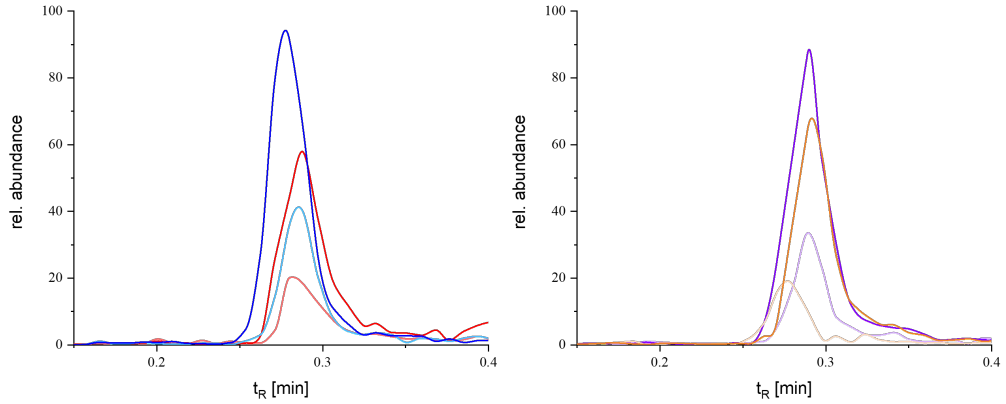


Figure 5.9: Coinjection verifying the formation of **1d** under prebiotic conditions from **79** and **35**. The reaction mixtures are depicted faded. left: 0.5 mmol **79**, 2.50 mmol **35** in *aq.*NH_3 1.00 mL (blue) 1 M, (red) 5 M; right: 0.5 mmol **79**, 2.50 mmol **35**, 0.65 mmol **28** in *aq.*NH_3 1.00 mL (violet) 1 M, (orange) 5 M.

heterocyclic compounds. Several interesting compounds were identified. A precise statement about their structure calls for intricate MS/MS analysis.

Covering the formation of all three pyrimidine nucleobases, the reaction mixture was analysed, regarding the presence of **1b**. Theoretically, **1b** is simply accessible as it is **1d**'s more stable hydrolysis product. In this case, no definite statement could be made about the presence of **1b** in the described reaction. Another pathway was pursued, by omitting ammonia. In theory, see Scheme 5.10, **1b** should be directly accessible *via* this pathway, not forming the formimidamide. The executed reactions did not substantiate this theory.

A long-discussed question is the emergence of **1e**. Was it present in an early set of organic molecules or did it evolve later *via* an enzyme-equivalent or an enzyme? Having a synthesis for **1d** at hand, simple variation of the aldehyde should furnish **1e** without the need for a methylating enzyme, like *thymidylate synthase*. Supposing the correctness of the proposed reaction mechanism, a simple change to propionaldehyde **80** would place the desired methyl group into the correct position. Subsequent expectable hydrolysis would then lead to the corresponding DNA nucleobase **1e**. In another scenario, **1e** could be reached by reacting **79** and **80** without the use of ammonia in an alkaline aqueous solution. The experiments were conducted in the same manner.

Both scenarios including and lacking ammonia did not give the expected results. The ionisation properties do not stand in the way of a sufficient detection, as **1e** is detected very easily. Arguments against the poor enoliseability of **80** are countless examples of proline-catalysed Michael-additions of **80**.[245] A possible explanation might be that the desired reaction is prevented by side reactions consuming the aldehyde. This theory is supported by the fact that a plethora of aldol condensation products in combination with **28** was found. Mono- and dialkylated species, derived from aldol-dimers of the corresponding aldehydes **35** and **80**, of **28** were found. Omitting **28** did not lead to the expected results, aldolcondensation still was a major issue, yet no **28**-adducts are involved. Opening the reaction vial, released a strong scent of over-ripe apples. This observation also points to the formation of aldol condensation products that resemble these fragrant compounds. Assuming the correctness of the presented pathway, the fact that **1e** is not formed, might be a hint for a later formation of **1e** during the evolution of DNA.

A holistic synthesis of **3** from simple starting materials is aimed. Therefore, it is necessary to link the presented synthesis of **1d** to the **3**-synthesis. As a consequence, it would be desirable to directly gain access to **68e** from a mixture of formylurea **79**, acetaldehyde **35** and guanidine **28**. The subsequent or initial addition of **10a** would then open up a direct pathway to pyrimidine nucleosides. Both scenarios were tested, however with no success. In none of the successful syntheses towards **1d**, the corresponding **68e** could be detected. Logically, directly starting from **79**, **35** and **10a** did not succeed either. A possible explanation might be that **1d** itself is formed in too small quantities, therefore the reaction to give **68e** is prevented in terms of concentration or prevens analysis as it is below the limit of detection.

A unified synthetic procedure to the pyrimidine nucleobases **1b**,**d**,**e** was pursued. The synthetic procedure should precede the synthesis of **3** by no major changes of the conditions and usage of a cut-set of starting materials. The key molecule to **3** synthesis was identified as the common building block. In reaction with **79** at 50 °C in aqueous ammonia solution, the formation of **1d** could be verified. The results provide a promising approach for pyrimidine nucleobase synthesis. By varying the reaction conditions or the corresponding aldehyde, possible access to **1b** and **1e** was delineated. Experimentally, however, these nucleobases could not be verified. A further effort

needs to be made to ensure a clear verification.

5.5 Complementary RNA Synthesis

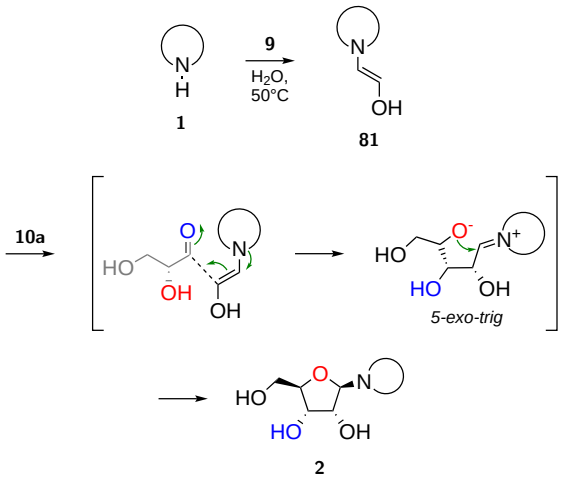
The above-described pathway to **3** is the first to perform in such high regio- and stereoselectivity. To transfer these properties to a concomitantly proceeding **2** pathway would imply a significant step ahead in explaining the emergence of our genetic code. By examining the mechanism closely, it came to mind that a minor change in the vinyl species, by adding a hydroxyl group, might open up a route to **2**-nucleosides. This would provide a unified pathway to all **2** under highly prebiotically plausible conditions. This approach is delineated theoretically in the following section and its experimental verification is presented. Parts of the investigation was executed in collaboration with J. S. Teichert.

5.5.1 Theoretical Considerations on 2

Retro analysis of **2**-nucleosides revealed glycolaldehyde **9** as the enamine forming C_2 synthon. In theory, the same proposed reaction mechanism leads to **2** when reacting **1** with **9** and subsequently **10a**, in an amino-catalysis like step. Compared to the **3**-synthesis, the corresponding transition state opens up a higher degree of freedom, in terms of regio- and stereo flexibility.

In complete analogy, the nucleobases **1** react with **9** to form the enamine **81**. The vinyl-species **81** then attacks **10a** from the Si-face in a stereoelectronically highly controlled fashion.[237] Concluding, the attack of the 4'-hydroxyl group at the intermediate iminium-ion closes the ring following a favoured *5-exo-trig* cyclisation,[238] furnishing the canonical ribonucleoside **2**, see Scheme 5.11.

The presence of the vinyl species **81** in E/Z isomers introduces an additional variance, leading either to ribo-**2** or arabino-**2**. Likewise in the mechanism leading to **3**, both the furanosides and the pyranosides, furnished by a *6-exo-trig*-cyclisation, are theoretically possible. Additionally, considering the D/L- isomers, the canonical nucleoside **2** needs to be selected out of six possible stereoisomers and four possible regioisomers, see Scheme 5.12.



Scheme 5.11: Proposed mechanism of the formation of **2**-nucleosides under prebiotic conditions (0.1 mmol **1**, 1.00 mmol **9**, 0.15 mmol **10a** in 1 mL H_2O for 4 days at 50 °C).

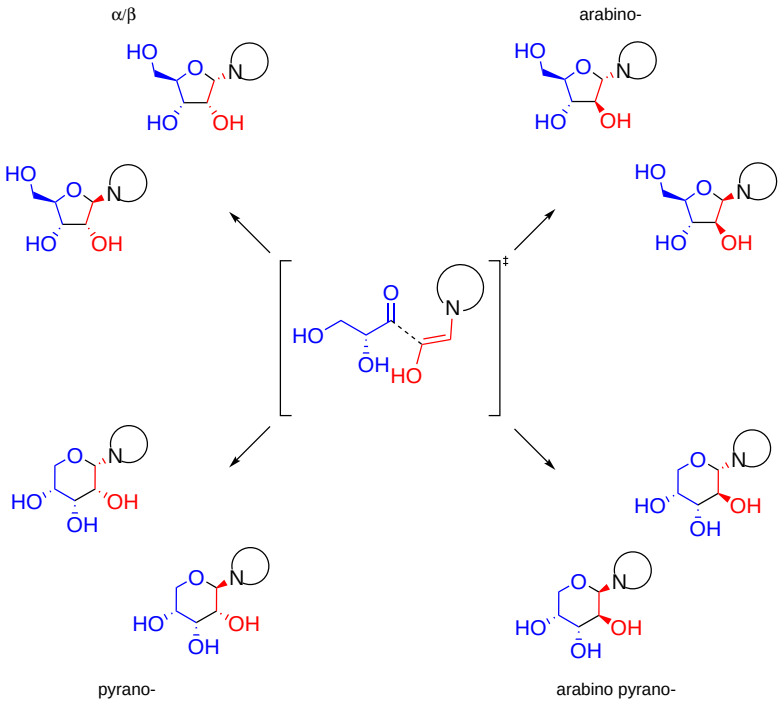
The reaction should again perform best in slightly alkaline media, to enhance the *N*1/*N*9 nucleophilic character of **1**. The choice of water or a mixture of methanol:water (99:1) should suffice as well. The temperature might influence the outcome of the reaction and thus places a possible selection tool, towards either **3** or **2** in the mixture.

5.5.2 Experimental Work

The same reaction conditions leading to **3** were applied. **1** and **9** were reacted in water or methanol:water (99:1) in the presence of sodium acetate. Further alkaline additives were screened evaluating their influence on the reaction. However, too alkaline conditions led to the formation of a detectable amount of aldol condensation products, derived from **9**. These aldol condensation products form adducts with the nucleobases as well. However, it was not examined whether **81** undergoes aldol condensation with **9** or **9** first condensates and then forms adducts with **1**.

Despite the observed side reactions, analysis *via* UPLC-QTOF-MS showed the initial formation of the nucleobase acetaldehyde **81**. To confirm the correctness of the data, synthetic reference samples were prepared.

The authentic reference compounds were synthesised according to the procedure described in the following, see the experimental section for further details. Nucleophilic attack of **1** at 2-bromo-1,1-diethoxyethane yielded ethoxy-acetal protected

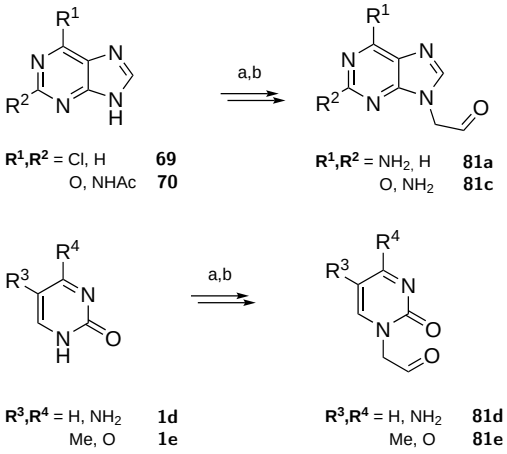


Scheme 5.12: Possible stereo- and regioisomers of **2**, the L-isomers are not depicted.

81. The reaction conditions differed quite substantially when using different **1**. **1a**,**d** reacted at 50 °C and also 2-bromo-1,1-dimethoxyethane could be used as electrophile. Whereas for **1e**,**c** higher temperatures were necessary to furnish the desired compound. In the case of **1e**, double alkylation at both endocyclic secondary amines was observed. Additionally, in the case of **1c**, further protection was needed to increase the solubility of the nucleobase. Subsequent deprotection of the methoxy- and ethoxyacetal by stirring in diluted hydrochloric acid in methanol at 60 °C, yielded the desired nucleobase acetaldehyde **81**, see Scheme 5.13. Other than expected, **81** were stable under ambient conditions.

Starting from a mixture of **1** and **9** led to the formation of **81** after a reaction time of five days.

Examining the additional degree of freedom present in **81**, it was anticipated that the coinjection with synthetic samples of **81** would not proceed as unambiguously as for the corresponding **68**. A distinct assignment could just be provided for **81c**, see Figure 5.10. In the prebiotic procedure towards **3**, the vinyl species **68** was found



Scheme 5.13: Synthesis of **81**. a) $BrCH_2CH(OEt)_2$, K_2CO_3, DMF, ; b) HCl, MeOH, reflux;

and verified as one possible isomer. In the case of **81d**, the transition to the hydroxy vinyl species led to the formation of at least two isomers. The species with the minor ionisation proved to be the correct species, however, the more intense signal showed the formation of an unidentifiable species. Likewise to the synthesis of **81**, these might be double-alkylated species. In the case of **81a** the coinjected reference peaks were close to the detected signals, thus indicating a related molecular structure, however not coeluting as one peak. The concept of prebiotic protection groups, earlier discussed in Section 5.2.2 might cause this slight change in retention times. The correct species might be present on the column as its *N*4/*N*6-**9** counterpart. Finally, during ESI-ionisation this unstable species is cut, detecting the correct *m/z* of its related **81**-species. The transition from **1b** to **81b** could not be confirmed. A distinct signal possessing the correct *m/z* could be identified, unfortunately, coinjection with the synthetic references showed a major discrepancy in the retention time.

Despite the inauspicious verification of **81**, the overall transition to **2** was studied. In analogy to **3** syntheses, a reaction mixture containing **1**, **9**, **10a** and sodium acetate, was stirred for 5 days at 50 °C. The verification of **2** was attempted by UPLC-QTOF-MS, using coinjections with reference samples. Further, the nucleosides were verified using UPLC-UV detection and high-resolution OrbiTrap-MS.

In contrast to the formation of **3**, the formation of **2** does not proceed comparatively easy. The selectivity of the reaction is much more complex, as multiple species can

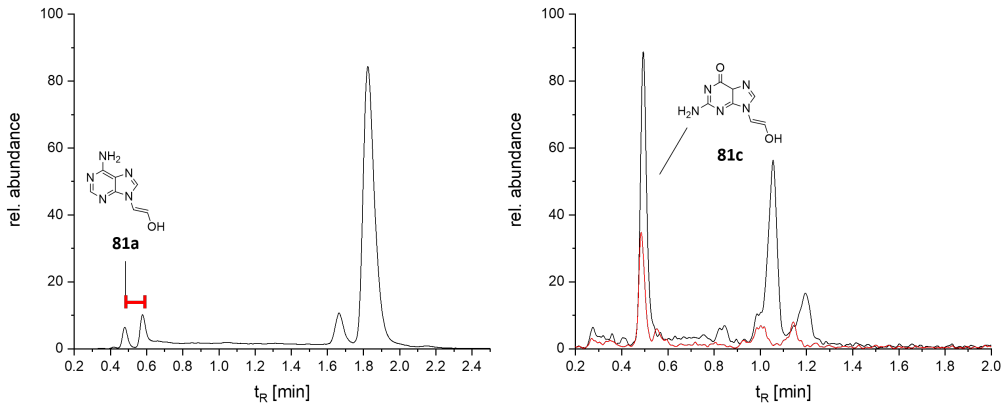
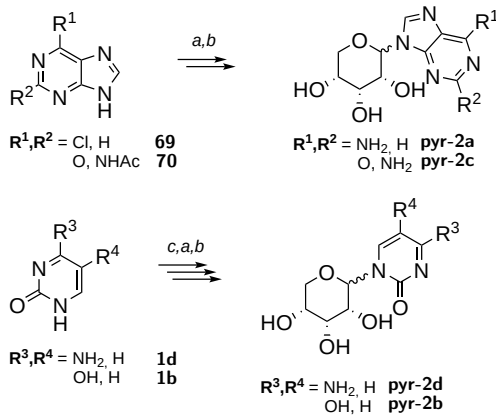


Figure 5.10: Coinjection of synthetic samples of **81a**,**c** confirmed the formation of **81c** under prebiotic conditions (right). The formation of **81a** could not be unambiguously confirmed as the synthetic reference showed a discrepancy in retention time (left) (0.1 mmol **1**, 0.2 mmol **9** in 1 mL H_2O for 5 days at 50 °C) (mass-filter *m/z* 178.070–178.074, 194.065–194.070).

be chromatographically observed. This is due to the described multiple possibilities in the transition state. A computational study would be beneficial to get a comprehension of the energy potential in the dependence of different configurations and variation of **1**. Several factors might influence the successful transition. The dimeric structure[246] of **9** and its general lower reactivity compared to **35** can prevent or slow down an adequate formation of the necessary **81**. As already outlined, **81** can be present as its E/Z isomers, which can influence the general reactivity. Unlike with **35**, with **9** the presence of various nucleophilic sites might be more relevant. The hypothesised mechanism leading to **2** is the same as for the synthetic procedure towards **3**, however, the introduction of a single hydroxy-group might alter the behaviour significantly. Further theoretical and experimental insights are beneficial, to verify all detected species.

As described above, for most cases the reaction yielded a non-selective mixture of different products, expressed by the presence of multiple signals in UPLC-QTOF-MS. In order to guarantee a definite verification of the formed **2** species, all possible isomers must be obtained. ***β*-fur-2** and **ara-2** were received commercially, whereas ***α*-fur-2**, ***α*-pyr-2** and ***β*-pyr-2** were obtained synthetically. The synthetic procedures leading to the latter are described in the following.

***α*,*β*-pyr-2** were synthesised from ribopyranose tetraacetate **82a** according to the Lewis-acid assisted silyl-Hilbert-Johnson procedure, see Scheme 5.14. **82a** was gained after the protection of **4a** with acetic anhydride, yielding a viscous oil of both the furanose- and the pyranose species. The pyranose species crystallised after several days at 4 °C and was gained after trituration with diisopropyl ether. ***α*-fur-2** was synthesised following the Lewis-acid assisted silyl-Hilbert-Johnson procedure from commercially available ribofuranose tetraacetate **82b**. To ensure a good conversion to the corresponding purine nucleosides, 6-chloro purine and *N*-acetyl-guanine needed to be implemented. Finally, the purification of all nucleoside-diastereomers after deprotection was performed by preparative HPLC.



Scheme 5.14: Synthesis of **pyr-2**. a) **82a**, HMDS, TMSCl, $SnCl_4$, acetronitrile, 50 °C; b) NH_3, MeOH, 60 °C ; c) HMDS, TMSCl, NH_4SO_4;

2a could be verified as its canonical isomer. Hence, the conceptual validity of the hypothesised mechanism, see Scheme 5.11, could be validated. Coinjection with all presented reference samples (***α*,*β*-pyr-2**, ***α*-2**) confirmed the formation of the canonical **2a**-nucleoside (t_R = 1.3 min), see Figure 5.11. A second major signal (t_R = 1.0 min) indicates the formation of **ara-2a**.

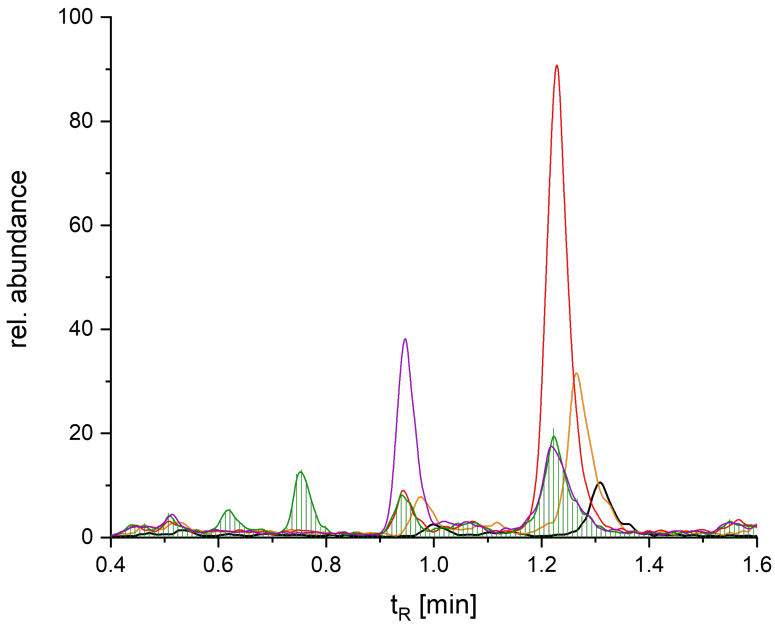


Figure 5.11: Coinjection of synthetic samples showed the sole formation of the canonical ***β*-fur-2a** under prebiotic conditions (0.1 mmol **1**, 0.2 mmol **9**, 0.15 mmol **10a** in 1 mL H_2O for 5 days at 50 °C). Coinjections to the reaction mixture (black) with ***α*-fur-2a** (orange), ***β*-fur-2a** (red), ***α*-pyr-2a** (blue), ***β*-pyr-2a** (green) and **ara-2a** (violet) enabled an unambiguous identification. (mass-filter *m/z* 268.102–268.106).

A second strong indication for the correctness of the proposed mechanism is provided by the exclusive formation of guanosine. It is formed as its β-anomer, however, a *6-exo-trig* cyclisation takes place. Although giving the non-canonical β-pyranose nucleoside ***β*-pyr-2c**, the reaction proceeds in remarkable stereo- and regioselectivity, see Figure 5.12. Additionally, a second product with the *m/z* of **2c** and its distinct fragmentation pattern could be identified. Yet the species could not be assigned to a characterised reference structure. Eventually, **ara-pyr-2c** is formed *via* this method and is expressed as the second signal. Synthesis of the corresponding reference structures is necessary to validate this hypothesis.

These results support the evidence that the presented pathways, relying on the scission of the 2'-3'-carbon bond is highly prebiotically plausible. **3** is obtained in exclusive stereoselectivity and high regioselectivity. Both **2** and **3** nucleosides can be formed from simple readily available starting materials and both are closely related by

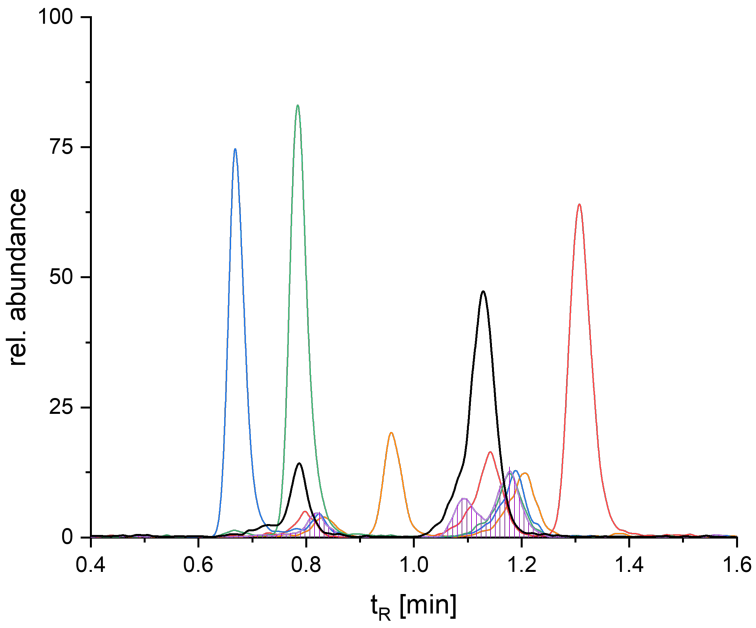


Figure 5.12: Coinjection of synthetic samples showed the sole formation of non-canonical ***β*-pyr-2c** under prebiotic conditions (0.1 mmol **1**, 0.2 mmol **9**, 0.15 mmol **10a** in 1 mL H_2O for 5 days at 50 °C). Coinjections to the reaction mixture (black) with ***α*-fur-2c** (orange), ***β*-fur-2c** (red), ***α*-pyr-2c** (blue), ***β*-pyr-2c** (green) and **ara-2c** (violet) enabled an unambiguous identification. (mass-filter *m/z* 284.097–284.100).

a minor change in the C_2-aldehyde. An evolutionary selection must have taken place at a later stage *e.g.* during the polymerisation of DNA.

Rethinking the assembly of nucleosides not only gave access to **3** but concomitantly to **2**. It is thereby not important to aim for a contribution of pure sugar starting materials as the major components of a formose reaction, **9** and **10a**, are both crucial to this pathway. The transition to both targets can be achieved from a mixture of all starting materials in one pot and is thus considered to be highly prebiotically plausible. Based on these results, the coevolution of RNA and DNA at the emergence of Life is advocated.

5.6 Cofactor Synthesis

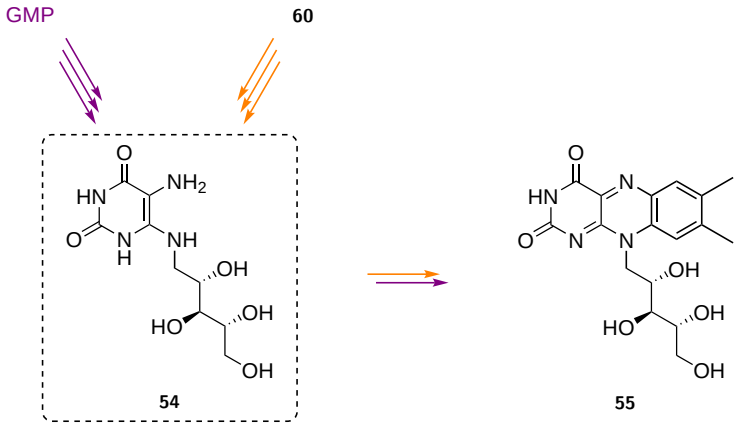
Cofactors play an eminent role in biosynthetic conversions. These are located in all three domains of Life. The fundamental functions argue for presence during the development of Life, otherwise, they would not be as inherently important. Despite their biological abundance, few prebiotic investigations were conducted to access cofactors. There are many interesting cofactors to examine, serving different tasks in biology, this work examines the prebiotic route towards riboflavin **55** more closely. **55** is one of the most potent and versatile redox cofactors and therefore quite appealing in the context of Life's formation.

Presuming the aspect that Nature rather optimises existing pathways than ever creating new ones, the delineated reaction cascade in this thesis to access **55** is closely related to its biosynthetic pathway, see Subsection 2.8.1.

A key intermediate in the biosynthetic procedure towards riboflavin is ribityl-diaminouracil **54**. Following logical thinking, it seemed feasible to intersect the biosynthesis at this point and find prebiotic routes towards **54**. Diaminouracil itself is thought to be easily available through heterocyclic reactions. The key challenge in the delineated pathway is thus the ribosylation at the heterocyclic unit and its subsequent reduction to the ribityl moiety. The formation of the pteridine ring and the lumazine dismutation to end up at **55** should proceed with not more than minor adjustments. In recent literature, the pteridine formation and its dismutation, applicable to prebiotic conditions, were studied intensively. However, the build-up of the ribityl side chain remains inexplicit.

Starting directly from **60** inherits regioselectivity-issues by two more or less equally reactive amino functionalities. To favour the sole reaction at the 6-amine, the 5-amine needs to be protected sufficiently, but reversibly. The protection group should be stable enough to withstand the surrounding conditions but should be easily cleavable to facilitate the build-up of the lumazine heterocycle.

Inspired by late works in the group of Eschenmoser[103,165] and by the investigation of the group of Carell[19,105] on **32**-compounds, 5-formamido-6-aminouracil **32a** was selected as a possible precursor. This heterocycle has already been studied, examin-



Scheme 5.15: Schematic outline of the presented synthesis of **55**.(Biosynthetic pathway (violet), prebiotic pathway (orange))

ing a possible prebiotic pathway to **55**, however, just the fusion with **4a** was studied. Further, it was hypothesised that an intramolecular reduction of the ribosyl-chain can be achieved *via* the formamido-hydrate.[165] Multiple conditions were tested, but **32a** did not react in the favoured way and the pathway was discarded. Applying our knowledge on the composition of sugars at heterocycles to this reaction series would reactivate **32a** as a possible starting material. The investigations selectively forming a ribosyl-chain at the 6-NH_2 are delineated in the following.

The reactions of **32a** with **9** and the reaction of **32a** with **9** and **10a** were tested in aqueous media, under the assistance of several alkaline additives. The outcome of the reaction was analysed by UPLC-QTOF-MS. Several alkaline additives were implemented, including NaOAc, K_2CO_3, NH_3, NaH_2PO_4, Na_2HPO_4, borax, phosphate buffer and guanidine **28**.

Although carefully tested under various conditions the conversion did not take place. Neither the **9**-adduct of **32a** could be found nor the **32a**-riboside was reachable. The most prominent species to be found in the reaction mixture is the corresponding xanthine **1g**, originating from the closure of the imidazole-ring.

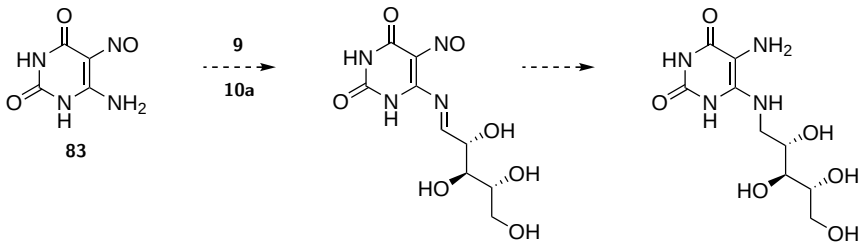
Accompanying, experiments were conducted examining the reactivity of **32a** with several sugar species. The screening confirmed its minor reactivity under the applied conditions. **32a** was the least reactive in this series, showing nearly no conversion with any of the investigated sugars after a long period of time. The low reactivity

might result from its insolubility in the solvent, as the well-established planarity of the molecule promotes $\pi - \pi$ stacking.

As a consequence, a purine ring-opening of **1g**, similar to the biosynthetic pathway was considered. This approach was motivated by the fact that ribosylation at **1** is reachable. After ribosylation, a purine ring opening might be achieved, implementing a titanium dioxid-mediated cut of the imidazole ring, reported in a prebiotic context.[247] The combination of the ring-opening with the sugar-forming reaction would open up another possibility to access the desired compound.

However, the reported ring-scission did not proceed in our hands, when starting from pure, available compounds. This applied for several tested nucleobases, as well as their ribonucleoside and deoxyribonucleoside counterparts. Trying to solve the problem, the presumption was made that incorporation of a Lewis-acids into the reaction mixture would destabilise the purine ring and promote its opening. However, screening of several Lewis-acids of different strength did not lead to a detection of the correct product. This also counted for the variation of the temperature. In consequence, a different starting molecule was needed, having higher reactivity towards the sugar forming reaction paired with the protection of the C6 amino group.

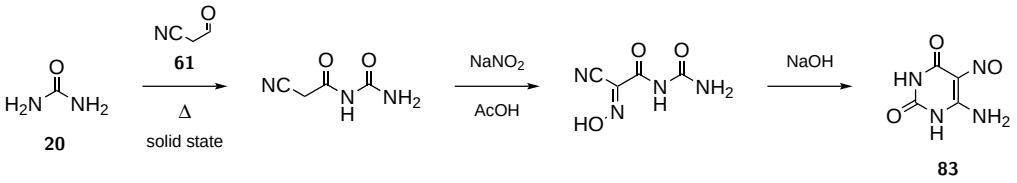
5-nitroso-6-aminouracil **83** was chosen as a new starting point. Nitrosopyrimidines were already considered in the context of prebiotic heterocyclic chemistry[104] and the formation of **2**.[19] The use of **83** would advantage a well-protected 5-position, enabling a more selective ribosylation. In the following reductional step, both the nitroso-moiety and the desired **4a**-side chain can be addressed by an orthogonal reduction in one step, see Scheme 5.16.



Scheme 5.16: Schematic depiction of the synthesis of **55** starting from **83**.

The synthesis of **83** can be easily achieved from available starting materials.[248,249]

Additionally, its translation to a prebiotic setting is facile. The conventional synthesis starts with a condensation of **20** with cyanoacetic acid **61**in acetic anhydride. As soon as the reaction temperature of 120 °C is reached, the product precipitates immediately. Rapid refrigeration of the mixture is needed to prevent overreaction. In a prebiotic context, the condensation could have been achieved by a urea-melt under desiccating conditions or *in vacuo*. Following this rationale, *N*-carbamoyl-2-cyanoacetamide **84** could be accessed both under standard and prebiotic conditions. All further steps are easily transferable to a prebiotic setting. **84** can be easily nitrosylated by using sodium nitrite under acidic conditions. The subsequent formation of the pyrimidine heterocycle is achieved by intramolecular ring closure under alkaline conditions, see Scheme 5.17. In the classic procedure, a solution of sodium hydroxide (5 M) in water is used. Under circumstances present on the early Earth, such harsh conditions might be disputable.



Scheme 5.17: Synthesis of 5-nitroso-6-aminouracil **83** under prebiotic conditions.

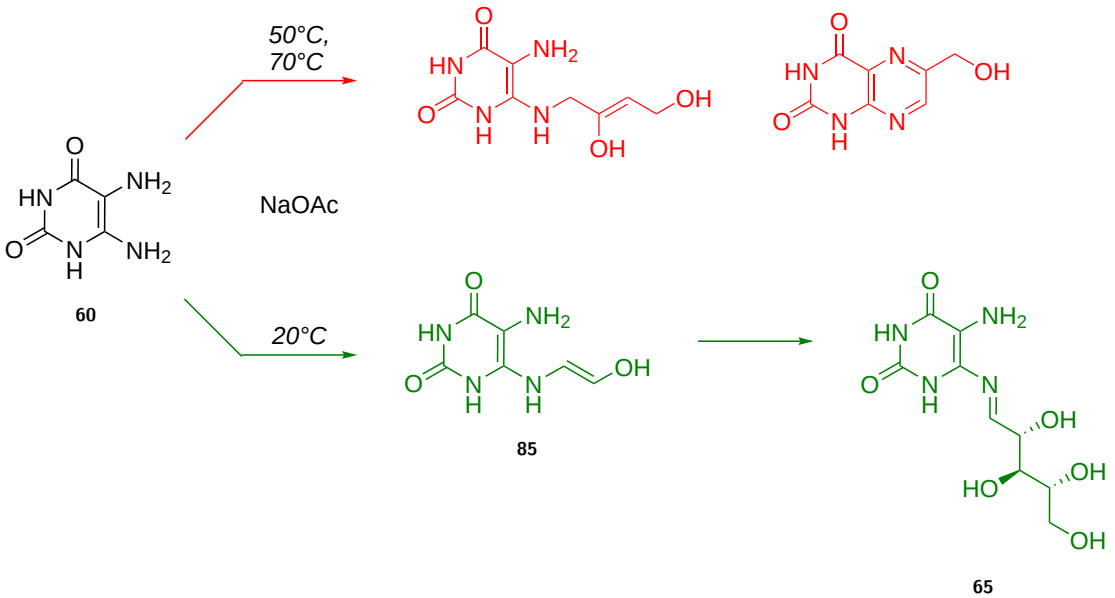
Contrary to the beauty of the described **83**, the reactions leading to a sugar-formation at the free-amino group did not succeed. To the best of our knowledge, not even the **9** derivatives could be found by analysis with UPLC-QTOF-MS and OrbiTrap-MS. As for the former starting material **32a**, several alkaline additives in different strengths were examined at different temperatures, ranging from 20–120 °C. Despite the variety of different conditions, it seemed that a major amount of starting material remained insoluble. Therefore, variation of the solvent to formamide, methanol and DMF and implementation of stronger alkaline additives was executed. Even under these applied circumstances **83** remained inert. The insolubility is a possible effect of its well-established planarity by $\pi-\pi$-stacking. Although **83** is quite favourable in terms of regioselective control, unfortunately, it seems to not react under any condition applied.

Although already described as unfavourable in the beginning **60** is considered next.

It is the least favourite of this series in terms of regiocontrol. Multiple possible reaction sites are present, complicating the regioselective ribosylation at the C6 amine. On the other hand, it is the most favourable compound in terms of simplicity. After the implementation of the ribosyl chain, no redundant make-up reaction is needed to transform or cleave artificially introduced protection groups.

60 was reacted in slightly altered conditions as applied to the synthesis of **2**, **3** and **67**. A higher equivalent ratio of alkaline additive was added to achieve slightly alkaline conditions. This is because **60** needs to be neutralised as it is commercially available as its stable sulfate-salt. All further boundary parameters are left the same.

Applying these conditions, the analysis by UPLC-QTOF-MS did not show the formation of the desired **9**-enamine of **60**, after a reaction time of 5 days. Interestingly, aldol-condensation products were detected in substantial quantities, see Scheme 5.18. These are the results of the aldol-condensation between diaminouracil acetaldehyde **85** and the aldol-condensation product of **9**. Further, the detected compound can undergo ring closure to form the corresponding pteridine ring.



Scheme 5.18: Temperature dependence of the proposed ribosylation of **60**

Not beneficial at first sight, it demonstrates that aldol-reactions under these circumstances are thinkable and possible monomers or aldoldimers react with **60**. In this context, it is not clear if **85** gets initially formed, attacking another **9** molecule or if **9**

first condenses and then reacts with **60**. In any case a cross-aldol-reaction of **85** with **10a** should be feasible.

It was tested if the conversion takes place at an elevated temperature of 70 °C. Further, the influence of formamide as solvent was tested. In both variations, UPLC-QTOF-MS experiments did not indicate the formation of **85** under the applied reaction conditions. However, still large quantities of aldol condensation products with **60** were detected. Possibly the high concentration of alkaline additives and the elevated temperature omit the sugar forming reaction at the heterocycle.

Applying milder conditions and decreasing the temperature to 20 °C did lead to the detection of **85** by UPLC-QTOF-MS in the aqueous mixture. In this case, the reaction in formamide did not show the conversion to **85**. The above-described side-products were still detectable in large quantities.

Based on the signals indicating the presence of **85**, it was tested if the established species would react with **10a** to lead to the corresponding 5,6-diaminouracil riboside **65**. Starting from **60**, **9** and **10a** did result in signals implying the formation of **65**. Remarkably, both species **85** and **65** were detectable after one day but increased in terms of relative abundance over a longer period of time.

In this case, the ribosylation performs under less mild conditions as a considerable amount of alkaline additive was added. Yet, the transformation proceeds in water at 20 °C. Analysis of the reaction was performed by UPLC-QTOF-MS, but this time the polarity of the gained compounds complicated the separation under C18-conditions as they eluted close to the dead time.

Several different methods were tested to achieve a separation of the described compounds and enable unambiguous identification. The performance of a reversed-phase-NH_2 stationary phase, of a HILIC-phase, an AdvanceBio Peptide-HILIC phase and a C30 phase were tested. Further, different mobile phases were tested, including gradients from water to methanol, water to acetonitrile, implementation of retaining salts like ammonium formate and ammonium acetate and variation of the pH from 5-9. However, none of the described stationary phases resulted in a satisfying separation. Finally, the application of an aqueous heptafluoro butyric acid solution (0.1 v%) as retaining ion-pairing reagent in combination with a C18 stationary phase, led to a reliable separation of the detected structures.

The above-described experiments were executed starting from **60** as its hemisulfate salt. A six-fold excess of the alkaline additive was applied, reaching a pH of 8–9. To examine the alteration of reactivity, caused by sulfate, pure **60** was prepared, by desalination in solution with $Ba(OH)_2$. In the following experiments the free base **60** was used.

Due to the basicity of the free **60** itself, the amount of alkaline additive was reduced to a two-fold excess in the corresponding experiments. The tested additives were NaOAc and Na_2HPO_4. The transition was pursued in both water and formamide. In contrast to its sulfate salt, the colour of the resulting solutions changed significantly. Using **60**-sulfate lead to a brownish solution, while free **60** in water has a bright orange colour and a bright yellow colour in formamide. The solubility in water is fairly poor, on the other hand, its solubility in formamide is higher, based on optical observations.

In consequence, the reaction behaviour varies substantially. In water, nearly no conversion to the glycolaldehyd adduct **85** was observed, whereas in formamide the transition to **85** was easily detectable. Subsequently, the transition to the ribosylated species **65** was just detected in the formamide experiments. Opposite to previously described experiments the chormatogram showed non-distinct peaks, indicating a possible mixture of isomers. Analysing the formed species by MS/MS experiments turned out to be ambiguous. In contrast to all previously fragmentation experiments with heterocycle-saccharides, the distinct heterocycle fragment was not detectable. Other fragments might evidence a formation of **85**, yet the absence of the otherwise main heterocyclic fragment questions the formation.

The ribosylations of **60**-sulfate and **60** were studied in water and formamide at temperatures ranging from 20–70 °C. Several concentrations of alkaline additives were tested. In some experiments, the transition to **65** seemed promising, yet further experiments need to be executed to give a definite answer about the key step of prebiotic **55** formation.

Aiming for completion of this reaction series, the reduction of the formed ribosyl-chain was studied as well. The proposed[165] intramolecular reduction *via* the ribosylated **32a**-hydrate was not taken into account as the ribosylation of **32a** itself was not feasible. Another possibility for reduction might be inorganic salts like sodium

thiosulfate or sodium dithionite. Another variant might be a Zn/formic acid system, producing hydrogen. The latter was already applied in recent prebiotic approaches to ribonucleosides.[19] These strong reducing agents are reachable under prebiotic conditions, yet it can be discussed about their availability and concentration. This thesis aims for a more elegant way, fitting to the *systems chemistry* approach.

Illuminating possible prebiotic networks, anteceding modern biochemistry, the class of Hantzsch esters was chosen. These are in close resemblance to NADH present in modern biochemistry, *cf.* Figure 5.13, and might have functioned as a mild heterocyclic reducing agent.

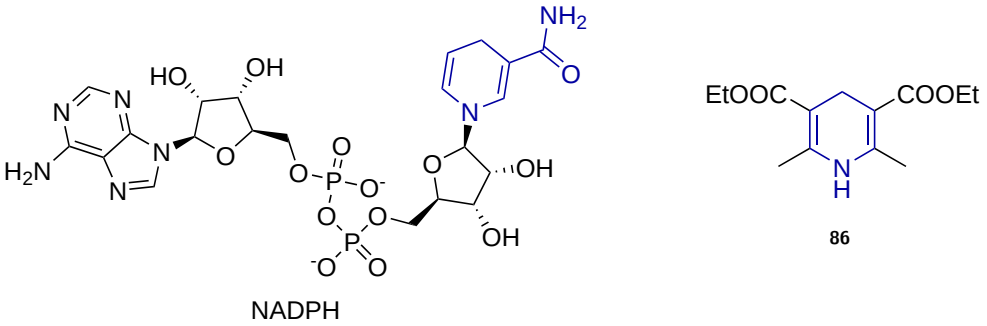


Figure 5.13: Comparison of the biosynthetic reducing agent NADH and its prebiotically discussable analogue **86**.

The simplest Hantzsch ester is composed of formaldehyde **24**, ammonia and ethyl acetoacetate **86**. All three precursors are easily discussable in a prebiotic context. Further, Hantzsch ester derivatives are reachable by the variation of the aldehyde or the ketone body.

The Hantzsch ester synthesised from **24**, ammonia and ethyl acetoacetate was chosen as prebiotic equivalent for the imine-reduction.

As **86** is reported to be a potent reductant,[250] its reduction potential was studied in the prebiotic context. **86** was added to experiments indicating the formation of **65** to study whether a reduction to a ribityl-chain would take place. Monitoring of **54** was performed *via* UPLC-QTOF-MS.

Two distinct methods were investigated. The addition of **86** at the beginning of the reaction and the addition of **86** after detection of the corresponding **65**. However, the desired **54** could not be verified in both cases. This might root back to two major issues. First, the solubility of **86** in water turned out to be less sufficient. Even in

formamide, the solubility was limited. Second, the found **65** might not be the correct species, hence a reduction to **54** is not performable. Intensive investigation of both ribosylation steps as well as the reduction is still in demand.

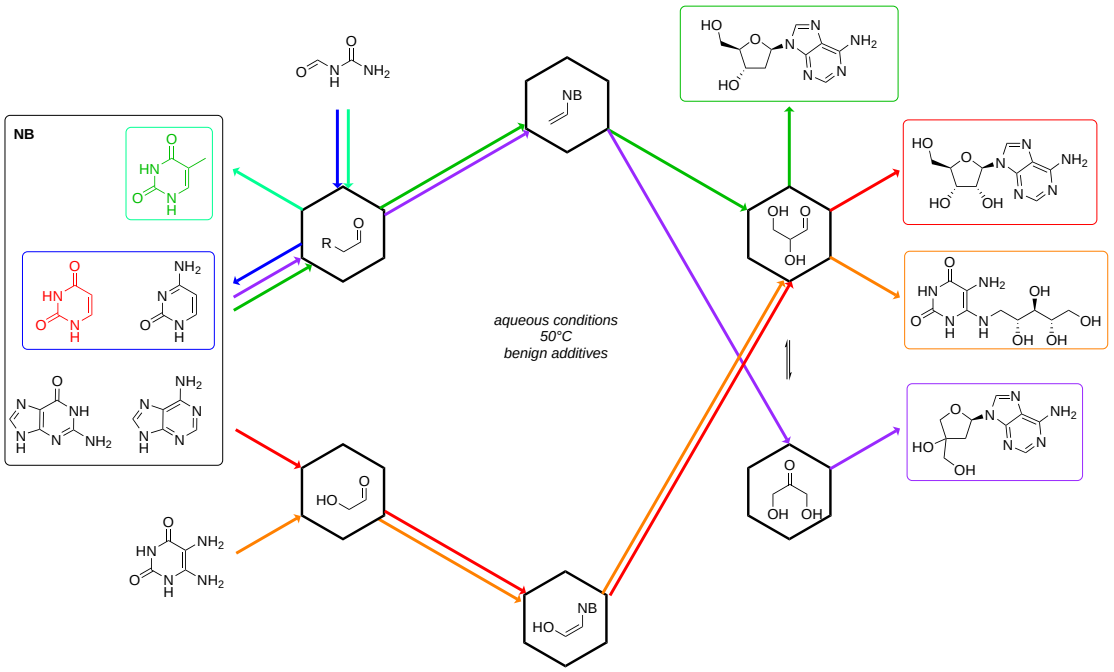
The last two steps of the reaction series leading to **55** are the lumazine formation and the dismutation reaction furnishing the fused pteridine heterocycle from two units of **56**. In case of a successful formation of **54** the lumazine formation was planned with diacetyl **63**. **63** is in close relationship to its biosynthetic analogue **57** and was tested positively earlier. Preliminary tests, examining the system were performed, starting from **60** and **63**. These confirmed the expected facile synthesis of the pteridine heterocycle, lacking the ribityl residue.

As both reactions were reported intensively in previous investigations, see Section 2.8.3, they are not addressed in further detail at this point.

This work aimed to demonstrate prebiotic access to riboflavin precursors or riboflavin itself. The outlined pathway mimics the biosynthetic procedure and is hence under strong consideration to be a plausible route towards riboflavin **55**. This task was executed, following the comprehension that Nature optimises existing pathways and does not invent new ones if not necessary. As for all multistep pathways in a prebiotic context, the major point of criticism is its multi-step nature, additionally implementing slight changes of the conditions. In these cases, sequential addition of the reagents is needed, which is not favourable, in terms of plausibility.

Although the delineated pathway seemed promising, at first sight, the ribosylation at **60** and its reduction could not be verified unambiguously. Extensive investigation is needed to affirm the posed evidence.

6 Conclusions



Scheme 6.1: Concluding scheme illustrating all pathways examined in this thesis. The investigation of the formation of the pyrimidine nucleobases uracil, cytosine and thymine **1b**,**d**,**e** is depicted (blue). The transition of all canonical nucleobases **1** to the vinyl nucleobases **68** (green and violet) and their reaction to DNA **3** (green) and DApiNA **67** (violet) was studied. Formation of the acetadehyde species of **1a**–**c**,**e** (red) and 5,6-diaminouracil **60** (orange) and their transition to RNA **2** (red) and riboflavin **55** (orange) was examined.

In the following, a summary of the results, presented in chapter 5, is given.

A different disconnection approach was considered, leading to a new perspective in the synthetic accessibility of **3**-nucleosides. Omitting the use of **d4a**, **10a** as a more easily accessible sugar-building block was placed at the start of the synthesis. Differing from the traditional procedure towards **2**, forming the anomeric bond in reaction with D-**4a**, this novel procedure is based on a scission between the 2'-3'-carbon. Hence,

not a pentose **4a**, but a triose **10a** is used.

This circumvents the problem of instability of the pentose-sugar series. The 1'- and 2'-carbon is provided by assembling a stable enamine **68** from **1** and **35**. The bijective formation of **68** was given by coinjection of synthetic samples of **68**. This places the crucial anomeric bond formation step at the beginning of the reaction sequence. In a subsequent reaction resembling amino catalysis, **68** nucleophilically attacks the carbonyl of **10a**, furnishing the deoxy nucleosides after a *5-exo-trig* cyclisation.

Synthetic access to all canonical **3** is possible under ambient conditions. The conditions are 50 °C in aqueous media, with the non-mandatory enhancing effect of a alkaline additive *e.g.* NaOAc. There is no need for drastic changes in temperature and pH. The deoxyribonucleosides are obtained in exclusive β-selectivity. Enhancing the pathways plausibility, **3** is formed in exclusive furano-regioselectivity. This could be demonstrated for **3a**–**c**. All evidence is given by UPLC-QTOF-MS coinjection with authentic, fully characterised samples. The **pyr-3** were synthesised and purified as their α/β-anomers by preparative HPLC. Highly-selective reactions like these might give an answer to the long-searched question on the origin of homochirality. Thus, it is concluded that this reaction series places a strong evidence in the synthetic accessibility to deoxyribonucleosides **3**.

Regarding the molecular evolution of **3**, possible nucleoside predecessors and a direct progenitor were found. Using **9** and **24** as the sugar forming building block, created various nucleoside-derivatives. Replacing **10a** by its tautomer **10b** led to the formation of DApiNA-nucleosides **67**. These are proposed as a possible, direct progenitor of **3**-nucleosides. Assuming that Nature is seeking the simplest possible setting, evolutionary selection might have favoured **3** over **67** due to their less constrained stereo- and regioflexibility. Synthetic access to all **67**-nucleosides was achieved in a new synthetic procedure. In analogy to the prebiotic procedure, it starts from **10b**. After a consecutive series of 6 steps furnishing the deoxyapiose precursor **76**, **67** is obtained *via* a Vorbrüggen nucleosidation from all four canonical nucleobases **1**. Synthetic access to **67** allowed the assignment of the formed **67** in the prebiotic samples. Concomitant development of **3** and **67** under the same prebiotic conditions is possible. Hence, a evolutionary selection must have taken place at a later stage *e.g.* during the polymerisation of DNA.

Aiming for a holistic approach to nucleoside synthesis, a different synthetic procedure towards **1d** is presented. In continuous manner, the composition of **1d** is in intersection with the same pool of starting materials later used in the formation of **3**. Further, the reaction conditions for both syntheses are set in a similar environment. The key molecule of both pathways is **35**. **1d** is assembled from **35** and **79**. The mechanism is described by an ammonia-mediated imine formation of **79**, attacked by the **35**-enol. This furnishes **1d** after an intramolecular *6-exo-trig* cyclisation and subsequent elimination of water. Verification was achieved *via* LC-MS coinjection with commercial samples. Unanticipated, large quantities of **1a** were detected. In theory, variation of the aldehyde to **80** should enable access to **1e**. However, its confirmation was not possible within this thesis. The same counts for **1b** Concluding, a synthesis furnishing two out of five canonical nucleobases is possible from simple starting materials, while the verification of two is still pending.

Not only unique access to **3**-nucleosides is delineated, but also a concomitant novel access to **2**-nucleosides is achieved. Exchanging the enamine forming aldehyde from **35** to **9** led to the formation of **2**-nucleosides. Synthetic samples of **81a–d** were prepared to confirm the formation of the nucleobase acetaldehyde **81**. Higher chemical flexibility complicated the unique verification of **81** in the prebiotic mixtures. Still, the reaction proceeds to give **2**. Compared to the procedure furnishing **3**, using **9** adds more degrees of freedom. In consequence, solely **2a** could be confirmed in its canonical β-furano-form. Contrary, **2c** is explicitly furnished as its β-pyrano-isomer. The **pyr-3** were synthesised and purified as their α/β-anomers by preparative HPLC. The reliability of the proposed mechanism is indicated by these two examples, but not as explicitly as for **3**. The scientific evidence was accomplished by UPLC-QTOF-MS coinjections with authentic samples, both commercial and synthetic. Except for this minor change of introducing a hydroxyl group at the aldehyde, all further parameters remained the same. Both syntheses of **3** and **2** can be conducted in the same chemical environment. These findings imply a coexistence of **3** and **2** at the origins of life.

The pathways towards **2**, **3** and **67** all run under the same reaction conditions. Thus, it can be concluded that the presented type of reaction towards nucleosides is ubiquitous.

Delineating a pathway to **4a** synthesised at heterocycles opened the door for prebiotic synthetic access to RNA cofactors. A prebiotic pathway culminating in riboflavin was delineated. The reaction sequence mimics the biosynthetic procedure. It intersects the biosynthesis by placing **60** as the starting molecule. Again the **4a**-formation is the key step of the riboflavin procedure. In analogy to **2**-nucleoside synthesis, the pentose should be established at **60** from **9** and **10a**. The reaction performs under alkaline conditions at mild conditions of 20 °C. Higher reaction temperatures prevented the desired reaction. So far, hints for the formation of **85** could be provided. Further examination is still needed to elucidate the complete pathway. Imitating a NADH reduction, a subsequent reduction with Hantzsch ester **86** was planned to furnish the ribityl-sidechain. This could not be accomplished in this thesis. The same applies to the pteridine ring formation by diacetyl **63**, leading to ribityl-lumazine **56**. The sequential addition of all presented starting materials is necessary. Yet, this reaction cascade benefits from the use of a single alkaline chemical environment. No major changes in pH or temperature are required to assemble the **55**-precursors. Regarding its close coherence with the **55**-biosynthesis and its benign reaction conditions, this pathway is considered as highly prebiotically plausible.

Simple, feasible and highly-selective pathways were revealed to give prebiotic access to **3**- and **2**-nucleosides. In addition, the evolution of **3**-nucleosides was delineated, leading to the discovery and verification of **67**-nucleosides. In line with the presented nucleoside-pathways, a scenario for the assembly of **1a**,**d** was achieved. Arguing a higher plausibility, it was aimed to use cut-sets of molecules. A pathway prebiotically accessing **55** is described, utilising our novel approach of sugar-assembly at heterocycles.

Concluding, distinct conditions are described, drawing a reproducible line from simple molecules to nucleosides and deoxynucleosides. These conditions clearly speak for a chemically ruled origin of Life rather than a derivation led by coincidence.

7 Future Work

Considerable effort in the context of prebiotic nucleoside synthesis is presented in this thesis. Executing the projects, several loose ends and consecutive projects arose. These are presented in the following.

So far, in none of the presented syntheses to **2** and **3**, the question on the correct stereochemical information of the attached **4a** moiety was as intensively examined. **3a**–**c** are exclusively formed as their canonical β-furanosides. This selectivity is unique to the proposed reaction mechanism. The unambiguous characterisation of **3d** in exclusive β-furano-selectivity still needs to be verified. Regarding **2**, a definite statement about the formation of the pyrimidine nucleosides needs to be given. Further, formation of **ara-pyr-2** should be investigated by the synthesis of characterised reference structures.

In this thesis, the most fundamental question of the origin of homochirality was just mentioned briefly. Starting from a racemic mixture of precursor sugar should lead to a racemic mixture of nucleoside. In the context of the origin of homochirality, however, it should be investigated if a D/L-deracemisation, caused by an imbalance in energies, occurs.

Further, it is of interest if the nucleosidation conditions can be chosen to either favour **2** or **3** nucleosides. This would provide a lever to selectively deflect the system.

Several methods were proposed in the literature, which describe phosphorylation of nucleosides and sugars. However, the question must be addressed whether downstream phosphorylation is necessary or if other methods might be feasible. Special attention should be paid to the regioselective phosphorylation of the nucleosides. Directly linked to this task is the subsequent production of **2** and **3** oligomers. Careful choice of the correct reaction conditions will be required to achieve polynucleotides.

A novel method was established to construct sugar chains at nucleobases. It would

be of interest to see whether other heterocyclic units perform equally during the course of the reaction. Further, variation of the building block, that creates the enamine is possible. In both a synthetic and prebiotic context, starting from acetylene might be feasible. Last, the attacked carbonyl functionality can be derived as well. Variation of these three parameters could open up new fields of research and introduce shorter routes to desired molecules.

Regarding the synthesis of the **67**-nucleosides, optimising the synthetic pathway would be desirable. This might open the strategy to large-scale applications. Access to **67a** and **67c** in significant amounts as well as full characterisation is pending. Variation at the beginning of the proposed synthesis also might open up access to the pyranose derivatives of **67**. Further, testing the anti-viral activity of the gained nucleosides from a medical-chemical point of view might be desirable.

Although investigated, the direct formation of the vinyl pyrimidine nucleobases **87** could not be detected in the reaction mixture. Aiming for a continuous pathway from simple starting materials to the nucleosides calls for further optimisation of the reaction conditions. A undirected analysis of the reaction mixtures, starting from a very simple molecule could enlarge the pool of possible heterocyclic compounds. These might be relevant for further applications. Modern analytical tools like the acquisition of undirected MS/MS spectra in combination with network analysis can be beneficial.

Translating the novel ribosylation to **88** opened an opportunity to synthesise riboflavin prebiotically. Although promising results indicated access to **55**, the unambiguous verification is still lacking. Further, the subsequent reduction needs to be investigated carefully. Therefore, a complete synthetic procedure from simple starting molecules, finally resulting in **55** was not yet delineated. Further, it is asked for a prebiotic procedure towards **88**.

Apart from riboflavin, a plethora of other important cofactors are catalysing modern biology. The task to find prebiotic routes towards *e.g.* NADH, pyridoxal phosphate and biotin will be challenging and interesting.

The synthetic access to the Origin of Life is around 150 years old, its fundamental

question possibly dates back to the beginning of reflective thinking. There is still much to discover in the mist of primordial times. Critical, holistic, scientific research will guide us to reveal one of Nature's most taunting mysteries.

Experimental Procedures

General Procedures

All chemicals and solvents were purchased from ABCR, ACROS ORGANICS, ALFA AESAR, FLUORO CHEM, JENA BIOSCIENCE, MERCK, SIGMA ALDRICH, EURISOTOP or TCI in the qualities *puriss.*, *p. a.* or *purum* and were used without further purification.
General techniques: Reactions with water- or air-sensitive compounds were performed, applying standard Schlenk techniques. The used glassware has been dried under high vacuum by using a heat gun and was consecutively flushed with argon. To ensure water- and air-free reaction conditions and a pressure compensation during the reaction, the equipment was closed with septa and an argon-filled balloon. Furthermore, syringes and needles were used to transfer solvents or reagents and were purged three times with argon prior to use. Purified products were dried under high vacuum.

Flash column chromatography was performed on ACROS ORGANICS Silica 60A (35-70 μm). Technical solvents for column chromatography were distilled with a HEIDOLPH rotaevaporator prior to use. Thin layer chromatography (TLC) was performed on MERCK 60 (silica gel F254) plates and the separated reaction products were visualized either with UV-light (254 nm or 366 nm), a Vanilin stain (8.6 g Vanillin and 2.5 mL H_2SO_4 in 200 mL EtOH) or a $KMnO_4$ stain (1.5 g $KMnO_4$, 10 g K_2CO_3, and 1.25 mL 10% NaOH in 200 mL water).

^{1}H and ^{13}C{^{1}H} NMR spectra were recorded in deuterated solvents on VARIAN VXR400S, VARIAN INOVA 400 and BRUKER AMX 600 spectrometers and calibrated to the residual solvent peak (methanol-d_4: $\delta_H = 3.31$ ppm, $\delta_C = 49.00$ ppm; DMSO-d_6 $\delta_H = 2.50$ ppm, $\delta_C = 39.52$ ppm; water-d_2: $\delta_H = 4.79$ ppm; chloroform-d_1: $\delta_H = 7.26$ ppm,

δ_C = 77.16 ppm; acetonitrile-d_3: δ_H = 1.94 ppm, δ_C = 1.32 ppm, 118.26 ppm;). The chemical shifts (δ) are given in ppm, the coupling constants (J) in Hz. Multiplicities are abbreviated as s = singlet, d = doublet, t = triplet, m = multiplet or combinations of these. Full elucidation of the structures was ensured by 2D NMR spectra (COSY, HSQC, HMBC). Please note that the numeration in the depicted structures may not accord to IUPAC rules and for clarity, residual solvent signals have not been assigned.[251]

High resolution electron impact ionization mass spectra (HR-EI-MS) were recorded on a THERMO FINNIGAN MAT 95Q or a THERMO FINNIGAN MAT 90.

High resolution electrospray ionisation mass spectra (HR-ESI-MS) were recorded on a THERMO FISHER Exactive Plus Orbitrap via direct injection. As injection solvent, iPrOH:water (0.05% formic acid) (80:20) was used. The ions were analysed in the scan range of 135–750 m/z with a resolution of 280000. No prior splitting was applied. The flow rate was set to 0.05 mL/min. The following parameters were applied during the mass analysis: Sheath gas flow: 2.00 L/min; auxiliary gas flow rate: 1.00 L/min; spray voltage: 4.00 kV; spray current: 0.80 µA; capillary temperature: 280 °C. Liquid Chromatography was performed on an AGILENT TECHNOLOGIES 1260 Infinity LC System. For separation, a AGILENT Extend C18 (2.10 mmx50.0 mm, 1.80 µm particle size) column was used. During analysis, the column was constantly set to 40.0 °C. The mobile phase was a gradient prepared from water containing 0.05% (v/v) formic acid (eluent A) and methanol (eluent B)[239] In a second methode separating strongly polar compounds, the mobile phase was a gradient prepared from water containing 0.1% (v/v) heptfluoro butyric acid (eluent A) and methanol (eluent B) . The gradient used for LC-MS was: 0% B held constant for 0.54 min, then increased to 40% B within 2.38 min and held for 0.59 min. Subsequent increase to 100% of eluent B within 0.03 min and held for 0.57 min to flush the column, subsequent reconstitution of the starting conditions within 0.03 min, and re-equilibration with 0% B for 6.81 min resulted in a total analysis time of 11.0 min. The liquid flow was set to 0.55 mL/min. In a second method, Liquid Chromatography was performed on an AGILENT TECHNOLOGIES 1290 Infinity LC System. A DIONEX Acclaim 120 C18 (3.00 mm x 150 mm, 3.00 µm particle size) column was used. During analysis, the column was constantly set to 40 °C. UV-spectra at the wavelength of 260 nm were recorded. The mobile phase

was a gradient prepared from water containing 0.1% (v/v) formic acid (eluent A) and methanol (eluent B) for LC-MS and for LC the mobile phase was a gradient prepared from water (eluent A) and methanol (eluent B). The gradient used for LC-MS or LC was: 0% B held constant for 2.00 min, then increased to 40% B within 6.00 min and held for 2.00 min. Subsequent increase to 100% of eluent B within 0.10 min and held for 1.90 min to flush the column, subsequent reconstitution of the starting conditions within 0.10 min, and reequilibration with 0% B for 5.90 min resulted in a total analysis time of 20.0 min. The liquid flow was set to 1.00 mL/min. The highly pure water was obtained from a VWR Puranity PU 15 UV water purification plant. The methanol and acetonitrile was purchased in HPLC-MS grade from ALDRICH. For method one, consecutive downstream mass analysis was performed by an AGILENT TECHNOLOGIES 6550 iFunnel Q-TOF/MS. The ions were analysed in the full scan range of 40–1700 m/z with a resolution of 3200 m/z, 2.00 GHz, extended dynamic range. No prior splitting was applied. The following parameters were applied during the analysis: gas temperature: 200 °C, drying gas flow: 14.00 L/min, nebulizer pressure: 2.41 bar, sheath gas temperature 275 °C, sheath gas flow: 11.00 L/min, capillary voltage: 4.00 kV, capillary current: 6.25 µA and nozzle voltage 0 kV or 1 kV, fragmentor voltage 180 V; for internal calibration the AGILENT TECHNOLOGIES ESI-TOF reference mass solution was used, diluted in acetonitrile:water (95:5). The internal reference masses were set to 121.050873 and 922.009798.

Preparative High-Performance Liquid Chromatography was performed on a AGILENT TECHNOLOGIES 1260 Infinity semi-preparative LC System. A PHENOMENEX Luna C18 (21.2 mm x 250 mm, 5.00 µm particle size) column was used. During analysis the column was constantly kept at room temperature. Automatic collection of the fractions took place. The mobile phase was a gradient prepared from water (eluent A) and methanol (eluent B). The gradient used for prep-LC was: 0% B held constant for 5.55 min, then increased to 40% B within 22.2 min and held for 5.55 min. Subsequent increase to 100% of eluent B within 0.10 min and held for 5.53 min to flush the column, subsequent reconstitution of the starting conditions within 0.10 min, and reequilibration with 0% B for 16.1 min resulted in a total analysis time of 50 min. The liquid flow was set to 30.00 mL/min. The given method may vary for the different nucleosides. Detailed descriptions of the methods used are placed at the corresponding

compounds.

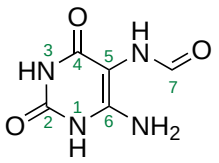
Lyophilisation was performed on a CHRIST Alpha 1-2 LDPlus. The samples were frozen in liquid nitrogen prior to lyophilisation.

X-ray intensity data was measured on a BRUKER D8 Venture TXS system equipped with a multilayer mirror monochromator and a Mo Kαrotating anode X-ray tube ($\lambda = 0.71073$Å). The frames were integrated with the BRUKER SAINT software package. Data were corrected for absorption effects using the Multi-Scan method (SADABS). The structure was solved and refined using the BRUKER SHELXTL Software Package. All C-bound hydrogen atoms have been calculated in ideal geometry riding on their parent atoms while the N-and O-bound hydrogen atoms have been refined freely. The figures have been drawn at the 50% ellipsoid probability level.[252]

Synthetic Procedures

N-(6-amino-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)formamide (32a)

The reaction was performed according to Koch[103]



5,6-Diaminouracil-hemisulfate (3.74 g, 9.80 mmol, 1 eq.) was suspended in 1 M NaOAc solution (14 mL, pH 1.1, acidified with HCl). Formic acid (0.40 mL, 10.6 mmol, 3.8 eq.) was added and stirred for 4 days at room temperature. The pH was adjusted with ammonia solution (5 M) to pH 8-9. The solution was kept at 4 °C overnight and the crystals were filtered off and washed with cold ethanol to give the product (0.31 g, 2.04 mmol, 73%) as colorless crystals.

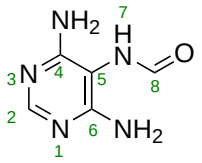
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 8.91 (s, 1 H, H_{7cis}), 8.18 (d, ^{3}J = 11.5 Hz, 1 H, H_{7trans}, 8.03 (d, ^{3}J = 1.5 Hz, 1 H, H_{8cis}), 7.94 (d, ^{3}J = 11.5 Hz, 1 H, H_{8trans}). 7.76 (s, 2 H, H_{3cis}), 7.75 (s, 2 H, H_{3trans}), 6.53 (s, 2 H, $H_{4–NH2trans}$), 6.25 (s, 2 H, $H_{4–NH2cis}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 165.9 (C_{8trans}), 160.2 (C_{8cis}), 159.6 (C_{4trans}), 159.3 (C_{6trans}), 158.8 (C_{4cis}), 158.2 (C_{6cis}), 147.3 (C_{2trans}), 147.0 (C_{2cis}), 97.4 (C_{5trans}), 97.3 (C_{5cis}).

HR-EIMS (*m/z* [M]$^{+}$): calc. for $[C_5H_7N_4O_3]^{+}$: 171.0513; **obs.:** 171.0515.

N-(4,6-diaminopyrimidin-5-yl)formamide (32b)

The reaction was performed according to *Becker et al.*[105]



4,5,6-triaminopyrimidine-sulfate (0.62 g, 2.80 mmol, 1 eq.) was suspended in 1 M NaOAc solution (14 mL, pH 1.1, acidified with HCl). Formic acid (0.40 mL, 10.6 mmol, 3.8 eq.) was added and stirred for 4 days at room temperature. The pH was adjusted with ammonia solution (5 M) to pH 8-9. The solution was kept at 4 °C overnight and the crystals were filtered off and washed with cold ethanol to give the product (0.31 g, 2.04 mmol, 73%) as colorless crystals.

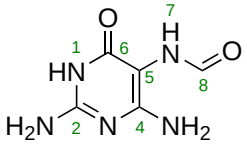
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 8.78 (s, 1 H, H_7), 8.08 (s, 1 H, H_8), 7.74 (s, 1 H, H_2, 5.96 (s, 4 H, $H_{4/6–NH2}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 160.9 (C_8), 159.5 ($C_{4/6}$), 155.7 (C_2), 93.8 (C_5).

HR-EIMS (*m/z* [M]$^+$): calc. for $[C_5H_8N_5O]^+$: 154.0723; **obs.:** 154.0720.

N-(2,4-diamino-6-oxo-1,6-dihydropyrimidin-5-yl)formamide (32c)

The reaction was performed according to *Becker et al.*[105]



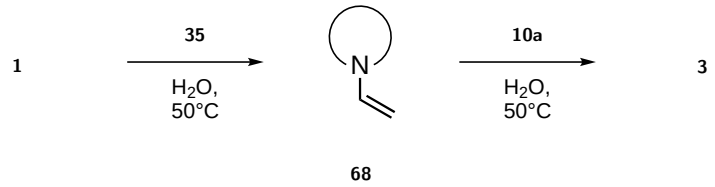
2,4,5-triamino-6-hydroxypyrimidine-sulfate (2.39 g, 2.50 mmol, 1.00 eq.) was suspended in formic acid (7.5 mL) and sodium formiate (0.17 g, 2.50 mmol, 1.00 eq.) was added. The mixture was heated to reflux for 2 h. The solvent was evaporated and water was added until everything was dissolved. Ammonia solution (5 M) was added to the solution until pH 8 was reached. The solution was cooled overnight at 4 °C. The precipitate was filtered off to obtain **32c** (0.30 g, 1.78 mmol, 71.4%) as light beige solid.

^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 10.11 (s, 2 H, $H_{1cis/trans}$), 8.49 (s, 1 H, H_{7cis}), 8.00 (d, ^{3}J = 1.6 Hz, 1 H, H_{8cis}, 7.86 (d, ^{3}J = 11.7 Hz, 1 H, H_{7trans}), 7.73 (d, ^{3}J = 11.6 Hz, 1 H, H_{8trans}). 6.26 (s, 2 H, $H_{2–NH2trans}$), 6.19 (s, 2 H, $H_{2–NH2cis}$), 6.04 (s, 2 H, $H_{4–NH2trans}$), 5.81 (s, 2 H, $H_{4–NH2cis}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 166.9 (C_{8trans}), 161.5 (C_{6trans}), 160.7 (C_{8cis}), 160.7 (C_{6cis}), 160.1 (C_{4trans}), 159.5 (C_{4cis}), 153.5 (C_{2trans}), 153.4 (C_{2cis}), 88.7 (C_{5trans}), 88.5 (C_{5cis}).

HR-EIMS (*m/z* [M]$^{+}$): calc. for $[C_5H_8N_5O]^{+}$: 170.0673; **obs.:** 170.0675.

General procedure leading to 3

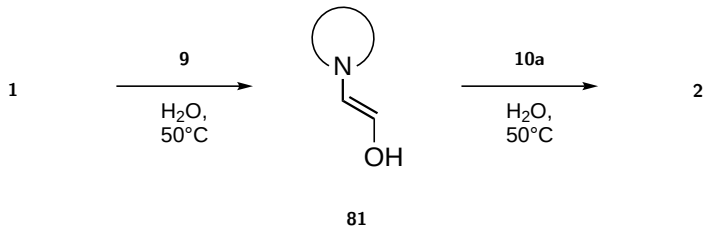


1 or a mixture of **1a**,**c**–**e**(0.10 mmol, 1.00 eq.) was suspended in H_2O (1.00 mL) and the corresponding additive (1.00 eq.) was added (s. Table 5.1). **35** (1.00 mmol, 10.00 eq.) and **10a** (0.15 mmol, 0.15 eq.) were added and the reaction mixture was stirred at 50 °C for 4 d. An aliquot of 5.00 µL was taken every 24 h and the sample was analysed via LC-QTOF-MS and Orbitrap-MS.

Table E.1: Summary of the synthesised vinylnucleobases **68** and the β-deoxyribonucleosides **3** after 4 d under prebiotic conditions. The retention times refer to the UPLC-column AGILENT Extend C18 (2.10 mmx50.0 mm, 1.80 µm particle size).

compound	t_R [min]	***m/z* ([M+H$^+$]) calc. for**	***m/z* ([M+H$^+$]) obs.**
68a	2.34	$[C_7H_7N_5]^+$ 162.0774	162.0774
68b	3.9	$[C_7H_9N_2O_2]^+$ 153.0659	153.0659
68c	3.1	$[C_7H_8N_5O_2]^+$ 178.0723	178.0723
68d	0.5	$[C_6H_8N_3O]^+$ 138.0662	138.0662
3a	1.03	$[C_{10}H_{14}N_5O_3]^+$ 252.1091	252.1087
3b	2.16	$[C_{10}H_{15}N_2O_4]^+$ 243.0975	243.0970
3c	1.24	$[C_{10}H_{14}N_5O_4]^+$ 268.1040	268.1036
3d	0.43	$[C_9H_{14}N_3O_4]^+$ 228.0979	228.0976

General procedure leading to 2

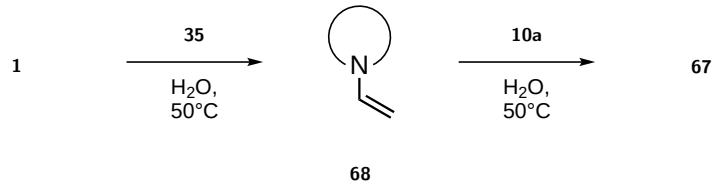


1 or a mixture of **1a**,**c**–**e** (0.10 mmol, 1.00 eq.) was suspended in H_2O (1.00 mL) and the corresponding additive (1.00 eq.) was added. **9** (0.20 mmol, 2.00 eq.) and **10a** (0.15 mmol, 0.15 eq.) were added and the reaction mixture was stirred at 50 °C for 4 d. An aliquot of 5.00 µL was taken every 24 h and the sample was analysed via LC-QTOF-MS and Orbitrap-MS.

Table E.2: Summary of the synthesised nucleobase acetaldehyde **81** and the β-ribonucleosides **2**. The retention times refer to the UPLC-column AGILENT Extend C18 (2.10 mmx50.0 mm, 1.80 µm particle size).

compound	t_R [min]	***m/z* ([M+H$^+$]) calc. for**	***m/z* ([M+H$^+$]) obs.**
81a	0.45	$[C_7H_7N_5O]^+$ 178.0723	178.0726
81b	0.7	$[C_6H_7N_2O_3]^+$ 155.0451	155.0454
81c	1.29	$[C_7H_7N_5O_2]^+$ 194.0673	194.00674
81d	0.30	$[C_6H_7N_3O_2]^+$ 154.061	154.0612
2a	1.25	$[C_{10}H_{14}N_5O_4]^+$ 268.1040	268.1045
2b	0.70	$[C_9H_{12}N_2O_6]^+$ 245.0768	245.0777
2c	1.29	$[C_{10}H_{14}N_5O_5]^+$ 284.0989	284.0996
2d	0.36	$[C_9H_{14}N_3O_5]^+$ 244.0928	244.0932

General procedure leading to 67

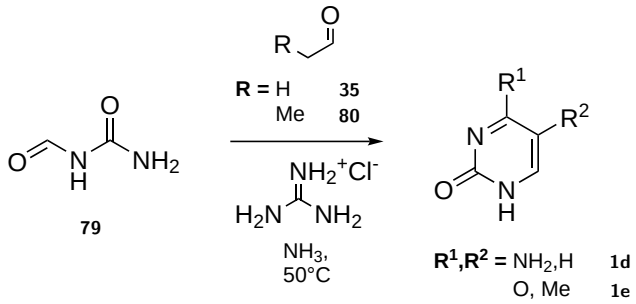


1 or a mixture of **1a**,**c**–**e** (0.10 mmol, 1.00 eq.) was suspended in H_2O (1.00 mL) and the corresponding additive (1.00 eq.) was added (s. Table 5.1). **35** (1.00 mmol, 10.00 eq.) and **10b** (0.15 mmol, 0.15 eq.) were added and the reaction mixture was stirred at 50 °C for 4 d. An aliquot of 5.00 µL was taken every 24 h and the sample was analysed via LC-QTOF-MS and Orbitrap-MS.

Table E.3: Summary of the synthesised **67**. The retention times refer to the UPLC-column AGILENT Extend C18 (2.10 mmx50.0 mm, 1.80 µm particle size).

compound	t_R [min]	***m/z* ([M+H$^+$]) calc. for**	***m/z* ([M+H$^+$]) obs.**
67a	1.66	$[C_{10}H_{14}N_5O_3]^+$ 252.1091	252.1087
67b	2.13/3.37	$[C_{10}H_{15}N_2O_4]^+$ 243.0975	243.0970
67c	0.96	$[C_{10}H_{14}N_5O_4]^+$ 268.1040	268.1036
67d	0.33/0.49	$[C_9H_{14}N_3O_4]^+$ 228.0979	228.0976

General procedure leading to the pyrimidine nucleobases

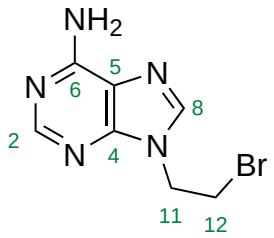


General procedure leading to either **1d** or **1e**

79 (0.5 mmol, 1.00 eq.) was suspended in (*aq.*NH_3) (1.00 mL, 1 M) and guanidine hydrochloride (0.65 mmol, 1.30 eq.) was added. The corresponding aldehyde (2.50 mmol, 5.00 eq.) was added and the reaction mixture was stirred at 50 °C for 4 d. An aliquot of 5.00 µL was taken every 24 h and the sample was analysed via LC-QTOF-MS and Orbitrap-MS.

9-(2-bromoethyl)-9*H*-purin-6-amine

The reaction was performed according to *Ciapetti, et al.*.[253]



9-(2-bromoethyl)-9H-purin-6-amine (10.0 mmol, 1.00 eq.) was dissolved in dry DMF (20 mL). Dibromethane (40.0 mmol, 4.00 eq.) and dry K_2CO_3 (25.0 mmol, 2.50 eq.) were added and the reaction mixture was stirred at room temperature for 48 h. The reaction mixture was filtered over Celite® . The solvent was removed under reduced pressure and the resulting crude product was purified via column chromatography (SiO_2,R_f: 0.26, 5% MeOH in DCM). The product was obtained as a colourless solid (1.51 g, 6.25 mmol, 63.0%).

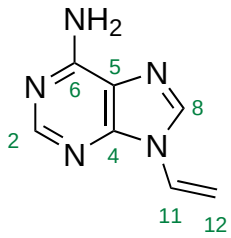
^{1}H-NMR (600 MHz, Methanol-d_4): δ (ppm) = 8.22 (s, 1 H, H_2), 8.17 (s, 1 H, H_8), 4.66 (t, ^3J = 6.1 Hz, 2 H, H_{11}, 3.88 (t, ^3J = 6.1 Hz, 2 H, H_{10}).

^{13}C-NMR (101 MHz, Methanol-d_4): δ (ppm) = 157.4 (C_6), 153.8 (C_2), 150.6 (C_5), 143.0 (C_8), 120.0 (C_4), 46.60 (C_{11}), 30.90 (C_{12}).

HR-EIMS (*m/z* [M]$^+$): calc. for $[C_7H_8BrN_5]^+$: 240.9958; **obs.:** 240.9963.

9-Vinyladenine (68a)

The reaction was performed according to *Ciapetti, et al.*.[253]



Adenine **1a** (6.25 mmol, 1.00 eq.) was dissolved in dry DMF (2.24 mL). The solution was diluted in EtOH (6.70 mL). Sodium ethoxide (12.5 mmol, 2.00 eq.) was added and the reaction mixture was stirred at room temperature for 2 h. The solvent was removed *in vacuo* and 9-vinyladenine **68a** was obtained as a colourless solid (0.791 g, 4.93 mmol, 79.0%), after precipitation with water.

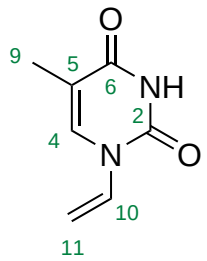
^{1}H-NMR (400 MHz, Methanol-d_4): δ (ppm) = 8.39 (s, 1 H, H_2), 8.23 (s, 1 H, H_8), 7.29 (dd, ^{3}J = 16.0 Hz, 1 H, H_{11}, 6.00 (dd, ^{3}J = 16.0 Hz, 1 H, H_{12}, 5.199 (dd, ^{3}J = 9.2 Hz, 1 H, H_{12}.

^{13}C-NMR (101 MHz, Methanol-d_4): δ (ppm) = 157.4 (C_6), 153.8 (C_2), 150.0 (C_4), 140.0 (C_8), 127.8 (C_{11}), 120.4 (C_5), 105.2 (C_{12}).

HR-EIMS (*m/z* [M]$^{+}$): calc. for $[C_7H_7N_5]^{+}$:161.0696; **obs.:** 161.0695.

1-Vinylthymine (68b)

The reaction was performed according to *Dalpozzo, et al.*.[240]



A suspension of thymine **1e** (10.0 mmol, 1.00 eq.) in HMDS (36.8 mmol, 3.68 eq.), TMSCl (4.50 mmol, 0.45 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in vinylacetate (610.00 mmol, 61.00 eq.). $Hg(OAc)_2$ (0.70 mmol, 0.07 eq.) and TMSOTf (2.50 mmol, 0.25 eq.) were added. The reaction mixture was stirred for 6 h under reflux. The reaction mixture was filtered over neutral activated alumina and the solvent was removed *in vacuo*. The resulting crude product was purified via column chromatography (SiO_2,R_f: 0.59, 10% MeOH in DCM). The product **68b** was obtained as a colourless solid (0.56 g, 3.70 mmol, 37%).

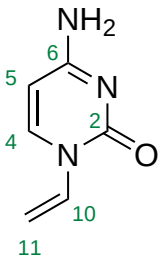
^{1}H-NMR (400 MHz, Methanol-d_4): δ (ppm) = 7.77 (d, ^4J = 1.2 Hz1 H, H_4), 7.18 (dd, ^3J = 16.1, 9.2 Hz, 1 H, H_{10}), 5.30 (dd, ^2J = 2.0 Hz, ^3J = 16.0 Hz 1 H, H_{11}), 4.92 (dd, ^2J = 2.0 Hz, ^3J = 9.2 Hz 1 H, H_{11}), 5.199 (dd, ^4J = 1.3 Hz, 3 H, H_9.

^{13}C-NMR (101 MHz, Methanol-d_4): δ (ppm) = 166.2 (C_6), 151.3 (C_2), 136.9 (C_4), 130.7 (C_{10}), 112.7 (C_5), 101.2 (C_{11}), 12.3 (C_9).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_7H_9N_2O_2]^+$: 153.0659; **obs.:** 153.0655.

1-Vinylcytosine (68d)

The reaction was performed according to *Dalpozzo, et al.*.[240,254]



A suspension of cytosine **1d** (4.40 mmol, 1.00 eq.) in HMDS (16.20 mmol, 3.68 eq.), TMSCl (1.90 mmol, 0.45 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in vinylacetate (305.00 mmol, 61.00 eq.). $Hg(OAc)_2$ (0.30 mmol, 0.07 eq.) and TMSOTf (1.10 mmol, 0.25 eq.) were added. The reaction mixture was stirred for 6 h under reflux. The reaction mixture was filtered over neutral activated alumina and the solvent was removed *in vacuo*. The resulting crude product was purified via column chromatography (SiO_2,R_f: 0.56, 9% MeOH in DCM). The product **68d** was obtained as a colourless solid (0.10 g, 0.75 mmol, 17%).

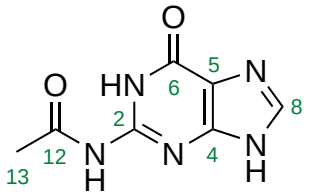
^{1}H-NMR (400 MHz, Methanol-d_4): δ (ppm) = 8.23 (d, ^3J = 7.5 Hz 1 H, H_4), 7.48 (d, ^3J = 7.5 Hz, 1 H, H_5), 7.30 (dd, ^3J = 16.0, 9.0 Hz, 1 H, H_9), 5.52 (dd, ^2J = 1.9 Hz, ^3J = 16.0 Hz 1 H, H_{10}), 5.16 (dd, ^2J = 1.9 Hz, ^3J = 9.0 Hz 1 H, H_{10}.

^{13}C-NMR (101 MHz, Methanol-d_4): δ (ppm) = 164.7 (C_6), 157.0 (C_2), 145.6 (C_4), 133.2 (C_9), 106.2 (C_{10}), 198.9 (C_5).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_6H_8N_3O]^+$: 138.0662; **obs.:** 138.0659.

N-(6-oxo-5,9-dihydro-*6H*-purin-2-yl)acetamide 70

The reaction was performed according to *Zou, et al.*.[254]



Acetic anhydride (92.00 mmol, 2.30 eq.) was added to a suspension of guanine **1c** (40.00 mmol, 1.00 eq.) in *N,N*-Dimethylacetamid (50 mL). The reaction mixture was stirred at 160 °C for 7 h. The brown clear solution was cooled down to 5 °C and the precipitate was filtered off, to give the product as a colourless solid (6.07 g, 31.40 mmol, 79%).

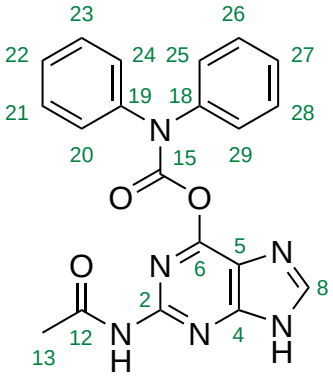
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 12.20 (s, 1 H, H_{NH}), 11.72 (s, 1 H, H_{NH}), 8.43 (s, 1 H, H_8)„ 2.19 (s, 3 H, H_{13}).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 173.8 (C_{12}), 168.1 (C_4), 154.7 (C_6), 147.9 (C_2), 137.5 (C_8), 121.5 (C_5), 23.9 (C_{13}).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_7H_8N_5O_2]^+$: 194.0673; **obs.:** 194.0671.

2-Acetamido-*9H*-purin-6-yl diphenylcarbamate

The reaction was performed according to *Zou, et al.*.[254]



70 (10.00 mmol, 1.00 eq.) was suspended in dry pyridine (50 mL) and DIPEA (21.50 mmol, 2.15 eq.) was added. After addition of diphenylcarbamic chloride (11.00 mmol, 1.10 eq.) the reaction mixture was stirred for 1 h at room temperature. The colour changed from colourless to dark red. H_2O (4 mL) was added and the reaction mixture was stirred for another 10 min. The solvent was removed *in vacuo*. The resulting crude product was stirred in EtOH/H_2O (1:1, 100 mL) at 100 °C for 2 h. The precipitate was filtered off and dried to give the product as a colourless solid (2.70 g, 6.95 mmol, 70%).

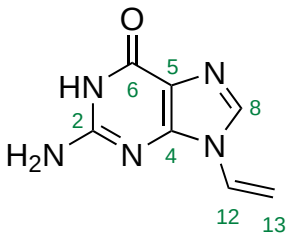
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 10.57 (s, 1 H, H_{NH}), 8.41 (s, 1 H, H_8), 7.60–7.19 (m, 10 H, H_{18}), 2.14 (s, 3 H, H_{13}).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 168.5 (C_{12}), 152.0 (C_{15}), 150.3 (C_2), 141.7 (C_4), 129.4 ($C_{18–29}$), 127.2 ($C_{18–29}$), 126.9 ($C_{18–29}$), 24.4 (C_{13}).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{20}H_{17}N_6O_3]^+$: 389.1357; **obs.:** 389.1355.

9-Vinylguanine (68c)

The reaction was performed according to *Dalpozzo, et al.*.[240]



2-acetamido-9H-purin-6-yl diphenylcarbamate (5.00 mmol, 1.00 eq.) was suspended in vinylacetate (305.00 mmol, 61.00 eq.). $Hg(OAc)_2$ (0.35 mmol, 0.07 eq.) and TMSOTf (1.25 mmol, 0.25 eq.) were added. The reaction mixture was stirred for 2 h under reflux. The reaction mixture was filtered over neutral activated alumina and the solvent was removed *in vacuo*. The resulting crude product was purified via column chromatography (SiO_2, R_f: 0.2, 2.5% MeOH in DCM), after deprotection in methanolic ammonia solution. The product **68c** was obtained as a colourless solid (0.10 g, 0.69 mmol, 14%).

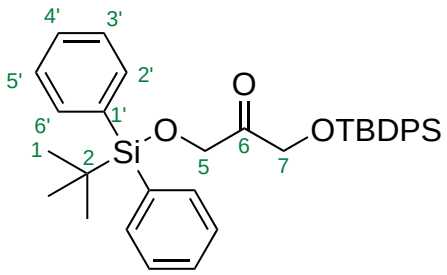
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 10.69 (s, 1 H, H_{NH}), 8.03 (s, 1 H, H_8), 7.00 (dd, ^3J = 16.1, 9.2 Hz, 1 H, H_{12}), 6.52 (s, 2 H, H_{NH2}), 5.80 (d, ^3J = 16.8 Hz, 1 H, H_{13}), 4.98 (d, ^3J = 9.9 Hz, 1 H, H_{13}).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 156.7 (C_6), 154.0 (C_2), 150.3 (C_4), 134.7 (C_8), 126.82 (C_{12}), 116.9 (C_5), 102.8 (C_{13}).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_7H_8N_5O]^+$: 178.0723; **obs.:** 178.0720.

2,2,10,10-Tetramethyl-3,3,9,9-tetraphenyl-4,8-dioxa-3,9-disilaundecan-6-one (72)

The reaction was performed according to Yang *et. al.*.[255]



Dihydroxyaceton **10b** (10.00 mmol, 1.00 eq.) and DMAP (2.30 mmol, 0.23 eq.) were dissolved in dry pyridine (11 mL) and cooled to 0 °C. *tert*-butylchlorodiphenylsilane (22.00 mmol, 2.20 eq.) was added slowly and the reaction mixture was stirred for 16 h. The reaction mixture was poured on ice, extracted with ethyl acetate and dried over $MgSO_4$. The solvent was removed *in vacuo*. The residue was purified by column chromatography (SiO_2, R_f: 0.48, 100% DCM) to yield the product as a white solid (4.14 g, 7.30 mmol, 73%).

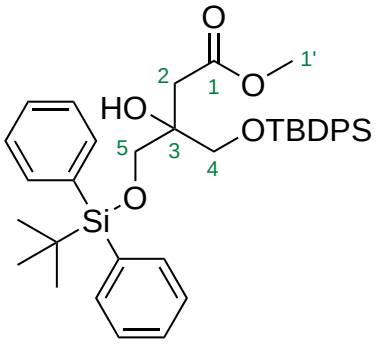
^{1}H-NMR (400 MHz, Chloroform-d_1): δ (ppm) = 7.63–7.59 (m, 4 H, $H_{2'}$), 7.48–7.42 (m, 2 H, $H_{4'}$), 7.41–7.35 (m, 4 H, $H_{3'}$), 4.44 (s, 2 H, $H_{5,7}$), 1.05 (s, 9 H, H_1).

^{13}C-NMR (101 MHz, Chloroform-d_1): δ (ppm) = 207.3 (C_2), 135.6 ($C_{2',6'}$), 132.7 ($C_{1'}$), 130.1 ($C_{4'}$), 127.9 ($C_{3',5'}$), 68.7 ($C_{1,3}$), 26.8 (C_1), 19.3 (C_2).

HR-ESIMS (*m/z* [M+Na]$^+$): calc. for $[C_{35}H_{43}NaO_3Si_2]^+$: 589.2565;
obs.: 589.2575.

Methyl 4-((*tert*-butyldiphenylsilyl)oxy)-3-(((tert-butyldiphenylsilyl)oxy)methyl)-3-hydroxybutanoate (73)

The reaction was performed according to Kanai *et. al.*.[243]



To a solution of wilkinson's catalyst (0.50 mmol, 0.05 eq.) in THF (20.00 mL) at 0 °C a solution of **72** (10.00 mmol, 1.00 eq.) in THF (20.00 mL) was added. After stirring for 5 min diethylzinc (20.00 mmol, 2.00 eq., 1.00 M in toluene) was added dropwise. Methyl 2-bromoacetate (10.00 mmol, 1.00 eq.) was added dropwise and the reaction mixture was stirred for 30 min. A saturated aqueous solution of $NaHCO_3$ was added and the reaction mixture was filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.45, 0% → 2% EtOAc in iso-hexanes). The product was obtained as a colourless oil (5.28 g, 8.24 mmol, 82.4%).

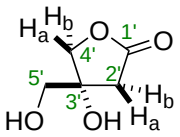
^{1}H-NMR (400 MHz, Chloroform-d_1): δ (ppm) = 7.63 (m, 8 H, $H_{2',6'}$), 7.45–7.38 (m, 4 H, $H_{4'}$), 7.38–7.31 (m, 8 H, $H_{3',5'}$), 3.75–3.66 (m, 4 H, $H_{4,5}$), 3.61 (s, 3 H, H_{Me}), 2.69 (s, 2 H, H_2), 1.03 (s, 18 H, H_{tBu}).

^{13}C-NMR (101 MHz, Chloroform-d_1): δ (ppm) = 172.6 (C_1), 135.7 ($C_{2',6'}$), 133.2 ($C_{1'}$), 129.9 ($C_{4'}$), 127.9 ($C_{3',5'}$), 74.3 (C_3), 66.8 ($C_{4,5}$) 51.7 (C_{Me}), 38.0 (C_2), 26.9 (C_{tBu}), 19.4 ($C_{tBu–quart}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{38}H_{49}O_5Si_2]^+$: 641.3113; **obs.:** 641.3109.

4-Hydroxy-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one (74)

The reaction was performed according to Nicolaou *et. al.*.[256]



73 (6.42 g, 10.0 mmol, 1.00 eq.) was dissolved in THF (50 mL). The solution was cooled to 0 °C and 70 wt.% HF·pyridine (2.60 mL, 100 mmol, 10.0 eq.) was added dropwise. The reaction was allowed to warm to room temperature and stirred overnight. The reaction was quenched with a solution of TMSCl (13.0 mL, 100 mmol, 10.0 eq.) in methanol (50.0 mL) and stirred for 1 h. The solution was concentrated *in vacuo*. The resulting solid was extracted with water and EtOAc. The combined aqueous layers were washed with EtOAc and then concentrated *in vacuo* at 60 °C. To the remaining pyridinic solution, toluene was added (10 mL). The solution was dried over HV overnight. A remaining concentrated solution of the product in pyridine (1.39 g) was used without further purification.

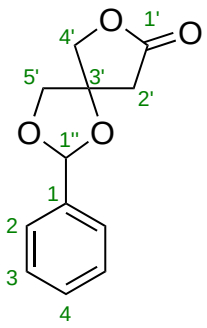
^{1}H-NMR (600 MHz, Deuteriumoxide-d_2): δ (ppm) = 4.45 (d, $^2J_{2a–2b}$ = 10.2 Hz, 1 H, $H_{4'b}$), 4.33 (dd, $^2J_{4a–4b}$ = 10.1 Hz, $^3J_{4'–2'a}$ = 0.9 Hz, 1 H, $H_{4'a}$), 3.72 (d, ^3J = 1.5 Hz, 2 H, $H_{5'}$), 2.93 (d, $^3J_{5–5'}$ = 18.0 Hz, 1 H, $H_{2'b}$), 2.61 (dd, $^3J_{2'a–2'b}$ = 18.0 Hz, $^3J_{2'a–4'}$ = 0.9 Hz, 1 H, $H_{2'a}$).

^{13}C-NMR (151 MHz, Deuteriumoxide-d_2): δ (ppm) = 179.3 ($C_{1'}$), 77.3 ($C_{3'}$), 76.9 ($C_{4'}$), 64.4 ($C_{5'}$), 38.7 ($C_{2'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_5H_9O_4]^+$: 133.0495; **obs.:** 133.0496.

2-Phenyl-1,3,7-trioxaspiro[4.4]nonan-8-one (75)

The reaction was performed according Zinner *et al.*[257]



74 (1.39 g, 10.0 mmol theor.) was dissolved in benzaldehyde (16.0 mL, 158 mmol, 16.0 eq.). Anhydrous $ZnCl_2$ (2.16 g, 15.8 mmol, 1.5 eq.) and anhydrous Na_2SO_4 (1.12 g, 7.91 mmol, 0.75 eq.) were dried in a Schlenk flask under vacuum using a heat gun. The salts were allowed to cool to room temperature and were added to the organic mixture. The reaction mixture was stirred at 50 °C overnight. The reaction mixture was allowed to cool to room temperature, diluted with EtOAc and filtered over Celite. The solution was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified via flash column chromatography (SiO_2, R_f: 0.55, 0% → 5% Methanol in DCM) to give the product as a pale yellow solid of 2 diastereoisomers containing 1.3 equiv. of pyridine (2.32 g, 43 mol %, 7.20 mmol, 72% over 2 steps).

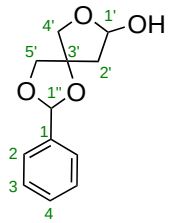
^{1}H-NMR (600 MHz, Chloroform-d_1): δ (ppm) = 7.47–7.44 (m, 2 H, H_2), 7.43–7.39 (m, 3 H, $H_{3/4}$), 5.94–5.91 (m, 1 H, $H_{1''}$), 4.60–4.28 (m, 2 H, $H_{4'}$), 4.24–4.00 (m, 2 H, $H_{5'}$), 3.03–2.70 (m, 2 H, $H_{2'}$).

^{13}C-NMR (151 MHz, Chloroform-d_1): δ (ppm) = 174.1 ($C_{1'}$), 136.4 (C_1), 129.9 (C_4), 128.7 (C_3), 126.5 (C_2), 104.6 ($C_{1''}$), 82.7 ($C_{3'}$), 76.4 ($C_{4'}$), 72.7 ($C_{5'}$), 39.3 ($C_{2'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{12}H_{13}O_4]^+$: 221.0808; **obs.:** 221.0807.

2-Phenyl-1,3,7-trioxaspiro[4.4]nonan-8-ol (78)

The reaction was performed according to Corey *et al.*[258]



75 (0.85 g, 4.02 mmol, 1.00 eq.) was dissolved in dry DCM (16.0 mL) and cooled to –78 °C. DIBAL-H (3.35 mL, 1.50 M in toluene, 1.25 eq.), precooled to –78 °C, was added dropwise. The reaction was stirred at –78 °C for 2 h and quenched with excess saturated aqueous solution of potassium sodium tartrate. The mixture was allowed to warm to room temperature, filtered over Celite and extracted with DCM. The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified via flash column chromatography (SiO_2, R_f 0.60, 0% → 11% Methanol in DCM) to give the product as a transparent oil of 4 diastereoisomers (0.42 g, 1.89 mmol, 47%).

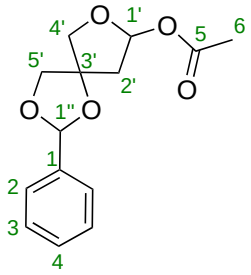
^{1}H-NMR (400 MHz, Chloroform-d_1): δ (ppm) = 7.50–7.44 (m, 2 H, H_2), 7.42–7.36 (m, 3 H, $H_{3/4}$), 5.95–5.79 (m, 1 H, $H_{1''}$), 5.73–5.48 (m, 1 H, $H_{1'}$), 4.73–3.33 (m, 4 H, $H_{5'/4'}$), 2.62–2.09 (m, 2 H, $H_{2'}$).

^{13}C-NMR (101 MHz, Chloroform-d_1): δ (ppm) = 137.0 (C_1), 129.0 ($C_{3,4}$), 126.6 (C_2), 103.9 ($C_{1''}$), 98.9 ($C_{1'}$), 87.2 ($C_{3'}$), 75.2 ($C_{4',5'}$), 44.2 ($C_{2'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{12}H_{15}O_4]^+$: 223.0965; **obs.:** 223.0964.

2-Phenyl-1,3,7-trioxaspiro[4.4]nonan-8-yl acetate (76)

The reaction was performed according to Gao *et al.*[259]



78 (0.38 g, 1.70 mmol, 1.00 eq.) was dissolved in pyridine (1.7 mL, 1.00 M) and cooled to 0 °C. Ac_2O (200 µL, 2.13 mmol, 1.25 eq.) was added dropwise. The solution was allowed to warm to room temperature and stirred overnight. The reaction was quenched with water and extracted with DCM. The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified via flash column chromatography (SiO_2, R_f 0.45, 0% → 2% Methanol in DCM) to give the product as a pale yellowish oil of 4 diastereoisomers (0.39 g, 1.49 mmol, 88%).

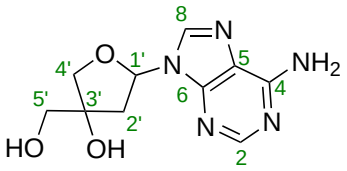
^{1}H-NMR (400 MHz, Chloroform-d_1): δ (ppm) = 7.47–7.41 (m, 2 H, H_2), 7.40–7.33 (m, 3 H, $H_{3/4}$), 6.45–6.29 (m, 1 H, $H_{1'}$), 5.94–5.78 (m, 1 H, $H_{1''}$), 4.40–3.90 (m, 4 H, $H_{5'/4'}$), 2.98–2.15 (m, 2 H, $H_{2'}$), 2.11–2.02 (m, 3 H, H_6).

^{13}C-NMR (101 MHz, Chloroform-d_1): δ (ppm) = 170.2 (C_5), 137.2 (C_1), 129.5 ($C_{3,4}$), 126.6 (C_2), 104.1 ($C_{1''}$), 98.5 ($C_{1'}$), 86.1 ($C_{3'}$), 76.0 ($C_{4',5'}$), 43.1 ($C_{2'}$), 21.2 (C_6).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{14}H_{17}O_5]^+$: 265.1071; **obs.:** 265.1069.

Deoxyapiofuranosyl adenosine (67a)

The reaction was performed according to Niedballa *et al.*[241]

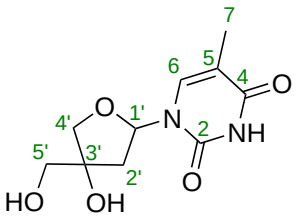


76 (264 mg, 1.00 mmol, 1.00 eq.) was dissolved in dry MeCN (6.30 mL). 6-Chloropurine (170 mg, 1.10 mmol, 1.10 eq.), HMDS (227 µL, 1.20 mmol, 1.2 eq.) and TMSCl (153 µL, 1.20 mmol, 1.20 eq.) were added. The mixture was stirred at room temperature for 15 min and $SnCl_4$ (145 µL, 1.25 mmol, 1.25 eq.) was added dropwise. The reaction mixture was stirred at 50 °C for 24 h. The reaction was quenched with a saturated aqueous solution of $NaHCO_3$ and filtered over Celite®. The filtrate was extracted with EtOAc. The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated in vacuo. Flash column chromatography (SiO_2, R_f 0.45, 0% → 5% Methanol in DCM) yielded the product as a yellowish solid. In a pressure tube, the product was suspended in a solution of 7 M NH_3 in MeOH (5 mL). The tube was sealed and stirred at 70 °C for 24 h. After cooling to room temperature, the mixture was concentrated in vacuo and dried under high vacuum. The product was then dissolved in 80% acetic acid in water (5 mL) at 0 °C and stirred at room temperature for 3 d. The solvent was removed under high vacuum. The remaining solid was suspended in water (5 mL), filtered and the filtrate purified via preparative HPLC.

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{10}H_{14}N_5O_3]^+$: 252.1091; **obs.:** 252.1090.

Deoxyapiofuranosyl thymidine(67b)

The reaction was performed according to Niedballa *et al.*[241]



1e(152 mg, 1.20 mmol, 1.10 eq.) was supended in HMDS (1.99 mL, 9.62 mmol, 8.8 eq.) and TMSCl (202 µL, 1.59 mmol, 1.45 eq.). Catalytic $(NH_4)_2SO_4$ was added and the suspension was heated to 130 °C for 4 h until a clear solution had formed. The solution was allowed to cool to room temperature and then concentrated under high vacuum. **76** (289 mg, 1.09 mmol, 1.00 eq.) was dissolved in dry MeCN (6.8 mL) and was added to the reaction vessel. The mixture was stirred at room temperature for 15 min and $SnCl_4$ (160 µL, 1.37 mmol, 1.25 eq.) was added. The reaction mixture was stirred at 50 °C for 24 h. The reaction was quenched with a saturated aqueous solution of $NaHCO_3$ and filtered over Celite®. The aqueous phase was extracted with EtOAc. The aqueous phase was then concentrated in vacuo at 60 °C. The remaining solid was dissolved in a minimal amount water. The concentrated solution was desalinated and eluted with acetonitrile:water (4:1,20 mL). The eluate was concentrated in vacuo, suspended in water (5 mL), filtered and the filtrate purified via preparative HPLC.

***trans*-67b**:

^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 11.22 (s, 1 H, H_{NH}), 7.76 (d, $^4J_{6–7}$ = 1.2 Hz, 1 H, H_6), 6.03 (dd, $^3J_{1'–2'}$ = 8.0 Hz,$^3J_{1'–2'}$ = 3.4 Hz, 1 H, $H_{1'}$), 5.20 (s, 1 H, $H_{5'–OH}$), 5.04 (s, 1 H, $H_{3'–OH}$), 3.89 (dd, $^2J_{4'–4'}$ = 9.1 Hz,$^3J_{4'–5'}$ = 1.3 Hz,1 H, $H_{4'}$), 3.77 (d, $^3J_{4'–4'}$ = 9.1 Hz, 1 H, $H_{4'}$), 3.39 (d, $^3J_{5'–4'}$ = 1.5 Hz, 2 H, $H_{5'}$), 2.47 (dd, $^2J_{2'–2'}$ = 14.2 Hz, $^3J_{2'–1'}$ = 8.1 Hz, 1 H, $H_{2'}$), 1.83 (ddd, $^2J_{2'–2'}$ = 14.2 Hz, $^3J_{2'–1'}$ = 3.5 Hz, ^3J = 1.3 Hz 1 H, $H_{2'}$), 1.78 (d, $^4J_{7–6}$ = 1.2 Hz, 1 H, H_7).

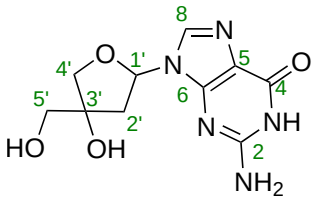
^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 164.6 (C_4), 150.9 (C_2), 136.9 (C_6), 108.6 (C_5),

85.2 ($C_{1'}$), 79.4 ($C_{3'}$), 76.4 ($C_{4'}$), 64.3 ($C_{5'}$), 40.9 ($C_{2'}$), 12.5 (C_7).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{14}H_{14}N_2O_5]^+$: 243.0975; **obs.:** 243.0977.

Deoxyapiofuranosyl guanosine (67c)

The reaction was performed according to Niedballa *et al.*[241]

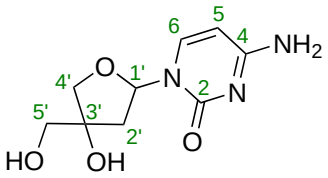


76 (289 mg, 1.00 mmol, 1.00 eq.) was dissolved in dry MeCN (6.80 mL). 2-Acetamido-6-hydroxypurine (230 mg, 1.19 mmol, 1.09 eq.), HMDS (248 µL, 1.31 mmol, 1.2 eq.) and TMSCl (167 µL, 1.20 mmol, 1.20 eq.) were added. The mixture was stirred at room temperature for 15 min and $SnCl_4$ (160 µL, 1.37 mmol, 1.25 eq.) was added dropwise. The reaction mixture was stirred at 50 °C for 24 h. The reaction was quenched with a saturated aqueous solution of $NaHCO_3$ and filtered over Celite®. The filtrate was extracted with EtOAc. The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated in vacuo. Flash column chromatography (SiO_2, R_f 0.30, 0% → 5% Methanol in DCM) yielded the product as a yellowish solid. In a pressure tube, the product was suspended in a solution of 7 M NH_3 in MeOH (5 mL). The tube was sealed and stirred at 70 °C for 24 h. After cooling to room temperature, the mixture was concentrated *in vacuo* and dried under high vacuum. The product was then dissolved in 80% acetic acid in water (5 mL) at 0 °C and stirred at room temperature for 3 d. The solvent was removed under high vacuum. The remaining solid was suspended in water (5 mL), filtered and the filtrate purified via preparative HPLC.

^{13}C-NMR (100 MHz, $CDCl_3$): δ (ppm) = 57.53, 73.18, 104.49;

Deoxyapiofuranosyl cytidine (67d)

The reaction was performed according to Niedballa *et al.*[241]



1d(147 mg, 1.30 mmol, 1.10 eq.) was supended in HMDS (2.15 mL, 10.4 mmol, 8.8 eq.) and TMSCl (217 µL, 1.71 mmol, 1.45 eq.). Catalytic $(NH_4)_2SO_4$ was added and the suspension was heated to 130 °C for 1 h until a clear solution had formed. The solution was allowed to cool to room temperature and then concentrated under high vacuum. **76** (312 mg, 1.18 mmol, 1.00 eq.) was dissolved in dry MeCN (7.4 mL) and was added to the reaction vessel. The mixture was stirred at room temperature for 15 min and $SnCl_4$ (172 µL, 1.48 mmol, 1.25 eq.) was added. The reaction mixture was stirred at 50 °C for 24 h. The reaction was quenched with a saturated aqueous solution of $NaHCO_3$ and filtered over Celite®. The aqueous phase was extracted with EtOAc. The aqueous phase was concentrated *in vacuo* at 60 °C. The remaining solid was dissolved in a minimal amount water. The concentrated solution was desalinated and eluted with acetonitrile:water (4:1,20 mL). The eluate was concentrated *in vacuo*, suspended in water (5 mL), filtered and the filtrate purified via preparative HPLC.

***cis*-67d**:

^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 7.57 (d, $^3J_{6–5}$ = 7.4 Hz, 1 H, H_6), 6.17 (dd, $^3J_{1'–2'}$ = 8.1 Hz, $^3J_{1'–2'}$ = 6.1 Hz, 1 H, $H_{1'}$), 5.74 (d, $^3J_{5–6}$ = 7.4 Hz, 1 H, H_5), 5.06 (s, 1 H, $H_{5'–OH}$), 4.99 (s, 1 H, $H_{4'–OH}$), 4.05 (d, $^3J_{4'–4'}$ = 8.8 Hz, 1 H, $H_{4'}$), 3.63 (d, $^3J_{4'–4'}$ = 8.8 Hz, 1 H, $H_{4'}$, 3.39 (s, 1 H, $H_{5'}$, 2.11 (dd, $^2J_{2'–2'}$ = 13.4 Hz, $^3J_{2'–1'}$ = 6.1 Hz, 1 H, $H_{2'a}$), 1.92 (dd, $^2J_{2'–2'}$ = 13.4 Hz, $^3J_{2'–1'}$ = 8.1 Hz, 1 H, $H_{2'b}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 165.5 (C_4), 155.1 (C_2), 140.7 (C_5), 94.1 (C_6), 86.1 ($C_{1'}$), 80.5 ($C_{3'}$), 76.2 ($C_{4'}$), 64.3 ($C_{5'}$), 41.9 ($C_{2'}$).

***trans*-67d**:

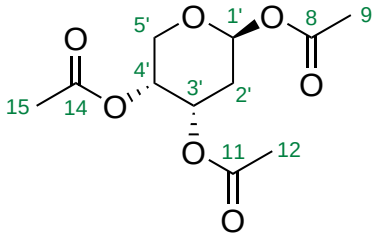
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 7.78 (d, $^{4}J_{6–5}$ = 7.4 Hz, 1 H, H_6), 5.98 (dd, $^{3}J_{1'–2'}$ = 7.8 Hz, $^{3}J_{1'–2'}$ = 3.0 Hz, 1 H, $H_{1'}$), 5.70 (d, $^{4}J_{5–6}$ = 7.4 Hz, 1 H, H_5), 5.01 (s, 2 H, $H_{5'/3'–OH}$), 3.87 (dd, $^{2}J_{4'–4'}$ = 9.0 Hz, $^{5}J_{4'–5'}$ = 1.3 Hz, 1 H, $H_{4'}$), 3.81 (dd, $^{2}J_{4'–4'}$ = 9.0 Hz, 1 H, $H_{4'}$), 3.36 (s, 2 H, $H_{5'}$, 2.44 (dd, $^{2}J_{2'–2'}$ = 14.1 Hz, $^{3}J_{2'–1'}$ = 7.8 Hz, 1 H, $H_{2'a}$), 1.74 (ddd, $^{2}J_{2'–2'}$ = 14.0 Hz, $^{3}J_{2'–1'}$ = 3.0 Hz, $^{5}J_{2'–4'}$ = 1.3 Hz, 1 H, $H_{2'b}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 165.6 (C_4), 155.2 (C_2), 141.8 (C_5), 93.2 (C_6) 86.4 ($C_{1'}$), 79.5 ($C_{3'}$), 76.6 ($C_{4'}$), 64.3 ($C_{5'}$), 41.6 ($C_{2'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_9H_{14}N_3O_4]^+$: 228.0979;s **obs.:** 228.0979.

1,3,4-Tri-*O*-acetyl-2-deoxy-*β*-d-ribopyranose (71a)

The reaction was performed according to Efange *et al.*.[260]



2-Deoxy-D-ribose **d4a** (10.00 mmol, 1.00 eq.) was dissolved in pyridine (4 mL) and the resulting solution was cooled to 0 °C. Acetic anhydride (35.00 mmol, 3.50 eq.) was added dropwise and the mixture was stirred over night. After the evaporation of the solvent the residue was quenched with saturated $NaHCO_3$-solution and extracted with DCM. The combined organic layers were washed with water in saturated NaCl-solution and dried over $MgSO_4$. The solvent was removed under reduced pressure and the residue was purified via column chromatography (SiO_2, R_f: 0.35, 0% → 1.6% Methanol in DCM). A viscous liquid, containing both the ribofuranose as well as the ribopyranose-derivatives, was obtained. Within one to a few days standing at ambient temperature colourless crystals precipitated. These were received by trituation of the residue with diisopropylether. 1,3,4-tri-*O*-acetyl-2-deoxy-d-ribopyranose was obtained as colourless crystals (0.442 g, 1.70 mmol, 17%). The remaining product was evaporated and a portion of it was purified via preparative HPLC. 1,3,5-tri-*O*-acetyl-2-deoxy-d-ribofuranose was obtained as a colourless viscous liquid and 1,3,4-tri-*O*-acetyl-2-deoxyd- ribopyranose was obtained as a colourless solid.

^{1}H-NMR (400 MHz, Chloroform-d_1): δ (ppm) = 6.27 (dd, $^3J_{1–2}$ = 3.4 Hz, $^3J_{1—2'}$ = 2.2 Hz, 1 H, H_1), 5.31 (ddd, $^3J_{3–2}$ = 11.9 Hz, $^3J_{3–2'}$ = 4.9 Hz, $^3J_{3–4}$ = 3.0 Hz, 1 H, H_3), 5.22 (dddd, $^3J_{4–3}$ = 3.0 Hz, $^3J_{4–5'}$ = 3.0 Hz, $^3J_{4–5}$ = 1.5 Hz, $^4J_{4–2'}$ = 1.5 Hz, 1 H, H_4), 4.02 (dd, $^3J_{5–5'}$ = 13.1 Hz, $^3J_{5–4}$ = 1.6 Hz, 1 H, H_5), 3.86 (dd, $^3J_{5–5'}$ = 13.1 Hz, $^3J_{5'–4}$ = 2.7 Hz, 1 H, $H_{5'}$), 2.26 (ddd, $^3J_{2–2'}$ = 13.3 Hz, $^3J_{2–3}$ = 12.0 Hz, $^3J_{2–1}$ = 3.5 Hz, 1 H, H_2), 2.14 (s, 3 H, H_{15}), 2.11 (s, 3 H, H_9), 2.03 (s, 3 H, H_{12}), 1.92 (dddd, $^3J_{2'–2}$ = 13.3 Hz, $^3J_{2'–3}$ = 5.0 Hz, $^3J_{2'–1}$ = 2.2 Hz, $^4J_{2'–4}$ = 1.2 Hz, 1 H, $H_{2'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 170.5 (C_{14}), 170.2 (C_{11}), 169.4 (C_8), 91.9 ($C_{1'}$), 67.3 ($C_{4'}$), 65.4 ($C_{3'}$), 63.2 ($C_{5'}$), 29.8 ($C_{2'}$) 21.2 (C_{15}), 21.1 (C_{12}), 21.0 (C_9).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{11}H_{17}O_7]^+$: 261.0969; **obs.:** 261.0968.

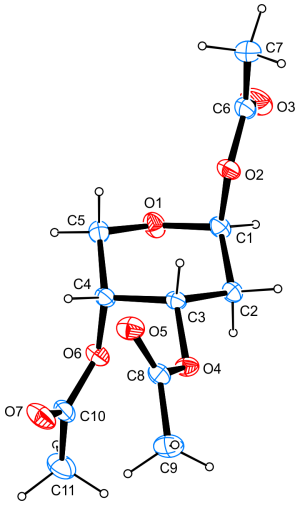
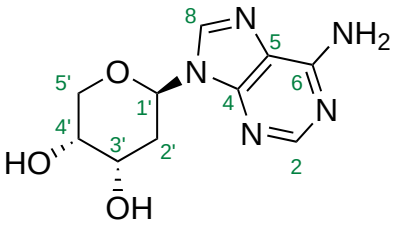


Figure E.1: Crystal structure of **71a**

net formula	$C_{11}H_{16}O_7$
M_r $[g/mol^1]$	260.24
crystal size [mm]	0.100 × 0.090 × 0.060
T[K]	103.(2)
radiation	Mo$K\alpha$
diffractometer	'Bruker D8 Venture TXS'
crystal system	orthorhombic
space group	'P 21 21 21'
a [Å]	7.1292(2)
b [Å]	8.0107(4)
c [Å]	22.2425(5)
α[°]	90
β[°]	90
γ[°]	90
V [Å^3]	1270.27(6)
Z	4
calc. density $[gcm^3]$	1.361
μ[gcm^3]	0.115
absorption correction	Multi-Scan
transmission factor range	0.96–0.99
refls. measured	26922
R_{int}	0.0369
mean $\sigma(I)/I$	0.0267
$\theta range$	3.000–33.142
observed refls.	4427
x, y (weighting scheme)	0.0421, 0.2751
hydrogen refinement	constr
Flack parameter	0.2(2)
refls in refinement	4834
parameters	166
restraints	0
R($F_o bs$)	0.0359
$R_w(F^2)$	0.0936
S	1.061
shift/$error_{max}$	0.001
max electron density [e Å^3]	0.281
min electron density[e Å^3]	0.208

Deoxyribopyranosyl adenosine (pyr-3a)

The reaction was performed according to Niedballa *et al.*[241]



71a (1.00 mmol, 1.00 eq.) and 6-chlorpurine (1.09 mmol, 1.09 eq.) were dissolved in acetonitrile. HMDS (1.10 mmol, 1.10 eq.), TMSCl (1.20 mmol, 1.20 eq.) and $SnCl_4$ (1.26 mmol, 1.26 eq.) were added and the resulting reaction mixture was stirred over night at 45 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.45, 0% → 1.6% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC, gave the β-product and the α-product were obtained as a white solid.

***β*-pyr-3a**:

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 8.33 (s, 1 H, H_8), 8.15 (s, 1 H, H_2), 7.28 (s, 2 H, H_{NH2}), 5.66 (dd, $^3J_{1'-2'}$ = 11.0 Hz,$^3J_{1'-2'}$ = 2.2 Hz, 1 H, $H_{1'}$), 4.96 (s, 2 H, $H_{3'-OH}$), 4.69 (s, 2 H, $H_{4'-OH}$), 3.91–3.81 (m, $^3J_{3'-2'}$ = 3.2 Hz, 2 H, $H_{3',5'}$), 3.69 (d, $^3J_{5'-5'}$ = 12.2 Hz, 1 H, $H_{5'}$), 3.64 (s, 1 H, $H_{4'}$, 2.51 (ddd, $^2J_{2'-2'}$ = 13.3 Hz, $^3J_{2'-1'}$ = 11.3 Hz, $^3J_{2'-3'}$ = 2.5 Hz, 1 H, $H_{a-2'}$), 1.93 (ddd, $^2J_{2'-2'}$ = 13.3 Hz, $^3J_{2'-1'}$ = 4.0 Hz, $^3J_{2'-3'}$ = 2.3 Hz, 1 H, $H_{e-2'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 156.5 (C_6), 153.2 (C_2), 149.5 (C_4), 139.2 (C_8), 118.9 (C_5), 79.4 ($C_{1'}$), 69.7 ($C_{4'}$), 68.0 ($C_{3'}$) 66.9 ($C_{5'}$), 34.0 ($C_{2'}$).

***α*-pyr-3a**:

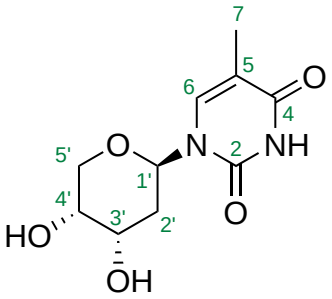
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 8.34 (s, 1 H, H_8), 8.15 (s, 1 H, H_2), 7.27 (s, 2 H, H_{NH2}), 5.95 (dd, $^{3}J_{1'-2'}$ = 11.0 Hz, $^{3}J_{1'-2'}$ = 2.2 Hz, 1 H, $H_{1'}$), 4.99 (s, 2 H, H_{OH}), 4.09 (d, $^{3}J_{3'-2'}$ = 3.2 Hz, 2 H, $H_{3',5'}$), 3.74 (q, $^{3}J_{5'-5'}$ = 8.8 Hz, $^{3}J_{5'-4'}$ = 8.1 Hz, 1 H, $H_{5'}$), 3.72–3.66 (m, 1 H, $H_{4'}$, 3.63 (dd, ^{2}J = 8.6 Hz, ^{2}J = 3.8 Hz, 1 H, $H_{5'}$), 2.63 (ddd, $^{2}J_{2'-2'}$ = 13.3 Hz, $^{3}J_{2'-1'}$ = 11.3 Hz, $^{3}J_{2'-3'}$ = 2.5 Hz, 1 H, $H_{a-2'}$), 2.02 (ddd, $^{2}J_{2'-2'}$ = 13.3 Hz, $^{3}J_{2'-1'}$ = 4.0 Hz, $^{3}J_{2'-3'}$ = 2.3 Hz, 1 H, $H_{e-2'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 159.5 (C_6), 153.2 (C_2), 149.7 (C_4), 139.6 (C_8), 119.2 (C_5), 76.8 ($C_{1'}$), 66.8 ($C_{4'}$), 66.6 ($C_{3'}$), 66.1 ($C_{5'}$), 36.6 ($C_{2'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{10}H_{14}N_5O_3]^+$: 252.1091; **obs.:** 252.1090.

Deoxyribopyranosyl thymidine (pyr-3b)

The reaction was performed according to Niedballa *et al.*[241]



A suspension of thymine **1e** (5.00 mmol, 1.00 eq.) in HMDS (40.00 mmol, 8.00 eq.), TMSCl (6.60 mmol, 1.32 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in acetonitrile (25 mL) and **71a** (5.00 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (1.26 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 45 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.3, 0% → 1.6% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC, gave the β-product and the α-product were obtained as a white solid.

***α*-pyr-3b**:

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 11.31 (s, 1 H, H_{NH}), 7.57 (q, $^4J_{6–7}$ = 1.2 Hz, 1 H, H_6), 5.51 (dd, $^3J_{1'–2'}$ = 11.3 Hz, $^3J_{1'–2'}$ = 2.3 Hz, 1 H, $H_{1'}$), 4.84 (s, 1 H, $H_{3'–OH}$), 4.61 (s, 1 H, $H_{4'–OH}$), 3.83 (dd, $^3J_{5'–5'}$ = 12.3 Hz, $^3J_{5'–4'}$ = 2.1 Hz, 1 H, $H_{5'}$), 3.77 (m, 1 H, $H_{3'}$, 3.60 (dd, $^2J_{5'–5'}$ = 12.3 Hz, $^2J_{5'–4'}$ = 1.3 Hz, 1 H, $H_{5'}$), 3.55 (s, 1 H, $H_{4'}$), 1.98 (q, $^2J_{2'–2'}$ = 11.6 Hz, 1 H, $H_{2'}$), 1.80 (d, $^4J_{7–6}$ = 1.2 Hz, 1 H, H_7), 5.51 (m, 1 H, $H_{2'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 164.1 (C_4), 150.5 (C_2), 136.7 (C_6), 109.9 (C_5), 79.7 ($C_{1'}$), 69.9 ($C_{5'}$), 67.8 ($C_{3'}$), 66.9 ($C_{4'}$), 33.5 ($C_{2'}$), 12.4 (C_7).

***β*-pyr-3b**:

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 11.31 (s, 1 H, H_{NH}), 7.58 (t, $^{4}J_{6-7}$ = 1.2 Hz, 1 H, H_6), 5.85 (dd, $^{3}J_{1'-2'}$ = 11.3 Hz, $^{3}J_{1'-2'}$ = 2.3 Hz, 1 H, $H_{1'}$), 4.90 (s, 1 H, $H_{3'-OH}$), 4.00 (s, 1 H, $H_{3'}$), 3.70–3.54 (m, 3 H, $H_{5',4'}$), 2.07 (ddd, $^{2}J_{2'-2'}$ = 13.6 Hz, $^{3}J_{2'-3'}$ = 11.3 Hz, $^{3}J_{2'-1'}$ = 2.5 Hz, 1 H, $H_{2'}$), 1.78 (d, $^{4}J_{7-6}$ = 1.2 Hz, 1 H, H_7), (m, 1 H, $H_{2'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 164.1 (C_4), 150.7 (C_2), 136.9 (C_6), 109.9 (C_5), 76.9 ($C_{1'}$), 66.8 ($C_{5'}$), 66.6 ($C_{3'}$), 66.1 ($C_{4'}$), 35.9 ($C_{2'}$), 12.4 (C_7).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{10}H_{15}N_2O_4]^+$: 243.0975; **obs.:** 243.0976.

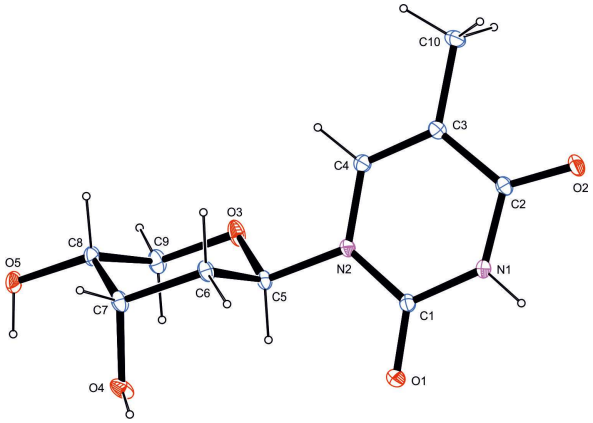
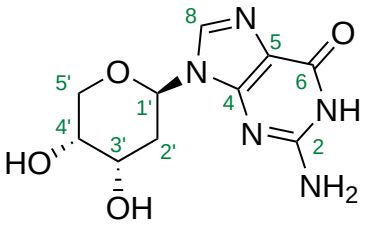


Figure E.2: Crystal structure of **pyr-3a**

net formula	$C_{10}H_{14}N_2O_5$
M_r $[g/mol^1]$	242.23
crystal size [mm]	0.100 × 0.090 × 0.030
T[K]	102.(2)
radiation	MoKα
diffractometer	'Bruker D8 Venture TXS'
crystal system	orthorhombic
space group	'P 21 21 21'
a [Å]	5.4250(2)
b [Å]	13.8445(4)
c [Å]	14.5229(5)
α[°]	90
β[°]	90
γ[°]	90
V [Å^3]	1090.76(6)
Z	4
calc. density $[gcm^3]$	1.475
μ[gcm^3]	0.119
absorption correction	Multi-Scan
transmission factor range	0.96–1.00
refls. measured	27766
R_{int}	0.0336
mean $\sigma(I)/I$	0.0194
$\theta range$	2.943–30.505
observed refls.	3204
x, y (weighting scheme)	0.0463, 0.1756
hydrogen refinement	mixed
Flack parameter	0.3(2)
refls in refinement	3322
parameters	167
restraints	0
R(F_Obs)	0.0290
$R_w(F^2)$	0.0809
S	1.094
shift/$error_{max}$	0.001
max electron density [eÅ^3]	0.294
min electron density[eÅ^3]	0.195

Deoxyribopyranosyl guanosine (pyr-3c)

The reaction was performed according to Niedballa *et al.*[241]



71a (1.00 mmol, 1.00 eq.) and *N*-(6-oxo-5,9-dihydro-*6H*-purin-2-yl)acetamide (1.09 mmol, 1.09 eq.) were dissolved in acetonitrile. HMDS (1.10 mmol, 1.10 eq.), TMSCl (1.20 mmol, 1.20 eq.) and $SnCl_4$ (1.26 mmol, 1.26 eq.) were added and the resulting reaction mixture was stirred over night at 45 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.21, 0% → 1.6% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC, gave the β-product and the α-product were obtained as a white solid.

***β*-pyr-3c**:

^{1}H-NMR (600 MHz, DMSO-*d*$_6$): δ (ppm) = 11.17 (s, 1 H, H_{NH}), 8.16 (s, 1 H, H_8), 6.45 (s, 2 H, H_{NH2}), 6.09 (dd, $^3J_{1'–2'}$ = 11.3 Hz,$^3J_{1'–2'}$ = 2.3 Hz, 1 H, $H_{1'}$), 4.99 (s, 1 H, $H_{3'–OH}$), 4.97 (s, 1 H, $H_{4'–OH}$), 4.04 (m, 1 H, $H_{3'}$), 3.71–3.65 (m, 1 H, $H_{5'}$), 3.65–3.60 (m, 1 H, $H_{4'}$), 3.57 (dd, $^2J_{5'–5'}$ = 9.9 Hz, $^3J_{5'–4'}$ = 4.8 Hz 1 H, $H_{5'}$), (m, 1 H, $H_{2'}$),), 2.00 (ddd, $^2J_{2'–2'}$ = 13.2 Hz, $^3J_{2'–1'}$ = 4.0 Hz, $^3J_{2'–3'}$ = 2.3 Hz, 1 H, $H_{2'}$).

^{13}C-NMR (151 MHz, DMSO-*d*$_6$): δ (ppm) =
159.9 (C_6), 153.4 (C_2), 149.7 (C_4), 141.2 (C_8), 107.5 (C_5), 78.3 ($C_{1'}$), 66.4 ($C_{4'}$), 66.3 ($C_{3'}$), 65.4 ($C_{5'}$), 36.5 ($C_{2'}$).

***α*-pyr-3c**:

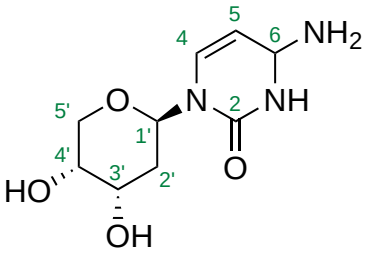
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 11.08 (s, 1 H, H_{NH}), 8.11 (s, 1 H, H_8), 6.26 (s, 2 H, H_{NH2}), 5.71 (dd, $^{3}J_{1'–2'}$ = 11.1 Hz,$^{3}J_{1'–2'}$ = 2.6 Hz, 1 H, $H_{1'}$), 4.86 (s, 1 H, $H_{3'–OH}$), 4.71 (s, 1 H, $H_{4'–OH}$), 3.80 (dd, $^{2}J_{5'–5'}$ = 12.2 Hz,$^{3}J_{5'–4'}$ = 2.3 Hz, 1 H, $H_{5'}$), 3.77–3.73 (m, 1 H, $H_{3'}$), 3.59 (d, $^{2}J_{5'–5'}$ = 12.3 Hz, 1 H, $H_{5'}$), 3.56 (s, 1 H, $H_{4'}$), 2.28 (q, $^{3}J_{2'–2'}$ = 11.6 Hz, 1 H, $H_{2'}$), 1.93 (ddd, $^{2}J_{2'–2'}$ = 12.1 Hz, $^{2}J_{2'–3'}$ = 4.7 Hz, $^{3}J_{2'–1'}$ = 2.7 Hz, 1 H, $H_{2'}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 159.9 (C_6), 153.3 (C_2), 149.7 (C_4), 141.2 (C_8), 107.5 (C_5), 81.2 ($C_{1'}$), 69.2 ($C_{4'}$), 67.4 ($C_{3'}$), 66.8 ($C_{5'}$), 34.4 ($C_{2'}$).

HR-ESIMS (*m/z* [M+H]$^{+}$): calc. for $[C_{10}H_{14}N_5O_4]^{+}$: 268.1040; **obs.:** 268.1040.

Deoxyribopyranosyl cytidine (pyr-3d)

The reaction was performed according to Niedballa *et al.*[241]



A suspension of cytosine **1d** (5.00 mmol, 1.00 eq.) in HMDS (40.00 mmol, 8.00 eq.), TMSCl (6.60 mmol, 1.32 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in acetonitrile (25 mL) and **71a** (5.00 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (1.26 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 45 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with brine, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.45, 0% → 1.6% Methanol in DCM). After protection in methanolic ammonia solution, purification *via* preparative HPLC, gave the β-product and the α-product were obtained as a white solid.

***β*-pyr-3d**:

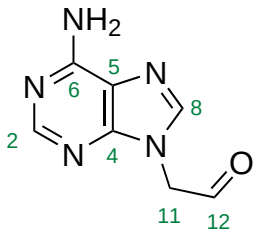
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 11.31 (s, 1 H, H_{NH}), 7.58 (d, $^4J_{6–5}$ = 7.4 Hz, 1 H, H_6), 5.91 (dd, $^3J_{1'–2'}$ = 11.3 Hz, $^3J_{1'–2'}$ = 2.4 Hz, 1 H, $H_{1'}$), 5.69 (d, $^4J_{5–6}$ = 7.4 Hz, 1 H, H_5), 4.83 (d, $^4J_{3'OH–3'}$ = 3.0 Hz, 1 H, $H_{3'–OH}$), 4.77 (d, $^4J_{4'OH–4'}$ = 4.7 Hz, 1 H, $H_{4'–OH}$), 3.97 (s, 1 H, $H_{3'}$), 3.68–3.54 (m, 3 H, $H_{4',5'}$, 1.86 (ddd, $^2J_{2'–2'}$ = 13.3 Hz, $^3J_{2'–3'}$ = 11.0 Hz, $^3J_{2'–1'}$ = 2.4 Hz, 1 H, $H_{a–2'}$), 1.74 (ddd, $^2J_{2'–2'}$ = 13.3 Hz, $^3J_{2'–3'}$ = 3.8 Hz, $^3J_{2'–1'}$ = 2.4 Hz, 1 H, $H_{e–2'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 165.4 (C_4), 154.5 (C_2), 141.3 (C_6), 94.0 (C_5), 77.1 ($C_{1'}$), 66.4 ($C_{5'}$), 66.3 ($C_{3'}$), 65.7 ($C_{4'}$), 36.3 ($C_{2'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_9H_{14}N_3O_4]^+$: 228.0979; **obs.:** 228.0980.

Adenine-*N*9-acetaldehyde (81a)

The reaction was performed according to *Palazzolo, et al.*[261] and *Solyev, et al.*.[262]



1a (2.00 mmol, 1.00 eq.) was suspended in DMF (2 mL). K_2CO_3 (4.00 mmol, 2.00 eq.) and 2-bromo-1,1-dimethoxyethane (2.20 mmol, 1.10 eq.) were added and the reaction mixture was stirred at 90 °C for 24 h. The reaction mixture was filtered over Celite®.The solvent was removed under reduced pressure and the resulting crude product was purified via column chromatography (SiO_2, R_f: 0.45, 0.6, 5% MeOH in DCM). The product was obtained as a colourless solid (0.13 g, 0.59 mmol, 30%). Subsequently, the product was suspended in aqueous HCl (1 M) and stirred at 90 °C for 16 h. The solvent was evaporated *in vacuo* and the desired product was obtained as a colourless solid (0.12 g, 0.72 mmol, 35%).

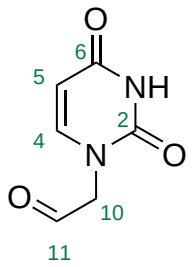
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 9.68 (s, 1 H, $H_{2'}$), 8.45 (s, 1 H, H_2), 8.41 (s, 1 H, H_8), 5.38 (s, 2 H, $H_{1'}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 196.2 ($C_{2'}$), 150.6 (C_6), 149.1 (C_2), 145.2 (C_8), 118.0 (C_5), 53.4 ($C_{1'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_7H_8N_5O]^+$: 178.0723; **obs.:** 178.0723.

Uracil-*N*1-acetaldehyde (81b)

The reaction was performed according to *Palazollo, et al.*[261] and *Solyev, et al.*.[262]



1b (2.00 mmol, 1.00 eq.) was suspended in DMF (2 mL). K_2CO_3 (4.00 mmol, 2.00 eq.) and 2-bromo-1,1-dimethoxyethane (2.20 mmol, 1.10 eq.) were added and the reaction mixture was stirred at 90 °C for 24 h. The reaction mixture was filtered over Celite®.The solvent was removed under reduced pressure and the resulting crude product was purified via column chromatography (SiO_2, R_f: 0.48, 2% MeOH in DCM). The product was obtained as a colourless solid (0.15 g, 0.76 mmol, 38%). Subsequently, the product was suspended in aqueous HCl (1 M) and stirred at 90 °C for 16 h. The solvent was evaporated *in vacuo* and the desired product was obtained as a colourless solid (0.13 g, 0.84 mmol, 42%).

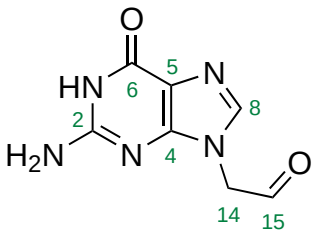
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 11.37 (s, 1 H, H_{NH2}), 9.53 (s, 1 H, $H_{2'}$), 7.49 (dd, ^3J = 7.8 Hz 1 H, H_6), 5.59 (dd, ^3J = 7.8, 2.3 Hz 1 H, H_5), 4.61 (s, 2 H, $H_{1'}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 197.4 ($C_{2'}$), 163.7 (C_4), 151.0 (C_2), 145.9 (C_6), 101.0 (C_5), 56.6 ($C_{1'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_6H_7N_2O_3]^+$: 155.0451; **obs.:** 155.0450.

Guanine-*N*9-acetaldehyde (81c)

The reaction was performed according to *Palazollo, et al.*.[261]



2-acetamido-*9H*-purin-6-yl diphenylcarbamate (2.00 mmol, 1.00 eq.) was stirred at 120 °C in DMF (2 mL) until everything was dissolved. K_2CO_3 (4.00 mmol, 2.00 eq.) and 2-bromo-1,1-dimethoxyethane (2.20 mmol, 1.10 eq.) were added and the reaction mixture was stirred at for 24 h. The reaction mixture was filtered over Celite®.The solvent was removed under reduced pressure and the resulting crude product was purified via column chromatography (SiO_2, R_f: 0.6, 0%–5% MeOH in DCM). Subsequently, the product was suspended in aqueous HCl (1 M) and stirred at 90 °C for 16 h. The solvent was evaporated *in vacuo* and the desired product was obtained as a colourless solid (0.13 g, 0.66 mmol, 33%).

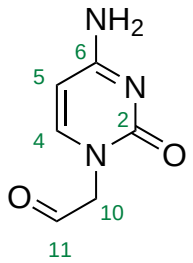
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 9.67 (s, 1 H, $H_{2'}$), 9.09 (s, 2 H, H_{NH2}), 8.87 (s, 1 H, H_8), 5.19 (s, 2 H, $H_{1'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 195.5 ($C_{2'}$), 155.7 (C_6), 153.3 (C_2), 149.8 (C_4), 137.8 (C_8), 107.2 (C_5), 53.2 ($C_{1'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_7H_8N_5O_2]^+$: 194.0673; **obs.:** 194.0673.

Cytosine-*N*1-acetaldehyde (81d)

The reaction was performed according to *Palazollo, et al.*.[261]



1d (2.00 mmol, 1.00 eq.) was suspended in DMF (2 mL). K_2CO_3 (4.00 mmol, 2.00 eq.) and 2-bromo-1,1-dimethoxyethane (2.20 mmol, 1.10 eq.) were added and the reaction mixture was stirred at 90 °C for 24 h. The reaction mixture was filtered over Celite®.The solvent was removed under reduced pressure and the resulting crude product was purified via column chromatography (SiO_2, R_f: 0.48, 5% MeOH in DCM). The product was obtained as a colourless solid (g, mmol, %). Subsequently, the product was suspended in aqueous HCl (1 M) and stirred at 90 °C for 16 h. The solvent was evaporated *in vacuo* and the desired product was obtained as a colourless solid (0.07 g, 0.48 mmol, 24%).

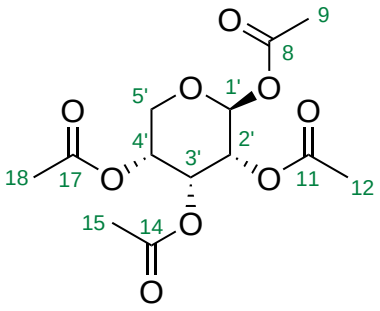
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 9.57 (s, 1 H, $H_{2'}$), 7.92 (d, ^3J = 7.6 Hz, 1 H, H_6), 6.21 (d, ^3J = 7.6 Hz, 1 H, H_5), 4.80 (s, 2 H, $H_{1'}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 196.9 ($C_{2'}$), 160.4 (C_4), 150.5 (C_6), 147.9 (C_2), 93.8 (C_5), 57.9 ($C_{1'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_6H_8N_3O_2]^+$: 154.0611; **obs.:** 154.0613.

(*3R,4R,5R*)-Tetrahydro-2H-pyran-2,3,4,5-tetrayl tetraacetate (82a)

The reaction was performed according to Efange, *et al.*.[260]



D-ribose **4a** (10.00 mmol, 1.00 eq.) was dissolved in pyridine (4 mL) and the resulting solution was cooled to 0 °C. Acetic anhydride (45.00 mmol, 4.50 eq.) was added dropwise and the mixture was stirred over night. After the evaporation of the solvent the residue was quenched with saturated $NaHCO_3$-solution and extracted with DCM. The combined organic layers were washed with water in saturated NaCl-solution and dried over $MgSO_4$. The solvent was removed under reduced pressure and the residue was purified via column chromatography (SiO_2, R_f: 0.9, 100% DCM). A viscous liquid, containing both the ribofuranose as well as the ribopyranose-derivatives, was obtained. Within one to a few days standing at ambient temperature colourless crystals precipitated. These were received by trituation of the residue with diisopropylether. (3R,4R,5R)-tetrahydro-2H-pyran-2,3,4,5-tetrayl tetraacetate was obtained as colourless crystals (1.16 g, 3.64 mmol, 35%).

^{1}H-NMR (400 MHz, Chloroform-d_1): δ (ppm) = 6.03 (d, $^3J_{1'-2'}$ = 4.8 Hz, 1 H, $H_{1'}$), 5.49 (t, $^3J_{3'-2',4'}$ = 3.4 Hz, 1 H, $H_{3'}$), 5.15 (dtd, $^3J_{4'-a-5'}$ = 5.8 Hz, $^3J_{4'-3'}$ = 3.4 Hz, $^3J_{4'-e-5'}$ = 0.8 Hz, 1 H, $H_{4'}$), 5.08 (ddd, $^3J_{2'-1'}$ = 4.5 Hz, $^3J_{2'-3'}$ = 3.5 Hz, 1 H, $H_{2'}$), 4.02 (dd, $^2J_{e-5'-a-5'}$ = 12.3 Hz, $^3J_{e-5'-4'}$ = 3.5 Hz, 1 H, $H_{e-5'}$), 3.91 (dd, $^2J_{a-5'-e-5'}$ = 12.3 Hz, $^3J_{a-5'-4'}$ = 5.8 Hz, 3 H, $H_{a-5'}$), 2.13 (s, 3 H, H_{Me}), 2.10 (s, 3 H, H_{Me}), 2.10 (s, 3 H, H_{Me}), 2.09 (s, 3 H, H_{Me}).

^{13}C-NMR (101 MHz, Chloroform-d_1): δ (ppm) = 170.6 (C_{14}), 169.9 (C_{17}), 169.6 (C_{11}), 168.9 (C_8), 91.1 ($C_{1'}$), 67.5 ($C_{2'}$), 66.3 ($C_{3'}$) 66.3 ($C_{4'}$), 62.8 ($C_{5'}$), 21.0 (C_{Me}), 20.9 (C_{Me}), 20.8 (C_{Me}), 20.8 (C_{Me}).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{13}H_{18}O_9]^+$: 318.0950; **obs.:** 318.0951.

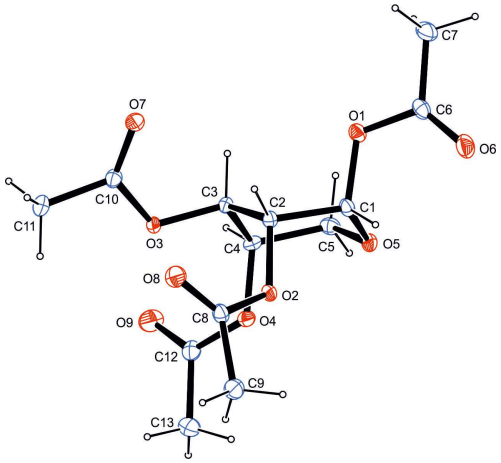
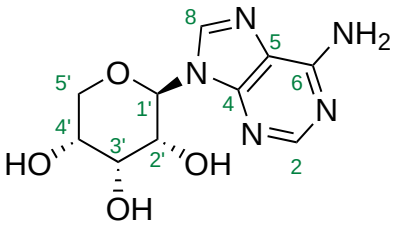


Figure E.3: Crystal structure of **82a**

net formula	$C_{13}H_{14}O_9$
M_r $[g/mol^1]$	318.27
crystal size [mm]	0.100 × 0.090 × 0.080
T[K]	102.(2)
radiation	Mo$K\alpha$
diffractometer	'Bruker D8 Venture TXS'
crystal system	tetragonal
space group	'P 41'
a [Å]	7.8016(5)
b [Å]	7.8016(5)
c [Å]	49.801(3)
α[°]	90
β[°]	90
γ[°]	90
V [Å^3]	3031.1(4)
Z	8
calc. density $[gcm^3]$	1.395
μ[gcm^3]	0.120
absorption correction	Multi-Scan
transmission factor range	0.88–0.99
refls. measured	18171
R_{int}	0.0391
mean $\sigma(I)/I$	0.0439
$\theta range$	2.885–27.101
observed refls.	6472
x, y (weighting scheme)	0.0250, 2.4456
hydrogen refinement	constr
Flack parameter	0.5(3)
refls in refinement	6625
parameters	405
restraints	1
R($F_o bs$)	0.0445
$R_w(F^2)$	0.1100
S	1.130
shift/$error_{max}$	0.001
max electron density [e Å^3]	0.209
min electron density[e Å^3]	0.211

Ribopyranosyl adenosine (pyr-2a)

The reaction was performed according to Niedballa *et al.*[241]



6-chloro purine (3.27 mmol, 1.09 eq.) was suspended in acetonitrile (25.00 mL). HMDS (3.30 mmol, 1.10 eq.), TMSCl (3.60 mmol, 1.20 eq.) was added. D-ribopyranose tetraacetate **82a** (1.00 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (3.78 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 50 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.38, 0% → 3.2% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC yielded the β-product and the α-product were obtained as a white solid.

***α*-pyr-2a**:

^{1}H-NMR (800 MHz, DMSO-*d*$_6$): δ (ppm) = 8.92 (s, 1 H, H_8), 8.13 (s, 1 H, H_2), 7.23 (s, 1 H, H_{NH2}), 5.62 (d, $^3J_{1'-2'}$ = 9.4 Hz, 1 H, $H_{1'}$), 5.15 (s, 1 H, $H_{2'-OH}$), 5.08 (s, 1 H, $H_{3'-OH}$), 4.90 (s, 1 H, $H_{5'-OH}$), 4.25–4.20 (m, 1 H, $H_{2'}$), 4.04 (t, $^3J_{3'-4'/2'}$ = 2.6 Hz, 1 H, $H_{3'}$), 3.72 (m, 1 H, $H_{4'}$), 3.67 (t, $^3J_{5'-5'}$ = 10.2 Hz, 1 H, $H_{5'}$), 3.58 (dd, $^2J_{5'-5'}$ = 10.2 Hz, $^3J_{5'-4'}$ = 5.2 Hz, 1 H, $H_{5'}$).

^{13}C-NMR (201 MHz, DMSO-*d*$_6$): δ (ppm) = 155.9 (C_4), 152.5 (C_2), 150.0 (C_6), 139.7 (C_8), 118.7 (C_5), 79.6 ($C_{1'}$), 71.2 ($C_{3'}$), 68.2 ($C_{2'}$) 66.7 ($C_{4'}$), 65.1 ($C_{5'}$).

***β*-pyr-2a**:

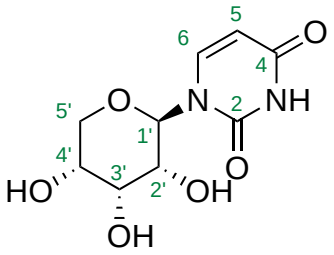
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 9.05 (s, 1 H, H_8), 8.86 (s, 1 H, H_2), 6.03 (d, $^{3}J_{1'-2'}$ = 9.2 Hz, 1 H, $H_{1'}$), 5.35 (s, 1 H, $H_{2'-OH}$), 5.32 (s, 1 H, $H_{3'-OH}$), 4.97 (s, 1 H, $H_{5'-OH}$), 4.18 (dd, $^{3}J_{2'-1'}$ = 9.2 Hz, $^{3}J_{2'-3'}$ = 2.7 Hz, 1 H, $H_{2'}$), 4.08 (s, 1 H, $H_{3'}$), 3.76 (d, ^{3}J = 5.2 Hz, 2 H, $H_{4'/5'}$), 3.71–3.63 (m, 1 H, $H_{5'}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 161.79 (C_4), 151.9 (C_2), 149.2 (C_8), 142.3 (C_5), 122.3 (C_6), 81.8 ($C_{1'}$), 71.1 ($C_{3'}$), 68.5 ($C_{2'}$) 66.4 ($C_{4'}$), 65.1 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{10}H_{14}N_5O_4]^+$: 268.1040; **obs.:** 268.1042.

Ribopyranosyl uridine (pyr-2b)

The reaction was performed according to Niedballa *et al.*[241]



A suspension of uracil **1b** (2.50 mmol, 1.00 eq.) in HMDS (20.00 mmol, 4.00 eq.), TMSCl (3.30 mmol, 1.32 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in acetonitrile (11.90 mL) and D-ribopyranose tetraacetate **82a** (2.50 mmol, 1.05 eq.) was added. After the addition of $SnCl_4$ (3.00 mmol, 1.26 eq.), the resulting reaction mixture was stirred over night at 50 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.45, 0% → 1.6% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC yielded the β-product and the α-product were obtained as a white solid.

***β*-pyr-2b**:

^{1}H-NMR (400 MHz, DMSO-*d*$_6$): δ (ppm) = 7.66 (d, $^3J_{6-5}$ = 8.1 Hz, 1 H, H_6), 5.60 (, $^3J_{5-6}$ = 8.5 Hz, 2 H, $H_{1',5}$), 5.12–5.03 (m, 1 H, $H_{3',4'-OH}$), 4.82 (s, 1 H, $H_{2'-OH}$), 3.97 (m, 1 H, $H_{3'}$), 3.69 (d, $^3J_{2'-1'}$ = 9.4 Hz, 3 H, $H_{2'}$), 3.68–3.59 (m, 1 H, $H_{4'}$), 3.62–3.49 (m, 2 H, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-*d*$_6$): δ (ppm) = 163.5 (C_4), 151.6 (C_2), 141.9 (C_6), 102.2 (C_5), 80.0 ($C_{1'}$), 71.7 ($C_{3'}$), 68.0 ($C_{2'}$) 66.9 ($C_{4'}$), 65.7 ($C_{5'}$).

***α*-pyr-2b**:

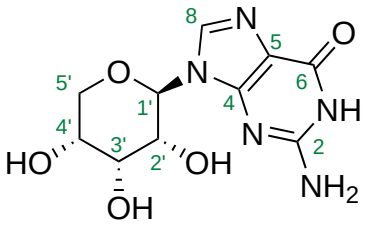
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 7.54 (d, $^{3}J_{6–5}$ = 7.5 Hz, 1 H, H_6), 7.15–7.08 (d, $^{3}J_{6–5}$ = 7.5 Hz, 1 H, H_6), 5.74–5.65 (m, 2 H, $H_{1',5}$), 4.98 (d, $^{3}J_{3'OH–3'}$ = 3.5 Hz, 1 H, $H_{3'–OH}$), 4.82 (d, $^{3}J_{2'OH–2'}$ = 7.4 Hz, 1 H, $H_{2'–OH}$), 4.77 (d, $^{3}J_{4'OH–4'}$ = 6.4 Hz, 1 H, $H_{4'–OH}$), 3.96 (q, $^{3}J_{3'–2'/4'}$ = 2.9 Hz, 1 H, $H_{3'}$), 3.62 (dt, $^{3}J_{3'–2'/4'}$ = 2.9 Hz, $^{3}J_{3'–2'/4'}$ = 2.9 Hz, 1 H, $H_{2'}$), 3.58 (dd, $^{3}J_{4'–3'/5'}$ = 5.4 Hz, $^{3}J_{4'–5'}$ = 3.1 Hz, 3 H, $H_{4'}$), 3.57–3.47 (m, 2 H, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 165.4 (C_4), 155.7 (C_2), 141.9 (C_6), 93.9 (C_5), 79.8 ($C_{1'}$), 71.5 ($C_{3'}$), 68.0 ($C_{2'}$) 66.7 ($C_{4'}$), 65.2 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_9H_{13}N_2O_6]^+$: 245.0768; **obs.:** 245.0769.

Ribopyranosyl guanosine (pyr-2c)

The reaction was performed according to Niedballa *et al.*[241]



N-acetyl guanine (0.98 mmol, 1.00 eq.) was suspended in acetonitrile (5.63 mL). HMDS (0.99 mmol, 1.10 eq.), TMSCl (1.08 mmol, 1.20 eq.) was added. D-ribopyranose tetraacetate **82a** (0.90 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (1.13 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 50 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.18, 0% → 3.2% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC yielded the β-product and the α-product were obtained as a white solid.

***α*-pyr-2c**:

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 7.81 (s, 1 H, H_8), 6.62 (s, 2 H, H_{NH2}), 5.45 (d, $^3J_{1'-2'}$ = 9.1 Hz, 1 H, $H_{1'}$), 5.16 (s, 1 H, $H_{3'-OH}$), 5.09 (s, 1 H, $H_{2'-OH}$), 4.90 (s, 1 H, $H_{4'-OH}$), 3.72–3.64 (m, 1 H, $H_{4'}$), 3.59 (t, $^3J_{5'-5'}$ = 10.2 Hz, 1 H, $H_{5'}$), 3.54 (m, 1 H, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 156.9 (C_6), 153.8 (C_2), 151.9 (C_4), 135.6 (C_8), 116.3 (C_5), 78.8 ($C_{1'}$), 71.2 ($C_{2'}$), 68.3 ($C_{3'}$) 66.6 ($C_{4'}$), 65.1 ($C_{5'}$).

***β*-pyr-2c**:

^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 8.13 (s, 1 H, H_8), 6.17 (s, 2 H, H_{NH2}), 5.73 (d, $^3J_{1'-2'}$ = 9.3 Hz, 1 H, $H_{1'}$), 5.10 (s, 1 H, $H_{3'-OH}$), 4.99 (d, $^3J_{2'OH-2'}$ = 7.4 Hz, 1 H, $H_{2'-OH}$),

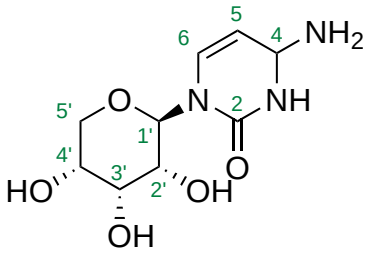
4.90 (s, 1 H, $H_{4'-OH}$), 4.14 (d, $^3J_{2'-2'}$ = 8.6 Hz, 1 H, $H_{2'}$), 3.99 (d, $^3J_{5'-5'}$ = 11.5 Hz, 1 H, $H_{3'}$), 3.67 (d, $^3J_{4'-5'}$ = 6.2 Hz, 1 H, $H_{4'}$), 3.65–3.58 (m, 1 H, $H_{5'}$), 3.58–3.52 (m, 1 H, $H_{5'}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 160.1 (C_6), 154.2 (C_2), 152.8 (C_4), 142.3 (C_8), 108.1 (C_5), 81.9 ($C_{1'}$), 71.3 ($C_{2'}$), 68.6 ($C_{3'}$) 66.6 ($C_{4'}$), 64.9 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{10}H_{14}N_5O_5]^+$: 284.0989; **obs.:** 284.0990.

Ribopyranosyl cytidine (pyr-2d)

The reaction was performed according to Niedballa *et al.*[241]



-0.5cm A suspension of cytosine **1d** (2.50 mmol, 1.00 eq.) in HMDS (20.00 mmol, 4.00 eq.), TMSCl (3.30 mmol, 1.32 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in acetonitrile (11.90 mL) and D-ribopyranose tetraacetate **82a** (2.50 mmol, 1.05 eq.) was added. After the addition of $SnCl_4$ (3.00 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 50 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.5, 0% → 1.6% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC yielded the β-product and the α-product were obtained as a white solid.

***β*-pyr-2d**:

^{1}H-NMR (400 MHz, DMSO-*d*$_6$): δ (ppm) = 7.55 (d, $^3J_{6-5}$ = 7.5 Hz, 1 H, H_6), 5.75–5.66 (m, 2 H, $H_{6,1'}$), 4.99 (d, $^3J_{3'-OH-3'}$ = 3.5 Hz, 1 H, $H_{3'-OH}$), 4.82 (d, $^3J_{2'-OH-2'}$ = 7.4 Hz, 1 H, $H_{2'-OH}$), 4.78 (d, $^3J_{4'-OH-4'}$ = 6.4 Hz, 1 H, $H_{4'-OH}$), 3.97 (q, $^3J_{3'-2',4'}$ = 2.9 Hz, 1 H, $H_{3'}$), 3.66–3.46 (m, 4 H, $H_{2',4',5'}$).

^{13}C-NMR (101 MHz, DMSO-*d*$_6$): δ (ppm) = 165.4 (C_4), 155.7 (C_2), 144.9 (C_6), 93.9 (C_5), 78.8 ($C_{1'}$), 71.5 ($C_{3'}$), 68.0 ($C_{2'}$) 66.7 ($C_{4'}$), 65.2 ($C_{5'}$).

***α*-pyr-2d**:

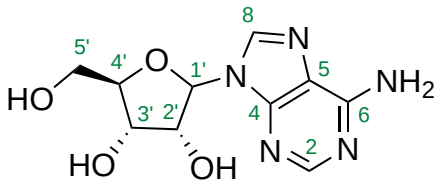
^{1}H-NMR (600 MHz, DMSO-d_6): δ (ppm) = 7.86 (d, $^{3}J_{6–5}$ = 7.4 Hz, 1 H, H_6), 7.27 (s, 1 H, H_{NH2}), 7.15 (s, 1 H, H_{NH2}), 5.76 (d, $^{3}J_{1'–2'}$ = 3.8 Hz, 1 H, $H_{1'}$), 5.76 (d, $^{3}J_{5–6}$ = 7.5 Hz, 1 H, H_5), 5.28 (d, $^{3}J_{2'–OH–2'}$ = 4.9 Hz, 1 H, $H_{2'–OH}$), 5.04 (t, $^{3}J_{4'–OH–4'}$ = 5.2 Hz, 1 H, $H_{4'–OH}$), 4.98 (d, $^{3}J_{3'–OH–3'}$ = 5.0 Hz, 1 H, $H_{3'–OH}$), 3.93 (h, $^{3}J_{3'–2',4'}$ = 4.9 Hz, 1 H, $H_{2'/3'}$), 3.81 (q, $^{3}J_{4'–5'/3'}$ = 3.6 Hz, 1 H, $H_{4'}$), 3.65 (ddd, $^{2}J_{5'–5'}$ = 12.1 Hz, $^{3}J_{5'–4'}$ = 5.1 Hz, ^{3}J = 3.1 Hz,1 H, $H_{5'}$), 3.54 (ddd, $^{2}J_{5'–5'}$ = 11.7 Hz, $^{3}J_{5'–4'}$ = 5.2 Hz, ^{3}J = 3.5 Hz,1 H, $H_{5'}$).

^{13}C-NMR (151 MHz, DMSO-d_6): δ (ppm) = 165.3 (C_4), 155.0 (C_2), 141.6 (C_6), 93.8 (C_5), 89.2 ($C_{1'}$), 84.1 ($C_{3'}$), 73.9 ($C_{2'}$), 69.4 ($C_{4'}$), 60.6 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_9H_{16}N_3O_5]^+$: 244.0928; **obs.:** 244.0926.

Ribofuranosyl adenosine (fur-2a)

The reaction was performed according to Niedballa *et al.*[241]



6-chloropurine (2.18 mmol, 1.09 eq.) was suspended in acetonitrile (12.5 mL). HMDS (2.20 mmol, 1.10 eq.), TMSCl (2.40 mmol, 1.20 eq.) was added. D-ribofuranose tetraacetate **82b** (2.00 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (2.52 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 50 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.48, 0% → 3.2% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC yielded the β-product and the α-product as white solid.

***β*-fur-2a**:

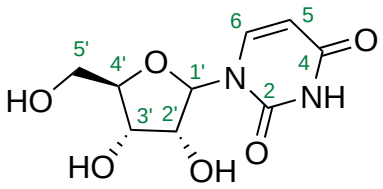
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 8.34 (s, 1 H, H_8), 8.13 (s, 1 H, H_2), 7.34 (s, 2 H, H_{NH2}), 5.87 (d, $^3J_{1'-2'}$ = 6.2 Hz, 1 H, $H_{1'}$), 4.61 (dd, $^3J_{2'-1'}$ = 6.2 Hz, $^3J_{2'-3'}$ = 5.0 Hz, 1 H, $H_{2'}$), 4.14 (dd, $^3J_{3'-2'}$ = 5.0 Hz, $^3J_{3'-4'}$ = 3.0 Hz, 1 H, $H_{3'}$), 3.96 (q, $^3J_{4'-3'/5'}$ = 3.4 Hz, 1 H, $H_{4'}$), 3.67 (dd, $^2J_{5'-5'}$ = 12.1 Hz, $^3J_{5'-4'}$ = 3.6 Hz, 1 H, $H_{5'}$), 3.52 (dd, $^2J_{5'-5'}$ = 12.1 Hz, $^3J_{5'-4'}$ = 3.6 Hz, 1 H, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 156.2 (C_4), 152.4 (C_2), 149.1 (C_5), 139.9 (C_8), 119.4 (C_6), 87.9 ($C_{1'}$), 85.9 ($C_{2'}$), 73.4 ($C_{3'}$), 70.7 ($C_{4'}$), 61.7 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{10}H_{14}N_5O_4]^+$: 268.1040; **obs.:** 268.1041.

Ribofuranosyl uridine (fur-2b)

The reaction was performed according to Niedballa *et al.*[241]



A suspension of uracil **1b** (2.50 mmol, 1.00 eq.) in HMDS (20.00 mmol, 8.00 eq.), TMSCl (3.30 mmol, 1.32 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in acetonitrile (12.5 mL) and D-ribofuranose tetraacetate **82b** (2.50 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (3.00 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 45 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with brine, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.45, 0% → 1.6% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC, the β-product and the α-product were obtained as a white solid.

***α*-fur-2b**:

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 7.88 (d, $^3J_{6–5}$ = 8.1 Hz, 1 H, H_6), 5.78 (d, $^3J_{1'–2'}$ = 5.4 Hz, 1 H, $H_{1'}$), 5.64 (dd, $^3J_{5–6}$ = 8.1 Hz, 1 H, H_5), 5.36 (s, 1 H, $H_{2'–OH}$), 5.07 (t, $^3J_{5'–OH–5'}$ = 5.2 Hz, 2 H, $H_{3'/5'–OH}$), 4.02 (s, 1 H, $H_{2'}$), 3.96 (t, $^3J_{3'–4'/2'}$ = 4.7 Hz, 1 H, $H_{3'}$), 3.84 (q, $^3J_{4'–5'/3'}$ = 3.4 Hz, 1 H, $H_{4'}$), 3.62 (d, $^2J_{5'–5'}$ = 12.2 Hz, 1 H, $H_{5'}$), 3.54 (d, $^2J_{5'–5'}$ = 12.1 Hz, 1 H, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 163.1 (C_4), 150.7 (C_2), 140.7 (C_6), 101.7 (C_5), 87.6 ($C_{1'}$), 84.8 ($C_{4'}$), 73.5 ($C_{2'}$), 69.8 ($C_{3'}$), 60.8 ($C_{5'}$).

***β*-fur-2b**:

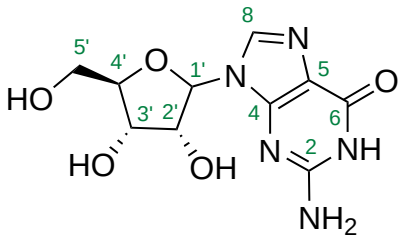
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 7.98 (d, $^{3}J_{6–5}$ = 8.2 Hz, 1 H, H_6), 6.09 (d, $^{3}J_{1'–2'}$ = 3.4 Hz, 1 H, $H_{1'}$), 5.77 (dd, $^{3}J_{5–6}$ = 6.5 Hz, 1 H, H_5), 5.07 (t, $^{3}J_{5'–OH–5'}$ = 5.2 Hz, 2 H, $H_{3'/5'–OH}$), 4.45 (dd, $^{3}J_{2'–3'}$ = 6.2 Hz, $^{3}J_{2'–1'}$ = 3.4 Hz, 1 H, $H_{2'}$), 4.09 (t, $^{3}J_{3'–4'/2'}$ = 6.5 Hz, 1 H, $H_{3'}$), 3.70–3.62 (m, 1 H, $H_{4'}$), 3.60 (dd, $^{2}J_{5'–5'}$ = 11.8 Hz, $^{3}J_{5'–4'}$ = 3.8 Hz, 1 H, $H_{5'}$), 3.42 (dd, $^{2}J_{5'–5'}$ = 11.8 Hz, $^{3}J_{5'–4'}$ = 5.8 Hz, 1 H, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 161.9 (C_4), 150.1 (C_2), 139.8 (C_6), 101.2 (C_5), 88.7 ($C_{1'}$), 84.8 ($C_{4'}$), 70.9 ($C_{2'}$), 70.0 ($C_{3'}$), 62.3 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_9H_{13}N_2O_6]^+$: 245.0768; **obs.:** 245.0766.

Ribofuranosyl guanosine (fur-2c)

The reaction was performed according to Niedballa *et al.*[241]



N-acetyl guanine (2.18 mmol, 1.09 eq.) was suspended in acetonitrile (12.5 mL). HMDS (2.20 mmol, 1.10 eq.) and TMSCl (2.40 mmol, 1.20 eq.) were added. D-ribofuranose tetraacetate **82b** (2.00 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (2.52 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 50 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with saturated NaCl-solution, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.2, 0% → 5% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC yielded the β-product and the α-product as white solid.

***β*-fur-2c**:

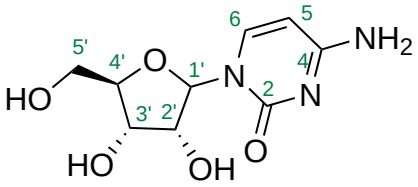
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 8.27 (s, 1 H, H_8), 6.30 (s, 2 H, H_{NH2}), 5.98 (d, $^3J_{1'-2'}$ = 5.7 Hz, 1 H, $H_{1'}$), 5.40 (s, 1 H, $H_{2'-OH}$), 5.11 (s, 1 H, $H_{3'-OH}$), 5.03 (s, 1 H, $H_{5'-OH}$), 4.36 (t, $^3J_{2'-1'/3'}$ = 5.5 Hz, $H_{2'}$), 4.07 (t, $^3J_{3'-2'/4'}$ = 4.5 Hz, $H_{3'}$), 3.88 (q, $^3J_{4'-3'/5'}$ = 3.8 Hz, 1 H, $H_{4'}$), 3.65 (d, $^2J_{5'-5'}$ = 11.9 Hz, 1 H, $H_{5'}$), 3.52 (dd, $^2J_{5'-5'}$ = 12.0 Hz, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 160.7 (C_6), 154.9 (C_2), 153.3 (C_5), 142.2 (C_8), 107.7 (C_4), 89.1 ($C_{1'}$), 85.2 ($C_{4'}$), 74.4 ($C_{2'}$), 69.7 ($C_{3'}$), 61.2 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_{10}H_{14}N_5O_5]^+$: 284.0989; **obs.:** 284.0990.

Ribofuranosyl cytidine (fur-2d)

The reaction was performed according to Niedballa *et al.*[241]



A suspension of cytosine **1d** (2.50 mmol, 1.00 eq.) in HMDS (20.00 mmol, 8.00 eq.), TMSCl (3.30 mmol, 1.32 eq.) and a small amount of NH_4SO_4 (cat.) was stirred at 150 °C, until everything was dissolved. The solvent was removed *in vacuo*. The resulting, protected pyrimidine base was suspended in acetonitrile (12.5 mL) and D-ribofuranose tetraacetate **82b** (2.50 mmol, 1.00 eq.) was added. After the addition of $SnCl_4$ (3.00 mmol, 1.26 eq.) the resulting reaction mixture was stirred over night at 45 °C. The reaction mixture was quenched with saturated $NaHCO_3$-solution and filtered over Celite®. The filtrate was extracted three times with EtOAc. The combined organic layers were washed with brine, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified via column chromatography (SiO_2, R_f: 0.35, 0% → 1.6% Methanol in DCM). After deprotection in methanolic ammonia solution, purification *via* preparative HPLC, the β-product and the α-product were obtained as a white solid.

***β*-fur-2d**:

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 7.84 (d, $^3J_{6–5}$ = 7.4 Hz, 1 H, H_6), 7.15–7.07 (m, 2 H, H_{NH2}), 5.76 (d, $^3J_{1'–2'}$ = 3.6 Hz, 1 H, $H_{1'}$), 5.70 (d, $^3J_{5–6}$ = 7.4 Hz, 1 H, H_5), 5.28 (s, 1 H, $H_{2'/3'–OH}$), 5.28 (t, $^3J_{5'OH–5'}$ = 5.2 Hz, 1 H, $H_{5'–OH}$), 4.98 (s, 1 H, $H_{4'–OH}$), 3.92 (s, 2 H, $H_{2'/3'}$), 3.81 (dt, ^3J = 4.7 Hz, ^3J = 3.2 Hz, 1 H, $H_{4'}$), 3.65 (ddd, $^2J_{5'–5'}$ = 12.1 Hz, $^3J_{5'–4'}$ = 5.1 Hz, $^3J_{5'–5'–OH}$ = 3.1 Hz, 1 H, $H_{5'}$), 3.63 (ddd, $^2J_{5'–5'}$ = 12.1 Hz, $^3J_{5'–4'}$ = 5.1 Hz, $^3J_{5'–5'–OH}$ = 3.1 Hz, 1 H, $H_{5'}$).

^{13}C-NMR (101 MHz, DMSO-d_6): δ (ppm) = 165.5 (C_4), 155.5 (C_2), 141.5 (C_6), 93.7 (C_5), 89.8 ($C_{1'}$), 84.0 ($C_{4'}$), 73.1 ($C_{2'}$), 69.4 ($C_{3'}$), 60.6 ($C_{5'}$).

HR-ESIMS (*m/z* [M+H]$^+$): calc. for $[C_9H_{16}N_3O_5]^+$: 244.0928; **obs.:** 244.0926.

List of Figures

List of Schemes

List of Tables

Bibliography

[1] A. Eschenmoser, *Tetrahedron* **2007**, *63*(52), 12821–12844.

[2] F. Miescher, *Medizinisch Chemische Untersuchungen* **1871**, *4*, 441–460.

[3] A. R. Todd, „Synthesis in the study of nucleotides“, **1957**.

[4] A. Kronberg, „The biologic synthesis of deoxyribonucleic acid“, **1959**.

[5] R. W. Holley, „Alanine transferRNA“, **1968**.

[6] H. G. Khorana, „Nucleic acid synthesis in the study of the genetic code“, **1968**.

[7] M. Nirenberg, „The genetic code“, **1968**.

[8] P. Berg, „Dissections and Reconstructions of Genes and Chromosomes“, **1980**.

[9] W. Gilbert, „DNA Sequencing and Gene Structure“, **1980**.

[10] B. McClintock, „The significance of responses of the genome to challange“, **1983**.

[11] S. Altman, „Enzymatic Cleaveage of RNA by RNA“, **1989**.

[12] T. R. Cech, „Self-Splicing and Enzymatic Ativity of an Intervening Sequence RNA from Tetrahymena“, **1989**.

[13] T. Lindahl, „The Intrinsic Fragility of DNA“, **2015**.

[14] A. Sanacar, „Mechanisms of DNA Repair by Photolyase and Excision Nuclease“, **2015**.

[15] F. M. Kruse, J. S. Teichert, O. Trapp, *Chem. Eur. J.* **2020**, *26*(65), 14776–14790.

[16] P. A. Levene, W. A. Jacobs, *Ber. Dtsch. Chem. Ges.* **1909**, *42*(2), 2703–2706.

[17] P. A. Levene, W. A. Jacobs, *Ber. Dtsch. Chem. Ges.* **1909**, *42*(2), 2703–2706.

[18] M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*(7244), 239–242.

[19] S. Becker, J. Feldmann, S. Wiedemann, H. Okamura, C. Schneider, K. Iwan, A. Crisp, M. Rossa, T. Amatov, T. Carell, *Science* **2019**, *366*(6461), 76–82.

[20] D. Canfield, *Annu. Rev. Earth Planet. Sci.* **2005**, *33*(1), 1–36.

[21] K. J. Zahnle, *Elements* **2006**, *2*(4), 217–222.

[22] R. M. Garrels, E. A. Perry, F. T. Mackenzie, *Econ. Geol.* **1973**, *68*(7), 1173–1179.

[23] H. D. Holland, *Geochem. News* **1999**, *100*, 20–22.

[24] A. A. Borisov, *Petrology* **2016**, *24*(2), 117–126.

[25] B. J. Wood, A. N. Halliday, *Nature* **2010**, *465*(7299), 767–770.

[26] J. F. Kasting, *Science* **1993**, *259*(5097), 920–926.

[27] D. Trail, E. B. Watson, N. D. Tailby, *Nature* **2011**, *480*(7375), 79–82.

[28] J. Wade, B. J. Wood, *Earth Planet. Sci. Lett.* **2005**, *236*(1), 78–95.

[29] R. Wordsworth, R. Pierrehumbert, *Science* **2013**, *339*(6115), 64–67.

[30] Z. R. Todd, A. C. Fahrenbach, C. J. Magnani, S. Ranjan, A. Björkbom, J. W. Szostak, D. D. Sasselov, *ChemComm* **2018**, *54*(9), 1121–1124.

[31] W. R. Kuhn, S. K. Atreya, *Icarus* **1979**, *37*(1), 207–213.

[32] M. G. Trainer, *Curr. Org. Chem.* **2013**, *17*(16), 1710–1723.

[33] A. R. Sarafian, S. G. Nielsen, H. R. Marschall, F. M. McCubbin, B. D. Monteleone, *Science* **2014**, *346*(6209), 623–626.

[34] S. J. Mojzsis, T. M. Harrison, R. T. Pidgeon, *Nature* **2001**, *409*(6817), 178–181.

[35] N. Kitadai, S. Maruyama, *Geosci. Front.* **2018**, *9*(4), 1117–1153.

[36] G. Feulner, *Rev. Geophys.* **2012**, *50*(2).

[37] V. S. Airapetian, A. Glocer, G. Gronoff, E. Hebrard, W. Danchi, *Nat. Geosci.* **2016**, *9*(6), 452–455.

[38] R. Ramirez, *Nat. Geosci.* **2016**, *9*(6), 413–414.

[39] M. P. Robertson, S. L. Miller, *Nature* **1995**, *375*(6534), 772–774.

[40] B. Burcar, M. Pasek, M. Gull, B. J. Cafferty, F. Velasco, N. V. Hud, C. Menor-Salvan, *Angew. Chem. Int. Ed.* **2016**, *55*(42), 13249–13253.

[41] P. A. Levene, *J. Biol. Chem.* **1917**, *31*(3), 591–598.

[42] P. A. Levene, E. S. London, *J. Biol. Chem.* **1929**, *83*, 793–802.

[43] E. Fischer, B. Helferich, *Ber. Dtsch. Chem. Ges.* **1914**, *47*(1), 210–235.

[44] W. Traube, *Liebigs Ann. Chem.* **1904**, *331*(1), 64–88.

[45] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1914**, *47*(2), 1377–1393.

[46] W. Traube, *Ber. Dtsch. Chem. Ges.* **1904**, *37*(4), 4544–4547.

[47] B. Lythgoe, H. Smith, A. R. Todd, *J. Chem. Soc.* **1947**, (0), 355–357.

[48] J. Davoll, B. Lythgoe, A. R. Todd, *J. Chem. Soc.* **1948**, (0), 1685–1687.

[49] G. A. Howard, G. W. Kenner, B. Lythgoe, A. R. Todd, *J. Chem. Soc.* **1946**, 855–861.

[50] J. Davoll, B. Lythgoe, A. R. Todd, *J. Chem. Soc.* **1946**, 833–838.

[51] G. A. Howard, B. Lythgoe, A. R. Todd, *J. Chem. Soc.* **1947**, 1052–1054.

[52] J. Oró, *Biochem. Biophys. Res. Commun.* **1960**, *2*(6), 407–412.

[53] K. A. P. Oró, J., *Arch. Biochem. Biophys.* **1961**, *94*(2), 217–227.

[54] W. Löb, *Ber. Dtsch. Chem. Ges* **1913**, *46*(1), 684–697.

[55] S. L. Miller, *J. Am. Chem. Soc.* **1955**, *77*(9), 2351–2361.

[56] S. Yuasa, D. Flory, B. Basile, J. Oró, *J. Mol. Evol.* **1984**, *20*(1), 52–58.

[57] R. A. Sanchez, J. P. Ferris, L. E. Orgel, *J. Mol. Biol.* **1967**, *30*(2), 223–253.

[58] J. P. Ferris, R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **1968**, *33*(3), 693–704.

[59] L. E. Orgel, *Orig. Life Evol. Biosph.* **2002**, *32*(3), 279–281.

[60] R. A. Sanchez, J. P. Ferris, L. E. Orgel, *J. Mol. Bio.* **1968**, *38*(1), 121–128.

[61] R. A. Sanchez, J. P. Ferris, L. E. Orgel, *Science* **1966**, *154*(3750), 784–785.

[62] H. Okamura, S. Becker, N. Tiede, S. Wiedemann, J. Feldmann, T. Carell, *Chem-Comm* **2019**.

[63] J. H. Hong, H. J. Pownall, R. S. Becker, *Photochem. Photobiol.* **1976**, *24*(3), 217–222.

[64] T. Sasaki, T. Yoshioka, K. Shoji, *J. Chem. Soc.* **1969**, (7), 1086–1088.

[65] M. K. Phibbs, P. A. Sipos, „CA19770291729“, **1977**.

[66] W. E. Truce, G. J. W. Tichenor, *J. Org. Chem.* **1972**, *37*(15), 2391–2396.

[67] K. I. Morita, S. Hashimoto, Naoto, „DE2506029“, **1975**.

[68] J. P. Ferris, G. Goldstein, D. J. Beaulieu, *J. Am. Chem. Soc.* **1970**, *92*(22), 6598–6603.

[69] J. P. Ferris, O. S. Zamek, A. M. Altbuch, H. Freiman, *J. Mol. Evol.* **1974**, *3*(4), 301–309.

[70] Y. L. Wang, H. D. Lee, M. W. Beach, D. W. Margerum, *Inorg. Chem.* **1987**, *26*(15), 2444–2449.

[71] Y. Hirose, K. Ohmuro, M. Saigoh, T. Nakayama, Y. Yamagata, *Orig. Life Evol. Biosph.* **1990**, *20*(6), 471–481.

[72] R. Shapiro, R. S. Klein, *Biochemistry* **1966**, *5*(7), 2358–2362.

[73] R. Shapiro, R. S. Klein, *Biochemistry* **1967**, *6*(11), 3576–3582.

[74] T. Lindahl, B. Nyberg, *Biochemistry* **1974**, *13*(16), 3405–3410.

[75] T. Lindahl, *Nature* **1993**, *362*(6422), 709–715.

[76] E. R. Garrett, J. Tsau, *J. Pharm. Sci.* **1972**, *61*(7), 1052–1061.

[77] L. A. Frederico, T. A. Kunkel, B. R. Shaw, *Biochemistry* **1990**, *29*(10), 2532–2537.

[78] L. Frick, J. P. Mac Neela, R. Wolfenden, *Bioorg. Chem.* **1987**, *15*(2), 100–108.

[79] B. K. Duncan, J. H. Miller, *Nature* **1980**, *287*(5782), 560–561.

[80] *Covalent Interactions of Nucleic Acids with Small Molecules and Their Repair*, S. 295–340, The Royal Society of Chemistry, **2006**.

[81] R. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*(8), 4396–4401.

[82] A. Butlerow, *Liebigs Ann. Chem.* **1861**, *120*(3), 295–298.

[83] R. Breslow, *Tetrahedron Lett.* **1959**, *1*(21), 22–26.

[84] A. Ricardo, M. A. Carrigan, A. N. Olcott, S. A. Benner, *Science* **2004**, *303*(5655), 196–196.

[85] C. Ponnamperuma, C. Sagan, R. Mariner, *Nature* **1963**, *199*(4890), 222–226.

[86] W. D. Fuller, R. A. Sanchez, L. E. Orgel, *J. Mol. Evol.* **1972**, *1*(3), 249–257.

[87] R. A. Horne, *Sea Water*, Bd. 6, S. 107–140, Elsevier, **1970**.

[88] W. D. Fuller, R. A. Sanchez, L. E. Orgel, *J. Mol. Bio.* **1972**, *67*(1), 25–33.

[89] G. Springsteen, G. F. Joyce, *J. Am. Chem. Soc.* **2004**, *126*(31), 9578–9583.

[90] R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **1970**, *47*(3), 531–543.

[91] D. H. Shannahoff, R. A. Sanchez, *J. Org. Chem.* **1973**, *38*(3), 593–598.

[92] O. Leslie E, *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*(2), 99–123.

[93] A. Ingar, R. W. A. Luke, B. R. Hayter, J. D. Sutherland, *ChemBioChem* **2003**, *4*(6), 504–507.

[94] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*(4356), 737–738.

[95] M. Böhringer, H.-J. Roth, J. Hunziker, M. Göbel, R. Krishnan, A. Giger, B. Schweizer, J. Schreiber, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1992**, *75*(5), 1416–1477.

[96] A. Eschenmoser, M. Dobler, *Helv. Chim. Acta* **1992**, *75*(1), 218–259.

[97] K. Groebke, J. Hunziker, W. Fraser, L. Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1998**, *81*(3-4), 375–474.

[98] J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diederichsen, M. Göbel, R. Krishnan, B. Jaun, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*(1), 259–352.

[99] G. Otting, M. Billeter, K. Wüthrich, H.-J. Roth, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*(8), 2701–2756.

[100] S. Pitsch, S. Wendeborn, B. Jaun, A. Eschenmoser, *Helvetica Chimica Acta* **1993**, *76*(6), 2161–2183.

[101] K.-U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science* **2000**, *290*(5495), 1347–1351.

[102] M. Böhringer, https://doi.org/10.3929/ethz-a-000592965, **1991**.

[103] E. Koch, Klemens, https://doi.org/10.3929/ethz-a-000694124, **1992**.

[104] U. P. Trinks, https://doi.org/10.3929/ethz-a-000413538, **1987**.

[105] S. Becker, I. Thoma, A. Deutsch, T. Gehrke, P. Mayer, H. Zipse, T. Carell, *Science* **2016**, *352*(6287), 833–836.

[106] B. E. Turner, H. S. Liszt, N. Kaifu, A. G. Kisliakov, *Astrophys. J.* **1975**, *201*, L149.

[107] A. Coutens, E. R. Willis, R. T. Garrod, H. S. P. Müller, T. L. Bourke, H. Calcutt, M. N. Drozdovskaya, J. K. Jørgensen, N. F. W. Ligterink, M. V. Persson, G. Stéphan, M. H. D. van der Wiel, E. F. van Dishoeck, S. F. Wampfler, *Astron. Astrophys.* **2018**, *612*, A107.

[108] C. Sleiman, G. El Dib, M. Rosi, D. Skouteris, N. Balucani, A. Canosa, *Phys. Chem. Chem. Phys.* **2018**, *20*(8), 5478–5489.

[109] P. Thaddeus, *Philos. Trans. R. Soc. B* **2006**, *361*(1474), 1681–1687.

[110] M. Powner, J. Sutherland, *Angew. Chem. Int. Ed.* **2010**, *49*(27), 4641–4643.

[111] F. Brotzel, Y. C. Chu, H. Mayr, *J. Org. Chem.* **2007**, *72*(10), 3679–3688.

[112] B. Maji, M. Baidya, J. Ammer, S. Kobayashi, P. Mayer, A. R. Ofial, H. Mayr, *Eur. J. Org. Chem.* **2013**, *2013*(16), 3369–3377.

[113] A. Choudhary, K. J. Kamer, M. W. Powner, J. D. Sutherland, R. T. Raines, *ACS Chem. Biol.* **2010**, *5*(7), 655–657.

[114] J. Xu, M. Tsanakopoulou, C. J. Magnani, R. Szabla, J. E. Sponer, J. Sponer, R. W. Gora, J. D. Sutherland, *Nat. Chem.* **2017**, *9*(4), 303–309.

[115] S. Islam, D.-K. Bučar, M. W. Powner, *Nat. Chem.* **2017**, *advance online publication.*

[116] M. W. Powner, J. D. Sutherland, J. W. Szostak, *J. Am. Chem. Soc.* **2010**, *132*(46), 16677–16688.

[117] B. Moffatt, H. Ashihara, *Purine and Pyrimidine Nucleotide Synthesis and Metabolism*, Bd. 1 von *The Arabidopsis book / American Society of Plant Biologists*, **2002**, S. e0018.

[118] J. P. Ferris, L. E. Orgel, *J. Am. Chem. Soc.* **1966**, *88*(5), 1074–1074.

[119] S. Stairs, A. Nikmal, D.-K. Bučar, S.-L. Zheng, J. W. Szostak, M. W. Powner, *Nat. Commun.* **2017**, *8*, 15270.

[120] M. M. Greenberg, *Acc. Chem. Res.* **2012**, *45*(4), 588–597.

[121] H. J. Kim, Y. Furukawa, T. Kakegawa, A. Bita, R. Scorei, S. A. Benner, *Angew. Chem. Int. Ed.* **2016**, *55*(51), 15816–15820.

[122] S. Becker, C. Schneider, H. Okamura, A. Crisp, T. Amatov, M. Dejmek, T. Carell, *Nat. Commun.* **2018**, *9*(1), 163.

[123] C. R. Woese, *PNAS* **1968**, *59*(1), 110–117.

[124] E. M. Silverman, R. J. Madix, C. R. Brundle, *J. Vac. Sci. Technol.* **1981**, *18*(2), 616–619.

[125] C. Schneider, S. Becker, H. Okamura, A. Crisp, T. Amatov, M. Stadlmeier, T. Carell, *Angew. Chem. Int. Ed.* **2018**, *57*(20), 5943–5946.

[126] B. T. Golding, C. Bleasdale, J. McGinnis, S. Müller, H. T. Rees, N. H. Rees, P. B. Farmer, W. P. Watson, *Tetrahedron* **1997**, *53*(11), 4063–4082.

[127] M. A. Machnicka, K. Milanowska, O. Osman Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz, K. M. Rother, M. Helm, J. M. Bujnicki, H. Grosjean, *Nucleic Acids Res.* **2012**, *41*(D1), D262–D267.

[128] T. Machinami, T. Suami, *Bull. Chem. Soc. Jpn.* **1975**, *48*(4), 1333–1334.

[129] S. Bhowmik, R. Krishnamurthy, *Nat. Chem.* **2019**, *11*(11), 1009–1018.

[130] T. Lindahl, B. Nyberg, *Biochemistry* **1972**, *11*(19), 3610–3618.

[131] J. Xu, N. J. Green, C. Gibard, R. Krishnamurthy, J. D. Sutherland, *Nat. Chem.* **2019**, *11*(5), 457–462.

[132] B.-M. Sjöberg, M. Sahlin, *Thiols in redox mechanism of ribonucleotide reductase*, Bd. 348, S. 1–21, Academic Press, **2002**.

[133] C. Gibard, S. Bhowmik, M. Karki, E.-K. Kim, R. Krishnamurthy, *Nat. Chem.* **2017**, *10*, 212.

[134] E. Vongerichten, *Liebigs Ann. Chem.* **1901**, *318*(1), 121–136.

[135] A. G. Darvill, M. McNeil, P. Albersheim, *Plant Physiol.* **1978**, *62*(3), 418–422.

[136] S. Pérez, M. A. Rodríguez-Carvajal, T. Doco, *Biochimie* **2003**, *85*(1), 109–121.

[137] M. A. O'Neill, T. Ishii, P. Albersheim, A. G. Darvill, *Ann. Rev. Plant Bio.* **2004**, *55*(1), 109–139.

[138] M. Hamburger, M. Gupta, K. Hostettmann, *Phytochemistry* **1985**, *24*(11), 2689–2692.

[139] M. Pičmanová, B. L. Møller, *Glycobiology* **2016**, *26*(5), 430–442.

[140] M. Mølhøj, R. Verma, W.-D. Reiter, *The Plant Journal* **2003**, *35*(6), 693–703.

[141] T. Eixelsberger, D. Horvat, A. Gutmann, H. Weber, B. Nidetzky, *Angw. Chemw Int. Ed.* **2017**, *56*(9), 2503–2507.

[142] J. C. Chaput, J. W. Szostak, *J. Am. Chem. Soc.* **2003**, *125*(31), 9274–9275.

[143] M. Kataoka, K. Sato, A. Matsuda, *Nucleic Acids Symposium Series* **2008**, *52*(1), 281–282.

[144] J. Wang, M. Sánchez-Roselló, J. L. Aceña, C. del Pozo, A. E. Sorochinsky, S. Fustero, V. A. Soloshonok, H. Liu, *Chem. Rev.* **2014**, *114*(4), 2432–2506.

[145] J. Kim, J. H. Hong, *Carbohydr. Res.* **2003**, *338*(8), 705–710.

[146] S. Kil Ahn, D. Kim, S. Kil Ahn, D. Kim, *Chem. Commun.* **1998**, (9), 967–968.

[147] S. Kim, J. H. Hong, *Bull. Korean Chem. Soc.* **2015**, *36*(10), 2484–2493.

[148] S.-Y. Park, T. Yokoyama, N. Shibayama, Y. Shiro, J. R. H. Tame, *J. Mol. Bio.* **2006**, *360*(3), 690–701.

[149] A. Bacher, S. Eberhardt, M. Fischer, K. Kis, G. Richter, *Annu. Rev. Nutr.* **2000**, *20*(1), 153–167.

[150] J. A. MacLaren, *J. Bacteriol.* **1952**, *63*(2), 233–241.

[151] A. Bacher, B. Mailänder, *J. Biol. Chem.* **1973**, *248*(17), 6227–6231.

[152] A. Bacher, Q. Le Van, P. J. Keller, H. G. Floss, *J. Biol. Chem.* **1983**, *258*(22), 13431–7.

[153] R. Beach, G. W. E. Plaut, *Tetrahedron Lett.* **1969**, *10*(40), 3489–3492.

[154] R. Kuhn, A. H. Cook, *Ber. Dtsch. Chem. Ges.* **1937**, *70*(4), 761–768.

[155] T. Kishi, M. Asai, T. Masuda, S. Kuwada, *Chem. Pharm. Bull.* **1959**, *7*(4), 515–519.

[156] E. G. Brown, T. W. Goodwin, O. T. G. Jones, *Biochem. J.* **1958**, *68*(1), 40–49.

[157] W. Traube, *Ber. Dtsch. Chem. Ges.* **1900**, *33*(1), 1371–1383.

[158] H. Katagiri, I. Takeda, K. Imai, *J. Vitaminol.* **1959**, *5*(4), 287–297.

[159] A. Bacher, Q. L. Van, P. J. Keller, H. G. Floss, *J. Am. Chem. Soc.* **1985**, *107*(22), 6380–6385.

[160] G. Neuberger, A. Bacher, *Biochem. Biophys. Res. Commun.* **1985**, *127*(1), 175–181.

[161] K. Kis, K. Kugelbrey, A. Bacher, *J. Org. Chem.* **2001**, *66*(8), 2555–2559.

[162] S. Gerhardt, A.-K. Schott, N. Kairies, M. Cushman, B. Illarionov, W. Eisenreich, A. Bacher, R. Huber, S. Steinbacher, M. Fischer, *Structure* **2002**, *10*(10), 1371–1381.

[163] R. Beach, G. W. E. Plaut, *Biochemistry* **1970**, *9*(4), 760–770.

[164] Y.-J. Zheng, D. B. Jordan, D.-I. Liao, *Bioorg. Chem.* **2003**, *31*(4), 278–287.

[165] C. J. Strupp, https://doi.org/10.3929/ethz-a-000667161, **1992**.

[166] G. M. Blackburn, A. W. Johnson, *J. Chem. Soc.* **1960**, (0), 4347–4358.

[167] G. M. Blackburn, A. W. Johnson, *J. Chem. Soc.* **1960**, (0), 4358–4365.

[168] H. Kuhn, *Angew. Chem. Int. Ed.* **1972**, *11*(9), 798–820.

[169] W. Gilbert, *Nature* **1986**, *319*(6055), 618–618.

[170] G. F. Joyce, *Nature* **1989**, *338*(6212), 217–224.

[171] S. A. Benner, A. D. Ellington, A. Tauer, *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*(18), 7054–7058.

[172] A. Eschenmoser, *Science* **1999**, *284*(5423), 2118–2124.

[173] R. Shapiro, *IUBMB Life* **2000**, *49*(3), 173–176.

[174] N. Lane, W. Martin, *Cell* **2012**, *151*(7), 1406–1416.

[175] W. F. Martin, F. L. Sousa, N. Lane, *Science* **2014**, *344*(6188), 1092–1093.

[176] R. F. Ludlow, S. Otto, *Chem. Soc. Rev.* **2008**, *37*(1), 101–108.

[177] J. Boekhoven, *Nat. Nanotechnol.* **2018**, *13*(11), 979–980.

[178] T. R. Cech, O. C. Uhlenbeck, *Nature* **1994**, *372*(6501), 39–40.

[179] H. W. Pley, K. M. Flaherty, D. B. McKay, *Nature* **1994**, *372*(6501), 68–74.

[180] J. R. Lorsch, J. W. Szostak, *Nature* **1994**, *371*(6492), 31–36.

[181] Y. Ura, J. M. Beierle, L. J. Leman, L. E. Orgel, M. R. Ghadiri, *Science* **2009**, *325*(5936), 73–77.

[182] N. V. Hud, B. J. Cafferty, R. Krishnamurthy, L. D. Williams, *Chem. Biol.* **2013**, *20*(4), 466–74.

[183] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*(1), 147–157.

[184] W. Gilbert, *Nature* **1978**, *271*(5645), 501–501.

[185] A. Eschenmoser, *Angew. Chem. Int. Ed.* **1988**, *27*(1), 5–39.

[186] A. M. Poole, D. T. Logan, B.-M. Sjöberg, *J. Mol. Evol.* **2002**, *55*(2), 180–196.

[187] M. W. Powner, J. D. Sutherland, J. W. Szostak, *Synlett* **2011**, *22*(14), 1956–1964.

[188] F. H. C. Crick, *J. Mol. Biol.* **1968**, *38*(3), 367–379.

[189] E. Kuruvilla, G. B. Schuster, N. V. Hud, *ChemBioChem* **2013**, *14*(1), 45–48.

[190] B. D. Heuberger, C. Switzer, *ChemBioChem* **2008**, *9*(17), 2779–2783.

[191] J.-M. Lehn, M. Mascal, A. Decian, J. Fischer, *ChemComm* **1990**, (6), 479–481.

[192] E. Meggers, L. Zhang, *Acc. Chem. Res.* **2010**, *43*(8), 1092–1102.

[193] M. Hernández-Rodríguez, J. Xie, Y. M. Osornio, R. Krishnamurthy, *Chem.: Asian J.* **2011**, *6*(5), 1252–1262.

[194] G. Mittapalli, K. Reddy, H. Xiong, O. Munoz, B. Han, F. DeRiccardis, R. Krishnamurthy, A. Eschenmoser, *Angew. Chem. Int. Ed.* **2007**, *46*(14), 2470–2477.

[195] H. D. Bean, F. A. L. Anet, I. R. Gould, N. V. Hud, *Orig. Life Evol. Biosph.* **2006**, *36*(1), 39–63.

[196] A. D. Keefe, S. L. Miller, *J. Mol. Evol.* **1995**, *41*(6), 693–702.

[197] F. Westheimer, *Science* **1987**, *235*(4793), 1173–1178.

[198] J. P. Dworkin, S. L. Miller, *Carbohydr. Res.* **2000**, *329*(2), 359–365.

[199] A. C. Rios, Y. Tor, *Astrobiology* **2012**, *12*(9), 884–891.

[200] R. Lohrmann, L. E. Orgel, *Science* **1971**, *171*(3970), 490–494.

[201] R. Saladino, C. Crestini, F. Ciciriello, G. Costanzo, E. DiMauro, *Chem. Biodiver.* **2007**, *4*(4), 694–720.

[202] S. A. Benner, A. Ricardo, M. A. Carrigan, *Curr. Op. Chem. Bio.* **2004**, *8*(6), 672–689.

[203] A. M. Schoffstall, E. M. Laing, *Orig. Life Evol. Biosph.* **1985**, *15*(2), 141–150.

[204] M. J. Russell, W. Martin, *Trends Biochem. Sci.* **2004**, *29*(7), 358–363.

[205] D. S. Kelley, J. A. Karson, D. K. Blackman, G. L. Früh-Green, D. A. Butterfield, M. D. Lilley, E. J. Olson, M. O. Schrenk, K. K. Roe, G. T. Lebon, P. Rivizzigno, A. T. S. P. the, *Nature* **2001**, *412*(6843), 145–149.

[206] W. Bach, H. Paulick, C. J. Garrido, B. Ildefonse, W. P. Meurer, S. E. Humphris, *Geophys. Res. Lett.* **2006**, *33*(13).

[207] G. Proskurowski, M. D. Lilley, J. S. Seewald, G. L. Früh-Green, E. J. Olson, J. E. Lupton, S. P. Sylva, D. S. Kelley, *Science* **2008**, *319*(5863), 604–607.

[208] J. C. Xavier, W. Hordijk, S. Kauffman, M. Steel, W. F. Martin, *Proc. R. Soc. B* **2020**, *287*(1922), 20192377.

[209] H. Dobbek, V. Svetlitchnyi, L. Gremer, R. Huber, O. Meyer, *Science* **2001**, *293*(5533), 1281–1285.

[210] P. W. Kudella, K. Preißinger, M. Morasch, C. F. Dirscherl, D. Braun, A. Wixforth, C. Westerhausen, *Sci. Rep.* **2019**, *9*(1), 18808.

[211] E. D. Agerschou, C. B. Mast, D. Braun, *Synlett* **2017**, *28*(01), 56–63.

[212] M. Morasch, J. Liu, C. F. Dirscherl, A. Ianeselli, A. Kühnlein, K. Le Vay, P. Schwintek, S. Islam, M. K. Corpinot, B. Scheu, D. B. Dingwell, P. Schwille, H. Mutschler, M. W. Powner, C. B. Mast, D. Braun, *Nat. Chem.* **2019**, *11*(9), 779–788.

[213] F. M. Möller, F. Kriegel, M. Kieß, V. Sojo, D. Braun, *Angew. Chem. Int. Ed.* **2017**, *56*(9), 2340–2344.

[214] G. Bruylants, K. Bartik, J. Reisse, *C. R. Chim.* **2011**, *14*(4), 388–391.

[215] A. Pross, V. Khodorkovsky, *J. Phys. Org. Chem.* **2004**, *17*(4), 312–316.

[216] P. Addy, *Pure Appl. Chem.* **2005**, *77*(11), 1905–1921.

[217] S. Lifson, *J. Mol. Evol.* **1997**, *44*(1), 1–8.

[218] R. Pascal, A. Pross, J. D. Sutherland, *Open Biol.* **2013**, *3*(11), 130156.

[219] R. Pascal, *J. Syst. Chem.* **2012**, *3*(1), 3.

[220] A. J. Lotka, *Proc. Natl. Acad. Sci. U.S.A.* **1922**, *8*(6), 147–151.

[221] J. Boekhoven, A. Brizard, K. Kowlgi, G. Koper, R. Eelkema, J. vanEsch, *Angew. Chem. Int. Ed.* **2010**, *49*(28), 4825–4828.

[222] J. Boekhoven, W. E. Hendriksen, G. J. M. Koper, R. Eelkema, J. H. van Esch, *Science* **2015**, *349*(6252), 1075–1079.

[223] P. B. Rimmer, O. Shorttle, *Life* **2019**, *9*.

[224] A. Pross, *Orig. Life Evol. Biosph.* **2004**, *34*(3), 307–321.

[225] F. Tian, J. F. Kasting, K. Zahnle, *Earth Planet. Sci. Lett.* **2011**, *308*(3), 417–423.

[226] W. R. Kuhn, J. F. Kasting, *Nature* **1983**, *301*(5895), 53–55.

[227] J. F. Kasting, M. T. Howard, *Philos. Trans. R. Soc. B* **2006**, *361*(1474), 1733–1742.

[228] R. Saladino, E. DiMauro, J. M. García-Ruiz, *Chem.: Eur. J.* **2019**, *25*(13), 3181–3189.

[229] A. Rimola, M. Sodupe, P. Ugliengo, *Life (2075-1729)* **2019**, *9*(1), 10–10.

[230] E. Knoevenagel, *Ber. Dtsch. Chem. Ges.* **1898**, *31*(1), 738–748.

[231] B. List, *Angew. Chem. Int. Ed.* **2010**, *49*(10), 1730–1734.

[232] R. Appel, H. Mayr, *J. Am. Chem. Soc.* **2011**, *133*(21), 8240–8251.

[233] R. Shapiro, *Origins Life Evol Biosphere* **1988**, *18*(1), 71–85.

[234] J. S. Teichert, F. M. Kruse, O. Trapp, *Angew. Chem. Int. Ed.* **2019**, *58*(29), 9944–9947.

[235] O. R. Wauchope, M. M. Mitchener, W. N. Beavers, J. J. Galligan, J. M. Camarillo, W. D. Sanders, P. J. Kingsley, H.-N. Shim, T. Blackwell, T. Luong, M. deCaestecker, J. P. Fessel, L. J. Marnett, *Nucleic Acids Res.* **2018**, *46*(7), 3458–3467.

[236] M. Breugst, F. Corral-Bautista, H. Mayr, *Chem. Eur. J.* **2012**, *18*(1), 127–137.

[237] N. T. Anh, O. Eisenstein, *Tetrahedron Lett.* **1976**, *17*(3), 155–158.

[238] J. E. Baldwin, *Chem. Comm.* **1976**, (18), 734–736.

[239] S. Neubauer, A. Rugova, D. B. Chu, H. Drexler, A. Ganner, M. Sauer, D. Mattanovich, S. Hann, G. Koellensperger, *Anal. Bioanal. Chem.* **2012**, *404*(3), 799–808.

[240] R. Dalpozzo, A. D. Nino, L. Maiuolo, A. Procopio, R. Romeo, G. Sindona, *Synthesis* **2002**, *2002*(02), 0172–0174.

[241] U. Niedballa, H. Vorbrueggen, *J. Org. Chem.* **1974**, *39*(25), 3654–3660.

[242] S. Reformatsky, *Ber. Dtsch. Chem. Ges.* **1887**, *20*(1), 1210–1211.

[243] K. Kanai, H. Wakabayashi, T. Honda, *Org. Lett.* **2000**, *2*(16), 2549–2551.

[244] F. G. Bordwell, G. Z. Ji, *J. Am. Chem. Soc.* **1991**, *113*(22), 8398–8401.

[245] K. Patora-Komisarska, M. Benohoud, H. Ishikawa, D. Seebach, Y. Hayashi, *Helv. Chim. Acta* **2011**, *94*(5), 719–745.

[246] V. A. Yaylayan, S. Harty-Majors, A. A. Ismail, *Carbohyd. Res.* **1998**, *309*(1), 31–38.

[247] R. Saladino, U. Ciambecchini, C. Crestini, G. Costanzo, R. Negri, E. Di Mauro, *ChemBioChem* **2003**, *4*(6), 514–521.

[248] E. Farbenfabriken vorm. F. Bayer, „DE175415“, **1906**.

[249] D. Firma F. Merck, „DE227390“, **1910**.

[250] S.-L. You, *Chem. Asian. J.* **2007**, *2*(7), 820–827.

[251] G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*(9), 2176–2179.

[252] L. Farrugia, *J. App. Cryst.* **2012**, *45*(4), 849–854.

[253] P. Ciapetti, M. Taddei, *Tetrahedron* **1998**, *54*(37), 11305–11310.

[254] R. Zou, M. J. Robins, *Can. J. Chem.* **1987**, *65*(6), 1436–1437.

[255] K.-T. Yang, L.-Y. Hsu, *Synth. Commun.* **2001**, *31*(6), 893–898.

[256] K. C. Nicolaou, S. E. Webber, *Synthesis* **1986**, *1986*(06), 453–461.

[257] H. Zinner, H. Voigt, J. Voigt, *Carbhydr. Res.* **1968**, *7*(1), 38–55.

[258] E. J. Corey, H. Cheng, C. H. Baker, S. P. T. Matsuda, D. Li, X. Song, *J. Am. Chem. Soc.* **1997**, *119*(6), 1277–1288.

[259] M. Gao, Y. Chen, S. Tan, J. H. Reibenspies, R. A. Zingaro, *Heteroatom Chem.* **2008**, *19*(2), 199–206.

[260] S. M. N. Efange, E. M. Alessi, H. C. Shih, Y. C. Cheng, T. J. Bardos, *J. Med. Chem.* **1985**, *28*(7), 904–910.

[261] M. A. Palazzolo, M. J. Nigro, A. M. Iribarren, E. S. Lewkowicz, *Eur. J. Org. Chem.* **2016**, *2016*(5), 921–924.

[262] P. N. Solyev, M. V. Jasko, A. A. Kleymenova, M. K. Kukhanova, S. N. Kochetkov, *Org. Biomol. Chem.* **2015**, *13*(44), 10946–10956.

Danksagung

Mein herzlichster Dank gilt *Prof. Dr. Oliver Trapp*, dafür dass ich die Masterarbeit und Promotion in seinem Arbeitskreis anfertigen durfte. Vielen Dank für die intensive und konstruktive Betreuung, die motivierenden und erhellenden Gespräche, die vielen Laborbesuche und die effektive, Spaß bringende Arbeit an den Papern. Deine gewinnende und motivierende Art hat maßgeblich zum Erfolg dieser Arbeit beigetragen.

Prof. Dr. Thomas Carell danke ich herzlich für die Übernahme des Zweitgutachtens und die externe Motivation seitens der präbiotischen Chemie.

Ich bedanke mich bei *Prof. Dr. Philip Tinnefeld*, *Prof. Dr. Franz Bracher*, *Prof. Dr. Ivan Huc*, und *Prof. Dr. Ochsenfeld* als Mitglieder der Prüfungskommission.

Mein Dank gilt insbesondere den vielen, fast unscheinbaren, Helfern des analytischen Departments und der Versorgung, ohne die diese Arbeit nicht möglich gewesen wäre. Vielen Dank an *C. Ober, Dr. Stephenson* für unzählige gemessene NMR, an *S. Kosak, Dr. Spahl* für die Massespektren, *Dr. Mayer* für viele erfreuliche Kristallstrukturen und *H. Buchholz, M.Ġayer, R. Schürer* für Ihre Hilfe. *Dr. Claudia Meier*, *Brigitte Janker*, *Carrie Louis Hildegard Lipfert* und *Heike Traub* danke ich für die Unterstützung in kleinen und großen organisatorischen Fragen und die vielen erheiternden Gesprächen.

Ein großer Dank geht an alle ehemaligen und aktuellen Mitglieder des Arbeitskreises ohne die diese Promotionszeit nur halb so viel Spaß gemacht hätte. Danke an Stefanie Auras, Maximilian Bechtel, Regina Berg, Gloria Betzenbichler, Anna Closs, Marian Ebeling, Roberta Franzini, Elina Hagelskamp, Maren Haas, Simone Heitsch, Laura Huber, Kai Kohler, Sabrina Kräh, Nathalie Kurrle, Saskia Lamour, Max Leopold,

Lena Mayer, Jan-Michael Menke, Sebastian Pallmann, Fabian Sauer, Jan-Felix Scholtes, Christoph Seifert, Max Siebert, Alexander Siegle, Constanze Sydow und Jennifer Teichert für unvergleichliche Erlebnisse und eine tolle Zeit.
Unserem Origins of Life Labor danke ich für die ansteckenden Fröhlichkeit, den unerschütterlichen Musikgeschmack, den ein oder anderen "nicht witzig" Stempel, reimende Stunden und für gewinnbringende und entspannende Gespräche in einem hervorragendem Team. Meinen F-PraktikantInnen *Christina*, *Can*, *Carina*, *Peter*, der "Subgroup" (*Florian, Matthias, Patrick*) *Mo* und *Jonas* danke ich für ihr Engagement, fröhliche Freitage, gute Gespräche über Gott und die Welt und die vielen Anregungen.

Den Mitgliedern der *Ottogruppe* danke ich für die kulinarischen Hochgenüsse, unzähligen Pasten, Erweiterung des Gewürzspektums, 10 Varianten von Reis aus dem Ofen, Mittagsdiskussionen und -gespräche, neue Entdeckungen und einem unvergleichbarem freitäglichem Ritual, das ich jetzt schon vermisse. Ich freue mich auf London 2023.

Ich bedanke mich insbesondere bei *Anna* für die unfassbar schnelle und konstruktive Korrektur meiner Arbeit. Ich danke *Kai* für konstruktive Gespräche, abwegige Ideen, wohlschmeckende Kochexperimente, Abende im Sommerwohnzimmer und Kaffeeverkostungen, *Max* für den kulinarische Austausch und deren Verkostung, *Laura*, *Anna*, *Elina* und *Felix* für spannende Konzertbesuche, *Alex* für Kaffepausengespräche und die Beantwortung der ein oder anderen HPLC Frage, *Gloria* für die Reime und die Musik, *Lena* für die Musik, die ein oder andere Teepause und die Orchesterfreizeiten und *Max* für viele gute Warum?-Fragen zur QToF. Bei meiner "Arbeitsehefrau" *Jenny* bedanke ich mich für eine wunderbare Zusammenarbeit seit Tag eins in unserer Box, die unermüdliche Jagd nach der Karotte, Papersessions mit starkem Kaffee, kleine und große Stümper-Momente, Pizzaabende und vieles mehr. Meiner großen Familie danke ich für den Rückhalt ohne den der Weg bis hierhin nicht möglich gewesen wäre, es ist wunderbar zu wissen, dass alles mit tragt. Mein besonderer Dank gebührt dir *Moni*, für deinen Rückhalt, deine Rückfragen, die Lehre der Wurschtigkeit und die unzähligen schönen Momente mit dir.