Synthesis of Adenosine 5'- (Aminoalkyl phosphonates and phosphinates) Analogues of Aminoacyl Adenylates

FLORIAN HAMON

Master’s Degree Project
Stockholm, Sweden 2005

TRITA-NA-E05054
Synthesis of Adenosine 5'-
(Aminoalkyl phosphonates and phosphinates)
Analogue of Aminoacyl Adenylates

FLORIAN HAMON

TRITA-NA-E05054

Master’s Thesis in Biomedical Engineering (20 credits)
at the School of Chemical Engineering,
Royal Institute of Technology year 2005
Supervisor at Nada was Erik Fransén
Examiner was Anders Lansner
Abstract

Translation of the genetic code and protein synthesis from amino acids takes place in the ribosome’s where tRNAs act as key intermediates. Correct translation of genetic information into the specific amino acids is dependent on aminoacyl-tRNA’s. The synthesis of these is catalyzed by aminoacyl-tRNA synthetase. The aminoacylation reaction, the so-called charging of tRNA is carried out in two steps. The first step is the formation of aminoacyl adenylates (I) by a specific reaction involving amino acids, adenosine triphosphate, and an aminoacyl-tRNA synthetase. The amino acids are thus activated, and subsequently transferred to their corresponding tRNA through a highly specific reaction catalyzed by the same enzymes.

It is known that Aminoalkyl adenylates (II) are inhibitors of these enzymes and are thought to occupy the binding sites for the aminoacyl adenylates. A drawback is that Aminoalkyl adenylates are sensitive to biological degradation. Adenosine 5’-(aminoalkyl phosphonates or phosphinates) (III) can prove valuable as stable inhibitors of amino acid activation and as substrates for studies of enzyme – substrate interaction.

My work consisted in developing synthesis routes to obtain the new analogues (III). Many reactions were performed and the general knowledge about the compound family increased. I did manage to reach several intermediary steps into the route towards the targets. Certain steps turned out to be harder and trickier than expected and many modifications had to be made to the initial synthetic routes.
Acknowledgements

I would like to thank my supervisor Roger Strömberg who accepted me in his lab. I want also to thank particularly Merita Murtola who took care of me in the lab work. I want also to thank the other members of the lab: Jessica, Charlotte, Hans and Hasmik and all the former PhD students. My stay in the lab was really nice thanks to all these people, and I had a great time, either in term of scientific or personal exchanges. I want also to thank Erik Fransen, my supervisor in KTH for his support.
# Table of content:

1 - INTRODUCTION .................................................................................................................. 1
  1-1: ROLE OF tRNA AND SYNTHESIS OF AMINOACYL ADENYLATES .................................. 1
  1-2: GOALS ............................................................................................................................ 1
  1-3: OUTLINE OF THIS THESIS ............................................................................................. 1

2 – BIOLOGICAL BACKGROUND ............................................................................................... 2
  2-1: THE MOLECULAR BUILDING BLOCKS: DNA, RNA, AMINO ACIDS AND PROTEINS ... 2
  2-2: THE DOGMA OF MOLECULAR BIOLOGY ...................................................................... 4
  2-3: THE 2 STEPS OF THE GENETIC INFORMATION FLOW .................................................. 5
    2-3.a: Transcription: From DNA to RNA .............................................................................. 5
    2-3.b: Translation: from mRNA to proteins ......................................................................... 5
  2-4: TRANSLATION OF THE GENETIC CODE AND PROTEIN SYNTHESIS .......................... 6
    2-4.a: The role of the tRNAs ................................................................................................. 6
    2-4.b: Synthesis of the tRNAs ............................................................................................... 9
  2-5: THE INTEREST OF THE PROJECT .................................................................................. 9

3 - METHODS AND TECHNIQUES .............................................................................................. 11
  3-1: THE SYNTHETIC STRATEGY ........................................................................................... 11
    3-1.a Protection reactions: ..................................................................................................... 11
    3-1.b Carbon extension: ......................................................................................................... 12
    3-1.c Synthesis of phosphinate: ............................................................................................. 14
    3-1.d Deprotection reaction: .................................................................................................. 16
  3-2: MAIN ANALYSIS AND PURIFICATION TECHNIQUES USED DURING THE PROJECT WORK .................................................................................................................. 17
    3-2.a: Chromatographic methods: ......................................................................................... 17
    3-2.b: Spectroscopic methods ............................................................................................... 18

4 - RESULTS ............................................................................................................................... 20
  4-1: RESULTS CONCERNING THE ROUTE TO TARGET MOLECULE 1: ................................. 20
    4-1.a: Protection of adenosine. ............................................................................................... 20
    4-1.b: Reactions on inosine ................................................................................................... 22
  4-2: RESULTS CONCERNING THE ROUTE TO TARGET MOLECULE 2: .................................. 23

5 – DISCUSSION AND FURTHER STUDIES .............................................................................. 26
  5-1: DISCUSSION .................................................................................................................... 26
  5-2: SUGGESTIONS FOR FURTHER WORK ............................................................................ 26

REFERENCES ............................................................................................................................... 28

APPENDIX I: SCHEME OF THE FIRST STRATEGIC ROUTE FOR THE TARGET MOLECULE 1 (USING ADENOSINE AS STARTING MATERIAL) ................................................. 29

APPENDIX II: SCHEME OF THE SECOND STRATEGIC ROUTE FOR THE TARGET MOLECULE 1 (USING INOSINE AS STARTING MATERIAL) ........................................ 30

APPENDIX III: SCHEME OF THE STRATEGIC ROUTE FOR THE TARGET MOLECULE 2 ................................................................................................................................. 31
List of Figures

Figure 2.1: The different DNA and RNA bases _______________________________ 2
Figure 2.2: The 20 amino acids ____________________________________________ 3
Figure 2.3: Central dogma of the molecular biology, the pathway from DNA to protein. 4
Figure 2.4: Scheme of tRNA – example of tRNA (With a phenylalanine amino acid) ___ 6
Figure 2.5.a: Initiation _____________________________________________________ 7
Figure 2.5.b: Binding of the ribosome _________________________________________ 7
Figure 2.5.c: Reading _______________________________________________________ 8
Figure 2.5.d: Termination of the synthesis ______________________________________ 8
Figure 2.6: Coupling of amino acid and RNA to build a tRNA __________________ 9
Figure 2.7.a: The natural adenylated glycine ________________________________ 10
Figure 2.7.b: The first synthesized analogues _________________________________ 10
Figure 2.7.c: The 2 molecules targeted to be synthesized _________________________ 10
Figure 2.7.d: The 3rd molecule that could be synthesized after development of the synthetic methods. ____________________________ 10
Figure 3.1: Scheme of the protection of the 2´ and 3´ alcohol functions __________ 11
Figure 3.2: Scheme of the protection of the amino function on the base of adenosine __ 12
Figure 3.3: Scheme of the protection by Z group of the amino function. ________________ 12
Figure 3.4: Scheme of the tosylation of the alcohol function ______________________ 13
Figure 3.5: Scheme of the replacement of the tosylate function by a cyanide function __ 13
Figure 3.6: Scheme of the reduction of the cyanide into an amine ________________ 14
Figure 3.7: Scheme of the transformation of amine into alcohol. __________________ 14
Figure 3.8: Scheme of the phosphorylation of protected inosine or adenosine. _______ 15
Figure 3.9: Scheme of the 2 possible pathways to the phosphinate ________________ 15
Figure 3.10: Scheme of the coupling and oxidation reaction. ________________ 16
Figure 3.11: Scheme of the deprotection step, example with target molecule 2. _______ 16
Figure 3.1: Scheme of Thin Layer Chromatography ____________________________ 17
Figure 3.2: Schematic presentation of column chromatography ____________________ 18
Figure 4.1: cyclonucleoside __________________________________________________________________________ 21
List of tables

Table 4.1: Results of the protection of inosine ................................................................. 24
1 - Introduction

1-1: Role of tRNA and synthesis of aminoacyl adenylates

Translation of the genetic code and protein synthesis from amino acids takes place in the ribosomes where tRNAs act as key intermediates. Correct translation of genetic information into the specific amino acids is dependent on aminoacyl-tRNAs. The synthesis of these is catalyzed by aminoacyl-tRNA synthetase. The aminoacylation reaction, the so-called charging of tRNA, is carried out in two steps. The first step is the formation of aminoacyl adenylates by a specific reaction involving amino acids, adenosine triphosphate, and an aminoacyl-tRNA synthetase. The amino acids are thus activated, and subsequently transferred to their corresponding tRNA through a highly specific reaction catalyzed by the same enzymes.

It is known that Aminoalkyl adenylates are inhibitors of these enzymes and are thought to occupy the binding sites for the aminoacyl adenylates. A drawback is that Aminoalkyl adenylates are sensitive to biological degradation. Adenosine 5’-(aminoalkyl phosphonates or phosphinates) can prove valuable as more stable inhibitors of amino acid activation and as substrates for studies of enzyme – substrate interaction.

1-2: Goals

The main goals were to develop methods for the synthesis of different new analogues of aminoacyl adenylates. The methods have to be reproducible and the obtained products have to be characterized by current analytical methods, like mass spectrometry and nuclear magnetic resonance.

1-3: Outline of this thesis

This report will present my participation in that very interesting project. The first part is a short introduction. The second part contains a general presentation of the background to permit the understanding by the non-initiated readers. Then techniques and methods used during the project will be described in the third part. The results that I obtained will constitute the next part and fourth part of the thesis. To conclude, a discussion about the results follows and an opening to further studies is suggested.
2 – Biological background

Genetic information in living cells is encoded in nucleic acid biopolymers, i.e. DNA and RNA. The translation of the genetic code and the protein synthesis from amino acids take place in ribosomes where tRNAs act as key intermediate.

A short general background in molecular biology is presented in this chapter. The end of the chapter is presented the specifics of the project.

2-1: The molecular building blocks: DNA, RNA, Amino acids and Proteins

The polymers DNA and RNA are the most common nucleic acids. Nucleic acids consist primarily of four different components, i.e. nucleosides linked together in a specific order through (3’ -> 5’) phosphodiester linkages, which constitute the backbone of the polymers. Furthermore each nucleoside contains a base and a sugar moiety (2’ deoxyribose in DNA and ribose in RNA). The heterocyclic bases in DNA are Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). For RNA, bases are the same except Thymine, which is replaced by Uracil (U).

![Figure 2.1: The different DNA and RNA bases](image)

It is interesting to note that the phosphodiester linkage gives directionality to the backbone. This directionality will guide the different molecular processes, which are responsible for copying and reading of the DNA and RNA in cells.
DNA usually consists of 2 complementary strands, which are held together by hydrogen bonds (see figure 1.3). The 2 strands are then twisted to form a double helix, which gives a very robust structure that can accommodate any sequence of nucleotides. RNA is most typically single stranded but often forms complex structures that also includes folding back of the strand to form double helical pairing with itself. RNA is more fragile than DNA, which is reinforced by the presence on the RNA of a free alcohol function (2’).

Proteins molecules are long unbranched polymer chains, constituted by monomeric building blocks drawn from a standard repertoire that is the same for most living cells. For the mammals, the standard repertoire consists of 20 different amino acids (compared to the 4 constituents of nucleic acids). Each amino acid is built around the same basic structure, which contains 2 connective groups (amino and carboxyl groups). A side chain is then attached. There are 20 common different side chains (see figure 2.2). Each side chain gives special chemical and physical properties to the amino acid.

![Figure 2.2: The 20 amino acids](image-url)
Protein molecules typically fold into a precise 3D structures, which is for example directed by hydrophilic and hydrophobic properties.

2-2: The dogma of molecular biology

The DNA in genomes does not directly direct the protein synthesis itself, but instead RNA is used as an intermediary molecule. When a particular protein is needed in the cell, the nucleotide sequence of the appropriate portion of the DNA is first copied into RNA (Transcription) in the nucleus of the cell. The obtained RNA copies of segments of DNA are after a maturation process used to direct the synthesis of the protein (Translation) in the cytoplasm of the cell. The flow of genetic information in all living cells, from the simplest bacteria to mammals, is therefore from DNA to RNA and from RNA to protein. As it is a universal pathway among living cells, it is known as the central dogma of molecular biology (see figure 2.3).

Figure 2.3: Central dogma of the molecular biology, the pathway from DNA to protein.

Despite its universality, there are important variations in the way information flows from DNA to proteins. For example RNA transcripts in eucaryotic cells are subject to a series of processing steps in the nucleus, like RNA splicing, before they are allowed to leave the nucleus. Such processing steps have of course a great influence on the meaning of the RNA. We can also remark that proteins synthesis is not always the only
use of RNA. RNA can be the final product after folding into a precise 3D structure. RNA can have structural, regulatory and catalytic properties in the cells.

2-3: The 2 steps of the genetic information flow

2-3.a: Transcription: From DNA to RNA

The RNA in a cell is coming from DNA transcription, which begins with the opening and unwinding of a small portion of the DNA double helix to expose the bases on each DNA strand. One of the 2 strands acts as a template for the synthesis of the RNA by complementary base pairing. During elongation, the incoming ribonucleotide is covalently linked to the growing RNA chain, by a phosphodiester bond. This reaction is catalyzed by an RNA transcription complex containing many different enzymes. Enzymes move along the DNA, unwinding the DNA helix just ahead of the active site and extending the RNA chain in the 5’ to 3’ direction, using high-energy triphosphated substrates (ATP, CTP, GTP and UTP). Since DNA and RNA are structurally similar, RNA can be directly synthesized using the DNA as base pairing template.

2-3.b: Translation: from mRNA to proteins

Once an mRNA has been produced in the nucleus, the information present in its nucleotide sequence can be used to synthesize a protein. In contrast to transcription, the translation of the information from RNA into protein represents a translation of the information into a different “language” that uses different structures. There are only 4 different coding nucleotides in RNA but 20 different amino acids in proteins. In the early 1960s, the genetic code, describing the relation between the 2 different “languages” was discovered. It is based on a 3 nucleotide association called codon. There are 64 possible codons, coding for 20 different amino acids. The code is said to be redundant: for example there are 4 different codons, coding for Valine (GAA, GAG, GAC, GAU) and there are 3 different stop codons. The genetic code is used universally in all present-day organisms. Only very few differences in the code have been found and only in the mitochondria. Mitochondria have their own transcription and protein synthesis systems that operate quite independently from those of the rest of the cell.

In principle, an RNA sequence can be translated in any one of the 3 different reading frames, depending on where the decoding process begins. Promoters are included in the sequence to help to select the good frame because only 1 of the 3 possible reading frames in an mRNA encodes the required protein.
2-4: Translation of the genetic code and protein synthesis

Translation of the genetic code and protein synthesis from amino acids takes place in the ribosome where tRNAs act as key intermediates.

2-4.a: The role of the tRNAs

The group of 3 nucleotides, which constitutes the codons, does not bind directly to the corresponding amino acid. Adaptor molecules are needed and these are tRNAs (t for Transfer RNA). Approximately 80 nucleotides long, tRNAs, like mRNAs, are able to fold into a specific structure. Due to 4 double helical segments, the tRNAs structure resembles a clover leaf. The structure is further folded by hydrogen bonds existing between the different regions of the tRNAs.

![The tRNA has a clover leaf form.](image)

Figure 2.4: Scheme of tRNA – example of tRNA (With a phenylalanine amino acid)

There are 2 unpaired regions, which are of special importance. The most important (“south”) is called the anticodon, which corresponds to the 3 complementary bases of the codon. The second region is the small single-stranded region (“north”) at the 3’ end, where the amino acid corresponding to the anticodon is attached to the tRNA.
The figure 2.5.a shows the initiation of the translation process. It begins with the small subunit of the ribosome, that binds to the mRNA. An initiator tRNA is attracted to the region (carrying a methionine), and binds to the triplet code AUG.

This then attracts the large ribosomal subunit, which will bind to the small subunit. Note that it has an A site and a P site. These are different binding sites for the tRNAs. The figure 2.5.c describes the next phase in the process.
The ribosome continues to read the code from the 5' to the 3' and amino acids are added to the growing peptide chain. This figure shows the tRNA carrying the glycine amino acid coded by CCA. Its complementary bases are GGU.

This continues until the stop codon is reached. This is highlighted in red in this figure and the next figure.

![Figure 2.5.c: Reading](image)

Figure 2.5.c: Reading

![Figure 2.5.d: Termination of the synthesis](image)

Figure 2.5.d: Termination of the synthesis

When the ribosome reaches a stop codon (red triplet), there is no tRNA attracted and the ribosome separates and leaves the mRNA.

The figure 2.5.d shows what happens when the stop codon is reached.
2-4.b: Synthesis of the tRNAs

As we have seen in the previous part, the tRNAs occupy a very important place in the synthesis of proteins. The next step of understanding, which will bring us directly to the content of our project, is the synthesis of these tRNAs, or more precisely the coupling of the RNA part with the amino acid. This coupling is shown in the next figure:

![Diagram of amino acid and RNA coupling](image)

*Figure 2.6: Coupling of amino acid and RNA to build a tRNA*

The release of monophosphate of ATP provides the necessary energy for the first step of the coupling, the adenylation of the amino acid. Both steps occur only in the presence of a special family of enzyme, which will catalyze the second step selectively depending on the anticodon (south part of the tRNA on the figure 2.4). There are as many different aminoacyl-tRNA synthetase enzymes as amino acids and each enzyme catalyzes only one type of coupling. It is interesting to note that adenine is the only base among the 4 possible, which is used during this reaction.

2-5: The interest of the project

The goal of the project is to develop synthesis routes for specific analogues of the adenylated amino acid.

The analogues are thought to replace the natural adenylated amino acids in the coupling with RNA to block the aminoacyl tRNA synthetase. By acting as an inhibitor the analogue will also block or slow the total protein synthesis. Many therapeutical applications can be found for such compounds.
Previously synthesized analogues of adenylated methionine, isoleucine and histidine have been made. Although some activity was observed, there was a need for improved efficacy. One hypothesis was that efficacy was limited by degradation of the analogues by phosphodiesterases. To try to minimize the degradation, new molecules were designed (see figure 2.7.c). The changes should prevent the hydrolysis but allow the molecule to act as an inhibitor.

The 2 initial target molecules present on figure 2.7.c are aimed to be synthesized. The method development should also be directed to synthesis of a third analogue that must be totally resistant to phosphodiesterase.

Figure 2.7.a: The natural adenylated glycine

Figure 2.7.b: The first synthesized analogues

Figure 2.7.c: The 2 molecules targeted to be synthesized

Figure 2.7.d: The 3rd molecule that could be synthesized after development of the synthetic methods.

target 1

Figure 2.7.d: The 3rd molecule that could be synthesized after development of the synthetic methods.
3 - Methods and techniques

The synthesis of the target molecules requires appropriate synthetic routes and different analytical methods to verify and characterize the compounds.

3-1: The synthetic strategy

To obtain the target molecules, we use different strategies (see appendix I – III). From the starting material, we first have to protect the sensitive functions, which should remain unreacted.

3-1.a Protection reactions:

To synthesize our analogues of adenosine, it is necessarily to protect the functions that should not react. On the adenosine and inosine, the alcohol functions situated in 2’ and 3’ positions are nearly as reactive as the 5’ alcohol functions and it will be difficult to perform the specific reactions on only the 5’ function. So the 2’ and 3’ positions have to be protected.

Furthermore, the amino functions, on the adenine base or on the aminopropanol, are reactive and are preferably protected.

- Protection of the 2’ and 3’ alcohol function on the ribose

\[ \text{Base: Adenosine} \]
\[ \text{Base: Inosine} \]

Figure 3.1: Scheme of the protection of the 2’ and 3’ alcohol functions
This reaction was important for the project because it allows us to “reach” the targeted function (to perform particularly the required reactions) on the molecule. The reaction had already performed with adenosine but it had never been done with inosine in our lab.

- Protection of the amino function on adenosine:

![Scheme of the protection of the amino function on the base of adenosine](image)

*Figure 3.2: Scheme of the protection of the amino function on the base of adenosine*

The amino function situated on the base in adenosine is reactive. To avoid side reactions on this function, we protected it with the described reaction. This reaction is also a reaction performed many times \(^1\) in our lab.

- Protection of the amino function of the amino acid analogues:

![Scheme of the protection by Z group of the amino function](image)

*Figure 3.3: Scheme of the protection by Z group of the amino function.*

The protection of the amino function on amino acids and their analogs is a widely used reaction, well known and easy to perform \(^1\). This protecting group for protection of the amino alcohol is called the Z group.

**3-1.b Carbon extension:**

The elongation of a molecule (to add 1 carbon unit to the 5’ extremity) can be done in different ways. This was a major difficulty of the synthesis of the target
molecule 1. To get our target, the “cyanide” pathway was chosen. This pathway contains 4 steps that will be described.

- The tosylation of the protected adenosine/inosine

![Figure 3.4: Scheme of the tosylation of the alcohol function](image)

Since OH does not leave from ordinary alcohols, it must be converted to a group that does leave much more easily. The chosen way is the conversion into a reactive ester, more precisely a sulfonic ester. The sulfonic ester groups are good leaving group. The para-toluene sulfonyl chloride is the most common reagent to perform this reaction. There are of course many possible conditions but the solubility properties of the starting material and the release of HCl during the reaction made us choose pyridine as solvent.

- The cyanide reaction

![Figure 3.5: Scheme of the replacement of the tosylate function by a cyanide function](image)

This reaction is the difficult step of the synthesis of target 1. We tried 2 different ways (1 and 2 with some variations of the temperature and the dryness of the solvent). According to reference 7, the first conditions should not give the cyclonucleoside byproduct (see section 4.1).
The reduction into an amine of the cyanide

This reaction is a reduction and according to the reference 9, Raney nickel is the best catalyst to perform this reaction.

The transformation of amine into an alcohol

The transformation of an amine into an alcohol can be performed following the procedure described in the reference 9.

3-1.c Synthesis of phosphinate:

The reactions chosen for our synthetic plan are known and previously used in the lab but only for quite different molecules. The reaction consists of 3 steps. First a better leaving group should replace an alcohol function. There are 2 different groups that are of choice: the triflate group or iodide. Then the phosphinate function replaces the leaving group. The introduced phosphinate can then be condensed with the alcohol function of another compound, and after that the phosphinate is oxidized to a phosphonate.
The phosphonylation of the alcohol function for target 1

![Chemical structure](image1.png)

*Figure 3.8: Scheme of the phosphonylation of protected inosine or adenosine.*

The phosphonylation of the alcohol function for target molecule 2

![Chemical structure](image2.png)

*Figure 3.9: Scheme of the 2 possible pathways to the phosphinate*

To make the phosphinate from the N\textsuperscript{\textalpha} -benzyloxycarbonyl-aminopropan-1-ol, there are 2 different possible ways. The 1\textsuperscript{st} way is using the same both steps as in the target molecule 1 synthesis. The alcohol function is in both cases replaced by a good leaving group: the triflate group in the 1\textsuperscript{st} way, by iodine in the 2\textsuperscript{nd} way. BSTP is then used in both cases to obtain the desired phosphinate. A base is needed in both cases to absorb the release of protons. The base is in the 1\textsuperscript{st} way the 2,6-ditertbutylpyridine, in the 2\textsuperscript{nd} way the Hünig base (diisopropyl ethylamine).
The coupling and the oxidation: synthesis of the phosphonate (example for target molecule 2)

![Diagram of coupling and oxidation reaction]

Figure 3.10: Scheme of the coupling and oxidation reaction.

The coupling and oxidative steps are quite well known steps, which were already performed and studied in the lab by a former PhD student, Anna Winqvist. She particularly studied the different coupling agent: OXP (\(N,N\)-bis (2-oxo-3-oxalizidin-1-yl) phosphinic chloride) was the most promising, giving at near equimolar amount very good coupling rate in a very short time and with no detected side reaction. The oxidative step can be performed according a procedure, that already was shown to be successful.

3-1.d Deprotection reaction:

The deprotection consists of several well-known reactions. These reactions can all be accomplished in two deprotection steps.

![Diagram of deprotection step]

Figure 3.11: Scheme of the deprotection step, example with target molecule 2.

The deprotection is accomplished in “mild” conditions, which means that the different functions should not be affected.
3-2: Main analysis and purification techniques used during the project work

3-2.a: Chromatographic methods:

- Thin layer Chromatography on silica (SiO$_2$)

Principle: The TLC is using the difference in affinity and thus mobility of the compounds in the mixture toward the eluent (mix of 2 or more solvents) and the silica layer. To perform a TLC, we just need a 3 cm * 6 cm plate and a container. The eluent in the container will “pull” the compounds along the plate. With organic solvents, solutes are exchanging between the solvent and the solid support.

![Scheme of Thin Layer Chromatography](image)

The TLC is a very practical, fast and easy method to monitor a reaction. But TLC has weaknesses. Good conditions (eluent, revelation, references…) have to be found and well reproduced to avoid problems.

- Chromatography on column:

The chromatography on column is a purification method using the same principles as TLC. The fixed solid phase is silica.
3-2.b: Spectroscopic methods

- **NMR**

  NMR is the reference method, widely used. In the lab, I had the chance to have ready access to a 400 MHz instrument.

  NMR is using the fact that the nuclei of all elements carry a charge. When the spins of the protons and neutrons comprising these nuclei are not paired, the overall spin of the charged nucleus generates a magnetic dipole, and the intrinsic magnitude of this dipole is a fundamental nuclear property called the nuclear magnetic moment, $\mu$. The symmetry of the charge distribution in the nucleus is a function of its internal structure and if this is spherical (i.e., analogous to the symmetry of a $1s$ hydrogen orbital), it is said to have a corresponding spin angular momentum number of $I=1/2$, of which the most common and used examples are $^1H$, $^{13}C$, $^{15}N$, $^{19}F$, $^{31}P$ etc.

  By applying an externally magnetic field $B_0$, the magnetic moment will partially align but not completely. By measuring the transitions between the different levels, it is then possible to determine the chemical structure of the analyzed compound.

- **Time of flight mass spectroscopy (TOF MS):**

  Time of flight mass spectroscopy is a very powerful method to determine the mass of the compounds. The principles of the mass spectrometer are the following: the substance is ionized and accelerated towards a detector. The time for the substance to hit the
detector is the converted to m/z ratio (mass/charge) values as the times of flights are dependent on mass and change.
4 - results

During the work in the lab, I performed many reactions, and the general knowledge about the compound family increased. I did manage to reach several steps into the route towards the target. Certain steps turned out harder and trickier than expected. There were 3 major difficulties during my project:

- Cyclisation of adenosine during the cyanide addition
- Difficulties with inosine derivates to develop the protection steps
- High apparent instability of the triflate

4-1: Results concerning the route to target molecule 1:

Target molecule 1 was first intended to be prepared according to the synthetic route, which is included as appendix I. There are however clues, confirmed in literature, that there may be problems with cyclisation of adenosine. So after trials always leading to the cyclic form without any visible desired product, we decided to follow another road, synthetically equivalent, but using inosine. Inosine is less sensitive to cyclisation and it is rather straightforward to transform the inosine to adenosine, in this case after the carbon extension.

4-1.a: Protection of adenosine.

- Protection of the amino-function on the adenine base

The adenosine was dried by evaporation of added pyridine 3 times. Adenosine (4,12g, 1 eq) was suspended in 50 ml of dry pyridine and trimethylchlorosilane (16ml, 8,2 eq) was slowly added; After 15 min, benzoic anhydride (4,58g, 1,3 eq) was added and the reaction was stirred for 2 hours at room temperature. The mixture was then cooled down with an ice bath, before adding 15 ml of water. After 15 min of stirring, 23,5 ml of aqueous NH$_3$ was added and the mixture was stirred for 80 min and then evaporated to dryness. 130 ml water and 100 ml of dichloromethane were added and stirred for 20 min. The mixture was filtered and dried overnight. The aqueous phase was reextracted with dichloromethane, filtered and dried. The combined crude product was 4,18g. Yield: 73%.

- Protection of the 2’ and 3’ alcohol functions

$N^6$-benzoyl Adenosine (3,20 g, 1 eq) after drying under vacuum overnight. was added to an oven dried flask. Benzaldehyde (13 ml, 14,9 eq) was added and after 5 min of stirring to dissolve the adenosine, zinc chloride (4,65g, 4,0 eq) was added. The reaction was stirred at room temperature for 72h. The reaction was followed with mass spectrometry and TLC. After 72h, excess benzaldehyde was evaporated and the obtained foam was extracted with ethyl acetate (3x200 ml). The organic phase was dried with MgSO$_4$ and concentrated. The product was then chromatographed on silica with
dichloromethane/methanol (95/5). Yield: 71% of a slightly yellow product that was containing minor impurities.

- The tosylation step:

\[ \text{N}^6\text{-benzoyl-2',3'-O,O-benzylidene adenosine (1g) was dried 2 times by evaporation of added dry pyridine and then dissolved in dry pyridine (3 ml). The mixture was cooled down to 0°C (ice bath) and para-toluenesulfonyl chloride (0,622g), previously dried under vacuum overnight and dissolved in 3 ml of dry pyridine, was added dropwise. The reaction was stirred overnight at room temperature. The mixture was evaporated to near dryness and the obtained orange oil was extracted with 150 ml of ethyl acetate/dichloromethane (1:1) and 30 ml of saturated NaCl solution. The organic layer was extracted with 50 ml of HCl (2 M) and then dried with Na}_2\text{SO}_4 \text{ and concentrated under reduced pressure. The product was then chromatographed on silica using as eluent Ethyl acetate/dichloromethane (1:2). 0,69 g of product was obtained. Yield: 52%. This experiment was actually realized by Merita Murtola with product I isolated.} \]

- The cyanide step:

The replacement of the tosylate group by cyanide ion was tried by Merita Murtola under different conditions/
- in DMSO
- in dry DMSO
- in dry DMSO with heating

In all these different conditions, the cyclonucleoside is the predominant product. There were no traces of desired product. It is known that adenosine derivatives are sensitive to cyclisation. There are also examples where cyclisation is not a problem. However under the conditions used, the cyclisation side-reaction is clearly promoted. The solvent polarity is likely to be a major reason for this. We then decided to use a similar route with less cyclisation-prone inosine (see figure) as starting point. We also opted for another protocol with an apolar solvent in order to suppress the cyclisation further. Thus, this synthetic route was abandoned but the samples of double protected product are kept stored in the freezer for other purposes and possibly alternative conditions to be tested later.

*Figure 4.1: cyclonucleoside*
The inosine and adenosine are 2 quite similar molecules (the only difference being one substituents on the base). The protocols could have been directly transposable from one to another. However it turned out that this was harder than expected. There were problems with solubility at several stages, which also proved to be problematic upon purification. So there were quite some changes to make. Several adaptations or changes were made to provide a material of good quality.

- Protection of the inosine: 2',3’-O,O-benzylidene-inosine

Inosine (1.56 g, 1 eq) was added to a dried flask. Benzaldehyde (8.40 ml, 14.2 eq) was added and after 5 min of stirring to dissolve the inosine, zinc chloride (3.10 g, 3.9 eq) was added. The reaction was stirred under argon at room temperature for 72 h. The reaction was followed with mass spectrometry and TLC. After 72 h, ethyl acetate (25 ml) was added slowly (to avoid precipitation) and then 25 ml of a solution of Na$_2$CO$_3$ (15% in water) was added. After filtration, the organic layer was separated and extracted with 25 ml of saturated NaCl aqueous solution. The organic phase was dried with MgSO$_4$ and concentrated. The product was then chromatographed on silica using a stepwise gradient of methanol in dichloromethane (5 → 20%). Yield: 46%. The product was then used to prepare the tosylated 2',3’-O,O-benzylidene-inosine. The structure had been checked with H-NMR and MS.

- Tosylation of the 2’,3’-O,O-benzylidene-inosine: 5’-tosyl-2’,3’-O,O-benzylidene-inosine ...

2’,3’-O,O-benzylidene-inosine (1.86 g, 1 eq) was dried 2 times by evaporation of added dry pyridine and then dissolved in dry pyridine (28 ml). The mixture was cooled down to 0°C (ice bath) and para-toluenesulfonyl chloride (1.58g, 1.58 eq), previously dried under vacuum overnight and dissolved in 4 ml of dry pyridine, was added dropwise over 20 min. The reaction was stirred overnight. An additional portion of para-toluenesulfonyl chloride (0.60g, 0.60 eq) was added drop wise after dissolution in 3 ml of dry pyridine. The reaction was stirred for another 24h. The mixture was evaporated to near dryness and the obtained orange oil was extracted with 125 ml of ethyl acetate and 30 ml of saturated NaCl solution. The organic layer was extracted with 50 ml of HCl (2 M) and then dried with MgSO$_4$ and concentrated. The product was chromatographed on silica using as eluent chloroform/ethyl acetate/methanol (12/7/1). 0.73 g of product was obtained. Yield: 27%. The structure had been checked with H-NMR and MS.

Concerning this synthesis, the eluent is not perfect and should be optimized (there are some unknown impurities in the product). The first eluent used was dichloromethane/ethyl acetate but the separation was worse either this.

- Trials to obtain the cyanide:

A trial in dry DMSO was performed, but the cyclonucleoside was the predominant product. So we decided to follow a different procedure. In this procedure...
18-crown-6-ether is used to capture the K+ cation. The less polar obtained solution should avoid the formation of the cyclonucleoside.

After drying overnight, tosylated inosine (0.23g, 1 eq) was dissolved in 5.5 ml of dioxane and with 18-crown-6-ether (0.05g, 0.42 eq). Potassium cyanide (0.11g, 3.75 eq) was added. After 24h, the mixture was filtered and the solid was washed 3 times with 1.5ml and the same amounts as initially of cyanide potassium and 18-crown-6-ether were added. The procedure was repeated after 24h. After 72h of reaction; dioxane was evaporated. TLC, MS and NMR were performed. The desired product was not seen on the MS or on the NMR, but TLC showed, there was a product different from starting material. The cyclonucleoside (M$^+$ = 339 g.mol$^{-1}$) was not detected on mass in these conditions. Further experiments and analysis should be performed.

Conclusion on target molecule 1:

The different difficulties due to the cyclisation of adenosine and the difference in solubility between adenosine and inosine were finally solved, even if some changes on the tosylation step are maybe needed in order to increase the purity and yield. I did not manage to perform the complete extension steps (only 2 trials on very small scale) but it seems that the conditions using dioxane and 18-crown-6- to capture the K$^+$ cation are promising but further developments are needed.

4-2: Results concerning the route to target molecule 2:

- Protection of the 2’ and 3’ alcohol functions: 2’,3’-O,O-benzylidene-adenosine

Adenosine (3.0 g, 1 eq) was added to a previously dried flask. Benzaldehyde (17 ml, 14.9 eq) was added and after 5 min of stirring to dissolve the adenosine, zinc chloride (6.2g, 4.1 eq) was added. The reaction was stirred at room temperature for 48h. The reaction was followed by TLC. After 48h, the benzaldehyde was evaporated and the obtained foam was extracted with ethyl acetate (3x200 ml). The organic layers were combined, dried with MgSO$_4$ and concentrated. The product was then chromatographed on silica using a stepwise gradient of methanol (0 to 10%) in chloroform. Yield: 56%.

- Protection of 3-aminopropan-1-ol: N-benzyloxycarbonyl-aminopropan-1-ol

3 aminopropan-1-ol (5.40ml, 1eq) was added to a previously dried flask containing 67 ml of an aqueous solution of Na$_2$CO$_3$ (3.2 mol.L$^{-1}$, 3.0 eq) and 50 ml of dioxane. The solution was cooled down with an ice bath. Benzyl chloroformate (11.90 ml, 1.2 eq) was added and the solution was stirred for 1 h at 0-2°C. The solution was neutralized with KH$_2$PO$_4$ to pH~7. After neutralization, the solution was extracted with 2x200 ml and 2x100 ml of ethyl acetate. The organic layers were combined, dried with MgSO$_4$ and concentrated. The product is then chromatographed on silica with chloroform/methanol (19/1) as eluent. The reaction was performed 2 times and the results are summarized in table 4.1:
Table 4.1: Results of the protection of inosine

<table>
<thead>
<tr>
<th>Alcohol</th>
<th>Alcohol eq</th>
<th>Benzyl chloroformate</th>
<th>BzCl eq</th>
<th>Na₂CO₃</th>
<th>Na₂CO₃ eq</th>
<th>Mprodut</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.02 g</td>
<td>1</td>
<td>14.23 g</td>
<td>1.25</td>
<td>67 ml</td>
<td>3.21</td>
<td>9.03</td>
<td>64%</td>
</tr>
<tr>
<td>5.30 g</td>
<td>1</td>
<td>14.22 g</td>
<td>1.18</td>
<td>67 ml</td>
<td>3.03</td>
<td>6.43</td>
<td>43%</td>
</tr>
</tbody>
</table>

- Synthesis of the N benzylxocarbonyl-3-amino propyl triflate

This synthesis was tried 7 times. I tried 2 times to isolate it but it seems it was a very unstable product so it was decided to directly use it in the next step (synthesis of the phosphinate).

\[ N^\alpha\text{-benzyloxycarbonyl-aminopropan-1-ol} \ (0.25g, \ 1 \text{ eq}) \text{ and } 2,6\text{-diterbutyl pyridine} \ (0.40 \text{ ml, 1.49 eq}) \text{ were dissolved in dry dichloromethane (15 ml) to a previously dried flask. The solution was cooled down to -78°C with an ethanol/dry ice bath under nitrogen flux. A solution of trifluoromethanesulfonic anhydride} \ (0.30 \text{ ml, 1.49 eq}) \text{ in dry dichloromethane (7 ml) was added dropwise over 15 min. The solution was stirred for 2 h at -78°C. The solution was then allowed to reach 0°C before being used for the next steps.} \]

- Synthesis of the N benzylxocarbonyl-3-amino-1-iodo propan

\[ N\text{-benzyloxycarbonyl-aminopropan-1-ol} \ (0.52g, \ 1 \text{ eq}) \text{ was added to a previously dried flask and dissolved in 20ml of dry acetonitrile/pyridine (95:5). The solution was cooled down to 0°C with an ice bath under argon. Triphenylphosphine} \ (0.98 \text{ g, 1.50 eq}) \text{ and iodine} \ (0.92g, \ 1.46 eq) \text{ were added and the solution was left on the melting bath. The mixture was stirred for 30h. TLC were performed to follow the reaction. The solution was extracted with 70 ml dichloromethane and 2x50 ml and 2x30 ml of a 0.5M Na₂SO₃ solution, containing also 5% of NaHCO₃. The organic layers were combined, dried with MgSO₄ and concentrated. The product is then chromatographed on silica with toluene/Ethyl acetate (9/1) as eluent. 0.34 g of product was obtained. Yield: 43%. The structure had been checked with H-NMR and MS.} \]

- Synthesis of bis(trimethylsilyl)phosphine (BSTP)

Ammonium hypophosphate was dried over P₂O₅ under vacuum. Hexamethyldisilsilazane (7.6 ml, 1.01 eq) was added to a previously dried flask which was kept for 1h under argon. Ammonium hypophosphate (3.05 g, 1.00 eq) was then added. The solution was stirred under nitrogen flux for 2h at 105°C. The solution was then cooled down to 0°C with an ice bath. The nitrogen flux was redirected to get rid of NH₃. BTSP was then used for the different trials of synthesis of phosphinate.

- Synthesis of the N benzylxocarbonyl-3-amino propyl phosphinate triethylammonium salt.

The synthesis of the phosphinate was tried many times following 2 procedures. The product was never formally identified, but there are good chances that the desired
product can be synthesized. Furthermore the BTSP is a very dangerous reagent, which can ignite in contact with oxygen.

✓ From the N benzyloxy carbonyl-3-amino propyl triflate

In the flask containing N benzyloxy carbonyl-3-aminopropyl triflate, 2,6-diterbutylpyridine (0,35 ml, ) was added. BSTP was added dropwise over 10 min while reaction was stirred on a melting ice bath and then allowed to reach RT while stirring overnight. The reaction was quenched by addition of 15ml of dichloromethane and 0,3ml of a triethylammonium bicarbonate solution (2M). The solution was evaporated and then suspended in chloroform and filtered. The chloroform layer was evaporated and dried under vacuum. The yellow oil obtained was chromatographed on silica with dichloromethane/methanol (19/1) as eluent.

There were many interesting trials but the desired product was never formally isolated and identified.

✓ From the N benzyloxy carbonyl-3-amino-1-iodo propan

This reaction was tried only once due to difficulties of obtaining a new batch of starting material.

N benzyloxy carbonyl-3-amino-1-iodo propan was added to a previously dried flask under argon with the Hünig base (1ml, 6,2 eq) in dry acetonitrile (14 ml). The solution was cooled down with an acetonitrile/dry ice bath (42°C). After 15 min of stirring under argon, BTSP (2ml, 8mmol, 8,5 eq) was added dropwise over 10 min and the solution was stirred overnight on the melting bath. The reaction was quenched by addition of 15ml of dichloromethane and 0,3ml of a triethylammonium bicarbonate solution (2M). The solution was evaporated and chromatographed on silica with dichloromethane/methanol (9/1) as eluent.

The TLC was showing many spots, which could be the desired product, but finally the collected fractions were only containing the Hünig base. The desired product may have been trapped on the column.
5 – Discussion and further studies

5-1: Discussion

There are many reactions, which were performed and our general knowledge about the compound family increased. The reactions that I managed to perform are mainly the protection reactions. Protection reactions were quite easy to perform for the target molecule 2 but there were much more difficulties for the target molecule 1. For the target molecule 1, the difficulties appeared with the change of starting material: from adenosine to inosine. The protection steps were well known with adenosine but it appears much more complicated with inosine. The protocols had to be adapted in order to obtain the desired product. The complication came from the lower solubility of inosine and its derivate. This difference in solubility could perhaps be used to purify the product by recrystallisation.

After the protection step, the carbon extension steps were another difficulty that I only partially could success with. The first step of the extension is replacement of the alcohol function by a tosyl function. This reaction was performed with a quite low yield and the purification was not 100% efficient. Some changes have to be done in order to enhance yield and purity. The replacement of the tosyl by a cyanide group was tried many times in different conditions but the desired product was not isolated. There was no directly interpretable evidence of the presence of the product but the cyclonucleoside, the most common side product was also not detectable in the last conditions we choose. This seems to be an interesting and promising approach that has to be analyzed further.

Concerning the target molecule 1, the replacement of the alcohol function on the N-benzyloxy carbonyl-aminopropan-1-ol by a better leaving group was tried in 2 different ways. The most tried was the replacement by a triflate group. The compound obtained seems to be really unstable so it was never properly isolated but rather directly used in the next step: the phosphonylation. On the other side, the iodo compound was isolated, purified and characterized with a correct yield. The phosphonylation of the iodo compound was only tried once. The phosphonylation was tried many times (7 times from the triflate group and 1 time from the iodo group) but the desired product was never properly detected. There were several indications that the desired product could be present. Some more work is needed in order to analyze the product and this reaction.

5-2: Suggestions for further work

The work done in the project is part of a collaboration with the cancer center at KI (CCK).

The protection reactions for the target molecule 2 can be done in good conditions and no further work is normally needed. For the target molecule 1, the change from adenosine to inosine was more complicated than expected due to the difference in solubility. The lower solubility of inosine and its derivates could be an interesting way to separate the desired product from the starting material and the other reagents. Some assays were done on the 2’,3’-O,O-benzylidene-adenosine and on the 5’-tosyl-2’,3’-O,O-
benzylidene-inosine. There were some interesting results (for both steps) but further work is needed, particularly in term of analysis to check all products.

Concerning the carbon extension steps and especially the replacement of the tosyl group by cyanide, only 2 small scale trials were performed. The conditions used in the last trial were particularly interesting (KCN with 18-crown-6-ether) as the cyclonucleoside was not detected. Some more trials has to be performed first by changing some parameters like the length and the number of cycle. In reference 7, it was advised to perform 3 times 20h of reaction, changing every time the 2 reagents. A first experiment would be to try to perform a single cycle. In these conditions, it should be easier to isolate the desired product. There is no risk to loose our product during the filtration. If the desired product was detected, it would be possible then to increase the number of cycles to try to increase the yield. Other further trials could also be to change the solvent (dioxane) for a less polar one or to change the reagent (KCN) for one with a bigger cation.

Concerning the phosphonylation (target molecule 2), it would be interesting first of all to try to do it again with the iodo pathway, as the iodo compound is much more stable. With this method the phosphonylation could be performed in 2 steps with a stable intermediary product, which seems not to be the case with the triflate pathway. In general some more trials need to be performed to know if the reaction is reasonable and to isolate some amount of the product. The next steps (coupling and oxidation) should be then rather easy to perform.

There are quite a few experiments that should be performed before the desired product could be obtained. Many indications made us think that I was quite close to obtain positive results in the cyanide and the phosphonylation but unfortunately there was not enough time left to pursue this in full. The steps after these steps are expected to be easier, but as chemistry is surprising…
References

Scientific Publications:

1. E Yousoufi-Salakdeh, M Murtola, A Zeterberg, E Yekesiely, R Stromberg, Synthesis of 8-aminoadenosine 5’-(aminoalkylphosphates), analogues of aminoacyl adenylates, manuscript

Thesis of former PhD student in the organic and bioorganic laboratory:

1. A Winquist, in studies towards a method for incorporation of 3’-deoxy-3’-C-methylene-phosphonate linkages into oligonucleotides (particularly paper II)
2. A Slaitas, in development of a new PNA analogue as a potential antisense drug and tool for life science studies
3. S Zaramella, in development of a novel methodology for the synthesis of oligonucleotide-peptide conjugates

Books:

Appendix I: scheme of the first strategic route for the target molecule 1 (using adenosine as starting material)
Appendix II: scheme of the second strategic route for the target molecule 1 (using inosine as starting material)
Appendix III: scheme of the strategic route for the target molecule 2

\[
\begin{align*}
1) & \quad \text{NH}_3/\text{MeOH} \\
2) & \quad \text{H}_2, \text{Pd} \quad \text{(coupling agent)} \\
\end{align*}
\]

in pyridine/water (96:4)

\[
\begin{align*}
1) & \quad \text{NH}_3/\text{MeOH} \\
2) & \quad \text{H}_2, \text{Pd} \\
\end{align*}
\]

in pyridine/water (96:4)

1) 1\text{h}, 0°C

2) 2\text{h}, -78°C

3) 2\text{h}, -78°C

4) 2\text{h}, -78°C

5) 2\text{h}, -78°C

6) 48\text{h}, \text{RT}

7) 48\text{h}, \text{RT}

Adenosine