Design of Thymidine-based Inhibitors of *Mycobacterium tuberculosis* Thymidylate Kinase

Apr. Ineke Van Daele

Thesis submitted to the Faculty of Pharmaceutical Sciences in order to obtain the degree of Doctor in Pharmaceutical Sciences

Promoter

Prof. dr. apr. S. Van Calenbergh

Academic year 2007-2008

Table of contents

Dankwoord

Nu ik toegekomen ben tot het schrijven van mijn doctoraat, kijk ik eens even terug naar de jaren die me tot hier hebben gebracht. Het waren 4 leuke en uitdagende jaren voor mij, niet altijd even gemakkelijk, maar ik heb geen moment spijt gehad dat ik eraan ben begonnen. En dit dankwoord mag dan misschien een 'verplicht' onderdeel zijn van dit werkje, toch zou ik een paar mensen oprecht willen bedanken.

Prof. Serge Van Calenbergh, mijn promotor, bedankt om me te verwelkomen in uw labo en me de kans te geven een doctoraat te volbrengen, om me te begeleiden en te stimuleren om zelfstandig te werken. Bedankt om me keer op keer op mijn smospoes-gedrag te wijzen, het begint stilaan zijn effect te hebben hoor.

Bedankt ook aan mijn collega's, Izet, Veerle, Philippe, Ulrik, Helga, Timo, Liesbet, Stijn, Steven, Matthias en Thomas. Jullie allemaal waren erbij als het leuk was, maar jullie waren er ook voor mij als het eens even wat minder ging. Merci, aan elkeen van jullie heb ik heb veel gehad. Veerle, bedankt om me te begeleiden toen ik hier als kersverse doctoraatstudente aankwam, voor de vele tips die je me gaf en om nadien als hulplijn altijd beschikbaar te zijn!

Thank you, Prof. Erik Bjerregaard Pedersen, to welcome me in your lab in Odense (DK) during my Marie Curie project. It was a very nice, interesting and stimulating period for me. I learned so many things, we discussed every problem we met. Thank you also to come to Belgium for my PhD defence. I am happy that we can meet again!

Thanks to my collegues in SDU, Slava, Daniel, Niels, Per, Theodora, Irina, Mads and so many others. Thanks for being my companions inside and outside the lab during my stay in Odense.

Een speciale bedanking verdienen ook de mensen die me geholpen hebben met allerlei zaken tijdens mijn onderzoek, Prof. Jef Rozenski voor het opnemen van de heel vele massaspectra, Mathy Froeyen, voor de modeling experimenten, Sofie Van Damme, voor de QSAR berekeningen, Dr. Hélène Munier-Lehmann, for testing my compounds for their biological activity en het Bijzondere Onderzoeksfonds van UGent voor de financiering. Bedankt, want zonder jullie had ik het niet tot hier gehaald!

Gert, jou zou ik ook heel speciaal willen bedanken voor de vele steun en de mooie jaren die je me gegeven hebt tijdens mijn doctoraat.

Mijn ouders, broer, zus en vrienden, bedankt voor de interesse en steun en om er altijd te zijn voor mij.

Summary

With the search for potent and selective inhibitors of *Mycobacterium tuberculosis* thymidine monophosphate kinase (TMPKmt), this PhD study describes the development of new anti-tuberculosis agents. As TMPKmt is an essential enzyme for the growth of the bacillus, inhibitors of this enzyme might be useful as anti-tuberculosis agents. The X-ray structure of the enzyme was determined in 2001. It shows the way of binding of the natural substrate, dTMP and was used to explain and predict the activities of modified thymidine derivatives.

As this work is a continuation of earlier research in our group, the structure-activity relationship derived from earlier synthesized compounds, was a useful tool for further development. The most interesting topics of the initial structure-activity relationship served as a starting point for this work.

The first chapter describes the influence of a F-substituent at the 2'-position on the conformational behaviour of the sugar ring and consequently, on TMPKmt activity. Synthesis of 2'-F-substituted thymidine derivatives allowed the biological evaluation and conformational analysis of those compounds.

A second important modification considers the discovery of two important new leads, a 2',3'-bicyclic nucleoside and a 3'-*C*-branched dinucleoside. Optimisation of those compounds led to a series of modified 2',3'-bicyclic thymidine derivatives and a series of 3'-*C*-branched thiourea substituted compounds. For those two new families of TMPKmt inhibitors, modelling, docking and conformational analysis are used to rationally design new compounds and to explain the observed activities. Finally this led to compounds with activities in the low micromolecular range with the most active compounds 1-(3-aminomethyl-3,5-dideoxy-2-*O*,6-*N*-(thiocarbonyl)-β-D-ribofuranosyl)thymine (K_i= 2.3 μM) and *N*-(5'-deoxy-α-D-thymidin-5'-yl)-*N'*-(3-trifluoromethyl-4-chlorophenyl)thiourea (K_i= 0.6 μM), with K_i-values, significantly lower than the K_m-value of dTMP (4.5 μM). Next to the good TMPKmt-affinities, those compounds showed a high selectivity index compared to the human enzyme and no significant toxicity. Moreover, for the first time, the most active TMPKmt-inhibitors showed inhibitory activities against growing *Mycobacterium bovis* and *Mycobacterium tuberculosis* strains.

An extension of the biological evaluation of modified thymidine derivatives in the field of antiviral agents, revealed the activity of 3'-*C*-branched thymidine derivatives on mitochondrial thymidine kinase 2.

The last chapter of this work describes the use of a 3'-*C*-branched thymidine derivative, obtained earlier as building block in oligonucleotide synthesis. As oligonucleotides have

found their application in DNA diagnosis, the fluorescent pyrene was attached to the 3'-position of thymidine, before this building block was incorporated in different oligonucleotides. Stability and fluorescence studies showed that those fluorescently labelled oligonucleotides are useful tools as nucleoside diagnostics. The fluorescently silent single stranded oligonucleotides emit their fluorescence after binding to a DNA duplex by Hoogsteen interactions.

Samenvatting

Met de zoektocht naar krachtige en selectieve inhibitoren van het thymidine monophoshate kinase van *Mycobacterium tuberculosis* (TMPKmt), beschrijft dit werk de ontwikkeling van nieuwe anti-tuberculose geneesmiddelen. Aangezien TMPKmt een essentieel enzyme is voor de groei van de bacterie, kunnen inhibitoren van dit enzyme verder ontwikkeld worden tot anti-tuberculose geneesmiddelen. De X-stralen structuur van dit enzyme is bepaald in 2001 and toont de bindingswijze van het natuurlijke substraat dTMP en werd gebruikt om de activiteiten van gemodificeerde thymidine derivaten te verklaren en te voorspellen.

Dit werk kadert in een eerder onderzoek in onze groep en bijgevolg kan de structuur-activiteitsrelatie die afgeleid werd uit vroeger gesynthesiseerde componenten gebruikt worden voor verdere ontwikkeling van nieuwe thymidine derivaten. De meest interessante topics van de reeds gekende structuur-activiteitsrelatie hebben gediend als startpunt van dit werk.

Een eerste belangrijke uitgangspunt betreft de invloed van een F-substituent op de 2'-plaats op de conformatie van het suikergedeelte, en bijgevolg ook op de biologische activiteit. Synthese van 2'-F gemodificeerde thymidine derivaten lieten de biologische evaluatie en conformationele analyse toe van deze verbindingen.

A tweede belangrijke modificatie volgde uit de ontdekking van twee belangrijke nieuwe lead-verbindingen, een 2',3'-bicyclische molecule en een 3'-C-vertakt dinucleoside. Optimalisatie van deze verbindingen gaf aanleiding tot een reeks van gemodificeerde 2',3'-bicyclische thymidine derivaten en een reeks van 3'-C-vertakte thiourea gesubstitueerde componenten. Docking, modelling en conformationele analyse werden gebruikt om uitgaande van deze deze twee nieuwe families van TMPKmt inhibitoren, op rationele wijze nieuwe componenten te ontwikkelen en hun activiteiten te verklaren. Uiteindelijk heeft dit aanleiding gegeven tot verbindingen met activiteiten in lage micromolaire range, zijnde 1-(3-aminomethyl-3,5-dideoxy-2-*O*,6-*N*-(thiocarbonyl)-β-D-ribofuranosyl)thymine (K_i= 2.3 μM) and *N*-(5'-deoxy-α-D-thymidin-5'-yl)-*N'*-(3-trifluoromethyl-4-chlorophenyl)thiourea (K_i= 0.6 μM), met K_i-waarden die significant lager zijn dan de K_m-waarde van dTMP (4.5 μM). Naast de goede affiniteiten voor TMPKmt, vertoonden deze componenten ook een hoge selectiviteit tegenover het humane enzyme en geen significante toxiciteit. Bovendien werden voor de eerste keer TMPKmt inhibitoren ontdekt, die activiteit vertonen tegenover groeiende *Mycobacterium bovis* en *Mycobacterium tuberculosis* stammen.

Een uitbreiding van de biologische evaluatie van de gemodificeerde thymidine derivaten naar antivirale activiteiten toe, toonden aan dat enkele 3'-C-vertakte thiourea gesubstitueerde thymidine derivaten een interessante activiteit vertoonden op het mitochondriale thymidine kinase 2.

Het laatste hoofdstuk van dit werk beschrijft het gebruik van een 3'-C-vertakte thymidine derivaat als bouwsteen in oligonucleotide synthese. Aangezien oligonucleotiden diagnostica hun toepassing hebben gevonden in het oligonucleotide onderzoeksdomein, werd pyrene via de 3'-plaats aan de nucleoside bouwsteen verbonden, alvorens deze werd ingebouwd in verschillende oligonucleotiden. Stabiliteits- en fluorescentie studies van triplexen gevormd door het gemodificeerde oligonucleotide dat bindt aan dsDNA, tonen aan dat de fluorescent gelabelde oligonucleotiden bruikbaar zijn als DNA diagnosticum.

List of Abbreviations

A	Adenosine
ACP	Acyl Carrier Protein
AIDS	Acquired Immuno-Deficiency Syndrome
ATP	Adenosine Triphosphate
AZT	3'-Azido-3'-deoxythymidine
BCG	Bacille Calmette Guérin
C	Cytidine
CFU	Colony-Forming Unit
CNS	Central Nervous System
COMBO-FISH	Combinatorial Oligonucleotide Fluorescence *In Situ* Hybridisation
CPZ	Caprazamycin
DARQ	Diarylquinolones
(ds)DNA	(double stranded) Deoxy ribo Nucleic Acid
DOTS	Directly Observed Treatment Short-Course
EXO	Exochelin
FAS	Fatty Acid Synthase
FTC	Emtricitabine, β-L-2',3'-dideoxy-5-F-3'-thiacytidine
G	Guanosine
HBV	Hepatitis B virus
HEPA	High Efficiency Particulate Air
HIV	Human Immuno-deficiency Virus
HTS	High-Througput Screening
INH	Isoniazid
LD_0	Letal Dose (with 0% of dead bodies)
LDH	Lactate Dehydrogenase
L-d4FC	β-L-2',3'-didehydro-2',3'-dideoxy-S-fluorocytidine
LNA	Locked Nucleic Acid
MDR	Multi-Drug Resistant
MIC	Minimum Inhibitory Concentration
mRNA	Messenger RNA
M. tuberculosis	*Mycobacterium tuberculosis*
MW	Molecular Weight
MYC	Mycobactin
NAD^+	Nicotinamide Adenine Dinucleotide (oxidised)

NADP$^+$	Nicotinamide Adenine Dinucleotide Phosphate (oxidised)
NDP	Nucleoside Diphosphate
NDPK	Nucleoside Diphosphate Kinase
NMP	Nucleoside Monophosphate
NMPK	Nucleoside Monophosphate Kinase
ON	Oligonucleotide
PEP	Phosphoenol Pyruvate
PK	Pyruvate Kinase
QSAR	Quantitative Structure-Activity Relationship
RNA	Ribo Nucleic Acid
SAR	Structure-Activity Relationship
SI	Selectivity Index
T	Thymidine
TB	Tuberculosis
TDPK	Thymidine Diphosphate Kinase
TFO	Triplex-Forming Oligonucleotide
TK	Thymidine Kinase
TMPK	Thymidine Monophosphate Kinase
TMPKh	Human Thymidine Monophosphate Kinase
TMPKmt	Thymidine Monophosphate Kinase of *M. tuberculosis*
U	Uracil
WHO	World Health Organisation
XDR	Extensively drug-resistant TB
3TC	Lamivudine, β-L-2',3'-dideoxy-3'-thiacytidine

Part I: Thymidine analogues as selective TMPKmt inhibitors

Chapter 1

Introduction

1   INTRODUCTION

1.1 GENERAL BACKGROUND

1.1.1 Tuberculosis on the rise

More than a century ago, tuberculosis (TB), an infectious disease usually affecting the lungs, was ranked among the most feared and dreaded of the numerous diseases afflicting mankind. The discovery of the causative agent of tuberculosis by Robert Koch in 1882 gave great hope that this 'white plague' would soon be vanquished.

During the 20th century new antituberculosis drugs were discovered. In the early 1980's the great majority of patients could be cured by treatment with short-course curative chemotherapy that was possible as a consequence of several newly introduced antituberculosis drugs.

Only one decade later, an outbreak of HIV-related **multidrug-resistant tuberculosis** in New York made tuberculosis suddenly the centre of attention. More than one century after Koch's discovery, far from celebrating the final conquest of tuberculosis, the World Health Organization (WHO) declared TB as a global emergency. Even in the beginning of the 21st century tuberculosis remains a major public health issue, killing over 1.6 million people annually.[1] Especially countries with poverty, like in Africa and Asia suffer from severe TB epidemics (Figure 1.1).
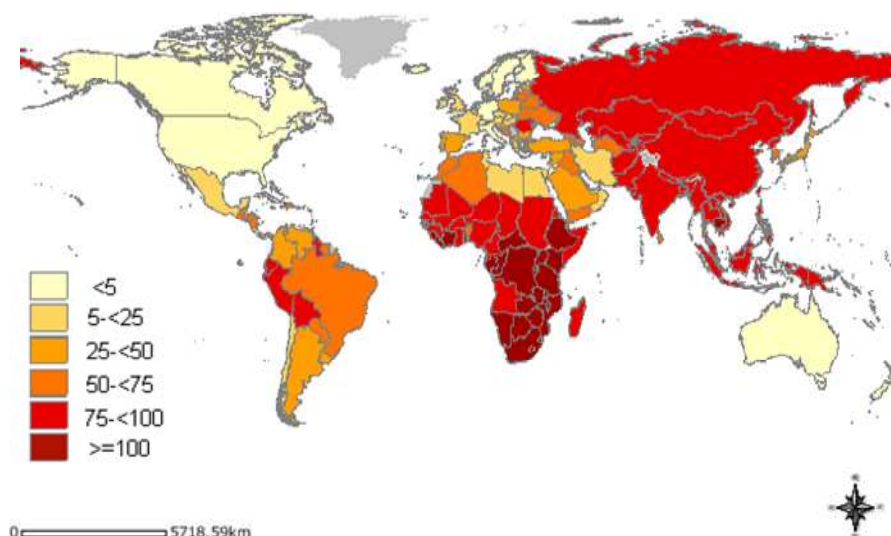
Figure 1.1 Estimated TB incidence per 100.000 inhabitants (2005)[1]

Mycobacterium tuberculosis, the causative agent of TB is a slow-growing bacillus, primarily transmitted via the respiratory system. The WHO estimates that one third of the world's population is infected by this organism, however usually this latent infection is not reactivated. HIV-infection is the major risk factor for converting latent TB into active transmissible TB. At the same time TB-bacteria assist in accelerating the progress of HIV-infection in the patient. Since both diseases combined are more destructive than each disease alone, the **TB- and HIV-epidemics form a vicious circle**. TB is the leading cause of death among people who are HIV-positive accounting for 13% of AIDS deaths per year worldwide.[1,2]

Nowadays TB can be cured with DOTS (Directly Observed Treatment Shortcourse) in which patients are subjected to standard treatment regimens that consist of **an exceedingly lengthy therapy of 6-9 months** under direct observation, often with a **cocktail consisting of three or four different drugs**.[3] Patients following this regimen become non-infectious after the first few weeks but the remaining treatment period is crucial to eradicate the slow growing bacilli and to allow the host immune system to achieve a clinical cure. The direct observation is needed because in addition to the significant toxicity, the lengthy therapy causes poor patient compliance, resulting in the selection of drug resistant and often deadly multi-drug-resistant (MDR)-TB bacteria.

For people infected with these MDR-TB strains, there is a **large need for more effective drugs acting on novel targets**.[1] The emergency for new TB-drugs is reflected in the amount of publications dealing with TB, which increased significantly during the last 20 years (See Figure 1.2).
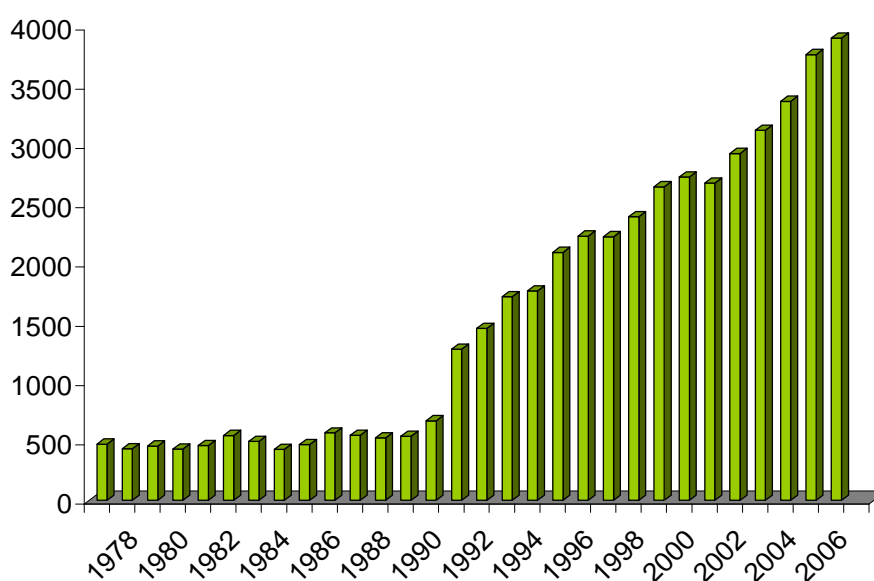
Figure 1.2 Number of publications about TB during the last 30 years[4]

Recently, the complete genome sequence of the best characterised strain of *M. tuberculosis* H37Rv, has been determined.[5] This was of great interest for the cloning and overexpression of mycobacterial proteins, needed for their structure elucidation via crystallisation and X-ray analysis.

1.1.2 Tuberculosis: the disease

1.1.2.1 The bacillus Mycobacterium tuberculosis

Human tuberculosis is caused by *Mycobacterium tuberculosis*, and less frequently by *M. bovis* (usually infecting cows), *M. avium* (usually infecting birds) or *M. africanum*, all being aerobic rod-shaped organisms.

The mycobacterial cell consists of cytoplasm surrounded by a plasma membrane and containing the single chromosome, tightly wrapped into a nuclear body, but not bounded by a nuclear membrane. Thus, like other bacteria, mycobacteria are prokaryots. Many environmental mycobacteria contain one or more small extrachromosomal elements of DNA known as plasmids. In contrast with *M. tuberculosis* where plasmids do not seem to occur, they have been found frequently in *M. avium*.

The main features, distinguishing the mycobacterial cells from other microorganisms, are **their typical cell wall** and the **presence of a unique type of iron chelating peptides**.

A) The mycobacterial cell wall
   i) Structure

A typical characteristic of *M. tuberculosis*, uniquely found in mycobacteria species, is the hardly permeable and complex cell wall with its **extremely high lipid content**. Lipids account for up to 60% of the cell wall weight and they consist of a wide range of components. Some, for example mycolic acids attached to the peptidoglycan by phosphodiester bonds, are unique for mycobacteria. These mycolic acids are a group of complex branched-chain hydroxy-lipids and offer the bacillus an extra layer for adhesion. The exact structure of the cell wall has not been fully elucidated but the generally accepted model[6,7] is shown in Figure 1.3.
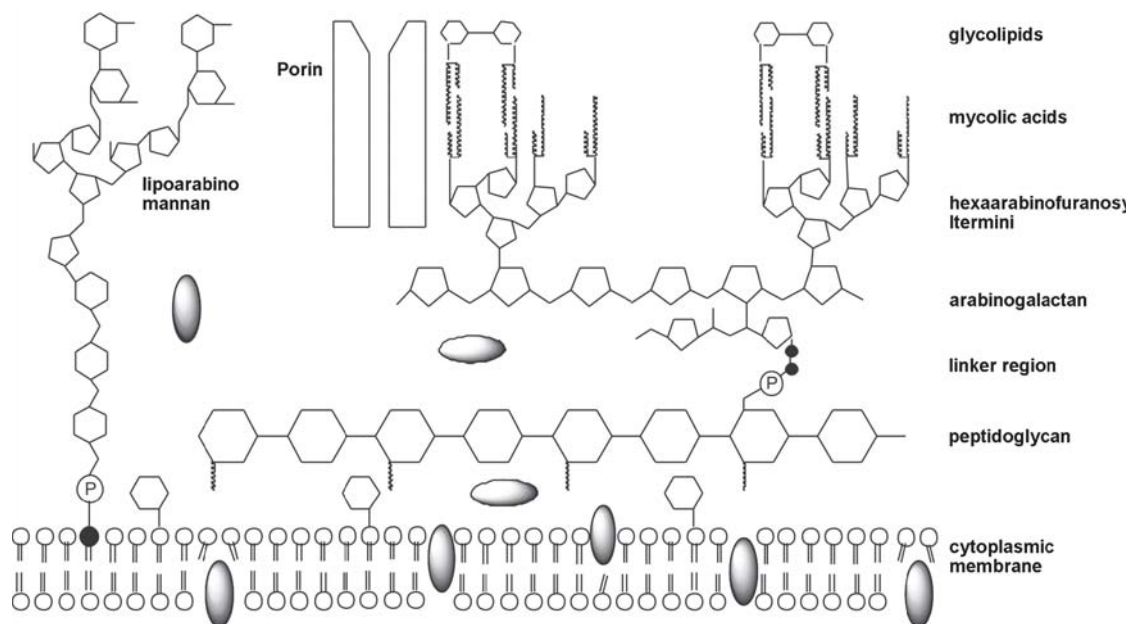
Figure 1.3 Schematic representation of the generally accepted structure of the mycobacterial cell wall.[8] The cytoplasmic membrane is encapsulated by a layer of peptidoglycan. The peptidoglycan backbone is attached to arabinogalactan. The hexaarabinofuranosyl termini of arabinogalactan are esterified to mycolic acids. Another major component associated to the mycobacterial cell wall is the immunogenic lipoarabinomannan. Small and hydrophilic solutes diffuse through water-filled protein channels, porins, whereas hydrophobic compounds use the lipid pathway. Proteins are represented by solid oval bodies.

To produce the typical branched fatty acids, two types of fatty acid synthases are used by mycobacteria: a eukaryotic type of fatty acid synthase (FAS I pathway) and a series of prokaryotic type of enzymes (FAS II pathway). The first synthase is a single multifunctional enzyme, producing C16 and C24/C26 fatty acids in four steps. The prokaryotic enzymes elongate these acid chains to lengths up to 56 carbons.

The synthesis of those characteristic mycolic acids represents an attractive target for developing antituberculosis drugs. In addition, these fatty acids render the cell wall extremely hydrophobic, **making the cell highly impermeable**. As a result, mycobacteria are resistant against dryness and disinfectants, and the low permeability for nutrients causes a very slow growth, with a generation time of 24 hours in synthetic media. This last feature contributes to the chronic nature of the disease requiring long treatments for complete sterilisation.

ii) Uptake through the mycobacterial cell wall

The hydrophobic mycobacterial cell wall acts as an efficient permeability barrier. Despite the low permeability of the cell wall, the penetration of small hydrophilic nutrient molecules into the cell occurs.

In Gram-negative bacteria the outer membrane serves as a permeability barrier. Hydrophilic channels are formed by **porins** across this outer membrane, allowing the transport of nutrients and small hydrophilic molecules. Although mycobacteria are Gram-positive organisms, known to lack this specific outer membrane, the extremely thick (10nm, 2.5-fold thicker than the one of Gram-negative bacteria), hydrophobic cell wall limits the uptake of nutrients significantly.

In *M. chelonae*, the presence of a porin in the cell wall was reported.[9] This porin forms a water-filled channel, allowing the diffusion of hydrophilic molecules through the lipid layers. Studies of this channel have shown a cation-selectivity and a strong voltage dependence of this channel.[10]

A similar porin, allowing the non-specific permeation of small molecules such as sugars and amino acids, was found in *M. smegmatis*.[11] In contrast with Gram-negative bacteria, that form trimeric porins, *M. smegmatis* forms a cone-like tetrameric complex of 10 nm in length with a single central pore, called MspA. The length of MspA is sufficient to span the outer membrane and contributes, in combination with the tapering end of the pore and the low number of pores (about 1000 pores per μm^2, 50-fold fewer compared to Gram-negative bacteria), to the low hydrophilic permeability of the cell wall.[12]

Both *M. smegmatis* and *M. chelonae* are fast-growing mycobacteria. For slow-growing bacteria, including *M. tuberculosis*, no similar porins have yet been described. A study of porins in the cell wall of *M. tuberculosis* revealed a lower channel-forming activity,[13] which might explain the slower growth of this strain.

Except for the transport of hydrophilic molecules through porins, other mechanisms for uptake through the mycobacterial cell wall are suggested. For isoniazid (INH), one of the currently most important antituberculosis drugs (see 1.1.5.2), the uptake in *M. tuberculosis* has been studied. This compound is likely to enter the cell by **passive diffusion through the bacterial envelope**.[14]

So far little information is known about the uptake of compounds into the mycobacterial cell, which can be used in the search for new antituberculosis agents. Still it is very hard to develop compounds that are able to penetrate this hydrophobic layer, as on one hand, passive diffusion promotes the search for hydrophobic compounds, while on the other hand the presence of hydrophilic channels allows the transport of very polar molecules.

Concerning these difficulties in drug development, Lipinski identified several physicochemical parameters that should be considered for small molecules, intended to be orally bioavailable. These parameters are defined as the 'Lipinski's rule of 5',[15] no more than one rule should be violated:

- molecular mass < 500 Daltons
- number of hydrogen-bond donors < 5
- number of hydrogen-bond acceptors < 10
- calculated octanol/water partition coefficient clogP <5

Known antibacterials however, do not generally follow 'Lipinski's rule of five' (Figure 1.4). Greater molecular diversity and better understanding of the physicochemical properties that are important for antibacterials is necessary.
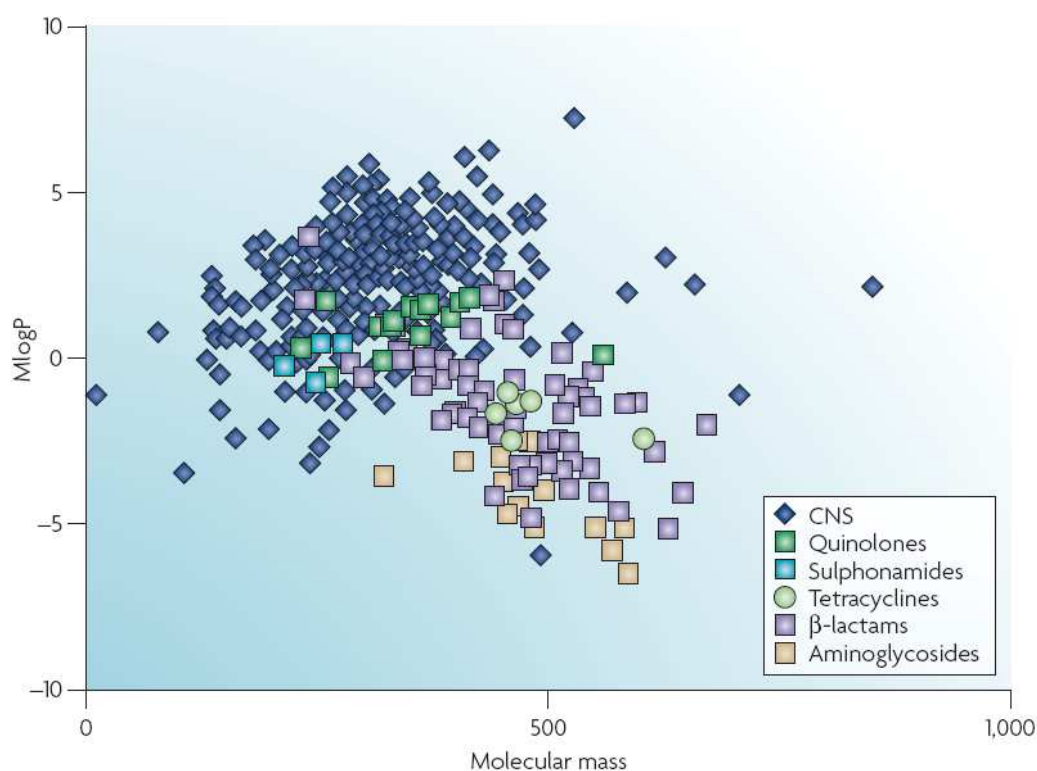
Figure 1.4 The chemical diversity of antibacterials is different to other drugs. A plot of calculated logP and molecular mass of marketed drugs for central nervous system disorders compared with marketed antibacterial classes. CNS drugs (similar observations for other mammalian target classes; data not shown) closely follow Lipinski's rule of five. Antibacterial molecules are on average more hydrophilic and slightly larger.[16]

Similarly to what is observed for all antibacterials, antituberculosis agents are often found not to follow the 'Lipinski's rule of 5'. This is illustrated by the recently discovered promising diarylquinolones (See 1.1.5.3), which would be flagged in a Lipinski analysis, with a molecular weight of 555.5 and a high clogP (7.72).[15]

B) Siderophores

Despite the very slow growth of mycobacteria and the complexity of their lipid-rich cell wall, their nutritional requirements are rather simple, including oxygen, carbon, nitrogen, phosphorus, sodium, potassium, sulfur, iron and magnesium. A wide range of nutrients are available for mycobacteria within cells and host tissues.[17] Free soluble iron however can only be found in a very small amount in lifeless environment and even less in the living host. As it is an essential requirement for bacterial growth and metabolism, micro-organisms need mechanisms for the capture and solubilisation of iron and for its transport to the interior of the cell. **The acquisition of iron by bacteria is mediated by chelating agents**, the so called 'siderophores'. In mycobacteria there are two unique classes of siderophores: exochelins and mycobactins.[18],[19] [19] Exochelins are water-soluble low-molecular weight peptides, secreted into the environment, while mycobactins are water-insoluble lipids, found within the cell wall. Iron is first solubilised and chelated from the external environment by exochelins and then transferred to mycobactin for transport across the cell wall (Figure 1.5).
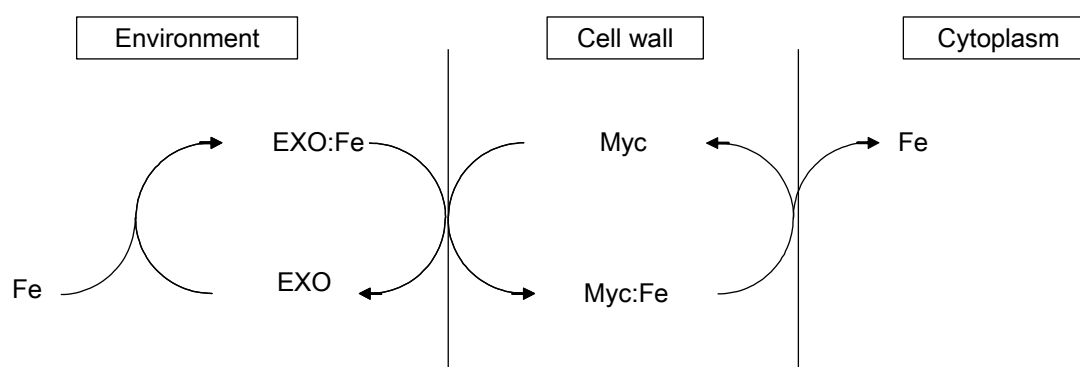
Figure 1.5 The role of exochelins (EXO) and mycobactins (MYC) in the acquisition of iron (Fe) by mycobacteria.

As exochelins and mycobactins are specific iron-chelating agents for mycobacteria species, they have recently been explored as promising target for antituberculosis agents.[19]

1.1.2.2   Infection with M. tuberculosis

Once a human host is infected with *M. tuberculosis*, the cell-mediated immunity is activated and battles the invading micro-organisms. The interaction between the human host and *M. tuberculosis* depends on the virulence of the strain, but it is supposed to depend more importantly on the specific or non-specific resistance of the host. Two types of infection occur: the **primary infection** and the **postprimary infection** (or reinfection).

During primary infection, the bacteria that reach the lungs after exogenous infection (e.g. from droplets containing viable bacteria from people with active pulmonary disease) are ingested by macrophages. As the fusion between phagosome and lysosomes can be hindered, the bacilli often survive this first attack of the human immune system. The multiplying bacteria in these macrophages cause a chemotactic response, attracting additional macrophages into the area. This flow of macrophages allows the formation of aggregates, called tubercles in which *Mycobacterium tuberculosis* continues its multiplication intracellularly. After a few weeks, many of the macrophages die, causing the release of TB bacteria and liquefying the centre of the tubercle, which is now surrounded by macrophages and lymphocytes. In this stage, the disease can become dormant.

In most cases, tuberculosis infection remains dormant and is usually unapparent. However, in people with a weakened immune system (e.g., HIV-infection, diabetes, malnutrition,…) the disease is not sufficiently controlled, leading to enlarging tubercles during liquefaction until the tubercle ruptures. At this point, the bacilli can move into a bronchiole and can then be disseminated to other organs in the body. This acute pulmonary infection can cause extensive destruction of lung tissue and can lead to the spread of bacteria to other parts of the body and finally to death. This extrapulmonary tuberculosis can affect the whole body: lymph nodes, bones, joints, central nervous system, urinary and genital tracts, abdomen and the skin. Especially in young children this can cause serious extrapulmonary forms, including meningitis. While mostly all cases of pulmonary tuberculosis are treated by antituberculosis drugs only, additional procedures are often required (e.g., surgical intervention) for extrapulmonary diseases.

1.1.3 Diagnosis

Due to the extremely slow growth of *M. tuberculosis*, **diagnosis is not easy**. Consequently, many different methods were developed to allow diagnosis and the onset of the required therapy as early as possible. The most relevant methods are described below.

1.1.3.1 Tuberculin skin test

Tuberculin is a protein factor extracted from *M. tuberculosis* and is used as an indicator for previous immunologically effective contact of the patient with the bacillus. During this test a small amount of tuberculin is intradermally injected. If the injected person has previously been exposed to the bacillus, a **localised immune reaction** (induration and swelling) appears within 1-3 days at the site of injection. The drawback of this diagnosis method is the absence of a difference between active disease or past infection and that past Bacille Calmette-Guérin vaccination also results in a positive tuberculin skin test. [20]

1.1.3.2 Culture

Sputum is by far the best material for the diagnosis of pulmonary mycobacterial infections. As most specimens submitted for **culture growth** contain other bacteria and possibly fungi, which will rapidly overgrow the medium, it is important to treat the specimens with reagents that selectively kill non-acid-fast organisms. Because of the very specific features of their cell wall, mycobacteria are relatively more resistant to acids, alkalis and certain disinfectants than other micro-organisms. Depending on the contamination grade of the specimen, 'hard' or 'soft' decontamination methods are used. A commonly used method consists of treatment of the specimen with 1N NaOH for 30 minutes, followed by neutralisation and inoculation on an isolation medium. Due to the slow growth of *M. tuberculosis*, it takes several weeks before colonies can be observed.

1.1.3.3 Sputum smear microscopy

As the isolation of *M. tuberculosis* by standard cultural methods takes several weeks, the use of **microscopy to obtain early diagnosis** is of great importance, especially for the detection of open or infectious cases. Sputum is examined either directly or after liquefaction and centrifugation and *M. tuberculosis* is visualised by the Ziehl-Neelsen method, based on carbol fuchsin with methylene blue as a counterstain.[21] Nevertheless, 25% of the smear negative patients appear to be culture positive.

1.1.3.4 Radiometric methods

In order to solve the detection problem of the very slow growing tubercle bacilli, attempts have been made to develop rapid methods for the detection of mycobacteria by their growth or metabolism. So far **the only rapid method** in routine clinical use is **radiometry**. In this technique, the bacillus uses [^{14}C]-palmitic acid as a nutrient and consequently releases radioactive labelled carbon dioxide, which can be detected within two or three days. Although this technique is clearly more rapid than conventional methods, the high price restricts the use in many countries.[22]

1.1.3.5 X-ray

Primary and reactivated tuberculosis have distinct radiographic patterns. When renewed pulmonary infection occurs, it is usually a chronic infection that involves destruction of lung tissue, followed by partial healing and slow spread of the lesions within the lungs. Those spots of damaged tissue may be observed by **X-ray examination**. However, intercurrent pneumonia, pulmonary embolism or supervening carcinoma may be confounding factors.[1]

1.1.3.6 Susceptibility to antibiotics

Smear negative patients can be treated with amoxicillin and/or erythromycin. Patients not responding to either of these antibiotics, have a 50-75% (depending on the chosen antibiotics) chance of having TB anyway.

1.1.3.7 Symptoms

Uncomplicated primary disease appears usually without any symptom. Post-primary infection causes **local and generalized symptoms**. Localized symptoms include coughing, sputum production and pleuritic pain. Constitutional symptoms include fatigue, fever, weight loss and night sweats.

1.1.4 Resistance

Many bacteria are known to acquire drug resistance through plasmid exchange. Although this is not the case for *M. tuberculosis*, resistance appears very often. Normal error rates of DNA replication make it obvious that even without antibiotic exposure, **spontaneous mutations** give rise to resistant strains.

Spontaneous mutation leading to isoniazid, streptomycine or ethambutol resistance occurs about once in 10^6 bacteria. About 1 of 10^8 bacteria is rifampicin-resistant. Consequently, treatment with a single drug in patients with a high bacterial load leads to suppression of only the susceptible bacteria, while the drug-resistant strains continue their growth. For this reason, TB is commonly **treated with cocktails of multiple drugs**.

1.1.4.1 Drug resistant tuberculosis

In cases of drug resistant TB, the bacteria are resistant to one or more anti-tuberculosis drugs. Two types can be distinguished. When resistant bacilli are found in people that have been treated before, it is called '**acquired resistance**'. In most cases, the resurgent organisms are resistant to the drug that was used before. Although especially in developing countries, the resurgence of the infection can be caused by several factors: the poor drug supply or the prescription of inadequate chemotherapy, combined with interruption of the lengthy therapy by the patient, are common reasons for failure of the treatment.

In patients without prior treatment with antituberculosis drugs, resistant bacilli can be observed as a consequence of the high level of acquired resistance in the community. This is called '**primary resistance**' and occurs less frequently than acquired resistance. Moreover it is less severe as it is very commonly restricted to just one drug (streptomycine or isoniazid).

1.1.4.2 Multi-drug resistant (MDR) tuberculosis

TB infections caused by MDR strains are resistant **to several antituberculosis drugs, and at least to isoniazid and rifampicin**. MDR tuberculosis affects usually patients after failed treatment regimens and represents a significant proportion of tuberculosis cases with acquired resistance.

In March 2006, the WHO even agreed to introduce a new "type" of infections: the '**extensively drug-resistant (XDR) TB**', now defined as TB cases in persons whose isolates are resistant to at least isoniazid and rifampicin (= MDR) and to any fluoroquinolone and at

least one of the three injectable second line drugs, such as capreomycin, kanamycin or amikacin.[1]

Data on the MDR and XDR-TB infections are lacking for many parts of the world, but since 1990 several outbreaks of MDR-TB associated with HIV have been reported in several locations, including London, Milan and New York. The WHO published some data about MDR and XDR-TB cases in March 2006. In a study of 17690 isolates spread over 49 countries during 2000-2004, 20% were found to be MDR, and 2% were XDR.

1.1.4.3 Cross-resistance

Patients showing resistance to one drug often show resistance against other drugs from the same family or even drugs from other families as well. This should be taken into account when choosing a drug regimen for a patient with multi-drug resistant TB.

1.1.5 Control and treatment

1.1.5.1 Chemoprophylaxis and preventive therapy

As vaccination with the Bacille Calmette-Guérin (BCG) vaccine, the only vaccine in use, often fails,[23] some more attention needs to be paid to the prevention and treatment of TB.

Chemoprophylaxis is defined as the treatment of uninfected people who are exposed to risk of infection, while preventive therapy refers to the treatment of people who are infected with tubercle bacilli but do not show evidence of active disease.

This preventive therapy is principally used to protect children, especially under the age of three, who are at risk to develop serious extrapulmonary forms of tuberculosis.

For adults, at risk to develop TB, like for example HIV-positive persons, the use of this preventive therapy is controversial. HIV-negative people have a 5% chance of developing active TB during their lifetime. Thus, if preventive chemotherapy is life-long and 100% effective, it requires the treatment of 20 patients to prevent the disease in 1 case. Problems of adherence and toxic side effects further limit the usefulness of this preventive approach.[24]

For HIV-infected patients, the situation is totally different as their risk for developing active TB increases to 50% during their life. But still, as the problem of adherence is quite large, therapy should be supervised, which is almost impossible especially in developing countries. This compromises the overall impact of preventive therapy.

Transplant recipients receiving steroids and other immunosuppressive drugs show a reduced immune response and are also at risk of developing active TB. It has been suggested to treat these people with a low dose of antituberculosis drugs.[25]

1.1.5.2 Current chemotherapy

The mycobacterial population in a patient can be divided into three stages:

1) freely dividing extracellular bacilli, mainly in the cavity walls;
2) slowly dividing bacilli within macrophages and in acidic, inflammatory lesions;
3) dormant and near-dormant bacilli, within cells and in firm caseous material.
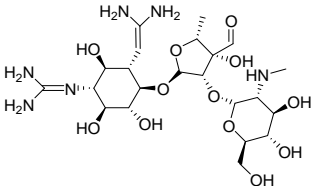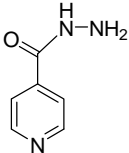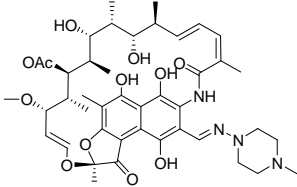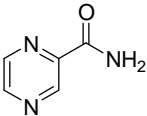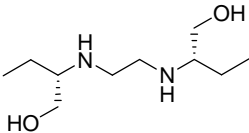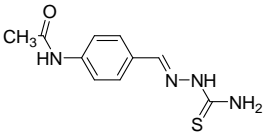
The aim of chemotherapy is to cure the patient, and at the same time prevent the emergence of drug resistance. Taking into account the problem of mutations, leading to resistance, it is essential to give at least two drugs to which the strain is susceptible as the chance of two mutations occurring simultaneously in one cell is almost non-existing. In regions with a high prevalence of drug resistance, a combination of three or more drugs should be given.

The currently used anti-tuberculosis drugs are divided into two categories: the **first-line drugs**, used in patients without drug resistant strains or strains resistant to one or two drugs (other than isoniazid and rifampicin) and the **second-line drugs**, used in cases of MDR infections, in combination with first-line drugs. They are often less effective, more toxic and more expensive and they are only used if the standard regimens are not effective anymore.
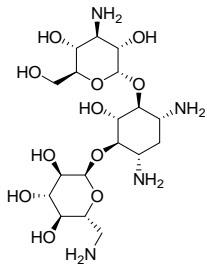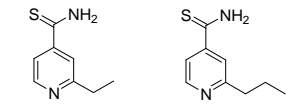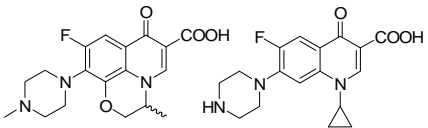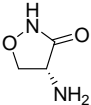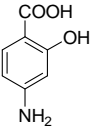
a) First-line antituberculosis drugs

The most important features of the first line antituberculosis drugs and their structure are shown in Table 1.1.

Table 1.1 First-line antituberculosis drugs

Compound	Structure	Mechanism of action	Activity	Drawback	Resistance
Streptomycine[26]		Inhibition of translation of mRNA via interaction with the 30S ribosomal unit.	Alkaline: bactericidal Acidic: ineffective	Given by injection.	Mutation on the genes coding for ribosomal protein S12.
Isoniazid[27]		Inhibition of mycolic acid synthesis. After activation by catalase-peroxidase, the obtained radical reacts with NAD^+ to yield an enoyl-ACP reductase inhibitor.	Bactericidal	Limited efficacy against metabolically less active bacilli.	Mutation on the gene coding for catalase-peroxidase.
Rifampicin[28]		Inhibition of RNA-transcription via interaction with RNA-polymerase.	Active against all 3 stages of bacterial population.	Resistance is developed with high frequency.	Mutation on the gene coding for RNA-polymerase β subunit.
Pyrazinamide[29]		Uncertain. After intracellular conversion to pyrazinoic acid by pyrazinamidase, an intake of protons leads to complete dysfunction of the pH balance. [30] Pyrazinoic acid might inhibit fatty acid synthetase I. [31, 32]	Alkaline or neutral: ineffective Acidic: effective. Kills semi-dormant bacilli	Toxicity to retinal ganglion cells and ocular tissue.	Correlated with pyrazinamidase activity.
Ethambutol[33, 34]		Inhibition of arabinan biosynthesis, via for example inhibition of arabinosyl transferase, thereby destroying the basis of the cell wall and facilitating the uptake of other compounds.[35, 36]	Bactericidal	No sterilizing agent, not very effective if used alone.	Mutation on the gene coding for arabinosyl transferase.
Thioacetazone[37]		Not elucidated yet.[38] Probably interfering with mycolic acid synthesis	Bacteriostatic	Low efficacy	

b) Second-line antituberculosis drugs (Table 1.2)

Table 1.2 Second-line antituberculosis drugs

Compound	Structure	Mechanism of action	Side effects	Resistance
Aminoglycosides (e.g. kanamycine, amikacine and capreomycin) [26,39]	 Kanamycin	Inhibition of translation of mRNA via interaction with the 30S ribosomal unit. Uptake through porins.	Ototoxicity, reversible nephrotoxicity.	Mutation on the genes coding for ribosomal protein S12.
Thioamides (e.g. ethionamide, prothionamide) [27]	 Ethionamide Prothionamide	Inhibition of mycolic acid synthesis. Binds enoyl-ACP reductase at a different part of the active site compared to isoniazid.	Nausea, psychotic reactions, hypoglycaemia, hepatitis.	
Fluoroquinolones (e.g. ofloxacin, ciprofloxacin) [26]	 Ofloxacin Ciprofloxacin	Inhibition of DNA-gyrase, responsible for supercoiling the bacterial DNA.	Rare. Gastrointestinal disturbance, dizziness, headache, mood changes.	Missense mutations in DNA gyrase.
Cycloserine [40]		Inhibition of mycobacterial peptidoglycan synthesis by inhibition of D-alanine racemase.	High toxicity. Dizziness, convulsions, headache, depression.	
***Para*-aminosalicylic acid** [41]		Inhibition of DNA-synthesis (mimic of *p*-aminobutyric acid). Mechanism unclear.	Dose related side effects: vomiting, diarrhea, hepatotoxicity, fever, hypocalcaemia, rash.	

c) WHO recommended regimens

In order to shorten the length of the illness, to lower the risk of death and to prevent the development of resistant strains, the WHO recommended a regimen known as Directly Observed Therapy, Short Course (DOTS). It focuses on five elements: political obligation, microscopic provision, drug supply, monitoring systems and a direct observation of the treatment. When implemented over a long period of time, this approach is expected to give a huge reduction in sources of infection and transmission.[1]

The fixed patterns as presented by the WHO regimens depend on the type of treatment.

♦ WHO standard first-line regimen

This regimen is used when a patient is treated for the first time with antituberculosis agents. It takes **at least 6 months** and requires a combination of different antibiotics. Usually, the treatment **starts with three drugs**: isoniazid, rifampicin and pyrazinamide, often with the inclusion of a fourth drug, e.g. ethambutol or streptomycin. These agents are administered daily in combination for two months. In a continuation phase, isoniazid and rifampicin are administered twice a week for 4 to 12 months. The advantage of using an early intensive phase of treatment is the possibility to cure a high proportion of patients resulting in a smaller chance of the disease relapsing if the patient absconds before completion of the less intensive continuation phase.

♦ WHO standard retreatment regimen

This retreatment regimen is used for patients returning after premature interruption of the first-line treatment and/or relapse. It combines **five drugs for a period of 8 months** to reduce the risk of failure due to acquired resistance. During the whole period, isoniazid, rifampicine and ethambutol are administered, combined with pyrazinamide during the first three months and streptomycine during the first two months. The vast majority of the patients will be cured with this retreatment regimen. Failures are generally due to the use of an incorrect regimen and/or failure to ensure that the regimen is fully administered and directly observed.

♦ Therapy of MDR-TB

Treatment of patients infected with MDR tubercle bacilli involves second-line drugs. Unfortunately, these agents are more expensive, less effective and cause more side effects compared to the standard drugs. The treatment consists of a mix of **at least four standard or second-line drugs**, distributed by a specialized unit and, if possible, selected on the basis of *in vitro* susceptibility, or else being likely to be fully sensitive, i.e. not previously used

by that patient. When the patient's sputum cultures have become negative, one or more drugs are withdrawn, preferably weaker drugs which are causing side effects. In order to prevent relapse, the treatment with this slimmed regimen is continued for at least 18 months after sputum conversion.

If therapy fails, the regimen should be revised in the light of repeated *in vitro* susceptibility tests. Additional drugs should never be added blindly to an ineffective regimen.

As in some parts of South-Africa, studies concerning extensively drug-resistant (XDR) tuberculosis in an HIV-positive population have shown alarmingly high mortality rates, alternative and new forms of treatment are urgently required.

d) Surgery

For patients with bacilli sensitive to only 2 or 3 relatively weak drugs, surgery is considered. Nevertheless, for many of these patients the infection has become too severe or the lung function has diminished too much for surgery to be possible. Nowadays, surgery is rarely considered in pulmonary tuberculosis but is often important for the diagnosis and management of extrapulmonary disease (e.g., bone lesions, lymph nodes,…).

If the patient has a large localised lesion with little other disease, reasonable lung function and only 2 or 3 weak drugs available, surgery can be considered. To avoid serious and potentially fatal TB complications, the operation is performed when the bacterial load is the lowest, i.e. after 2 months of intensive treatment. After surgery, the same drug regimen is continued for at least 18 months.

1.1.5.3 *Patent developments in small-molecule antimycobacterials*

As the clinical management of patients infected with *M. tuberculosis* faces difficult problems such as the worldwide emergence of multidrug resistant tuberculosis (MDR-TB) and the increase of AIDS-associated infections, the development of new drugs with greater or distinct antimycobacterial activity than those currently available, is urgently required. As this message is well understood in scientific research area, **a large amount of patents** have been published during the last years discussing the discovery of small-molecule agents that appear useful in the treatment of TB. Three main strategies toward this effort can generally be distinguished: structural modification of existing antibiotics, discovery of new targets and the identification of natural resources for novel antibiotics. Some of major remarkable developments in recently patented antimycobacterials are described below.

a) Derivatives of existing drugs

♦ Using **thioacetazone** (**1.1**) and ***p*-aminosalicylic acid** (**1.2**) as a template, the University of Sciences in Philadelphia prepared some **halogenated derivatives**.[42] A 3-fluorinated derivative **1.3** emerged as the most active derivative of thioacetazone. The compound showed a MIC-value < 0.1 µg/mL and a selectivity index (SI) > 312.5. Moreover, it proved active against TB strains resistant to isoniazid, rifampicin, ethambutol, kanamycin and ciprofloxacin. Compared to the MIC-value of thioacetazone against *M. tuberculosis* H37Rv (>2.0 µg/mL) this fluorinated derivative is about 20-times more effective.

In the series of *p*-aminosalicylic acid derivatives, the 5-fluorinated *p*-aminosalicylic acid (**1.4**) showed the highest activity with a MIC-value of < 3.13 µg/mL and an SI of > 10. Tests for activity against drug-resistant strains revealed a lower but still significant activity. Only strains resistant to isoniazid are less susceptible (MIC= 25 µg/mL). Compared with the MIC of *p*-aminosalicylic acid (1.25 µg/mL), this derivative seems less active than the original lead compound.

♦ Using **ethambutol** (**1.5**) as the lead, a large library of ethylene diamines with a variety of amine substituents, as well as substituents in the linker region, has been produced using split-and-pool procedures on a solid support.[43,44] Initial evaluation of the library was done by a high-throughput screening (HTS) assay using bioluminescent reporter strains that produce light in response to inhibition of the cell wall synthesis.[45] A number of hits resulted from this HTS assay. One of the interesting derivatives, flanked by a 2-adamantyl and a geranyl substituent (**1.6**), gave a MIC-value of 0.2 µM against *M. tuberculosis*, compared to 9 µM for ethambutol in the same BACTEC assay.[43] This derivative also proved to be active against three MDR patient isolates that were highly resistant to ethambutol and was well tolerated by mice in concentrations up to 600 mg/kg. *In vivo* assays in mice infected with *M. tuberculosis* H37Rv showed that 10 mg of compound **1.6** had similar activity as 100 mg of ethambutol in reducing colony-forming units (CFUs) in the spleen, whereas the same dose had superior activity than 100 mg ethambutol in reducing CFUs in lung.
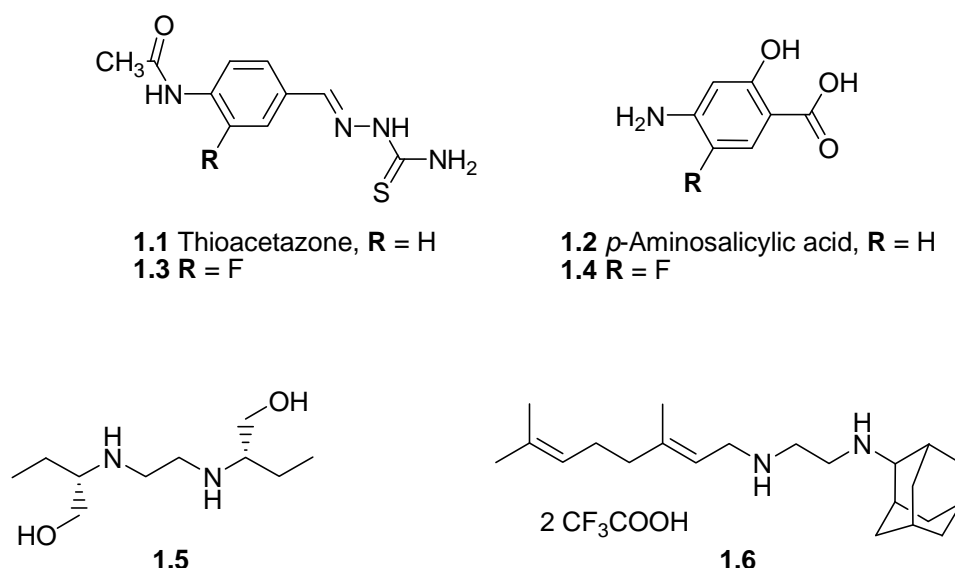
Figure 1.6 Derivatives of existing drugs

♦ **Rifampicin** (**1.7**), a semisynthetic antimicrobial drug derived from rifamycine, is a front-line drug in the treatment of TB. As it kills persistent semidormant bacilli, it is crucial to achieve sterilisation. Since its introduction in the treatment of susceptible TB in the beginning of the 70's, new derivatives have been synthesized and evaluated for the treatment of TB.[46] The best headway has been made with rifapentine (**1.8**), a long-acting rifamycin derivative characterised by a lower MIC-value (0.06 µg/mL versus 0.25 µg/mL), which was approved for the treatment of TB in the US in 1998.

WOO3O84965 describes *N*-(3-rifamycinyl)carbamates (**1.9**) as new substances for treating and preventing TB.[47] Compounds have been tested *ex vivo*. The intracellular activity of growth inhibition in mouse macrophages, infected with *M. tuberculosis* H37Rv, was determined by counting the CFU/mL. The most active compounds (R= ethyl, methyl) exhibit an eightfold higher activity than rifampicin.

WO04005298[48] highlights new derivatives of rifabutine Ia (**1.10**), a spiropiperidyl-rifamycine, which has been recommended for HIV-infected TB patients who cannot receive rifampicin because of interactions with antiretroviral agents. Antimycobacterial activity was determined against *M. avium*. While the known lead displayed a MIC-value of 0.15 µg/mL, the most promising derivatives were slightly less active. Only rifabutin IIIa (**1.11**) had comparable activity to compound **1.10** (MIC-value= 0.2 µg/mL).

Several patents from Hokuriku Seiyaki are related to novel series of 14-membered macrolide derivatives. Among the 1265 erythromycin analogues disclosed in WO0226753,[49] several compounds are claimed useful for preventing and/or treating TB. Analogues **1.13**-**1.16** showed MIC-values in the 0.10-0.20 µg/mL range against different *M. avium* and *M. intercellulare* strains. Another typical analogue **1.17** showed a MIC-value of 1.56 µg/ml

against *M. tuberculosis* H37Rv and some other *M. tuberculosis* strains, [50] while **1.12** exhibits the same MIC-value against *M. avium* strains.

1.7 Rifampi(ci)n, **R** = CH$_3$
1.8 Rifapentine, **R** =

1.9

1.10 Rifabutin Ia, **R** = H
1.11 Rifabutin IIIa, **R** = COCH$_3$

1.12 R1 = H   **R2** = H

1.13 R1 = H   **R2** =

1.14 R1 = H   **R2** =

1.15 R1 = H   **R2** =

1.16 R1 = CH$_3$ **R2** =

1.17 R1 = H   **R2** =
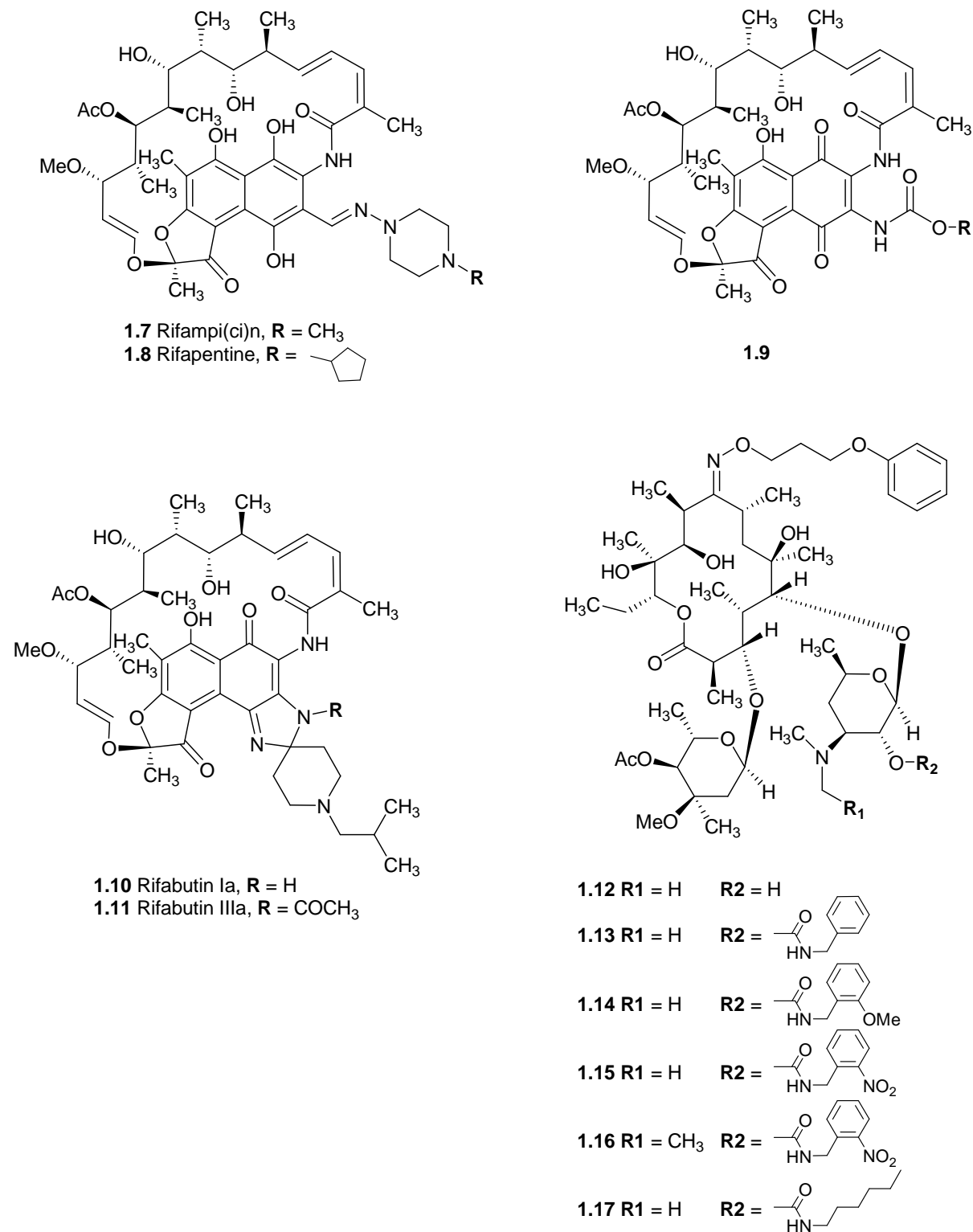
Figure 1.7 Rifampicin-analogues and macrolide derivatives

♦ For a long period, the mechanism of action of **isoniazid** has remained uncertain. Now it is established that isoniazid first requires activation by Kat G into an unstable intermediate which reacts with NAD to give biologically active species as **1.18**. This adduct then binds the NAD(H) recognition site of enoyl ACP-reductase and this leads to the antimycobacterial effect. Further analysis and biological evaluation of the adducts arising from the reaction between isoniazid, an oxidizing agent and NAD proved that compounds structurally related to the generated adducts can also act as ACP-reductase inhibitors.[51] Few synthetic analogues of NAD bearing a double substitution on the nicotinamide moiety such as **1.19** have been patented for their antimycobacterial properties.[52]
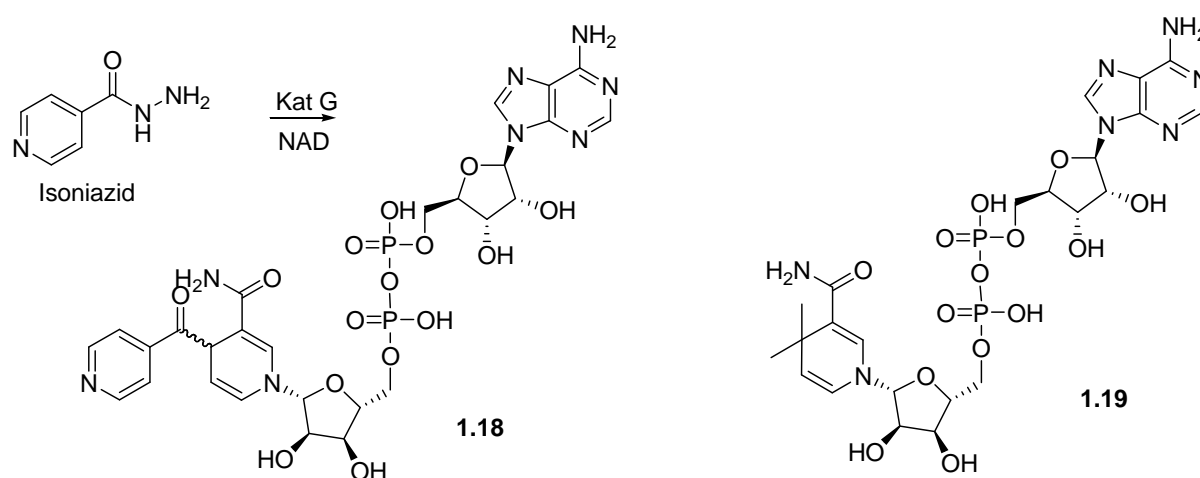
Figure 1.8 Mechanism of action of isoniazid and derived intermediate-analogue **1.19**

b) Drugs on known targets

♦ A Janssen Pharmaceutica patent discloses novel 6-bromoquinoline derivatives with potent antimycobacterial activity.[53] These so-called **diarylquinolines** (DARQ) are structurally distinct from known broad-spectrum fluoroquinolone antibiotics. With a MIC-value of 0.01 µg/ml against *M. tuberculosis* H37Rv and *M. smegmatis* strains, one of the four possible diastereoisomers of compound **1.20** showed potent antituberculous activity. Unfortunately, the absolute configurations of the two stereocenters were not determined. A follow-up paper of Andries at al., however, attributes the activity to the (1*R*,2*S*) diastereomer (called TMC207, previously called R207910).[54] This paper further reveals that TMC207 does not inhibit *M. tuberculosis* DNA gyrase, the target for fluoroquinolones. A gene commonly affected in isolated mutants resistant to TMC207 encodes for atpE, a part of the F0 subunit of ATP synthase. This indicates that the atpE gene product (*i.e.*, the proton pump of *M.*

tuberculosis ATP synthase) is inhibited by TMC207. This distinct target of TMC207 implies that there is no cross-resistance with existing anti-TB drugs. Pharmacodynamic studies in mice, where TMC207 was used individually or where one of the first-line drugs of the triple combination therapy was replaced with 25 mg/kg TMC207, proved that the activity of each combination containing TMC207 was significantly better than the standard regimen (e.g. culture-negative lungs after two months). The mouse studies also showed that this new compound quickly enters the bloodstream and is actually concentrated in lung cells, which harbor the TB bacilli, killing the bacilli soon after they enter the body. Also, TMC207 lingers in the body for days continuing to kill bacilli even when administered only once a week in mice. At this moment, this compound is going into clinical phase II, to evaluate the anti-bacterial activity, safety and tolerability in subjects with sputum smear-positive pulmonary infection with MDR-TB.[55]

♦ **Oxazolidinones** represent the first completely new class of synthetic antibacterial agents to achieve regulatory approval in over 30 years, as exemplified by linezolid (**1.21**, Zyvox®). In addition to its potent activity against Gram-positive pathogens, this class is of great interest because it exerts its antibacterial action by a mechanism distinct from other antibacterial agents. Oxazolidinones inhibit bacterial protein synthesis (translation) at a very early step. They inhibit the formation of a ribosomal initiation complex involving 30S and 50S ribosomes. Because of their unique mechanism, the oxazolidinones are not cross-resistant with any known antibiotic. Since this class was found to be endowed with promising antibacterial properties, many pharmaceutical companies worldwide began research programmes in the oxazolidinone area, which resulted in an impressive number of patents.[56] However, most patents are silent about antimycobacterial activity of the disclosed compounds and only a few of the disclosures present compounds that have demonstrated potential as antimycobacterials. A general feature of the oxazolidinone derivatives is the fact that only enantiomers with a (5*S*)-acetamidomethyl configuration are known to exhibit antibacterial activity. Another feature is that almost all oxazolidinones endowed with antibacterial activity carry a phenyl ring attached to the nitrogen atom of the oxazolidinone. Linezolid forms an attractive starting point to design anti-TB agents, since it displays a MIC-value of 0.5 µg/ml against *M. tuberculosis* and it is known that resistant bacterial mutants to linezolid arise at low frequency. Some of the potential side effects associated with extended linezolid treatment are toxic optic neuropathy, peripheral neuropathy and myelosuppression. Ranbaxy disclosed novel oxazolidinone analogues related to eperezolid.[57] All analogues have a diazine moiety attached to the phenyloxazolidinone, which is further substituted by a heterocyclic or aromatic ring. Compound **1.22** with a 5-nitrothien-2-yl moiety directly

attached to the piperazinyl oxazolidinone core was very effective against the different *M. tuberculosis*, *M. intracellulare*, *M. avium* and *M. bovis* strains tested and MIC-values compared well with those of established anti-tuberculosis agents.
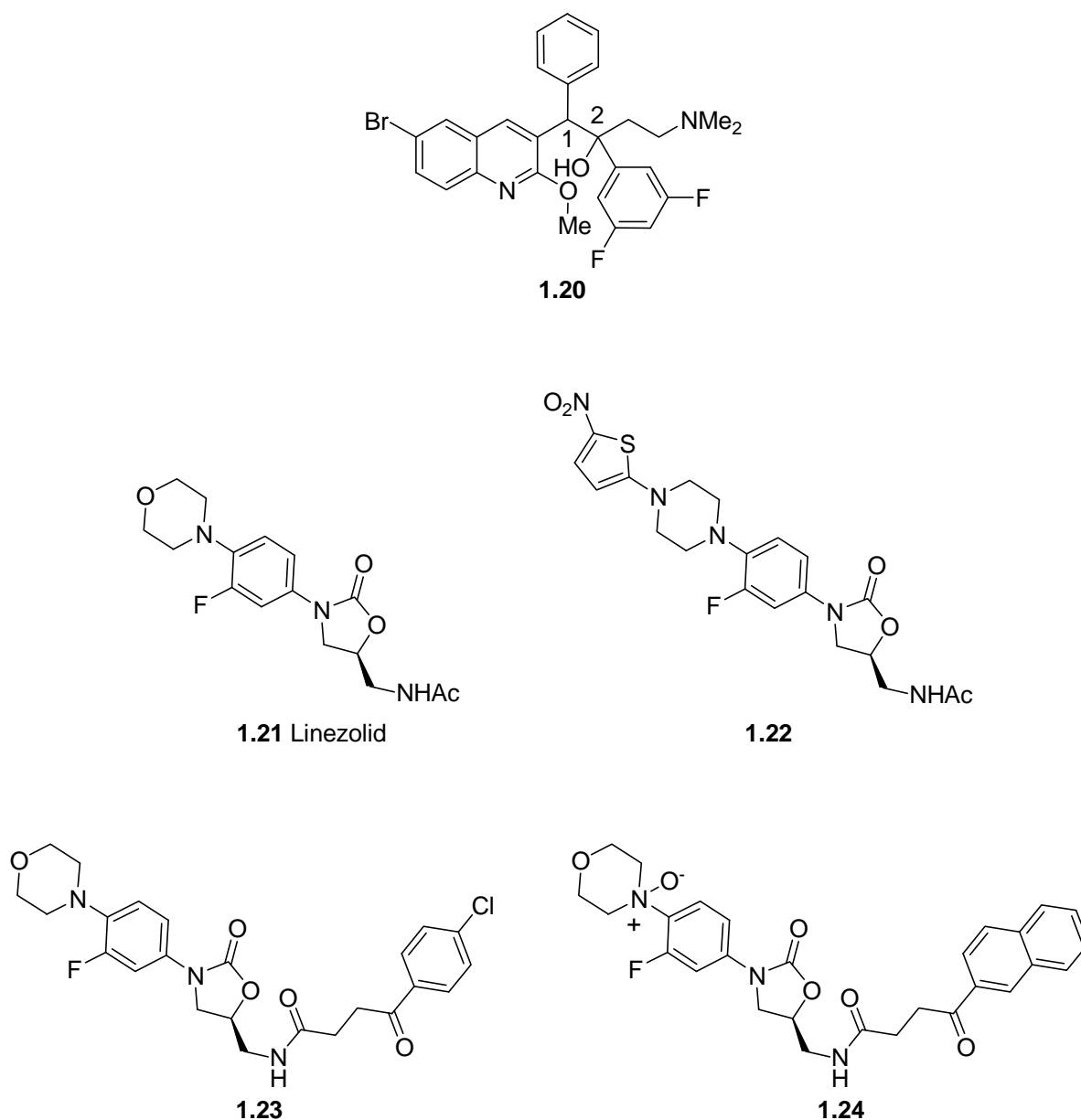
Figure 1.9 Diarylquinolone **1.20** and oxazolidinones **1.21**-**1.24**

♦ A Lupin Patent disclosed 78 novel **oxazolidinones** as antibiotic agents against *M. tuberculosis*.[58] Elongation of the acetamido moiety (as it occurs in linezolid) is common to all claimed analogues. Among the most potent analogues are **1.23** and its *N*-oxide (MIC-value of 0.25 µg /ml for both compared to 0.50 µg /ml for linezolid) and the related **1.24** (MIC-value of 0.50 µg /ml). The latter analogue was also potent against resistant clinical isolates and showed a favourable acute toxicity profile (LD$_0$ > 1000 mg/kg P.O.). When administered at

12.5 or 25 mg/kg to mice infected with *M. tuberculosis* ATCC27294 5 days/week for 4 weeks, a mean $\log_{10}$ reduction in colony forming units of 0.20 and 2.3 (in lung) and of 0.26 and 2.49 (in spleen) was observed. The *N*-oxide derivatives are expected to be highly water soluble.

c) Agents with unspecified mechanisms of action

1'-Acetoxychavicol acetate (**1.25**) is a natural compound, found in some plants of the family *Zingiberaceae*. The mechanisms of action of this compound are not clear. It could inhibit the function of xanthine oxidase and NADPH oxidase. These enzymes are involved in the production of superoxide anion, which is one of the spontaneously occurring toxic substances in the body. It may also inhibit nitric oxide synthase production. The MIC of 1'-acetoxychavicol acetate against *M. tuberculosis* H37Rv is 0.1 µg/ml (0.1 to 0.5 µg/ml to 30 clinical isolates), which is well below the toxic concentration against various mammalian cells.[59]

1.25

1.26 Pleuromutilin, R = OH
1.27 Valnemulin, R =

1.28

1.29 (+)-Calanolide A

1.30

Figure 1.10 Agents with unspecified mechanism of action

Pleuromutilin (**1.26**) is a tricyclic diterpenoid and a naturally occurring antibiotic produced by the basidiomycetes *Pleurotus muti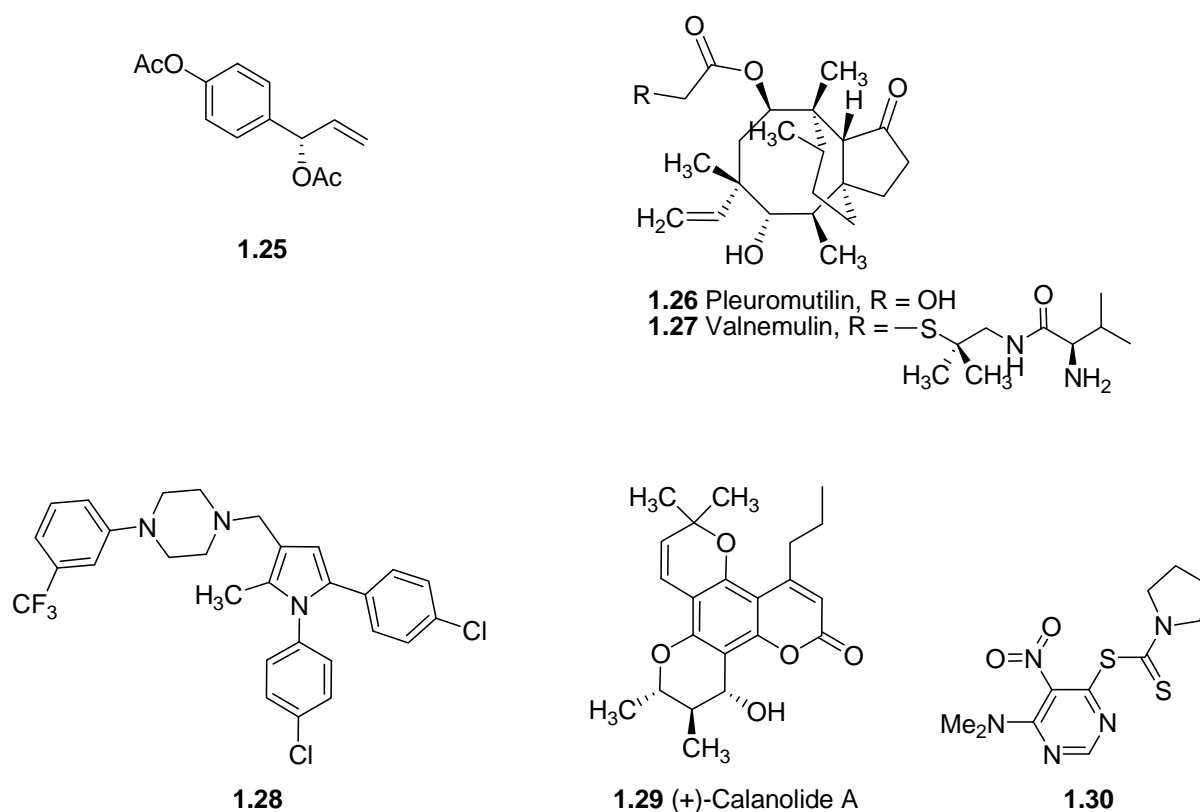lus* and *P. passeckerianus.* A number of further pleuromutilin derivatives have been claimed by Sandoz.[60] The most active derivatives, e.g., Valnemulin (**1.27**), displayed MIC-values in the range of 0.5 to 8 µg/ml depending on the *M. tuberculosis* strain considered.

By using the existing lead BM212,[61] researchers at Lupin have discovered novel substituted pyrroles endowed with antimycobacterial activity.[62] Compared to the plethora of pyrrole derivatives reported earlier and reviewed in this patent, several of the ca. 90 disclosed derivatives possess higher antimycobacterial activity against clinically sensitive as well as resistant strains. The MIC-value of compound **1.28**, the most active derivative, against *M. tuberculosis* 27294 is 0.125 µg/ml. When assayed *in vivo* in Swiss albino mice that were challenged with the same strain, **1.28** compared favourably (on a mg/kg basis) with isoniazid in reducing CFUs in lung and spleen, while its LD_0 was > 2000 mg/kg, compared to a reported LD_{50} of isoniazid of 139 mg/kg in mice.

US6268393 contains disclosures on the synthesis and the antimycobacterial activity of calanolide analogues.[63] (+)-Calanolide A (**1.29**), a known anti-HIV agent which was originally isolated from the rain forest tree Calophyllum lanigerum, was moderately active (MIC-value of 3.13 µg/ml against *M. tuberculosis* H37Rv).

In WO03042186, Medac revealed dithiocarbamate derivatives endowed with activity against different *Mycobacterium* species. The most active agent is 4-dimethylamino-6-tetramethylenedithiocarbamoyl-5-nitropyrimidine (**1.30**) with a MIC-value of 3.12 µg/ml against *M. tuberculosis.*[64]

1.1.5.4 Nucleosides as anti- tuberculosis agents

Whereas nucleoside derivatives are used to treat HIV, so far no nucleosides are registered as antibacterial or antituberculosis drugs. Nevertheless, during the last years, some interesting nucleosides were discovered with potent activity against *M. tuberculosis.*

Almost all nucleoside antibiotics have non-specific antibacterial activity because nucleoside and nucleotide metabolism are essential for DNA and RNA biosynthesis in all organisms.

A simple nucleoside named **toyocamycin** (**1.31**)[65] is isolated from the fermentation broth of *Streptomyces toycaensis* and shows good antibacterial activity against *Mycobacterium tuberculosis* H37Rv (MIC-value = 2.0 mg/mL). Next to this promising activity, it has many other biological activities such as antitumor, antivirus and herbicidal activities, but it is also cytotoxic against eukaryotic cells and showed toxicity in mouse, thus limiting their use as specific antibiotic.

Aldrich et al.[66,67] described the rational development of compounds inhibiting the growth of *M. tuberculosis* H37Rv under iron-limiting conditions, with the most active compounds **1.32** (MIC_{99}= 0.29 µM) and **1.33** (MIC_{99} = 0.19 µM) (See Figure 1.11). These compounds target MbtA, the enzyme that initiates the biosynthesis of mycobactins, essential iron-chelating siderophores. The main drawback of these compounds is their hydrolysis to 5'-*O*-(sulfamoyl)adenosine, an extremely cytotoxic compound. Thus further exploration for compounds with similar activities but with lower toxicities is required.

1.31

1.32: X= O
1.33: X=NH

1.34

1.35
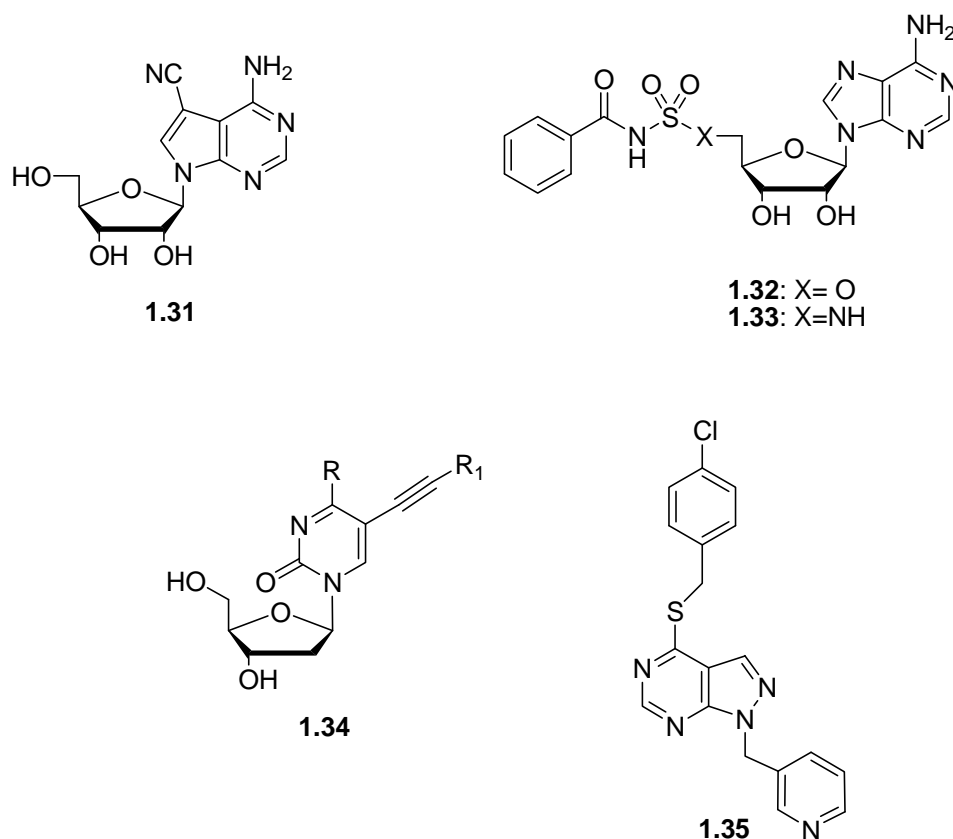
Figure 1.11 Nucleosides as antituberculosis agents

Kumar et al.[68,69] explored a large family of **5-substituted alkynylpyrimidine** nucleosides **1.34** (Figure 1.11). 5-Decynyl, 5-dodecynyl and 5-tetradecynyl derivatives showed the highest antimycobacterial potency against *M. bovis* with MIC_{90}-values below 10 µg/mL. Next to normal ribose and 2'-ribose sugars, also acyclic compounds have been synthesized and evaluated for their activity. The target of these nucleosides remains unknown.

Very recently, a series of **modified adenine derivatives** was reported as potent antimycobacterial agents.[70] Based on earlier results, this paper starts from the isomeric pyrazolo [3,4-*d*]pyrimidine nucleus and explores a large amount of *S*- and *N*-alkylations, to

end up with a very potent compound **1.35** with a MIC-value of 0.5-1 μg/mL. As is the case for the 5-alkylated pyrimidines, no target is known for this new class of antituberculosis agents.

Caprazamycins (CPZs) **1.36** were isolated from a culture broth of the Actinomycete strain *Streptomyces* sp. MK730-62F2.[71] Different CPZs were isolated, each with a different fatty acyl side chain (R). Caprazamycin B (with R= 11-methyl-dodecyl) showed excellent anti-mycobacterial activity in vitro against drug-susceptible *Mycobacterium tuberculosis* H37Rv (MIC= 3 μg/mL) and multi-drug resistant strains (MIC= 6.25-12.5 μg/mL). Furthermore, CPZ B proved equipotent as clarithromycin in inhibiting *M. avium* growth and did not show any toxicity in mice. Caprazamycins are believed to disrupt the formation of the bacterial cell wall peptidoglycan by strong inhibition of phosphor-MurNAc-pentapeptide translocase (MraY, translocase I). Being one of the key enzymes for peptidoglycan biosynthesis, MraY is essential for the growth of bacteria.[72]
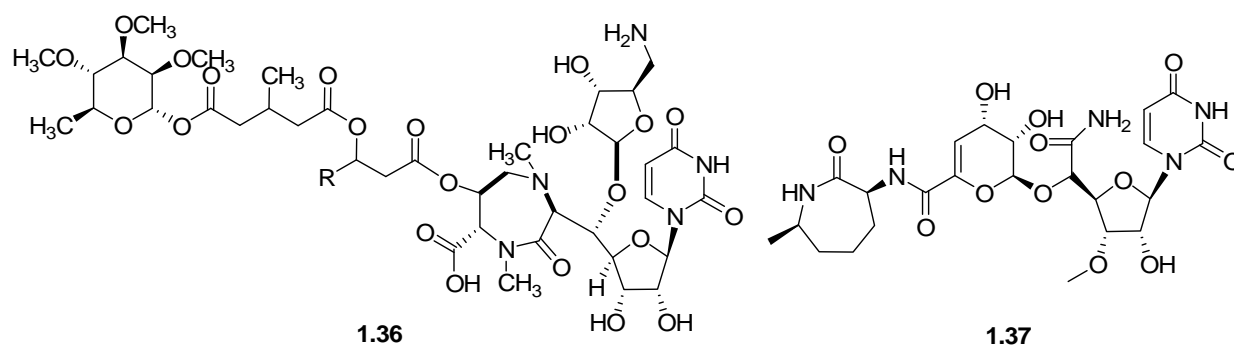
Figure 1.12 Caprazamycins and A-500359A

In 2003, **A-500359A** (**1.37**) was isolated from the fermentation broth of *Streptomyces griseus* 446-S3 by Sankyo Co. Ltd.[73] It consists of a 3'-*O*-methylated uridine coupled to a dehydrohexopyranuronate and a caprolactam fragment and strongly inhibits MraY (IC$_{50}$= 0.010 μg/mL). Due to its cell wall biosynthesis inhibiting activity, it showed potent antibacterial activity against *Mycobacterium smegmatis* (MIC = 9 μg/mL) as well as *Mycobacterium avium* (MIC = 2 μg/mL) and *Mycobacterium intracellulare* (MIC = 6.25 μg/mL). Efforts were made to increase the lipophilic nature of this compound by introducing an acyl group at the 2'-position to enhance uptake through the lipophilic cell wall.[74] An A-500359A derivative possessing a decanoyl group at the 2'-position exhibited far superior activity against *M. avium* (MIC < 0.063 μg/mL) and *M. intracellulare*, despite a strongly

decreased MraY inhibitory activity (IC_{50}= 0.55 mg/mL). Most probably, the 2'-decanoyl substituted compound acts as a prodrug of A-500359A.

A more detailed overview of nucleoside-derived antibiotics is given in several reviews.[75]

1.1.5.5 Modified nucleosides as therapeutic agents

Especially in the antiviral field synthetic nucleoside analogues have found clinical applications. Most therapeutically useful nucleosides are based on the structure of natural β-D-nucleosides (i.e. AZT, d4T).

However also a significant number of L-nucleosides such as 3TC (**1.38**; Lamivudine, β-L-2',3'-dideoxy-3'-thiacytidine)[76], FTC (**1.39**; Emtricitabine, β-L-2',3'-dideoxy-5-F-3'-thiacytidine),[77] L-d4FC (**1.40**; Elvucitabine, β-L-2',3'-didehydro-2',3'-dideoxy-5-fluorocytidine)[78] and β-L-2',3'-didehydro-2',3'-dideoxy-2'-F-4'-thio-cytidine (**1.41**)[79], have been discovered as potent antiviral agents. All compounds show activity against Human Immunodeficiency Virus-1 (HIV-1) and hepatitis B virus (HBV). These nucleosides are in fact prodrugs that need to be converted to their triphosphorylated forms to achieve pharmacological activity. 3TC and FTC are both phosphorylated by deoxycytidine kinase, indicating that this enzyme possesses the ability to phosphorylate nucleosides with unnatural L-chirality.[80] 3TC, also known as Lamivudine was found to be the first drug active against HBV. Its D-isomer proved less active and more cytotoxic. Several of these L-nucleosides belong to nowadays antiviral therapy.

1.38: X= H
1.39: X= F   **1.40**   **1.41**
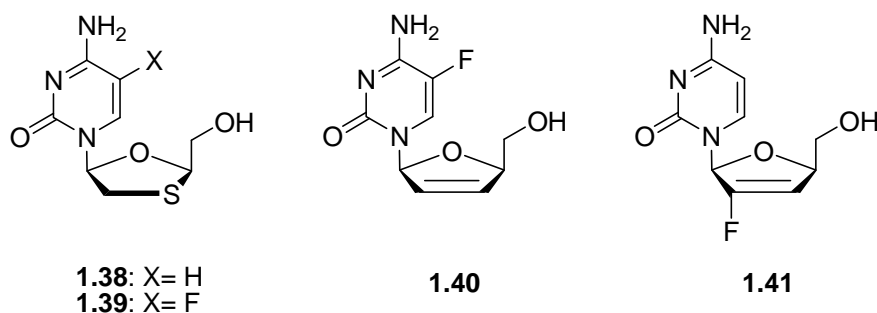
Figure 1.13 L-nucleosides with potent antiviral activity

While the above mentioned L-nucleosides are very well studied, only few reports were made on pharmacologically active α-nucleoside analogues.[81,82] α-L-lyxofuranosyl benzimidazoles showed no activity against HIV-1, but proved active against human cytomegalovirus (HCMV). However, while those reports date from 1998, no further explorations have been published in this area, indicating that only little success was obtained with those compounds.
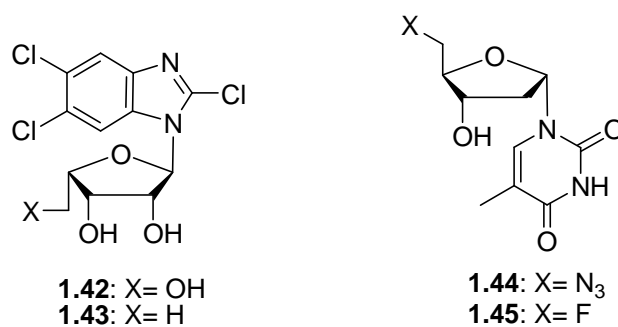
Figure 1.14 α-L-lyxofuranosyl benzimidazoles and α-D-thymidine derivatives

Schmidt et al. described the synthesis of 5'-azido (**1.44**) and 5'-fluoro (**1.45**) substituted α-D-thymidine. No antiviral (HIV-1 and HSV-1) activity was observed for those compounds.[83] Remarkably, while variation between D- or L- sugars of modified nucleosides proved possible while retaining antiviral activity, modification of the stereochemistry at the anomeric centre, from β to α, is generally associated with complete loss of bioactivities of the modified nucleosides. So far, α-nucleosides are solely well-known for their recognition of complementary DNA and RNA sequences forming stable parallel stranded duplexes.[84,85]

1.2 THYMIDINE MONOPHOSPHATE KINASE

A prerequisite for a valuable target enzyme in antibiotic research is that it is essential for bacterial growth and intracellular survival. Consequently it should catalyze a reaction necessary for growth and survival of the micro-organisms, which can not succeed without the presence of only this enzyme.

For antiviral and anticancer chemotherapy, classical molecular targets are often key enzymes involved in the salvage and *de novo* pathways of DNA synthesis, such as DNA polymerase, dihydrofolate reductase, thymidylate synthetase, ribonucleotide reductase and 2'-deoxynucleoside kinases.[86,87] Despite the fact that prokaryotes rely on the same pathways for DNA synthesis,[88] antibiotics targeting enzymes of these pathways have hardly been discovered. Except for trimethoprim, an antibiotic targeting dihydrofolate reductase,[89] no antibiotics targeting other enzymes of the DNA synthesis-pathways have entered the clinical stage.

For the use of 2'-deoxynucleosides, as antibiotics, two strategies can be followed. Similar to the anti-HIV drug AZT,[86] activation into its triphosphate by three successive kinases (TK, TMPK and TDPK) could yield a toxic 2'-deoxynucleoside triphosphate, serving as a DNA-polymerase inhibitor or resulting in a disruption in the DNA-duplex when built in, thus leading to cell death. Alternatively, 2'-deoxynucleosides could serve as an inhibitor of one of the kinases instead of being recognized as a substrate, in this way causing its toxic effect.

So far, substrates or inhibitors specifically targeting bacterial TKs or TMPKs have not yet been developed for therapeutic purposes.

1.2.1 *Mycobacterium tuberculosis* TMPK: An ideal target

Thymidine monophosphate kinase (TMPK) is one of the many nucleoside monophosphate kinases responsible for important steps during DNA-synthesis. With ATP as its preferred phosphoryl donor, it catalyzes the reversible conversion of thymidine monophosphate (dTMP) to thymidine diphosphate (dTDP), as important intermediate for DNA-building blocks.[90]

$$\text{dTMP} + \text{ATP} \overset{\text{TMPK}}{\Rightarrow} \text{dTDP} + \text{ADP}$$
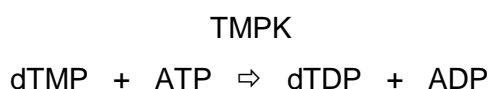
TMPK lies at the junction of the ***de novo* and salvage pathway of the thymidine triphosphate** (dTTP) metabolism (Figure 1.15) and it is the last specific enzyme for its synthesis, making it an attractive target for drug design.
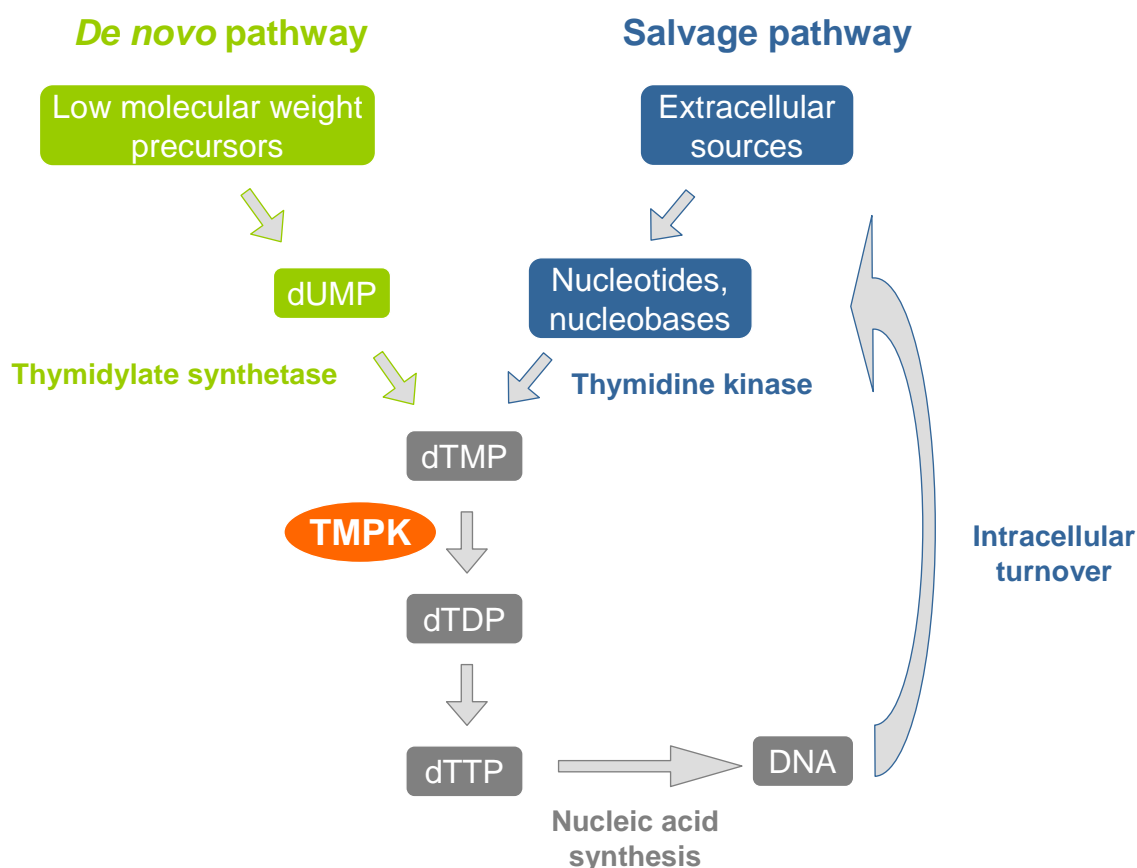
Figure 1.15 Salvage and *de novo* pathway of dTTP metabolism.

In TMPK-deficient *Saccharomyces cerevisiae* cells, Jong et al. proved that yeast TMPK is essential for DNA-replication and thus for cellular growth.[91,92] Moreover, a correlation between thymidylate kinase activity and the rate of DNA-synthesis during rapid growth has been documented, while in non-proliferating tissues, enzyme activity is barely detectable.[93]

Although the global folding of this enzyme is similar to that of the human TMPK (TMPKh), there is **only 22% of sequence identity between the human and the mycobacterial isozyme**. This strongly supports the possibility to design inhibitors which are **selective for the mycobacterial enzyme**, thereby reducing or excluding toxicity for human cells and side effects. The first selective competitive inhibitor found to bind to mycobacterial TMPK (TMPKmt) was AZTMP.[94] By now, several nucleoside- and nucleotide-derivatives have been found to selectively inhibit mycobacterial TMPK.

1.2.2 The structure of TMPKmt

In 2001, Li de la Sierra et al.[95] described the crystal structure of TMPKmt on a 1.95 Å resolution X-ray image. This structure shows TMPKmt, bound to its natural substrate dTMP, with a sulphate ion mimicking for the β-phosphoryl group of ATP.

1.2.2.1 Primary, secondary, tertiary and quaternary structure

TMPKmt has a molecular mass of 24 kDa and consists of 214 amino acid residues. It is a homodimer in solution, with each monomer having nine α-helices surrounding a five-stranded β-sheet core. At the interface of the dimer, three pairs of helices are kept together through hydrophobic interactions and an ion pair between Glu50 and Arg127.[95]

The global folding of TMPKmt is similar to that of other TMPKs. Nevertheless only a low similarity of their amino acid sequences is found:

- 26% with TMPK of *E. coli*
- 25% with TMPK of yeast
- 22% with TMPKh

During binding of the substrate into the active site of TMPKmt, large conformational changes can be observed.   In the absence of substrate, the enzyme appears in the open conformation.   This changes to a partially closed conformation in the presence of one substrate, and finally to a fully closed and active conformation when both substrates are available (dTMP and ATP).

Within the 214 amino acid residues, the essential residues for the function of the TMPKmt can be found in three regions:

♦ The P-loop motif (Gly $X_1X_2X_3X_4$ Gly Lys X_5)[96]

This motif of 8 amino acid residues is responsible for the positioning of the phosphoryl groups of the phosphate donor.   Through interactions between the amide backbone hydrogens and the phosphate oxygen atoms, it binds and positions the α- and β-phosphoryl groups.   Unique for TMPKs is an interaction between the P-loop and the phosphoryl acceptor dTMP. This is achieved through a hydrogen bond between the 3'-hydroxyl group of dTMP and a carboxylic acid residue at position X_2 of the P-loop motif (Asp9 in TMPKmt). Although the exact function of this acidic amino acid is still unclear, it appears to be crucial for the catalytic mechanism of TMPK. Any mutation of this residue in the yeast, human or *E. coli* enzyme disables the phosphoryl transfer activity.[97]

♦ The DR(Y/H/F) motif

This region contains a strictly conserved arginine residue acting as a clip which favourably orients the phosporyl donor and acceptor via an interaction with the phosphate group of dTMP and the γ-phosphate of ATP.

♦ The LID region

Upon binding of the phosporyl donor to the enzyme, the LID region closes as a flexible stretch. Important conformational changes appear when ATP is fixed onto the enzyme. At the same moment, there is a transition from a coil to a helical conformation in the LID region. Phosphoryl groups of both substrates interact with arginine side chains in the LID region, revealing the importance of these residues in catalysing phosphoryl transfer.[95]

The presence of those three important amino acid regions is an important feature of all TMPKs. Uniquely in the mycobacterial TMPK, **a magnesium ion** binds the enzyme transiently during catalysis, establishing a link between the P-loop (Asp9) and the LID region (Glu166). Giving rise to a strong electrostatic potential, the γ-phosphate of the phosphoryl donor ATP is attracted very close the α-phosphate of dTMP, thereby acting as a clamp between phosphor donor and acceptor. For the human enzyme, one magnesium ion binding site has been reported, located further away along the ATP-binding site, between the β- and γ-phosphate groups. For the yeast and the *E. coli* enzyme, no magnesium atoms have been reported. Next to its unique presence in TMPKmt, the essential role of this magnesium ion during catalysis, makes it an attractive target to destabilise it in order to disturb catalysis.

a) The dTMP-binding site

In order to design competitive inhibitors for TMPKmt in a rational way, the main interactions with its substrate have to be taken into account. Substrate derivatives envisioned to interact more strongly with the active site than the substrate itself, can in this way be constructed, synthesized and generate interesting lead compounds for enzyme inhibition. As observed by the X-ray structure of TMPKmt, the main bonding forces between dTMP and the enzyme are the following:

- A stacking interaction between the pyrimidine nucleobase and Phe70;
- A hydrogen bond between O-4 of thymine and the Arg74 side-chain. This results in greater affinity for the thymidine compared to cytidine;
- A hydrogen bond between Asn100 and N-3 of the thymine ring;
- A hydrogen bond between 3'-hydroxyl of dTMP and the terminal carboxyl of Asp9, that in its turn interacts with the magnesium ion, responsible for positioning the phosphate oxygen of dTMP;
- Hydrogen bonds and ionic interaction between the 5'-O-phosphoryl and Tyr39, Phe36, Arg95 and Mg^{2+} respectively;
- The presence of Tyr 103 close to the 2'-position is believed to render the enzyme catalytically selective for 2'-deoxy-nucleotides versus *ribo*-nucleotides.
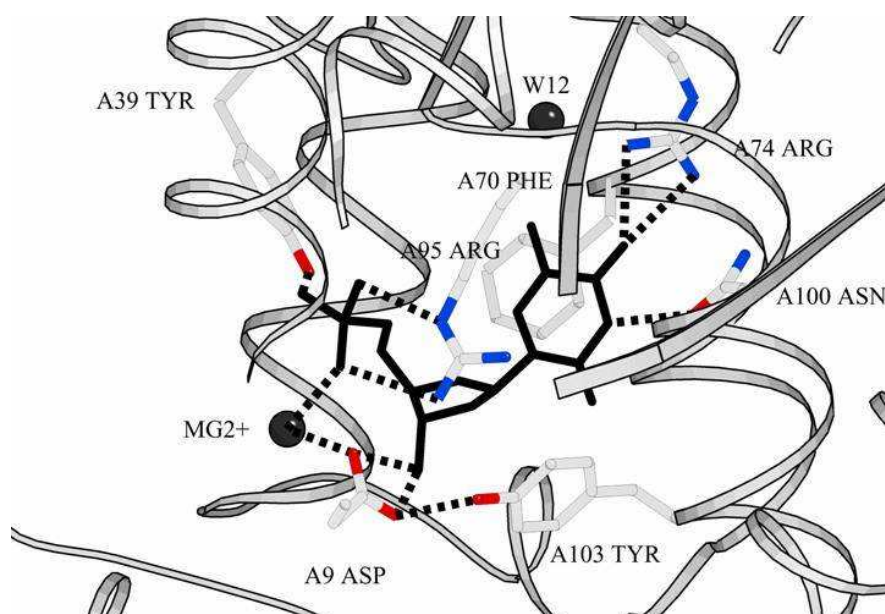
Figure 1.16 The active pocket of TMPKmt

In order to obtain selective inhibition, the differences with other TMPK's have to be taken into account:

-   The hydrogen bond between the N-3 atom of the pyrimidine ring with the Asn100 sidechain (corresponding to a glycine in the yeast and human enzyme and to a threonine in the *E. coli* enzyme);
-   The Tyr39 residue makes two polar contacts with dTMP: one with the phosphate oxygen atom and the other with the oxygen atom at the 5'-position. Arginine and glycine residues replace this tyrosine in the human and yeast enzymes respectively;
-   The 3'-hydroxyl group of the ribose moiety makes three polar contacts: one with water molecule involved in the Mg^{2+}-coordination and two others with Asp9.[95]

b) The ATP binding site

The crystal structure of TMPKmt, co-crystallised as a complex with dTMP and SO_4^{2-}, in the absence of ATP, is similar to those observed for yeast and *E. coli* enzymes, where ATP is fixed at its donor position. This indicates that this crystal structure represents the closed conformation, Arg149 and Thr15 interacting with the adenine ring of ATP, accommodating it at the correct position for enzyme activity.[95]

TMPKmt has broader substrate specificity for nucleoside triphosphates than other nucleotide monophosphates kinases. The reaction rates with ATP or dATP as phosphate donors are similar. ITP, GTP, CTP and UTP also serve as efficient phosphate donors (order of efficacy:

ITP > GTP > CTP > UTP). Even dGTP, dCTP and dTTP can act as phosphate donors, although less efficient (1% for dTTP, 8% for dCTP and 35% for dGTP compared to ATP).[93]

1.3 RATIONAL DRUG DESIGN

While in the past drugs were discovered rather than being developed, the scientific evolution afforded several helpful tools that changed the general drug discovery process significantly. First of all, numerous target structures have been elucidated by X-ray or NMR, clarifying the requirements for effective binding to the enzyme, which can be useful to improve the potency and specificity of new lead compounds. In addition to providing structural information, computer models are used to simulate the binding mode of ligands to the known targets. In this way, new ligands with a higher probability of effectively binding a target can be predicted and synthesized. Combined with **combinatorial chemistry**, **computer-aided molecular design** is now routinely used in lead identification and optimisation.

In cases where the target of a series of compounds is unknown or a reliable structure of a target is not available yet, molecular modelling is a useful tool to analyse the structure-activity relationship of a series of ligands. This can lead to more active compounds. A technique implementing this strategy is the Quantitative Structure-Activity Relationship (QSAR), which is based on a set of analogues of a biologically active compound. By the use of descriptors (e.g. MW, log P,…), the physicochemical properties of a compound are identified and quantified. A correlation is determined between these descriptors and the corresponding activities, which is then used for the prediction of the activities of new analogues.

Based on 47 earlier synthesized thymidine derivatives, Aparna et al.[98] performed a 3D-QSAR study for TMPKmt inhibitors. From this study, the following structural requirements for strong TMPKmt inhibition could be concluded:

- C-5: Moderately bulky substituents, preferably electronegative groups are favored.
- C-2': A hydrogen bond donor at this position may lead to drop in activity.
- C-3': Moderately bulky substituents are favored.
- C-5': Either hydrogen bond acceptors or donors can be accommodated. Bulky and highly polar substituents are favored.

Both strategies are used in this work in order to find new potent TMPKmt inhibitors in a rational way.

To illustrate the relevance of structure based drug design, the successful development of HIV-protease inhibitors, based on crystal structures of inhibitor/HIV protease complexes is a textbook example. Several papers[99,100] describe the rational design of protease inhibitors, which even resulted in compound **1.38**, a protease inhibitor with a K_i in the low picomolar range (4.5 pM) and active against drug-resistant HIV.[101]
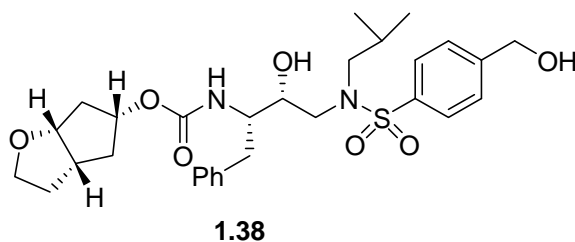
1.38

Figure 1.17 HIV-protease inhibitor **1.38**, the result of successful rational design

1.4 INITIAL STRUCTURE ACTIVITY RELATIONSHIP

Earlier evaluation of thymidine analogues, synthesized in our lab as inhibitors of TMPKmt has generated useful information about their structure-activity relationship (SAR).

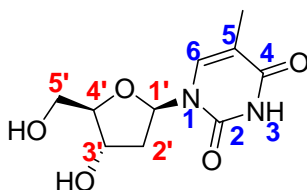The generally accepted numbering system for nucleosides is as depicted in Figure 1.18:

Figure 1.18 Generally used numbering system for nucleosides

The ribofuranose ring numbering can be distinguished from the pyrimidine numbering by the use of the prime after the numbers.

1.4.1 Base-modifications
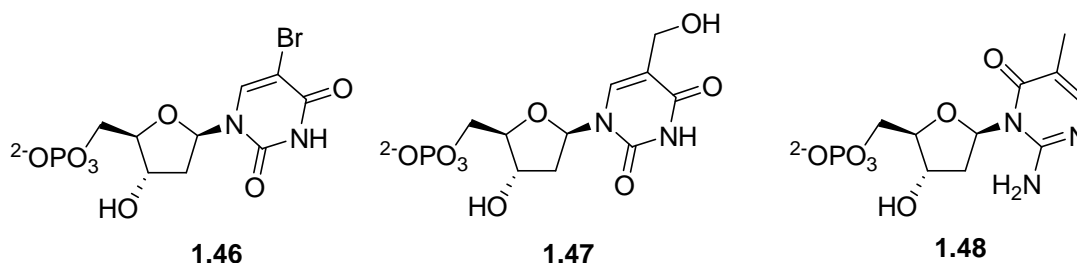
Figure 1.19 Base modifications

♦ 5-position

The thymine nucleobase has been modified at its 5- and 2-position. Modifying the 5-methyl into several halogens, resulted in TMPKmt-substrates with the 5-Br-dUMP (**1.46**) being the most active one (K_m= 30 μM).[95] Other halogens at this position like iodine or fluorine extremely decreased the affinity for the enzyme. Complete removal of the methyl group (dUMP) affected the affinity even more, resulting in a 50-fold increase in the K_m-value. The difference in kinetic parameters between the different 5-substituted dUMP-analogues seemed related to the ability of the differently sized halogen atoms to fill a cavity.[102] Later, larger substituents at the 5-position have been introduced in an effort to form stacking interactions with Phe36 (e.g. furanyl, thienyl or benzyloxymethyl) or ionic interactions with Arg74 (5-hydroxymethyl). In contrast to the previously synthesized 5-halogen substituted

compounds, none of these derivatives behaved as a substrate for TMPKmt. Moreover, since 5-hydroxymethyl (**1.47**) and 5-furan-2-yl dUMP were the most active compounds with K_i-values of 110 and 140 μM respectively, it was concluded that **the volume of this cavity cannot be stretched too much**, limiting further explorations at this position.[103]

♦ 2-position

At the 2-position of the pyrimidine ring, modifications have been introduced in order to enhance the inhibitory activity on the enzyme, which resulted in 5-methyl-iso-dCMP (**1.48**) as a competitive inhibitor with a K_i of 130 μM.[95]

1.4.2 Sugar modifications

♦ 3'-position

1.49

1.50: R_1= NH_2
1.51: R_1= F

1.52: R_2= N_3
1.53: R_2= NH_2
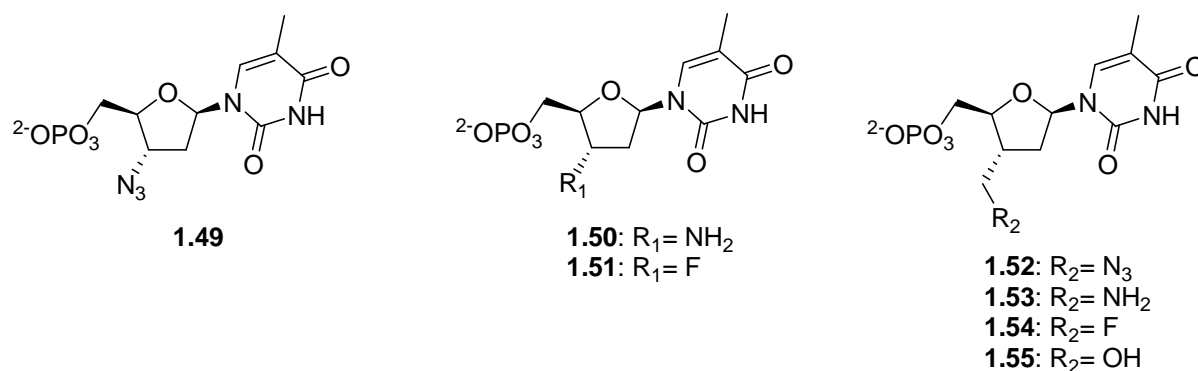**1.54**: R_2= F
1.55: R_2= OH

Figure 1.20 3'-modified thymidine derivatives

The first reported selective competitive inhibitor of TMPKmt is AZTMP (**1.49**), having an azide at the 3'-position instead of a hydroxyl in the natural substrate. While TMPKs from eukaryotes and other prokaryotes recognize it as a substrate, AZTMP showed a K_i of 10 μM for TMPKmt.[95] The complex of the natural substrate dTMP to TMPKmt is characterized by an important hydrogen bond between 3'-OH and the carboxylic acid of Asp9. In the case of AZTMP however, no stabilizing hydrogen bond can be formed. The X-ray structure of TMPKmt complexed with AZTMP demonstrated that binding of AZTMP prevents the positioning of Mg^{2+} in the active site and consecutive LID-closure. Moreover a direct interaction between the azidofunction and Asp9 causes an inadequate positioning of this important amino acid residue. As a result, catalysis abolishes.[104]

Considerable efforts to find strong inhibitors were directed towards a strong ionic interaction between a 3'-substituent and the carboxylic acid of Asp9. Despite the fact that modelling pointed towards a 3'-amino group, 3'-aminothymidine monophosphate (**1.50**) showed only

disappointing inhibitory activity (K_i= 235 µM). Replacement of a 3'-OH by a 3'-F afforded analogue **1.51** that behaved as a substrate with a K_m of 30 µM.[105]

As it was anticipated that the 3'-NH$_2$ group was unable to approach Asp 9 close enough to establish the envisaged interaction, a series of 3'-branched-chain nucleotides was synthesized, i.e. 3'-*C*-azidomethyl (**1.52**; K_i= 12 µM), 3'-aminomethyl (**1.53**; K_i= 10.5 µM), 3'-fluoromethyl (**1.54**; K_i= 15 µM) and 3'-hydroxymethyl derivatives (**1.55**; K_i= 29 µM) of dTMP. Their activities indicated **the possibility of the enzyme to accommodate sterically more demanding substituents at the 3'-position**. Since no significant difference in K_i-values among these 3'-*C*-branched chain analogues was obtained, it remained uncertain if the 3'-CH$_2$NH$_2$ was really involved in an ionic interaction with Asp9.[106]

In an attempt of Veerle Vanheusden to synthesize 3'-aminomethyl-3'-deoxythymidine, Barton deoxygenation of a 2'-O-phenylthionocarbonate substituted 3'-azidomethyl-ribothymidine, resulted in three peculiar nucleoside analogues, being two bicyclic nucleosides **1.56** and **1.57** (discussed later) and one dinucleoside **1.58**.[107]

T = thymin-1-yl

Ki (TMPKmt) =	3.5 µM	13.5 µM	37 µM
Ki (TMPKh) =	700 µM	1100 µM	n.i. at 1 mM
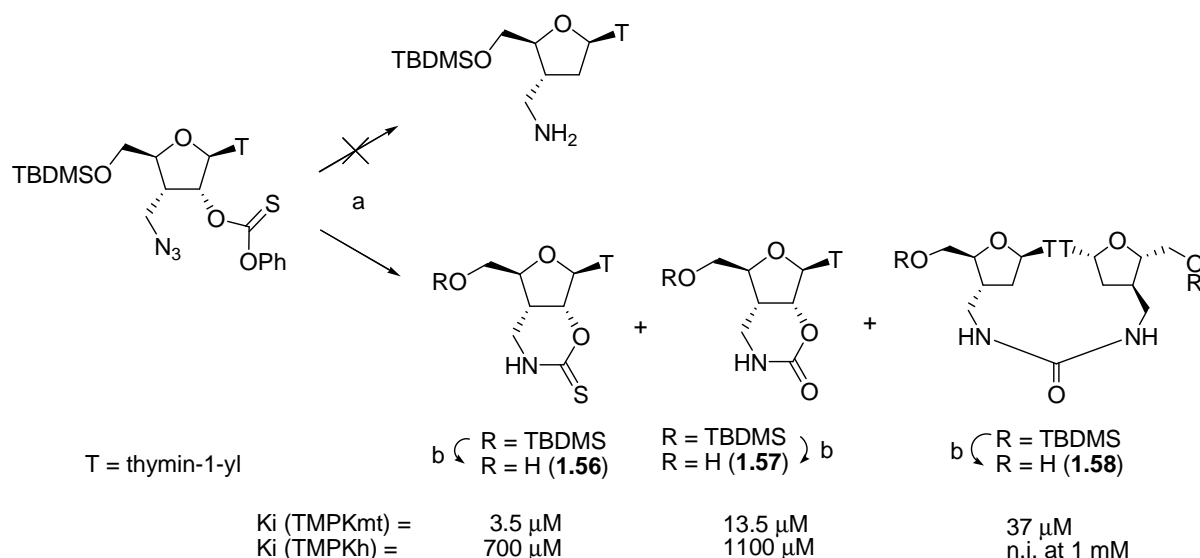
Km (TMPKmt) of the natural substrate: 4.5 µM

Reagents: (a) AIBN, Bu$_3$SnH, toluene; (b) TBAF, THF

Scheme 1.1 Serendipitous discovery of three potent TMPKmt inhibitors

As the X-ray structure of TMPKmt did not point towards the presence of a binding site for a second nucleoside at the 3'-position, the activity of the dinucleoside (K_i= 37 µM) was very surprising. Modelling experiments in GOLD showed that the sugar ring of the first monomer I binds the TMP-pocket upside down, allowing the hydrogen of its 5'-hydroxyl to interact with Asp9 (Figure 1.21). This binding mode permits the sugar ring of the second monomer (II) to be directed towards the outside of the enzyme where the phosphoryl donor normally binds

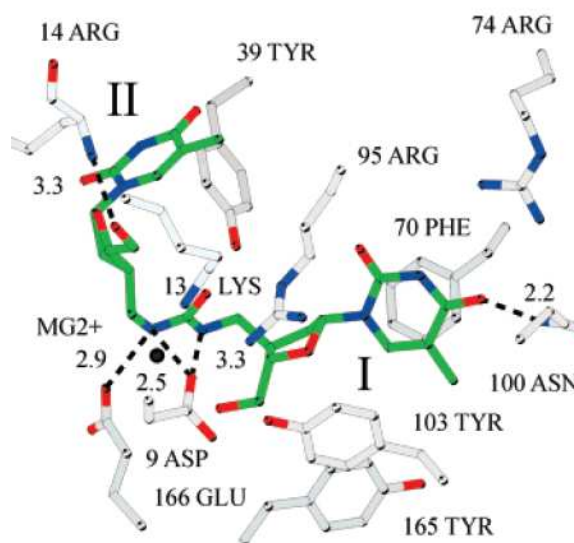and opened new possibilities for the design of much larger inhibitors for TMPKmt than explored earlier.

Figure 1.21 Predicted binding mode of dinucleoside **1.58** with TMPKmt

♦ 5'-position

At the onset of our research group's search for TMPKmt inhibitors, the aim was to synthesize modified nucleoside 5'-*O*-phosphates that were not prone to further phosphorylation in order to avoid cell toxicity. A main disadvantage of this approach however, concerned the uptake of 5'-*O*-phosphorylated analogues in cells, generally accepted to be very difficult. Normally, nucleosides cross the cell membrane barrier in the non-phosphorylated form and are then phosphorylated by intracellular kinases. However, in this particular case, intracellular activation is not possible, as no gene coding for TK can be identified in the *M. tuberculosis* genome.

Fortunately, previous work by Vanheusden et al. demonstrated that certain non-phosphorylated thymidine analogues (**1.52**-**1.55**) were equally active compared to their phosphorylated analogues. At the same time, these analogues showed only a low affinity for the human enzyme resulting in selective TMPKmt inhibitors.[106] Our research will consequently be directed toward the synthesis of non-phosphorylated TMPKmt inhibitors.

Other 5'-modified dT-analogues showed a positive influence of the functional groups at the 5'-position on the interaction between inhibitor and the enzyme. The K_i-value of 5'-NH$_2$dT (**1.59**) and 5'-N$_3$dT (**1.60**) are 12 and 7 µM respectively, both close to the K_m of the natural substrate. For this reason, **azido- and amino-modifications seem to be attractive substitution patterns** to result in potent inhibitors.[108]
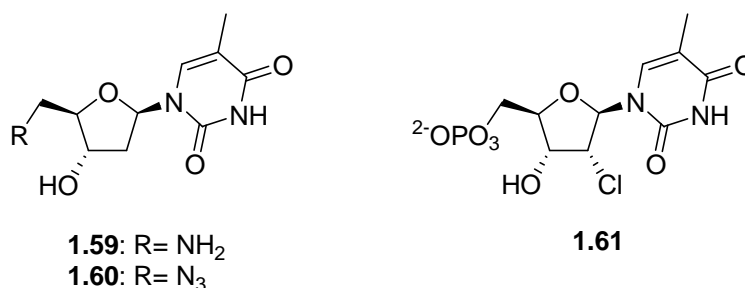
1.59: R= NH$_2$
1.60: R= N$_3$

1.61

Figure 1.22 5'- and 2'- modified dT and dTMP analogues

♦ 2'-position

As described earlier, in the active site of the enzyme Tyr103 is responsible for a sterical hindrance at the 2'-position resulting in **a higher activity for 2'-deoxynucleosides** compared to their *ribo* analogues. Only in the case of a 3'-amino modification, the 2'-ribonucleotide (K_i= 45 µM) proved more active than its 2'-deoxy analogue (K_i= 235 µM).[104]

Surprisingly, the introduction of halogens appeared to be better tolerated than a hydroxyl group at this position. 2'-Chloro-dTMP (**1.61**), for example, showed a K_i of 19 µM for the mycobacterial enzyme. [105] Modelling indicated that the introduction of a chlorine influenced the puckering of the sugar. This allowed the chlorine to occupy the pocket where the 3'-hydroxyl group normally resides.

Van Rompaey et al.[109] described the synthesis and conformational analysis of 2'- and 3'-modified thymidine analogues. This study showed that ribo- configured **halogens at the 2'-position limited the conformational freedom** of the nucleoside and forced it into a Northern conformation which is opposite to the conformation of the natural substrate, explaining the low affinity for the enzyme. A more detailed explanation of the sugar puckering is discussed below.

♦ 1'-position

The natural substrate dTMP has the nucleobase at the 1'-position above the sugar plane, which makes it, like most natural nucleosides, a β-nucleoside. Earlier, an **undesired α-anomer** (3'-azidomethyl-α-thymidine) had been synthesized and tested for TMPKmt affinity. It showed an **unexpected high affinity** for the enzyme with a K_i of 29 µM compared to 40 µM for its β-derivative. At that time, no efforts were made to further explore this interesting observation.[107]

♦ Bicyclic nucleosides (2'-3'-substituted)

The two bicyclic nucleosides, thiocarbamate **1.56** and carbamate **1.57** were found to be very potent inhibitors of TMPKmt. With a K_i of 3.5 µM, the thiocarbamate showed the highest

TMPKmt inhibitory activity at that time. Moreover, these compounds showed a very high selectivity.[107]

Due to the extra space needed to fit the bicyclic compounds in the active site, the high activity of these compounds was unexpected. Modelling indicated that the 2',3'-substituted 6-membered ring forces the 6'-nitrogen into the most appropriate position for interaction with the Asp9 residue in the enzyme cavity. The sulfur atom, in its turn, forms hydrophobic interactions with Tyr103 and Tyr165 residues. The smaller oxygen in the carbamate compound fills the 3'-cavity less efficiently, which resulted in a lower affinity for the enzyme. However, despite their high affinities for the enzyme both bicyclic nucleosides failed to inhibit the growth of *M. tuberculosis* H37Rv strain, limiting further potential of these compounds as antituberculosis agents. Nevertheless we believed these hits were **valuable starting points for further optimisation**.

♦ Sugar puckering

Apart from the substitution pattern of the sugar, **the conformational behaviour of nucleosides is considered to be of great importance** for their interaction with target proteins and hence for their biological activity.[110] As a planar furanose conformation is energetically highly unfavourable, the ring atoms are displayed out of the plane, giving rise to ring puckering.

The complete definition of the conformation of a nucleoside involves the determination of three principal structural parameters: (1) the glycosyl torsion angle χ, which determines the relative position of the base to the sugar moiety (syn or anti); (2) the torsion angle γ, determining the orientation of the 5'-OH with respect to C3'; and (3) the puckering of the furanose ring (pseudorotation P) and its deviation from planarity (maximum out-of-plane pucker ν_{max}).[111] In the present study the most important determinant is the ring puckering because it is able to influence the χ- and γ-angle and it is established that TMPKmt prefers to bind its substrate and related analogues with the thymine moiety in the anti-position and prefers a specific rotamer (γ-angle) around the C4'-C5'- bond upon binding. By convention, a pseudorotation angle P = 0° corresponds to the ab solute Northern conformation *(C2'-exo-C3'-endo)*, whereas the Southern conformation (*C2'-endo-C3'-exo*) is represented by P= 180°.[112] The sugar puckering of natural ribo- and deoxyribonucleosides in solution appears to be a dynamic equilibrium between those two major conformers, the North (N) and the South (S) type, with a very low energy barrier between the N- and the S-state. These two conformations are depicted in the pseudorotational cycle, shown in Figure 1.23.
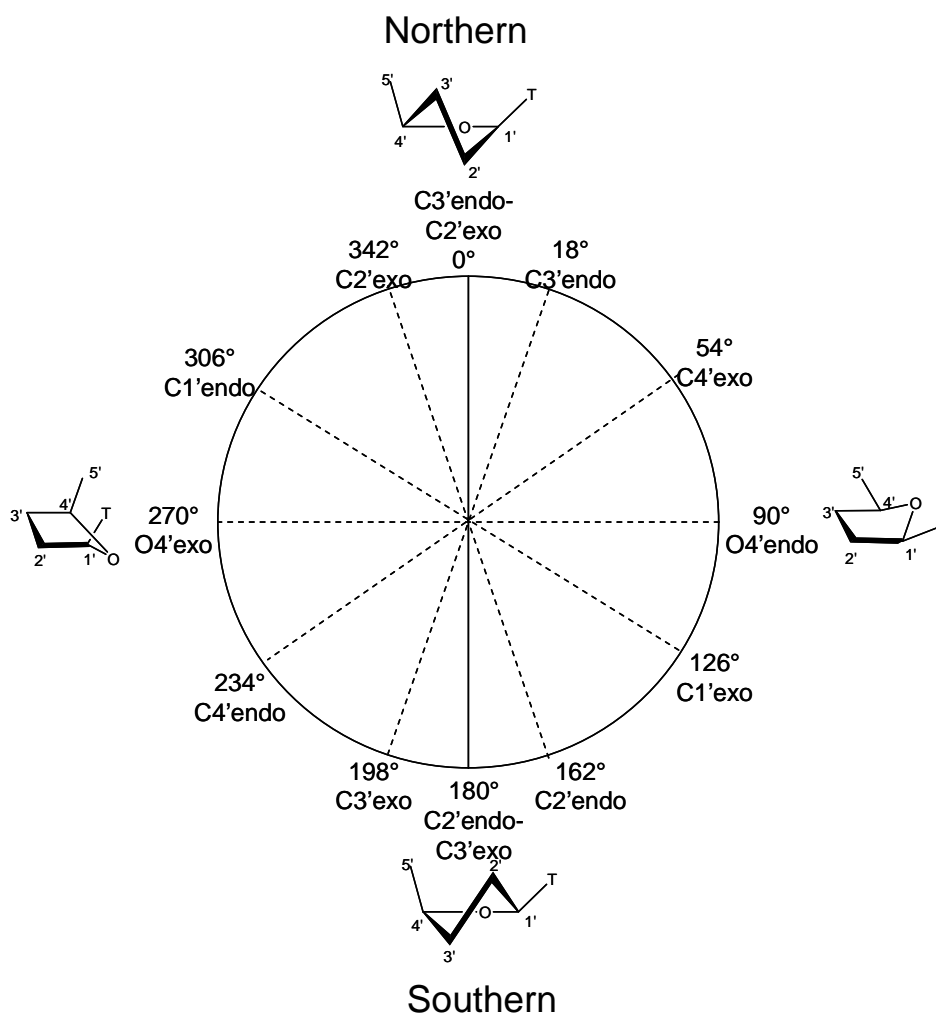
Northern

Figure 1.23 Pseudorotational cycle

In the solid state, one of the two solution conformations is usually present, determined mainly by crystal packing forces. The activity on the enzyme is mainly determined by the adopted sugar puckering of the nucleoside in solution. The more this puckering corresponds to the conformation in which the nucleoside binds to the enzyme, the easier this nucleoside fits into the active site.

Additionally, other structural features of nucleosides like rotamers about C1'-N1 and C4'-C5' and C5'-O5' bonds affect the affinity for the enzyme. As a result, conformation-activity studies based exclusively on solid-state conformational parameters are often not very significant.

Since the X-ray analysis of the dTMP-TMPKmt complex indicates that **the substrate adopts a Southern conformation** with the base in an equatorial orientation,[95] one would expect that nucleoside analogues that exhibit a similar S-type conformation would have an optimal affinity for TMPKmt.[109]

Each of the substituents influences the conformation of the sugar which complicates the prediction of the preferred conformation of new nucleoside analogues. Additionally, the

sugar has a constant conformational mobility between a North and South configuration. This makes it first of all difficult to determine the general conformation of a molecule in solution, and when the energy needed to change between the two conformations is very low, this diminishes the influence of the conformation on activity. Moreover, the amino acid chains in the active site have some flexibility, allowing the enzyme to be favourable for one certain conformation for one type of modification and to be disadvantageous for another one. This is illustrated by the activity of the earlier described 3'-*C*-branched chain nucleosides. Modelling indicated that these nucleosides bind to TMPKmt in the N- (northern, 2'-exo-3'-endo) conformation. While this conformation is known to be the opposite of the natural substrate binding to the enzyme, in this case it is favourable as it enables interactions between the 3'-C-substituent and Asp9. Thus, although the conformational behaviour of nucleosides is an important feature, it renders the prediction of binding affinities of nucleoside analogues difficult.

For this reason, conformationally restrained nucleosides have been synthesized in which an extra ring locks the conformation in one preferred conformation. This makes it possible to predict and determine this one conformation, and to study the influence of the conformation on the biological activity.

1.5 OBJECTIVES

1. Rational design of TMPKmt inhibitors

By making use of different techniques of rational drug design, described in 1.3, we aimed to synthesize potent and selective inhibitors of TMPKmt. Both modelling, docking and Quantitative Structure-Activity-Relationship analysis were used throughout this work to facilitate the search for interesting TMPKmt inhibitors.

2. Profound understanding of enzyme binding

As the X-ray structure of the enzyme is known, the way of binding of new inhibitors can be studied in detail. By making use of docking techniques, clear and innovative design hypotheses could be stated, and consequently be tested and proven by the synthesis of thymidine derivatives.

3. Applicability as useful drugs

The overall ambition of this thesis goes beyond enzyme inhibition. Although bacterial uptake puts an extra challenge on molecule synthesis, our clear goal was to integrate and incorporate the different aspects of useful drug design. For this reason, we focused not only the inhibitory activity and selectivity, but also the low toxicity and bacterial uptake were taken into account.

1.6 STRUCTURE OF THIS THESIS

In chapter 2 a number of hybrid nucleosides were designed, based on modifications that were earlier discovered as favourable for good affinity for TMPKmt, i.e. 5'-amino-3'-azido and 2'-fluoro-2'-deoxy-arabino thymidine analogues.
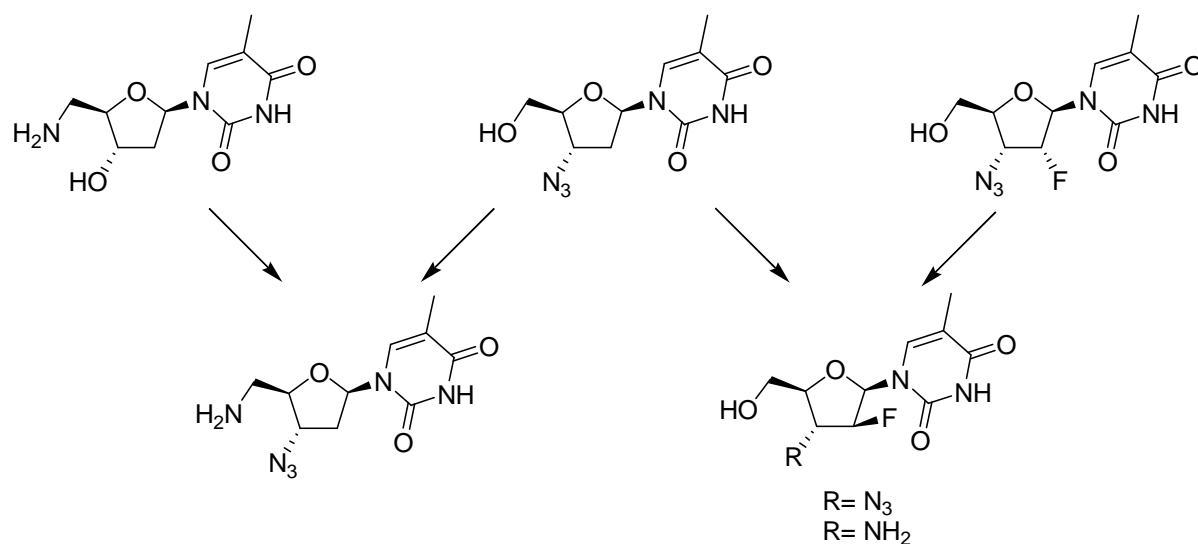

Figure 1.24 Target molecules, described in chapter 2

As described in the initial structure-activity relationship, the presence of sterically more demanding nitrogen containing substituents at the 3'-position resulted in good activities.

Moreover some interesting compounds have been discovered recently, two bicyclic compounds and a dinucleoside, all showing a good inhibitory activity, and characterized by an extra carbon atom at the 3'-position. Some extra attention will therefore be paid to the introduction of an extra carbon at this position and is described in chapter 3.

Two interesting recently discovered potent TMPKmt inhibitors are the 2'-3'-bicyclic nucleosides. The synthesis of alternative fused 5- and 6-membered rings intends the optimisation of those compounds and can yield a better understanding of the binding mode of those compounds.
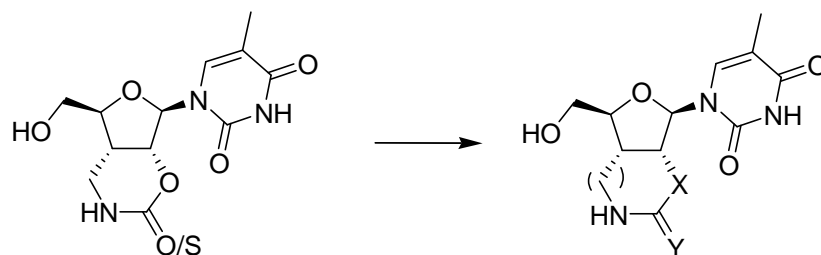
The discovery of dinucleoside **1.51** as TMPKmt inhibitor illustrated the possibility of TMPKmt to accommodate sterically more demanding substituents at the 3'-position.   In order to explore the feasibility to replace the second nucleoside, a series of 3'-branched thiourea substituted derivatives was envisaged.   Modelling and the studied structure-activity relationship can be used to develop new modifications in a rational way.

Figure 1.26 Exploration of dinucleoside **1.58**, described in chapter 5

Finally, the synthesized compounds were tested for their activity against other kinases than TMPK, which is described in Chapter 6.

Chapter 2

Synthesis of 2'- and 5'-modified thymidine derivatives
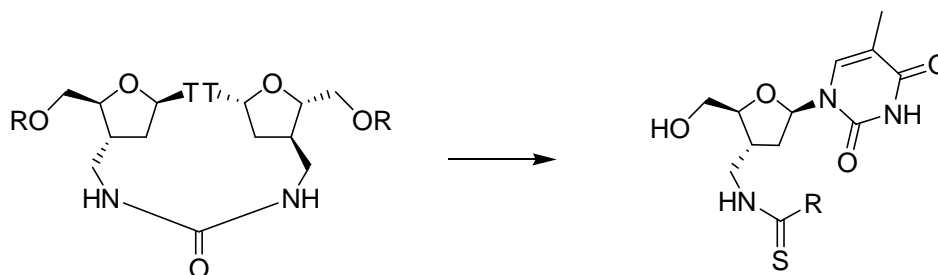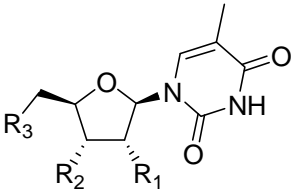
---

2 SYNTHESIS AND BIOLOGICAL EVALUATION OF 5'-AMINO AND 2'-FLUORO SUBSTITUTED 3'-AZIDO-3'-DEOXYTHYMIDINE DERIVATIVES

2.1 OBJECTIVE

Combining modifications at different positions of the thymidine scaffold, individually leading to potent inhibitors, may result in more active compounds. Consequently, the introduction of a 5'-amino function (as in **2.4**) together with a 3'-azido group (as in AZT) on thymidine, was proposed as an interesting target molecule. Apart from the expected synergistic activity of the two modifications, the **5'-amino-3'-azido-substituted nucleoside** opens interesting perspectives for further modifications at the 5'-position (e.g. amideformation, ureum, carbamates,…).

Table 2.1 Kinetic parameters of TMPKmt with compounds **2.4-2.6**

Compound	R_1	R_2	R_3	K_i (µM) TMPK mt
dTMP	H	OH	OPO_3^{2-}	4.5 [a]
AZT	H	N_3	H	28
2.4[108]	H	OH	NH_2	12
2.5[109]	F	N_3	OH	118
2.6[109]	F	NH_2	OH	190

[a] K_m-Value

A second objective is based on **the influence of fluorine** on the conformation of a furanose ring. Fluorine is known to serve as a hydroxyl mimic and to participate in hydrogen bonding.[113] Its van der Waals radius (1.47 Å) is intermediate between that of hydrogen (1.2 Å) and oxygen (1.6 Å) and its electronegativity is similar to that of hydroxyl, making it act as a H-bond acceptor. Replacing a 2'-H atom by a fluorine has a significant influence on the resulting nucleoside. Such modification not only stabilizes the glycosidic bond,[114] it also influences the conformation of the furanose ring.[115]

The two main influences which affect the conformation of the furanose ring are the gauche effect and an additive anomeric effect.

The **anomeric effect** favours the pseudoaxial position of an electronegative ligand at the C1' position. For the β-D-nucleosides, a pseudoaxial position at the C1' position forces the sugar to adopt the North conformation (Figure 2.1). Additionally, the anomeric effect is influenced by other effects like the stereo-electronic competition at the anomeric centre[116] and the influence of the chemical nature of nearby substituents on the anomeric strength.[117]

The **gauche effect** causes the tendency of two vicinal electronegative ligands on a C-C bond to adopt a gauche conformation along this C-C bond. This effect drives the sugar to the South conformation.

As described earlier, a 2'-fluoro substituent in the ribo configuration (**3.5** and **3.6**) locks the sugar pucker in the North or C3'-endo conformation. In a 2'-fluoro-ara-nucleoside, the gauche effect opposes the much weaker anomeric effect, resulting in a favourable South conformation.[114]
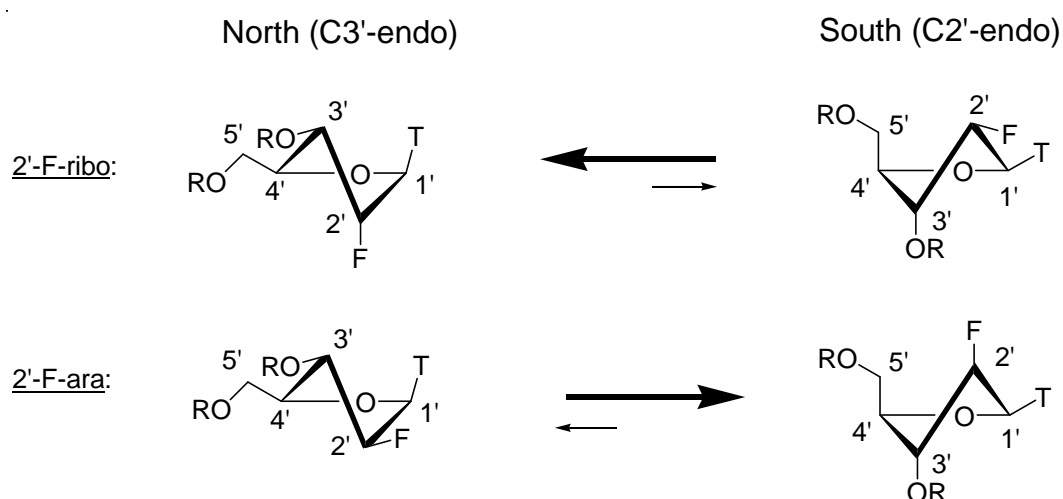

Figure 2.1 Conformational equilibrium for 2'-F-ribo and 2'-F-arabino thymidine. The anomeric effect in the 2'-F-ribo configuration dominates, leading to a preferred North conformation. For the 2'-F-arabino configuration, the gauche effect opposes the weaker anomeric effect, resulting in a favourable South conformation.

As dTMP binds TMPKmt in the South conformation, this conformation is expected to be preferable for compounds to exhibit a good binding affinity for the enzyme. The opposite was illustrated by the weak activity of 3'-NH$_2$ and 3'-N$_3$ substituted derivatives containing a 2'-F in ribo configuration (**2.5** and **2.6**). Conformational analysis based on NMR coupling constants proved these compounds to be biased in North conformation in solution, which explains their moderate activity.[109]

Consequently, we considered the synthesis of analogues that combine a 3'-NH_2 or 3'-N_3 modification and a 2'-F-ara-substituent an interesting objective.
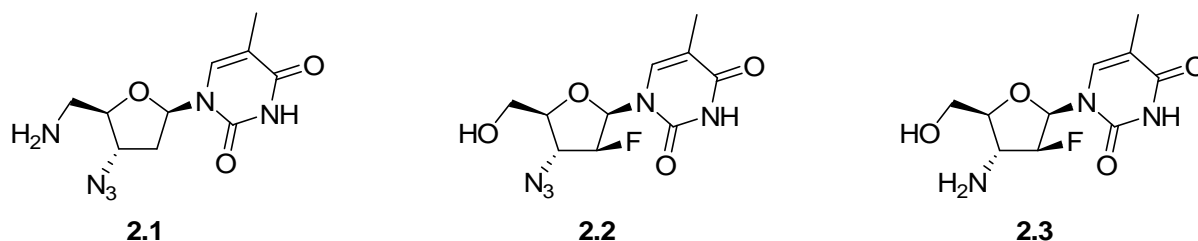

Figure 2.2 Target molecules

2.2 CHEMISTRY

For the synthesis of 5'-amino-3'-azido-3',5'-dideoxythymidine (See Scheme 2.1), we selected the natural nucleoside β-D-thymidine as a starting compound. The main difficulty is the presence of two different nitrogen-substituents in the target analogue, a 5'-amino group and a 3'-azido function. An azido group can be introduced rather easily and is inert to a lot of reaction conditions. It is often used as a protected amine that can be 'deprotected' easily by Staudinger reduction or hydrogenolysis. Regioselective azido reduction however is not possible, and consequently the two nitrogen containing substituents need to be introduced separately into this molecule.

Based on the work of McGuigan et al.[118] we used a tandem Mitsunobu reaction. Triphenylphosphine combines with DEAD to generate a phosphonium intermediate that binds to the 3'- and 5'-alchohol oxygen, activating them as a leaving group. The 3'-leaving group immediately leads to 2,3'-*O*-anhydrothymidine, while the 5'-leaving group is substituted by phtalimide. Opening of this 2,3'-*O*-anhydroring with NaN_3 and subsequent deprotection of the 5'-phtalimide with hydrazine gave final compound **2.1** in an overall yield of 31%.
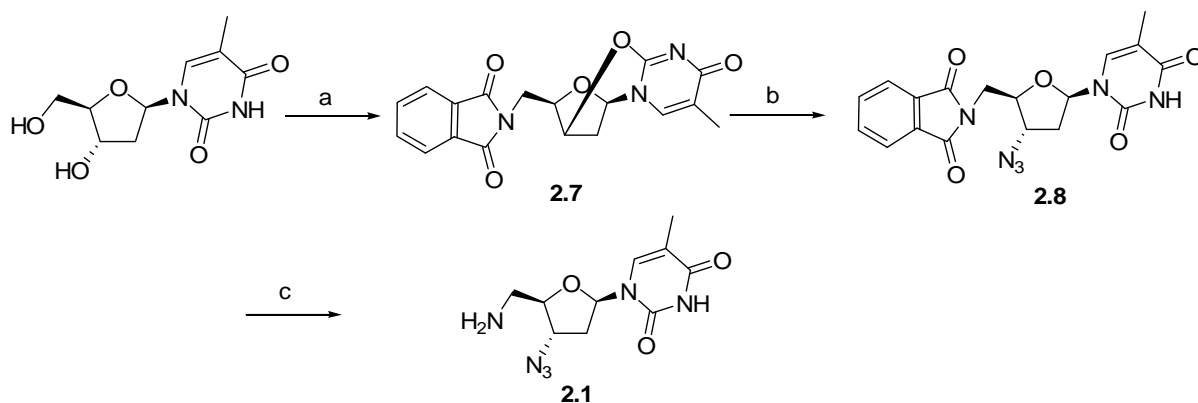
Scheme 2.1, reagents and conditions: (a) phtalimide, Ph_3P, DEAD, THF, rt, 16h, 60%; (b) NaN_3, DMF, benzoic acid, 150 °C, 2h, 70%; (c) hydrazine, EtOH, 28 °C, 2d, 73%.

For the 2'-fluoro arabino compounds, the starting material of the synthesis was the 1,2-*O*-isopropylidene-α-D-xylofuranose **2.8**.

The synthetic pathway depicted in Scheme 2.2 proved high-yielding and straight forward.
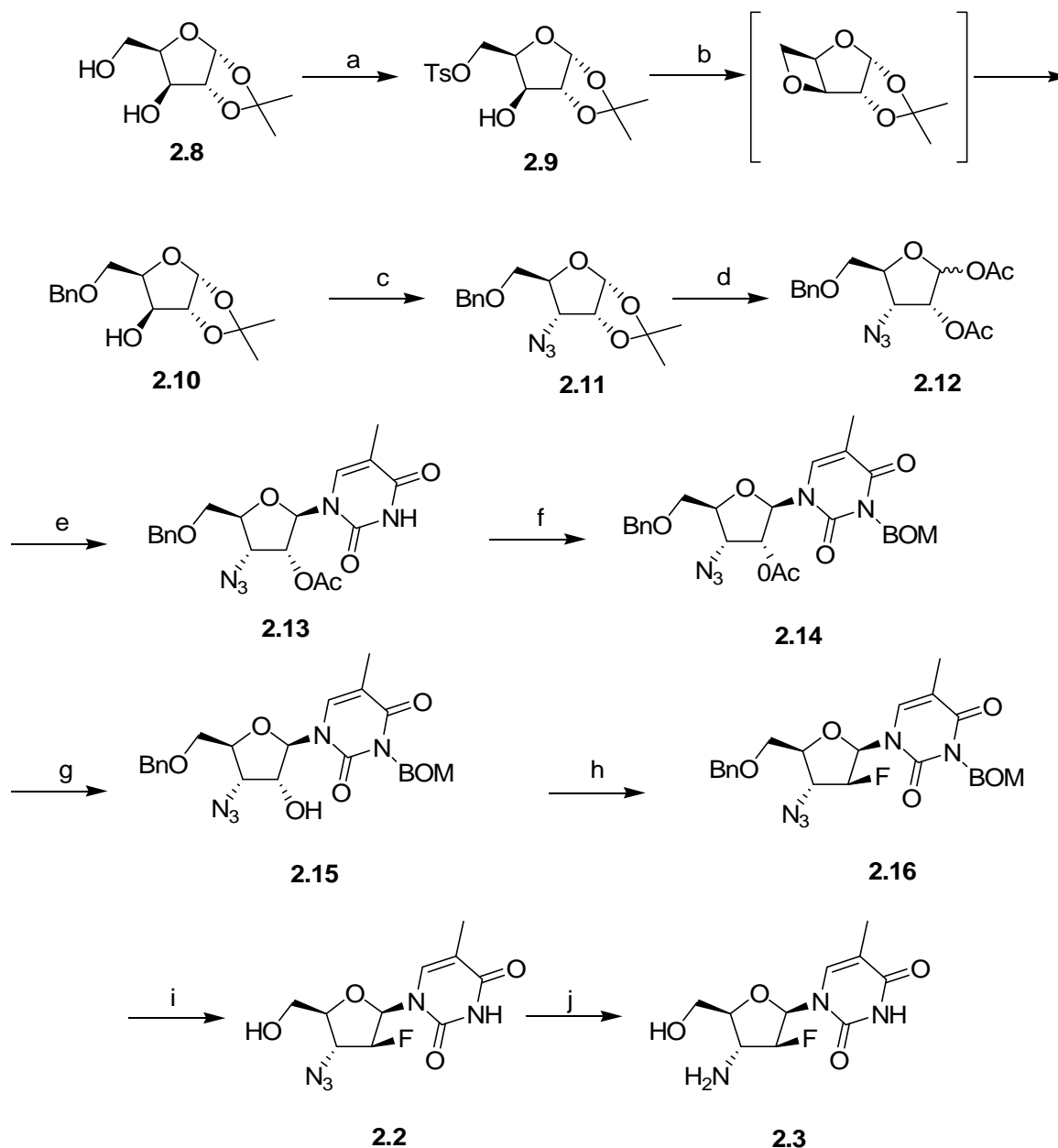
Scheme 2.2 , reagents and conditions: (a) TsCl, CH_2Cl_2, rt, 1h, 94%; (b) Na^0, BnOH, 100 °C, 16h, 78%; (c) (i) Tf_2O, CH_2Cl_2, pyridine, 0 °C, 1h, (ii) TMGA, DMF, rt, 16h, 59%; (d) (i) $HOAc/H_2O$ 3:1, 80 °C, 16h (ii) Ac_2O, pyridine rt, 3h, 95%; (e) silylated thymine, TMSOTf, $(CH_2)_2Cl_2$, 40 °C, 5h, 90%; (f) BOMCl, DBU, DMF, 0 °C, 40 min, 91%; (g) 1N NH_3 in methanol, rt, 3h, 98%; (h) DAST, toluene, pyridine, 50 °C, 3h, 61%; (i) 1M BCl_3, CH_2Cl_2 -78 °C, 15 min, 89%; (j) 10% Pd/C, MeOH, rt, 4h, 94%.

The most limiting step for an overall high yield is the introduction of the azidogroup on the 3-position of the ribose-sugar. Due to the antiperiplanar position of the triflate leaving group and the proton at the 4-position, elimination competes with S_N2-substitution (See Scheme 2.3), especially when the poorly soluble NaN_3 is used. In order to enlarge the amount of the desired **2.11,** the better soluble N,N,N',N'-tetramethylguanidinium azide was used. As a result of faster substitution, a lower amount of elimination product is obtained.
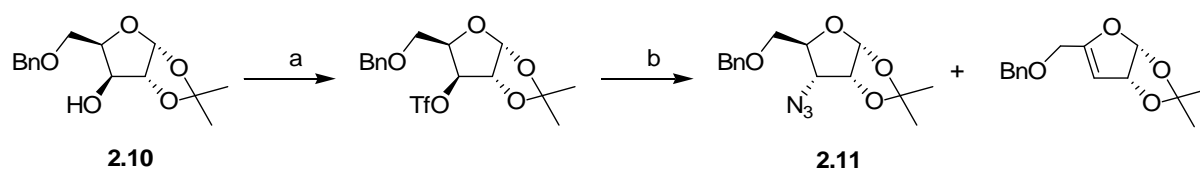
Scheme 2.3, reagents and condititons: (a) Tf_2O, CH_2Cl_2, pyridine, -78 °C, 1h; (b) TMGA, DMF, rt, 16h, 59% from **2.13**.

Deprotection of the 1,2-*O*-isopropylidene group in 75% acetic acid and subsequent protection of the alcohol functionalities resulted in the formation of the 3'-azido sugar building block, which after Vorbrüggen coupling, afforded 2'-*O*-acetyl-3'-azido-5'-*O*-benzyl-substituted thymidine in 90% yield.

Deacetylation was performed with 1N ammonia in methanol. The fluorine-atom is introduced by the fluorinating agent diethylaminosulfur trifluoride (DAST). During this reaction, DAST is introduced as a leaving group at the free alcohol function, which can then be substituted by the fluorine anion. In pyrimidine nucleosides however, a good leaving group at the 2'-ribo position is known to stimulate the intramolecular attack of the *O*-2 of the nucleobase, resulting in 2,2'-*O*-anhydrothymidine **2.17** (Scheme 2.4).
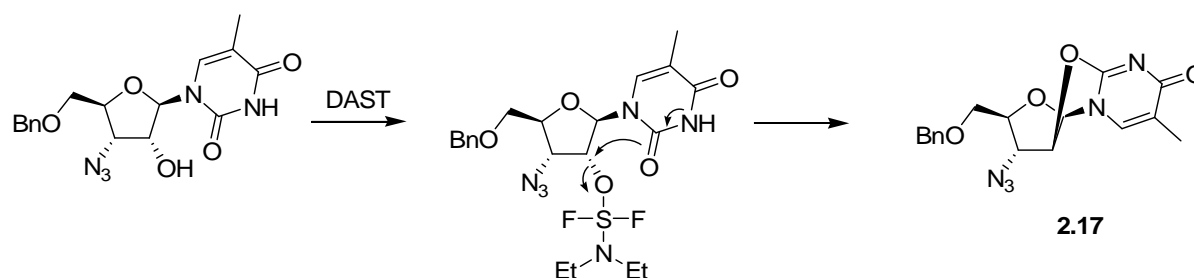
Scheme 2.4 Undesired 2,2'-anhydrothymidine formation
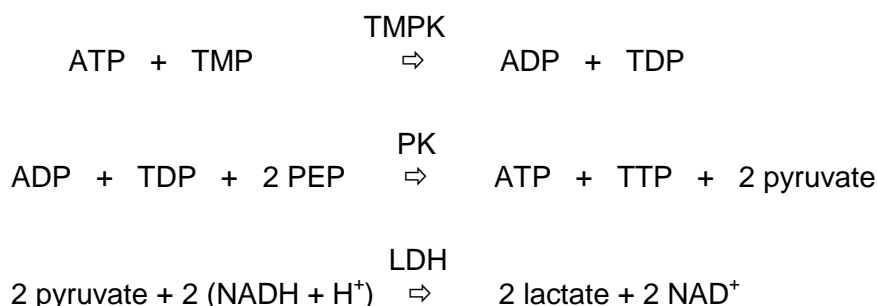
For this reason, the *N*-3 of thymine was protected with a benzyloxymethyl group, thereby preventing anhydro formation. Treatment of compound **2.15** with DAST resulted in the 2'-arabino F-substituted compound. After simultaneous deprotection of the benzyl and the benzyloxymethyl groups, the 3'-azido (**2.2**) and 3'-amino (**2.3**) substituted compounds were obtained.
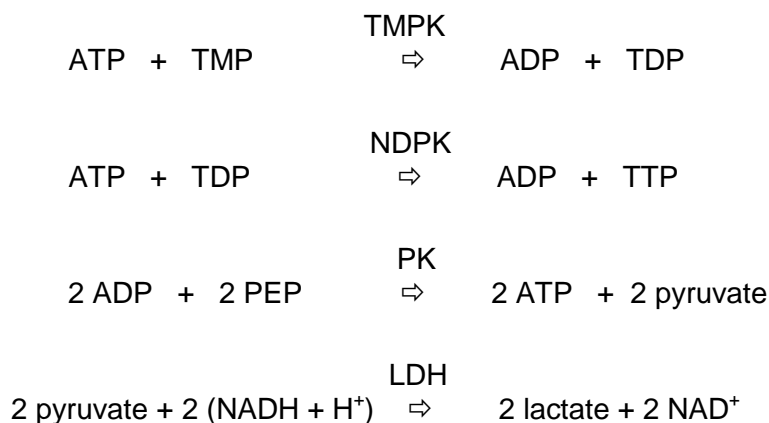
2.3 BIOLOGICAL EVALUATION

The three synthesized derivatives were tested for their affinity for TMPKmt via a reported spectrometric assay. [119]

In this spectrophotometric assay, the formation of thymidine diphosphate is coupled to reactions catalyzed by pyruvate kinase (PK) and lactate dehydrogenase (LDH) in the presence of phosphoenolpyruvate (PEP) and NADH, using NADH as indicator. Each mole of transferred phosphoryl group generates two moles NDP (ADP and TDP) and consequently two moles of NADH are oxidised to NAD^+.

$$\text{ATP} + \text{TMP} \quad \xrightarrow{\text{TMPK}} \quad \text{ADP} + \text{TDP}$$

$$\text{ADP} + \text{TDP} + 2\,\text{PEP} \quad \xrightarrow{\text{PK}} \quad \text{ATP} + \text{TTP} + 2\,\text{pyruvate}$$

$$2\,\text{pyruvate} + 2\,(\text{NADH} + \text{H}^+) \quad \xrightarrow{\text{LDH}} \quad 2\,\text{lactate} + 2\,\text{NAD}^+$$
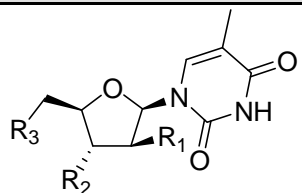
However, as the reactivity of various NDP's towards PK can vary a lot, the coupling system may become rate limiting, which leads to underestimation of the phosphorylation rate. Adding NDP-kinase, with broad substrate specificity circumvents this problem.

$$\text{ATP} + \text{TMP} \quad \xrightarrow{\text{TMPK}} \quad \text{ADP} + \text{TDP}$$

$$\text{ATP} + \text{TDP} \quad \xrightarrow{\text{NDPK}} \quad \text{ADP} + \text{TTP}$$

$$2\,\text{ADP} + 2\,\text{PEP} \quad \xrightarrow{\text{PK}} \quad 2\,\text{ATP} + 2\,\text{pyruvate}$$

$$2\,\text{pyruvate} + 2\,(\text{NADH} + \text{H}^+) \quad \xrightarrow{\text{LDH}} \quad 2\,\text{lactate} + 2\,\text{NAD}^+$$

From the spectrometrically measured decrease in NADH, the phosphorylation rate can be calculated. In case the synthesised compounds are no substrates, their inhibitory capacity is determined. At a fixed concentration of dTMP, the decrease in phosphorylation rate was measured, allowing the calculation of the K_i-value.

According to this biological assay, compounds **2.1**, **2.2** and **2.3** were tested for their inhibitory activity against TMPKmt. Their activities are shown in Table 2.2.

Table 2.2 Kinetic parameters of TMPKmt with compounds **2.1-2.3**

Compound	R_1	R_2	R_3	K_i (µM) TMPK mt
Dtmp				4.5 [a]
2.1	H	N_3	NH_2	152
2.2	F	N_3	OH	157
2.3	F	NH_2	OH	> 2000

[a] K_m-Value

Compound **2.1** can be considered as a hybrid structure of AZT (K_i = 28 µM) and 5'-NH_2-5'-deoxythymidine (**2.4**; K_i = 12 µM). The combination of these two beneficial modifications however leads to a compound with only weak inhibitory activity (K_i = 152 µM). One modification probably limits the good positioning of the other modification, thereby preventing this compound to optimally interact with the enzyme.

Compounds **2.2** and **2.3**, containing a 2'-F-substituent in the arabino configuration were designed to force the sugar into the South conformation. The activities of **2.2** and **2.3** appeared to be very weak. To explain these disappointing results, a conformational analysis was performed.

For each compound, three vicinal proton-proton and two vicinal proton-fluor coupling constants were measured at 500 MHz in DMSO at four different temperatures in a 298-328 K range (See Table 2.3).

Table 2.3 $^3J_{H,H}$ and $^3J_{H,F}$ coupling constants (Hz) of **2.2** and **2.3** at four different temperatures

	$^3J_{1'2'}$	$^3J_{2'3'}$	$^3J_{3'4'}$	$^3J_{1'F}$	$^3J_{3'F}$
2.2					
298	1.76	4.60	8.79	19.12	23.26
308	1.82	4.65	8.64	19.22	23.00
318	1.85	4.70	8.63	19.27	22.71
328	1.95	4.76	8.63	19.37	22.48
2.3					
298	1.60	5.49	8.78	21.24	23.61
308	1.73	5.44	8.80	21.18	23.11
318	1.73	5.55	8.74	21.06	22.91
328	1.78	5.61	8.73	21.06	22.81

$^3J_{H,H}$ recorded at 300 MHz in DMSO. X_S is the mole fraction in South conformation at 30 °C

Subsequently, in collaboration with the Lab for NMR and structural analysis of the UGent, the Matlab graphical user interface for 5-membered rings was used to calculate the best fit for the coupling constants at the different temperatures for **2.2** and **2.3**. For both compounds, a best fit was found for a North conformation (P ~ 320, ν_m~70) and an East conformation (P ~ 90; ν_m= 90).

For five-membered ring systems, this ν_m is normally between 30 and 60 degrees. A value of 90 degrees, obtained for the East conformation is physically impossible, which indicates that the used system is underdetermined. This can be explained by the fact that not enough coupling constants are available for a reliable result.

To answer this uncertainty, a systematic scan of the full pseudorotational cycle is performed (see Figure 2.3). This scan shows the regions of the conformations for which a good similarity between the experimental and calculated coupling constants can be found. This indicates that the Northern and Eastern part of the cycle enclose the regions that contain the most probable conformations of compounds **2.2** and **2.3**, which is in agreement with the calculations of the graphical user interphase. Most probably, the North and East conformation appear, but with different ν_m values as those calculated above. The calculated ratio North/East is 63/37 for **2.2** and 50/50 for compound **2.3**.
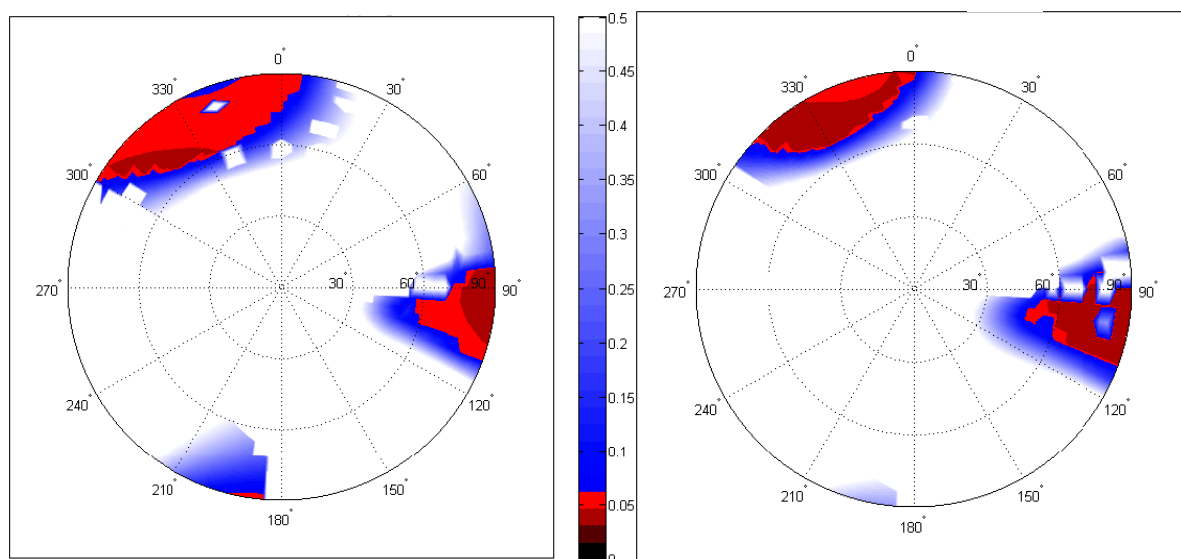
Figure 2.3 Pseudorotational mapping of compound **2.2** (left) and compound **2.3** (right); the red areas contain the conformations for which there is a very small difference between the expected and calculated NMR-data.

While a 2'-F-arabino-substitution on β-D-thymidine is known to force the sugar into a South conformation, our results show that the additional influence of an N-containing substituent on the 3'-position dominates the effect of the 2'-F-substitution, leading to compounds with a furanose in the less preferable North and East conformations, which might explain the low affinities of compounds **2.2** and **2.3** for TMPKmt.

2.4 CONCLUSION

This chapter describes the synthesis and biological activities of compounds **2.1**, **2.2** and **2.3**. Compound **2.1** combines two beneficial substituents at the 3'- and 5'- position in an attempt to optimise the TMPKmt inhibitory activity. This compound, however, showed only a moderate activity with a K_i of 152 μM, indicating that the two individually interesting modifications do not act synergistically.

Compounds **2.2** and **2.3**, which combine an amino or azido function at the 3'-position, with a 2'-F-arabino group, intend to force the sugar ring in the preferable South conformation. Conformational analysis, however, indicated that these analogues did not adopt the intended conformation in solution. Apparently, the presence of an additional N-containing substituent at the 3'-position favours the North or East conformation, which might explain the poor activities observed for **2.2** and **2.3**.

2.5 EXPERIMENTAL PART

Nucleoside nomenclature

Since IUPAC nomenclature does not allow rapid visualisation of modified nucleosides, the typical nucleoside notation starts from the name of the natural nucleoside, to which the chemical modifications at the different positions are added. This nomenclature is generally accepted in the nucleoside chemistry and is therefore generally used throughout this work. When an extra carbon is introduced at the 3'-position, it is described as C-6'.

As these conventions can lead to misinterpretations if changes on the nucleoside structure modify the parent nucleoside too much, sometimes the IUPAC nomenclature should be used. This is the case for the 2',3'-bicyclic thymidine derivatives, described in chapter 4 which differ too much from the natural nucleoside.

Synthesis

General

NMR spectra were obtained with a Varian Mercury 300 spectrometer. Chemical shifts are given in ppm (δ) relative to the residual solvent peak, in the case of DMSO-d_6 2.54 ppm for ^{1}H and 40.5 ppm for ^{13}C, in the case of CDCl$_3$ 7.26 ppm for ^{1}H and 77.4 ppm for ^{13}C. All signals assigned to hydroxyl groups were exchangeable with D$_2$O. Mass spectra and exact mass measurements were performed on a quadrupole/orthogonal-acceleration time-of-flight (Q/oaTOF) tandem mass spectrometer (qT of 2, Micromass, Manchester, U.K.) equipped with a standard electrospray ionization (ESI) interface. Samples were infused in a 2-propanol/water (1:1) mixture at 3µL/min. Precoated Merck silica gel F254 plates were used for TLC and spots were examined under UV light at 254 nm and revealed by sulfuric acid-anisaldehyde spray. Column chromatography was performed on ICN silica gel (63-200 µm, ICN, Asse Relegem, Belgium).

Spectrophotometric binding assay

The in vitro tests were conducted with recombinant enzymes overexpressed in *E. coli*: TMPKmt and TMPKh. TMPKmt and TMPKh activity was determined using the coupled spectrophotometric assay described by Blondin et al.[119] at 334 nm in an Eppendorf ECOM 6122 photometer. The reaction medium (0,5 ml final volume) contained 50 mM Tris-HCl pH 7.4, 50 mM KCl, 2 mM MgCl$_2$, 0.2 mM NADH, 1 mM phosphoenol pyruvate kinase and 2 units each of lactate dehydrogenase, pyruvate kinase and nucleoside diphosphate kinase. The concentrations of ATP and dTMP were kept constant at 0.5 mM and 0.05 mM respectively, whereas the concentrations of the analogues varied between 0.01 and 2 mM.

Pseudorotational analysis

Compounds **2.2** and **2.3** were dissolved in DMSO and high resolution 1D ^{1}H spectra were recorded on a Bruker Avance 500MHz spectrometer on the standard TXI-probe. Spectra were recorded at four temperatures between 298 and 328 K.

For the calculation of the pseudorotational parameters, the Karplus equation as described by Donders et al.[120] was used. The group electronegativity-values used in this equation are shown below.

1'- 2'	0.68	1.37	0.56	1.40
2'- 3'	1.24	0.68	1.37	0.68
3'- 4'	1.40	0.68	0.65	1.24

The pseudorotational analysis was performed using the Matlab graphical user interface for five-membered rings, at the Laboratory for NMR and structural analysis at UGent (Prof. Martins).

2',3'-*O*-Anhydro-5'-*N*-phtalimidyl-5'-deoxy-β-D-thymidine (2.7)

To a solution of thymidine (300 mg, 1.24 mmol), phtalimide (200 mg, 1.36 mmol) and triphenylphosphine (357 mg, 1.36 mmol) in THF (8 mL), was added diisopropylazodicarboxylate (536 μL, 2.72 mmol) at 0 °C. The reaction mixture was stirred at room temperature overnight. Water (20 mL) was added and the reaction mixture was extracted with ethylacetate (3 x 20 mL). The combined organic layers were dried over MgSO$_4$, evaporated and purified by column chromatography (CH$_2$Cl$_2$/MeOH 95:5) to yield compound **2.7** (264 mg, 60 %) as a slightly yellow solid.
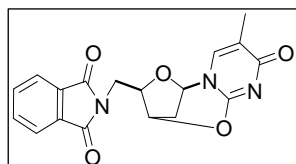
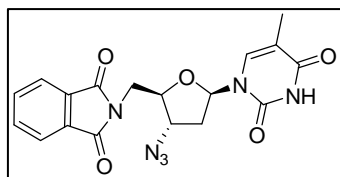**^{1}H NMR (300 MHz, DMSO-d$_6$):** δ 1.79 (3H, d, 5-CH$_3$), 2.48 (1H, m, H-2'), 2.58 (1H, d, *J*= 12.9 Hz, H-2"), 3.71 (1H, dd, *J*= 8.7 and 14.4 Hz, H-5'), 3.87 (1H, dd, J= 4.8 and 14.4 Hz, H-5"), 4.51 (1H, m, H-4'), 5.36 (1H, s, H-3'), 5.77 (1H, d, *J*= 3.6 Hz, H-1'), 7.55 (1H, s, H-6), 7.84 (4H, m, arom-H).

HRMS (ESI-MS) for C$_{18}$H$_{16}$N$_3$O$_5$ [M+H]$^+$ found, 354.1096; calcd, 354.1089.

3'-Azido-3',5'-dideoxy-5'-*N*-phtalimidyl-β-D-thymidine (2.8)

A mixture of **2.7** (160 mg, 0.45 mmol) and NaN_3 (146 mg, 2.25 mmol) was dissolved in DMF (3 mL) and heated to 150 °C. After the addition of benzoic acid (55 mg, 0.45 mmol), the reaction was refluxed for 2 hours. The reaction was quenched with water (10 mL) and the reaction mixture was extracted with ethylacetate (3 x 10 mL). The combined organic layers were dried over $MgSO_4$ and evaporated to dryness. The residue was purified by column chromatography (CH_2Cl_2/MeOH 95:5), affording **2.8** (124 mg, 70%).
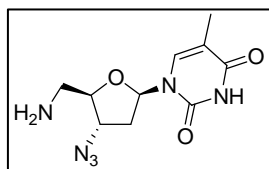

^{1}H NMR (300 MHz, DMSO-d_6): δ 1.64 (3H, d, 5-CH_3), 2.29 (1H, m, H-2'), 2.48 (1H, m, H-2"), 3.89 (2H, m, H-5' and H-5"), 3.98 (1H, m, H-4'), 4.43 (1H, m, H-3'), 6.04 (1H, t, *J*= 6.3 Hz, H-1'), 7.44 (1H, s, H-6), 7.84 (4H, m, arom-H).

HRMS (ESI-MS) for $C_{18}H_{16}N_6O_5Na$ [M+Na]$^+$ found, 419.1084; calcd, 419.1080.

5'-Amino-3'-azido-3',5'-dideoxy-β-D-thymidine (2.1)

Compound **2.8** (100 mg, 0.25 mmol) was dissolved in ethanol (3 mL) and hydrazine monohydrate (49 μl, 1 mmol) was added. The reaction mixture was stirred at 28 °C during 2 days, evaporated to dryness and purified by column chromatography (CH_2Cl_2/MeOH 90:10) to afford final compound **2.1** (48 mg, 73%) as a white solid.

^{1}H NMR (300 MHz, DMSO-d_6): δ 1.77 (3H, d, 5-CH_3), 2.23 (1H, m, H-2'), 2.37 (1H, m, H-2"), 2.75 (2H, d, *J*= 5.1 Hz, H-5' and H-5"), 5.69 (1H, dd, J= 5.1 and 9.9 Hz, H-3'), 4.39 (1H, m, H-4'), 6.05 (1H, t, *J*= 6.9 Hz, H-1'), 7.64 (1H, s, H-6).

**HRMS (ESI-MS)** for $C_{10}H_{15}N_6O_3$ [M+H]$^+$ found, 267.1198; calcd, 267.1205.

^{13}C NMR (75 MHz, DMSO-d_6): δ 12.57 (5-CH_3), under DMSO (C-2'), 41.39 (C-5'), 61.37 (C-3'), 85.61 (C-1'), 86.84 (C-4'), 110.79 (C-5), 138.56 (C-6), 151.11 (C-4), 164.82 (C-2).

1,2-*O*-Isopropyl-5-tosyl-α-D-xylofuranose (**2.9**)

1,2-*O*-isopropyl-α-D-xylofuranose (10 g, 53 mmol) was dissolved in pyridine (53 mL) and cooled to 0 °C. A solution of *para*-toluenesulfonyl chloride (10.11 g, 53 mmol) in CH_2Cl_2 (25 mL) was added. After warming up to room temperature, the reaction was allowed to stir for 1 h. H_2O (0.2 mL) was added and the obtained mixture was stirred for 30 minutes. After the reaction mixture was evaporated *in vacuo*, the residue was partitioned between H_2O

(100 mL) and CH_2Cl_2 (100 mL), and the aqueous layer was extracted with CH_2Cl_2 (4 x 100 mL). The combined organic layers were washed with water and brine, dried over $MgSO_4$ and concentrated *in vacuo*. The resulting solid was washed with ether and filtered to afford **2.9** (17.22 g, 94%) as a white powder.
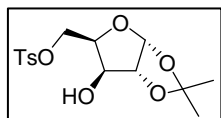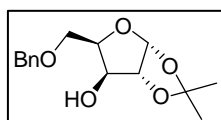

^{1}H NMR (300 MHz, DMSO-d_6): δ 1.21 (3H, s, CH_3), 1.33 (3H, s, CH_3), 2.42 (3H, s, CH_3Ph), 4.02 (2H, m, H-5 and H-5'), 4.13 (1H, m, H-3), 4.24 (1H , dd, *J*= 3.2 and 10.5 Hz, H-4), 4.37 (1H, m, H-2), 5.41 (1H, d, *J*= 4.5 Hz, 3-OH), 5.81 (1H, d, *J*= 4.5 Hz, H-1), 7.49 (2H, d, *J*= 8.1 Hz, 2 ar H), 7.79 (2H, d, *J*= 8.1 Hz, 2 ar H);

Exact mass (ESI-MS) for $C_{15}H_{21}O_7S$ [M+H]$^+$ found, 345.1011; calcd, 345.1008.

5-*O*-Benzyl-1,2-*O*-isopropyl-α-D-xylofuranose (2.10)

Sodium metal (6.9 g, 300 mmol) was dissolved in benzyl alcohol (190 mL) at 100 °C and solid **2.9** (17.22 g, 50 mmol) was added. The reaction mixture was heated at 100 °C overnight, cooled to room temperature and treated with H_2O and glacial acetic acid till pH= 7. The mixture was partitioned between ether (150 mL) and H_2O (100 mL). The organic layer was washed twice with H_2O (100 mL), dried over $MgSO_4$, and evaporated. The oily residue was purified by column chromatography (pentane/ethylacetate 85:15) to yield **2.10** (11 g, 78%) as a white crystalline powder.

^{1}H NMR (300 MHz, DMSO-d_6): δ 1.23 (3H, s, CH_3), 1.37 (3H, s, CH_3), 3.52 (1H, dd, *J*= 7.1 and 10.5 Hz, H-5), 3.68 (1H, dd, *J*= 3.4 and 10.5 Hz, H-5'), 3.98 (1H, m, H-3), 4.14 (1H, m, H-4), 4.38 (1H, m, H-2), 4.50 (2H, s, CH_2Ph), 5.22 (1H, d, *J*= 4.5 Hz, 3-OH), 5.82 (1H, d, *J*= 4.1 Hz, H-1), 7.33 (5H, m, CH_2*Ph*).

Exact mass (ESI-MS) for $C_{15}H_{21}O_5$ [M+H]$^+$ found, 281.1388; calcd, 281.1389.

3-Azido-5-*O*-benzyl-1,2-*O*-isopropyl-α-D-ribofuranose (2.11)

Compound **2.10** (6g, 20 mmol) was dissolved in CH_2Cl_2 (70 mL) and pyridine (5.5 mL). At -78 °C, triflic anhydride (4 mL, 24 mmol) was added slowly. The reaction was allowed to warm up to 0 °C. After 1 hour, the reaction was quenched with saturated $NaHCO_3$-solution (75 mL) and extracted with CH_2Cl_2 (3 x 100 mL). The combined organic layers were dried over $MgSO_4$, evaporated to dryness and used immediately in the next reaction.

The crude compound was dissolved in DMF (50 mL) and N,N,N',N'-tetramethylguanidium azide was added (6.32 g, 40 mmol). The reaction was stirred overnight, evaporated to

dryness and purified by column chromatography (hexane/ethylacetate 97:3) to obtain 3.79 g (59%) of the pure compound **2.11** as a white solid.
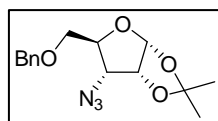
¹H NMR (300 MHz, DMSO-d₆): δ 1.27 (3H, s, CH₃), 1.43 (3H, s, CH₃), 3.57 (2H, m, H-5a and H-3), 3.67 (1H, dd, *J*= 2.6 and 11.1 Hz, H-5b), 4.06 (1H, m, H-4), 4.51 (2H, s, *CH₂*Ph), 4.77 (1H, app t, *J*= 4.1 Hz, H-2), 5.78 (1H, d, *J*= 3.9 Hz, H-1), 7.32 (5H, m, CH₂*Ph*).

HRMS (ESI-MS) for $C_{15}H_{20}N_3O_4$ [M+H]⁺ found, 306.1449; calcd, 306.1454.

Elimination procuct:
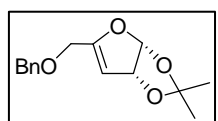
¹H NMR (300 MHz, CDCl₃): δ 1.47 (3H, s, CH₃), 1.50 (3H, s, CH₃), 4.83 (2H, s, H-5a and H-5b), 5.31 (2H, m, H-2 and H-3), 4.51 (2H, s, *CH₂*Ph), 6.13 (1H, d, *J*= 4.1 Hz, H-1), 7.37 (5H, m, CH₂*Ph*).

HRMS (ESI-MS) for $C_{15}H_{18}O_4$ [M+H]⁺ found, 263.1289; calcd, 263.1283.

1,2-Di-*O*-acetyl-3-azido-5-*O*-benzyl-α-D-ribofuranose (2.12)

A solution of **2.11** (3.48 g, 10.83 g) in acetic acid/water (3:1, 120 mL) was heated overnight at 80 °C. The reaction mixture was evaporated to dryness and coevaporated with toluene (2x 15 mL). The obtained compound was dissolved in pyridine (40 mL) and acetic anhydride (13 mL) was added. This reaction was stirred for 3 hours under nitrogen atmosphere. After evaporation, no further purification was needed to obtain the pure compound **2.12** (3.76 g, 95%) as a white solid.
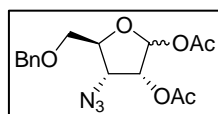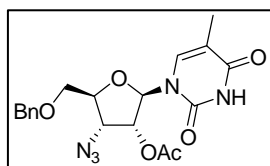
¹H NMR (300 MHz, DMSO-d₆): δ 1.94 (3H, s, COCH₃), 2.10 (3H, s, COCH₃), 3.59 (2H, m, H-5a and H-5b), 4.09 (1H, m, H-4), 4.42 (1H, dd, J= 4.8 and 7.8 Hz, H-3), 4.54 (2H, s, *CH₂*Ph), 5.31 (1H, d, *J*= 4.5 Hz, H-2), 5.97 (1H, s, H-1), 7.33 (5H, m, CH₂*Ph*).

HRMS (ESI-MS) for $C_{16}H_{20}N_3O_6$ [M+H]⁺ found, 350.1354; calcd, 350.1352.

2'-*O*-Acetyl-3'-azido-5'-*O*-benzyl-β-D-thymidine (2.13)

Chlorotrimethylsilane (1.04 mL, 8.2 mmol) was added to a suspension of thymine (1.43 g, 11.32 mmol) in HMDS (120 mL) and pyridine (10 mL). The suspension was heated overnight at 130 °C. The clear solution was cooled to room temperature, evaporated to dryness and coevaporated with toluene (2x 10 mL). The silylated base was then dissolved

in dry dichloroethane (40 mL). Compound **2.12** was dissolved in 45 mL dichloroethane and added to the silylated base, followed by the dropwise addition of TMSOTf. The clear solution was stirred at 40 °C for 5 hours. After cooling down to room temperature, the reaction mixture was extracted with a saturated NaHCO$_3$-solution (50 mL). The water layer was extracted three times with CH$_2$Cl$_2$ (50 mL). The organic layers were dried over MgSO$_4$, evaporated and purified by column chromatography (hexane/ethylacetate 65:35) yielding 3.85 g (90%) pure nucleoside **2.13** as a white foam.


^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.79 (3H, s, 5-CH$_3$), 2.09 (3H, s, COCH$_3$), 3.67 (1H, dd, *J*= 3.3 and 11.1 Hz, H-5'), 3.78 (1H, dd, *J*= 2.5 and 11.1 Hz, H-5"), 4.06 (1H, m, H-4'), 4.59 (2H, s, *CH$_2$*Ph), 4.62 (1H, app t, *J*= 5.5 Hz, H-3'), 5.47 (1H, app t, *J*= 5.5 Hz, H-2'), 5.91 (1H, d, *J*= 4.8 Hz, H-1'), 7.32 (5H, m, CH$_2$*Ph*), 7.53 (1H, s, H-6), 11.39 (1H, s, N(3)H).

HRMS (ESI-MS) for C$_{19}$H$_{22}$N$_5$O$_6$ [M+H]$^+$ found, 416.1572; calcd, 416.1569.

2'-*O*-Acetyl-3'-azido-5'-*O*-benzyl-3-*N*-benzyloxymethyl-β-D-thymidine (2.14)

Compound **2.13** (830 mg, 2 mmol) was dissolved in anhydrous DMF (20 mL). DBU (301 μL, 2 mmol) was added, followed by the addition of benzyloxymethylchloride (334 μl, 2.4 mmol) at 0 °C. After stirring at this temperature for 40 minutes, the reaction was quenched with saturated NaHCO$_3$-solution (25 mL) and extracted three times with CH$_2$Cl$_2$ (30 mL). After drying over MgSO$_4$ and evaporation, the crude material was purified by column chromatography (hexane/ethylacetate 85:15) yielding 974 mg **2.14** (91%) as a white foam.
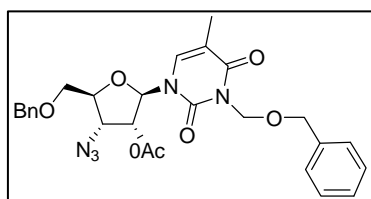

^{1}H NMR (300 MHz, CDCl$_3$): δ 1.55 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 2.11 (3H, s, COCH$_3$), 3.66 (1H, dd, *J*= 2.2 and 11.0 Hz, H-5'), 3.89 (1H, dd, *J*= 2.2 and 11.0 Hz, H-5"), 4.11 (1H, m, H-4'), 4.35 (1H, app t, *J*= 5.7 Hz, H-3'), 4.68 and 4.64 (4H, 2 x s, 2 x *CH$_2$*Ph), 5.34 (2H, s, NC*H$_2$*O), 5.46 (1H, d, *J*= 0.9 Hz, H-2'), 6.09 (1H, d, *J*= 4.2 Hz, H-1'), 7.33 (10H, m, 2 x CH$_2$*Ph*), 7.44 (1H, d, *J*= 1.2 Hz, H-6).

HRMS (ESI-MS) for C$_{27}$H$_{30}$N$_5$O$_7$ [M+H]$^+$ found, 536.2142; calcd, 536.2145.

3'-Azido-5'-*O*-benzyl-3-*N*-benzyloxymethyl-β-D-thymidine (2.15)

Compound **2.14** (970 mg, 1.81 mmol) was dissolved in 1N NH_3 in methanol (20 mL). The reaction was stirred at room temperature during 3 hours, evaporated to dryness and purified by column chromatography (hexane/ethylacetate 80:20) to yield 875 mg **2.15** (98%) as a slightly yellow foam.
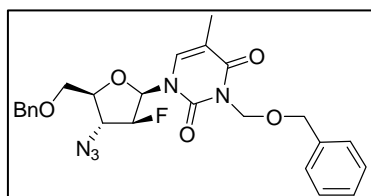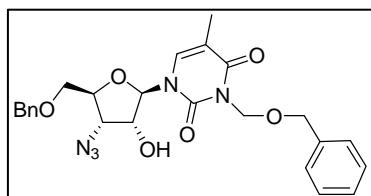
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.51 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 3.66 (1H, dd, *J*= 3.1 and 10.9 Hz, H-5'), 3.79 (1H, dd, *J*= 2.7 and 10.8 Hz, H-5"), 4.08 (1H, m, H-4'), 4.18 (1H, app t, *J*= 5.3 Hz, H-3'), 4.44 (1H, m, H-2'), 4.58 and 4.55 (4H, 2 x s, 2 x *CH$_2$*Ph), 5.30 (2H, d, *J*= 1.8 Hz, NC*H$_2$*O), 5.81 (1H, d, *J*= 5.2 Hz, 2'-OH), 6.21 (1H, d, *J*= 5.1 Hz, H-1'), 7.33 (10H, m, 2 x CH$_2$*Ph*), 7.56 (1H, d, *J*= 1.1 Hz, H-6).

HRMS (ESI-MS) for $C_{25}H_{28}N_5O_6$ [M+H]$^+$ found, 494.2032; calcd, 494.2039.

1-(3-Azido-5-*O*-benzyl-2-deoxy-2-fluoro-β-D-arabinofuranosyl)3-*N*-benzyloxymethyl thymine (2.16)

A solution of **2.15** (200 mg, 0.41 mmol) in toluene (4 mL) and pyridine (0.4 mL) was treated with DAST (201 μl, 1.52 mmol). The reaction was stirred at room temperature during 2 hours before heating it at 50 °C for 3 additional hours. Ethylacetate (10 mL) was added to the reaction and the organic layer was washed with a saturated $NaHCO_3$-solution (15 mL) and water (10 mL). After drying over $MgSO_4$, the combined organic layers were evaporated to dryness and purified by column chromatography (hexane/ethylacetate 80:20) to yield 126 mg pure **2.16** (61%) as a white foam.
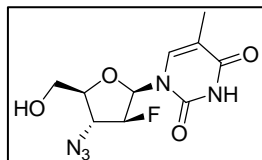
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.50 (3H, d, *J*= 1.1 Hz, 5-CH$_3$), 3.68 (1H, dd, *J*= 3.3 and 11.7 Hz, H-5'), 3.88 (1H, dd, *J*= 2.1 and 11.7 Hz, H-5"), 4.18 (1H, m, H-4'), 4.32-4.45 (1H, ddd, J= 4.5, 9.0 and 24.0 Hz, H-3'), 4.54 and 4.58 (4H, s and d, *J*= 2.4 Hz, 2 x *CH$_2$*Ph), 5.28 (2H, s, NC*H$_2$*O), 5.29-5.48 (1H, dd, *J*= 4.5 and 52.2 Hz, H-2'), 5.92 (1H, d, *J*= 17.7 Hz, H-1'), 7.29 (10H, m, 2 x CH$_2$*Ph*), 7.57 (1H, d, *J*= 1.1 Hz, H-6).

HRMS (ESI-MS) for $C_{25}H_{27}N_5O_5F$ [M+H]$^+$ found, 496.1991; calcd, 496.1996.

1-(3-Azido-2-deoxy-2-fluoro-β-D-arabinofuranosyl)thymine (2.2)

Compound **2.16** (126 mg, 0.26 mmol) was dissolved in anhydrous CH_2Cl_2 (10 mL) and at -78 °C a 1M BCl_3 in CH_2Cl_2 (2.07 mL) was added dropwisely. After 15 minutes, the reaction was quenched with methanol (10 mL). The resulting mixture was evaporated to dryness and purified by column chromatography (CH_2Cl_2/MeOH 97:3 to 95:5) yielding 66 mg pure compound **2.2** (89%) as a white foam.

¹H NMR (300 MHz, DMSO-d₆): δ 1.72 (3H, d, *J*= 1.0 Hz, 5-CH_3), 3.58 (1H, d, *J*= 12.3 Hz, H-5'), 3.78 (1H, d, *J*= 12.3 Hz, H-5"), 3.97 (1H, m, H-4'), 4.22-4.35 (1H, ddd, *J*= 4.5, 8.7 and 23.4 Hz, H-3'), 5.30-5.50 (1H, ddd, *J*= 1.5, 4.5 and 53.1 Hz, H-2'), 5.42 (1H, br s, (5'-OH), 5.87 (1H, dd, *J*= 1.5 and 17.7 Hz, H-1'), 7.57 (1H, d, *J*= 1.1 Hz, H-6), 11.45 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.76 (5-CH_3), 58.67 (J= 15.6 Hz, C-3'), 59.76 (C-5'), 81.58 (C-4'), 88.38 (*J*= 34.6 Hz, C-2'), 94.68 (*J*= 186.4 Hz, C-1'), 110.10 (C-5), 136.89 (C-6), 150.84 (C-4), 164.52 (C-2).

¹⁹F NMR (300 MHz, DMSO-d₆): δ -197.65.

HRMS (ESI-MS) for $C_{10}H_{13}N_5O_4F$ [M+H]⁺ found, 286.0955; calcd, 286.0952.

1-(3-Amino-2-deoxy-2-fluoro-β-D-arabinofuranosyl)thymine (2.3)

Compound **2.2** (47 mg, 0.16 mmol) was dissolved in methanol (5 mL) and after the addition of 10% Pd/C (5 mg), the reaction was stirred under H_2-atmosphere for 4 hours. After filtration over a celite path, the filtrate was evaporated to dryness and the crude residue was purified by column chromatography (CH_2Cl_2/MeOH 93:7) to yield 40 mg **2.3** (94%) as a white foam.
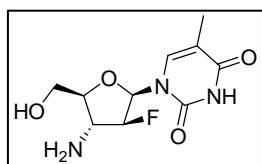
¹H NMR (300 MHz, DMSO-d₆): δ 1.73 (3H, s, 5-CH_3), 3.61 (1H, m, H-5'), 3.78 (1H, m, H-5"), 4.06 (1H, m, H-4'), 4.35 (1H, app t, *J*= 4.8 H-3'), 5.30-5.50 (1H, dd, *J*= 4.6 and 52.4 Hz, H-2'), 5.31 (1H, br s, (5'-OH), 5.80 (1H, d, *J*= 21.0 Hz, H-1'), 7.70 (1H, s, H-6), 8.16 (2H, br s, 3'-NH_2), 11.44 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.82 (5-CH_3), 51.63 (J= 18.2 Hz, C-3'), 59.59 (C-5'), 84.88 (C-4'), 88.36 (*J*= 35.7 Hz, C-2'), 96.88 (*J*= 180.5Hz, C-1'), 109.57 (C-5), 136.82 (C-6), 150.82 (C-4), 164.51 (C-2).

HRMS (ESI-MS) for $C_{10}H_{15}N_3O_4F$ [M+H]⁺ found, 260.1048; calcd, 260.1046.

Chapter 3

Synthesis of 3'-C-branched chain thymidine intermediates

3 SYNTHESIS OF 3'-*C*-BRANCHED CHAIN THYMIDINE INTERMEDIATES

Many of the envisaged analogues in this work are structurally derived from a couple of hits, which are characterized by a 3'-CH$_2$-X moiety (cf. supra). As a consequence, **3'-*C*-branched thymidine and ribo-thymidine** constitute precious intermediates for the synthesis of these analogues. As illustrated in Figure 3.1, 3'-aminomethyl ribothymidine and its 2'-deoxy analogue constitute important building blocks throughout this work. Therefore, this chapter is devoted to possible strategies for the stereoselective introduction of a 3'-*C*-branching on pyrimidine nucleosides.
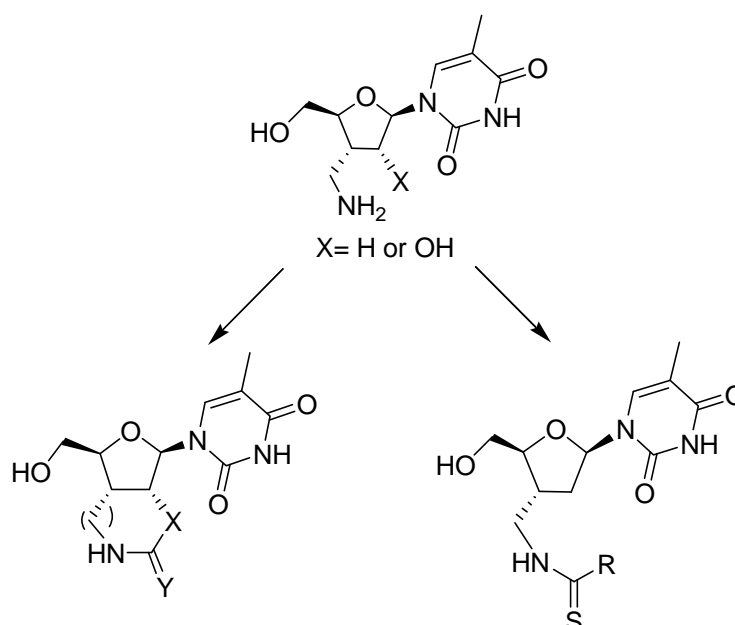

Figure 3.1 Application of 3'-*C*-branched thymidine derivatives in this work

3.1 NUCLEOSIDE SYNTHESIS

The intense search for clinically useful nucleoside derivatives has resulted in a wealth of approaches for their synthesis. In general, two main strategies for the synthesis of nucleoside analogues are followed. In a first approach, the **natural and readily available nucleoside** is used to introduce the intended modifications. Unfortunately, the reaction conditions required to introduce the desired modifications are not always compatible with the nucleoside. Sometimes one does not profit from starting from the natural nucleoside. In

these cases a second approach is followed, generally consisting of the **glycosylation of a suitable sugar synthon** with the desired nucleobase.

A critical feature of the glycosylation step is stereoselectivity. Generally, attack of the nucleobase from the β-side of the sugar, resulting in the formation of a β-nucleoside, is desirable. A method developed by Vorbrüggen, uses silylated nucleobases and strong Lewis acids and provides reproducibly high yields of ribofuranose derivatives with reliable and predictable stereochemical outcome.[121] As depicted in Scheme 3.1, the mechanism of the Vorbrüggen coupling method involves the conversion of a 1,2-diacetylated ribofuranosyl sugar **3.1** into a stable 1,2-acetyloxonium salt **3.3** by both assistance of the 2'-acetate group and a Lewis acid catalyst, trimethylsilyl triflate (**3.2**). The resulting triflate anion reacts with the persilylated nucleobase to regenerate trimethylsilyl triflate, along with the simultaneous attack of the thymine from the β-face of the electrophilic 'ribofuranose cation' to afford exclusively the β-nucleoside **3.4**.[122]
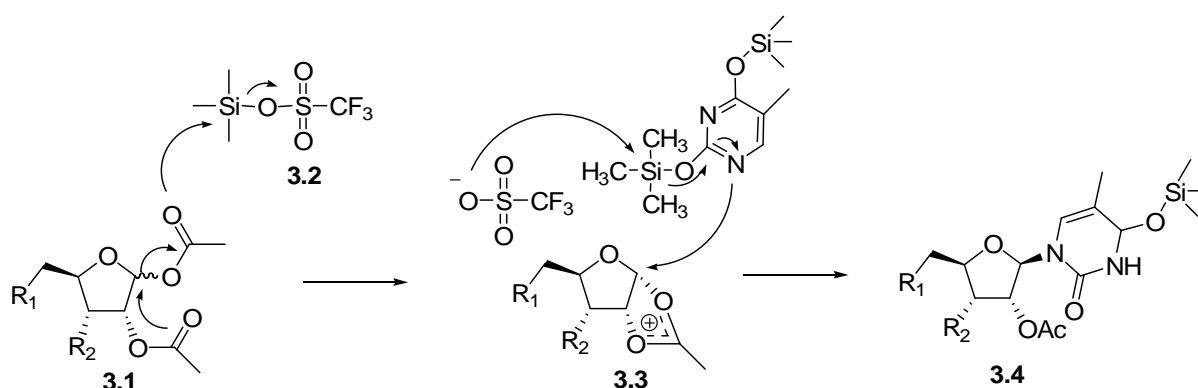
Scheme 3.1 The Vorbrüggen coupling reaction for the formation of β-ribonucleosides

As the esterified 2'-OH-group contributes to the stereoselectivity, Vorbrüggen's method can only be used for the synthesis of ribonucleosides.

For the **preparation of 2'-deoxynucleosides**, where the driving force in the stereoselectivity of the glycosylsation is lacking, a different strategy is needed. If attainable, starting from an available 2'-deoxynucleoside is the method of choice. An alternative strategy involves the glycosylation of 2'-deoxyribose derivatives. A major drawback of this approach is that it often produces hardly separable mixtures of α- and β-nucleosides.

With the synthesis of ribo- as well as 2'-deoxy-3'-branched-thymidine analogues in view, strategies towards either type will be described in the following part.
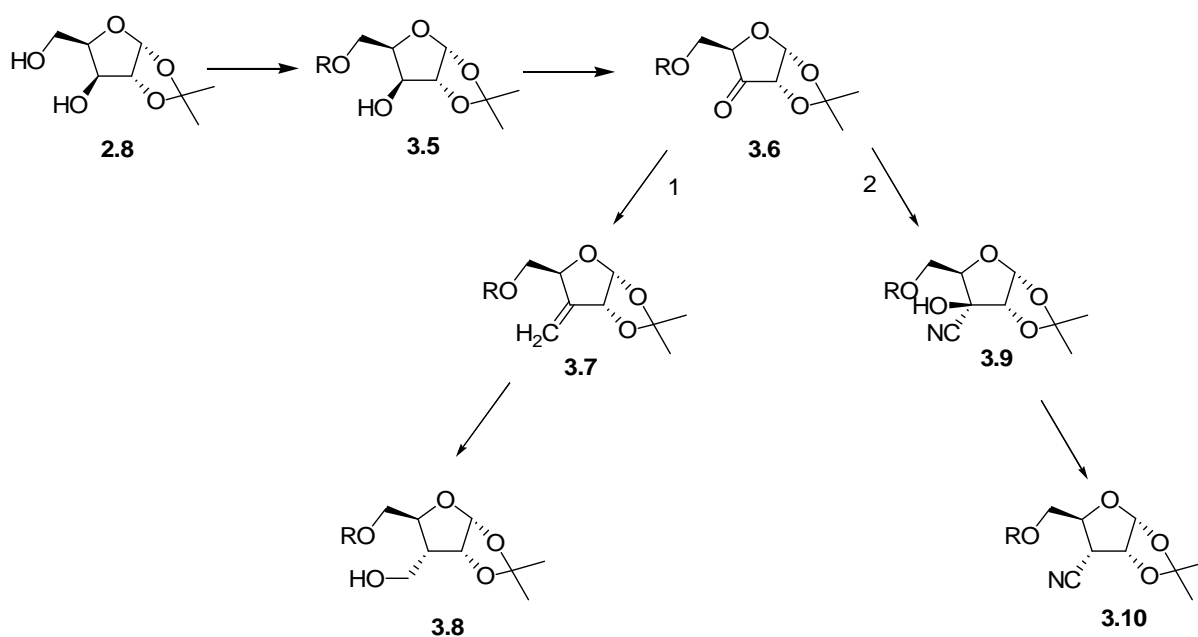
3.2 SYNTHESIS OF 3'-*C*-BRANCHED RIBONUCLEOSIDES

3.2.1 Synthesis of 3-*C*-branched ribofuranose

Different methods for the formation of a 3-α-oriented carbon-containing substituent on ribofuranose have been reported.[123,124,125] We will discuss the two most interesting procedures. Both methods start from the commercially available or readily available 1,2-*O*-isopropylidene-α-D-xylofuranose (**2.8**), which needs to be protected at the 5-position (**3.5**) (Scheme 3.2) prior to oxidation (Swern or Cr(VI)O$_3$) of the 3-alcohol function to afford the pentofuranos-3-ulose **3.6**. In a first method the 3-keto group of **3.6** is converted into an exocyclic 3-methylene moiety (**3.7**) using a Wittig reaction. Due to the sterical hindrance of the 1,2-*O*-isopropylidine protecting group, hydroboration of this methylene intermediate **3.7**, selectively generates a 3-α-hydroxymethylated derivative **3.8** in an overall yield of 55%, starting from **2.8**.[123]

A second method[124] relies on the stereoselective addition of KCN to the 3-keto function, leading to the xylo cyanoisomer **3.9**. Barton deoxygenation proved effective to obtain the 3-α-cyano-sugar **3.10**. Next to the much lower yield described for this reaction pathway (22%), Filichev et al. also reported the presence of the 3-β-epimer, leading to hardly separable mixtures.[124]

Consequently we chose to synthesize the sugar building block of 3'-azidomethyl-3'-deoxythymidine according to the first method.


Scheme 3.2 Synthesis of 3'-branched ribofuranose: overview of two discussed methods

3.3 SYNTHESIS OF 3'-*C*-BRANCHED 2'-DEOXYNUCLEOSIDES

3.3.1 Earlier observations

For 2'-deoxynucleosides, which are lacking the 2'-OH, the stereoselective Vorbrüggen coupling is excluded. Previous attempts in our laboratory to convert a 3'-azidomethyl-ribonucleoside into a 2'-deoxynucleoside proved problematic. During attempted **Barton deoxygenation** of a 2'-phenylthionocarbonate substituted 3'-azidomethyl-thymidine, no trace of the desired 2'-deoxygenated compound could be detected. Instead, **rearrangements** gave rise to **three unexpected nucleoside analogues**, being two bicyclic nucleosides and one dinucleoside (See Scheme 1.1).[107] This eliminates the use of Barton deoxygenation on 3'-azidomethyl substituted nucleosides.

Vanheusden et al.[106] described a divergent synthesis which utilizes a *p*-anisyl group to protect the 3-hydroxymethyl group of compound **3.8**. Coupling of this suitably protected sugar, afforded the ribo-thymidine derivative **3.11**, for which uncomplicated 2'-deoxygenation proved feasible (Scheme 3.3).
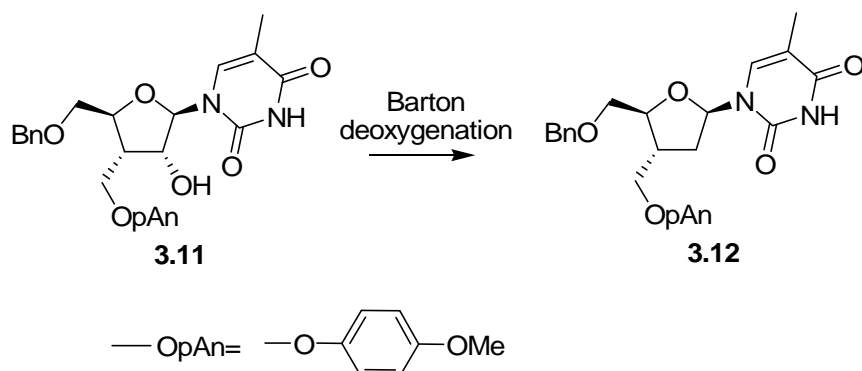
Scheme 3.3 Barton deoxygenation on suitably protected 3'-*C*-branched ribothymidine

An interesting feature of this reaction sequence was that it gave access to both thymidine and ribo-thymidine derivatives. However this synthetic route proved very laborious. Due to the introduction and removal of different protecting groups, 12 steps were required to obtain intermediate **3.12**, which, we felt, is a major drawback of this approach.

For this reason, Vorbrüggen coupling followed by 2'-deoxygenation was not considered as a convenient method to obtain 3'-*C*-branched thymidine analogues.

In search for an alternative synthesis, we carried out a literature survey, from which the most interesting methods are described below.

3.3.2 <u>Reported methods</u>

3.3.2.1 De novo sugar synthesis

One of the oldest methods reported is the synthesis of Samuelsson starting from 1,4-butenediol (Scheme 3.4).[126] This method is not starting from a naturally available nucleoside and consequently it is not a method of first choice. As it is a generally known old method and it illustrates the **major drawback of non-stereoselective glycosylation** of a 2'-deoxynucleoside, it is described in this part.

The furanose part was constructed using *de novo* sugar synthesis involving Sharpless epoxidation (**3.13**),[127] regioselective alkylation of a 2,3-epoxy alcohol using allylmagnesium bromide (**3.14**), followed by oxidative cleavage of the double bond (**3.16**). The resulting protected furanose derivative was condensed with silylated base, which after deprotection and separation of α- and β-isomers, provided the nucleoside analogues (**3.17**). Although this route was explored for the synthesis of 3'-*C*-branched nucleosides, it is obvious that the poor stereoselectivity observed in the late-stage glycosylation step (1.4:1 α:β) prevents efficient preparation of 3'-*C*-branched β-nucleosides.[128]
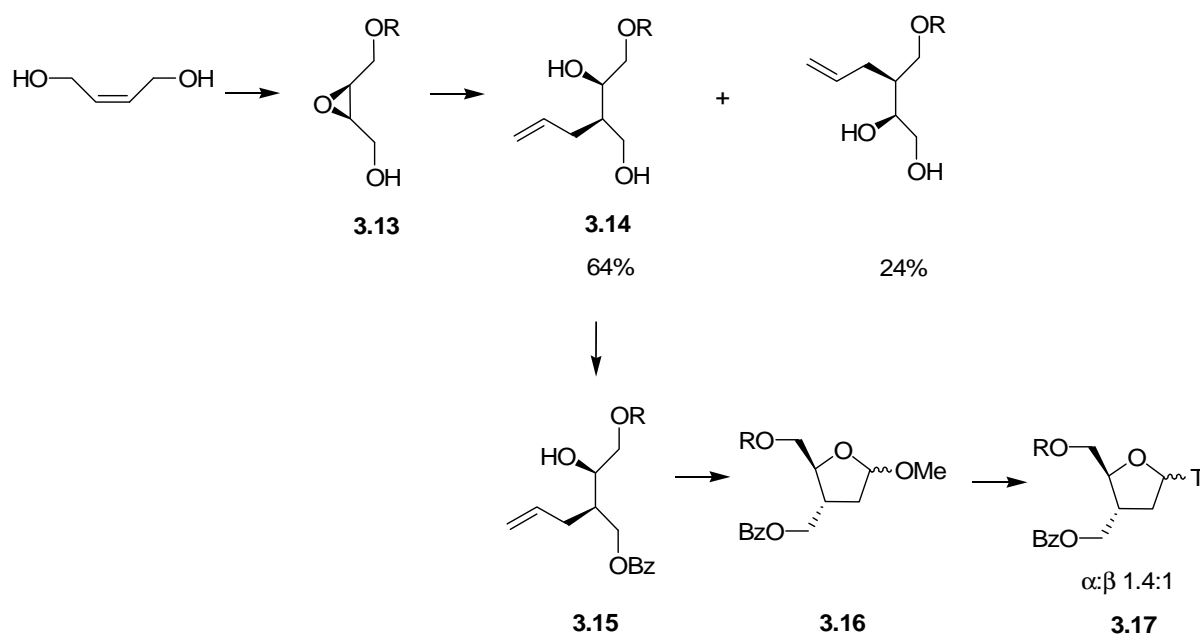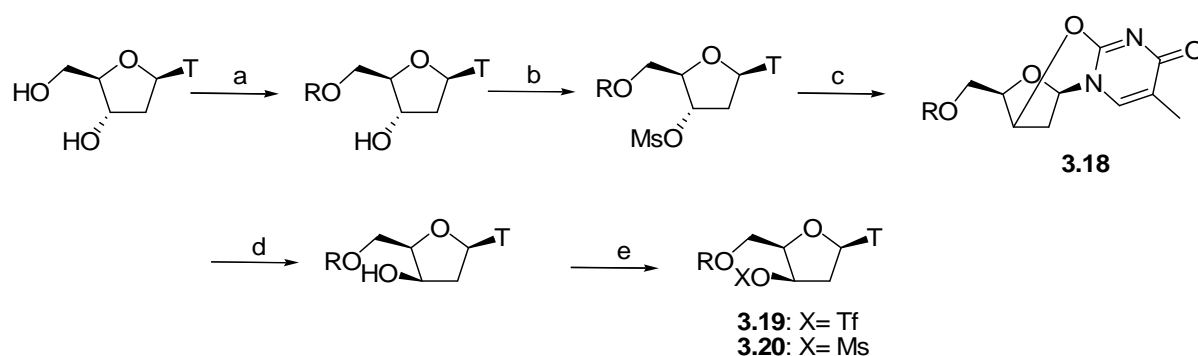

Scheme 3.4, reagents and conditions: (a) (i) RBr, NaH, THF, TBAI, rt, 15h; (ii) Ti(*O*-iPr)$_4$, diisopropyltartrate, *tert*-butyl hydroxyperoxide, -20 °C, 16h; (b) AllylMgBr, Et$_2$O, -50 °C, 30 min; (c) BzCl, pyridine, 0 °C, 15 min; (d) (i) OsO$_4$, NMNO, THF/H$_2$O, rt, 16h; (ii) NaIO$_4$, THF/H$_2$O, rt, 30 min; (iii) Na, NH$_3$, rt, 30 min; (e) silylated base, TBDMSOTf, CH$_2$Cl$_2$, rt, 24h.

To circumvent problems associated with non-stereoselective glycosylation using 2'-deoxy sugars, we decided to use the **2'-deoxynucleoside thymidine** with intrinsic β-stereochemistry at the anomeric position as starting material.

3.3.2.2 Substitution of 3'-mesylate or 3'-triflate ester or opening of 2,3'-O-anhydro with cyanide

Substitution of the β-oriented 3'-triflate (**3.19**) or 3'-mesylate esters of 2'-deoxynucleosides (**3.20**) with azide, amine and few carbon nucleophiles (e.g. cyano-anion) or opening of 2,3'-O-anhydro-derivative (**3.18**) of 2'-deoxynucleosides with amines or azides has proven successfull.[129] The synthesis of compounds **3.19** and **3.20** requires the inversion of the stereochemistry at the 3'-position, which is performed over 2,3'-O-anhydro derivative **3.18** (Scheme 3.5).
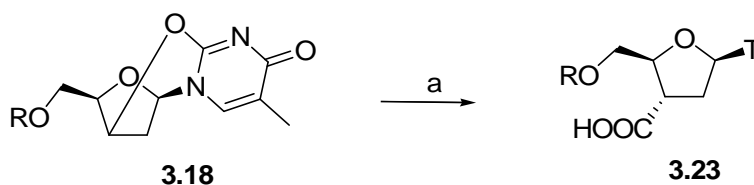

Scheme 3.5, Reagents and conditions: (a) RCl; (b) MsCl, pyridine, 0 °C, 2h; (c) DBU, acetonitrile, reflux, 16h; (d) NaOH, EtOH/H$_2$O, THF, rt, 4h; (e) MsCl or (CF$_3$SO$_2$)$_2$O

Although Schreiber[130] reported the reaction of a 3'-triflate ester with NaCN to provide 3'-cyano thymidine **3.21** along with the 2',3'-olefin **3.22** in 71% yield (**3.21**:**3.22** in a ratio 3:2), this method seemed not reproducible.[128] Others[131] tried to perform the same reaction using NaCN or Et$_4$NCN but obtained only the undesired 2',3'-didehydronucleoside **3.22**.
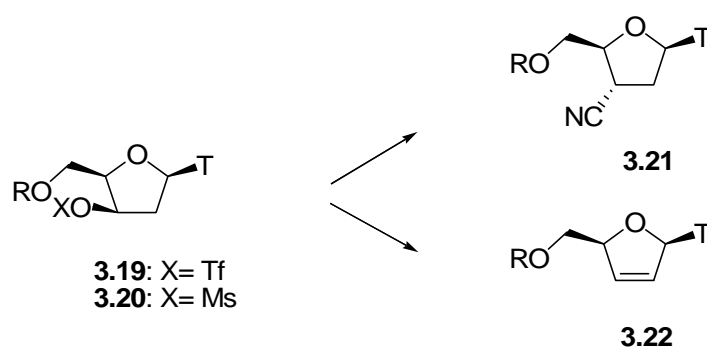
Scheme 3.6 Strategy for nucleophilic attack of a cyanide ion

Saha[132] reported the reaction of anhydronucleoside **3.18** with CuCN in the presence of MeOTf to provide **3.23**, the nitrile addition product being hydrolyzed to the carboxylic acid in the presence of cuprous ion (Scheme 3.7). According to Faul et al. this reaction was not reproducible.[128]

3.18   **3.23**

Scheme 3.7, Reagents and conditions: (a) MeOTf, CuCN, rt

As conflicting results were published concerning those reactions, compounds **3.20** and **3.18**, obtained as intermediates during other synthesis pathways (see further), were used to repeat those reactions. Compound **3.18** was reacted with trimethylsilyl cyanide in DMF. Neither heating, nor the addition of acid or Lewis acids appeared successful in the opening of the 2,3'-*O*-anhydro ring. Compound **3.20** was treated with tetrabutylammonium cyanide in acetonitrile. Elimination resulted in complete conversion to **3.22**.

From these results it coulcarbd be concluded that the cyanide appeared to be **too weak as carbon nucleophile** to open 2,3'-*O*-anhydrothymidine or to attack 3'-mesylated or triflated compounds.

3.3.2.3 *Metal-catalyzed hydroformylation of 2',3'-didehydro-2',3'-dideoxynucleosides*

A **very direct approach** involves the functionalization of 2',3'-didehydro-2',3'-dideoxynucleoside **3.24** by hydrometalation-carbonylation, which provides the extra carbon at the 3'-position in one single step.[128] Moreover, compound **3.24** can very easily be obtained from the antiviral drug d4T, which is commercially available and very cheap ($136/50g).[133] The major problem however is the regioselectivity. As the obtained aldehydes are easily equilibrated to the desired *erythro*-isomer, the facial selectivity is less important.

Using Rh(CO)$_2$acac.PPh$_3$ as a catalyst in the case of a uracil base regiomers **3.25** and **3.26** were obtained in a 2.7:1 ratio in a total yield of 37%. For cytidine counterpart a 3.2:1 ratio of **3.25** and **3.26** was achieved in 42% yield.

Although this reaction efficiently introduces the desired extra carbon, the **resulting regioisomers** are difficult to purify. Not only the regioisomers need to be separated. The presence of a very small amount of the unwanted epimers (<5%) further complicates the resolution of the reaction mixture.

d4T   **3.24**

Rh(CO)$_2$acac
Ph$_3$P, THF, 60 °C, 48h
CO (40psi), H$_2$ (40psi)

3.25   **3.26**
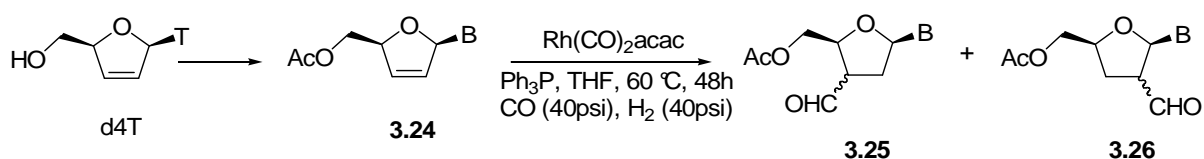
**Scheme 3.8** Hydrometalation-carbonylation

3.3.2.4 Free radical introduction of a 3'-substituent of 2'-deoxynucleosides

A third method described in the literature uses a 3'-iodo or a 3'-phenoxythiocarbonyl substituted nucleoside, which is activated to a free radical after treatment with 2,2'-azobisisobutyronitrile (AIBN) in benzene.

It was demonstrated that β-tributylstannylstyrene-mediated C-C bond formation was able to generate a 3'-formyl functionality in ribonucleosides.[134] Also the isocyanide radical can be used as the C-containing radical reagent, attacking the 3'-position.[135]

The sterical hindrance of the selected 5'-O-protecting group as well as the nucleobase favours the radical attack from below, resulting in the desired *erythro*-configuration.

The two different methods will be used in this work. Reduction of the 3'-cyano compound leads to 3'-aminomethyl-3'-deoxythymidine, while 3'-styrene-3'-deoxythymidine can be converted to 3'-hydroxymethyl-3'-deoxythymidine by successive *cis*-dihydroxylation, periodate cleavage and final reduction. For both procedures, the yield of the free radical reaction is around 60%.
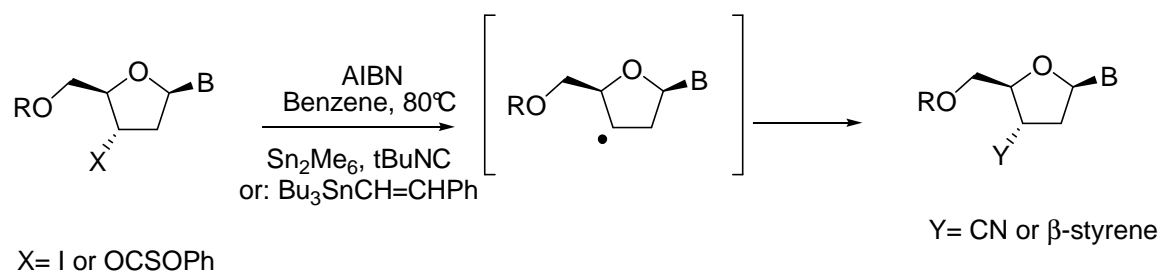

Scheme 3.9 Radical introduction of a 3'-carbon

3.4 RESULTS AND DISCUSSION

3.4.1 3'-*C*-branched ribonucleosides

For the synthesis of 3'-azidomethyl-3'-deoxythymidine, we started from 1,2-*O*-isopropyl-α-D-xylofuranose (Scheme 3.10). Reaction of the primary alcohol with TsCl gave **3.27**, which was treated with the sodium salt of benzyl alcohol at 100 °C to afford **3.28**, presumably via the intermediate oxetane. At 70 °C only the intermediate oxetane was formed. As this compound is volatile, the *in situ* conversion to the benzylated **3.28** at 100 °C is highly recommended.

On this 5-protected sugar a 3-*C*-branching was introduced by successive Swern and Wittig reaction, followed by hydroboration. After conversion of the 3-hydroxymethyl to an azidofunction by substituting the corresponding mesylate ester with NaN₃ in DMF, the sugar

was prepared for Vörbruggen coupling, using a one-pot conversion of the 1,2-isopropylidene-protected sugar into the 1,2-*O*-diacetylated protected sugar. Coupling with persilylated thymine to afford compound **3.32** was realized in 95 % yield. This precious ribo-intermediate will be used in Chapter 4, for the synthesis of a series of 2',3'-bicyclic nucleosides.
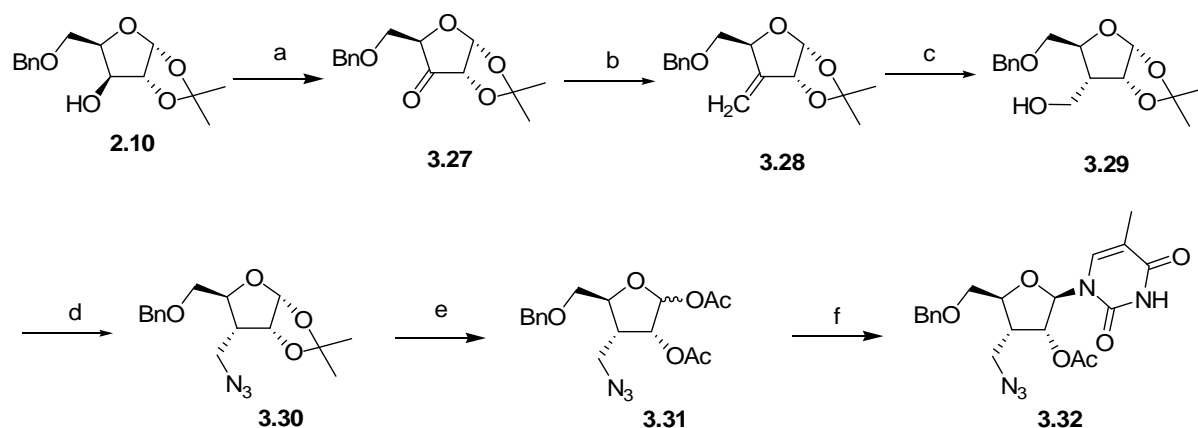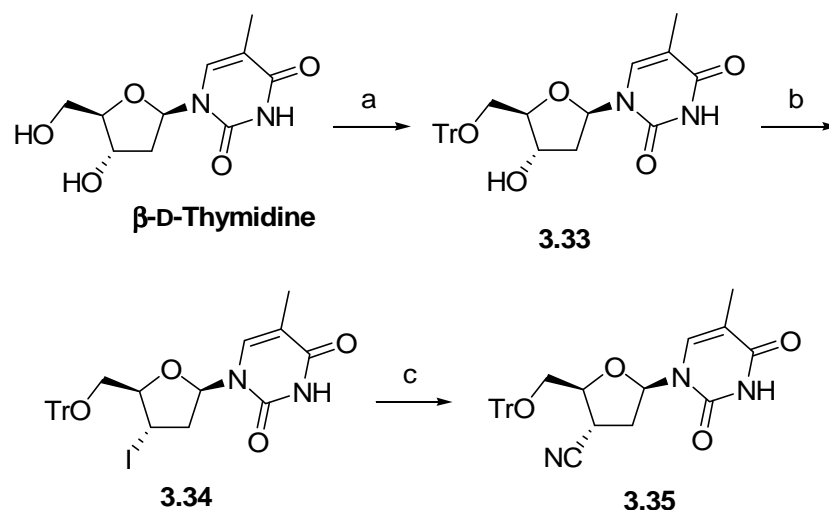

Scheme 3.10, reagents and conditions: (a) (COCl)$_2$, CH$_2$Cl$_2$, DMSO, -78 °C, 2h, then, TEA, rt, 2.5 h; (b) CH$_3$P(C$_6$H$_5$)$_3$Br, NaH, DMSO, 70%, rt, 2.5h, 70% from **2.10**; (c) 1M Borane/THF, 2N NaOH, 35% H$_2$O$_2$, rt, 1.5 h, 79%; (d) i) CH$_3$SO$_2$Cl, pyridine, 0 °C, 1h; ii) NaN$_3$, DMF, 60 °C, 16h, 77%; (e) i) 75% CH$_3$COOH, 80 °C, 16h; ii) (CH$_3$CO)$_2$O, rt, 3h, 74%; (f) silylated base, trimethylsilyltriflate, (CH$_2$)$_2$Cl$_2$, 40 °C, 5h, 95%.

3.4.2 3'-*C*-branched 2'-deoxynucleosides

3.4.2.1 Radical cyanation

We tried to reproduce the free radical introduction of a 3'-substituent. The synthesis starts with protection of the 5'-hydroxyl function of β-D-thymidine, followed by iodination of the 3'-hydroxyl function in **3.33** with methyltriphenoxyphosphonium iodide.

Although a yield of around 60% has been described for the free radical reaction, in our hands compound **3.35** was only obtained in 32% yield. This intermediate can further be reduced to obtain the 3'-aminomethyl substituted compound, used in chapter 5.

Scheme 3.11 , reagents and conditions: (a) TrCl, DMAP, pyridine, 65 °C, 16h, 83%; (b) $CH_3(PhO)_3P^+I^-$, DMF, rt, 16h, 60%; (c) *t*-BuNC, AIBN, $(CH_3)_6Sn_2$, benzene, 80 °C, 13h, 32%.

While Parkes et al. described stereoselective formation of the *erythro*-isomer,[135] other groups reported the presence of the *threo*-isomer during this reaction.[136] In our hands, the *threo*-isomer was not isolated, but as the reaction mixture contained several side products, it is not possible to exclude its formation. A more polar compound was isolated in a 14% yield. NMR and mass spectroscopy elucidated that this compound was amide **3.36**. A possible mechanism for its formation is proposed in Scheme 3.12. After the radical attack of *tert*-butylisonitrile, the radical on the cyano-carbon is captured by water (during workup), or by radicals present in the reaction mixture, leading to amide **3.36** during workup of the reaction.
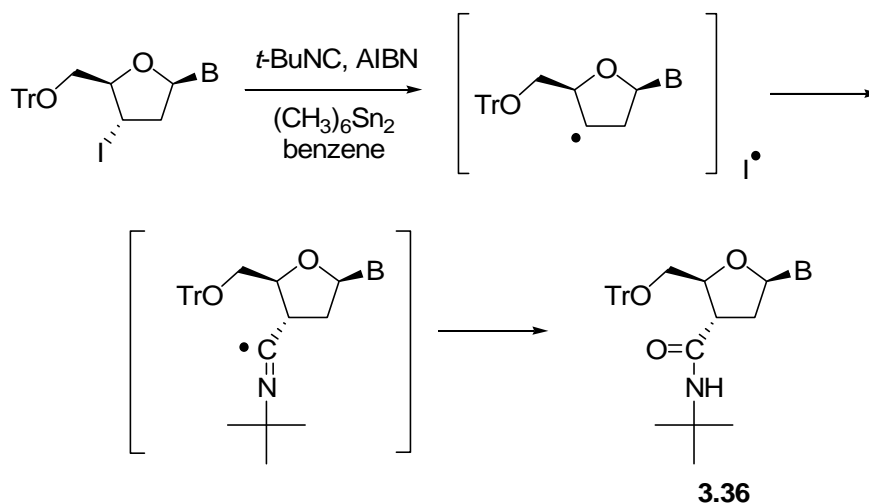
Scheme 3.12 Mechanism of the formation of the main side product **3.36** during radical cyanation

3.4.2.2 Radical introduction of β-styrene

We followed the method of Shangvi et al.[134] to generate a 3'-*C*-formyl functionality via an intermolecular radical C-C bond formation reaction. The synthesis started from β-D-thymidine. Reaction of the primary alcohol with *tert*-butyldiphenylsilyl chloride gave **3.37**, which was treated with phenylchlorothionoformate in acetonitrile to afford compound **3.38**. Treatment with β-tributylstannyl styrene and 2,2'-azobisisobutyronitrile in benzene gave **3.39**. A two-step one pot reaction involving *cis*-dihydroxylation with osmium tetroxide and 4-methylmorpholine N-oxide as cooxidans, followed by sodium periodate cleavage of the diol, resulted in an unstable aldehyde, which was immediately reduced with sodium borohydride in aqueous ethanol to afford **3.40**. This compound forms an alternative strategy for the synthesis of 3'-*C*-branched thymidine derivatives in Chapter 5.
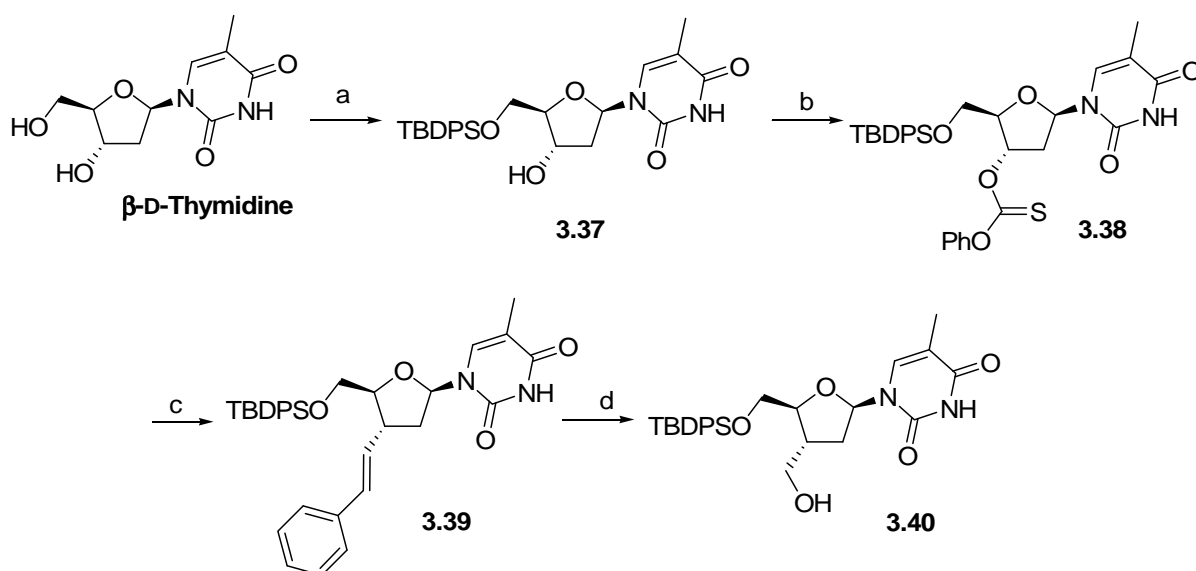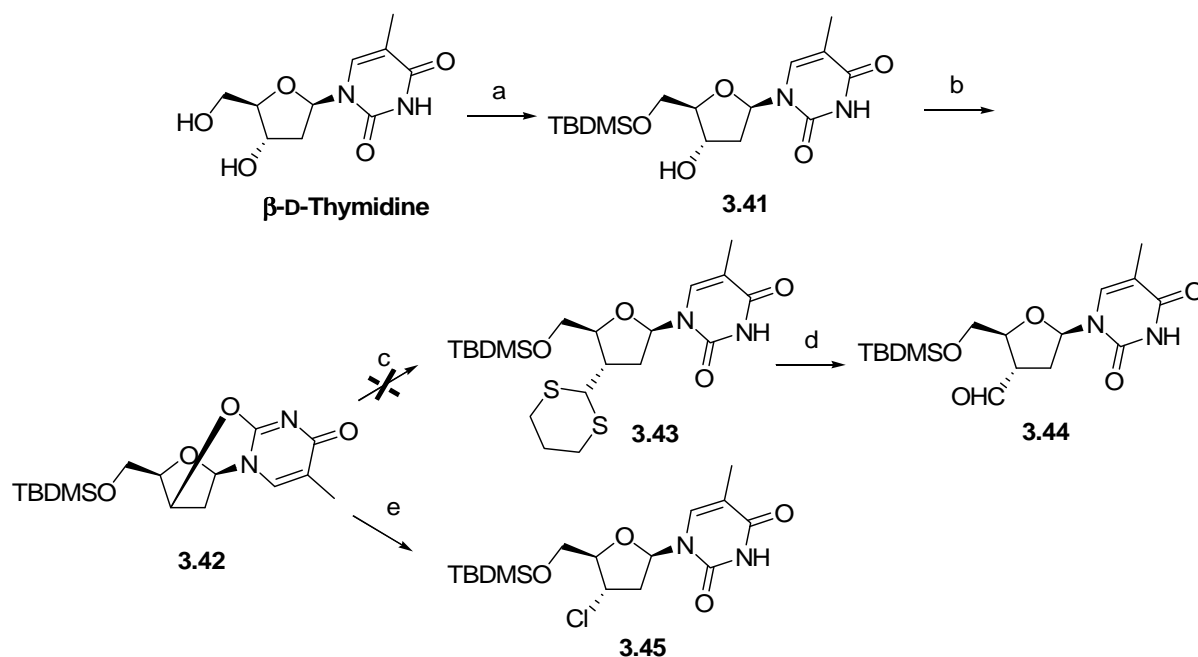

Scheme 3.13 , Reagents and conditions: (a) TBDPSCl, imidazole, DMF, rt, 16h, 80%; (b) PhOCSOCl, DMAP, acetonitrile, rt, 16h, 64%; (c) β-Bu$_3$SnCH=CHPh, AIBN, benzene, reflux, 72h, 24%; (d) (i) 4-NMO, OsO$_4$, dioxane, rt, 16h; (ii) NaIO$_4$, rt, 2h; (iii) NaBH$_4$, EtOH/H$_2$O, rt, 2h, 59%;

3.4.2.3 Attempted opening of 2,3'-O-anhydrothymidine by 1,3-dithiane

1,3-Dithiane acts as a sulfur-stabilized acyl anion equivalent and has been widely used as a masked nucleophilic acylating agent.[137]

Treatment of 1,3-dithiane with *n*-BuLi, results in the formation of the (p,d)π-stabilized anion. After reaction with an electrophile, hydrolysis of the dithioacetal moiety provides the corresponding aldehyde.

A review describing the use of dithiane in natural product synthesis illustrates the utility of epoxides, alkyl halides, sulphonates and triflates as ideal electrophiles to react with the 1,3-dithiane anion.[138]   Since the use of triflates and sulphonates was anticipated to give elimination products, we decided to explore the use of the 2,3'-*O*-anhydrothymidine derivative **3.42** as electrophile (

Scheme 3.14).

Scheme 3.14 , Reagents and conditions: (a) TBDMSCl, DMAP, pyridine, 0 °C, 2h, 88%; (b) (i) MsCl, pyridine, 0 °C, 2h; (ii) DBU, acetonitrile, reflux, 16h, 95%; (c) 1,3-dithiane, *n*-BuLi, DMPU, THF, -20 °C to rt, 16h; (d) MeI, CaCO₃, acetonitrile, H₂O; (e) 1,3-dithiane, n-BuLi, BF₃.etherate, THF, CH₂Cl₂, -20 °C to -10 °C, 2h.

Unfortunately, no reaction could be observed. Attempts to activate the anhydroring by the addition of Lewis acids (e.g. BF₃) or metals (e.g. CuSO₄) and the use of different solvents all failed.

Interestingly, in case CH₂Cl₂ was used as solvent and BF₃ as activating agent, the anhydroring was opened by a chloride anion, resulting in the undesired 3'-chloro substituted compound **3.45**.

3.5 CONCLUSION

So far, no short, high-yield procedure is known for the synthesis of nucleosides with an additional carbon at the 3'-position. For the synthesis of a suitable 3'-*C*-branched chain ribonucleoside building block, we first constructed the sugar part, which was coupled with silylated thymine under Vorbrüggen conditions.

2'-Deoxynucleosides were synthesized from β-D-thymidine using a radical introduction of a 3'-*C*-containing substituent. Attempts to use the 1,3-dithiane-anion to open 2,3'-*O*-anhydrothymidine remained unsuccessful.

3.6 EXPERIMENTAL PART

5-O-Benzyl-3-deoxy-1,2-*O*-isopropylidene-3-methylene-α-D-ribofuranose (3.28)

To a solution of oxalyl chloride (1.24 mL, 14.5 mmol) in CH_2Cl_2 (15 mL) at -78 °C, DMSO (2.06 mL, 29 mmol) in CH_2Cl_2 (9 mL) was added. After stirring for 15 minutes, compound **2.10** (3.9 g, 13.2 mmol) in CH_2Cl_2 (29 mL) was added dropwise. The reaction was stirred at -78 °C for 2 h. Triethylamine (3.54 mL, 43.4 mmol) was added and the reaction mixture was warmed to room temperature, stirred for 2.5 h, then poured into H_2O (75 mL) and extracted with CH_2Cl_2 (3 x 75 mL). The organic layer was washed with brine (50 mL), dried over $MgSO_4$ and evaporated to dryness. The crude ketone **3.27** was used in the next step without any further purification.

A suspension of NaH (1.1 g, 46.3 mmol) in DMSO (60 mL) was heated at 65 °C under nitrogen atmosphere until all the sodium hydride was dissolved. The solution was cooled to room temperature and methyltriphenylphosphonium bromide (18.2 g, 51 mmol) was added while stirring. After 1.5 h, compound **3.27** (13.2 mmol) in DMSO (20 mL) was added and the mixture was stirred for 2.5 h. The mixture was then poured into ice water (450 mL) and extracted with pentane (5 x 300 mL). The organic layer was dried over $MgSO_4$ and evaporated to give a syrup, which was purified by column chromatography (pentane/ethylacetate 97:3) to afford compound **3.28** (2.5 g, 70%) as a slightly yellow oil.
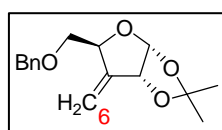
^{1}H NMR (300 MHz, $CDCl_3$): δ 1.37 (3H, s, CH_3), 1.50 (3H, s, CH_3), 3.55 (1H, dd, *J*= 5.2 and 10.5 Hz, H-5), 3.66 (1H, dd, *J*= 3.5 and 10.5 Hz, H-5'), 4.57 (2H, s, *CH₂*Ph), 4.87 (1H, m, H-4), 4.89 (1H, d, *J*= 4.1 Hz, H-2), 5.18 (1H, m, H-6), 5.41 (1H, m, H-6'), 5.86 (1H, d, *J*= 4.1 Hz, H-1), 7.24-7.33 (5H, m, CH_2*Ph*).

Exact mass (ESI-MS) for $C_{16}H_{20}O_4Na$ [M+Na][+] found, 299.1261; calcd, 299.1259.

### 5-*O*-Benzyl-3-hydroxymethyl-1,2-*O*-isopropyl-α-D-ribofuranose (3.29)

To a solution of compound **3.28** (2.2 g, 7.9 mmol) in THF (16 mL) at 0 °C, a 1M borane/THF-solution (17 mL) was added under vigorous stirring. The solution was stirred at room temperature for 1.5 h. After cooling to 0 °C, the mixture was subsequently treated with THF/H_2O (8.1 mL, 1:1, v/v), 2 N NaOH (24.4 mL) and 35% H_2O_2 (10.5 mL). The turbid mixture was stirred at room temperature for 1h. Ether (50 mL) was added to the reaction mixture, which was then washed twice with ice water (30 mL) and brine (30 mL). After drying over $MgSO_4$, the solvents were evaporated and the residue was purified by column chromatography (CH_2Cl_2/MeOH 99:1) to yield compound **3.29** (1.83 g, 79%) as a white solid.
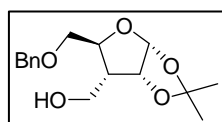
¹H NMR (300 MHz, CDCl₃): δ 1.32 (3H, s, CH₃), 1.51 (3H, s, CH₃), 2.17 (1H, m, H-3), 3.59 (1H, dd, *J*= 6.6 and 10.6 Hz, H-6), 3.76-3.83 (3H, m, H-6', H-5, H-5'), 4.21 (1H, m, H-4), 4.59 (2H, s, *CH₂*Ph), 4.75 (1H, dd, *J*= 3.5 and 4.2 Hz, H-2), 5.81 (1H, d, *J*= 3.5 Hz, H-1), 7.28-7.35 (5H, m, CH₂*Ph*).

Exact mass (ESI-MS) for $C_{16}H_{22}O_5Na$ [M+Na]⁺ found, 317.1362; calcd, 317.1365.

Next to desired compound **3.29**, 10 % (230 mg) of compound **3.29b** was obtained.
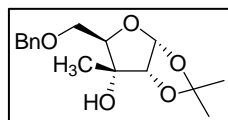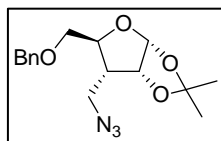

¹H NMR (300 MHz, DMSO-d₆): δ 1.12 (3H, s, CH₃), 1.17 (3H, s, CH₃), 1.32 (3H, s, CH₃), 3.46 (1H, dd, *J*=7.8 and 11.1 Hz, H-5), 3.62 (1H, dd, *J*= 5.4 and 11.2 Hz, H-5'), 3.82 (1H, m, H-4), 4.10 (1H, d, *J*= 3.6, H-2), 4.47 (2H, s, *CH₂*Ph), 4.97 (1H, d, *J*= 0.6 Hz, 3-OH), 5.76 (1H, d, *J*= 3.5 Hz, H-1), 7.28-7.35 (5H, m, CH₂*Ph*).

3-Azidomethyl-5-*O*-benzyl-1,2-*O*-isopropyl-α-D-ribofuranose (3.30)

To a solution of **3.29** (8.35 g, 28.4 mmol) in pyridine (100 mL) at 0 °C, methanesulfonyl chloride (5.7 mL, 73.8 mmol) was added. The reaction mixture was stirred for 1 h at 0 °C. CH_2Cl_2 (75 mL) was added, and the organic layer was washed with $NaHCO_3$ (75 mL) and H_2O (75 mL), dried over $MgSO_4$ and evaporated.

A solution of the crude mesylated compound (28.4 mmol) and NaN_3 (18.4 g, 284 mmol) in DMF (70 mL) was heated at 60 °C overnight. The reaction mixture was evaporated *in vacuo*. The residue was resolved in CH_2Cl_2 (75 mL), and washed with H_2O (3 x 75 mL). The organic layer was dried over $MgSO_4$, evaporated and purified by column chromatography (pentane/ethylacetate 95:5) to afford compound **3.30** (9.1 g, 77%) as a colourless oil.

^{1}H NMR (300 MHz, DMSO-d_6): δ 1.27 (3H, s, CH$_3$), 1.39 (3H, s, CH$_3$), 2.85 (1H, m, H-3), 3.38 (1H, m, H-6), 3.49 (2H, m, H-6' and H-5), 3.62 (1H, m, H-5'), 3.88 (1H, m, H-4), 4.48 (2H, s, *CH$_2$*Ph), 4.69 (1H, dd, *J*= 2.7 and 3.6 Hz, H-2), 5.67 (1H, d, *J*= 3.7 Hz, H-1), 7.31 (5H, m, CH$_2$*Ph*).
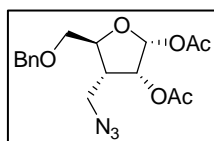
Exact mass (ESI-MS) for C$_{16}$H$_{21}$N$_3$O$_4$Na [M+Na]$^+$ found, 342.1436; calcd, 342.1430.

1,2-Di-*O*-acetyl-3-azidomethyl-5-*O*-benzyl-α-D-ribofuranose (3.31)

A solution of **3.30** (1.19 g, 3.73 mmol) in acetic acid/H$_2$O (3:1, 40 mL) was heated at 80 °C overnight. The reaction mixture was evaporated to dryness and coevaporated twice with toluene (20 mL).

The obtained residue was dissolved in pyridine (29 mL). Acetic anhydride (4.44 mL) was added and the reaction was allowed to stir for 3 h at room temperature. The resulting mixture was evaporated and purified by column chromatography (CH$_2$Cl$_2$/MeOH, 99:1) to obtain **3.31** (1.00 g, 74%) as a mixture of two anomers.


NMR of the most polar anomer:
^{1}H NMR (300 MHz, CDCl$_3$): δ 1.96 (3H, s, CH$_3$), 2.17 (3H, s, CH$_3$), 2.78 (1H, m, H-3), 3.52 (1H, dd, *J*= 6.6 and 12.4 Hz, H-6), 3.65 (1H, dd, J= 8.5 and 12.4 Hz, H-6'), 4.33-4.60 (3H, m, H-4, H-5 and H-5'), 4.54 (2H, s, *CH$_2$*Ph), 5.29 (1H, d, *J*= 4.7 Hz, H-2), 6.14 (1H, s, H-1), 7.31 (5H, m, CH$_2$*Ph*).

Exact mass (ESI-MS) for C$_{17}$H$_{21}$N$_3$O$_6$Na [M+Na]$^+$ found, 386.1332; calcd, 386.1328.

2'-*O*-Acetyl-3'-azidomethyl-5'-*O*-benzyl-β-D-thymidine (3.32)

Thymine (2.22 g, 17.62 mmol) was dissolved in HMDS (100 mL) and pyridine (15.87 mL). TMSCl (1.61 mL) was added and the mixture was heated at 130 °C overnight. After cooling of the clear reaction mixture, it was evaporated to dryness and coevaporated twice with toluene.

The obtained silylated base was dissolved in dry (CH$_2$)$_2$Cl$_2$ (10 mL). Compound **3.31** was dissolved in (CH$_2$)$_2$Cl$_2$ (10 mL) and added to the base. TMS triflate (479 µL) was added and the clear solution was stirred at 40 °C during 5 h. The reaction mixture was then extracted with CH$_2$Cl$_2$ (3 x 50 mL) and NaHCO$_3$ (50 mL). The organic layer was dried and evaporated to dryness. The dry residue was purified by column chromatography to yield compound **3.32** (6.88 g, 95%) as a white foam.
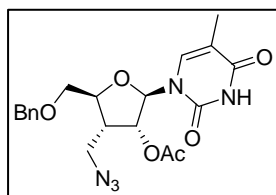
¹H NMR (300 MHz, DMSO-d₆): δ 1.53 (3H, s, CH₃), 2.07 (3H, s, COCH₃), 2.79 (1H, m, H-3'), 3.48 (1H, m, H-6'), 3.61 (2H, m, H-6" and H-5'), 3.80 (1H, m, H-5"), 4.11 (1H, m, H-4'), 4.55 (2H, s, *CH₂*Ph), 5.39 (1H, m, H-2'), 5.67 (1H, m, H-1'), 7.33 (5H, m, CH₂*Ph*), 7.55 (1H, s, H-6), 11.29 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{20}H_{24}N_5O_6$ [M+H]⁺ found, 430.1729; calcd, 430.1726.

5'-*O*-Trityl-β-D-thymidine (**3.33**)

To a solution of β-D-thymidine (5 g, 21.1 mmol) in anhydrous pyridine (50 mL) was added DMAP (2.8 g, 23.2 mmol) and trityl chloride (6.9 g, 24.8 mmol). The mixture was heated to 65 □C and stirred overnight. The reaction mixture was then diluted with CH₂Cl₂ (50 mL), washed with saturated aqueous NaHCO₃ (3 x 100 mL), and dried over anhydrous MgSO₄. The solvent was removed under reduced pressure and the resulting residue was purified by column chromatography (CH₂Cl₂/MeOH 98:2), affording **3.33** as a white foam (8.3 g, 83 %).
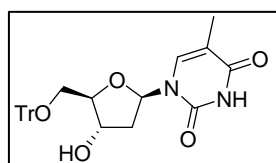

¹H NMR (300 MHz, CDCl₃): δ 1.43 (3H, s, 5-CH₃), 2.25-2.35 (1H, m, H-2'), 2.40-2.51 (1H, m, H-2"), 3.36 (2H, m, H-5' and H-5"), 4.09 (1H, d, *J*= 2.2 Hz, H-3'), 4.58 (1H, t, *J*= 2.7, H-4'), 6.44 (1H, dd, *J*= 6.2 and 7.4 Hz, H-1'), 7.21-7.41 (15H, m, Ph), 7.60 (1H, s, H-6).

Exact mass (ESI-MS) for $C_{29}H_{28}N_2O_5Na$ [M+Na]⁺ found, 507.1899; calcd, 507.1896.

3'-Deoxy-3'-iodo-5'-*O*-trityl-β-D-thymidine (**3.34**)

To a solution of **3.33** (3.1 g, 6.29 mmol) in 25 mL anhydrous DMF, methyltriphenoxy-phosphonium iodide (3.7 g, 8.18 mmol), freshly washed with EtOAc and subsequently dried, was added slowly. The mixture was stirred at room temperature for 16 h. The reaction was cooled to 0 °C and treated with 20 mL 10% aqueous Na₂S₂O₃. Addition of water was needed to effect separation of an oil. The solvent was carefully decanted. The residual oil was dissolved in 30 mL CH₂Cl₂ and washed three times with 75 mL water. The water layer was extracted with CH₂Cl₂ and the combined organic layers were dried (MgSO₄) and crystallised from CH₂Cl₂/MeOH 2/1, yielding pure compound **3.34** (2.24 g, 60%).
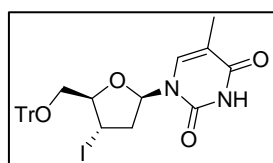
¹H NMR (300 MHz, CDCl₃): δ 1.47 (3H, d, *J*= 0.8 Hz, 5-CH₃), 2.77-2.82 (2H, m, H-2' and H-2"), 3.47-3.60 (2H, m, H-5' and H-5"), 4.30 (1H, dt, *J*= 2.0 and 8.6 Hz, H-4'), 4.48 (1H, dd, *J*= 8.5 and 8.8 Hz, H-3'), 6.15 (1H, dd, *J*= 4.2 and 5.9 Hz, H-1'), 7.24-7.45 (15H, m, Ph), 7.68 (1H, d, *J*= 0.9 Hz, H-6), 9.63 (1H, br s, NH).

Exact mass (ESI-MS) for $C_{29}H_{27}N_2O_4INa$ [M+Na]⁺ found, 617.0919; calcd, 617.0915.

3'-Cyano-3'-deoxy-5'-*O*-trityl-β-D-thymidine (3.35)

To a solution of **3.34** (3.00 g, 5.05 mmol) in benzene (120 mL) was added *tert*-butyl isocyanide (5.46 g, 65.65 mmol) and hexamethylditin (1.65 g, 5.05 mmol). Before adding AIBN (415 mg, 2.53 mmol) to the reaction mixture, 40 mL of benzene was distilled under argon and removed. The reaction mixture was heated at 80 °C for 13 hours. After cooling, it was concentrated and purified by column chromatography (cyclohexane/ethylacetate 70:30) to give 800 mg of the pure **3.35** (32%).
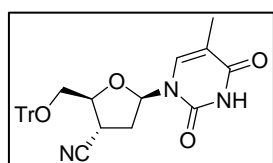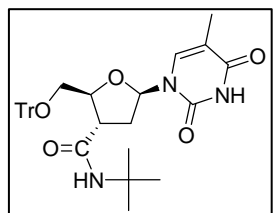

^{1}H NMR (300 MHz, CDCl$_3$): δ 1.57 (3H, s, 5-CH$_3$), 2.60 (1H, m, H-2'), 2.74 (1H, m, H-2''), 3.42-3.64 (3H, m, H-3', H-5' and H-5''), 4.28 (1H, m, H-4'), 6.18 (1H, dd, *J*= 4.2 and 7.0 Hz, H-1'), 7.26-7.46 (16H, m, Ph and H-6), 9.61 (1H, br s, NH).

^{13}C NMR (300 MHz, CDCl$_3$): δ 12.02 (5-CH$_3$), 28.12 (C-3'), 36.58 (C-2'), 61.88 (C-5'), 81.86 (C-4'), 85.36 (*C*Ph$_3$), 87.60 (C-1'), 111.38 (C-5), 118.05 (CN), 127.61, 128.18, 128.52 (C*Ph$_3$*), 135.11 (C-6), 142.97 (C*Ph$_3$*), 150.19 (C-2), 163.87 (C-4).

Exact mass (ESI-MS) for C$_{30}$H$_{27}$N$_3$O$_4$Na [M+Na]$^+$ found, 516.1905; calcd, 516.1899.

3'-*tert*-Butylcarbamate-3'-deoxy-5'-*O*-trityl-β-D-thymidine (3.36)

Compound **3.36** was obtained as an impurity of radical cyanation reaction in a yield of 14%.

^{1}H NMR (300 MHz, CDCl$_3$): δ 1.21 (9H, s, (C(C*H$_3$*)$_3$)), 1.54 (3H, s, 5-CH$_3$), 2.25-2.35 (1H, m, H-2'), 2.77-2.88 (1H, m, H-2''), 3.10 (1H, m, H-3'), 3.27 (1H, dd, *J*= 3.0 and 10.9 Hz, H-5'), 3.66 (1H, dd, *J*= 2.7 and 11.2 Hz, H-5''), 4.22 (1H, dt, *J*= 2.7 and 9.0 Hz, H-4'), 6.13 (1H, dd, *J*= 3.0 and 7.2 Hz, H-1'), 7.26-7.46 (15H, m, Ph), 7.65 (1H, s, H-6), 8.69 (1H, br s, NH).

^{13}C NMR (300 MHz, CDCl$_3$): δ 12.05 (5-CH$_3$), 28.60 (C-3'), 28.61 (C(*CH$_3$*)$_3$), 36.95 (C-2'), 45.01 (*C*(CH$_3$)$_3$), 62.40 (C-5'), 83.08 (C-4'), 85.79 (*C*Ph$_3$), 87.18 (C-1'), 110.51 (C-5), 127.49, 128.08, 128.59 (C*Ph$_3$*), 135.75 (C-6), 143.24 (C*Ph$_3$*), 150.02 (C-2), 163.76 (C-4), 169.13 (NHC=O).

Exact mass (ESI-MS) for C$_{34}$H$_{37}$N$_3$O$_5$ [M+Na]$^+$ found, 590.2000; calcd, 590.2631.

5'-*O*-*tert*-Butyldiphenylsilyl-β-D-thymidine (3.37)

To a mixture of β-D-thymidine (12 g, 42.9 mmol) and imidazole (8.76 g, 128.7 mmol) in dry DMF (75 mL) was added slowly *tert*-butyldiphenylsilyl chloride (11.16 mL, 42.9 mmol). The reaction was stirred at room temperature overnight and evaporated to dryness. The organic residue was resolved in water (150 mL) and extracted with ethyl acetate (3 x 150 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated. The residue was purified by column chromatography (CH$_2$Cl$_2$/MeOH 97:3) to yield **3.37** (16.9 g, 80%) as a white foam.
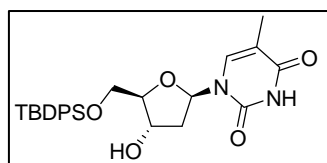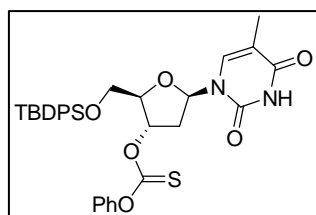

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 0.99 (9H, s, *tert*-butyl), 1.48 (3H, s, 5-CH$_3$), 2.14 (2H, m, H-2' and H-2"), 3.77 (1H, dd, *J*= 5.1 and 12.6 Hz, H-5'), 3.86 (2H, m, H-4' and H-5"), 4.32 (1H, app s, H-3'), 5.33 (1H, app s, 3'-OH), 6.19 (1H, t, *J*= 6.9 Hz, H-1'), 7.37-7.46 (7H, m, 6 ar-H and H-6), 7.58-7.62 (4H, m, ar-H), 11.35 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{26}$H$_{32}$N$_2$O$_5$SiNa [M+Na]$^+$ found, 503.1979; calcd, 503.1978.

5'-*O*-*tert*-Butyldiphenylsilyl-3'-deoxy-3'-*O*-(phenoxythiocarbonyl)-β-D-thymidine (3.38)

To a solution of **3.37** (10.39 g, 21.61 mmol) and DMAP (5.28 g, 43.22 mmol) in dry acetonitrile (100 mL) at 0 °C was added dropwise phenylchlorothionoformate (5 g, 29 mmol). The reaction mixture was stirred at room temperature for 16 hours. The next day, acetonitrile was removed *in vacuo* and the residue was purified by column chromatography (hexane/ethyl acetate 75:25) to afford **3.38** (8.5 g, 64%).

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 0.98 (9H, s, *tert*-butyl), 1.48 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 2.49-2.64 (2H, m, H-2' and H-2"), 3.94 (2H, m, H-5' and H-5"), 4.33 (1H, m, H-4'), 5.82 (1H, d, *J*= 6.0 Hz, H-3'), 6.22 (1H, dd, *J*= 5.7 and 9.0 Hz, H-1'), 7.16-7.19 (2H, m, 2 ar-H), 7.32-7.46 (10H, m, 9 ar-H and H-6), 7.59-7.62 (4H, m, ar-H), 11.35 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{33}$H$_{36}$N$_2$O$_6$SSiNa [M+Na]$^+$ found, 639.1957; calcd, 639.1961.

5'-*O-tert*-Butyldiphenylsilyl-3'-deoxy-3'-*O*-((*E*)-phenylethenyl)-β-D-thymidine (3.39)

To a solution of carbonothioate **3.38** (7.3 g, 11.85 mmol) in benzene (40 mL) was added β-tributylstannylstyrene (11.64 g, 29.62 mmol) in 20 mL benzene. The resulted solution was degassed twice with nitrogen during 30 minutes (1 x at room temperature; 1 x at 45 °C). 2,2'-azo-bis-isobutyronitrile (AIBN; 583 mg, 3.56 mmol) was added and the solution was refluxed for 2 h. A second portion of AIBN (583 mg, 3.56 mmol) was added after cooling the reaction mixture to 40 °C. The reaction mixture was refluxed again for 2 h. This procedure was repeated 10 times over a period of 72 hours (total amount of AIBN: 6.99 g). After evaporation of the solvent, the residue was purified by column chromatography (hexane/ethyl acetate 85:15) to give **3.39** as an oil (1.8 g, 24%).
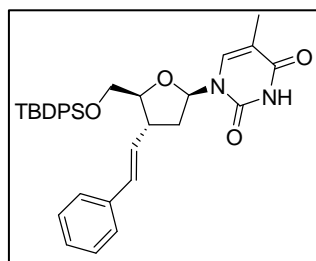
**^{1}H NMR** (300 MHz, CDCl$_3$): δ 1.05 (9H, s, *tert*-butyl), 1.54 (3H, s, 5-CH$_3$), 2.22-2.40 (2H, m, H-2' and H-2"), 3.26 (1H, m, H-3'), 3.85 (1H, m, H-4'), 4.10 (2H, m, H-5' and H-5"), 5.98 (1H, dd, *J*= 8.4 and 15.9 Hz, CH-6'), 6.19 (1H, dd, *J*= 3.3 and 7.2 Hz, H-1'), 6.50 (1H, d, *J*= 15.6 Hz, CH-Ph), 7.25-7.41 (11H, m, 11 ar-H), 7.55 (1H, d, *J*= 1.2 Hz, H-6), 7.67 (4H, m, ar-H), 8.21 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{34}$H$_{38}$N$_2$O$_4$SiNa [M+Na]$^+$ found, 589.2492; calcd, 589.2499.

5'-*O-tert*-Butyldiphenylsilyl-3'-deoxy-3'-(hydroxymethyl)-β-D-thymidine (3.40)

To a mixture of **3.39** (1.8 g, 3.11 mmol) and 4-methylmorpholine *N*-oxide (546 mg, 4.66 mmol) in dioxane (50 mL) was added an aqueous solution of OsO$_4$ (1 mL, 0.16 mmol, 1% in water). After stirring overnight at room temperature under light protection, the reaction was completed. Sodium periodate (1.33 g, 6.22 mmol) was added and after 2 h the oxidative cleavage was completed. The mixture was diluted with ethyl acetate, filtered through a celite path and solids were washed with ethyl acetate. The combined filtrates were washed with brine, dried over MgSO$_4$ and evaporated under reduced pressure. To the crude resulting aldehyde dissolved in ethanol/water (4:1, 45 mL) at 0 °C was added NaBH$_4$ (540.8 mg, 14.3 mmol) in small portions. After 2 h at room temperature, the reaction mixture was diluted with ethyl acetate and washed with water. The organic layer was dried over MgSO$_4$ and concentrated to dryness. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 97:3), compound **3.40** (900 mg, 59% from **3.39**) was isolated as a white foam.
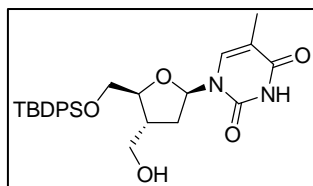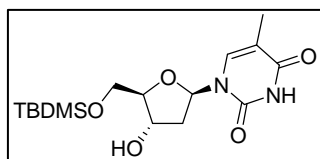
¹H NMR (300 MHz, DMSO-d₆): δ 1.00 (9H, s, tert-but), 1.49 (3H, s, 5-CH₃), 2.03-2.19 (2H, m, H-2' and H-2"), 2.48 (under DMSO-signal, H-3'), 3.43 (2H, d, *J*= 5.1 Hz, H-6' and H-6"), 3.77 (1H, dd, *J*= 4.2 and 10.5 Hz, H-5'), 3.86-3.95 (2H, m, H-4' and H-5"), 4.77 (1H, br s, 6'-OH), 6.03 (1H, dd, *J*= 5.4 and 6.6 Hz, H-1'), 7.37-7.44 (7H, m, 6 ar-H and H-6), 7.61-7.64 (4H, m, 4 ar-H), 11.23 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₂₇H₃₄N₂O₅SiNa [M+Na]⁺ found, 517.2139; calcd, 517.2134.

5'-*O*-*tert*-Butyldimethylsilyl-β-D-thymidine (3.41)

β-D-thymidine (13.47 g, 55.6 mmol) and DMAP (1.7 g, 13.9 mmol) were dissolved in pyridine (110 mL). *Tert*-butyldimethylsilyl chloride (7.96 g, 52.8 mmol) was added at 0 °C and the reaction was stirred at this temperature for 2 hours. 10 mL of water was added to quench the reaction. The reaction mixture was evaporated to dryness, resolved in ethylacetate (75 mL), and washed with water (75 mL). After drying over MgSO₄, the organic layer was evaporated and the crude oil was purified by column chromatography (CH₂Cl₂/MeOH 95:5) to obtain compound **3.41** (19.82 g, 88%) as a white foam.

¹H NMR (300 MHz, DMSO-d₆): δ 0.00 (6H, s, 2x CH₃), 0.92 (9H, s, *tert*-butyl), 1.73 (3H, s, 5-CH₃), 2.04 (2H, m, H-2' and H-2"), 3.73 (3H, m, H-4', H-5' and H-5"), 4.16 (1H, app s, H-3'), 5.26 (1H, app s, 3'-OH), 6.13 (1H, t, *J*= 6.9 Hz, H-1'), 7.44 (1H, d, *J*= 1.2 Hz, H-6), 11.31 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₆H₂₈N₂O₅SiNa [M+Na]⁺ found, 379.1663; calcd, 379.1665.

5'-*O*-*tert*-Butyldimethylsilyl-2,3'-anhydro-β-D-thymidine (3.42)

To an ice-cooled solution of **3.41** (1g, 2.8 mmol) in pyridine (7.5 mL) was added methanesulfonyl chloride (326 μl, 4.2 mmol). The reaction mixture was stirred for 2 hours, quenched with NaHCO₃ (15 mL) and extracted with ethyl acetate (3 x 25 mL). The combined organic layers were dried over MgSO₄, evaporated to dryness and the obtained white foam was dissolved in acetonitrile (12 mL). After the addition of DBU (453 μL, 3.0 mmol) the reaction was refluxed overnight. The reaction mixture was evaporated to dryness and purified by column chromatography (CH₂Cl₂/MeOH 97:3) to yield **3.42** (899 mg, 95%) as a white foam.
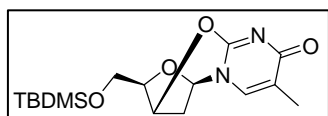
¹H NMR (300 MHz, DMSO-d₆): δ 0.00 (6H, s, 2x CH₃), 0.83 (9H, s, *tert*-butyl), 1.73 (3H, d, *J*= 1.2 Hz, 5-CH₃), 2.44 (1H, ddd, *J*= 2.7 and 12.9 Hz, H-2'), 2.53 (1H, ddd, *J*= 1.5 and 12.9 Hz, H-2"), 3.66 (2H, m, H-5' and H-5"), 4.22 (1H, ddd, *J*= 2.4 and 6.3 Hz, H-4'), 5.21 (1H, d, *J*= 2.1 Hz, H-3'), 5.81 (1H, d, *J*= 3.9 Hz, H-1'), 7.55 (1H, d, *J*= 1.2 Hz, H-6).

Exact mass (ESI-MS) for C₁₆H₂₆N₂O₄SiNa [M+Na]⁺ found, 361.1555; calcd, 361.1560;

Attempted opening of 3.42 with 1,3-dithiane

To a stirred solution of dithiane (254 mg, 2.11 mmol) in anhydrous THF (4 mL) at -20 °C, *n*-BuLi (1.26 mL, 2.01 mmol) was added over a period of 30 minutes. This mixture was stirred for 2 hours, while the temperature increased to -15 °C. DMPU (506 mL, 4.2 mmol) was added, followed by the dropwise addition of **3.42** (150 mg, 0.42 mmol) in anhydrous THF (5 mL). The reaction was stirred at -10 °C for 10 hours and then at room temperature overnight. No reaction could be observed.

When performing this reaction in CH₂Cl₂ with the addition of BF₃.etherate, a product was isolated which arose from opening of the anhydroring by a chlorine anion (**3.45**).
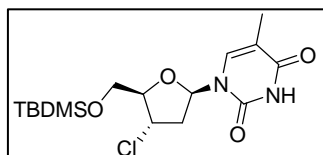
¹H NMR (300 MHz, CDCl₃): δ 0.00 (6H, s, 2x CH₃), 0.80 (9H, s, *tert*-butyl), 1.79 (3H, s, 5-CH₃), 2.32 (1H, m, H-2'), 2.48 (1H, m, H-2"), 3.85 (1H, dd, J= 2.2 and 16.0 Hz, H-5'), 3.95 (1H, dd, J= 2.1 and 15.8 Hz, H-5"), 4.21 (1H, m, H-4'), 4.30 (1H, m, H-3'), 6.35 (1H, t, *J*= 6.3 Hz, H-1'), 7.47 (1H, s, H-6), 8.17 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₆H₂₇N₂O₄SiClNa [M+Na]⁺ found, 397.1326; calcd, 397.1326.

Chapter 4

Synthesis of
bicyclic nucleosides

4 SYNTHESIS AND BIOLOGICAL EVALUATION OF BICYCLIC NUCLEOSIDES

4.1 OBJECTIVE

As described in the initial structure-activity relationship, two interesting bicyclic nucleosides with promising TMPKmt inhibitory potency have been discovered. The structure of these nucleosides is shown in Figure 4.1. The second ring is fused to the sugar ring involving its 2',3'-bond. The fused six-membered ring is built up as a thiocarbamate (**4.1**) or a carbamate (**4.2**). Except for their high affinity for the mycobacterial TMPK, they also show only a low affinity for the human enzyme, resulting in a very selective activity.
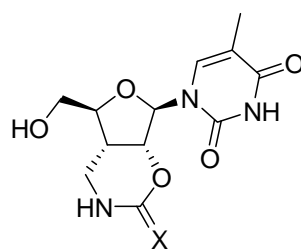
4.1. X= S
K_i(TMPKmt) = 3.5 µM
K_i(TMPKh) = 700 µM

4.2. X= O
K_i(TMPKmt) = 13.5 µM
K_i(TMPKh) = 1100 µM

Figure 4.1 Bicyclic TMPKmt inhibitors, used as leads in this chapter

The present synthetic work was directed towards functional, structural and conformational variations of these bicyclic derivatives to gain further insight into their binding mode and to improve inhibitory activity.

First, a series of bicyclic six-membered derivatives **4.3**, **4.4** and **4.5** (Figure 4.2) was envisaged in which the thiourethane moiety of **4.1** was replaced by a thiourea, a urea and a guanidine function respectively. The goal of introducing a guanidine functionality was to investigate if delocalisation of a positive charge might account for an optimal interaction with the carboxylate residue of Asp 9, which is one of the important amino acids in the active site of TMPKmt.[95] A second modification consisted of a ring contraction, resulting in five-membered bicyclic nucleosides **4.6** and **4.7**.
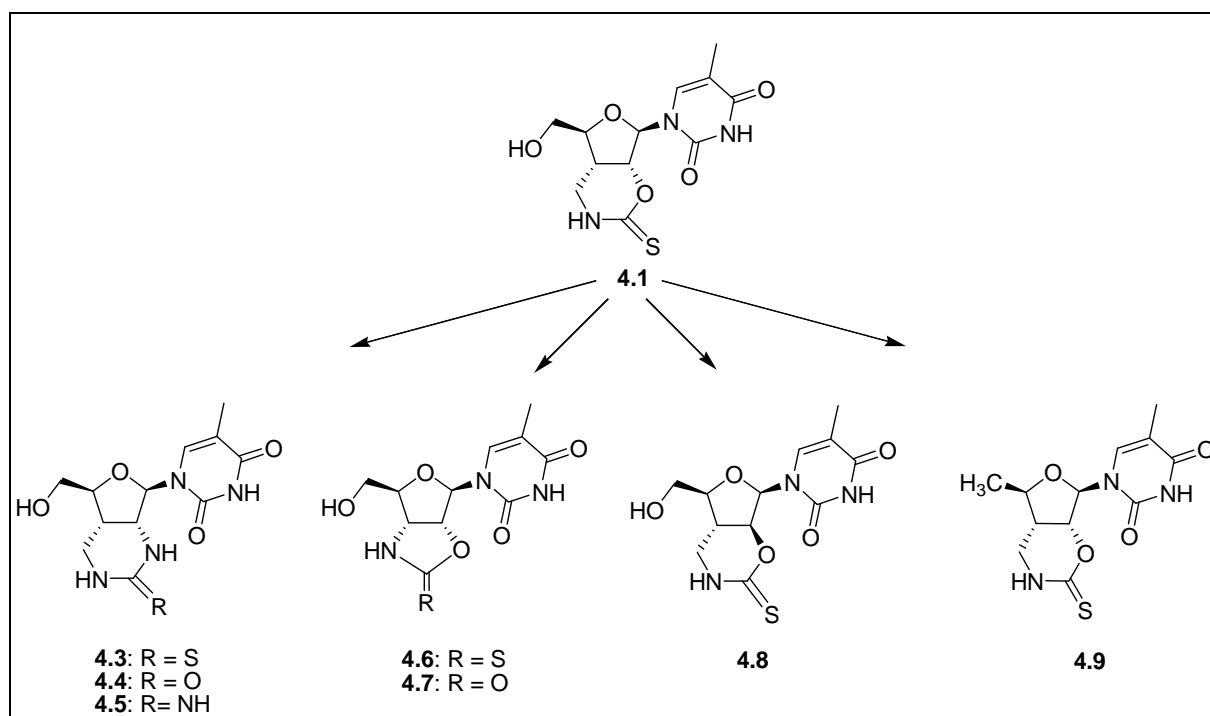
Figure 4.2 Overview of the synthesized bicyclic nucleosides

To investigate the influence of a different sugar puckering on binding the enzyme, an analogue with inverted configuration at C-2' (**4.8**) was considered.

Despite the fact that compound **4.1** is a potent inhibitor of TMPKmt, it failed to inhibit the growth of the *M. tuberculosis* H37Rv strain. As described earlier, an important hurdle in drug development for TB is the very low permeability of the cell wall. Hence, a more apolar derivative (**4.9**) was designed in order to penetrate the lipophilic bacterial cell wall more easily. This 5'-deoxy analogue also allows to further investigate the contribution of the 5'-hydroxyl group on enzyme affinity. Munier-Lehmann et al.[139] already demonstrated that removal of the 5'-OH group of thymidine improved the affinity for TMPKmt.

As mentioned before, the conformational behaviour of nucleosides is considered to be of great importance for their interactions with target proteins and thus for their biological activity.[110]

Conformation-activity studies based exclusively on solid-state conformational parameters are often not very significant. This explains why the evaluation of conformationally restrained nucleosides can be very useful, as they adopt certain restricted, geometrical shapes.[140] Kifli et al.[141] described the synthesis and conformational analysis of bicyclic nucleosides related to those reported here.

4.2 CHEMISTRY

4.2.1 Synthesis of six-membered bicyclic nucleosides **4.3**, **4.4** and **4.5**

The synthesis of the bicyclic nucleosides **4.3**, **4.4** and **4.5** started from compound **3.32**, which was synthesized as described in chapter 3, starting from 1,2-*O*-isopropyl-α-D-xylofuranose.
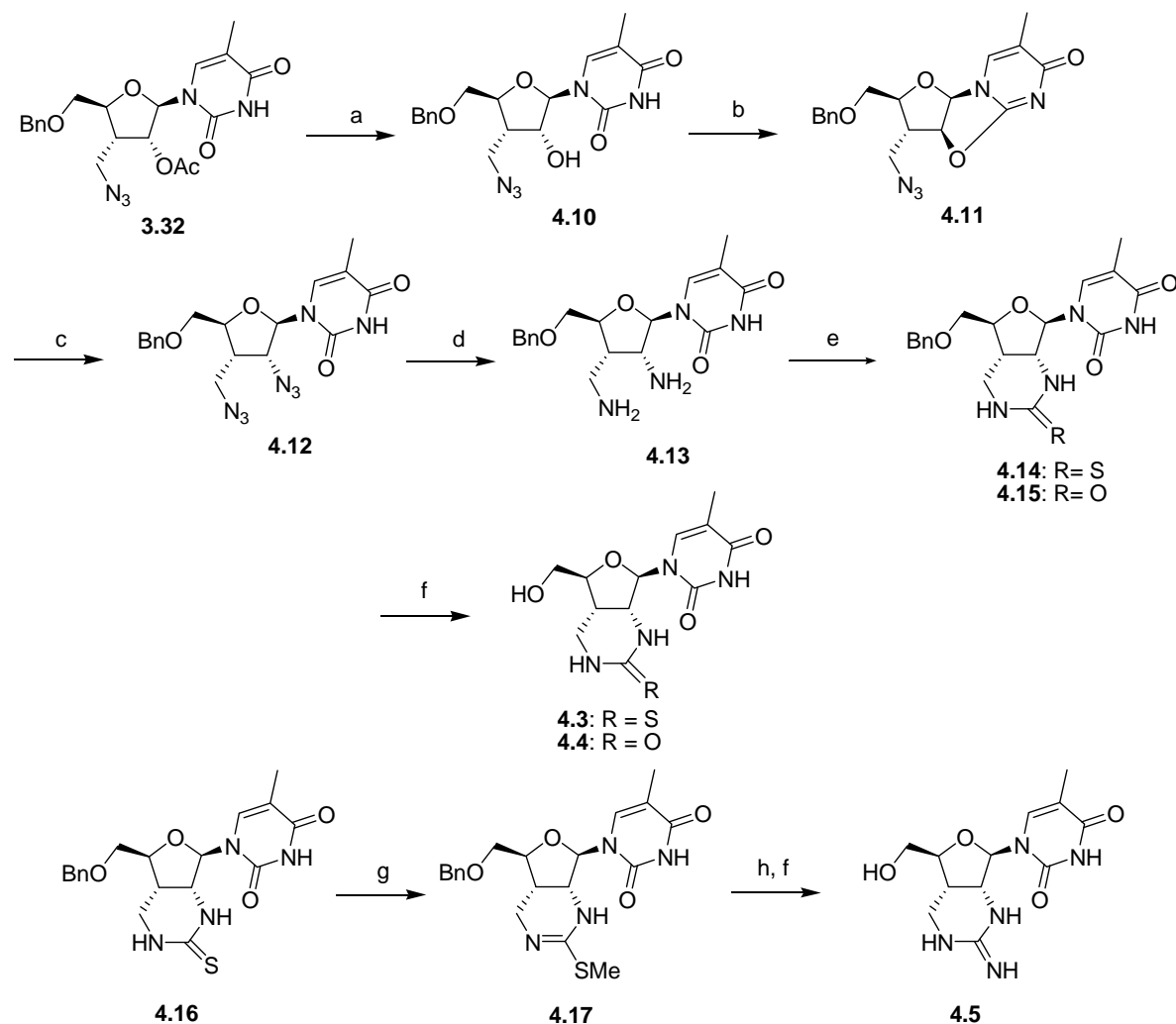
Scheme 4.1, Reagents and conditions: (a) NH_3, MeOH, rt, 4h, 95%; (b) TfCl, DMAP, CH_2Cl_2, rt, 2h, 93%; (c) NaN_3, benzoic acid, DMF, reflux, 7h, 74%; (d) Pd/C, H_2, MeOH, rt, 5h, 90%; (e) for 4.14: thiocarbonylimidazole, THF, rt, 1.5h, 60%; for 4.15: carbonyldiimidazole, THF, rt, 3h, 60%; (f) BCl_3, CH_2Cl_2, -78 °C, 15 min, 47% (4.3), 60% (4.4) and 38% (4.5); (g) MeI, MeOH, rt, 8h, 78%; (h) NH_3, MeOH, 100 °C, 16h, 50%.

The 5'-*O*-benzyl protected compound **3.32** was deacetylated and converted into anhydronucleoside **4.11** upon treatment with trifluoromethanesulfonyl chloride and DMAP. Opening of the anhydro ring with NaN$_3$ afforded the bis-azido derivative **4.12**. Reduction of both azido groups was accomplished by hydrogenolysis (Pd/C), without loss of the 5'-benzyl group. Ring closure was realized by reacting **4.13** with 1,1'-carbonyldiimidazole or its thiocarbonyl analogue in THF. Finally, deprotection with boron trichloride gave the desired bicyclic nucleosides **4.3** and **4.4**.

The thiourea compound **4.3** also served as a starting point to prepare the guanidine analogue **4.5** by subsequent methylation using iodomethane, substitution with ammonia in MeOH at 100 °C and removal of the benzyl groups.

4.2.2 Synthesis of five-membered bicyclic nucleosides **4.6** and **4.7**

The synthesis of the five-membered bicyclic nucleosides (Scheme 4.2) started from **2.13** (obtained from 1,2-*O*-isopropylidene-α-D-xylofuranose as described in chapter 2 with an overall yield of 37%).

Deacetylation and subsequent azidoreduction led to compound **4.19** that was used to synthesize two five-membered bicyclic analogues.

Ring closure and final deprotection were carried out following the same procedures as described for the six-membered rings.
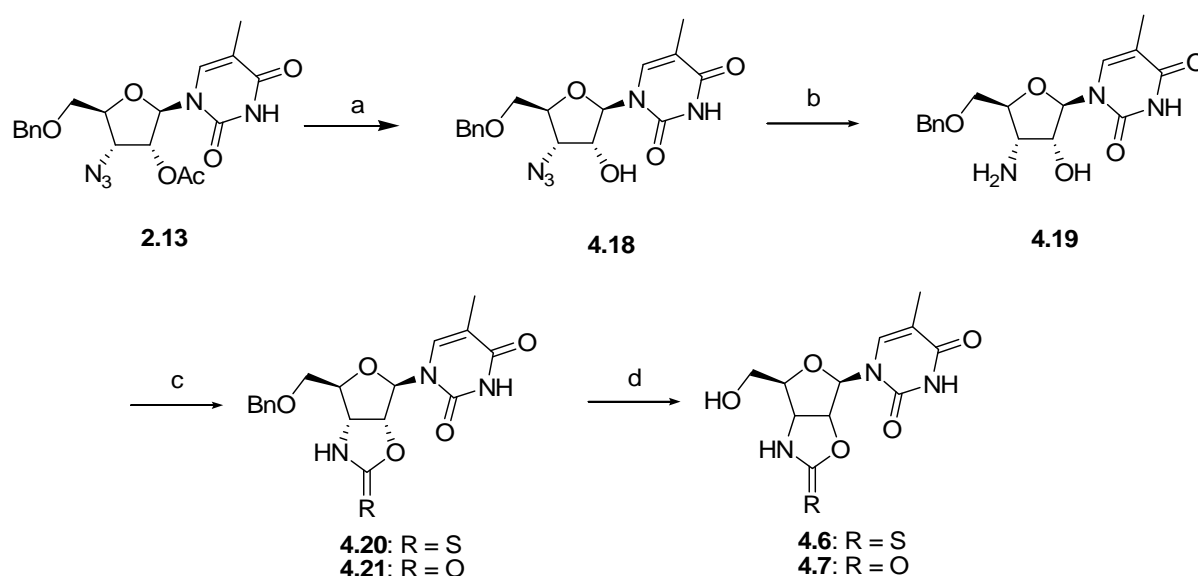
Scheme 4.2, reagents and conditions: (a) 1N NH$_3$ in MeOH, rt, 3h, 98%; (b) Pd/C, H$_2$, MeOH, rt, 3h, 96%; (c) for **4.20**: 1,1'-TCDI, THF, 50 °C, 16h, 50%; for **4.21**: 1,1'-CDI, THF, 50 °C, 16h, 72%; (d) BCl$_3$, CH$_2$Cl$_2$, -78 °C, 15 min, 70%.

4.2.3 Synthesis of nucleoside **4.8**

The synthesis of compound **4.8**, with a reversed configuration at the 2'-position, started from the previously used 2,2'-*O*-anhydronucleoside **4.11** (Scheme 4.3). The anhydro ring was opened under basic conditions to afford the arabino sugar **4.22**. Reduction of the azide then gave compound **4.23**. Ring closure with 1,1'-thiocarbonyldiimidazole was sluggish but gave access to the desired compound **4.8** after final debenzylation.

Scheme 4.3, Reagents and conditions: (a) NaOH, dioxane, EtOH/H$_2$0, rt, 3h, 92%; (b) Pd/C, H$_2$, MeOH, rt, 5h; (c) 1,1'-TCDI, THF, rt, 3h, 43%; (d) BCl$_3$, CH$_2$Cl$_2$, -78 °C, 15 min, 38%.

4.2.4 Synthesis of 5'-deoxygenated nucleoside **4.9**

Several attempts to convert **4.1** to its 5'-deoxy analogue remained unsuccessful. After introduction of a *p*-toluenesulphonyl group at the 5'-position of compound 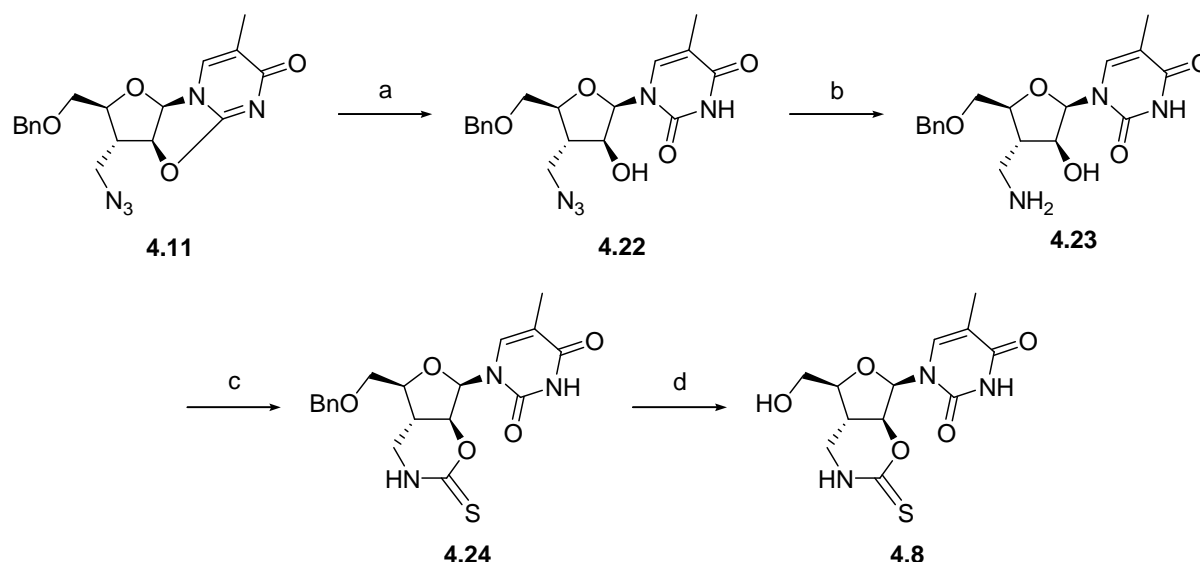**4.25**, substitution by iodine and subsequent reduction should lead to the 5'-deoxygenated compound. However this intended synthesis failed, as the first step in this pathway, the introduction of the *p*-toluenesulponyl-group, could not be achieved, most likely due to the poor sterical availability of the 5'-hydroxygroup.

Therefore we were forced to perform the deoxygenation step before the ring closure. 5'-deprotection was therefore performed on the earlier obtained compound **3.32**. During removal of the 5'-*O*-benzylgroup with BCl$_3$ under standard conditions (15 minutes, -78 °C), partial removal of the 2'-*O*-acetyl group occurred. Extension of the reaction time to 30 minutes resulted in complete 2'-deacetylated compound **4.26**.
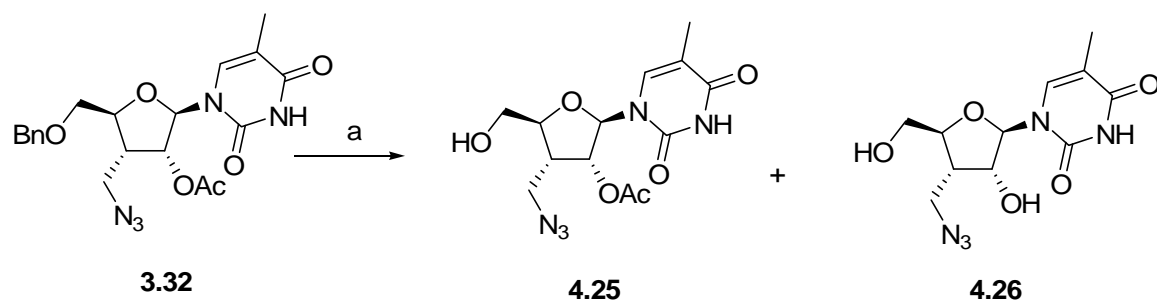
Scheme 4.4, Reagents and conditions: (a) BCl_3, CH_2Cl_2, -78 °C, 15 min, **4.25**: 30% and **4.26**: 45%.

5'-Tosylation and iodination of 5'-tosylate using NaI, gave the 5'-iodinated compound **4.28**. However, during simultaneous reduction of the 6'-azido function and the 5'-iodine, acylmigration resulted in the N-6'-acylated derivative **4.30**.
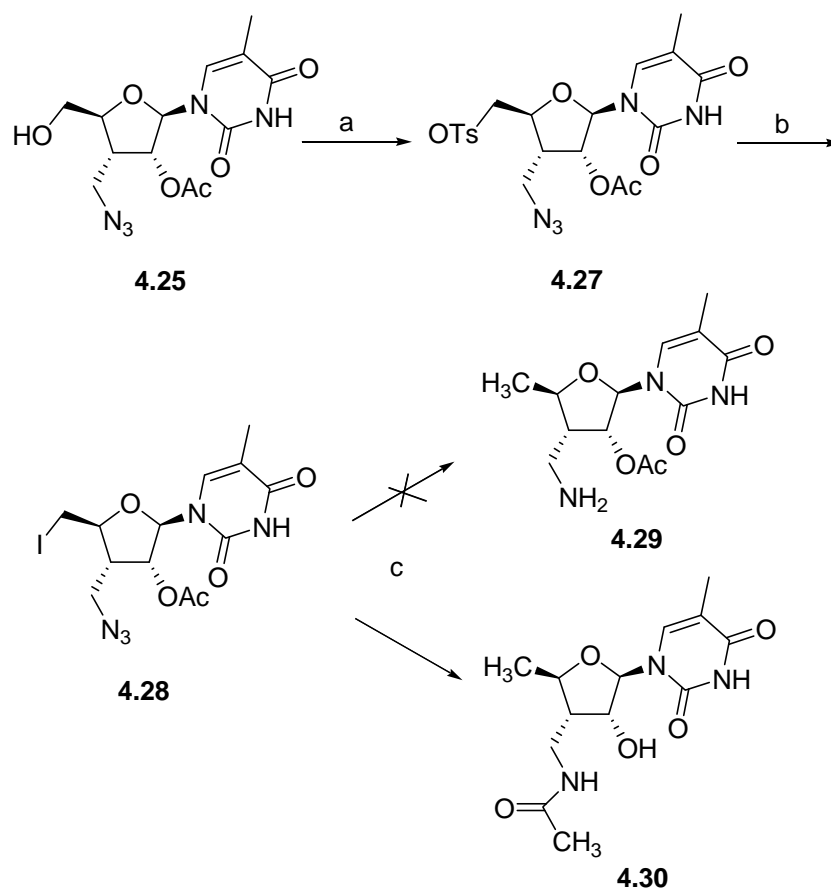
Scheme 4.5, Reagents and conditions: (a) *p*-toluenesulfonyl chloride, pyridine, rt, 2d, 87%; (b) NaI, acetone, 60 °C, 16h, 97%; (c) Pd/C, H_2, MeOH, TEA, rt, 3h, 98%.

Several attempts were made to reduce selectively the 5'-iodine without 6'-azidoreduction. The use of NaBH$_4$ resulted in 2'-deacetylation without reduction of 5'-iodine. The use of NaCNBH$_3$ or LiAlH$_3$ both remained unsuccesfull.

For this reason, the completely deprotected compound **4.26** was used. Selective tosylation of the primary alcohol of **4.26**,[142] followed by nucleophilic displacement of the tosylate with sodium iodide gave the 5'-iodo analogue **4.32**. Simultaneous reduction of the azide and the iodide using catalytic hydrogenation at atmospheric pressure led to compound **4.33**, which gave target compound **4.9** after ring closure.
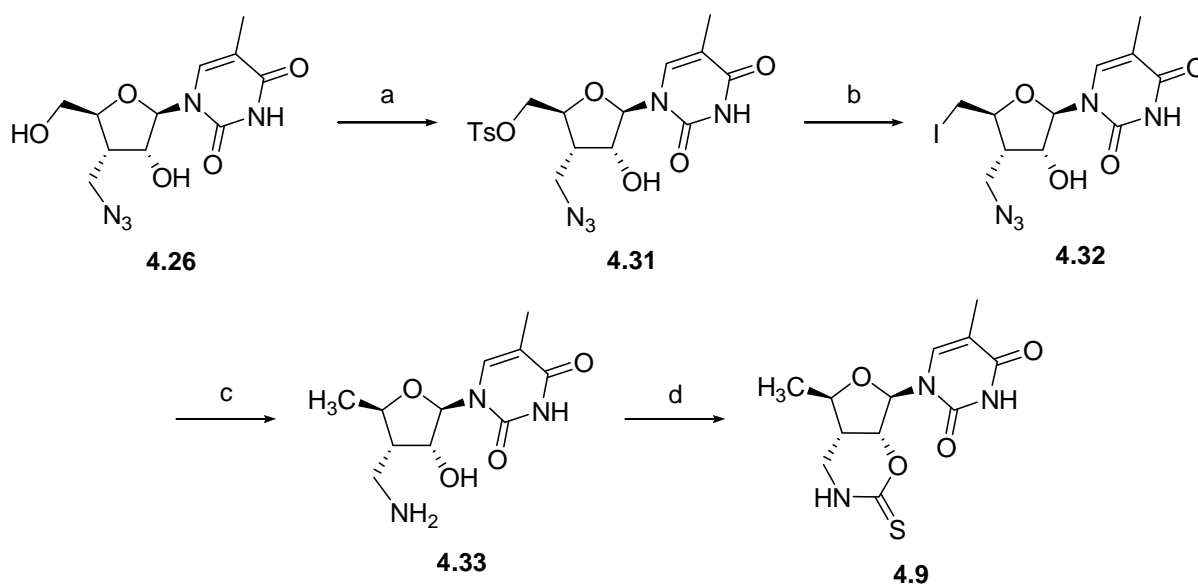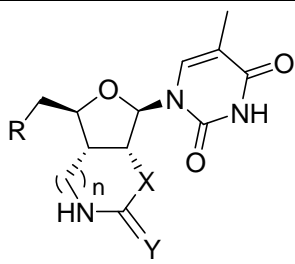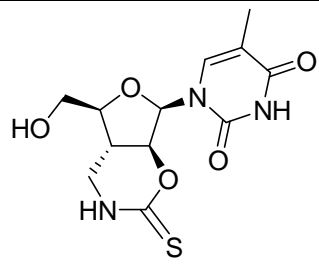
Scheme 4.6, Reagents and conditions: (a) *p*-toluenesulfonyl chloride, pyridine, rt, 2d, 46%; (b) NaI, acetone, TBAI, 60 °C, 16h, 100%; (c) Pd/C, H$_2$, MeOH, TEA, rt, 16h, 69%; (d) thiocarbonyldiimidazole, THF, rt, 5h, 43%.

4.3 BIOLOGICAL EVALUATION

All compounds were tested for TMPKmt inhibition as described in the experimental section of chapter 3. This series of bicyclic thymidine derivatives revealed several potent inhibitors. (See Table 4.1) The rationalisation of the observed structure-activity relationship involving molecular modelling and conformational analysis is described in the following part.

Table 4.1 Kinetic parameters of TMPKmt with bicyclic sugar derivatives.

Compound	X	Y	R	N	K_i (µM) TMPKmt	IC$_{99}$ *Myc. Bovis*
4.1	O	S	OH	1	3.5[106]	
4.2	O	O	OH	1	13.5[106]	
4.3	NH	S	OH	1	21	
4.4	NH	O	OH	1	59	
4.5	NH	NH	OH	1	46	
4.6	O	S	OH	0	8	500 µg/mL
4.7	O	O	OH	0	27	
4.8 (arabino)	O	S	OH	1	150	
4.9	O	S	H	1	2.3	100 µg/mL

The compounds with the highest inhibitory activity on the enzyme (i.e. **4.6** and **4.9**) were evaluated for their *in vitro* inhibitory activity against *Mycobacterium bovis* BCG. While compound **4.6** showed 100% inhibition of bacterial growth at 500 µg/mL, derivative **4.9** displayed a stronger inhibitory activity: no growth was observed at 100 µg/mL. No compound demonstrated toxicity against VERO cell lines in a 0-500 µg/mL concentration range.

Since compound **4.1** failed to show any inhibition of bacterial growth at 64 µg/mL, the activity of **4.9** represents a significant progress in our search for new anti-tuberculosis agents.

4.4 STRUCTURE-ACTIVITY RELATIONSHIP (S.A.R.) AND MODEL BUILDING

Compared to the bicyclic nucleosides **4.1** and **4.2**, isosteres **4.3** and **4.4**, with the 2' oxygen replaced by a 2'-NH, showed a significant drop in affinity. Modelling experiments indicate a repulsion between 2'-NH and the aromatic ring of Tyr 103.

The five-membered bicyclic derivatives **4.6** and **4.7** are 2-fold less active than the corresponding leads **4.1** and **4.2**. This result confirms earlier observations that a 3'-branching enhances the interaction with the enzyme.[107] Upon superimposition of the five- and six-membered bicyclic (thio)urethane derivatives, differences in binding TMPKmt

become perceptible, i.e. the H-bond between 3'-NH and Asp 9 is weaker in the five-membered ring derivative. In both series the distance between NH and the O-atoms of Asp 9 is very small (<3.5 Å), which is a requirement for a strong H-bonding. In the five-membered ring derivatives, however, the angle of the H-bond is too large. For a strong H-bond the angle between the N-H-bond and the O-H-bonds should be smaller than 35 degrees.[143] For compound **4.1** this angle is 23 degrees, for compound **4.6** 139 degrees, thus explaining the weakness of the H-bond in compounds **4.6** and **4.7** (Figure 4.3).
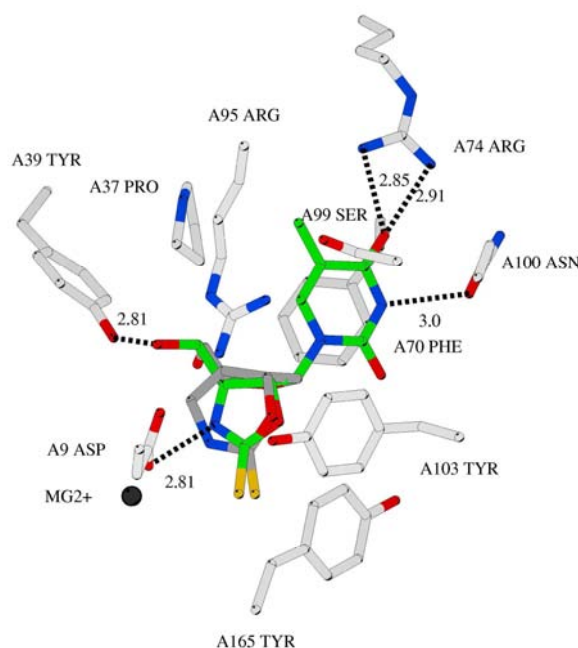
Figure 4.3 Superimposition of compounds **4.1** and **4.6** in their predicted binding modes to the enzyme

Comparison of all final compounds shows that the sulfur-derivatives are consistently more active than the corresponding oxygen-derivatives. Presumably, the large lipophylic sulfur atom gives a stronger hydrophobic interaction with the enzyme.

Compound **4.8** was synthesized in order to gain insight on the influence of the puckering of the sugar on binding affinity. Conformational analysis provides an explanation for the low binding affinity of this compound (K_i = 150 µM), as will be explained further.

Deletion of 5'-OH, the only hydroxyl function in compound **4.1,** strongly reduces polarity. As deoxygenation comes with the loss of a strong H-bonding between 5'-OH and Tyr 39, a lower binding affinity for this compound was expected. Surprisingly this 5'-deoxy derivative **4.9** was almost twice as active as compound **4.1** toward inhibition of TMPKmt. The same phenomenon was observed for the 5'-methyl derivative of dT accompanied by a 6-fold increase in binding affinity.[139] This higher activity can be explained by a higher hydrophobic

interaction between 5'-CH$_3$ and Pro 37 (See Figure 4.4 for a modelled binding mode of compound **4.9** and the LIGPLOT interaction map in Figure 4.5).
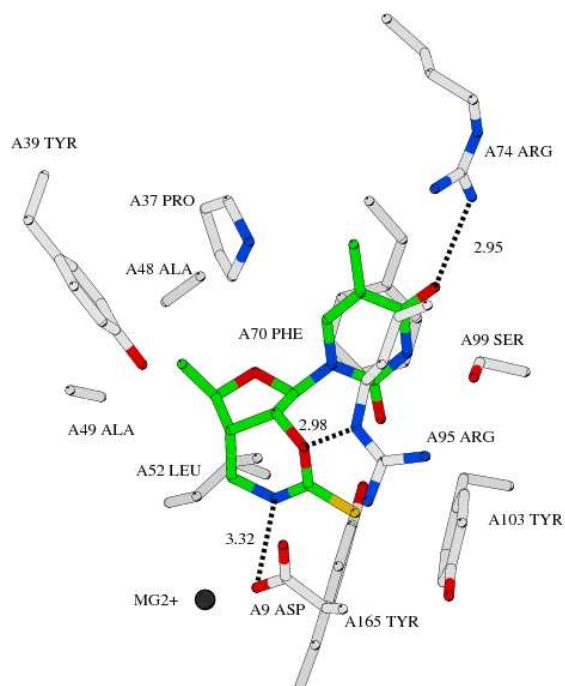
Figure 4.4 The modeled binding mode of compound **4.9** to TMPKmt.
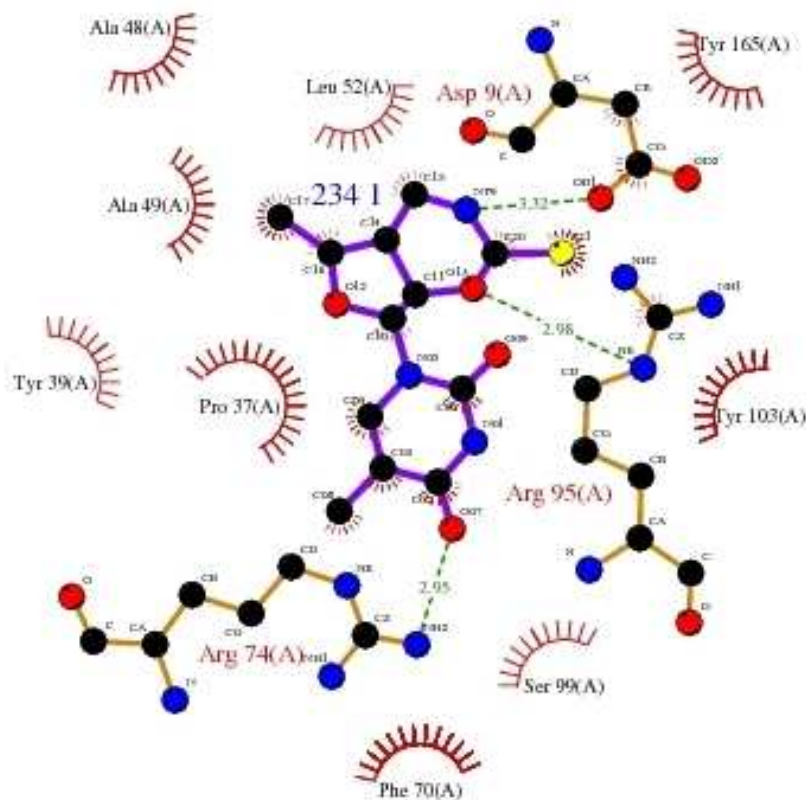
Figure 4.5 LIGPLOT interaction map of compound **4.9**.

4.5 CONFORMATIONAL ANALYSIS OF COMPOUNDS 4.6 AND 4.9

The solution state conformation of compounds **4.6** and **4.9** was analyzed using Pseurot 6.2.[144] Consequently, vicinal proton-proton coupling constants ($^3J_{1',2'}$, $^3J_{2',3'}$ and $^3J_{3',4'}$) of **4.6** and **4.9** were accurately measured at 500 MHz in D_2O using the Selective Refocussing (SERF) technique[145] at five different temperatures ranging from 293 K to 325 K. In the case of compound **4.6**, the *J* values were virtually unaffected within the temperature interval studied. This observation can be explained by assuming that the sugar adopts one peculiar conformation independent of temperature. Alternatively, an equilibrium might exist between two conformations, which are already present in their T_∞ ratio at the lowest temperature, indicating a small conformational energy barrier. The latter scenario proved the only satisfactory one with the following conformations found: one with P equal to 77° and a ν_{max} of about 18° (O4' endo), the other with P of − 88° and a ν_{max} of about 17° (O4' exo). These results extend a previously reported conformational analysis of the thymine congener of **4.6** using Macromodel potential energy calculations, which only pointed towards a single O4'-endo-C4'-exo puckering.[141]

Pseurot analysis of **4.9** reveals a predominant sugar puckering with P value of 40° and a ν_{max} of 39°, corresponding to a puckering between C4' ex o and C3' endo. Although this puckering is very similar to the one deduced from modelling of compound **4.8**, the inverted configuration at C2' probably excludes a favourable interaction of the fused ring of **4.8** with TMPKmt.

Since the X-ray analysis of the dTMP-TMPKmt complex indicates that the substrate adopts a Southern conformation with the base in an equatorial orientation,[95] one would expect that nucleoside analogues exhibiting a similar S-type conformation would have an optimal affinity for TMPKmt.[110] In the present case, however, the extra ring fused to the sugar part via 2' and 3' seems to exclude such conformation. The established rigidification in compound **4.9** seems to contribute to its enhanced affinity for TMPKmt by allowing favourable interactions with Asp9 and Arg95.

4.6 CONCLUSIONS

This chapter describes the synthesis, the biological and conformational analysis of a series of bicyclic sugar nucleosides as inhibitors of TMPKmt. Most compounds showed an inhibitory potency in the low micromolar range, with compound **4.9** being the most active one (K_i = 2.3 μM). With an IC_{99} of 100 μg/mL, this compound gives a strong indication to be capable of inhibiting M. tuberculosis growth, promoting TMPKmt as an attractive target for further inhibitor design.

4.7 EXPERIMENTAL PART

3'-Azidomethyl-5'-*O*-benzyl-β-D-thymidine (4.10)

Compound **2.35** (810 mg, 1.89 mmol) was dissolved in 7N NH_3 in MeOH (20 mL) and was stirred at room temperature during 4 hours. The reaction mixture was evaporated, suspended in H_2O (20 mL) and extracted with CH_2Cl_2 (3 x 25 mL). The organic layer was dried and evaporated to dryness. The residue was purified by column chromatography (CH_2Cl_2/MeOH 95/5) to afford compound **4.10** (700 mg, 95%).
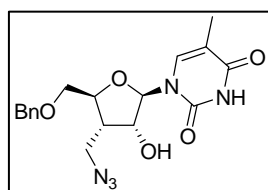
^{1}H NMR (300 MHz, DMSO-d_6): δ 1.41 (3H, d, 5-CH_3), 2.44 (1H, m, H-3'), 3.64 (2H, m, H-6' and H-6"), 3.86 (1H, dd, *J*= 2.4 and 11.1 Hz, H-5'), 4.08 (1H, m, H-5"), 4.20 (1H, m, H-4'), 4.56 (2H, s, *CH₂*Ph), 5.65 (1H, d, *J*= 2.1, H-2'), 5.91 (1H, d, 4.8 Hz, H-1'), 7.31 (5H, m, CH₂*Ph*), 7.64 (1H, d, *J*= 1.2 Hz, H-6).

Exact mass (ESI-MS) for $C_{18}H_{21}N_5O_5Na$ [M+Na]$^+$ found, 410.1445; calcd, 410.1441.

2,2'-*O*-Anhydro-1-(3-azidomethyl-3-deoxy-5-*O*-benzyl-β-D-arabinofuranosyl)thymine (4.11)

A solution of **4.10** (3.00 g, 7.74 mmol) and DMAP (3.78 g, 30.96 mmol) in dichloromethane (77 mL) was stirred at room temperature. After 30 minutes the mixture was cooled to 0 °C and trifluoromethanesulfonyl chloride (1.65 mL, 15.50 mmol) was added. After stirring for 2 hours at room temperature, the reaction mixture was quenched with water. The mixture was partioned between CH_2Cl_2 and a saturated $NaHCO_3$-solution. The aqueous layer was washed with CH_2Cl_2 and the combined organic layers were dried over $MgSO_4$, filtered and evaporated to dryness. The residue was purified by column chromatography (CH_2Cl_2/MeOH 98:2), affording **4.11** (2.86 g, 93%) as a white foam.
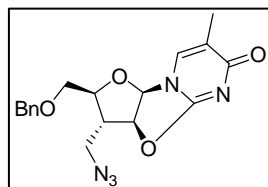
^{1}H NMR (300 MHz, DMSO-d_6): δ 1.79 (3H, d, *J*= 1.2 Hz, 5-CH_3), 2.66 (1H, m, H-3'), 3.19 (1H, dd, *J*= 6.6 and 10.2 Hz, H-5'), 3.35 (1H, dd, *J*= 3.9 and 10.5 Hz, H-5"), 3.63 (2H, m, H-6'/H-6"), 4.21 (1H, m, H-4'), 4.31 (2H, s, *CH₂*Ph), 5.28 (1H, dd, *J*= 2.1 and 5.4 Hz, H-2'), 6.20 (1H, d, *J*= 6.0 Hz, H-1'), 7.26 (5H, m, CH₂*Ph*), 7.74 (1H, d, *J*= 1.5 Hz, H-6).

Exact mass (ESI-MS) for $C_{18}H_{20}N_5O_4$ [M+H]$^+$ found, 370.1512; calcd, 370.1515.

1-(2-Azido-3-azidomethyl-2,3-dideoxy-5-*O*-benzyl-β-D-ribofuranosyl)thymine (4.12)

A solution of **4.11** (2.46 g, 6.66 mmol), NaN$_3$ (2.16 g, 33 mmol) and benzoic acid (0.813 g, 6.66 mmol) in DMF (40 mL) was heated to 150 °C. The reaction mixture was refluxed for 7 h and evaporated to dryness. The residue was partitioned between water and CH$_2$Cl$_2$. The organic layer was dried over MgSO$_4$, filtered and evaporated. Further purification by column chromatography (CH$_2$Cl$_2$/MeOH 99:1) yielded **4.12** (2.01 g, 74%).
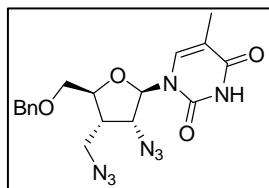
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.40 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 2.63 (1H, m, H-3'), 3.43 (1H, dd, *J*= 6.6 and 12.6 Hz, H-6'), 3.60 (2H, m, H-6" and H-5'), 3.88 (1H, dd, *J*= 2.1 and 11.4 Hz, H-5"), 4.01 (1H, m, H-4'), 4.57 (2H, s, *CH$_2$*Ph), 4.55 (1H, d, *J*= 1.5 Hz, H-2'), 5.79 (1H, d, *J*= 1.8 Hz, H-1'), 7.32 (5H, m, CH$_2$*Ph*), 7.62 (1H, d, *J*= 0.9 Hz, H-6), 11.34 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{18}$H$_{20}$N$_8$O$_4$Na [M+Na]$^+$ found, 435.1508; calcd, 435.1505.

1-(2-Amino-3-aminomethyl-2,3-dideoxy-5-*O*-benzyl-β-D-ribofuranosyl)thymine (4.13)

A solution of **4.12** (200 mg, 0.48 mmol) in methanol (12 mL) was hydrogenated at atmospheric pressure for 5 h in the presence of 10% Pd/C (20 mg). The catalyst was removed by filtration through a celite path and the filtrate was evaporated. The residue was purified by column chromatography (CH$_2$Cl$_2$/MeOH 85:15) to give compound **4.13** (155 mg, 90%).
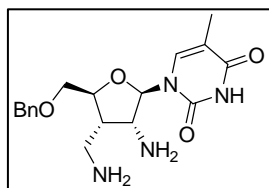
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.45 (3H, s, 5-CH$_3$), 2.14 (1H, m, H-3'), 2.56 (1H, dd, *J*= 6.0 and 12.6 Hz, H-6'), 3.71 (1H, dd, *J*= 7.5 and 12.3 Hz, H-6"), 3.41 (1H, dd, *J*= 3.6 and 6.9 Hz, H-2'), 3.56 (1H, dd, *J*= 3.6 and 11.1 Hz, H-5'), 3.79 (1H, dd, *J*= 2.4 and 11.4 Hz, H-5"), 4.07 (1H, m, H-4'), 4.55 (2H, s, *CH$_2$*Ph), 5.55 (1H, d, *J*= 3.6 Hz, H-1'), 7.32 (5H, m, CH$_2$*Ph*), 7.64 (1H, d, *J*= 1.2 Hz, H-6).

Exact mass (ESI-MS) for C$_{18}$H$_{25}$N$_4$O$_4$ [M+H]$^+$ found, 361.1871; calcd, 361.1875.

1-[2-Amino-3-aminomethyl-2,3-dideoxy-2-*N*,6-*N*-(thiocarbonyl)-5-*O*-benzyl-β-D-ribofuranosyl]thymine (4.14)

A solution of **4.13** (200 mg, 0.56 mmol) and 1,1'-thiocarbonyldiimidazole (109 mg, 0.61 mmol) in THF (6 mL) was stirred at room temperature for 1.5 h. The mixture was evaporated to dryness and purified by column chromatography (CH$_2$Cl$_2$/MeOH 95:5) to give compound **4.14** (135 mg) in 60 % yield.
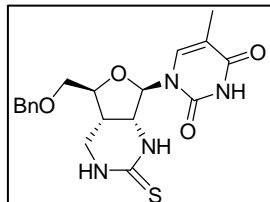
¹H NMR (300 MHz, DMSO-d$_6$): δ 1.50 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 2.48 (1H, m, H-3'), 3.06 (1H, dd, *J*= 9.3 and 12.0 Hz, H-6'), 3.32 (1H, m, H-6"), 3.63 (1H, dd, *J*= 3.3 and 10.5 Hz, H-5'), 3.75 (1H, dd, *J*= 3.0 Hz and 11.1 Hz, H-5"), 3.86 (1H, m, H-2'), 4.03 (1H, dd, *J*= 3.0 and 6.3 Hz, H-4'), 4.56 (2H, s, *CH$_2$*Ph), 5.79 (1H, d, *J*= 6.3 Hz, H-1'), 7.33 (5H, m, CH$_2$*Ph*), 7.47 (1H, d, *J*= 1.2 Hz, H-6), 8.33 (2H, m, N(2')H and N(6')H), 11.27 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{19}$H$_{23}$N$_4$O$_4$S [M+H]$^+$ found, 403.1447; calcd, 403.1439.

1-[2-amino-3-aminomethyl-2,3-dideoxy-2-*N*,6-*N*-(carbonyl)-5-*O*-benzyl-β-D-ribofuranosyl]thymine (4.15)

A solution of **4.13** (150 mg, 0.41 mmol) and 1,1'-carbonyldiimidazole (73 mg, 0.45 mmol) in THF (4 mL) was stirred at room temperature for 3 h. The mixture was evaporated to dryness and purified by column chromatography (CH$_2$Cl$_2$/MeOH 95:5) to give compound **4.15** (95 mg) in 60 % yield.

¹H NMR (300 MHz, DMSO-d$_6$): δ 1.50 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 2.52 (1H, m, H-3'), 3.07 (1H, m, H-6'), 3.21 (1H, m, H-6"), 3.63 (1H, dd, *J*= 3.3 and 10.5 Hz, H-5'), 3.75 (1H, dd, *J*= 3.0 Hz and 11.1 Hz, H-5"), 3.86 (1H, m, H-2'), 4.06 (1H, m, H-4'), 4.56 (2H, s, *CH$_2$*Ph), 5.74 (1H, d, *J*= 5.4 Hz, H-1'), 6.45 (1H, br s, N(2')H), 6.55 (1H, br s, N(6')H), 7.33 (5H, m, CH$_2$*Ph*), 7.51 (1H, d, *J*= 1.2 Hz, H-6), 11.29 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{19}$H$_{22}$N$_4$O$_5$Na [M+Na]$^+$ found, 409.1488; calcd, 409.1498.

1-[2-Amino-3-aminomethyl-2,3-dideoxy-2-*N*,6-*N*-(thiocarbonyl)-β-D-ribofuranosyl]thymine (4.3) and 1-[2-amino-3-aminomethyl-2,3-dideoxy-2-*N*,6-*N*-(carbonyl)-β-D-ribofuranosyl]thymine (4.4)

To a solution of **4.14** or **4.15** (1 eq) in dry CH_2Cl_2 (40 mL/mmol), a 1 M BCl_3 solution in CH_2Cl_2 was added (4 eq) at -78 °C. After 15 minutes the reaction was quenched with methanol, the mixture was evaporated and the residue was purified by column chromatography (CH_2Cl_2/MeOH 95:5) to yield final compounds **4.3** (47%) and **4.4** (60%).
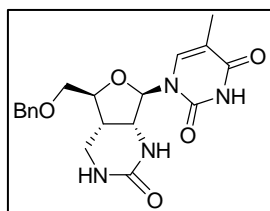

¹H NMR (300 MHz, DMSO-d₆): δ 1.75 (3H, d, *J*= 1.2 Hz, 5-CH₃), 2.43 (1H, m, H-3'), 3.04 (1H, dd, *J*= 9.6 and 12.3 Hz, H-6'), 3.25 (1H, m, H-6"), 3.60 (2H, m, H-5' and H-5"), 3.85 (2H, m, H-4' and H-2'), 5.18 (1H, t, *J*= 5.1 Hz, 5'-OH), 5.76 (1H, d, *J*= 6.3 Hz, H-1'), 7.71 (1H, d, *J*= 1.2 Hz, H-6), 8.31 (2H, br s, N(6')H and N(2')H), 11.27 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.94 (5-CH₃), 33.14 (C-3'), 37.58 (C-6'), 57.28 (C-5'), 62.71 (C-2'), 82.50 (C-4'), 89.08 (C-1'), 110.12 (C-5), 137.04 (C-6), 151.40 (C-4), 164.64 (C-2), 177.18 (C=S).

Exact mass (ESI-MS) for $C_{12}H_{17}N_4O_4S$ [M+H]⁺ found, 313.0986; calcd, 313.0970; Anal. ($C_{12}H_{16}N_4O_4S$) C, H, N, S.
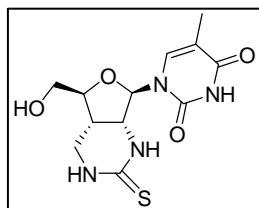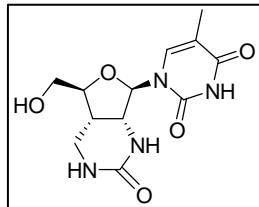

¹H NMR (300 MHz, DMSO-d₆): δ 1.75 (3H, d, *J*= 0.9 Hz, 5-CH₃), 2.45 (1H, m, H-3'), 3.05 (1H, m, H-6'), 3.19 (1H, m, H-6"), 3.54 (1H, m, H-5'), 3.65 (1H, m, H-5"), 3.85 (2H, m, H-4' and H-2'), 5.18 (1H, t, *J*= 5.1 Hz, 5'-OH), 5.71 (1H, d, *J*= 5.7 Hz, H-1'), 6.42 (1H, br s, N(6')H), 6.55 (1H, br s, N(2')H), 7.74 (1H, d, *J*= 0.9 Hz, H-6), 11.29 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.93 (5-CH₃), 34.96 (C-3'), hidden by DMSO-signal (C-6'), 57.51 (C-5'), 62.57 (C-2'), 82.22 (C-4'), 90.02 (C-1'), 109.90 (C-5), 137.02 (C-6), 151.39 (C-4), 156.49 (C=O), 164.56 (C-2).

Exact mass (ESI-MS) for $C_{12}H_{16}N_4O_5Na$ [M+Na]⁺ found, 319.1015; calcd, 319.1018; Anal. ($C_{12}H_{16}N_4O_5$) C, H, N.

1-[2-Amino-3-aminomethyl-2,3-dideoxy-2-*N*,6-*N*-(methylthiomethylidyne)-5-*O*-benzyl-β-D-ribofuranosyl]thymine (4.16)

To a solution of compound **4.14** (778 mg, 1.93 mmol) in methanol (60 mL), MeI (168 μL, 2.71 mmol) was added and the solution was stirred for 8 h. The reaction mixture was evaporated to dryness and the residue was purified by column chromatography (CH₂Cl₂/MeOH 98:2) yielding **4.16** (630 mg, 78 %).
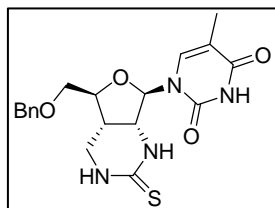
¹H NMR (300 MHz, DMSO-d₆): δ 1.46 (3H, s, 5-CH₃), 2.61 (1H, m, H-3'), 3.18 (1H, m, H-6'), 3.41 (1H, dd, *J*= 4.8 and 12.9 Hz, H-6"), 3.66 (3H, s, S-CH₃), 3.79 (1H, d, *J*= 2.4 Hz, H-5'), 3.83 (1H, d, *J*=2.7 Hz, H-5"), 3.96 (1H, m, H-2'), 4.06 (1H, m, H-4'), 4.58 (2H, s, *CH₂*Ph), 5.74 and 5.81 (1H, 2 tautomers, d, *J*= 2.7 and 4.8 Hz, H-1'), 7.32 (5H, m, CH₂*Ph*), 7.61 and 7.65 (1H, 2 tautomers, s, H-6), 11.32 and 11.36 (1H, 2 tautomers, s, N(3)H).

Exact mass (ESI-MS) for C₂₀H₂₅N₄O₄S [M+H]⁺ found, 417.1589; calcd, 417.1596.

1-[2-Amino-3-aminomethyl-2,3-dideoxy-2-*N*,6-*N*-(imine)-5-*O*-benzyl-β-D-ribofuranosyl]thymine (4.17)

Compound **4.16** (189 mg, 0.45 mmol) was dissolved in 7N NH₃ in methanol (20 mL). The solution was allowed to stir overnight at 100 °C. The reaction mixture was evaporated to dryness. The residue was purified by column chromatography (CH₂Cl₂/MeOH 95:5), affording **4.17** (86 mg, 50 %).
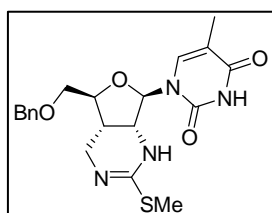
¹H NMR (300 MHz, DMSO-d₆): δ 1.47 (3H, d, *J*= 1.2 Hz, 5-CH₃), 2.71 (1H, m, H-3'), 3.22 (1H, dd, *J*= 7.5 and 12.9 Hz, H-6'), 3.43 (1H, dd, *J*= 5.4 and 13.2 Hz, H-6"), 3.67 (1H, dd, *J*= 3.0 and 11.1 Hz, H-5'), 3.82 (1H, dd, *J*= 2.7 and 10.8 Hz, H-5"), 4.04 (1H, t, *J*= 5.1 Hz, H-2'), 4.12 (1H, m, H-4'), 4.58 (2H, s, *CH₂*Ph), 5.77 (1H, d, *J*= 4.5 Hz, H-1'), 7.04 (1H, br s, NH), 7.32 (5H, m, CH₂*Ph*), 7.57 (1H, d, *J*= 1.2 Hz, H-6), 7.99 (1H, br s, N(6')H), 8.24 (1H, br s, N(2')H), 11.40 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₂₀H₂₄N₄O₄ [M+H]⁺ found, 386.1817; calcd, 386.1953.

1-[2-Amino-3-aminomethyl-2,3-dideoxy-2-*N*,6-*N*-(imine)-β-D-ribofuranosyl]thymine (4.5)

The title compound was synthesized from **4.17** (72 mg, 0.19 mmol) using the same procedure as described for the deprotection of **4.14**, in a yield of 21 mg or 38 %.
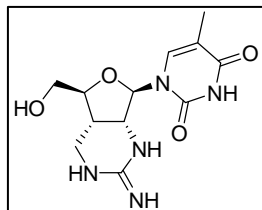
¹H NMR (300 MHz, DMSO-d_6): δ 1.76 (3H, s, 5-CH₃), 2.65 (1H, m, H-3'), 3.22 (1H, dd, *J*= 6.9 and 13.2 Hz, H-6'), 3.41 (1H, dd, *J*= 5.1 and 13.5 Hz, H-6"), 3.61 (1H, d, *J*= 11.7 Hz, H-5'), 3.72 (1H, d, *J*= 11.7 Hz, H-5"), 3.93 (1H, m, H-4'), 4.03 (1H, m, H-2'), 5.41 (1H, s, 5'-OH), 5.74 (1H, d, *J*= 3.6 Hz, H-1'), 7.23 (1H, br s, C=NH), 7.75 (2H, br s, N(6')H and N(2')H), 7.88 (1H, s, H-6).

¹³C NMR (75 MHz, DMSO-d_6): δ 12.97 (5-CH₃), 33.19 (C-6'), 36.89 (C-3'), 56.42 (C-2'), 61.50 (C-5'), 82.58 (C-4'), 89.46 (C-1'), 109.76 (C-5), 136.58 (C-6), 151.29 (C-4), 154.31 (C=NH), 164.52 (C-2).

Exact mass (ESI-MS) for $C_{13}H_{18}N_4O_4$ [M+H]⁺ found, 296.1366; calcd, 296.1358; Anal. ($C_{13}H_{17}N_4O_4 \cdot H_2O$) C, H, N.

3'-Azido-5'-benzyl-β-D-thymidine (4.18)

Compound **3.11** (3.02 g, 7.27 mmol) was dissolved in 1N NH₃ in methanol (70 mL) and stirred at room temperature during 3 hours. The solvent was removed and the resulting crude residue was dissolved in CH₂Cl₂ (50 mL) and extracted with water (50 mL). The water layer was washed three times with CH₂Cl₂ (50 mL). Further workup of the organic layers (drying and evaporation) resulted in 2.66 g pure compound **4.18** (98%).
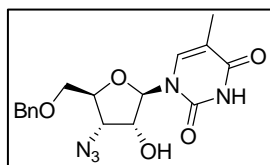
¹H NMR (300 MHz, DMSO-d_6): δ 1.48 (3H, d, *J*= 1.0 Hz, 5-CH₃), 3.63 (1H, dd, *J*= 3.1 and 10.8 Hz, H-5'), 3.73 (1H, dd, *J*= 2.8 and 10.8 Hz, H-5"), 4.02 (1H, m, H-4'), 4.21 (1H, app t, *J*= 4.8 Hz, H-3'), 4.43 (1H, m, H-2'), 4.56 (2H, d, *J*= 2.1 Hz, *CH₂*Ph), 5.77 (1H, d, *J*= 6.0 Hz, H-1'), 6.14 (1H, d, J= 4.8 Hz, 2'-OH), 7.33 (5H, m, CH₂*Ph*), 7.47 (1H, d, *J*=1.2 Hz, H-6), 11.32 (1H, s, N(3)H).

Exact mass (ESI-MS) for $C_{17}H_{20}N_5O_5$ [M+H]⁺ found, 374.1456; calcd, 374.1465.

3'-Amino-5'-benzyl-β-D-thymidine (4.19)

A solution of **4.18** (852 mg, 2.28 mmol) and 10% Pd/C (85 mg) in methanol was hydrogenated under atmospheric pressure during 3 hours. The Pd/C was removed from the solution by filtration over celite. After evaporation, 826 mg (96%) of compound **4.19** was obtained.
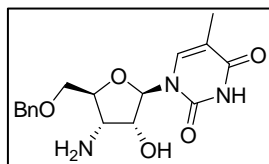
¹H NMR (300 MHz, DMSO-d₆): δ 1.43 (3H, d, *J*= 1.0 Hz, 5-CH₃), 3.27 (1H, dd, *J*= 5.5 and 7.7 Hz, H-3'), 3.64 (1H, dd, *J*= 3.4 and 10.8 Hz, H-5'), 3.73 (1H, ddd, *J*= 2.1, 3.3 and 7.8 Hz, H-4'), 3.80 (1H, dd, *J*= 1.9 and 10.9 Hz, H-5"), 3.86 (1H, dd, *J*= 2.5 and 5.1 Hz, H-2'), 4.55 (2H, s, *CH₂*Ph), 5.69 (1H, d, *J*= 2.7 Hz, H-1'), 7.33 (5H, m, CH₂*Ph*), 7.59 (1H, d, *J*=1.2 Hz, H-6).

Exact mass (ESI-MS) for C₁₇H₂₂N₃O₅ [M+H]⁺ found, 348.1559; calcd, 348.1559.

1-(3-Amino-3-deoxy-2-*O*,3-*N*-(thiocarbonyl)-5-*O*-benzyl-β-D-ribofuranosyl)thymine (4.20)

Compound **4.19** (410 mg, 1.10 mmol) and 1,1'-thiocarbonyldiimidazole (215 mg, 1.21 mmol) in THF (18 mL) were stirred at 50 °C overnight. The reaction mixture was evaporated to dryness, and the residue was purified by column chromatography (CH₂Cl₂/MeOH 97:3) to yield **4.20** (228 mg, 50%).
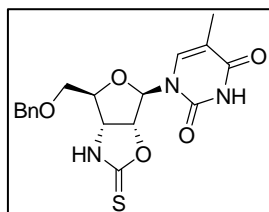

¹H NMR (300 MHz, DMSO-d₆): δ 1.65 (3H, d, *J*= 1.2 Hz, 5-CH₃), 3.59 (1H, dd, *J*= 5.4 and 10.8 Hz, H-5'), 3.67 (1H, dd, *J*= 3.9 and 11.1 Hz, H-5"), 4.21 (1H, m, H-4'), 4.50 (2H, d, *CH₂*Ph), 4.54 (1H, d, *J*= 3.9 Hz, H-3'), 5.59 (1H, dd, *J*= 3.0 and 8.7 Hz, H-2'), 5.92 (1H, d, *J*= 2.7, H-1'), 7.31 (5H, m, CH₂*Ph*), 7.51 (1H, d, *J*= 1.2 Hz, H-6), 10.56 (1H, br s, N(3')H), 11.39 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₈H₂₀N₃O₅S [M+H]⁺ found, 390.1123; calcd, 390.1123.

1-(3-Amino-3-deoxy-2-*O*,3-*N*-(carbonyl)-5-*O*-benzyl-β-D-ribofuranosyl)thymine (4.21)

Compound **4.19** (325 mg, 0.87 mmol) and 1,1'-carbonyldiimidazole (156 mg, 0.96 mmol) in THF (14 mL) were stirred at 50 °C overnight. The reaction mixture was evaporated to dryness, and the residue was purified by column chromatography (CH₂Cl₂/MeOH 97:3) to yield **4.21** (250 mg, 72%).
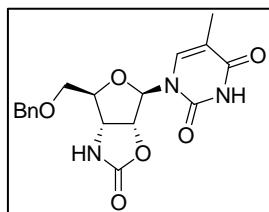
¹H NMR (300 MHz, DMSO-d₆): δ 1.65 (3H, d, *J*= 1.2 Hz, 5-CH₃), 3.59 (1H, dd, *J*= 5.4 and 10.8 Hz, H-5'), 3.67 (1H, dd, *J*= 3.6 and 10.8 Hz, H-5"), 4.13 (1H, dd, *J*= 5.4 and 9.3 Hz, H-4'), 4.25 (1H, dd, *J*= 4.8 and 8.4 Hz, H-3'), 4.51 (2H, d, *CH₂*Ph), 5.20 (1H, dd, *J*= 3.3 and 8.7 Hz, H-2'), 5.90 (1H, d, *J*= 3.3, H-1'), 7.32 (5H, m, CH₂*Ph*), 7.49 (1H, d, *J*= 1.2 Hz, H-6), 8.27 (1H, br s, N(3')H), 11.42 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₈H₂₀N₃O₆ [M+H]⁺ found, 374.1354; calcd, 374.1351.

1-(3-Amino-3-deoxy-2-*O*,3-*N*-(thiocarbonyl)-β-D-ribofuranosyl)thymine (4.6) and 1-(3-Amino-3-deoxy-2-*O*,3-*N*-(carbonyl)-β-D-ribofuranosyl)thymine (4.7)

These compounds were synthesized from **4.20** (250 mg, 0.67 mmol) and **4.21** (215 mg, 0.52 mmol) using the same procedure as described for the deprotection of **4.14**, to yield **4.6** (103 mg, 70 %) and **4.7** (140 mg, 70%).
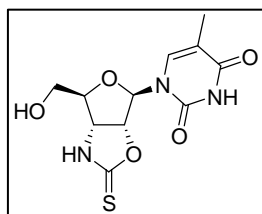

¹H NMR (300 MHz, DMSO-d₆): δ 1.74 (3H, d, *J*= 0.9 Hz, 5-CH₃), 3.54 (2H, m, H-5' and H-5"), 3.97 (1H, dd, *J*= 4.5 and 9.3 Hz, H-4'), 4.45 (1H, dd, *J*= 4.5 and 8.7 Hz, H-3'), 5.09 (1H, t, *J*= 5.1 Hz, 5'-OH), 5.35 (1H, dd, *J*= 3.0 and 8.7 Hz, H-2'), 5.90 (1H, d, *J*= 3.0, H-1'), 7.58 (1H, d, *J*= 0.9 Hz, H-6).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.75 (5-CH₃), 56.60 (C-3'), 61.71 (C-5'), 87.96 (C-4'), 88.85 (C-2'), 91.87 (C-1'), 110.34 (C-5), 138.25 (C-6), 150.99 (C-4), 164.51 (C-2), 187.56 (C=S).

Exact mass (ESI-MS) for $C_{11}H_{14}N_3O_5S$ [M+H]⁺ found, 300.0659; calcd, 300.0654; Anal. ($C_{11}H_{13}N_3O_5S$) C, H, N, S.
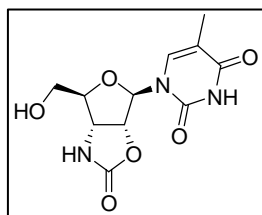
¹H NMR (300 MHz, DMSO-d₆): δ 1.74 (3H, d, *J*= 0.6 Hz, 5-CH₃), 3.56 (2H, m, H-5' and H-5"), 3.90 (1H, dd, *J*= 4.5 and 9.3 Hz, H-4'), 4.18 (1H, dd, *J*= 4.8 and 8.7 Hz, H-3'), 5.08 (1H, s, 5'-OH), 5.15 (1H, dd, *J*= 3.6 and 8.7 Hz, H-2'), 5.87 (1H, d, *J*= 3.3, H-1'), 7.56 (1H, d, *J*= 0.9 Hz, H-6), 8.24 (1H, br s, N(3')H), 11.42 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.79 (5-CH₃), 56.85 (C-3'), 61.69 (C-5'), 82.24 (C-4'), 88.08 (C-2'), 91.27 (C-1'), 110.41 (C-5), 138.06 (C-6), 151.01 (C-4), 157.90 (C=O), 164.48 (C-2).

Exact mass (ESI-MS) for $C_{11}H_{13}N_3O_6Na$ [M+Na]⁺ found, 306.0709; calcd, 306.0702; Anal. ($C_{11}H_{13}N_3O_6$) C, H, N.

1-(3-Azidomethyl-3-deoxy-5-*O*-benzyl-β-D-arabinofuranosyl)thymine (4.22)

A solution of **4.11** (530 mg, 1.43 mmol), 1 M NaOH (4.24 mL), dioxane (58 mL) and EtOH/H₂O 1:1 (58 mL) was allowed to stir for 3 h at room temperature. The resulting mixture was extracted with CH₂Cl₂ (3 x 100 mL). The organic layer was dried over MgSO₄, filtered and evaporated, and the crude material was purified by column chromatography (CH₂Cl₂/MeOH 98:2) to yield **4.22** (509 mg, 92%).
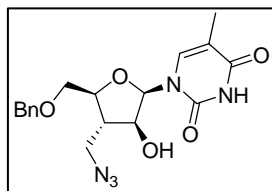
¹H NMR (300 MHz, DMSO-d₆): δ 1.46 (3H, d, *J*= 0.9 Hz, 5-CH₃), 2.27 (1H, m, H-3'), 3.48 (1H, dd, *J*= 6.9 and 12.6 Hz, H-6'), 3.66 (2H, m, H-5' and H-6''), 3.82 (2H, m, H-4' and H-5''), 4.16 (1H, dd, *J*= 6.0 and 12.6 Hz, H-2'), 4.55 (2H, s, *CH₂*Ph), 5.62 (1H, d, *J*= 5.4 Hz, 2'-OH), 5.95 (1H, d, *J*= 5.7 Hz, H-1'), 7.33 (5H, m, CH₂*Ph*), 7.47 (1H, d, *J*= 1.2 Hz, H-6), 11.22 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₈H₂₁N₄O₅Na [M+Na]⁺ found, 435.1508; calcd, 435.1505.

1-(3-Aminomethyl-3-deoxy-5-*O*-benzyl-β-D-arabinofuranosyl)thymine (4.23)

A solution of **4.22** (195 mg, 0.50 mmol) in methanol (12 mL) was hydrogenated at

atmospheric pressure for 5 hours in the presence of 10 % Pd/C (20 mg). The catalyst was removed by filtration through a celite path and the filtrate was evaporated. The residue was used in the next step without further purification.
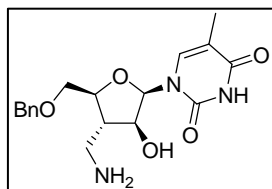
1-(3-Aminomethyl-3-deoxy-2-*O*,6-*N*-(thiocarbonyl)-5-*O*-benzyl-β-D-arabinofuranosyl)thymine (4.24)

Compound **4.23** (43 mg, 0.12 mmol) was dissolved in THF (6 mL), and the resulting solution was slowly added to a solution of 1,1'-thiocarbonyldiimidazole (64 mg, 0.36 mmol) in THF (2 mL). After 1 hour, 5 mL THF was added, and the solution was allowed to stir another 2 h. Evaporation of the solvent and purification of the residue by column chromatography (CH₂Cl₂/MeOH 97:3) gave **4.24** (20 mg, 43%).
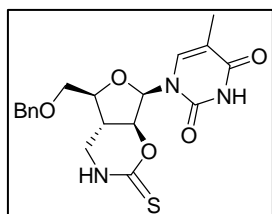
¹H NMR (300 MHz, DMSO-d₆): δ 1.43 (3H, d, *J*= 0.6 Hz, 5-CH₃), 2.40 (1H, m, H-3'), 3.70 (1H, dd, *J*= 3.9 and 11.3 Hz, H-6'), 3.81 (1H, dd, *J*= 3.3 and 11.1 Hz, H-6''), 3.88 (3H, m, H-4', H-5' and H-5''), 4.19 (1H, m, H-2'), 4.55 (2H, s, *CH₂*Ph), 5.98 (1H, d, *J*= 6.3 Hz, H-1'), 7.32 (5H, m, CH₂*Ph*), 7.45 (1H, d, *J*= 1.2 Hz, H-6), 11.23 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₉H₂₁N₃O₅SNa [M+Na]⁺ found, 426.1093; calcd, 426.1099.

1-(3-Aminomethyl-3-deoxy-2-*O*,6-*N*-(thiocarbonyl)-β-D-arabinofuranosyl)thymine (4.8)

This compound was synthesized from **4.24** using the same procedure as described for the deprotection of **4.14**, in a yield of 38 % (15 mg).
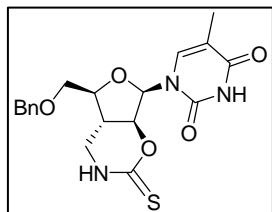
¹H NMR (300 MHz, DMSO-d₆): δ 1.72 (3H, d, *J*= 0.9 Hz, 5-CH₃), 2.99 (1H, m, H-3'), 3.59 (1H, m, H-6'), 2.67 (2H, m, H-5' and H-6''), 3.88 (2H, m, H-4' and H-5''), 4.17 (1H, m, H-2'), 5.18 (1H, app s, 5'-OH), 5.66 (1H, d, *J*= 6.0 Hz, H-1'), 7.67 (1H, d, *J*= 0.9 Hz, H-6), 11.22 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.99 (5-CH₃), 31.38 (C-3'), 33.31 (C-6'), 60.04 (C-5'), 79.82 (C-4'), 80.14 (C-2'), 81.07 (C-1'), 109.16 (C-5), 136.63 (C-6), 150.90 (C-4), 164.38 (C-2), 186.61 (C=S).
Exact mass (ESI-MS) for C₁₂H₁₅N₃O₅SNa [M+Na]⁺ found, 336.0630; calcd, 336.0636; Anal. (C₁₂H₁₅N₃O₅S) C, H, N.

1-(2-*O*-Acetyl-3-azidomethyl-3-deoxy-β-D-ribofuranosyl)thymine (4.25)
1-(3-Azidomethyl-3-deoxy-β-D-ribofuranosyl)thymine (4.26)

To a solution of **2.19** (1.515 g, 3.91 mmol) in dry CH₂Cl₂ (156 mL), a 1 M BCl₃ solution in CH₂Cl₂ was added at -78 °C. After 15 minutes the reaction was quenched with methanol, evaporated and purified by column chromatography (CH₂Cl₂/MeOH 98:2) to yield compounds **4.25** (399 mg, 30%) and **4.26** (518 mg, 45%).
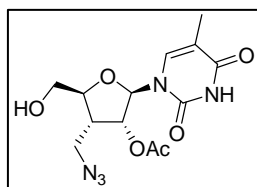
¹H NMR (300 MHz, DMSO-d₆): 1.76 (3H, d, 5-CH₃), 2.08 (3H, s, COCH₃), 2.70 (1H, m, H-3'), 3.50-3.59 (3H, m, H-6', H-6'' and H-5'), 3.71 (1H, d, *J*=12.4, H-5''), 3.92 (1H, m, H-4'), 5.17 (1H, t, OH-5'), 5.35 (1H, m, H-2'), 5.77 (1H, d, *J*= 3.3 Hz, H-1'), 7.75 (1H, d, *J*= 1.2, H-6), 11.31 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₃H₁₈N₅O₆ [M+H]⁺ found, 340.1263; calcd, 340.1257
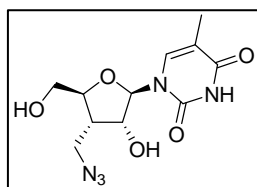
¹H NMR (300 MHz, DMSO-d₆): 1.74 (3H, d, 5-CH₃), 2.35 (1H, m, H-3'), 3.34-3.38 (1H, m, H-6'), 3.51-3.64 (2H, m, H-6'' and H-5'), 3.73-3.80 (1H, m, H-5''), 3.89 (1H, m, H-4'), 4.18 (1H, m, H-2'), 5.19 (1H, t, OH-5'), 5.65 (1H, d, *J*=2.1 Hz, OH-2'), 5.88 (1H, d, *J*= 5.1 Hz, H-1'), 7.92 (1H, d, *J*= 1.2, H-6), 11.26 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₁H₁₆N₅O₅ [M+H]⁺ found, 298.1146; calcd, 298.1151

1-(2-*O*-Acetyl-3-azidomethyl-3-deoxy-5-*O*-tosyl-β-D-ribofuranosyl)thymine (4.27)

To a solution of **4.25** (340 mg, 1.00 mmol) in pyridine (3 mL) at 0 °C, *p*-toluenesulfonyl chloride (381 mg, 2.00 mmol) was added. The reaction mixture was stirred at room temperature over 2 days. The reaction mixture was evaporated *in vacuo* to remove pyridine and partitioned between CH_2Cl_2 and H_2O. This mixture was extracted three times, dried over $MgSO_4$ and evaporated. The residue was purified by column chromatography (CH_2Cl_2/MeOH 99:1) to give compound **4.27** (430 mg, 87%).
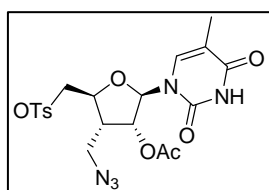
^{1}H NMR (300 MHz, CDCl$_3$): 1.87 (3H, d, 5-CH$_3$), 2.09 (3H, s, COCH$_3$), 2.46 (3H, s, Ph*CH$_3$*), 2.74 (1H, m, H-3'), 3.38 (1H, dd, *J*= 7.2 and 12.6, H-6'), 3.57 (1H, dd, *J*=6.3 and 12.6, H-6''), 4.16-4.24 (2H, m, H-5' and H-4'), 4.41-4.45 (1H, m, H-5''), 5.40 (1H, dd, *J*= 3.0 and 6.9, H-2'), 5.78 (1H, d, *J*= 3.0 Hz, H-1'), 7.37 (2H, m, ar H), 7.81 (2H, m, ar H), 7.93 (1H, s, H-6).

Exact mass (ESI-MS) for $C_{20}H_{23}N_5O_9SNa$ [M+Na]$^+$ found, 532.1114; calcd, 532.1114

1-(2-*O*-Acetyl-3-azidomethyl-3,5-dideoxy-5-iodo-β-D-ribofuranosyl)thymine (4.28)

A mixture of tosylate **4.27** (430 mg, 0.87 mmol), NaI (640 mg, 4.27 mmol) and a catalytic quantity of tetra-*n.*butylammonium iodide (5 mg) in acetone (34.8 mL) was refluxed at 60 °C overnight. Ice-cold water was poured into the reaction mixture. This mixture was extracted with CH_2Cl_2, dried over $MgSO_4$ and evaporated. The obtained residue was purified by column chromatography (CH_2Cl_2/MeOH: 99:1) to yield compound **4.28** (381 mg, 97%).
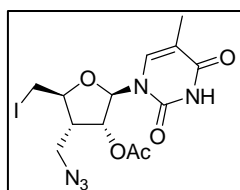
^{1}H NMR (300 MHz, CDCl$_3$): 1.88 (3H, d, 5-CH$_3$), 2.09 (3H, s, COCH$_3$), 2.61 (1H, m, H-3'), 3.35-3.41 (2H, m, H-6' and H-5'), 3.52-3.61 (2H, m, H-6'' and H-5''), 3.75 (1H, m, H-4'), 5.37 (1H, dd, *J*= 3.9 and 7.8, H-2'), 5.72 (1H, d, *J*= 3.6 Hz, H-1'), 7.25 (1H, m, H-6), 7.97 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{13}H_{16}N_5O_5INa$ [M+Na]$^+$ found, 472.0099; calcd, 472.0096

1-(3-Acetaminomethyl-3-deoxy-4-methyl-β-D-ribofuranosyl)thymine (4.30)

A mixture of **4.28** (41 mg, 0.091 mmol) and triethylamine (30 µL) in methanol (3 mL) was hydrogenated at atmospheric pressure for 3 hours in the presence of 10% Pd/C (12.3 mg). The catalyst was removed by filtration through celite and the filtrate was evaporated. The residue was purified by column chromatography (CH_2Cl_2/MeOH 95:5) to give the final compound **4.30** (31 mg, 98%).
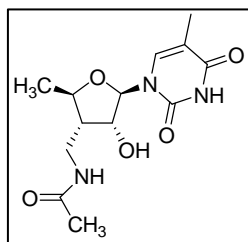
¹H NMR (300 MHz, $CDCl_3$): 1.44 (1H, m, 5'-CH$_3$), 1.92 (3H, d, 5-CH$_3$), 1.93-2.00 (1H, m, H-3'), 2.00 (3H, s, COCH$_3$), 3.12-3.21 (2H, m, H-6' and H-6''), 4.14 (1H, m, H-4'), 4.31 (1H, m, H-2'), 4.95 (1H, d, *J*= 2.9 Hz, H-1'), 5.59 (1H, d, *J*= 0.88 Hz, 2'-OH), 6.13 (1H, br t, NH-3'), 7.27 (1H, m, H-6), 8.85 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{13}H_{20}N_3O_5$ [M+H]$^+$ found, 297.1328; calcd, 297.1325.

1-(3-Azidomethyl-3-deoxy-5-*O*-tosyl-β-D-ribofuranosyl)thymine (4.31)

To a solution of **4.26** (398 mg, 1.34 mmol) in pyridine (3.2 mL) cooled at 0 °C, *p*-toluenesulfonyl chloride (307 mg, 1.61 mmol) was added. The reaction mixture was stirred at room temperature during 2 days and evaporated *in vacuo* to remove pyridine. The residue was partitioned between CH_2Cl_2 (10 mL) and H_2O (10 mL), separated, and the aqueous layer was extracted two times more with 15 mL of CH_2Cl_2. The combined organic layers were dried over MgSO$_4$, filtered and evaporated. The residue was purified by column chromatography (CH_2Cl_2/MeOH 97:3) to yield compound **4.31** (280 mg, 46%).
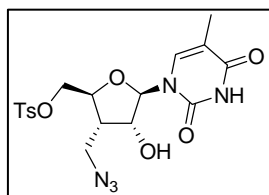
¹H NMR (300 MHz, DMSO-d$_6$): δ 1.77 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 2.29 (1H, m, H-3'), 2.41 (3H, s, PhC*H₃*), 3.31 (1H, m, H-6'), 3.58 (1H, dd, *J*= 5.7 and 12.3, H-6''), 4.06 (1H, m, H-4'), 4.20 (2H, m, H-5' and H-2'), 4.36 (1H, dd, *J*= 2.1 and 11.1, H-5''), 5.62 (1H, d, *J*= 2.4 Hz, H-1'), 5.90 (1H, d, *J*= 5.4 Hz, 2'-OH), 7.36 (1H, d, *J*= 1.5, H-6), 7.45 (2H, m, arom H), 7.77 (2H, m, arom H), 11.31 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{18}H_{22}N_5O_7S$ [M+H]$^+$ found, 452.1240; calcd, 452.1239.

1-(3-Azidomethyl-3,5-dideoxy-5-iodo-β-D-ribofuranosyl)thymine (4.32)

A mixture of **4.31** (342 mg, 0.76 mmol), NaI (557 mg, 3.71 mmol) and a catalytic amount of tetra-*n*.butylammonium iodide (5 mg) in acetone (30 mL) was refluxed at 60 °C overnight. Ice-cold water (5 mL) was poured into the reaction mixture. This mixture was extracted with CH_2Cl_2. The organic layer was dried over $MgSO_4$, filtered and evaporated, and the residue was purified by column chromatography (CH_2Cl_2/MeOH: 97:3) to yield compound **4.32** (311 mg, 100%).
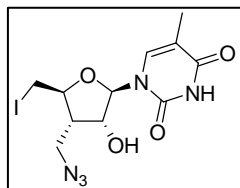
¹H NMR (300 MHz, DMSO-d$_6$): δ 1.79 (3H, d, *J*= 1.8 Hz, 5-CH$_3$), 2.26 (1H, m, H-3'), 3.45 (2H, m, H-6' and H-5'), 3.62 (2H, m, H-6'' and H-5''), 3.80 (1H, m, H-4'), 4.35 (1H, m, H-2'), 5.43 (1H, d, *J*= 3.6 Hz, H-1'), 5.88 (1H, d, *J*= 5.1 Hz, 2'-OH), 7.48 (1H, d, *J*= 1.2, H-6), 11.34 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{11}H_{14}N_5O_7INa$ [M+Na]$^+$ found, 429.9999; calcd, 429.9990.

1-(3-Aminomethyl-3,5-dideoxy-β-D-ribofuranosyl)thymine (4.33)

A mixture of **4.32** (311 mg, 0.764 mmol) and triethylamine (250 µL) in methanol (20 mL) was hydrogenated at atmospheric pressure overnight in the presence of 10% Pd/C (95 mg). The catalyst was removed by filtration through a celite path, and the filtrate was evaporated. The residue was purified by column chromatography (CH_2Cl_2/MeOH 85:15 to 80:20) to give compound **4.33** (134 mg, 69%).
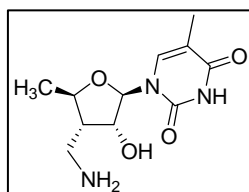
¹H NMR (300 MHz, DMSO-d$_6$): δ 1.32 (3H, d, *J*= 6.0 Hz, 5'-CH$_3$), 1.79 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 2.03 (1H, m, H-3'), 2.76 (1H, dd, *J*= 5.1 and 12.6, H-6'), 2.95 (1H, m, H-6''), 3.96 (1H, m, H-4'), 4.29 (1H, dd, *J*= 2.1 and 6.0 Hz, H-2'), 5.60 (1H, d, *J*= 2.1 Hz, H-1'), 7.32 (1H, d, *J*= 1.2 Hz, H-6).

Exact mass (ESI-MS) for $C_{11}H_{18}N_3O_4$ [M+H]$^+$ found, 256.1294; calcd, 256.1297.

1-(3-Aminomethyl-3,5-dideoxy-2-*O*,6-*N*-(thiocarbonyl)-β-D-ribofuranosyl)thymine (4.9)

A solution of **4.33** (127 mg, 0.497 mmol) and 1,1'-thiocarbonyldiimidazole (97.47 mg, 0.547 mmol) in THF (7 mL) was stirred at room temperature for 5 h. The mixture was evaporated to dryness and the residue was purified by column chromatography (CH_2Cl_2/MeOH 97:3) to give compound **9** (64 mg, 43%) as a white solid.
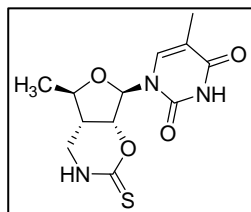
¹H NMR (300 MHz, DMSO-d$_6$): δ 1.40 (3H, d, J= 6.3 Hz, 5'-CH$_3$), 1.80 (3H, s, 5-CH$_3$), 2.61 (1H, m, H-3'), 3.16 (1H, d, J= 4.2 Hz, H-6'), 3.46 (1H, dd, J= 5.7 and 13.8 Hz, H-6''), 3.96 (1H, m, H-4'), 5.02 (1H, d, J= 5.7, H-2'), 5.80 (1H, s, H-1'), 7.37 (1H, s, H-6), 9.85 (1H, br s, N(6')H), 11.40 (1H, br s, N(3)H).

¹³C NMR (300 MHz, DMSO-d$_6$): δ 12.83 (5-CH$_3$), 18.67 (C-5'), 37.18 (C-6'), 38.52 (C-3'), 77.83 (C-4'), 84.14 (C-2'), 90.36 (C-1'), 110.24 (C-5), 136.54 (C-6), 150.88 (C-2), 164.49 (C-4), 184.70 (6-NHCS).

Exact mass (ESI-MS) for C$_{12}$H$_{16}$N$_3$O$_4$S [M+H]$^+$ found, 298.0858; calcd, 298.0861; Anal. (C$_{12}$H$_{15}$N$_3$O$_4$S) C, H, N.

Chapter 5

Synthesis of thiourea substituted thymidine derivatives

5 RATIONAL DESIGN OF 3'-BRANCHED-THIOUREA SUBSTITUTED β-THYMIDINE DERIVATIVES AND 5'-THIOUREA SUBSTITUTED α-THYMIDINE ANALOGUES

5.1 OBJECTIVE

Simultaneously with the discovery of the 2',3'-bicyclic thymidine derivatives, described in chapter 4, an unusual dinucleoside **5.1** (K_i= 37 μM) was revealed as TMPKmt inhibitor.[107] The second thymidine monomer attached to the 3'-position suggested the possibility to introduce large substituents at this position to fit in the dTMP-pocket. Next to its surprisingly inhibitory activity, compound **5.1** did not inhibit human TMPK (TMPKh) at 1 mM, indicating that further explorations of this type of analogues might result in potent and selective TMPKmt inhibitors.

5.1
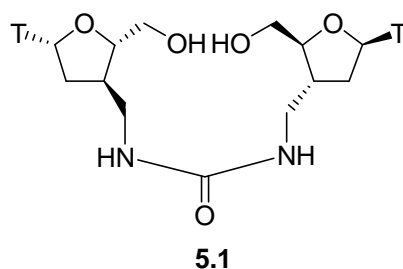
Figure 5.1 Dinucleoside **5.1** (K_i= 37 μM)

5.2 3'-*C*-BRANCHED THIOUREA SUBSTITUTED β-THYMIDINE DERIVATIVES

5.2.1 Rational design

Starting from the structure of dinucleoside **5.1**, we investigated the possibility to replace one thymidine monomer by substituted phenyl groups. Concurrently, thiocarlide,[146] a drug that formerly proved efficient in treating TB,[147,148] inspired us to replace the connecting urea group by a thiourea linker, resulting in 3'-*C*-branched thiourea substituted β-thymidine derivatives.

By carefully selecting a limited set of compounds based on the Topliss' tree, [149] we aimed to assess the relative importance of the lipophilic, electronic and sterical properties of the aryl substituents.
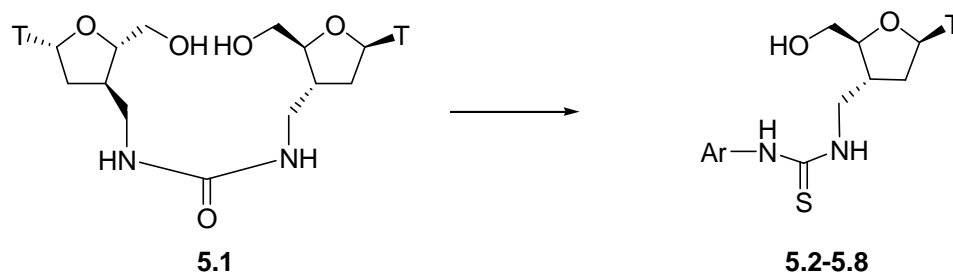
Figure 5.2 Dinucleoside **5.1** and envisaged derivatives

5.2.2 Chemistry

The synthesis of a series of 3'-branched-β-thiourea-derivatives started from β-D-thymidine. As depicted in chapter 2, different methods are described for the synthesis of 3'-branched nucleosides. A radical introduction of a β-styrene residue, followed by *cis*-dihydroxylation, oxidative cleavage and reduction resulted in compound **2.43**. After mesylation of the primary alcohol, the azido group was introduced using sodium azide in DMF. Hydrogenation of **5.35** with Pd/C gave amine **5.36**. This precursor was reacted with different isothiocyanates in a parallel fashion to give target compounds **5.2-5.8** after final deprotection of the silyl group.

5.2: $R_1 = H$; $R_2 = H$
5.3: $R_1 = Cl$; $R_2 = H$
5.4: $R_1 = OMe$; $R_2 = H$
5.5: $R_1 = CH_3$; $R_2 = H$
5.6: $R_1 = Cl$; $R_2 = Cl$
5.7: $R_1 = OCH_2Ph$; $R_2 = H$
5.8: $R_1 = Cl$; $R_2 = CF_3$
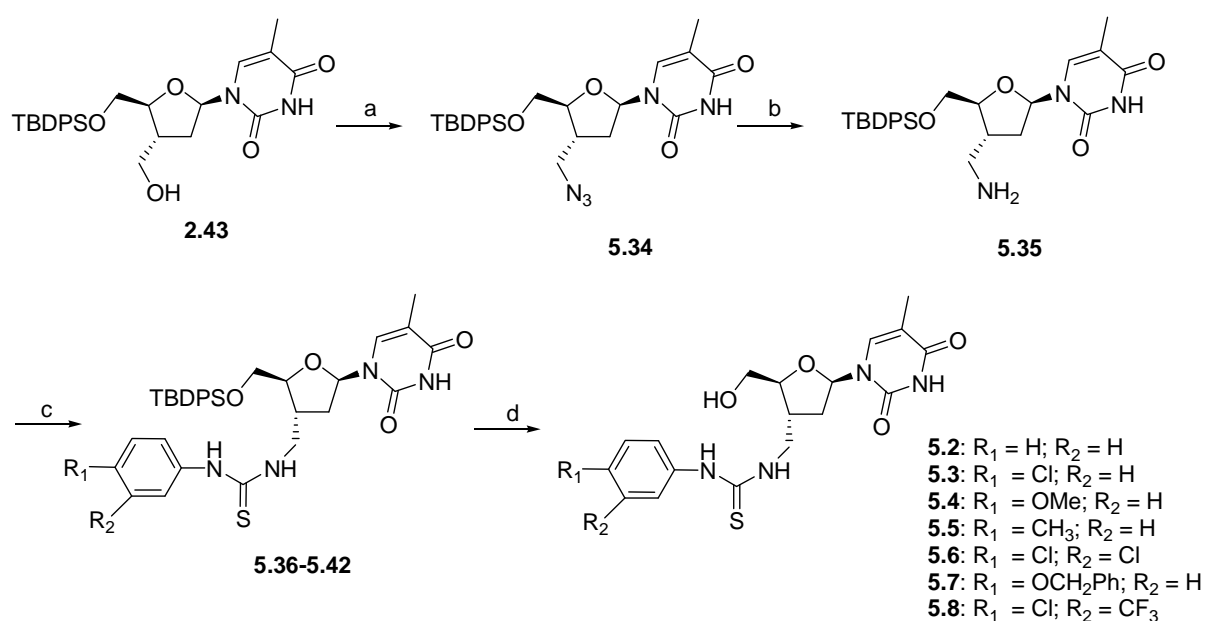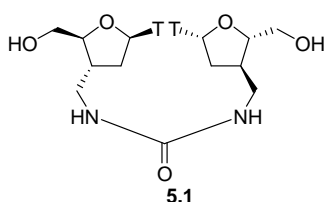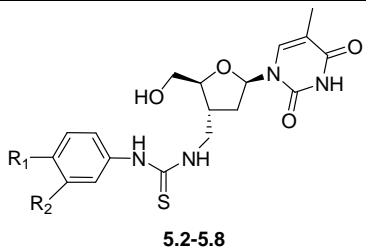
Scheme 5.1, Reagents and conditions: (a) (i) MsCl, pyridine, 0 °C, 1h; (ii) NaN$_3$, DMF, 60 °C, 7h, 97%; (b) Pd/C, H$_2$, MeOH, rt, 3h, 99%; (c) 3-R$_2$-4-R$_1$-phenylisothiocyanate, DMF, 0 °C, 1h, 70-88%; (d) TBAF, THF, rt, 1h, 87-92%.

5.2.3 Biological evaluation

All compounds were tested for TMPKmt inhibition as described in the experimental section of chapter 3. The results of the β-derivatives are shown in Table 5.1.

Table 5.1 Kinetic parameters of TMPKmt with compounds **5.1-5.8**

Compound	R_1	R_2	K_i (µM) TMPK mt	K_i (µM) TMPK h	MCC (µM) Vero cell cultures[b]
dTMP			4.5 [a]	5.0 [a]	
5.1			37.0	NI 1mM	
5.2	H	H	69.0		
5.3	Cl	H	21.0		> 100
5.4	OCH_3	H	46.0		> 100
5.5	CH_3	H	36.0		> 100
5.6	Cl	Cl	7.2		> 100
5.7	OCH_2Ph	H	12.0		> 100
5.8	Cl	CF_3	5.0	NI 1mM	> 100

[a] K_m-Value

[b] Minimum cytotoxic concentration in Vero cell cultures, i.e. the concentration required to cause microscopically detectable alteration of normal cell morphology.

For the 3'-*C*-branched-β-thymidine derivatives, the thiourea analogue **5.2** with an unsubstituted phenyl ring showed a K_i-value of 69.0 µM. This value indicates a weaker binding to the enzyme than dinucleoside **5.1** (K_i= 37 µM) and the natural substrate (K_m = 4.5 µM). However, this result suggests that this analogue can still be accommodated in the active pocket. In order to optimize the potency of this phenyl analogue, Topliss' scheme[149] (Figure 5.3) was used.

This decision tree is based on the concept, pioneered by Hansch,[150] that principally the lipophilic, electronic and steric properties of the introduced substituent influence the activity.
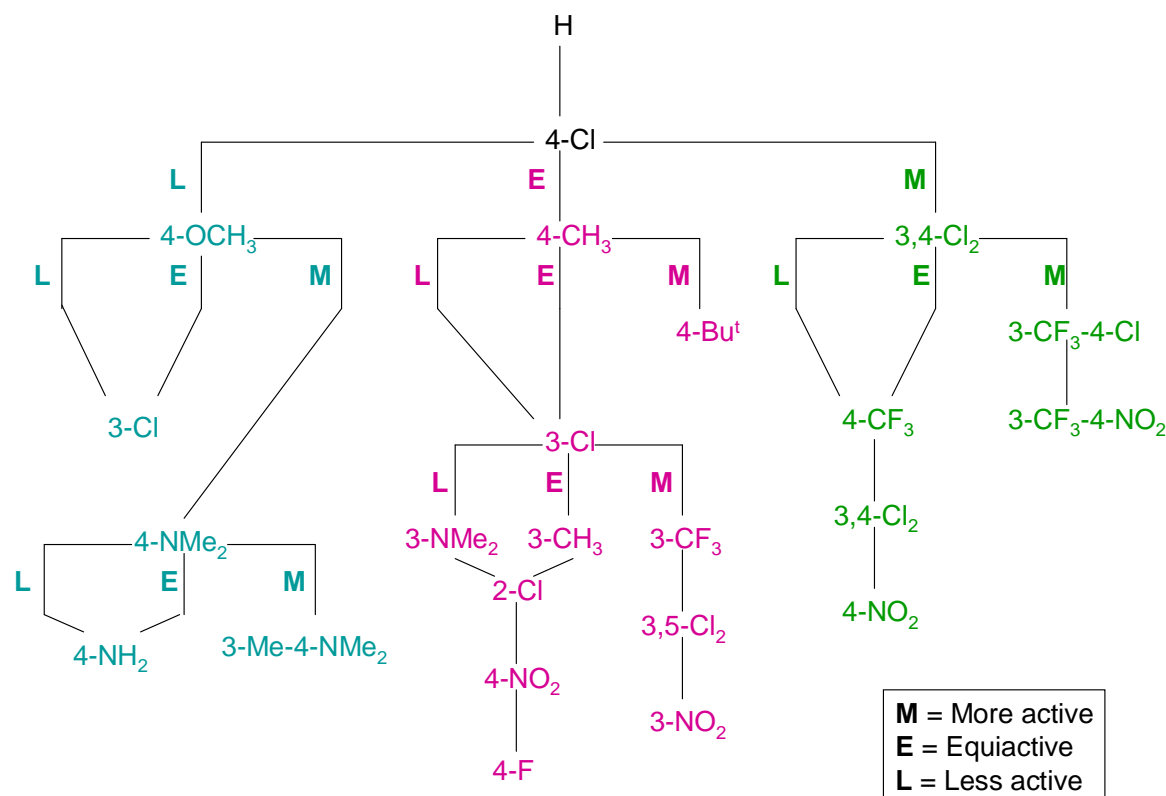
Figure 5.3 Topliss' tree

A 4-chloro-substitution of the phenyl ring (**5.3**) resulted in a 3-fold increased activity. According to Topliss' decision tree, this increase confirms the positive effect of lipophilic and electron withdrawing substituents on the inhibitory activity. This trend was validated by further improvement of the activity with a 3,4-dichloro-substitution pattern (**5.6**) which further increased activity by another 3-fold (K_i= 7.2 μM). On the other hand, a 4-methoxy- (**5.4**) or 4-methyl-substituent (**5.5**) resulted in a decreased activity. To further validate electron withdrawing properties, compound **5.8** was synthesized, while **5.7** should indicate if sterically more demanding substituents were also allowed. Compound **5.8** with a 3-CF$_3$-4-Cl-substitution pattern was found to be the most potent inhibitor of this series with a K_i-value of 5.0 μM, indicating that mainly the electronic features of the substituents positively influenced the inhibitory activity.

The good affinity observed among these β-thiourea derivatives suggests a binding mode analogous to that described for dinucleoside **5.1**[107] with the arylthiourea moiety occupying the phosphoryl donor binding site.

5.3 5'-*O*-PHOSPHORYLATED α-THYMIDINE

5.3.1 Rational design

Modelling studies on the binding mode of dinucleoside **5.1** to its target suggested that the second monomer binds to the area where normally the phosphoryl donor binds, in this way forcing the sugar ring to tilt over 180° compared t o the natural substrate dTMP. Similarly, the envisaged β-thiourea derivatives are expected to bind the dTMP-pocket upside down, positioning the aromatic 3'-substituent into the phosphoryl donor binding area, and the nucleobase below the sugar plane.
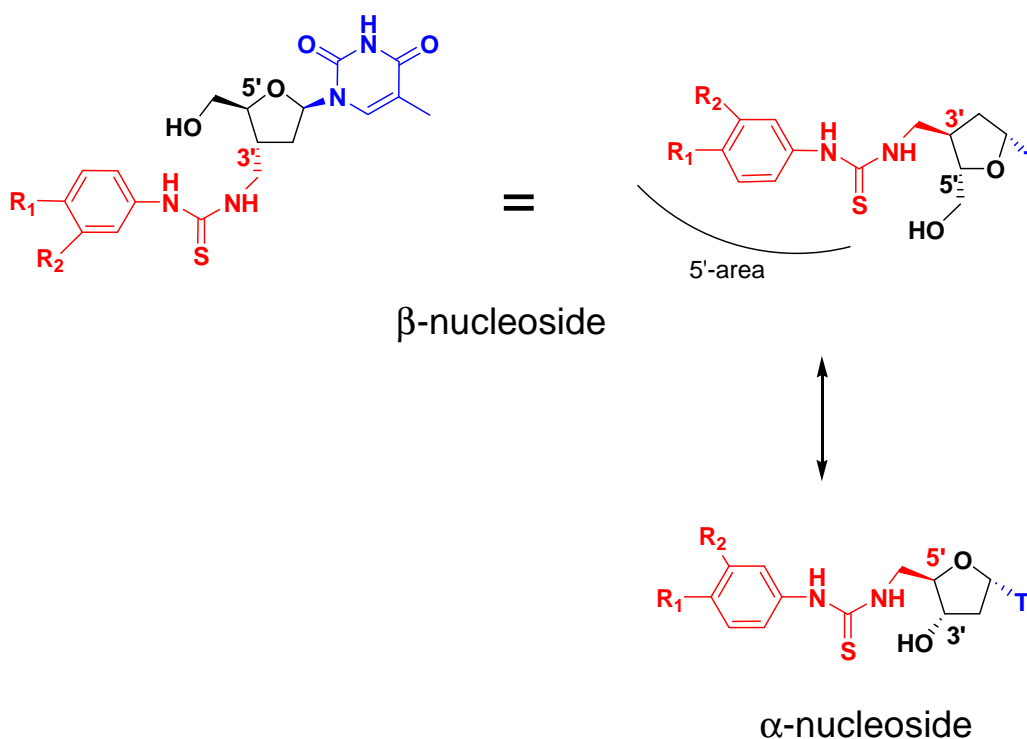
Figure 5.4 Suggested inverse sugar binding of 3'-*C*-arylthiourea-modified β-thymidines and anticipated similar relative orientation of the coloured moieties in 5'-deoxy-5'-arylthiourea modified α-thymidines.

As depicted in Figure 5.4, we hypothesized that the relative orientation of the nucleobase and the arylthiourea moiety in these β-thiourea derivatives might be imitated by 5'-substituted α-thymidine analogues.

In order to sustain this hypothesis, 5'-*O*-phosphorylated α-thymidine was synthesized to assess if it can be accepted by TMPKmt as a substrate.

5.3.2 Chemistry

For the synthesis of α-D-thymidine 5'-*O*-monophosphate **5.9**, 3',5'-*O*-diacetyl-β-D-thymidine was first anomerized to its α-anomer **5.44** according to the procedure of Ward et al.,[151] followed by deprotection and conversion to the 5'-monophosphate using classical phosphorylation conditions.
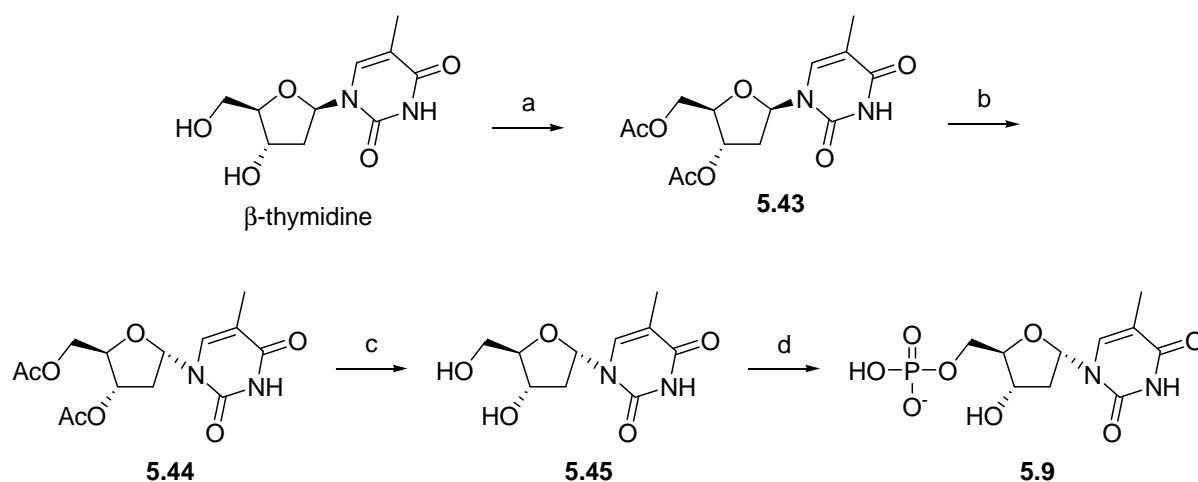
Scheme 5.2, Reagents and conditions: (a) Ac_2O, Pyridine, rt, 2h, 99%; (b) H_2SO_4, Ac_2O, acetonitrile, rt, 2h, 45%; (c) NH_3, MeOH, rt, 16h, 98%; (d) $POCl_3$, $P(OMe)_3$, 0 °C, 4h then rt, 0.5h, 62%.

5.3.3 Biological evaluation

The reported spectrophotometric assay[119] confirmed that α-thymidine monophosphate was accepted as a substrate by TMPKmt (K_m= 450 μM, V_m= 0.77 U/mg), thereby competing with dTMP (K_i= 15.0 μM). Compared to the natural substrate however, the binding constant of this anomer was 10 times higher and its maximum velocity (V_m) only 7.5% of that of dTMP.

The same broad stereoisomeric substrate specificity has been observed for mitochondrial thymidine kinase (TK2), deoxyguanosine kinase (dGK) and deoxycytosine kinase (dCK), which proved able to recognize β-D and β-L, as well as α-D and α-L nucleosides as a substrate.[152]

The fact that 5'-*O*-phosphorylated α-thymidine was accepted by TMPKmt as a substrate confirmed that α-nucleosides proved able to adopt the proposed binding mode, which incited us to synthesize the readily accessible 5'-deoxy-5'-*N*-arylthiourea α-thymidine derivatives.

5.4 5'-THIOUREA SUBSTITUTED α-THYMIDINE DERIVATIVES

5.4.1 Objective

As discussed above, the validation of 5'-*O*-phosphorylated α-D-thymidine monophosphate as a TMPKmt substrate, combined with the encouraging results obtained with the 3'-branched β-thymidine series opened interesting perspectives for the synthesis of 5'-thiourea substituted α-nucleosides.

A detailed exploration of the 5'-thiourea substituent pattern was envisaged, followed by a study about the importance of the deoxyribofuranose moiety by the synthesis of 2',3'-dideoxy, 2',3'-dideoxydidehydro- and acyclic derivatives.

Very recently,[153,154] a series of 5'-modified adenosine derivatives has been discovered as potent siderophore biosynthesis inhibitors. Some of these analogues showed MIC_{99}-values against *M. tuberculosis* H37Rv comparable to that of isoniazid, suggesting the presence of active adenosine-uptake mechanisms. In an attempt to use adenosine to enhance the uptake in the mycobacterial cell, a dinucleoside was designed, consisting of a α-thymidine monomer connected to adenosine by a thioureum linker.

The synthesis and biological evaluation of those compounds allows the determination of a detailed structure-activity relationship.

5.4.2 Chemistry

The synthesis of 5'-substituted α-thymidine derivatives started from α-thymidine (**5.45**) which was converted to the 5'-deoxy-5'-(thio)urea-derivatives **5.10**-**5.27** as depicted in Scheme 5.3. Amide **5.30** (Table 5.2) was isolated as a side product upon treatment of **5.47** with benzoylisothiocyanate.
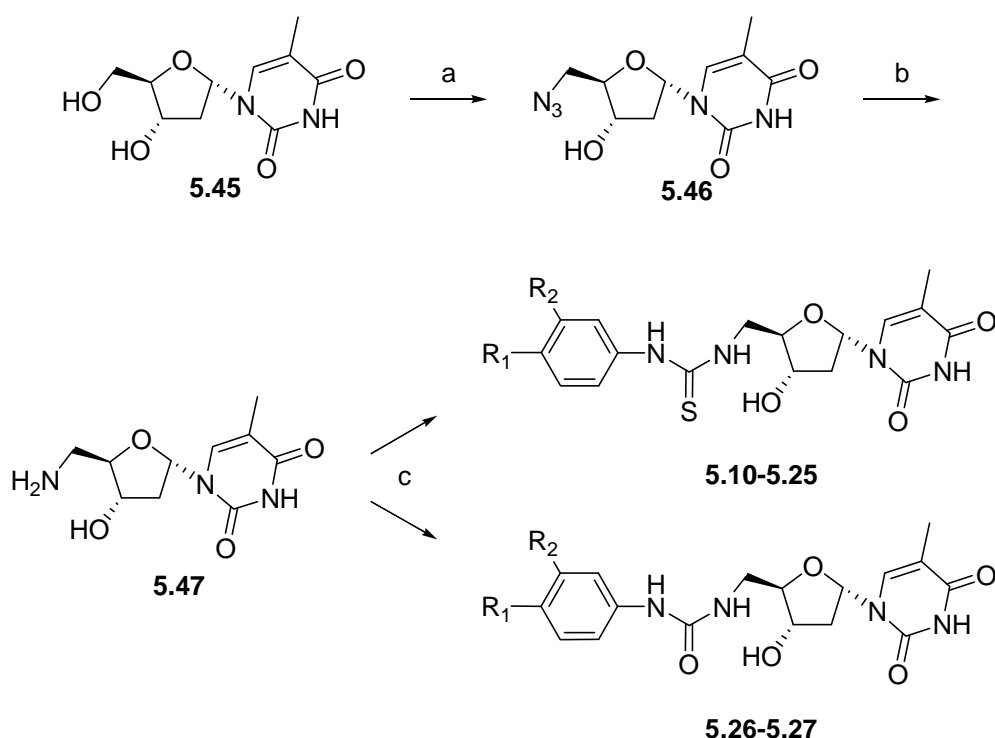
Scheme 5.3, Reagents and conditions: (a) (i) MsCl, Pyridine, 0 °C, 1h (ii) NaN$_3$, DMF, 60 °C, 16h, 68%; (b) H$_2$, Pd/C, MeOH, rt, 6h, 98%; (c) suitable iso(thio)cyanate, DMF, 71-91%; or: suitable amine, 1,1'-TCDI, DMF, rt, 3h, 75-82%.

3'-Deoxygenation was performed by converting the 3'-hydroxyl group to its mesylate ester (Scheme 5.4). Upon treatment with base, elimination yielded the unsaturated sugar in nucleoside **5.49**. Hydrogenation over Pd/C gave the 2',3'-dideoxy-analogue **5.50**, while Staudinger reduction selectively converted the azide, affording the unsaturated amine **5.51**. Final treatment of these amines with 3-CF$_3$-4-Cl-isothiocyanate afforded derivatives **5.28** and **5.29**.
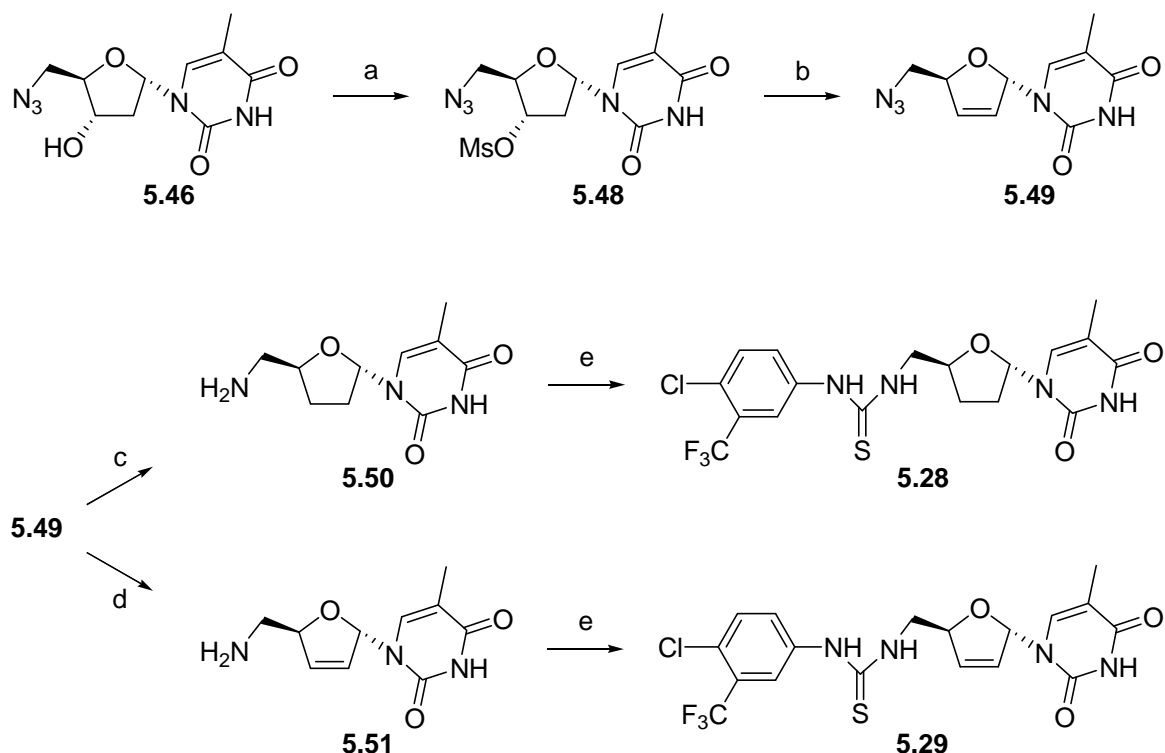
Scheme 5.4, Reagents and conditions: (a) MsCl, Pyridine, 0 °C, 2h, 81%; (b) DBU, THF, 80 °C, 16h, 95%; (c) H$_2$, Pd/C, MeOH, rt, 5h, 95%; (d) PPh$_3$, Pyridine, rt, 3h, 92% (e) 3-CF$_3$-4-Cl-isothiocyanate, DMF, rt, 3h, 78%.

Acyclic derivatives were synthesized following a procedure as reported for uridine-derivatives by Danel et al.[155] (see Scheme 5.5). The primary alcohol of **5.52** was converted to different thiourea-compounds as described above.
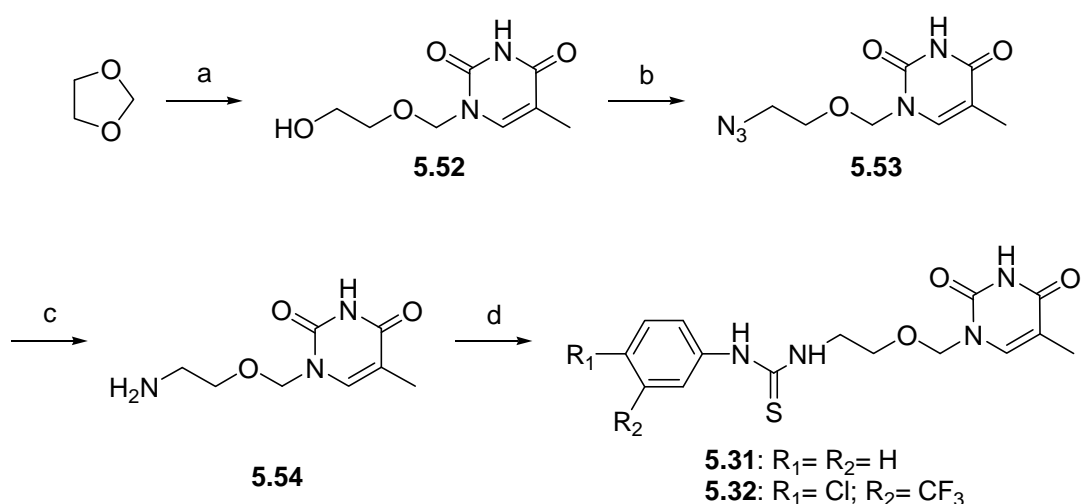
Scheme 5.5, **Reagents and conditions**: (a) Silylated thymine, TMSOTf, acetonitrile, -45 °C, 2h then rt, 16h, 76 %; (b) (i) MsCl, Pyridine, 0 °C, 2h (ii) NaN$_3$, DMF, 60 °C, 16h, 96%; (c) Pd/C, H$_2$, MeOH, rt, 16h, 99%; (d)) 3-R$_2$-4-R$_2$ isothiocyanate, DMF, rt, 3h, 83-86%.

For the preparation of the adenosine-conjugate **5.33**, adenosine was protected with a 2',3'-isopropylidene group, followed by mesylation and azido-substitution to compound **5.56**. This compound was reduced to obtain 5'-amino-2',3'-isopropylidene adenosine.[156] Upon treatment with 1,1'-thiocarbonyldiimidazole, 5'-amino-2',3'-isopropylidene adenosine **5.56** was converted *in situ* to the isothiocyanate, which was further reacted with 5'-amino-α-thymidine to afford the envisaged dinucleoside after deprotection of the acetonide.
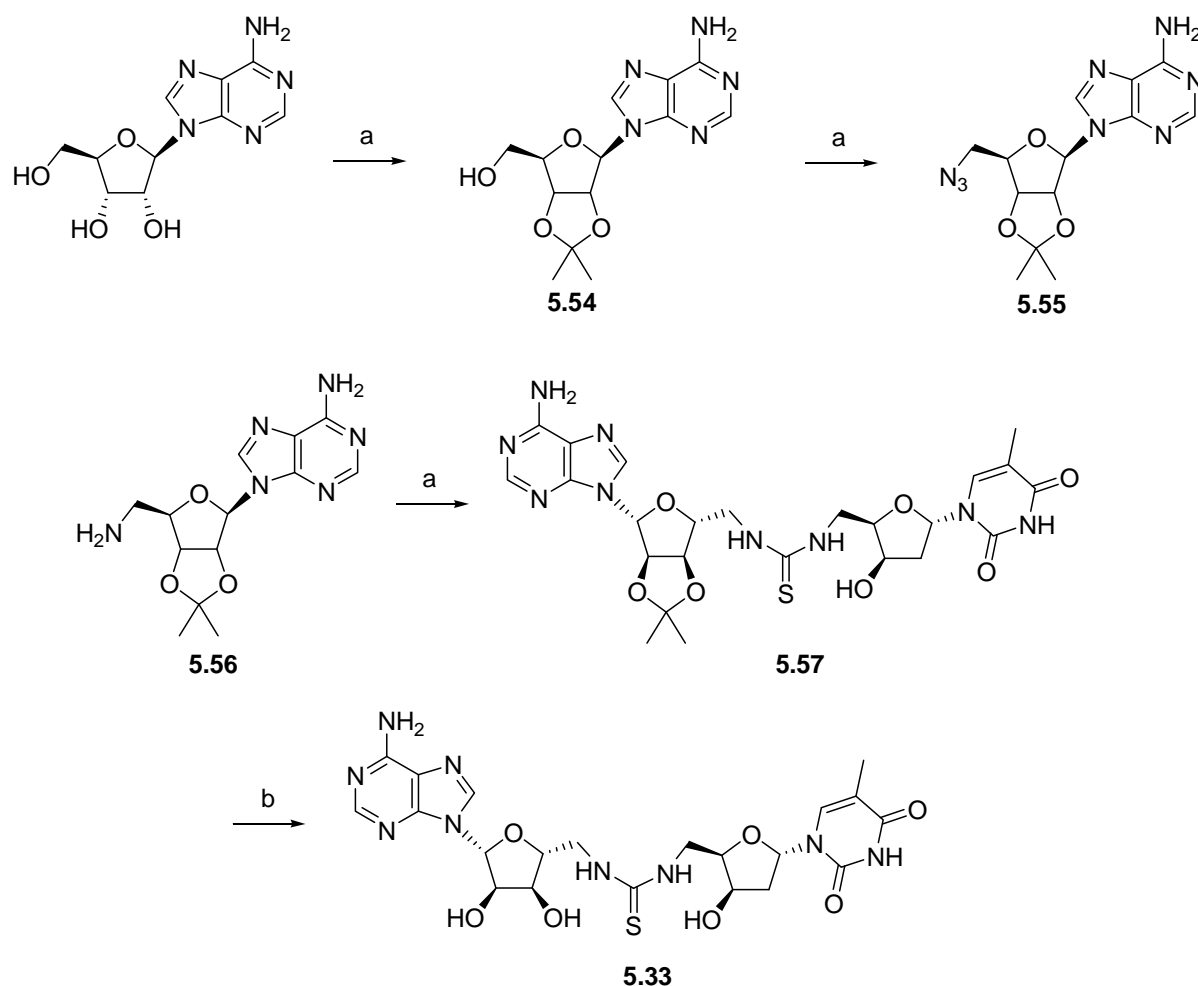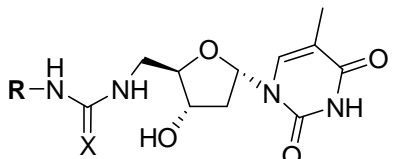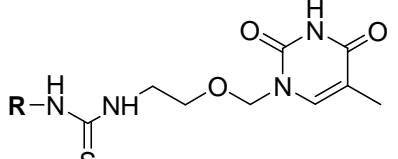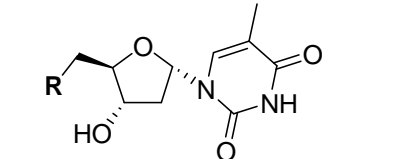
Scheme 5.6, Reagents and conditions: (a) Dimethoxypropane, Acetone, PTSA, rt, 1h, 77%; (b) (i) MsCl, Pyridine, 0 °C, 1h, (ii) NaN$_3$, DMF, 60 °C, 16h, 71%; (b) Pd/C, H$_2$, MeOH, rt, 16h; (c) (i) 5'-NH$_2$-α-thymidine, TCDI, DMF, rt, 2h; (ii) 50% TFA, rt, 2h, 73%.

5.4.3 Biological evaluation

The biological activities of the synthesized α-thymidine derivatives are represented in Table 5.2. In analogy to our observations for the 3'-*C*-branched β-thymidine thiourea analogues, lipophilic and electronic withdrawing substituents favour high activity. With a K_i of 0.6 μM, analogue **5.15** (R$_1$=Cl, R$_2$= CF$_3$) proved to be the most potent TMPKmt inhibitor reported so far.

Table 5.2: Kinetic parameters of TMPKmt with compounds **5.9-5.33 and 5.46-5.47**

5.10-5.29,5.33   **5.31-5.32**   **5.30, 5.46-5.47**

Comp N°	R	X	K_i (µM) TMPKmt	K_i (µM) TMPKh	SI	MIC$_{99}$ (µg/ml) (*M. bovis*)	% Inh (*M. tub* H37Rv)[d]	MCC µM Vero cell cultures[e]
Dtmp			4.5 [a]	5.0 [a]				
A-Dtmp			15.0					
5.10	Phenyl	S	16.0					> 100
5.11	4-Cl-phenyl	S	3.2	> 1000				
5.12	4-MeO-phenyl	S	10.0					
5.13	4-Me-phenyl	S	7.8					
5.14	3,4-diCl-phenyl	S	1.0	274	270	25	38%	> 100
5.15	3-CF$_3$-4-Cl-phenyl	S	0.6	362	600	20	39%	> 100
5.16	4-morpholinophenyl	S	19.2					
5.17	1-adamantyl	S	15.0					> 100
5.18	3-pyridyl	S	18.0					
5.19	Fluoresceinyl	S	5.4			> 150		
5.20	Phenylmethyl	S	5.4					
5.21	Benzhydryl	S	4.9	> 500				> 100
5.22	Benzoyl	S	7.2					
5.23	3-CF$_3$-4-Cl-phenylmethyl	S	2.6	50	20	40		> 100
5.24	Phenylethyl	S	3.8	> 800				> 100
5.25	3,4-diCl-phenylethyl	S	2.2	79	40	50		> 100
5.26	3,4-diClphenyl	O	1.1	116	105	40 (MIC$_{50}$)	41%	
5.27	3-CF$_3$-4-Cl-phenyl	O	1.9	166	90	40 (MIC$_{50}$)	49%	
5.28 [b]	3-CF$_3$-4-Cl-phenyl	S	2.3	> 500		40	34%	
5.29 [c]	3-CF$_3$-4-Cl-phenyl	S	3.8	190	50	100	55%	
5.30	Benzamido		35					
5.31	Phenyl		260					> 100
5.32	3-CF$_3$-4-Cl-phenyl		37.0				34%	100
5.33	5'-deoxy-β-D-adenosin-5'-yl	S	26	280	10			
5.46	N$_3$		26.5					
5.47	NH$_2$		16.0			> 40		

[a] K_m-Value; [b] 3'-deoxy-ribonucleoside; [c] 3'-deoxy-2',3'-didehydronucleoside; [d] inhibition at 6.25 µg/mL; [e] Minimum cytotoxic concentration in Vero cell cultures, i.e. the concentration required to cause microscopically detectable alteration of normal cell morphology.

As the corresponding α- and β-thiourea derivatives were designed to bind in a similar fashion, a comparison between the corresponding activities is presented in Figure 5.5.
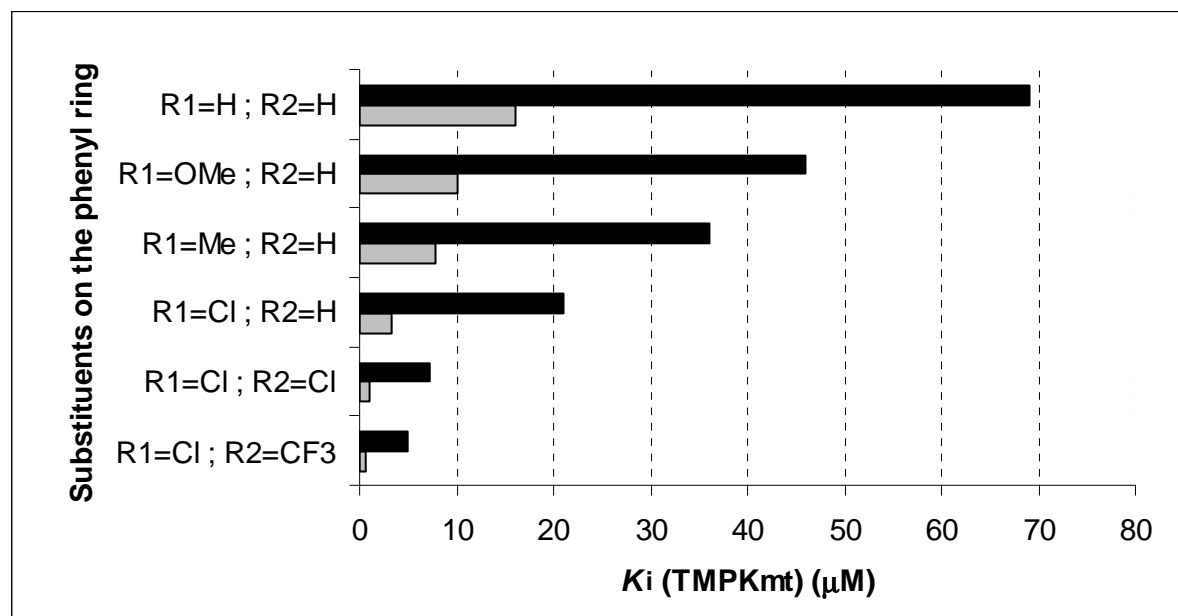
Figure 5.5 Comparison of the parallel synthesized alpha- and beta- thiourea derivatives; (■ β-derivatives, ▪ α-derivatives)

The graph shows that the α-derivatives are consistently (~ 4- to 5-fold) more active than the β-congeners, and clearly reveals a similar trend in both series, further supporting our design hypothesis.

To illustrate the binding mode of the two main series of inhibitors, docking of two representative compounds **5.2** (β) and **5.10** (α) was performed. This experiment resulted in 50 conformations, mainly containing two different orientations: one with the nucleobase stacking with Phe 70 and one with the aromatic tail stacking to Phe 70. As opposed to the former, the latter conformations preclude certain amino acids (*e.g.*, Arg 74, Asn 100) to form H-bonds and were therefore rejected. This rejection is justified as docking showed that the space close to Phe 70 is too small to accommodate bulky substituents like a Cl or a CF_3 group. Moreover the docked conformations with the aromatic ring occupying the base site are shifted, reducing the stacking interaction and thus drastically limiting the number of meaningful conformations found.
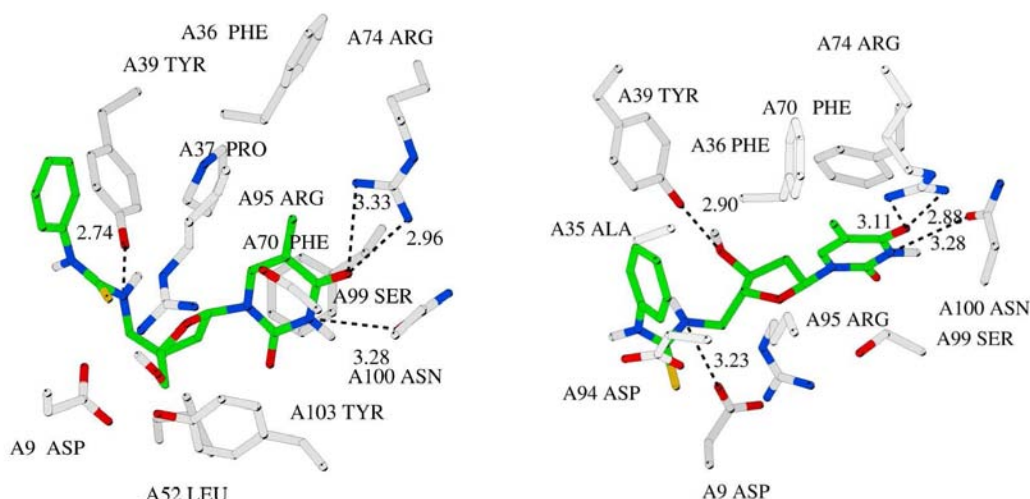
Figure 5.6 Compound **5.2** (left) and compound **5.10** (right) docked in the active site of TMPKmt: All residues interacting with the inhibitors (hydrophobic contact) are shown as well as the hydrogen bonds (calculated using Ligplot and HBPlus.[157, 158] Drawings created using Molscript.[159]

When comparing corresponding analogues from both series such as **2** and **10** (Figure 5.6), both the nucleobase and the arylthiourea-moiety occupy the same areas. For the α-derivatives, the pyrimidine ring and Phe 70 are arranged in a more parallel fashion compared to the β-derivatives, resulting in a stronger stacking. Moreover compound **5.10** is more favourably positioned to form an extra H-bond with Asp 9 through a nitrogen of its thiourea function, which may contribute to the higher affinity of the α-analogues for TMPKmt.

Further structural modifications elucidated that small structural changes, such as introduction of an alkyl chain of 1 or 2 carbon atoms between the thioureum and phenyl, are well tolerated by TMPKmt, yet not enhancing the affinity for the enzyme.

It is found that the tail of molecules **5.10** and **5.24** points to the outside of the enzyme through a channel, the same orientation as dinucleoside **5.1**.[107] This exit channel is surrounded by residues Arg14, Ala35, Phe36, Pro37, Tyr39, Arg160 (Figure 5.7). Adding functional groups in *para-* and *meta*-position on the phenyl ring seems to improve the activity, by increasing hydrophobic contact with these residues. However since longer chains extend more out of the enzyme into the solvent (Figure 5.7), this indicates that no further improvement can be obtained by increasing the linker length.
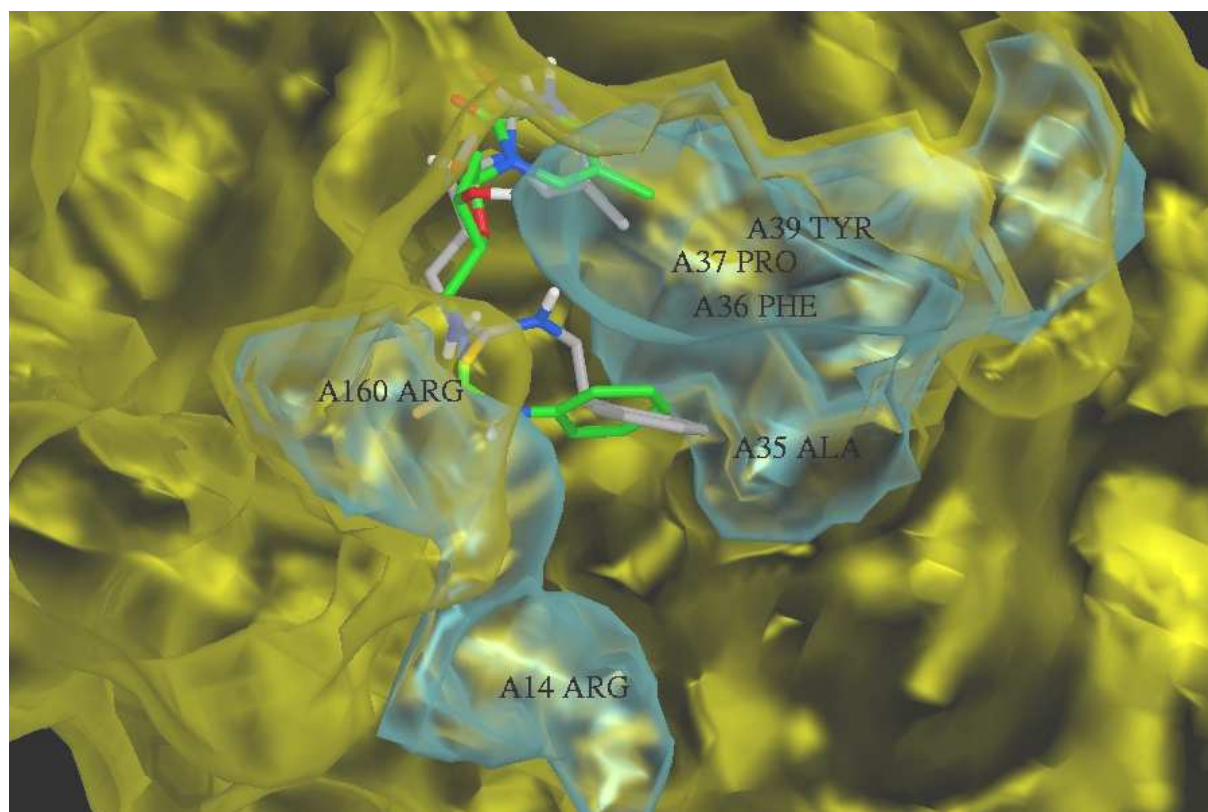
Figure 5.7 Docking and superposition of α-derivatives **5.10** (green carbons) and **5.24** (grey carbons) in the active site of TMPKmt, **5.24** having a longer linker (n=2). The phenyl tails of the inhibitors point to the front and seem to exit the enzyme through a channel lined up by residues Arg14, Ala35, Phe36, Pro37, Tyr39, Arg160. The "channel" residues are coloured blue. Picture created using, Molscript,[160] Bobscript,[161] Raster3D.[162]

Surprisingly, sterically demanding substituents as in **5.19** and **5.21** cause almost no change in inhibitory activity, so most likely a large pocket around the phosphoryl donor area is available for the inhibitor to interact with. The activity of the fluoresceine-labelled compound **5.19** opens interesting perspectives for its use in the enzymatic activity determination or in mycobacterial uptake-studies. The lower activity of compounds **5.16** and **5.17** suggests a role for the aromatic moieties to stabilize the enzyme-inhibitor complex.

To check the importance of the hydrophobic sulfur of the thiourea linkage, urea derivatives **5.26** and **5.27** were synthesized. As those compounds retain their good affinity for the enzyme, a larger structural freedom for the linkage between the aromatic moiety and the nucleobase seems to be tolerated.

Only a small drop in affinity was observed upon removal of the 3'-hydroxyl (**5.28**) or upon additional introduction of a double bond between 2' and 3' (**5.29**). We therefore conclude that the most important feature for TMPKmt inhibition of this series is the relative orientation

of the nucleobase and the 5'-substituent, combined with the ability for the 5'-substituent to interact with the enzyme active pocket.

This led us to synthesize acyclic derivatives **5.31** and **5.32**, which exhibited only a very weak affinity for the enzyme. Despite the fact that the number of bonds between the thymine and the aromatic moiety is retained in these structures, the higher entropy changes required to form the complex are likely to be unfavourable for binding.

The selectivity of the most potent TMPKmt vis-à-vis the human isozyme was investigated. Most tested compounds proved highly selective for the mycobacterial enzyme. Especially compound **5.15**, the most active compound of the series, showed a selectivity index of 600, indicating that the inverse binding mode of the sugar is not tolerated by the human enzyme.

Consequently, we assessed the potential of the most promising analogues to restrain mycobacterial growth. Whereas compound **5.15** proved to be the most potent TMPKmt inhibitor, it also proved superior in a *M. bovis* growth inhibition assay. Selected compounds were also tested for their capability to reduce *M. tuberculosis* H_{37}Rv growth at a concentration of 6.25 µg/mL, showing an inhibition of bacterial growth between 34 and 55% at this concentration.

For the first time, TMPKmt inhibitors proved capable to inhibit the growth of both *M. bovis* and *M. tuberculosis*, confirming that TMPKmt indeed represents a valuable target for designing antituberculosis drugs. Moreover, the compounds were found non-toxic in Vero cell cultures at 100 µM and also devoid of appreciable inhibitory activity against a broad variety of viruses including herpes simplex virus type 1 (KOS) and type 2 (G), vaccinia virus, vesicular stomatitis virus, Sindbis virus, reovirus-1, parainfluenza virus-3, Coxsackie B4, respiratory syncytial virus, Punta Toro virus, feline coronavirus, influenza virus A (H1N1 and H3N1) and influenza virus B. This lack of activity further points to the highly selective activity of several compounds against mycobacteria.

Very recently, a series of 5'-modified adenosine derivatives has been discovered as powerful siderophore biosynthesis inhibitors.[153,154] Some of these analogues showed MIC_{99}-values against *M. tuberculosis* H37Rv comparable to that of isoniazid, suggesting the presence of active adenosine-uptake mechanism. For this reason, conjugate **5.33** was synthesized to exploit this proposed adenosine uptake mechanism to enhance antimycobacterial activity. While still showing moderate TMPKmt inhibitor activity, dinucleoside **5.33** failed to inhibit mycobacterial growth.

Finally, all compounds tested in Vero cell cultures where devoid of cytotoxicity at 100 µM, while another assay proved **5.15** to have a minimum cytotoxic concentration of 300-400 µg/mL, which is more than 10 times the concentration needed to kill *M. bovis*. The urea analogues (**5.26** and **5.27)** of the most active compounds (**5.14** and **5.15**), which lack the thiourea moiety, earlier described as the perpetrator of toxicity,[163] showed nearly equal inhibitory activity against TMPKmt, while still retaining a sufficient selectivity versus the human enzyme (with SI= 90-105) and interesting activity on the growing *M. bovis* and *M. tuberculosis* strains.

5.5 QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP ANALYSIS

As explained in the introduction (See 1.3) Quantitative Structure-Activity Relationship (QSAR) analysis, is a technique developed to design new active compounds in a rational way. As the target is not considered in this method, QSAR can also be used for compounds for which the target enzyme remains unknown. This technique is based on a set of analogues of a biologically active compound for which the physicochemical properties are identified and quantified by making use of descriptors (e.g. MW, log P). A correlation is determined between these descriptors and the corresponding activities, which is then used for the prediction of the activities of new analogues.

In collaboration with the laboratory for Quantum Chemistry of Ugent (Prof. P. Bultinck) a dataset of 22 thiourea derivatives (involving both $\alpha-$ and β-nucleosides) was submitted to a QSAR analysis. For this study it is important that all K_i-values are obtained using the same assay method, and the pK_i-values should span a sufficiently wide range from 2 to 6. As the analysis was not performed in our group, only a brief result is described, without detailed description of the methods used.

A good correlation could be obtained for a model with the following 5 descriptors:
- Log P: Partition coefficient between octanol and water (positively correlated).
- Electrophilicity/TPSA (= Total Polar Surface Area) (positively correlated).
- $Q_{max}O$: maximum charge on an oxygen atom according to mulliken population analysis (positively correlated).
- Alpha: alpha- or beta-substituted compounds (positively correlated).
- Simi: similarity with dTMP calculated on quantumchemical level (negatively correlated).

The fitting between the observed and predicted activity, obtained with this model, is shown in Figure 5.8, in which the crossing of the axes lies at the $-\log(K_i)$-value of dTMP. With a $R^2 \gg$ 0.75, the model gives a good fitting, resulting in a high predictability.
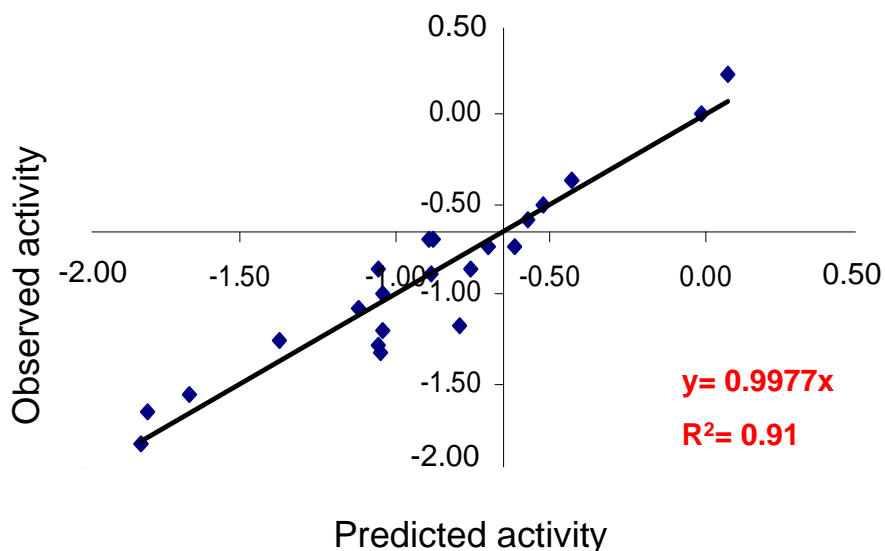
Figure 5.8 Fitting between observed and predicted activity as obtained with the 5-descriptor model

Unfortunately, when this model was extended with a wider range of earlier synthesized compounds, the good correlation as observed for the thiourea analogues disappeared. In attempt to build a new model with a good fitting between observed and predicted activities, only models with a low predictability and based on meaningless descriptors (e.g. the difference in mass between atoms on a distance of 3 bonds from each other) could be obtained, further limiting the use of QSAR to develop new compounds.

5.6 CONCLUSION

Based on the structure of the recently discovered dinucleoside **5.1** as a selective TMPKmt inhibitor, this chapter describes the synthesis and biological evaluation of a series of 3'-*C*-aryl thiourea derivatives of β-D-thymidine. Optimisation of the aryl thiourea led to analogue **5.7** that, with a K_i-value of 5.0 μM, clearly surpassed **5.1** in inhibitory activity.

Modelling experiments suggested a distinct binding mode for these analogs, as compared to the natural substrate: the sugar moiety binds the active site upside down. The acceptance of α-thymidine-5'-phosphate as a substrate for TMPKmt supported the synthesis of 5'-substituted α-thymidine thiourea derivatives, mainly characterized by a similar relative orientation of the 5'-aryl moiety and the nucleobase.

This led to the discovery of compound **5.15** as the most promising compound of this series with a K_i-value of 0.6 μM, a selectivity index (TMPKh) of 600, and a good inhibitory activity on the growing *M. bovis* (MIC$_{99}$ = 20 μg/mL) and *M. tuberculosis* (39% inhibition at 6.25 μg/mL) strains.

Next to the relative orientation between the aryl moiety and the nucleobase, structural exploration of the α-thymidine derivatives, revealed the positive impact of electronic withdrawing and lipophilic substituents on the 5'-aryl moiety and the need for aromatic residues at this 5'-position.

In conclusion, we have designed, synthesized, and evaluated a series of nucleoside inhibitors of *M. tuberculosis* TMPK, which resulted in the identification of 5'-arylthiorea α-thymidine analogues endowed with significant inhibitory activity against *M. tuberculosis* and *M. bovis* growth and low cytotoxicity. This strategy represents a promising approach for the development of a new class of antibiotics effective for the treatment of TB.

In a broader sense, this study opens interesting perspectives for using sugar-modified α-nucleosides as readily accessible scaffolds for the rational design of biologically important tool compounds acting on kinase targets or in other areas in chemical biology.

5.7 EXPERIMENTAL SECTION

Biological assays on Mycobacterium bovis (BCG)

The different compounds were assayed for their inhibitory potency on *Mycobacterium bovis* var. BCG growth *in vitro*.[164] A micro-method of culture was performed in 7H9 Middlebrook broth medium containing 0.2% glycerol, 0.5% Tween 80 and supplemented with oleic acid, albumin, dextrose and catalase (Becton-Dickinson). Serial two fold dilutions of each compound were prepared directly in 96 well plates. The bacterial inoculum was prepared previously at concentration in the range of 10^7 bacteria (*M. bovis* BCG 1173P2) in 7H9 medium and stored at -80 °C until used. The bacteria, adjusted at 10^5 per ml, were delivered under a volume of 100µl per well. The covered plates were sealed with Parafilm and incubated at 37 °C in plastic boxes containing humidified normal atmosphere. At day 8 of incubation, 30 µl of a resazurin (Sigma) solution at 0.01% (wt/vol) in water were added to each well. After overnight at 37 °C the plates were assessed for colour development using the optical density difference at 570 nm and 630nm on a ELISA reader. The change from blue to pink indicates reduction of resazurin and therefore bacterial growth. The lowest

compound concentration that prevented the colour change determined the MIC for the assayed compound.

Biological assays on Mycobacterium tuberculosis H$_{37}$Rv

Antimycobacterial data were provided by the Tuberculosis Antimicrobial Acquisition and Coordinating Facility (TAACF) through a research and development contract with the U.S. National Institute of Allergy and Infectious Diseases.

The screening test against *M. tuberculosis* H37Rv was conducted at 6.25 mg/mL in BACTEC 12B medium using the Microplate Alamar Blue Assay (MABA).[165]

Cytotoxic activity of test compounds in Vero cell cultures

To monolayers of confluent Vero cell cultures in 96-well microtiter plates were added serial dilutions of the test compounds (total volume of compound-containing culture medium: 200 µl). After 3 days of incubation at 37 °C in a humidified CO_2-controlled atmosphere, the cell cultures were microscopically inspected for morphological alteration. The MCC was defined as the compound concentration required to cause a microscopical alteration of the mock-infected cell cultures at day 3 post compound administration.

Molecular Modelling

The published X-ray structure of the TMPKmt (PDB entry 1G3U)[95] was used in all docking experiments. The inhibitors **5.2** and **5.10** were drawn using JChemPaint[166] and BUILD3D.[167] The molecular geometry was fed into gamess for geometry optimization using the AM1 force field.[168] The sugar conformation in **5.2** was modelled as C2'-endo with the base in an equatorial orientation by replacing the sugar and base by the sugar-base fragment in thymidine monophosphate from the pdb file 1G3U.[169] The sugar-base in **5.10** was replaced by the sugar-base in CSD[170] entry LEDRIV[171] having the base in alpha position and equatorial orientation. Polar hydrogen atoms were added to the enzyme and inhibitor structures using autodocktools.[172] The compounds were docked in the cavity close to Tyr-70 by means of the Autodock 3.05 software.[173] The top 50 of docked ligand conformations was examined and a manual selection procedure was used to validate the docked conformations.

Synthesis

3'-(Azidomethyl)-5'-*O*-*tert*-butyldiphenylsilyl-3'-deoxythymidine (5.34)

Methanesulfonyl chloride (0.366 mL, 4.73 mmol) was added to a solution of **2.43** (900 mg, 1.82 mmol) in pyridine (10 mL) at 0 °C. The reaction was stirred at 0 °C during 1 h. The mixture was diluted with CH_2Cl_2 (25 mL), washed with saturated aqueous $NaHCO_3$ and dried over $MgSO_4$. The solvent was evaporated *in vacuo* to give the crude mesylate. The obtained residue was dissolved in DMF (50 mL) and treated with NaN_3 (1.18 g, 18.2 mmol) at 60 °C. After 7 h the reaction was completed. The reaction mixture was evaporated to dryness and the residue was dissolved in CH_2Cl_2 (25 mL). The organic layer was washed with water, dried over $MgSO_4$ and evaporated to give a syrup which was purified by column chromatography (CH_2Cl_2/MeOH 99:1) yielding **5.34** (915 mg, 97%) as a white foam.
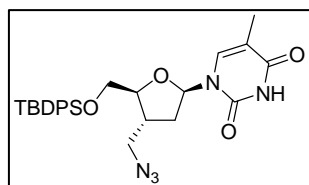
¹H NMR (300 MHz, DMSO-d_6): δ 1.00 (9H, s, *tert*-but), 1.50 (3H, d, *J*= 0.6 Hz, 5-CH_3), 2.15 (2H, m, H-2' and H-2"), 2.62 (1H, m, H-3'), 3.42-3.57 (2H, m, H-6' and H-6"), 3.77-3.85 (2H, m, H-4' and H-5'), 3.93 (1H, m, H-5"), 6.06 (1H, t, *J*= 6.6 Hz, H-1'), 7.36-7.45 (7H, m, 6 ar-H and H-6), 7.61-7.65 (4H, m, 4 arom H), 11.28 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{27}H_{33}N_5O_4SiNa$ [M+Na]⁺ found, 542.2197; calcd, 542.2199.

3'-(Aminomethyl)-5'-*O*-*tert*-butyldiphenylsilyl-3'-deoxythymidine (5.35)

A solution of **5.34** (915 mg, 1.76 mmol) in methanol (50 mL) was hydrogenated under atmospheric pressure for 5 hours in the presence of 10 % Pd/C (90 mg). The catalyst was removed by filtration through celite and the filtrate was evaporated to give pure compound **5.35** (860 mg, 99%).

¹H NMR (300 MHz, DMSO-d_6): δ 1.00 (9H, s, *tert*-but), 1.48 (3H, d, *J*= 0.9 Hz, 5-CH_3), 2.03-2.19 (2H, m, H-2' and H-2"), 2.32 (1H, m, H-3'), 2.57 (2H, dd, *J*= 3.9 and 5.7 Hz, H-6' and H-6"), 3.74-3.85 (2H, m, H-4' and H-5'), 3.93 (1H, dd, *J*= 2.4 and 10.8 Hz, H-5"), 6.02 (1H, dd, *J*= 5.1 and 6.9 Hz, H-1'), 7.37-7.43 (7H, m, 6 arom H and H-6), 7.61-7.65 (4H, m, 4 arom H).

Exact mass (ESI-MS) for $C_{27}H_{36}N_2O_4Si$ [M+H]⁺ found, 494.2478; calcd, 494.2474.

N-[(5'-*O-tert*-Butyldiphenylsilyl-3'-deoxythymidin-3'-yl)methyl]-*N'*-phenylthiourea (5.36)

To a solution of **5.35** (55 mg, 0.11 mmol) in DMF (1 mL), phenylisothiocyanate (15 mg, 0.11 mmol) in 1 mL DMF was added at 0 °C. The reaction mixture was stirred during 1 h. The solvent was evaporated to dryness and the residue was purified by column chromatography (CH$_2$Cl$_2$/MeOH 98:2) affording thiourea **5.36** (48 mg) in 70% yield.

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 0.99 (9H, s, *tert*-but), 1.45 (3H, s, 5-CH$_3$), 2.17 (2H, m, H-2' and H-2"), 2.83 (1H, m, H-3'), 3.51-3.67 (2H, m, H-6' and H-6"), 3.80-3.93 (3H, m, H-5', H-4' and H-5"), 6.08 (1H, t, *J*= 5.4, H-1'), 7.10 (1H, t, J= 7.2 Hz, 1 arom H), 7.26-7.43 (11H, m, 10 arom H and H-6), 7.61-7.66 (4H, m, 4 arom H), 7.93 (1H, br s, N(6')H), 9.55 (1H, br s, N(ar)H), 11.27 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{34}$H$_{41}$N$_4$O$_4$SSi [M+H]$^+$ found, 629.2618; calcd, 629.2617.

N-[(5'-*O-tert*-Butyldiphenylsilyl-3'-deoxythymidin-3'-yl)methyl]-*N'*-(4-chlorophenyl)-thiourea (5.37)

Thiourea **5.37** was synthesized from **5.35** (138 mg, 0.28 mmol) using the same procedure as described for the synthesis of **5.36**. The solid 4-chlorophenylisothiocyanate (52 mg, 0.3 mmol) was dissolved in DMF (1 mL) before adding it to the reaction mixture at 0 °C. After purification, compound **5.37** (143 mg) was obtained in 77% yield.

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 0.99 (9H, s, *tert*-but), 1.45 (3H, d, *J*= 0.6 Hz, 5-CH$_3$), 2.17 (2H, m, H-2' and H-2"), 2.83 (1H, m, H-3'), 3.51-3.67 (2H, m, H-6' and H-6"), 3.80 (1H, dd, *J*= 3.3 and 10.8 Hz, H-5'), 3.88-3.96 (2H, m, H-4' and H-5"), 6.08 (1H, t, *J*= 5.4, H-1'), 7.31-7.43 (11H, m, 10 arom H and H-6), 7.61-7.66 (4H, m, 4 arom H), 8.04 (1H, br s, N(6')H), 9.63 (1H, br s, N(ar)H), 11.28 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{34}$H$_{40}$N$_4$O$_4$SClSi [M+H]$^+$ found, 663.2230; calcd, 663.2227.

N-[(5'-*O-tert*-Butyldiphenylsilyl-3'-deoxythymidin-3'-yl)methyl]-*N'*-(4-methoxyphenyl) -thiourea (5.38)

Compound **5.38** was synthesized from **5.35** (130 mg, 0.26 mmol) using the same procedure as described for the synthesis of **5.36**. The solid 4-methoxyphenylisothiocyanate (48 mg, 0.29 mmol) was dissolved in DMF (1 mL) before adding it to the reaction mixture at 0 °C. After purification, compound **5.38** (150 mg) was obtained in 88% yield.
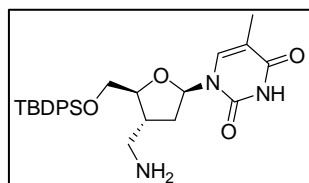
¹H NMR (300 MHz, DMSO-d₆): δ 0.99 (9H, s, *tert*-but), 1.45 (3H, d, J= 0.6 Hz, 5-CH₃), 2.15 (2H, m, H-2' and H-2''), 2.83 (1H, m, H-3'), 3.48-3.65 (2H, m, H-6' and H-6''), 3.71 (3H, s, OCH₃), 3.80 (1H, dd, J= 3.6 and 10.8 Hz, H-5'), 3.86-3.95 (2H, m, H-4' and H-5''), 6.07 (1H, t, J= 5.7, H-1'), 6.86 (2H, m, 2 arom H), 7.17 (2H, m, 2 arom H), 7.34-7.46 (7H, m, 6 arom H and H-6), 7.63-7.67 (4H, m, 4 arom H), 7.73 (1H, br s, N(6')H), 9.38 (1H, br s, N(ar)H), 11.28 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₃₅H₄₃N₄O₅SSi [M+H]⁺ found, 659.2726; calcd, 659.2723.

N-[(5'-*O*-*tert*-Butyldiphenylsilyl-3'-deoxythymidin-3'-yl)methyl]-*N'*-(4-methylphenyl)-thiourea (5.39)

Compound **5.39** was synthesized from **5.35** (160 mg, 0.32 mmol) using the same procedure as described for the synthesis of **5.36**. The solid 4-methylphenylisothiocyanate (53 mg, 0.36 mmol) was dissolved in DMF (1 mL) before adding it to the reaction mixture at 0 °C. After purification, compound **5.39** (164 mg) was obtained in 79% yield.
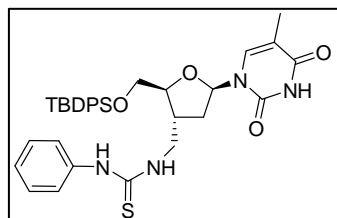

¹H NMR (300 MHz, DMSO-d₆): δ 0.99 (9H, s, *tert*-but), 1.44 (3H, d, J= 0.9 Hz, 5-CH₃), 2.17 (2H, m, H-2' and H-2''), 2.25 (3H, s, ar-CH₃), 2.87 (1H, m, H-3'), 3.48-3.65 (2H, m, H-6' and H-6''), 3.77-3.94 (3H, m, H-4', H-5' and H-5''), 6.08 (1H, t, J= 5.7, H-1'), 7.09 (2H, d, J= 9.2 Hz, 2 arom H), 7.21 (2H, d, J= 9.2 Hz, 2 arom H), 7.34-7.44 (7H, m, 6 arom H and H-6), 7.62-7.67 (4H, m, 4 arom H), 8.05 (1H, br s, N(6')H), 9.69 (1H, br s, N(ar)H), 11.31 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₃₅H₄₃N₄O₄SSi [M+H]⁺ found, 643.2776; calcd, 643.2774.

N-[(5'-*O*-*tert*-Butyldiphenylsilyl-3'-deoxythymidin-3'-yl)methyl]-*N'*-(3,4-dichlorophenyl)thiourea (5.40)

Compound **5.40** was synthesized from **5.35** (172 mg, 0.35 mmol) using the same procedure as described for the synthesis of **5.36**. After purification, compound **5.40** (192 mg) was obtained in 79% yield.
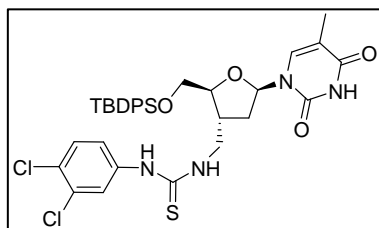
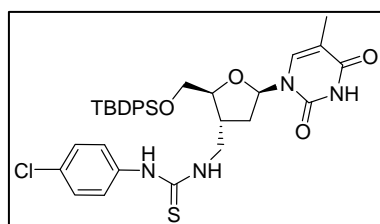**¹H NMR** (300 MHz, DMSO-d₆): δ 0.98 (9H, s, *tert*-but), 1.44 (3H, s, 5-CH₃), 2.19 (2H, m, H-2'), 2.85 (1H, m, H-3'), 3.50-3.69 (2H, m, H-6' and H-6"), 3.80 (1H, dd, *J*= 3.9 and 12.2 Hz, H-5'), 3.89-3.96 (2H, m, H-4' and H-5"), 6.09 (1H, t, *J*= 5.7, H-1'), 7.34-7.44 (8H, m, arom H), 7.52 (1H, d, J= 8.7 Hz, arom H), 7.63-7.67 (4H, m, 4 arom H), 7.89 (1H, s, H-6), 8.38 (1H, br s, N(6')H), 9.97 (1H, br s, N(ar)H), 11.31 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₃₄H₃₉N₄O₄SCl₂Si [M+H]⁺ found, 697.1838; calcd, 697.1838.

N-(4-(Benzyloxy)phenyl)-*N'*-[(5'-*O*-*tert*-butyldiphenylsilyl-3'-deoxythymidin-3'-yl)methyl]thiourea (5.41)

Compound **5.41** was synthesized from **5.35** (95 mg, 0.19 mmol) using the same procedure as described for the synthesis of **5.36**. The solid 4-(benzyloxy)phenylisothiocyanate (51 mg, 0.36 mmol) was dissolved in DMF (1 mL) before adding it to the reaction mixture at 0 °C. After purification, compound **5.41** (116 mg) was obtained in 83% yield.
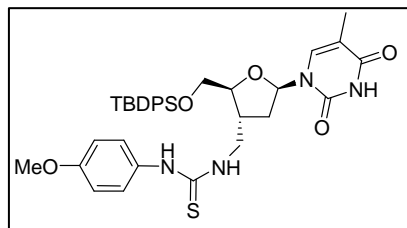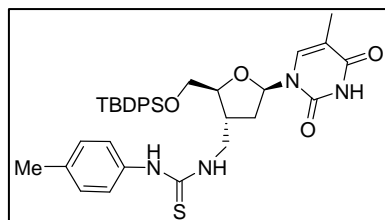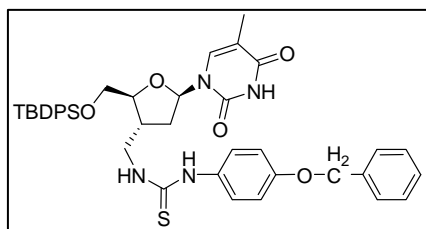

¹H NMR (300 MHz, DMSO-d₆): δ 0.99 (9H, s, *tert*-but), 1.44 (3H, d, *J*= 0.6 Hz, 5-CH₃), 2.17 (2H, m, H-2'), 2.88 (1H, m, H-3'), 3.46-3.64 (2H, m, H-6' and H-6"), 3.79-3.89 (3H, m, H-4', H-5' and H-5"), 5.07 (2H, s, *CH₂*Ph), 6.07 (1H, t, *J*= 5.7, H-1'), 6.94 (2H, dd, *J*= 2.4 and 7.2 Hz, 2 arom H), 7.10 (2H, dd, *J*= 2.4 and 6.9 Hz, 2 arom H), 7.25-7.37 (12H, m, arom H and H-6), 7.56-7.63 (4H, m, arom H), 7.76 (1H, br s, N(6')H), 9.43 (1H, br s, N(ar)H), 11.26 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₄₁H₄₇N₄O₅SSi [M+H]⁺ found, 735.3047; calcd, 735.3036.

N-[(5'-*O*-*tert*-Butyldiphenylsilyl-3'-deoxythymidin-3'-yl)methyl]-*N'*-(3-chloro-4-trifluoromethylphenyl)thiourea (5.42)

Compound **5.42** was synthesized from **5.35** (90 mg, 0.18 mmol) using the same procedure as described for the synthesis of **5.36**. After purification, compound **5.42** (100 mg) was obtained in 75% yield.
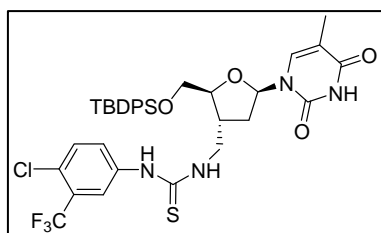
¹H NMR (300 MHz, DMSO-d₆): δ 0.99 (9H, s, *tert*-but), 1.44 (3H, s, 5-CH₃), 2.17 (2H, m, H-2'), 2.87 (1H, m, H-3'), 3.52-3.68 (2H, m, H-6' and H-6"), 3.80-3.96 (3H, m, H-4', H-5' and H-5"), 6.09 (1H, t, *J*= 5.7, H-1'), 7.34-7.48 (9H, m, 9 arom H), 7.63-7.74 (4H, m, 4 arom H), 8.06 (1H, s, H-6), 8.24 (1H, br s, N(6')H), 9.97 (1H, br s, N(ar)H), 11.23 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₃₅H₃₉N₄O₄SCl₂F₃Si [M+H]⁺ found, 731.2103; calcd, 731.2101.

N-[(3'-Deoxythymidin-3'-yl)methyl]-*N'*-phenylthiourea (5.2)

Compound **5.36** (48 mg, 0.08 mmol) was dissolved in THF (4 mL). A solution of 1 M tetra-*n*butylammonium fluoride in THF (4 mL) was added. After 1 h at room temperature the reaction was completed. The solvent was evaporated, and the dry residue was purified by column chromatography (CH_2Cl_2/MeOH 97:3) to give pure compound **5.2** (27 mg) in 87% yield.
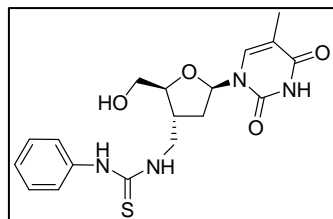
^{1}H NMR (300 MHz, DMSO-d_6): δ 1.75 (3H, s, 5-CH_3), 2.11 (2H, m, H-2' and H-2"), 2.57 (1H, m, H-3'), 3.51-3.78 (5H, m, H-4', H-5', H-5", H-6' and H-6"), 5.10 (1H, t, *J*= 5.1 Hz, 5'-OH), 5.99 (1H, t, *J*= 4.8, H-1'), 7.09 (1H, t, *J*= 7.2 Hz, arom H), 7.30 (2H, t, *J*= 7.5 Hz, 2 arom H), 7.38 (2H, d, *J*= 7.5 Hz, 2 arom H), 7.85 (1H, s, H-6), 7.95 (1H, br s, N(6')H), 9.60 (1H, br s, N(ar)H), 11.22 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d_6): δ 12.99 (5-CH_3), 36.73 (C-2'), 38.10 (C-3'), 45.85 (C-6'), 62.09 (C-5'), 84.57 and 84.63 (C-4' and C-1'), 109.38 (C-5), 124.02 (arom C), 125.03 (arom C), 129.34 (arom C), 137.03 (C-6), 139.76 (arom C), 151.07 (C-4), 164.54 (C-2), 181.39 (C=S).

Exact mass (ESI-MS) for $C_{18}H_{22}N_4O_4SNa$ [M+Na]$^+$ found, 413.1246; calcd, 413.1259.

N-(4-Chlorophenyl)-*N'*-[(3'-deoxythymidin-3'-yl)methyl]thiourea (5.3)

Compound **5.37** (143 mg, 0.22 mmol) was deprotected using the same procedure as described for the synthesis of compound **5.2** to afford analogue **5.3** (81 mg) in 89% yield.
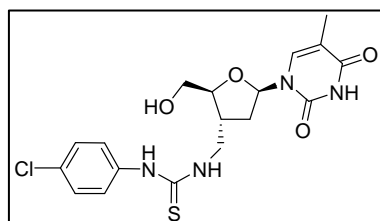
^{1}H NMR (300 MHz, DMSO-d_6): δ 1.75 (3H, d, *J*= 0.9 Hz, 5-CH_3), 2.12 (2H, m, H-2' and H-2"), 2.57 (1H, m, H-3'), 3.51-3.65 (2H, m, H-6' and H-6"), 3.67-3.78 (3H, m, H-4', H-5' and H-5"), 5.10 (1H, t, *J*= 5.1 Hz, 5'-OH), 6.00 (1H, t, *J*= 4.5, H-1'), 7.32-7.45 (4H, m, 4 arom H), 7.85 (1H, d, *J*= 1.2 Hz, H-6), 7.93 (1H, br s, N(6')H), 9.57 (1H, br s, N(ar)H), 11.22 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d_6): δ 12.97 (5-CH_3), 36.74 (C-2'), 38.08 (C-3'), 45.84 (C-6'), 62.07 (C-5'), 84.54 and 84.64 (C-4' and C-1'), 109.38 (C-5), 125.55 (arom C), 128.73 (arom C), 129.12 (arom C), 137.00 (C-6), 138.90 (arom C), 151.07 (C-4), 164.51 (C-2), 181.51 (C=S).

Exact mass (ESI-MS) for $C_{18}H_{21}N_4O_4SClNa$ [M+Na]$^+$ found, 447.0864; calcd, 447.0869.

N-[(3'-Deoxythymidin-3'-yl)methyl]-*N'*-(4-methoxyphenyl)thiourea (5.4)

Compound **5.38** (150 mg, 0.23 mmol) was deprotected using the same procedure as described for the synthesis of compound **5.2** to afford analogue **5.4** (87 mg) in 91% yield.
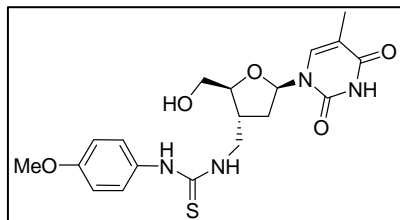
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.75 (3H, s, 5-CH$_3$), 2.09 (2H, m, H-2' and H-2''), 2.57 (1H, m, H-3'), 3.51-3.62 (2H, m, H-6' and H-6''), 3.66-3.78 (6H, m, OCH$_3$, H-4', H-5' and H-5''), 5.11 (1H, t, J= 5.1 Hz, 5'-OH), 5.99 (1H, t, J= 4.5, H-1'), 6.93 (2H, d, J= 8.7 Hz, 2 arom H), 7.20 (2H, dd, J= 2.7 and 9.0 Hz, 2 arom H), 7.70 (1H, br s, N(6')H), 7.86 (1H, s, H-6), 9.40 (1H, br s, N(ar)H), 11.25 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$):δ 12.94 (5-CH$_3$), 36.73 (C-2'), 38.10 (C-3'), 45.78 (C-6'), 55.93 (OCH$_3$), 62.20 (C-5'), 84.55 and 85.36 (C-4' and C-1'), 109.42 (C-5), 120.30 (arom C), 126.91 (arom C), 132.19 (arom C), 135.86 (arom C), 137.03 (C-6), 150.37 (C-4), 164.54 (C-2), 181.78 (C=S).

Exact mass (ESI-MS) for C$_{19}$H$_{25}$N$_4$O$_5$S [M+H]$^+$ found, 421.1539; calcd, 421.1545.

N-[(3'-Deoxythymidin-3'-yl)methyl]-*N'*-(4-methylphenyl)thiourea (5.5)

Compound **5.39** (164 mg, 0.26 mmol) was deprotected using the same procedure as described for the synthesis of compound **5.2** to afford analogue **5.5** (95 mg) in 92% yield.
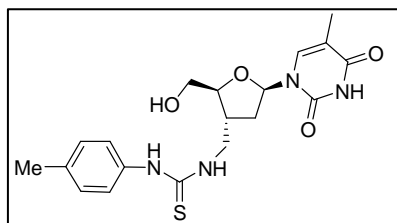
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.75 (3H, s, 5-CH$_3$), 2.11 (2H, m, H-2'), 2.26 (3H, s, ar-CH$_3$), 2.57 (1H, m, H-3'), 3.51-3.62 (2H, m, H-6' and H-6''), 3.66-3.78 (3H, m, H-4', H-5' and H-5''), 5.11 (1H, t, J= 5.1 Hz, 5'-OH), 5.99 (1H, t, J= 4.5, H-1'), 7.12 (2H, d, J= 8.1 Hz, 2 arom H), 7.21 (2H, d, J= 8.1 Hz, 2 arom H), 7.76 (1H, br s, N(6')H), 7.86 (1H, d, J= 1.2 Hz, H-6), 9.45 (1H, br s, N(ar)H), 11.25 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$):δ 12.99 (5-CH$_3$), 21.18 (ar-CH$_3$), 36.67 (C-2'), 38.16 (C-3'), 45.86 (C-6'), 62.13 (C-5'), 84.56 and 84.63 (C-4' and C-1'), 109.39 (C-5), 124.62 (arom C), 1298.89 (arom C), 134.47 (arom C), 137.03 (C-6), 151.08 (C-4), 164.54 (C-2), 181.42 (C=S).

Exact mass (ESI-MS) for C$_{19}$H$_{25}$N$_4$O$_4$S [M+H]$^+$ found, 405.1586; calcd, 405.1596.

N-(3,4-Dichlorophenyl)-*N'*-[(3'-deoxythymidin-3'-yl)methyl]thiourea (5.6)

Compound **5.40** (192 mg, 0.28 mmol) was deprotected using the same procedure as described for the synthesis of compound **5.2** to afford analogue **5.6** (109 mg) in 87% yield.

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.75 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 2.12 (2H, m, H-2'), 2.58 (1H, m, H-3'), 3.51-3.62 (2H, m, H-6' and H-6"), 3.66-3.79 (3H, m, H-4', H-5' and H-5"), 5.12 (1H, t, *J*= 5.1 Hz, 5'-OH), 6.00 (1H, t, *J*= 4.5, H-1'), 7.35 (1H, dd, *J*= 2.4 and 8.7 Hz, arom H), 7.53 (1H, d, *J*= 8.7 Hz, arom H), 7.86 (2H, m, H-6 and arom H), 8.16 (1H, br s, N(6')H), 9.78 (1H, br s, N(ar)H), 11.23 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 12.99 (5-CH$_3$), 36.82 (C-2'), 37.93 (C-3'), 45.81 (C-6'), 61.98 (C-5'), 84.55 and 84.64 (C-4' and C-1'), 109.37 (C-5), 123.58 (arom C), 124.74 (arom C), 130.98 (arom C), 131.22 (arom C), 137.00 (C-6), 140.29 (arom C), 151.07 (C-4), 164.53 (C-2), 181.41 (C=S).

Exact mass (ESI-MS) for C$_{18}$H$_{20}$N$_4$O$_4$SCl$_2$Na [M+Na]$^+$ found, 481.0480; calcd, 481.0487.

N-(4-(benzyloxy)phenyl)-*N'*-[(3'-deoxythymidin-3'-yl)methyl]thiourea (5.7)

Compound **5.41** (116 mg, 0.16 mmol) was deprotected using the same procedure as described for the synthesis of compound **5.2** to afford analogue **5.7** (69 mg) in 88% yield.
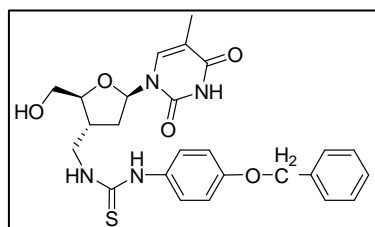

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.75 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 2.10 (2H, m, H-2'), 2.56 (1H, m, H-3'), 3.49-3.60 (2H, m, H-6' and H-6"), 3.64-3.79 (3H, m, H-4', H-5' and H-5"), 5.07 (2H, s, *CH$_2$*Ph), 5.09 (1H, t, J= 5.1 Hz, 5'-OH), 5.98 (1H, t, *J*= 4.5, H-1'), 6.95 (2H, dd, *J*= 2.4 and 7.2 Hz, 2 arom H), 7.20 (2H, dd, *J*= 2.4 and 6.9 Hz, 2 aromH), 7.31-7.45 (5H, m, CH$_2$*Ph*), 7.76 (1H, br s, N(6')H), 7.85 (1H, d, *J*= 1.2 Hz, H-6), 9.42 (1H, br s, N(ar)H), 11.22 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 12.99 (5-CH$_3$), 36.66 (C-2'), 38.18 (C-3'), 45.92 (C-6'), 62.17 (C-5'), 70.06 (*CH$_2$*Ph), 84.60 (C-4' and C-1'), 109.37 (C-5), 115.52 (arom C), 126.64 (arom C), 128.35 (arom C), 128.52 (arom C), 129.13 (arom C), 132.52 (arom C), 137.02 (C-6), 137.80 (arom C), 151.09 (C-4), 156.34 (arom C), 164.53 (C-2), 181.69 (C=S).

Exact mass (ESI-MS) for C$_{25}$H$_{28}$N$_4$O$_5$S [M+H]$^+$ found, 497.1851; calcd, 497.1858.

N-(3-Chloro-4-trifluoromethylphenyl)-*N'*-[(3'-deoxythymidin-3'-yl)methyl]thiourea (5.8)

Compound **5.42** (100 mg, 0.14 mmol) was deprotected using the same procedure as described for the synthesis of compound **5.2** to afford analogue **5.8** (60 mg) in 89% yield.
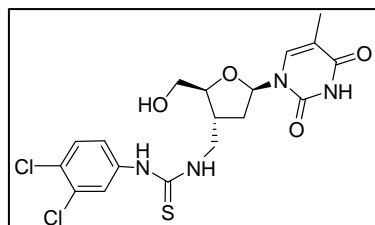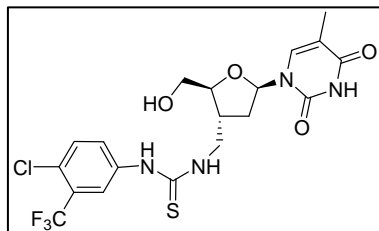

¹H NMR (300 MHz, DMSO-d₆): δ 1.75 (3H, d, *J*= 0.9 Hz, 5-CH₃), 2.13 (2H, m, H-2'), 2.58 (1H, m, H-3'), 3.51-3.63 (2H, m, H-6' and H-6"), 3.68-3.79 (3H, m, H-4', H-5' and H-5"), 5.11 (1H, s, 5'-OH), 6.00 (1H, t, *J*= 5.4, H-1'), 7.61 (1H, d, *J*= 8.7 Hz, arom H), 7.71 (1H, dd, *J*= 1.2 and 8.7 Hz, arom H), 7.86 (1H, s, arom H), 8.07 (1H, s, H-6), 8.24 (1H, br s, N(6')H), 9.90 (1H, br s, N(ar)H), 11.23 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.93 (5-CH₃), 36.81 (C-2'), 37.95 (C-3'), 45.78 (C-6'), 61.98 (C-5'), 84.52 and 84.69 (C-4' and C-1'), 109.41 (C-5), 110.00 (CF₃), 121.58 (arom C), 122.15 (arom C), 125.20 (arom C), 126.77 (arom C), 128.28 (arom C), 132.34 (arom C), 137.00 (C-6), 139.78 (arom C), 151.06 (C-4), 164.56 (C-2), 181.62 (C=S).

Exact mass (ESI-MS) for $C_{19}H_{21}N_4O_4SClF_3$ [M+H]⁺ found, 493.0920; calcd, 493.0924.

3'-5'-Di-*O*-acetyl-β-D-thymidine (5.43)

β-thymidine (4 g, 16.5 mmol) was dissolved in pyridine (70 mL) and acetic acid anhydride (70 mL). The reaction mixture was stirred at room temperature during 2 h, evaporated to dryness, and coevaporated with toluene to obtain pure compound **5.43** in a 99% yield (5.3 g).
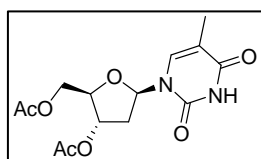
¹H NMR (300 MHz, CDCl₃): δ 1.94 (3H, d, *J*= 1.2 Hz, 5-CH₃), 2.11 (3H, s, COCH₃), 2.13 (3H, s, COCH₃), 2.12 (1H, dd, *J*= 6.6 and 8.4 Hz, H-2'), 2.43-2.51 (1H, ddd, *J*= 2.1, 5.7 and 14.1 Hz, H-2"), 4.24 (1H, m, H-4'), 4.35 (2H, m, H-5' and H-5"), 5.21 (1H, dt, *J*= 2.1 and 6.6 Hz, H-3'), 6.31 (1H, dd, *J*= 5.7 and 8.7, H-1'), 7.29 (1H, d, *J*= 1.2 Hz), 8.62 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{14}H_{19}N_2O_7$ [M+H]⁺ found, 327.1194; calcd, 327.1192.

3'-5'-Di-*O*-acetyl-α-D-thymidine (5.44)

To a solution of compound **5.43** (10 g, 41,28 mmol) in 20 mL dry acetonitrile, was added the freshly prepared solution, containing 0.710 mL H_2SO_4 and 2.70 mL acetic acid anhydride in 20 mL dry acetonitile. After 2 h, the mixture was quenched with saturated $NaHCO_3$-solution, and extracted with ethylacetate three times. The alpha anomer was crystallised from this mixture, giving 4.535 g (45%) pure **5.44**, and 5.1 g of an anomeric mixture.
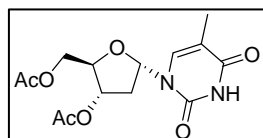
^{1}H NMR (300 MHz, $CDCl_3$): δ 1.95 (3H, d, *J*= 1.2 Hz, 5-CH_3), 2.05 (3H, s, $COCH_3$), 2.12 (3H, s, $COCH_3$), 2.21-2.28 (1H, ddd, *J*= 1.8 and 15.6 Hz, H-2'), 2.81 (1H, dt, *J*= 6.9 and 15.1 Hz, H-2"), 4.19 (2H, m, H-5' and H-5"), 4.59 (1H, m, H-4'), 5.23 (1H, dt, *J*= 1.6 and 6.4 Hz, H-3'), 6.26 (1H, dd, *J*= 2.4 and 7.2 Hz, H-1'), 7.30 (1H, d, *J*= 1.2 Hz), 8.62 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{14}H_{19}N_2O_5$ [M+H]$^+$ found, 327.1195; calcd, 327.1192.

α-D-Thymidine (5.45)

3'-5'-Di-*O*-acetyl-α-thymidine (4.54 g, 13.91 mmol) was dissolved in 150 mL NH_3 in MeOH (7N solution). The reaction mixture was stirred at room temperature overnight. The solvent was evaporated to dryness, the residue was dissolved in ethylacetate and extracted with water three times, yielding 3.30 g of the pure compound **5.45** (98%).
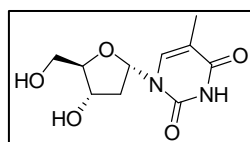
^{1}H NMR (300 MHz, DMS0): δ 1.78 (3H, d, *J*= 1.2 Hz, 5-CH_3), 1.84-1.91 (1H, ddd, *J*= 3.0 and 14.4 Hz, H-2'), 2.48-2.60 (1H, m, H-2"), 3.39 (2H, t, *J*= 5.2 Hz, H-5' and H-5"), 4.13 (1H, m, H-3'), 4.23 (1H, m, H-4'), 4.81 (1H, t, *J*= 5.7 Hz, 5'-OH), 5.30 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.12 (1H, dd, *J*= 3.3 and 7.5 Hz, H-1'), 7.75 (1H, d, *J*= 1.2 Hz), 11.20 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{10}H_{15}N_2O_5$ [M+H]$^+$ found, 243.0975; calcd, 243.0981.

α-D-Thymidine 5'-monophosphate (5.9)

A solution of α-D-thymidine (150 mg, 0.62 mmol) in trimethyl phosphate (6.2 mL) was cooled to 0 °C, $POCl_3$ (369 μL, 4.03 mmol) was added dropwise and the mixture was stirred for 4 hours at 0 °C and for 30 minutes at room temperature. The mixture was poured into ice-water (12 mL), neutralised with concentrated NH_4OH and evaporated to dryness. The resulting residue was purified by column chromatography (iPrOH/NH_4OH/H_2O 77.5:15:2.5). Further purification was performed by HPLC (C-18, CH_3CN/MeOH/0.05% HCOOH in H_2O 45:45:10, 3 mL/min). After lyophilisation of the collected pure fractions compound **5.9** was obtained (123 mg, 62%) as a white powder.
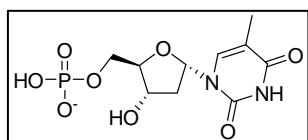
^{1}H NMR (300 MHz, D_2O-d_6): δ 1.81 (3H, d, *J*= 0.9 Hz, 5-CH_3), 2.02-2.09 (2H, ddd, *J*= 3.0 and 14.7 Hz, H-2'), 2.66-2.76 (1H, m, H-2"), 3.78 (2H, app t, *J*= 5.1 Hz, H-5' and H-5"), 4.43 (2H, m, H-3' and H-4'), 6.12 (1H, dd, *J*= 3.0 and 7.2 Hz, H-1'), 7.68 (1H, d, *J*= 0.9 Hz, H-6).

^{31}P-NMR (500 MHz, D_2O): δ 2.97.

^{13}C NMR (75 MHz, D_2O-d_6): δ 11.73 (5-CH_3), 39.49 (C-2'), 64.67 (C-5'), 71.20 (C-3'), 87.38 and 87.81 (C-4' and C-1'), 110.75 (C-5), 138.39 (C-6), 151.76 (C-4), 166.97.

Exact mass (ESI-MS) for $C_{10}H_{14}N_2O_8PNa$ [M+Na]$^+$ found, 345.0477; calcd, 345.0464; Anal. ($C_{10}H_{14}N_2O_8P$) C, H, N.

5'-Azido-5'-deoxy-α-D-thymidine (5.46)

To a solution of α-thymidine **5.45** (886 mg, 3.66 mmol) in pyridine (13.5 mL) at -78 °C, methanesulfonylchloride (256 μL, 0.41 mmol) was added. The reaction mixture was stirred for 1 h at 0 °C. The reaction was quenched with saturated aqueous $NaHCO_3$-solution and extracted with CH_2Cl_2 three times, dried over $MgSO_4$ and evaporated. The residue was purified by column chromatography (CH_2Cl_2/MeOH 98:2) to give the mesylated compound (916 mg, 78%).
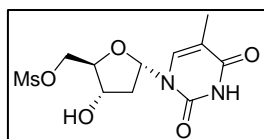
^{1}H NMR (300 MHz, DMSO-d_6): 1.75 (3H, s, 5-CH_3), 1.95 (1H, m, H-2'), 2.55 (1H, m, H-2"), 3.19 (3H, s, CH_3SO_2), 4.12-4.26 (3H, m, H-5', H-5" and H-3'), 4.33 (1H, m, H-4'), 5.58 (1H, br s, 3'-OH), 6.12 (1H, dd, *J*= 4.4 and 7.6 Hz, H-1'), 7.73 (1H, s, H-6), 11.27 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{11}H_{17}N_2O_7S$ [M+H]$^+$ found, 321.0759; calcd, 321.0756.

A solution of 5'-mesylated α-D-thymidine (916 mg, 2.88 mmol) and NaN$_3$ (1.87 g, 29 mmol) in DMF (50 mL) was heated to 60 °C overnight. The reaction mixture was evaporated *in vacuo.* The residue was resolved in CH$_2$Cl$_2$ and washed with H$_2$O. The organic layer was dried over MgSO$_4$, evaporated and purified by column chromatography (CH$_2$Cl$_2$/MeOH 98:2) to afford compound **5.46** (672 mg, 87%).
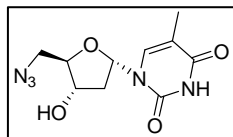
^{1}H NMR (300 MHz, DMSO-d$_6$): 1.76 (3H, s, 5-CH$_3$), 1.93 (1H, m, H-2'), 2.51 (1H, m, H-2"), 3.49 (2H, m, H-5' and H-5"), 4.12 (1H, dd, H-3'), 4.24 (1H, m, H-4'), 5.50 (1H, br s, 3'-OH), 6.13 (1H, dd, *J*= 4.5 and 7.5 Hz, H-1'), 7.72 (1H, s, H-6), 11.26 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{10}$H$_{14}$N$_5$O$_4$ [M+H]$^+$ found, 268.1045; calcd, 268.1046.

5'-Amino-5'-deoxy-α-D-thymidine (5.47)

A solution of azide **5.46** (531 mg, 1.99 mmol) in methanol (30 mL) was hydrogenated under atmospheric pressure for 6 hours in the presence of 10% Pd/C (53.1 mg). The catalyst was removed by filtration through celite and the filtrate was evaporated to yield pure compound **5.47** (471 mg, 98%).
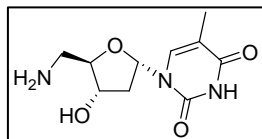
^{1}H NMR (300 MHz, DMSO-d$_6$): 1.74 (3H, s, 5-CH$_3$), 1.85 (1H, m, H-2'), 2.49 (1H, m, H-2"), 3.43 (2H, m, H-5' and H-5"), 4.02 (1H, m, H-3'), 4.14 (1H, m, H-4'), 5.27 (1H, br s, 3'-OH), 6.05 (1H, dd, *J*= 3.3 and 7.5 Hz, H-1'), 7.69 (1H, s, H-6), 11.24 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_{10}$H$_{16}$N$_3$O$_4$ [M+H]$^+$ found, 242.1134; calcd, 242.1141.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-phenylthiourea (5.10)

For the synthesis of compound **5.10**, compound **5.47** (54 mg, 0.22 mmol) was dissolved in DMF (2 mL). At 0 °C, phenyl isothiocyanate (36 mg, 0.26 mmol) was added and the reaction mixture was allowed to stir at room temperature during 3 h. After completion of the reaction, the reaction mixture was evaporated to dryness and the residue was purified by column chromatography (CH$_2$Cl$_2$/MeOH 97:3) to obtain the pure final compound **5.10** (69 mg, 83%).
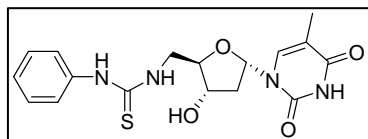
¹H NMR (300 MHz, DMSO-d₆): δ 1.77 (3H, d, *J*= 1.2 Hz, 5-CH₃), 1.89-1.98 (1H, ddd, *J*= 3.3 and 14.1 Hz, H-2'), 2.53-2.63 (1H, m, H-2"), 3.49 (1H, m, H-5'), 3.64 (1H, m, H-5"), 4.22 (1H, m, H-3'), 4.34 (1H, m, H-4'), 5.47 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.18 (1H, dd, *J*= 3.6 and 8.1, H-1'), 7.10 (1H, t, *J*= 7.2 Hz, arom H), 7.32 (2H, t, *J*= 7.5 Hz, 2 arom H), 7.77 (2H, d, *J*= 1.2 Hz, H-6 + N(5')H), 9.63 (1H, br s, N(ar)H), 11.26 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 13.03 (5-CH₃), under DMSO (C-2'), 46.18 (C-5'), 71.58 (C-3'), 85.40 (C-1'), 86.57 (C-4'), 109.51 (C-5), 123.83 (arom C), 124.97 (arom C), 129.29 (arom C), 137.68 (C-6), 139.84 (arom C), 151.21 (C-4), 164.55 (C-2), 181.41 (C=S).

Exact mass (ESI-MS) for $C_{17}H_{21}N_4O_4S$ [M+H]⁺ found, 377.1279; calcd, 377.1283; Anal. ($C_{17}H_{20}N_4O_4S$) C, H, N, S.

N'-(4-Chlorophenyl)- *N*-(5'-Deoxy-α-D-thymidin-5'-yl)thiourea (5.11)

Compound **5.11** was synthesized from **5.47** (49 mg, 0.20 mmol) and 4-chlorophenyl isothiocyanate (40 mg, 0.24 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH₂Cl₂/MeOH 97:3), compound **5.11** (66 mg, 81%) was obtained.

¹H NMR (300 MHz, DMSO-d₆): δ 1.76 (3H, d, *J*= 0.9 Hz, 5-CH₃), 1.88-1.95 (1H, ddd, *J*= 3.0 and 14.4 Hz, H-2'), 2.53-2.63 (1H, m, H-2"), 3.48 (1H, m, H-5'), 3.65 (1H, m, H-5"), 4.22 (1H, m, H-3'), 4.33 (1H, m, H-4'), 5.48 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.19 (1H, dd, *J*= 3.3 and 7.8, H-1'), 7.34 (2H, dd, *J*= 2.4 and 8.7 Hz, 2 arom H), 7.48 (2H, dd, *J*= 2.4 and 9.3 Hz, 2 arom H), 7.77 (1H, d, *J*= 1.2 Hz, H-6), 7.93 (1H, br s, N(5')H), 9.77 (1H, br s, N(ar)H), 11.25 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 13.01 (5-CH₃), under DMSO (C-2'), 46.13 (C-5'), 71.59 (C-3'), 85.43 (C-1'), 86.45 (C-4'), 109.56 (C-5), 125.41 (arom C), 128.73 (arom C), 129.10 (arom C), 137.69 (C-6), 138.92 (arom C), 151.22 (C-4), 164.60 (C-2), 181.50 (C=S).

Exact mass (ESI-MS) for $C_{17}H_{20}N_4O_5SCl$ [M+H]⁺ found, 411.0889; calcd, 411.0893; Anal. ($C_{17}H_{19}N_4O_5S$) C, H, N, S.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(4-methoxyphenyl)thiourea (5.12)

Compound **5.12** was synthesized from **5.47** (57 mg, 0.24 mmol) and 4-methoxyphenyl isothiocyanate (47 mg, 0.29 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH₂Cl₂/MeOH 97:3), compound **5.12** (84 mg, 86%) was obtained.

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.78 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 1.88-1.93 (1H, ddd, *J*= 3.0 and 14.1 Hz, H-2'), 2.53-2.63 (1H, m, H-2''), 3.44-3.60 (2H, m, H-5' and H-5''), 3.72 (3H, s, OCH$_3$), 4.22 (1H, m, H-3'), 4.33 (1H, m, H-4'), 5.45 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.16 (1H, dd, *J*= 3.0 and 7.5, H-1'), 6.87 (2H, dd, *J*= 2.1 and 9.0 Hz, 2 arom H), 7.23 (2H, dd, *J*= 2.1 and 8.7 Hz, 2 arom H), 7.64 (1H, br s, N(5')H), 7.76 (1H, d, *J*= 1.5 Hz, H-6), 9.51 (1H, br s, N(ar)H), 11.22 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 13.05 (5-CH$_3$), 43.02 (C-2'), 46.22 (C-5'), 55.92 (OCH$_3$), 71.52 (C-3'), 85.41 (C-1'), 86.69 (C-4'), 109.45 (C-5), 114.60 (arom C), 126.56 (arom C), 132.37 (arom C), 137.69 (C-6), 151.21 (C-4), 157.26 (arom C), 164.55 (C-2), 181.74 (C=S).

Exact mass (ESI-MS) for C$_{18}$H$_{23}$N$_4$O$_5$S [M+H]$^+$ found, 407.1389; calcd, 407.1389; Anal. (C$_{18}$H$_{22}$N$_4$O$_5$S) C, H, N, S.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(4-methylphenyl)thiourea (5.13)

Compound **5.13** was synthesized from **5.47** (56 mg, 0.23 mmol) and 4-methoxyphenyl isothiocyanate (41 mg, 0.27 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 97:3), compound **5.13** (74 mg, 82%) was obtained.

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.78 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 1.89-1.96 (1H, ddd, *J*= 3.0 and 14.1 Hz, H-2'), 2.27 (3H, s, ar-CH$_3$), 2.54-2.64 (1H, m, H-2''), 3.49 (1H, m, H-5'), 3.63 (1H, m, H-5''), 4.23 (1H, m, H-3'), 4.36 (1H, m, H-4'), 5.47 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.19 (1H, dd, *J*= 3.3 and 7.5, H-1'), 7.13 (2H, d, *J*= 8.1 Hz, 2 arom H), 7.27 (2H, d, *J*= 8.7 Hz, 2 arom H), 7.67 (1H, br s, N(5')H), 7.78 (1H, d, *J*= 1.2 Hz, H-6), 9.55 (1H, br s, N(ar)H), 11.28 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 12.93 (5-CH$_3$), 21.16 (*p*-CH$_3$), under DMSO (C-2'), 46.17 (C-5'), 71.55 (C-3'), 85.42 (C-1'), 86.63 (C-4'), 109.49 (C-5), 124.30 (arom C), 129.82 (arom C), 134.38 (arom C), 137.07 (arom C), 137.68 (C-6), 151.21 (C-4), 164.57 (C-2), 181.44 (C=S).

Exact mass (ESI-MS) for C$_{18}$H$_{23}$N$_4$O$_4$S [M+H]$^+$ found, 391.1431; calcd, 391.1439; Anal. (C$_{18}$H$_{22}$N$_4$O$_4$S) C, H, N, S.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(3,4-dichlorophenyl)thiourea (5.14)

Compound **5.14** was synthesized from **5.47** (80 mg, 0.33 mmol) and phenyl isothiocyanate (82 mg, 0.40 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH_2Cl_2/MeOH 97:3), compound **5.14** (147 mg, 82%) was obtained.
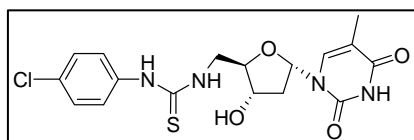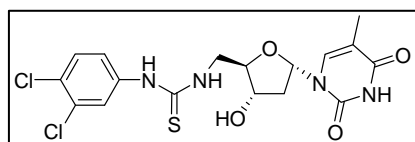

^{1}H NMR (300 MHz, DMSO-d_6): δ 1.77 (3H, d, *J*= 1.2 Hz, 5-CH_3), 1.89-1.97 (1H, ddd, *J*= 3.0 and 14.1 Hz, H-2'), 2.54-2.63 (1H, m, H-2''), 3.48 (1H, m, H-5'), 3.65 (1H, m, H-5''), 4.21 (1H, m, H-3'), 4.33 (1H, m, H-4'), 5.48 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.19 (1H, dd, *J*= 3.6 and 7.8, H-1'), 7.36 (1H, dd, *J*= 2.7 and 8.7 Hz, arom H), 7.53 (1H, d, *J*= 8.7 Hz, arom H), 7.77 (1H, d, *J*= 1.2 Hz, H-6), 7.94 (1H, d, J= 2.1 Hz, arom H), 8.02 (1H, br s, N(5')H), 9.83 (1H, br s, N(ar)H), 11.26 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d_6): δ 12.96 (5-CH_3), under DMSO (C-2'), 46.15 (C-5'), 71.61 (C-3'), 85.52 (C-1'), 86.31 (C-4'), 109.62 (C-5), 123.51 (arom C), 124.65 (arom C), 124.67 (arom C), 130.96 (arom C), 131.23 (arom C), 137.73 (C-6), 140.21 (arom C), 151.21 (C-4), 164.73 (C-2), 181.40 (C=S).

Exact mass (ESI-MS) for $C_{17}H_{19}N_4O_4SCl_2$ [M+H]$^+$ found, 445.0503; calcd, 445.0503; Anal. ($C_{17}H_{18}N_4O_5$) C, H, N.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(3-trifluoromethyl-4-chlorophenyl)thiourea (5.15)

Compound **5.15** was synthesized from **5.47** (59 mg, 0.25 mmol) and 4-chloro-3-trifluoromethylphenyl isothiocyanate (70 mg, 0.29 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH_2Cl_2/MeOH 97:3), compound **5.15** (107 mg, 91%) was obtained.
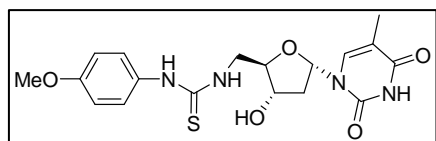

^{1}H NMR (300 MHz, DMSO-d_6): δ 1.78 (3H, d, *J*= 0.9 Hz, 5-CH_3), 1.91-1.99 (1H, ddd, *J*= 3.3 and 14.4 Hz, H-2'), 2.56-2.66 (1H, m, H-2''), 3.50 (1H, m, H-5'), 3.69 (1H, m, H-5''), 4.23 (1H, m, H-3'), 4.36 (1H, m, H-4'), 5.49 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.21 (1H, dd, *J*= 3.6 and 7.8, H-1'), 7.63 (1H, d, *J*= 9.0 Hz, arom H), 7.74 (1H, dd, *J*= 2.7 and 8.9 Hz, arom H), 7.79 (1H, d, *J*= 1.2 Hz, H-6), 8.17 (2H, br s, arom H and N(5')H), 10.05 (1H, br s, N(ar)H), 11.27 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d_6): δ 13.01 (5-CH_3), under DMSO (C-2'), 46.23 (C-5'), 71.63 (C-3'), 85.37 (C-1'), 86.25 (C-4'), 109.58 (C-5), 116.26 (CF_3), 121.59 (arom C), 121.92 (arom C), 125.20 (arom C), 128.00 (arom C), 132.33 (arom C), 137.63 (C-6), 139.85 (arom C), 151.22 (C-4), 164.55 (C-2), 181.47 (C=S).

Exact mass (ESI-MS) for $C_{18}H_{19}N_4O_4SClF_3$ [M+H]$^+$ found, 479.0766; calcd, 479.0767; Anal. ($C_{18}H_{18}N_4O_4SClF_3$) C, H, N, S.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(4-morpholinophenyl)thiourea (5.16)

Compound **5.16** was synthesized from **5.47** (100 mg, 0.42 mmol) and morpholinophenyl isothiocyanate (110 mg, 0.50 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 98:2), compound **5.16** (152 mg, 80%) was obtained.
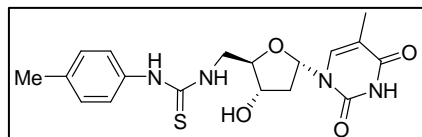

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.78 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 1.89-1.96 (1H, ddd, *J*= 3.2 and 14.4 Hz, H-2'), 2.53-2.63 (1H, m, H-2"), 3.07 (4H, t, *J* = 5.3 Hz, 2x C*H$_2$*N), 3.49 (1H, m, H-5'), 3.59 (1H, m, H-5"), 3.73 (4H, t, *J* = 5.1 Hz, 2x C*H$_2$*O), 4.23 (1H, m, H-3'), 4.45 (1H, m, H-4'), 5.44 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.17 (1H, dd, *J*= 3.3 and 7.7, H-1'), 6.90 (2H, d, *J* = 9.0 Hz, 2 arom H), 7.20 (2H, d, *J* = 9.0 Hz, 2 arom H), 7.55 (1H, br s, N(5')H), 7.77 (1H, *J* = 1.2 Hz, H-6), 9.42 (1H, br s, N(ar)H), 11.21 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 13.03 (5-CH$_3$), 46.22 (C-2'), 46.32 (*CH$_2$*N), 49.32 (C-5'), 66.80 (*CH$_2$*O), 71.53 (C-3'), 85.40 (C-1'), 86.73 (C-4'), 109.43 (C-5), 115.93 (arom C), 125.93 (arom C), 131.13 (arom C), 137.68 (C-6), 149.16 (arom C), 151.20 (C-4), 164.51 (C-2), 181.57 (C=S).

Exact mass (ESI-MS) for C$_{21}$H$_{28}$N$_5$O$_5$S [M+H]$^+$ found, 462.1803; calcd, 462.1732; Anal. (C$_{21}$H$_{27}$N$_5$O$_5$S) C, H, N, S.

N-(Adamant-1-yl)- *N'*-(5'-deoxy-α-D-thymidin-5'-yl)thiourea (5.17)

Compound **5.17** was synthesized from **5.47** (120 mg, 0.50 mmol) and 4-methoxyphenyl isothiocyanate (114 mg, 0.60 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 99:1), compound **5.17** (178 mg, 82%) was obtained.


^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.60 (6H, m, CH and CH$_2$), 1.76 (3H, d, *J* = 1.2 Hz, 5-CH$_3$), 1.86-1.94 (1H, ddd, *J*= 3.1 and 14.1 Hz, H-2'), 2.00 (2H, m, 2.53-2.63, CH$_2$), 2.12 (5H, m, C*H* and C*H$_2$*), 2.49 (1H, m, H-2"), 3.38 (1H, m, H-5'), 3.56 (1H, m, H-5"), 4.19 (2H, m, H-3' and H-4'), 5.42 (1H, d, *J*= 2.8 Hz, 3'-OH), 6.16 (1H, dd, *J*= 3.4 and 7.6, H-1'), 7.14 (1H, br s, N(ar)H), 7.30 (1H, t, *J* = 5.4 Hz, N(5')H), 7.75 (1H, *J* = 1.2 Hz, H-6), 11.21 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 12.96 (5-CH$_3$), 29.67, 36.62, 41.82 (*C*H and *C*H$_2$), 45.36 (C-2'), 53.46 (C-5'), 71.59 (C-3'), 85.42 (C-1'), 87.10 (C-4'), 109.58 (C-5), 137.66 (C-6), 151.22 (C-4), 164.61 (C-2), 181.49 (C=S).

Exact mass (ESI-MS) for C$_{21}$H$_{31}$N$_4$O$_4$S [M+H]$^+$ found, 435.2065; calcd, 435.2065; Anal. (C$_{21}$H$_{30}$N$_4$O$_4$S) C, H, N, S.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(3-pyridyl)thiourea (5.18)

Compound **5.18** was synthesized from **5.47** (40 mg, 0.17 mmol) and 3-pyridyl isothiocyanate (27 mg, 0.20 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH_2Cl_2/MeOH 95:5), compound **5.18** (52 mg, 83%) was obtained.
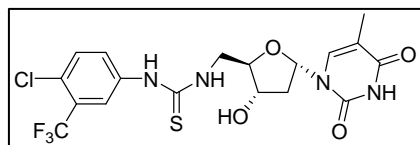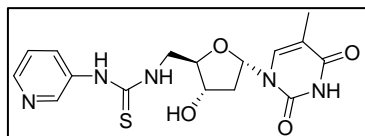

¹H NMR (300 MHz, DMSO-d_6): δ 1.79 (3H, d, *J*= 0.9 Hz, 5-CH_3), 1.90-1.99 (1H, ddd, *J*= 3.0 and 14.1 Hz, H-2'), 2.54-2.64 (1H, m, H-2"), 3.50 (1H, m, H-5'), 3.65 (1H, m, H-5"), 4.23 (1H, m, H-3'), 4.34 (1H, m, H-4'), 5.47 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.19 (1H, dd, *J*= 3.6 and 7.8, H-1'), 7.32 (1H, dd, *J*= 4.5 and 8.7 Hz, arom H), 7.77 (1H, d, *J*= 1.2 Hz, H-6), 7.95 (1H, d, *J*= 8.1 Hz, arom H), 8.08 (1H, br s, N(5')H), 8.27 (1H, dd, *J*= 1.5 and 4.5 Hz, arom H), 8.55 (1H, d, *J*= 2.4 Hz, arom H), 9.80 (1H, br s, N(ar)H), 11.22 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d_6): δ 13.02 (5-CH_3), 44.30 (C-2'), 46.33 (C-5'), 71.64 (C-3'), 85.39 and 86.44 (C-1' and C-4'), 109.56 (C-5), 123.86 (arom C), 125.16 (arom C), 135.26 (arom C), 136.95 (arom C), 137.65 (C-6), 145.61 (arom C), 151.23 (C-4), 164.53 (C-2), 182.11 (C=S).

Exact mass (ESI-MS) for $C_{16}H_{20}N_5O_4S$ [M+H]⁺ found, 378.1229; calcd, 378.1235; Anal. ($C_{16}H_{19}N_5O_4S$) C, H, N, S.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(fluoresceinyl)thiourea (5.19)

Compound **5.19** was synthesized from **5.47** (70 mg, 0.29 mmol) and fluoresceine isothiocyanate (135 mg, 0.35 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH_2Cl_2/MeOH 95:5), compound **5.19** (129 mg, 71%) was obtained.

¹H NMR (300 MHz, DMSO-d_6): δ 1.78 (3H, d, *J*= 0.9 Hz, 5-CH_3), 1.92-2.00 (1H, ddd, *J*= 3.3 and 14.4 Hz, H-2'), 2.57-2.68 (1H, m, H-2"), 3.57 (1H, m, H-5'), 3.72 (1H, m, H-5"), 4.27 (1H, m, H-3'), 4.38 (1H, m, H-4'), 5.53 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.23 (1H, dd, *J*= 3.3 and 7.2 H-1'), 6.54-6.66 (1H, m, 6 arom H), 7.18 (1H, d, *J*= 8.4 Hz, arom H), 7.75 (1H, d, *J*= 6.6 Hz, arom H), 7.80 (1H, d, *J* = 1.2 Hz, H-6), 8.19 (1H, br s, N(5')H), 8.31 (1H, s, arom H), 10.12 (1H, br s, N(ar)H), 11.23 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d_6): δ 13.10 (5-CH_3), 42.96 (C-2'), 46.87 (C-5'), 72.74 (C-3'), 82.84 and 89.14 (C-1' and C-4'), 84.91 (*C*OPh₃), 103.29 (arom C), 106.30 (arom C), 110.98 (C-5), 115.49 (arom C), 123.86 (arom C), 123.97 (arom C), 126.36 (arom C), 129.32 (arom C), 133.76 (arom C), 139.95 (arom C), 136.65 (C-6), 151.69 (C-4), 152.15 (arom C), 152.94 (arom C), 153.29 (arom C), 164.33 (C-2), 181.18 (C=S).

Exact mass (ESI-MS) for $C_{31}H_{27}N_4O_9S$ [M+H]⁺ found, 631.1498; calcd, 631.1498; Anal. ($C_{31}H_{26}N_4O_9S$) C, H, N, S.

N-Benzyl- *N'*-(5'-deoxy-α-D-thymidin-5'-yl)thiourea (5.20)

Compound **5.20** was synthesized from **5.47** (100 mg, 0.42 mmol) and benzyl isothiocyanate (74 mg, 0.50 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH₂Cl₂/MeOH 97:3), compound **5.20** (120 mg, 74%) was obtained.
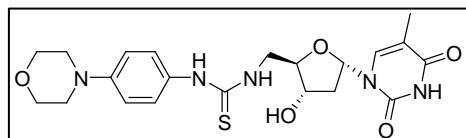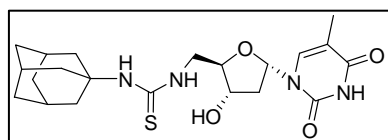

¹H NMR (300 MHz, DMSO-d₆): δ 1.78 (3H, d, *J*= 1.2 Hz, 5-CH₃), 1.87-1.96 (1H, ddd, *J*= 3.0 and 14.5 Hz, H-2'), 2.49-2.60 (1H, m, H-2"), 3.49 (1H, m, H-5'), 3.60 (1H, m, H-5"), 4.21 (1H, m, H-3'), 4.28 (1H, m, H-4'), 4.66 (2H, d, *J*= 4.8 Hz, C*H*₂NH), 5.43 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.17 (1H, dd, *J*= 3.6 and 7.7 Hz, H-1'), 7.21-7.35 (5H, m, 5 arom H), 7.58 (1H, br s, N(5')H), 7.76 (1H, d, *J*= 1.2 Hz, H-6), 7.98 (1H, br s, CH₂N*H*), 11.23 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 13.01 (5-CH₃), 36.45 (*C*H₂NH), 46.24 (C-2'), 47.77 (C-5'), 71.56 (C-3'), 85.42 (C-1'), 87.03 (C-4'), 109.48 (C-5), 127.54 (arom C), 127.98 (arom C), 128.96 (arom C), 137.61 (C-6), 139.86 (arom C), 151.19 (C-4), 164.51 (C-2), 181.91 (C=S).

Exact mass (ESI-MS) for C₁₈H₂₃N₄O₄S [M+H]⁺ found, 391.1440; calcd, 391.1439; Anal. (C₁₈H₂₂N₄O₄S) C, H, N, S.

N-(Benzhydryl)-*N*-(5'-deoxy-α-D-thymidin-5'-yl)thiourea (5.21)

Compound **5.21** was synthesized from **5.47** (70 mg, 0.29 mmol) and benzhydryl isothiocyanate (78 mg, 0.35 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH₂Cl₂/MeOH 97:3), compound **5.21** (114 mg, 84%) was obtained.
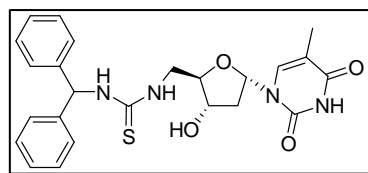

¹H NMR (300 MHz, DMSO-d₆): δ 1.76 (3H, s, 5-CH₃), 1.87-1.94 (1H, ddd, *J*= 3.4 and 14.2 Hz, H-2'), 2.52-2.61 (1H, m, H-2"), 3.47 (1H, m, H-5'), 3.66 (1H, m, H-5"), 4.17 (1H, m, H-3'), 4.25 (1H, m, H-4'), 5.42 (1H, d, *J*= 3.6 Hz, 3'-OH), 6.18 (1H, dd, *J*= 3.9 and 7.5 Hz, H-1'), 6.67 (1H, d, *J*= 8.2 Hz, C*H*NH), 7.20-7.35 (10H, m, 10 arom H), 7.55 (1H, br s, N(5')H), 7.75 (1H, s, H-6), 8.45 (1H, d, *J*= 8.4 Hz, CHN*H*), 11.21 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.99 (5-CH₃), 46.39 (C-2'), 49.28 (C-5'), 61.36 (*C*HNH), 71.61 (C-3'), 85.36 (C-1'), 86.98 (C-4'), 109.60 (C-5), 127.66 (arom C), 127.70 (arom C), 127.88 (arom C), 129.13 (arom C), 137.58 (C-6), 143.18 (arom C), 151.21 (C-4), 164.49 (C-2), 183.34 (C=S).

Exact mass (ESI-MS) for C₂₄H₂₇N₄O₄S [M+H]⁺ found, 467.1757; calcd, 467.1752; Anal. (C₂₄H₂₆N₄O₄S) C, H, N, S.

N-Benzoyl-N'-(5'-deoxy-α-D-thymidin-5'-yl)thiourea (5.22) and **5'-benzoylamino-5'-deoxy-α-D-thymidine (5.30)**

Treatment of amine **47** (140 mg, 0.58 mmol) and benzoyl isothiocyanate (114 mg, 0.70 mmol) using the same procedure as described for the synthesis of **5.10**, afforded, after purification by column chromatography (CH₂Cl₂/MeOH 97:3), the expected thiourea **5.22** (120 mg, 51 %) as well as benzamide **5.30** (39 mg, 20%).
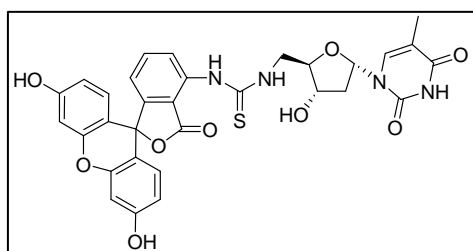

¹H NMR (300 MHz, DMSO-d₆): δ 1.76 (3H, d, *J*= 0.9 Hz, 5-CH₃), 1.90-1.98 (1H, ddd, *J*= 2.9 and 14.5 Hz, H-2'), 2.56-2.67 (1H, m, H-2"), 3.63 (1H, m, H-5'), 3.82 (1H, m, H-5"), 4.23 (1H, m, H-3'), 4.43 (1H, m, H-4'), 5.50 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.20 (1H, dd, *J*= 3.3 and 7.5, H-1'), 7.49 (2H, m, 2 arom H), 7.62 (1H, m, arom H), 7.76 (1H, *J* = 0.9 Hz, H-6), 7.91 (2H, m, 2 arom H), 11.00 (1H, t, *J*= 5.1 Hz, N(5')H), 11.23 (1H, br s, N(ar)H), 11.38 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 13.03 (5-CH₃), 41.77 (C-2'), 47.17 (C-5'), 71.72 (C-3'), 85.59 and 85.55 (C-1' and C-4'), 109.57 (C-5), 129.10 (arom C), 129.25 (arom C), 132.84 (arom C), 133.71 (arom C), 137.62 (C-6), 151.18 (C-4), 164.54 (C-2), 168.86 (C=O), 181.42 (C=S).

Exact mass (ESI-MS) for C₁₈H₂₀N₄O₅SNa [M+Na]⁺ found, 427.1051; calcd, 427.1052; Anal. (C₁₈H₂₀N₄O₅S) C, H, N, S.
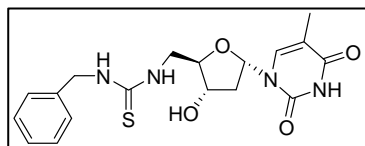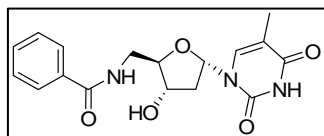

¹H NMR (300 MHz, DMSO-d₆): δ 1.75 (3H, d, *J*= 1.2 Hz, 5-CH₃), 1.87-1.94 (1H, ddd, *J*= 2.5 and 14.4 Hz, H-2'), 2.51-2.62 (1H, m, H-2"), 3.30 (2H, m, H-5' and H-5"), 4.22-4.32 (2H, m, H-3' and H-4'), 5.39 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.16 (1H, dd, *J*= 2.8 and 7.8, H-1'), 7.42-7.52 (3H, m, 3 arom H), 7.75 (1H, *J* = 1.3 Hz, H-6), 7.84 (2H, m, 2 arom H), 8.57 (1H, t, *J*= 5.9 Hz, N(5')H), 11.20 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 13.04 (5-CH₃), 41.96 (C-2'), 49.12 (C-5'), 71.72 (C-3'), 85.54 (C-1'), 87.24 (C-4'), 109.33 (C-5), 127.91 (arom C), 128.99 (arom C), 131.97 (arom C), 134.91 (arom C), 137.69 (C-6), 151.18 (C-4), 164.56 (C-2), 167.21 (C=O).

Exact mass (ESI-MS) for C₁₇H₁₉N₃O₅Na [M+Na]⁺ found, 368.1224; calcd, 368.1222; Anal. (C₁₇H₁₉N₃O₅) C, H, N.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-N'-(3-trifluoromethyl-4-chlorobenzyl)thiourea (5.23)

4-Chloro-3-trifluoromethylbenzylamine (65 mg, 0.31 mmol) was added at 0 °C to a stirred solution of 1,1'-thiocarbonyldiimidazole (61 mg, 0.34 mmol) and imidazole (6.3 mg, 0.09 mmol) in 4 mL acetonitril. After 10 minutes at 0 °C, the reaction was allowed to stir during 3 hours at room temperature. A solution of **5.47** (75 mg, 0.31 mmol) in 2 mL DMF was added and the reaction mixture was stirred at room temperature during overnight. The reaction

mixture was evaporated to dryness and purified by column chromatography (CH$_2$Cl$_2$/MeOH 99:1) to obtain compound **5.23** (114 mg, 75%).
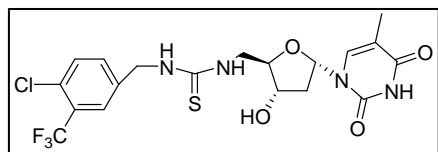
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.77 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 1.87-1.95 (1H, ddd, *J*= 2.9 and 14.4 Hz, H-2'), 2.50-2.60 (1H, m, H-2"), 3.47 (1H, m, H-5'), 3.55 (1H, m, H-5"), 4.18 (1H, m, H-3'), 4.26 (1H, m, H-4'), 4.71 (2H, d, *J*= 4.8 Hz, C*H$_2$*NH), 5.41 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.14 (1H, dd, *J*= 3.2 and 7.6 Hz, H-1'), 7.56 (1H, dd, *J* = 2.0 and 8.8 Hz, arom H), 7.65 (1H, d, *J* = 8.7 Hz, arom H), 7.73 (2H, app s, arom H and H-6), 7.78 (1H, t, *J* = 4.7 Hz, N(5')H), 8.14 (1H, d, *J*= 4.9 Hz, CH$_2$N*H*), 11.20 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 13.00 (5-CH$_3$), 31.47 (*C*H$_2$NH), 46.48 (C-2'), 46.58 (C-5'), 71.58 (C-3'), 85.46 (C-1'), 86.98 (C-4'), 109.44 (C-5), 116.28 (*C*F$_3$), 127.10 (arom C), 127.17 (arom C), 129.49 (arom C), 132.14 (arom C), 133.57 (arom C), 137.58 (C-6), 140.54 (arom C), 151.19 (C-4), 164.52 (C-2), 182.28 (C=S).

Exact mass (ESI-MS) for C$_{19}$H$_{21}$N$_4$O$_4$SClF$_3$ [M+H]$^+$ found,; calcd, 493.0924; Anal. (C$_{19}$H$_{20}$N$_4$O$_4$SClF$_3$) C, H, N, S.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(phenethyl)thiourea (5.24)

Compound **5.24** was synthesized from **5.47** (108 mg, 0.45 mmol) and phenethyl isothiocyanate (130 mg, 0.54 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 97:3), compound **5.24** (137 mg, 76%) was obtained.
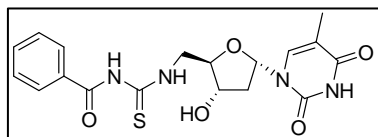

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.77 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 1.87-1.95 (1H, ddd, *J*= 3.0 and 14.3 Hz, H-2'), 2.49-2.60 (1H, m, H-2"), 2.80 (2H, t, *J*= 7.5 Hz, C*H$_2$*Ph), 3.49 (1H, m, H-5'), 3.60 (2H, m, H-5" and C*H$_2$*NH), 4.19-4.29 (2H, m, H-3' and H-4'), 5.39 (1H, d, *J*= 2.8 Hz, 3'-OH), 6.14 (1H, dd, *J*= 3.4 and 7.7 Hz, H-1'), 7.16-7.31 (5H, m, 5 arom H), 7.52 (2H, m, N(5')H and CH$_2$N*H*), 7.74 (1H, d, *J*= 1.2 Hz, H-6), 11.20 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 13.01 (5-CH$_3$), 35.47 (*C*H$_2$Ph), 45.89 (*C*H$_2$NH), 46.09 (C-2'), 47.82 (C-5'), 71.56 (C-3'), 85.43 (C-1'), 87.10 (C-4'), 109.46 (C-5), 126.80 (arom C), 129.04 (arom C), 129.34 (arom C), 137.63 (C-6), 140.01 (arom C), 151.20 (C-4), 164.51 (C-2), 184.28 (C=S).

Exact mass (ESI-MS) for C$_{19}$H$_{25}$N$_4$O$_4$S [M+H]$^+$ found, 405.1595; calcd, 405.1596; Anal. (C$_{19}$H$_{24}$N$_4$O$_4$S) C, H, N, S.

N-(3,4-Dichlorophenethyl)- N'-(5'-deoxy-α-D-thymidin-5'-yl)thiourea (5.25)

3,4-dichlorophenethylamine (59 mg, 0.31 mmol) was added at 0 °C to a stirred solution of 1,1'-thiocarbonyldiimidazole (61 mg, 0.34 mmol) and imidazole (6.3 mg, 0.09 mmol) in 4 mL acetonitril. After 10 minutes at 0 °C, the reaction was allowed to stir during 3 hours at room temperature. A solution of **5.47** (75 mg, 0.31 mmol) in 2 mL DMF was added and the reaction mixture was stirred at room temperature during overnight. The mixture was evaporated and after purification by column chromatography (CH$_2$Cl$_2$/MeOH 97:3), compound **5.25** (120 mg, 82%) was obtained.
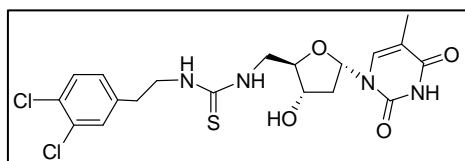
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.75 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 1.85-1.93 (1H, ddd, *J*= 3.0 and 14.7 Hz, H-2'), 2.48-2.57 (1H, m, H-2"), 2.80 (2H, t, *J*= 7.1 Hz, C*H*$_2$Ph), 3.49 (1H, m, H-5'), 3.61 (2H, m, H-5" and C*H*$_2$NH), 4.19-4.31 (2H, m, H-3' and H-4'), 5.40 (1H, d, *J*= 3.0 Hz, 3'-OH), 6.14 (1H, dd, *J*= 3.2 and 7.6 Hz, H-1'), 7.21 (1H, dd, *J* = 2.1 and 8.1 Hz, arom H), 7.49 (1H, d, *J* = 1.8 Hz, arom H), 7.51-7.54 (1H, d, *J* = 8.2 Hz, arom H), 7.53 (2H, br s, CH$_2$N*H* and N(5')H), 7.74 (1H, d, *J*= 1.2 Hz, H-6), 11.20 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 13.02 (5-CH$_3$), 34.37 (*C*H$_2$Ph), 45.19 (*C*H$_2$NH), 46.13 (C-2'), 47.82 (C-5'), 71.56 (C-3'), 85.42 (C-1'), 87.07 (C-4'), 109.46 (C-5), 129.44 (arom C), 129.91 (arom C), 131.08 (arom C), 131.42 (arom C), 131.50 (arom C), 137.63 (C-6), 141.38 (arom C), 151.20 (C-4), 164.51 (C-2), 183.65 (C=S).

Exact mass (ESI-MS) for C$_{19}$H$_{23}$N$_4$O$_4$SCl$_2$ [M+H]$^+$ found, ; calcd, 473.0816; Anal. (C$_{19}$H$_{22}$N$_4$O$_4$SCl$_2$) C, H, N, S.

N-(3,4-Dichlorophenyl)-N'-(5'-deoxy-α-D-thymidin-5'-yl)urea (5.26)

Compound **5.26** was synthesized from **5.47** (85 mg, 0.35 mmol) and 3,4-dichlorophenyl isocyanate (79 mg, 0.42 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 97:3), compound **5.26** (113 mg, 75%) was obtained.
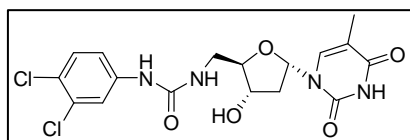
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.78 (3H, s, 5-CH$_3$), 1.89-1.97 (1H, ddd, *J*= 3.2 and 14.3 Hz, H-2'), 2.51-2.61 (1H, m, H-2"), 3.09 (1H, m, H-5'), 3.21 (1H, m, H-5"), 4.16 (2H, m, H-3' and H-4'), 5.42 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.15 (1H, dd, *J*= 3.6 and 7.8, H-1'), 6.44 (1H, t, *J*= 5.1 Hz, N(5')H), 7.22 (1H, dd, *J*= 2.4 and 9.0 Hz, arom H), 7.43 (1H, d, *J*= 9.0 Hz, arom H), 7.75 (1H, d, *J*= 0.9 Hz, H-6), 7.82 (1H, d, *J*= 2.4 Hz, arom H), 8.92 (1H, br s, N(ar)H), 11.24 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.99 (5-CH₃), 41.77 (C-2'), 49.28 (C-5'), 71.57 (C-3'), 85.35 (C-1'), 87.32 (C-4'), 109.54 (C-5), 118.40 (arom C), 119.36 (arom C), 122.97 (arom C), 131.12 (arom C), 131.61 (arom C), 137.60 (C-6), 141.33 (arom C), 151.21 (C-4), 155.56 (C=O), 164.51 (C-2).

Exact mass (ESI-MS) for C₁₇H₁₈N₄O₅Cl₂Na [M+Na]⁺ found, 451.0548; calcd, 451.0552; Anal. (C₁₇H₁₈N₄O₅Cl₂) C, H, N.

N-(5'-Deoxy-α-D-thymidin-5'-yl)-*N'*-(3-trifluoromethyl-4-chlorophenyl)urea (5.27)

Compound **5.27** was synthesized from **5.47** (79 mg, 0.33 mmol) and 4-chloro-3-trifluorophenyl isocyanate (88 mg, 0.40 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH₂Cl₂/MeOH 97:3), compound **5.27** (120 mg, 79%) was obtained.
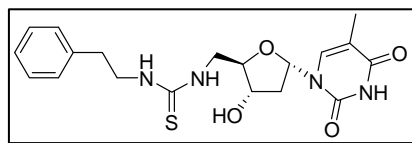

¹H NMR (300 MHz, DMSO-d₆): δ 1.78 (3H, d, *J*= 1.2 Hz, 5-CH₃), 1.89-1.97 (1H, ddd, *J*= 3.1 and 14.3 Hz, H-2'), 2.52-2.62 (1H, m, H-2"), 3.15 (1H, m, H-5'), 3.24 (1H, m, H-5"), 4.18 (2H, m, H-3' and H-4'), 5.44 (1H, d, *J*= 3.3 Hz, 3'-OH), 6.17 (1H, dd, *J*= 3.6 and 7.8, H-1'), 6.53 (1H, t, *J*= 5.1 Hz, N(5')H), 7.55 (2H, app s, 2 arom H), 7.77 (1H, d, *J*= 1.2 Hz, H-6), 8.06 (1H, s, arom H), 9.14 (1H, br s, N(ar)H), 11.23 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 13.00 (5-CH₃), 41.81 (C-2'), 49.35 (C-5'), 71.51 (C-3'), 85.31 (C-1'), 87.26 (C-4'), 109.53 (C-5), 116.57 (*C*F₃), 122.11 (arom C), 123.00 (arom C), 123.05 (arom C), 128.05 (arom C), 132.56 (arom C), 137.60 (C-6), 140.66 (arom C), 151.20 (C-4), 155.57 (C=O), 164.51 (C-2).

Exact mass (ESI-MS) for C₁₈H₁₉N₄O₅ClF₃ [M+H]⁺ found, 463.0993; calcd, 463.0995; Anal. (C₁₈H₁₈N₄O₅ClF₃) C, H, N.

5'-Azido-5'-deoxy-α-D-thymidine 3'-methanesulfonate (5.48)

Compound **5.46** (400 mg, 1.5 mmol) was dissolved in pyridine (5 mL) and methanesulphonylchloride (150 μL, 1.95 mmol) was added at 0 °C. After 2h, the reaction was quenched by saturated NaHCO₃-solution (5 mL) and extracted with CH₂Cl₂ (3x 10 mL). The combined organic layers were dried with MgSO₄ and evaporated to dryness to obtain pure compound **5.48** (420 mg, 81%).
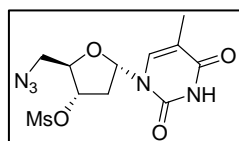
¹H NMR (300 MHz, DMSO-d₆): δ 1.77 (3H, d, *J*= 1.0 Hz, 5-CH₃), 2.32-2.40 (1H, ddd, *J*= 3.6 and 15.0 Hz, H-2'), 2.81-2.91 (1H, m, H-2"), 3.27 (3H, s, SO₂CH₃), 3.52 (2H, d, *J*= 5.1 Hz, H-5' and H-5"), 4.70 (1H, m, H-4'), 5.17 (1H, m, H-3'), 6.16 (1H, dd, *J*= 3.6 and 6.9, H-1'), 7.52 (1H, d, J= 1.1 Hz, H-6), 11.31 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₁₁H₁₅N₅O₆S [M+H]⁺ found, 346.0818; calcd, 346.0821.

5'-Azido-3',5'-dideoxy-2',3'-didehydro-α-D-thymidine (5.49)

To a solution of compound **5.48** (420 mg, 1.21 mmol) in THF (15 mL) was added DBU (906 µL, 6.07 mmol), and the reaction was refluxed at 80 °C during overnight. After cooling down, the reaction was poured into NH$_4$Cl (15 mL) solution, and extracted with CH$_2$Cl$_2$ (3 x 25 mL). The organic layers were dried over MgSO$_4$, evaporated and purified by column chromatography (CH$_2$Cl$_2$/MeOH 99:1) to yield 290 mg (95%) pure compound **5.49**.
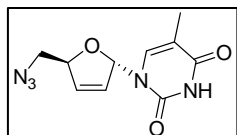
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.77 (3H, d, J= 1.0 Hz, 5-CH$_3$), 3.36 (1H, dd, J= 4.8 and 13.2 Hz, H-5'), 3.62 (1H, dd, J= 3.3 and 13.2 Hz, H-5"), 5.32 (1H, m, H-4'), 6.03 (1H, dd, J= 1.5 and 5.7, H-1'), 6.40 (1H, dd, J= 1.5 and 6.0 Hz, H-2"), 6.89 (1H, dd, J= 1.5 and 5.4 Hz, H-3'), 7.52 (1H, d, J= 1.1 Hz, H-6), 11.31 (1H, br s, N(3)H).

Exact mass(ESI-MS) for C$_{10}$H$_{11}$N$_5$O$_3$Na [M+Na]$^+$ found, 272.0761; calcd, 272.0759.

5'-Amino-3',5'-dideoxy-α-D-thymidine (5.50)

To a solution of azide **5.49** (150 mg, 0.6 mmol) in methanol (4 mL) was added 10% Pd/C (15 mg) and the reaction placed under a H$_2$ atmosphere. After 5 hours, the reaction mixture was filtered through a plug of Celite, which was further washed with MeOH and the combined filtrates were evaporated to yield pure **5.50** as a white solid (128 mg, 95%).
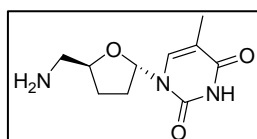
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.70 (1H, m, H-3'), 1.77 (3H, d, J= 1.2 Hz, 5-CH$_3$), 1.89-2.06 (2H, m, H-2' and H-3"), 2.27 (1H, m, H-2"), 2.56 (2H, d, J= 5.1 Hz, H-5' and H-5"), 4.27 (1H, m, H-4'), 5.98 (1H, dd, J= 5.1 and 6.6, H-1'), 7.41 (1H, d, J= 1.1 Hz, H-6).

Exact mass (ESI-MS) for C$_{10}$H$_{16}$N$_3$O$_3$ [M+H]$^+$ found, 226.1188; calcd, 226.1191.

N-(3-Trifluoromethyl-4-chlorophenyl)-*N'*-(3',5'-dideoxy-α-D-thymidin-5'-yl)thiourea (5.28)

Compound **5.28** was synthesized from **5.50** (78 mg, 0.35 mmol), and 4-chloro-3-trifluoromethylphenyl isothiocyanate (100 mg, 0.42 mmol) using the same procedure described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 97:3), compound **5.28** (127 mg, 78%) was obtained.
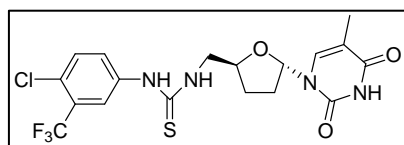
¹H NMR (300 MHz, DMSO-d₆): δ 1.78 (3H, d, *J*= 1.2 Hz, 5-CH₃), 1.76 (1H, m, H-3'), 1.97-2.19 (2H, m, H-2' and H-3"), 2.22-2.39 (1H, m, H-2"), 3.52 (1H, m, H-5'), 3.67 (1H, m, H-5"), 4.57 (1H, m, H-4'), 6.10 (1H, dd, *J*= 5.4 and 6.0, H-1'), 7.48 (1H, d, *J*= 1.2 Hz, arom H), 7.60 (1H, d, *J*= 8.7 Hz, arom H), 7.73 (1H, dd, *J*= 2.4 and 8.7 Hz, arom H), 8.19 (1H, s, H-6), 8.32 (1H, br s, N(5')H), 10.15 (1H, br s, N(ar)H), 11.21 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.76 (5-CH₃), 27.99 (C-3'), 31.45 (C-2'), 48.26 (C-5'), 79.22 (C-4'), 86.11 (C-1'), 110.17 (C-5), 117.86 (*C*F₃), 121.02 (arom C), 125.21 (arom C), 127.32 (arom C), 128.28 (arom C), 136.95 (C-6), 139.90 (arom C), 151.14 (C-4), 164.56 (C-2), 180.98 (C=S).

Exact mass (ESI-MS) for $C_{18}H_{18}N_4O_3SClF_3Na$ [M+Na]⁺ found, 485.0641; calcd, 485.0638; Anal. ($C_{18}H_{18}N_4O_3SClF_3$) C, H, N, S.

5'-Amino-3',5'-dideoxy-2',3'-didehydro-α-D-thymidine (5.51)

Compound **5.49** (80 mg, 0.32 mmol) was dissolved in dry pyridine (4 mL) and triphenylphospine (135 mg, 1.6 mmol) was added to the solution. The reaction was stirred during 3 hours at room temperature, evaporated to dryness and purified by column chromatography (CH₂Cl₂/MeOH 90:10), yielding 65 mg pure compound **5.51** (92%).

¹H NMR (300 MHz, DMSO-d₆): δ 1.69 (3H, d, *J* = 1.2 Hz, 5-CH₃), 2.59 (1H, dd, *J* = 4.7 and 13.5 Hz, H-5'), 2.70 (1H, dd, *J* = 4.7 and 13.7 Hz, H-5"), 4.99 (1H, m, H-4'), 5.83 (1H, app dt, *J* = 2.4 and 5.7, H-1'), 6.33 (1H, app dt, *J* = 1.5 and 6.0 Hz, H-2'), 6.81 (1H, app dt, *J* = 1.5 and 4.9 Hz, H-3'), 7.10 (1H, d, *J*= 1.2, H-6).

¹³C NMR (75 MHz, DMSO-d₆): δ 11.96 (5-CH₃), 45.68 (C-5'), 88.49 (C-1'), 89.41 (C-4'), 109.75 (C-5), 125.37 (C-2'), 135.72 (C-3'), 137.42 (C-6), 150.53 (C-4), 163.75 (C-2).

Exact mass (ESI-MS) for $C_{10}H_{14}N_3O_3$ [M+H]⁺ found, 224.1039; calcd, 224.1035.

N-(3',5'-dideoxy-2',3'-didehydro-α-D-thymidin-5'-yl)-*N*-(3-trifluoromethyl-4-chlorophenyl)thiourea (5.29)

Compound **5.29** was synthesized from **5.51** (70 mg, 0.31 mmol) and 4-chloro-3-trifluoromethylphenyl isothiocyanate (90 mg, 0.38 mmol) using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH₂Cl₂/MeOH 99:1), compound **5.29** (112 mg, 78%) was obtained.
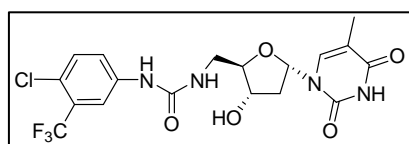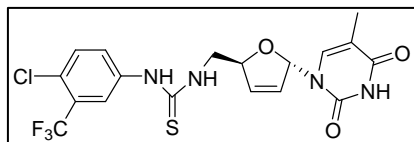
¹H NMR (300 MHz, DMSO-d₆): δ 1.75 (3H, d, *J*= 1.2 Hz, 5-CH₃), 3.73 (2H, m, H-5' and H-5"), 4.29 (1H, m, H-4'), 5.97 (1H, dt, *J*= 2.4 and 6.0, H-1'), 6.45 (1H, dt, *J*= 1.8 and 6.0 Hz, H-2'), 6.98 (1H, dt, *J*= 1.8 and 5.4 Hz, H-3'), 7.13 (1H, d, *J*= 1.2 Hz, arom H), 7.61 (1H, d, *J*= 9.0 Hz, arom H), 7.72 (1H, dd, *J*= 2.4 and 8.4 Hz, arom H), 8.07 (1H, br s, N(5')H), 8.17 (1H, s, H-6), 9.96 (1H, br s, N(ar)H), 11.30 (1H, br s, N(3)H).

¹³C NMR (75 MHz, DMSO-d₆): δ 12.66 (5-CH₃), 47.80 (C-5'), 85.82 (C-4'), 90.11 (C-1'), 110.71 (C-5), 115.26 (*C*F₃), 121.59 (arom C), 125.21 (arom C), 126.67 (arom C), 126.84 (C-2'), 127.07 (arom C), 127.85 (arom C), 132.29 (C-3'), 135.44 (arom C), 136.42 (C-6), 139.87 (arom C), 151.34 (C-4), 164.49 (C-2), 181.68 (C=S).

Exact mass (ESI-MS) for C₁₈H₁₆N₄O₃SClF₃Na [M+Na]⁺ found, ; calcd, 483.0481; Anal. (C₁₈H₁₆N₄O₃SClF₃) C, H, N, S.

1-[(2-Hydroxyethoxy)methyl]thymine (5.52)

A mixture of thymine (400 mg, 3.17 mmol), HMDS (16 mL) and ammoniumsulphate (16 mg) was refluxed during 2 hours under nitrogen. After cooling to room temperature, the mixture was evaporated to dryness and resolved in acetonitrile (32 mL). At -45 °C trimethylsilyltriflate (602 μL, 3.33 mmol) was added dropwise, followed by the dropwise addition of dioxolane (470 mg, 6.34 mmol). After 2 hours the reaction was allowed to warm to room temperature and stirred overnight. To quench the reaction, a saturated aqueous NaHCO₃-solution (20 mL) was added at -45 °C. The resulting mixture was extracted three times with diethylether (25 mL). The combined organic layers were dried over MgSO₄ and evaporated to dryness. The crude residue was purified by column chromatography (CH₂Cl₂/MeOH 95:5) to obtain 480 mg (76%) of pure compound **5.52**.
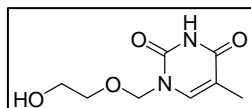
¹H NMR (300 MHz, DMSO-d₆): δ 1.75 (3H, d, *J*= 1.2 Hz, 5-CH₃), 3.45 (4H, m, CH₂CH₂), 4.61 (1H, t, *J*=5.1 Hz, OH), 5.03 (2H, s, OCH₂N), 7.54 (1H, d, *J*= 1.2, H-6), 11.26 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C₈H₁₂N₂O₄ [M+Na]⁺ found, 223.0721; calcd, 223.0694.

1-[(2-Azidoethoxy)methyl]thymine (5.53)

Compound **5.52** (480 mg, 1.82 mmol) was dissolved in anhydrous pyridine (14 mL) and methanesulfonylchloride (183 μL, 2.36 mmol) was slowly added at 0 °C. After 2 hours, the reaction was quenched with aqueous NaHCO₃-solution (10 mL) and the mixture was extracted three times with ethylacetate (20 mL). The combined organic layers were dried

over MgSO$_4$ and evaporated to dryness. The obtained residue was dissolved in DMF (40 mL) and 1.46 g (22.5 mmol) of NaN$_3$ was added. The reaction mixture was heated at 60 °C during the night, evaporated and purified by column chromatography (CH$_2$Cl$_2$/MeOH 97:3), yielding 393 mg (96%) of **5.53**.
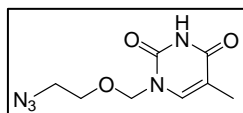
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.75 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 3.39 (2H, t, *J*= 4.8 Hz, CH$_2$O), 4.64 (2H, t, *J*= 4.8 Hz, CH$_2$N$_3$), 5.07 (2H, s, OCH$_2$N), 7.56 (1H, d, *J*= 1.2 Hz, H-6), 11.30 (1H, br s, N(3)H).

Exact mass (ESI-MS) for C$_8$H$_{12}$N$_5$O$_3$ [M+H]$^+$ found, 226.0948; calcd, 226.0940.

1-[(2-Aminoethoxy)methyl]thymine (5.54)

To a solution of azide **5.53** (393 mg, 1.75 mmol) in methanol (15 mL) was slowly added 50 mg of 10% Pd/C. The reaction mixture was submitted to a hydrogen atmosphere overnight. After filtration of the catalyst over a plug of celite and evaporation of the solvent in vacuo, resulting compound **5.54** (345 mg, 99%) was obtained and used without further purification.
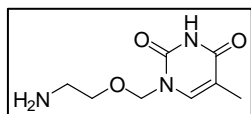
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.77 (3H, s, 5-CH$_3$), 2.65 (2H, t, *J*= 5.7 Hz, CH$_2$NH$_2$), 3.43 (2H, t, *J*= 4.8 Hz, CH$_2$O), 5.05 (2H, s, OCH$_2$N), 7.48 (1H, s, H-6).

Exact mass (ESI-MS) for C$_8$H$_{14}$N$_3$O$_3$ [M+H]$^+$ found, 200.1034; calcd, 200.1035.

N-[(Thymin-1-yl)methoxyethyl]-*N'*-phenylthiourea (5.31)

Compound **5.31** was synthesized from **5.54** (65 mg, 0.33 mmol) and phenyl isothiocyanate (57 mg, 0.42 mmol) in 2 mL DMF using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 99:1), thiourea **5.31** (94 mg, 86%) was obtained as a white powder.
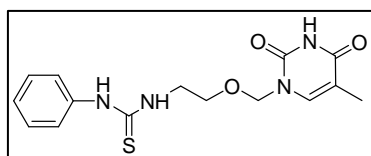
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.74 (3H, d, *J*= 0.9 Hz, 5-CH$_3$), 3.62 (4H, app s, CH$_2$CH$_2$), 5.06 (2H, s, OCH$_2$N), 7.07 (1H, m, arom H), 7.28 (2H, m, 2 arom H), 7.37 (2H, d, *J*= 7.5 Hz, 2 arom H), 7.56 (1H, d, *J*= 1.2 Hz, H-6), 7.76 (1H, br s, N(5')H), 9.61 (1H, br s, N(ar)H), 11.28 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 12.56 (5-CH$_3$), 44.10 (*C*H$_2$NH'), 67.35 (*C*H$_2$O), 76.84 (O*C*H$_2$N), 109.92 (C-5), 123.79 (arom C), 124.93 (arom C), 129.33 (arom C), 139.77 (C-6), 141.23 (arom C), 151.78 (C-4), 164.98 (C-2), 181.18 (C=S).

Exact mass (ESI-MS) for C$_{15}$H$_{19}$N$_4$O$_3$S [M+H]$^+$ found, 335.1175; calcd, 335.1177.

N-[(Thymin-1-yl)methoxyethyl]-*N'*-(4-chloro-3-trifluoromethylphenyl)thiourea (5.32)

Compound **5.32** was synthesized from **5.54** (97 mg, 0.49 mmol) and 4-chloro-3-trifluoromethylphenyl isothiocyanate (138 mg, 0.58 mmol) in 3 mL DMF using the same procedure as described for the synthesis of **5.10**. After purification by column chromatography (CH$_2$Cl$_2$/MeOH 99:1), compound **5.32** (178 mg, 83%) was obtained.
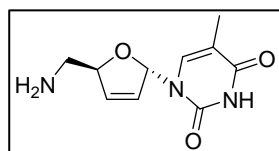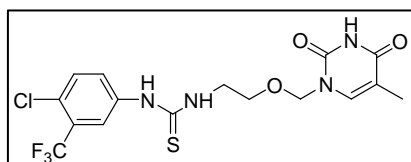

^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.76 (3H, d, *J*= 1.2 Hz, 5-CH$_3$), 3.65 (4H, app s, CH$_2$CH$_2$), 5.09 (2H, s, OCH$_2$N), 7.58 (1H, d, *J*= 1.2 Hz, arom H), 7.62 (1H, d, *J*= 8.7 Hz, arom H), 7.73 (1H, dd, *J*= 2.4 and 8.7 Hz, arom H), 8.12 (2H, app d, *J*= 1.2 Hz, H-6 and N(5')H), 9.94 (1H, br s, N(ar)H), 11.29 (1H, br s, N(3)H).

^{13}C NMR (75 MHz, DMSO-d$_6$): δ 12.52 (5-CH$_3$), 44.15 (*C*H$_2$NH'), 67.19 (*C*H$_2$O), 76.87 (O*C*H$_2$N), 109.92 (C-5), 117.96 (*C*F$_3$), 121.58 (arom C), 125.21 (arom C), 126.73 (arom C), 127.14 (arom C), 132.26 (arom C), 139.86 (C-6), 141.14 (arom C), 151.78 (C-4), 164.90 (C-2), 181.33 (C=S).

Exact mass (ESI-MS) for C$_{16}$H$_{17}$N$_4$O$_3$SF$_3$Cl [M+H]$^+$ found, 437.0659; calcd, 437.0661.

2',3'-Isopropylideneadenosine (5.54)

Adenosine (1.00 g, 3.74 mmol) and *p*-toluenesulphonic acid (710 mg, 3.74 mmol) were suspended in acetone (25 mL) and dimethoxypropane (5 mL). After 10 minutes the reaction mixture became clear and after 1 hour, the mixture was evaporated to dryness, resolved in aqueous NaHCO$_3$ (25 mL) and extracted three times with ethyl acetate (25 mL). The organic solvents were dried over MgSO$_4$, evaporated to dryness and purified by column chromatography (CH$_2$Cl$_2$/MeOH 95:5) to yield 880 mg (77%) **5.54**.
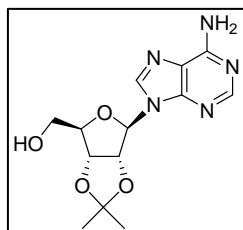
^{1}H NMR (300 MHz, DMSO-d$_6$): δ 1.33 (3H, s, CCH$_3$), 1.55 (3H, s, CCH$_3$), 3.47-3.60 (2H, m, H-5' and H-5"), 4.21 (1H, m, H-4'), 4.96 (1H, dd, *J*= 2.7 and 6.0 Hz, H-3'), 5.23 (1H, d, *J*= 5.4 Hz, 5'-OH), 5.35 (1H, dd, *J*= 3.3 and 6.0 Hz, H-2'), 6.12 (1H, d, *J*= 3.3 Hz, H-1'), 7.34 (2H, s, NH$_2$), 8.15 and 8.34 (2H, 2 x s, H-2 and H-8).

Exact mass (ESI-MS) for C$_{13}$H$_{17}$N$_8$O$_3$ [M+H]$^+$ found, 308.1353; calcd, 308.1359.

5'-Azido-5'-deoxy-2',3'-isopropylideneadenosine (5.55)

Compound **5.54** (880 mg, 2.87 mmol) was dissolved in pyridine (12 mL) and methanesulfonyl chloride was added at 0 °C. After 1 hour the reaction was finished. Aqueous $NaHCO_3$ solution (20 mL) was added and the reaction was extracted three times with ethyl acetate (20 mL). The combined organic layers were dried over $MgSO_4$ and evaporated to dryness. The obtained residue was resolved in DMF (40 mL), NaN_3 (1.527 g, 23.5 mmol) was added, and the reaction was heated to 60 °C overnight. The mixture was evaporated to dryness and purified by column chromatography (CH_2Cl_2/MeOH 98:2), yielding 677 mg (71%) of compound **5.55**.
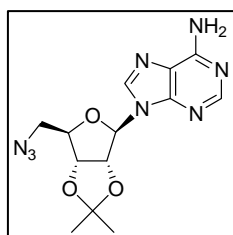
¹H NMR (300 MHz, DMSO-d_6): δ 1.34 (3H, s, CCH_3), 1.54 (3H, s, CCH_3), 3.57-3.72 (2H, m, H-5' and H-5"), 4.24 (1H, m, H-4'), 4.98 (1H, dd, J= 2.7 and 6.0 Hz, H-3'), 5.35 (1H, dd, J= 3.3 and 6.0 Hz, H-2'), 6.11 (1H, d, J= 3.3 Hz, H-1'), 7.38 (2H, s, NH_2), 8.14 and 8.36 (2H, 2 x s, H-2 and H-8).

Exact mass (ESI-MS) for $C_{13}H_{17}N_8O_3$ $[M+H]^+$ found, 333.1419; calcd, 333.1424.

5'-Amino-5'-deoxy-2',3'-isopropylideneadenosine (5.56)

A solution of **5.55** (677 mg, 2.04 mmol) and Pd/C (68 mg) in MeOH (25 mL) was hydrogenated overnight. The reaction mixture was filtered over celite, evaporated to dryness and used in the next reaction without further purification.

N-(5'-Deoxy-2',3'-*O*-isopropylidene-β-D-adenosin-5'-yl)-*N'*-(5'-deoxy-α-D-thymidin-5'-yl)thiourea (5.57)

5'-Amino-2',3'-isopropylidene adenosine (**5.56**) (130 mg, 0.43 mmol), imidazole (6mg, 0.08 mmol) and 1,1-thiocarbonyldiimidazole (83 mg, 0.47 mmol) were dissolved in DMF (7 mL) at 0 °C and after 10 minutes the reaction was allowed to warm up to room temperature. After 2 hours, **5.47** (112 mg, 0.43 mmol) in 2 mL DMF was added. After 3 hours, the reaction was finished and evaporated to dryness. The obtained crude compound was purified by column chromatography (CH_2Cl_2/MeOH 95:5) to obtain 205 mg (82%) **5.57**.
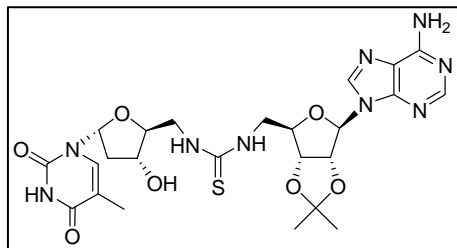
¹H NMR (300 MHz, DMSO-d₆): δ 1.32 (3H, s, CCH₃ (A)), 1.53 (3H, s, CCH₃ (A)), 1.76 (3H, d, J= 0.9 Hz, 5-CH₃ (T)), 1.87-1.94 (2H, ddd, J= 2.6 and 14.2 Hz, H-2' (T)), 2.49-2.59 (1H, m, H-2" (T)), 3.41-3.67 (3H, m, H-5'(T), H-5" (T) and H-5' (A)), 3.78-3.90 (1H, m, H-5" (A)), 4.13-4.34 (3H, m, H-3' (T), H-4' (T) and H-4' (A)), 5.01 (1H, dd, J= 3.4 and 6.2 Hz, H-3' (A)), 5.41 (1H, d, J= 3.4 Hz, 3'-OH (T)), 5.46 (1H, dd, J= 2.8 and 6.3 Hz, H-2' (A)), 6.15 (2H, m, H-1' (T) and H-1'

(A)), 7.34 (2H, s, NH₂), 7.55 and 7.64 (2H, 2 x br s, 2 x N(5')H), 7.73 (1H, d, J= 1.2 Hz, H-6), 8.18 and 8.33 (2H, 2 x s, H-2 and H-8 (A)), 11.24 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{24}H_{32}N_9O_7S$ [M+H]⁺ found, 390.2148; calcd, 390.2145.

N-(5'-deoxy-β-D-adenosin-5'-yl)-N'-(5'-deoxy-α-D-thymidin-5'-yl)thiourea (5.33)

Compound **5.58** (120 mg, 0.20 mmol) was dissolved in 50% trifluoroacetic acid in H₂O (10 mL) and stirred for 2 hours at room temperature. The reaction mixture was evaporated to dryness and purified by column chromatography (CH₂Cl₂/MeOH 93:7) to obtain 98 mg of the pure compound **5.33** (89%).
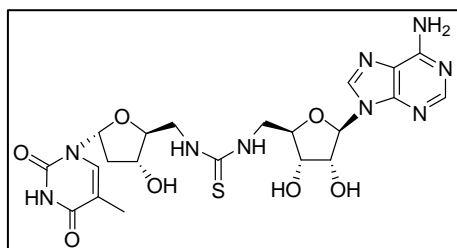
¹H NMR (300 MHz, DMSO-d₆): δ 1.76 (3H, d, J= 1.2 Hz, 5-CH₃ (T)), 1.86-1.94 (2H, ddd, J= 2.6 and 14.6 Hz, H-2' (T)), 2.49-2.59 (1H, m, H-2" (T)), 3.41-3.67 (3H, m, H-5'(T), H-5" (T) and H-5' (A)), 3.82-3.93 (1H, m, H-5" (A)), 4.01-4.29 (4H, m, H-3' (T), H-4' (T), H-4' (A) and H-3' (A)), 4.70 (1H, t, J= 5.4 Hz, H-2' (A)), 5.29 (1H, br s, 3'-OH (A), 5.43 (1H, br s, 3'-OH (T)), 5.49 (1H, br s, 2'-OH (A)), 5.88 (1H, d, J= 6.0 Hz, H-1' (A)), 6.14 (1H, dd, J= 3.0 and 7.5 Hz, H-1' (T)),

7.43 (2H, s, NH₂), 7.59 (1H, br s, N(5')H), 7.74 (2H, app d, J= 1.2 Hz, H-6 and N(5'(A)H)), 8.20 and 8.38 (2H, 2 x s, H-2 and H-8 (A)), 11.25 (1H, br s, N(3)H).

Exact mass (ESI-MS) for $C_{21}H_{28}N_9O_7S$ [M+H]⁺ found, 550.1833; calcd, 550.1832.

Chapter 6

Evaluation on other thymidine kinases

6 EVALUATION OF THYMIDINE ANALOGUES ON OTHER THYMIDINE KINASES

Nucleosides (N) and deoxynucleosides (dN) form an important class of biological building blocks, which are essential for the synthesis of DNA and RNA, and thus for cell survival. Before being built in DNA or RNA, nucleosides require the activation into their triphosphates by different enzymes. The nucleoside kinase (NK) is responsible for the first phosphorylation, followed by the phosphorylation by nucleoside monophosphate kinase (NMPK) to the nucleoside diphosphate (NDP), which is then converted to the nucleoside triphosphate (NTP) by nucleoside diphosphate kinase (NDPK). Similar enzymes exist to convert the corresponding deoxynucleosides into their activated triphosphate forms.

$$N \xrightarrow{\text{NK}} NMP \xrightarrow{\text{NMPK}} NDP \xrightarrow{\text{NDPK}} NTP$$

Despite the structural differences found in bacterial and human TKs and TMPKs, [174,175,176,177] all these enzymes catalyze very similar biochemical processes. Both kinases possess a binding site for dT or dTMP embedded in the interior of the enzyme, an adenosine binding site located on the surface of the enzymes, and a bridging phosphate binding site which accommodates either three (TK) or four (TMPK) phosphate moieties. Interestingly, herpes viruses use TKs which have both TK- and TMPK activity.[178,179]

In this thesis a wide variety of thymidine derivatives was synthesized and evaluated for their activity against the mycobacterial thymidine monophosphate kinase. Due to the high similarity between TKs and TMPKs, those compounds might also interfere with thymidine kinases. After a description of the role, appearance and characteristics of the different deoxynucleoside kinases, the activities of selected analogues against different kinases will be reported.

6.1 DEOXYNUCLEOSIDE KINASES

In mammalian cells four different deoxynucleoside kinases can be found: thymidine kinase 1 (TK-1), thymidine kinase 2 (TK-2), deoxycytidine kinase (dCK) and deoxyguanosine kinase (dGK).

As explained before, the main role of deoxynucleoside kinases is the activation of deoxynucleosides to their monophosphates, which is essential for their conversion to DNA-building blocks. Deoxynucleoside kinases play also a fundamental role in the activation of

nucleoside drugs. The anticancer agents Gemcitabine, Fludarabine, Clofarabine and Cladribine require activation by dCK before being biologically active. The same activation is required for Zalcitabine (ddC), Lamivudine (3TC), didanosine (ddI), abacavir (ABC) and Emtricitabine ((-)-FTC), used for the treatment of HIV-infection. Similarly, TK-1 is crucial for the activation of drugs used in HIV-infected patients, like Zidovudine (AZT) and Stavudine (d4T).[180]

After further conversion to their triphosphates, these agents serve as alternative substrates in the reverse transcriptase reaction, where upon their incorporation they act obligatory as chain terminators.

In contrast to these nucleoside reverse transcriptase inhibitors, nucleotide reverse transcriptase inhibitors such as adefovir (PMEA) and tenofovir (PMPA) already possess a phosphate mimicking phosphonate group. Consequently, only two phosphorylation steps are required to convert these antivirals to the active metabolites.

6.1.1 Thymidine kinase 1 and 2

Among the different mammalian deoxynucleoside kinases, two enzymes phosphorylate thymidine, TK-1 and TK-2. The main differences between the two kinases are demonstrated by their amino acid sequences, their substrate specificities, their location and their levels of expression during the different cell cycle phases. [181, 182]

Table 6.1 The most important features of TK-1 and TK-2		
	TK-1	**TK-2**
Amino acid sequence	Unique in its group	Belongs to a larger group of nucleoside kinases: similar to dCK, dGK, Dm-dNK[a] (40% analogy), HSV-1 TK
Substrate specificity:		
- Natural nucleosides	Thymidine, 2'-deoxyuridine	Thymidine, 2'-deoxyuridine, 2'-deoxycytidine
- Antiviral nucleoside analogues	- AZT: excellent substrate	AZT: poor substrate, but in non-dividing cells TK-2 phosphorylation becomes significant.
	- Ara-T: weak substrate	- Ara-T: good substrate
Location	Cytosol	Mitochondria
Levels of expression	Highest expression in S phase cells; very low activity in resting cells	Constitutively expressed, physiologically active in non-proliferating and resting cells.

[a] Dm-dNK: multisubstrate deoxynucleoside kinase of the fruitfly Drosophilia melanogaster

So far, no crystallographic determination of TK-2 has been reported. The strong similarity with Dm-dNK (40% similarity in amino acid sequence and no long amino acid sequence insertions or deletions) allowed the construction of three-dimensional models of TK-2, using the Dm-dNK structure as a template.[183,184]

6.1.2 The role of thymidine kinase 2

6.1.2.1 TK-2 deficiency

Throughout the whole cell cycle, mitochondrial DNA replication takes place, thus constantly requiring the production of dNTPs for mtDNA synthesis. Being active in non-proliferating tissues, TK-2 provides the nucleotides for mitochondrial DNA (mtDNA) synthesis. TK-2 deficiency would therefore lead to mitochondrial disorders, designated as mtDNA depletion syndromes, mostly affecting skeletal muscles.[185] Some of these disorders might be associated to mutations in mitochondrial dNKs.[186]

6.1.2.2 Mitochondrial toxicity

Next to the mitochondrial disorders associated to TK-2 deficiency, severe mitochondrial toxicity is also associated to long-term treatment with antiviral nucleoside analogues such as AZT.[187,188] Although the mechanism by which these nucleoside analogues exert their mitochondrial toxicity is not fully understood,[187] it has been suggested that after phosphorylation of the nucleoside analogues by TK-2, their triphosphates might accumulate in the mitochondria. Although AZT is not an excellent substrate for TK-2, in non-replicating cells AZT phosphorylation by TK-2 is significant. The accumulation of AZT-TP is suggested to affect DNA-polymerase-γ, resulting in mtDNA depletion.[187,188]

In order to prevent this TK-2 related mitochondrial toxicity, TK-2 inhibitors might be useful to be given as a co-treatment with nucleoside analogues. Additionally, the identification of TK-2 inhibitors can be a valuable tool to answer the many open questions regarding the real contribution of TK-2 in the maintenance and homeostasis of mitochondrial dNTP pools and in the phosphorylation of certain antiviral drugs which is suggested to cause their mitochondrial toxicity.

6.1.3 Known thymidine kinase 2-inhibitors

6.1.3.1 Ara-, ribo- and 2'-deoxyribo- nucleoside analogues

In the past, several TK-2 inhibitors have been identified. An interesting finding concerns the inhibitory activity of different ribofuranosyl nucleosides (e.g. 5-(*E*)-(2-bromovinyl)uridine **6.1**; K_i = 10.4 µM), while their 2'-deoxy analogues are phosphorylated by the enzyme.[189] TK-2 can also be inhibited by certain nucleosides, including 3'-*O*-alkyl thymidine analogues and 3'-hexanoylamino-3'-deoxythymidine **6.2**, a potent TK-2 inhibitor (K_i-value = 0.15 µM).[190] While Ara-T behaves as a good substrate for TK-2, the introduction of long chain acyl substituents at the 2'-OH affords TK-2 inhibitors (e.g. **6.3**: K_i = 2.3 µM). As these 2'-*O*-acyl derivatives are unstable in cell culture and readily converted to the parent nucleoside,[191] they can not be used as tools to study TK-2 in intact cells.
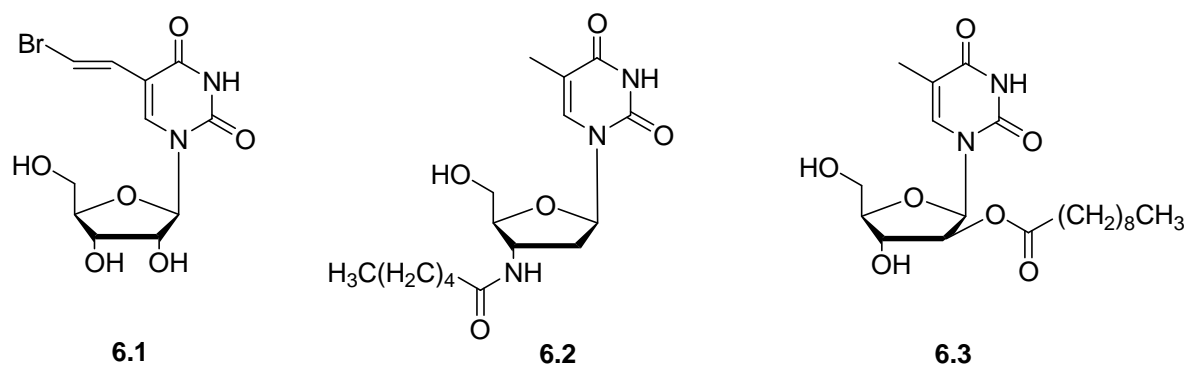
Figure 6.1 Ara-, ribo- and 2'-deoxyribo- nucleoside analogues as TK-2 inhibitors.

6.1.3.2 Acyclic nucleoside analogues

The identification of 5'-*O*-trityl-thymidine as a moderately active inhibitor of TK-2 (IC_{50} = 33 µM), encouraged Pérez-Pérez et al. to replace the sugar moiety by acyclic spacers, as the main role of the sugar moiety was suggested to be the correct positioning of the trityl substituent and the thymine base.[192] This led to from the acyclic 1-[(*Z*)-4-(triphenylmethoxy)-2-butenyl]thymine **6.4** (IC_{50} = 1.5 µM), which, after further optimisation efforts, resulted in **6.5** with a IC_{50}-value as low as 0.4 µM.[193]
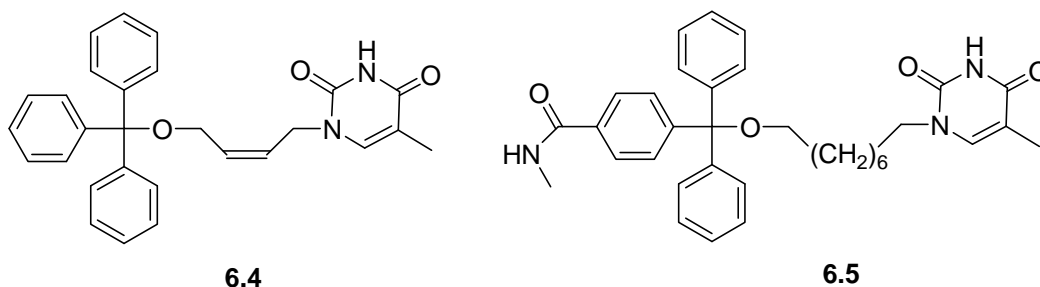
Figure 6.2 Acyclic nucleoside analogues as TK-2 inhibitors.

Enzyme kinetic analysis showed that these acyclic nucleoside derivatives are competing with the thymidine substrate for TK-2 binding, but behave as uncompetitive inhibitors with respect to ATP, the co-substrate of the TK-2 catalyzed reaction. ATP-binding to the enzyme results in a conformational change. Only after this conformational change, the inhibitors show their affinity for TK-2.

Consequently, it can be concluded that the acyclic nucleoside derivatives bind to the thymidine-binding site of the enzyme, but only after ATP was bound to the enzyme.[192]

6.2 ACTIVITIES OF SYNTHESIZED THYMIDINE DERIVATIVES

Selected 2',3'-bicyclic thymidine derivatives, 3'-branched β-thymidine analogues and 5'-thiourea substituted α-thymidine derivatives were tested for their inhibitory activities against TK-1 and TK-2, the thymidine kinases of *Herpes simplex* virus and *Varicella zoster* virus, and the multifunctional 2'-deoxynucleoside kinase from *Drosophilia melanogaster*. The activities are shown in Table 6.2.

Table 6.2 Kinase activity of synthesized thymidine derivatives (IC$_{50}$ in µM)

	TK-1	TK-2	HSV-1 TK	VZV TK	Dm dNK
4.3	> 500	408 ± 20	> 500	≥500	> 500
4.6	> 500	**8.2 ± 2.8**	208 ± 31	37 ± 1	78 ± 8
4.8	> 500	146 ± 66	> 500	237 ± 72	> 500
4.9	> 500	> 500	≥500	268 ± 19	> 500
5.3	> 500	**1.2 ± 1.1**	400 ± 54	31 ± 5	84 ± 51
5.4	> 500	**3.1 ± 0.3**	≥500	77 ± 51	72 ± 33
5.5	> 500	**1.2 ± 0.1**	444 ± 18	37 ± 2	38 ± 4
5.10	> 500	405 ± 78	> 500	40 ± 3	> 500
5.14	> 500	57 ± 20	422 ± 6	30 ± 7	258 ± 13
5.17	> 500	31 ± 8	392 ± 8	45 ± 1	249 ± 24
5.21	> 500	110 ± 13	> 500	95 ± 66	38 ± 5
5.23	> 500	38 ± 1	472 ± 40	40 ± 4	266 ± 49
5.24	> 500	99 ± 36	> 500	150 ± 38	474 ± 17
5.25	> 500	41 ± 3	365 ± 22	45 ± 6	253 ± 5
5.31	> 500	≥500	≥500	> 500	> 500
5.32	> 500	176 ± 6	19 ± 0	282 ± 5	242 ± 15
5.33	> 500	> 500	> 500	330 ± 2	> 500
5.45	99 ± 51	46 ± 2	34 ± 3	47 ± 3	232 ± 4
5.46	> 500	466 ± 47	> 500	284 ± 6	> 500
5.49	> 500	490 ± 14	> 500	304 ± 0	≥500
5.53	> 500	> 500	> 500	> 500	> 500
5.54	> 500	> 500	> 500	> 500	> 500

While, with exception of the α-D-thymidine (**5.45**), the tested compounds were virtually not inhibitory active against TK-1, several analogues showed good inhibitory activity against TK-2, resulting in selective TK-2 inhibition.

Especially in the series of the 3'-*C*-branched thiourea-substituted β-thymidine derivatives, several analogues (**5.3**, **5.4** and **5.5**) displayed potent and highly selective inhibitory activity

for TK-2 versus TK-1. With IC$_{50}$-values around 1 μM, these compounds approach the high activities found for some acyclic nucleoside analogues.

Initial kinetic studies indicated that these compounds act as competitive inhibitors of the phosphorylation of thymidine to TMP, catalyzed by TK-2.

In contrast to our observations on TMPKmt, the thiourea-substituted α-D-thymidine derivatives showed weaker activities compared to their analogous 3'-branched β-thymidine derivatives.

In the series of 2',3'-bicyclic derivatives, the five-membered thiocarbamate derivative **4.6** with an IC$_{50}$-value of 8.2 μM, appeared to be the most active TK-2 inhibitor.

In order to determine the way of binding of those compounds to the enzyme, mor detailed enzyme kinetics will be performed. These studies, combined with molecular modelling studies, should reveal if the identified hits bind the TK-2 enzyme in a similar way as TMPKmt.

6.3 CONCLUSION

With IC$_{50}$-values in the low micromolar range, 3'-*C*-branched thiourea-substituted β-D-thymidine derivatives (e.g. **5.3**, **5.4** and **5.5)** represent a new class of potent and selective TK-2 inhibitors. This finding opens the way for further structure-based optimisation. This approach may ultimately lead to the discovery of agents that are useful for studying the physiological role of the mitochondrial TK-2 enzyme. Furthermore, new leads may be discovered that are capable to suppress the mitochondrial toxicity observed after prolonged treatment with nucleoside reverse transcriptase inhibitors, whilst maintaining cytosolic production of the active triphosphate metabolites.

Part II: Extension of the applicability of a 3'-*C*-branched thymidine analogue

Chapter 7

Fluorescently labelled oligonucleotides

7 A FLUORESCENT 3'-*C*-BRANCHED THYMIDINE DERIVED COMPOUND AS BUILDING BLOCK IN OLIGONUCLEOTIDE SYNTHESIS

7.1 INTRODUCTION

3'-*C*-branched thymidine derivatives play a major role throughout this thesis. A project in collaboration with the University of Southern Denmark (SDU, Odense, DK), raised the opportunity to use such a compound as modified building block for oligonucleotide synthesis, in this way illustrating the broad applicability of modified nucleosides in medicinal chemistry. This chapter describes the synthesis of a 3'-*C*-branched pyrene substituted thymidine derivative, its incorporation into different oligonucleotides and a study of the most important properties of the obtained oligonucleotides.

7.1.1 Secondary and tertiary structure of DNA and RNA

Genetic information in cells is carried by 2'-deoxyribonucleic acids (DNA) and ribonucleic acids (RNA).[194,195]

Two complementary strands of DNA/RNA may self-assemble and coil around a common axis, forming a right-handed double stranded helix. The nucleobases pair with each other following the Watson-Crick base pairing rules. This pairing is determined by the formation of hydrogen bonds, i.e. two between adenosine (A) and thymidine or uridine (T/U), and three between guanosine (G) and cytidine (C) and is further stabilized by π-π stacking between the aromatic nucleobases. The two polymers in the duplex are hybridized with each other in an antiparallel orientation of the polymers, bringing the 5'-end of one polymer opposite to the 3'-end of the other polymer.[194,195]

AT-pair

GC-pair
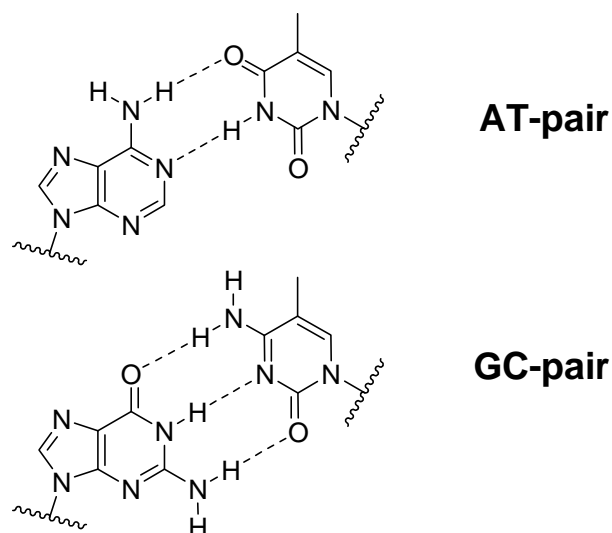
Figure 7.1 Watson-Crick baseparing between adenine and thymine and between guanine and cytosine

The nucleobase moieties are buried in the hydrophobic core of the duplex while the sugar moieties and the phosphate backbone are on the exterior of the duplex and therefore more exposed to the solvent. Depending on the duplex type and environmental conditions, different helix geometries are observed, i.e. A-, B- or Z-type helices. Their main differences are the amount of base-pairs per helical turn, and the wideness of their major and minor groove (see Figure 7.2).

Minor groove

Major groove
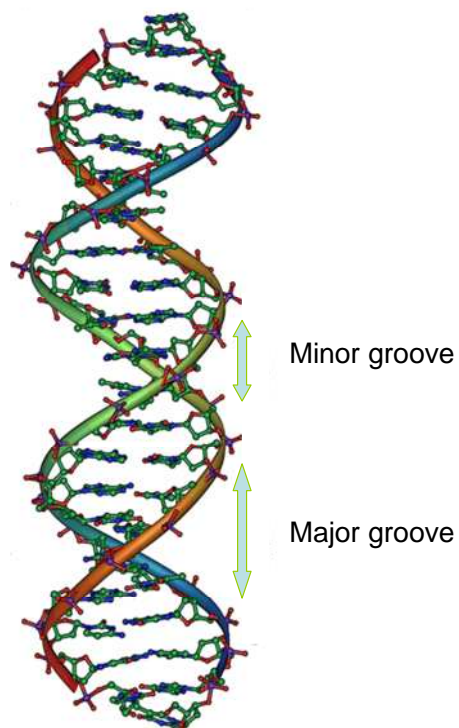
Figure 7.2 Right-handed double stranded DNA helix[196]

Other structures than duplexes may be formed by nucleic acids, such as DNA triplexes.[197] In the formation of triplexes, two different motifs are known, depending on the base composition and the orientation of the third strand, i.e. the triplex forming oligonucleotide (TFO).

A TFO consisting of only pyrimidines may bind to dsDNA (having a polypurine tract) via base triplets where a thymine (T) or protonated cytidine (C$^+$) form weak hydrogen bonds with an AT- or GC-base-pair, respectively, called Hoogsteen bonds. In this triplex, the TFO and the duplex purine strand are aligned in a parallel orientation. As the cytidine of the TFO needs to be protonated for a strong interaction, these Hoogsteen bonds are clearly pH-dependent.

In the antiparallel motif, a TFO consisting of only purines binds a dsDNA via reverse Hoogsteen bonding (G-GC and A-AT). In both cases, the TFO is positioned in the major groove of the dsDNA.
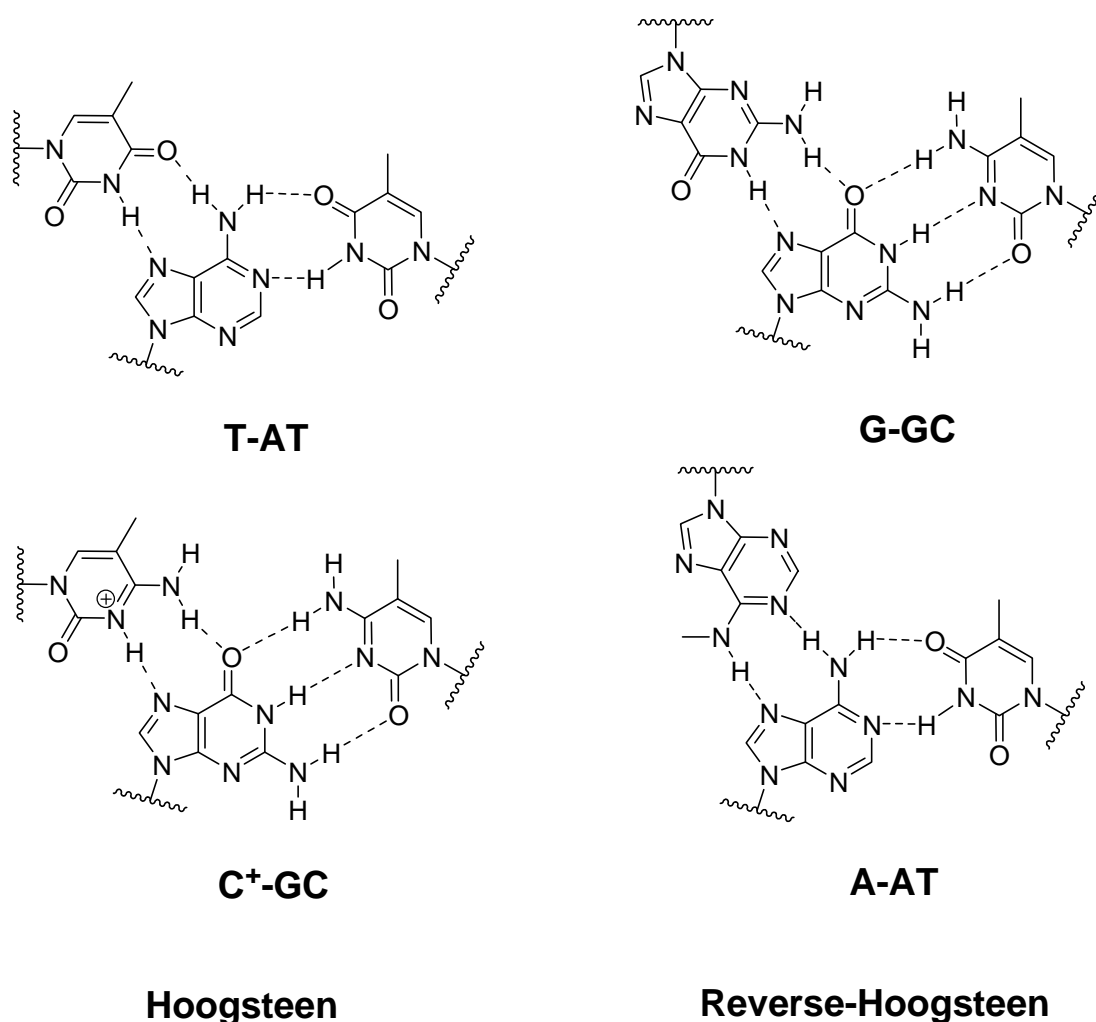
Figure 7.3 Illustration of Hoogsteen and reverse Hoogsteen base triplets.

7.1.2 <u>Nucleic acid diagnostics</u>

During the last decade, the sequencing of the human genome has led to a high need for the development of probes, able to detect specific sequences of nucleic acids in cytogenetic research and clinical diagnosis.[198,199] The demand for such assays stems from a need of clinics to reduce time in the detection of diseases. Two principle approaches to achieve this, consist of heterogenous or homogenous assays. In heterogenous assays, a target is immobilized on a surface (e.g. gel or membrane) and a solution of various probes is poured over the surface. Only complementary probes bind to the target, while non-complementary probes are removed by washing. In order to detect hybridization, a reporter group (e.g. fluorescent or radioactive molecule, termed label) is attached to the probe, causing a signal upon hybridization. The main disadvantages of these heterogenous assays are the very time-consuming and labor-intensive work and above this, the inability to investigate living organisms.

In homogenous assays however, the excess of non-binding probes cannot be washed out, thus requiring a strategy in which a change in a spectral property (e.g. intensity, emission wavelength) is observed upon hybridization to the nucleic acid target, relative to other non-binding probes. The most popular homogeneous assays rely on fluorescence detection.

7.1.2.1 *Binding of small molecules to dsDNA*

The simplest design of homogenous fluorescence assays involves the binding of small molecules (e.g. ethidium bromide or cyanine dyes) to dsDNA in the minor groove or by intercalation. Upon binding to dsDNA, these dyes become fluorescent.[200] A main disadvantage of these molecules is the non-specific binding to dsDNA and moreover, they cannot detect single stranded probes.
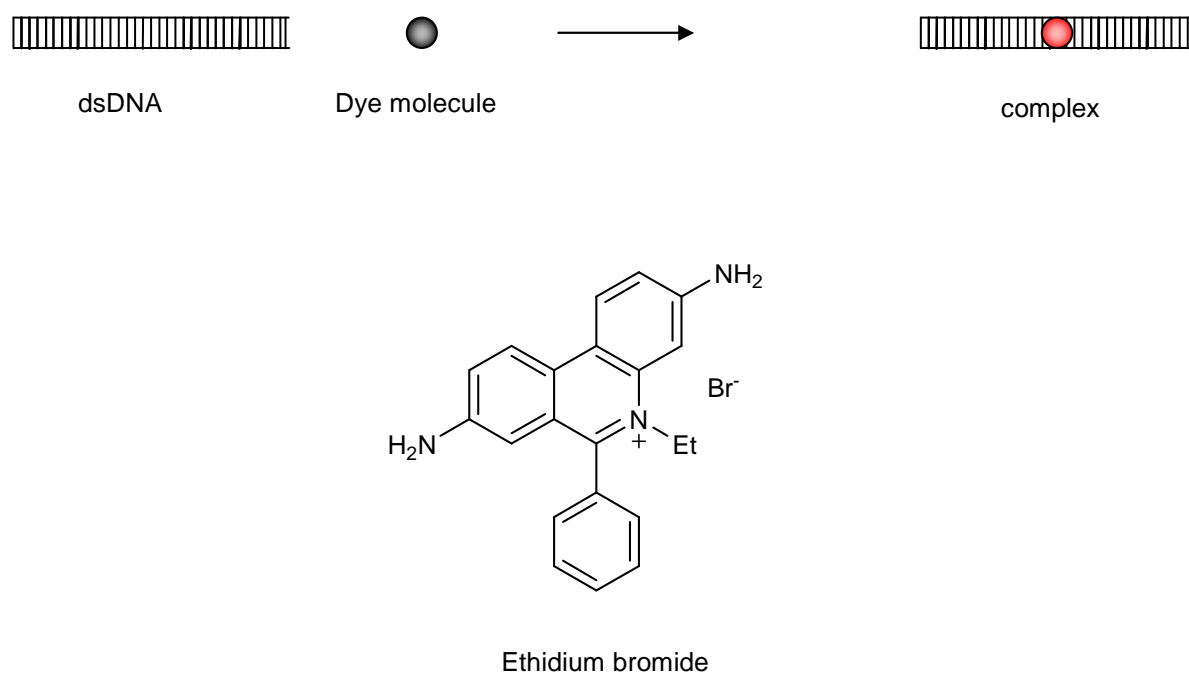
Ethidium bromide

Figure 7.4 The simplest homogenous fluorescence assay based on ethidium bromide

7.1.2.2   Intercalating agents

In order to obtain specific detection of single stranded nucleic acids, reporter groups have been connected to oligonucleotides,[201,202] using the Watson-Crick mediated hybridisation. Intercalating agents have been attached to ON's, which might lead to increases in fluorescence intensity upon hybridization to DNA/RNA complements. However, very often these single stranded probes already show a high level of fluorescence, especially if the intercalating dye is attached to the ON via a long linker,[203] in this way causing only a weak hybridization-induced fluorescence increase.

A similar approach to facilitate detection of nucleic acids in solution is the attachment of a fluorophore-group giving a strong fluorescence signal when it is projected away from the nucleobase stacking in a nucleic acid duplex. Partial success was achieved with ONs labelled with fluoresceins,[204] dansyl[205] or pyrenes.[206,207] However, instead of hybridisation of the probe to its target, other substances might cause quenching of fluorescence. Moreover, the highest decrease that can be obtained is maximum 100%, indicating that hybridization-induced increase in fluorescence intensity is preferred.

The change in the fluorescence intensity observed is mainly depending on the chemical nature of the fluorophore and on the position of the fluorescent label relative to the nucleobases before and after hybridization. Nucleobase moieties in the close proximity of

the fluorophore can quench the fluorescence emitted by the fluorescent label.[208]   Both the increase and decrease of fluorescence intensity upon hybridization are based on this feature of fluorophores.

7.2 TRIPLEX FORMATION OF PYRENE LABELLED PROBES FOR NUCLEIC ACID DETECTION IN FLUORESCENCE ASSAYS

7.2.1 General requirements for nucleic acid detection

A technique that requires the earlier described typical features is the combinatorial oligonucleotide fluorescence *in situ* hybridisation (COMBO-FISH) introduced by Hausmann et al. in 2003.[209]   COMBO-FISH uses oligonucleotides (ONs) that form triple helices with intact duplexes to label chromosomes in a cell nucleus under non-denaturing conditions. The triplex formation, upon binding of the triplex forming oligonucleotides (TFOs) to a homopurine/homopyrimidine DNA duplex, is highly specific.   It can be used for sequence-specific recognition of dsDNA, without prior denaturation of it.[210]   As described above, one of the disadvantages is the use of nucleic acid probes possessing a fluorophore, which is not sensitive enough to changes in the microenvironment, especially after hybridization of the probe to the complement.   This leads to a high background signal and the requirement of washing out excess fluorescent probes which makes labelling very difficult.   To overcome these problems, novel fluorescent modified nucleic acids are needed.   Among the available fluorophores, pyrene and its derivatives are attractive fluorescent dyes due to their inherent chemical and photochemical characteristics, rather long fluorescence lifetimes and sensitivity to the microenvironment.[211]   The ideal ON probe should have the fluorophore incorporated in the middle of the sequence and display a very low fluorescence signal as single stranded ON.   Upon hybridisation a considerable increase of the fluorescence signal (positive signal) must indicate binding to the target.

7.2.2 Pyrene in modified nucleoside building blocks

In previous research, pyrene derivatives have been incorporated at different positions of the modified nucleoside building blocks.   The 2'-N-position of 2'-amino-DNA[212] and –LNA[213] or 2'-O-position of RNA[214] were substituted with pyrene via a short tether.   In both cases, increase of fluorescence upon hybridisation to RNA is strong and is much weaker in the case of binding to DNA.
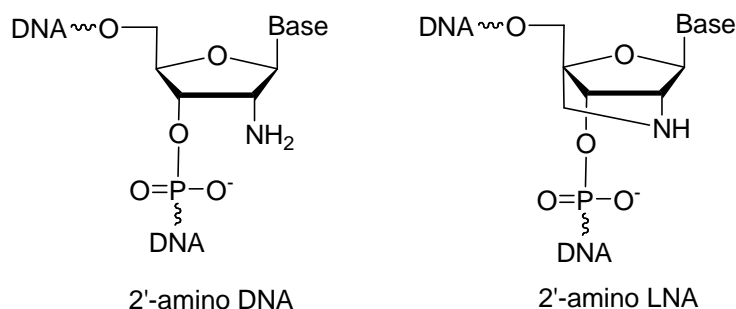
Figure 7.5 Structures of 2'-amino DNA and LNA

Incorporation of pyrene at an internal or 5'-terminal phosphorus via a relatively long linker ($\geq$ 4 carbons) resulted in considerable quenching of fluorescence upon hybridisation due to intercalation of pyrene,[215] while incorporation at the 3'- or 5'-end via a relatively short linker (1 or 2 atoms) gave a slight increase of fluorescence.[216] Bryld et al.[217] describe the synthesis and evaluation of 4'-pyrene-substituted DNA via a piperazinomethyl linker. They could not observe any difference in fluorescence upon hybridisation. Finally the pyrene unit has been coupled to sugar moieties to intercalate as bulge in the middle of the TFO.[218, 219]

None of the described probes are known to give increased fluorescence intensities upon triplex formation in which the single-stranded triplex-forming oligonucleotide (TFO) binds to dsDNA through specific major groove interactions.

Since the 3'-position is needed as a part of the phosphate-backbone, this position has not been substituted with pyrene moieties, except at a 3'-end monomer.

7.2.3 <u>Objectives</u>

Obika et al.[220] demonstrated that several 2'-5'-linkages in a 3'-5' TFO strand led to stabilisation of triplexes. As this change in the phosphate-backbone opens possibilities for the introduction of a pyrene unit at the 3'-position, we decided to synthesize a modified nucleoside with the fluorophore attached to the 3'-position via a short linker.

A novel fluorescent pyrene-containing nucleoside was prepared by click chemistry between 1-ethynylpyrene and 3'-azidomethyl-3'-deoxyribothymidine and incorporated into several ONs.   Thermal stabilities and fluorescence spectra of the different probes and their corresponding duplexes and triplexes with complementary ssDNA/RNA and dsDNA's were examined.   Modelling was used as a tool to explain the results of thermal stability studies and the fluorescence properties observed.

7.2.4 Chemistry

The synthesis of phosphoramidite **7.3** (Scheme 7.1) started from compound **4.26**, which was obtained in 11 steps from 1,2-*O*-isopropylidine-α-D-xylofuranose with an overall yield of 17%.[221] The formation of the 1,4-triazol was performed in the microwave cavity at 125 °C during 15 minutes with Cu (I) as a catalyst in 79% yield. 5'-*O*-Dimethoxytrityl-protection followed by phosphytilation at the 2'-position were performed under standard conditions in 59% yield over two steps.

DNA-synthesis of ONs containing compound **7.3** was performed on a 0.2 µmol scale under standard conditions except for an increased coupling time (10 min) and an extended deprotection step (100 sec), using 4,5-dicyanoimidazole as an activator. This resulted in a coupling efficiency of 98%. The obtained ONs were purified by reverse-phase HPLC, their composition was verified by MALDI-TOF (Table S1) and the purity was found to be over 82% by ion-exchange HPLC.
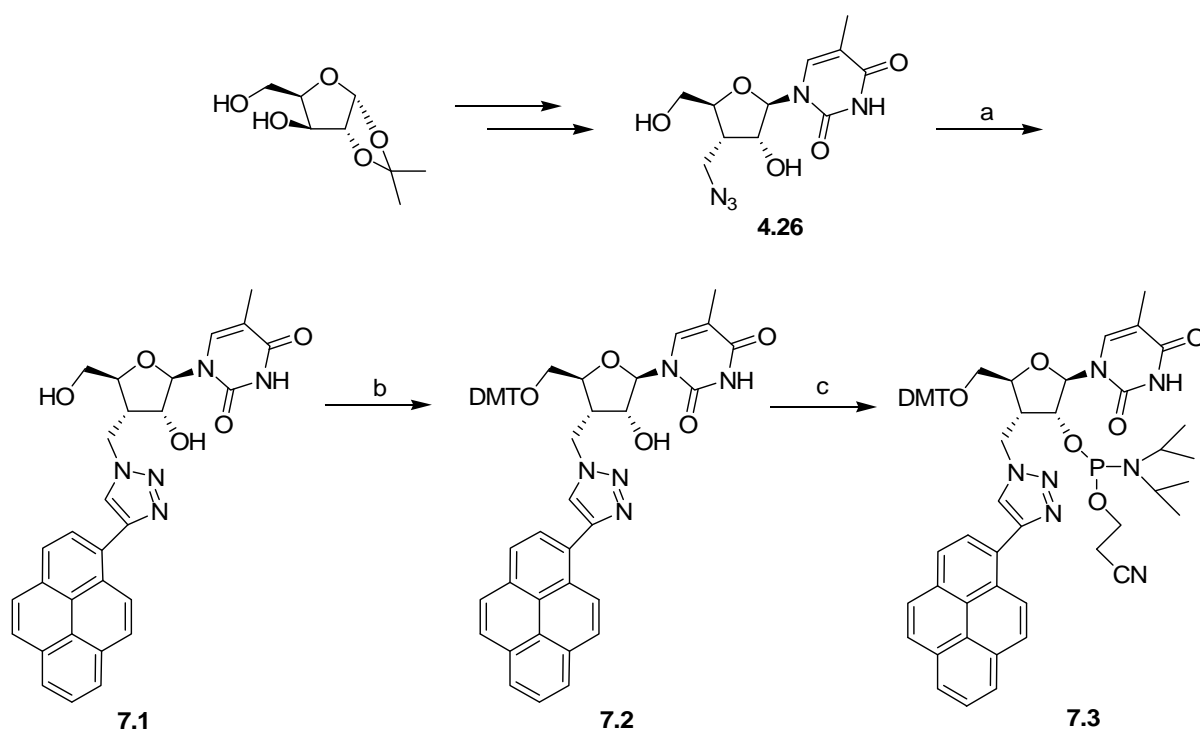
Scheme 7.1, **Reagents and conditions:** a) 1-ethynylpyrene, CuSO$_4$, ascorbic acid, DMF/H$_2$O 19:1, 125 °C, 15 min, microwave, 79%; b) DMT-Cl, pyridine, rt, 16h, 91%; c) NC(CH$_2$)$_2$OP(NPri$_2$)$_2$, diisopropylammonium tetrazolide, CH$_2$Cl$_2$, rt, 16h, 65%.

7.2.5 Homopyrimidine oligonucleotides

7.2.5.1 Thermal stability

The thermal stability of triplexes and duplexes (DNA/DNA and DNA/RNA) using synthesized oligonucleotides was determined by thermal denaturation studies. The melting temperatures (T_m, °C) determined as the first derivatives of the melting curves at 260 nm are listed in Table 7.1, 7.3 and 7.4.

For homopyrimidine sequences, pH-dependent Hoogsteen-type base-pairing was studied, by determining the thermal stability of parallel triplexes toward duplex **D1** and of parallel duplexes toward **ON15** (Table 7.1). T_m values of Watson-Crick type antiparallel duplexes formed by **ON1**-**10** and **ON16** are also shown. Since the latter type of duplexes can also be studied by mixed pyrimidine/purine sequences, mixed 9-mer sequences were used for hybridization with ssDNA and ssRNA (Table 7.4).

Table 7.1 T_m [°C] Data for Duplex Melting of **ON1-10** and **D1**, **ON15** and **ON16**, Taken from UV Melting Curves ($\lambda = 260$ nm)[a]

		Triplex[a] 5' GACGGGGAAAGAAAAAA 3' CTGCCCCTTTCTTTTTT (**D1**)			Parallel duplex[b] 5' GACGGGGAAAGAAAAAA (**ON15**)			Antiparallel duplex[b] 3' GGGGAAAGAAAAAAA (**ON16**)		
pH		5.0	6.0	7.2	5.0	6.0	7.2	5.0	6.0	7.2
ON1	3' TTT TTT CTT TCC CC	55.0[d]	28.0	<5.0	29.5	19.5	*n.d.*	47.5	48.5	47.0
ON2	3' TTT TTT CT**X** TCC CC	55.0[c]	25.0	<5.0	20.0	<5.0	*n.d.*	45.0	47.5	45.5
ON3	3' TTT TTT C**XX** TCC CC	33.0	19.5	*n.d.*[e]	32.0	21.5	*n.d.*	36.0	38.0	*n.d.*
ON4	3' TTT TTT C**XX** **X**CC CC	32.0	< 5.0	*n.d.*	31.5	24.0	*n.d.*	36.5	38.5	*n.d.*
ON5	3' TTT T**X**T CTT TCC CC	55.5[c]	16.0	*n.d.*	23.5	<5.0	*n.d.*	39.5	43.0	*n.d.*
ON6	3' TTT TTT C**X**T **X**CC CC	25.5	<5.0	*n.d.*	21.5	<5.0	*n.d.*	31.5	37.5	*n.d.*
ON7	3' TTT T**X**T CT**X** TCC CC	27.0	<5.0	*n.d.*	22.5	<5.0	*n.d.*	34.5	39.5	*n.d.*
ON8	3' TTT TTT CT**pX** TCC CC	54.0[c]	20.5	*n.d.*	27.5	16.5	*n.d.*	43.0	47.5	*n.d.*
ON9	3' TTT TTT CT**Xp** TCC CC	53.0[c]	22.5	*n.d.*	29.0	20.0	*n.d.*	44.0	46.5	*n.d.*
ON10	3' TT**T**^{L}T TT**T**^{L}T **C**^{L}T**X** TC**C**^{L}C **C**^{L}C	65.5[d]	49.5[c]	35.5[c]	53.0	43.5	38.5	58.5	59.0	58.5

[a] *C*= 1.5 µM of **ON1-10** and 1.0 µM of each strand of dsDNA (**D1**) in 20 mM sodium cacodylate, 100 mM NaCl, 10 mM MgCl$_2$, pH 5.0, 6.0 and 7.2; duplex T_m = 56.5 °C (pH 5.0), 58.5 (pH 6.0) and 57.0 (pH 7.2); [b] *C*= 1.0 µM of each strand in 20 mM sodium cacodylate, 100 mM NaCl, 10 mM MgCl$_2$, pH 5.0, 6.0 and 7.2; [c] Third strand and duplex melting overlaid. T_m was determined at 350 nm; [d] Third strand and duplex melting overlaid; [e] *n.d.* not determined; **p** = TINA, **T**^{L} = thymin-1-yl α-L-LNA monomer, **C**^{L} = 5-methylcytosin-1-yl α-L-LNA monomer. Structure of monomer **X** is shown in Scheme 1 as the 5'-DMT-2'-phosphoramidite-derivative **4**, structures of **p** and α-L-LNA are shown in Figure 7.6.
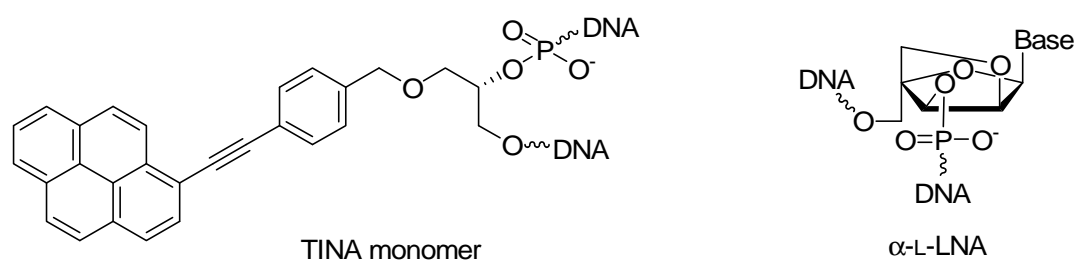
Figure 7.6 Structures of the incorporated triplex-stabilisators

Obika et al.[220] described the stabilising effect of the discontinuous replacement of 3',5'-phosphodiester linkages in TFOs by 2',5'-linkages. As demonstrated by the T_m values for the **ON2/D1** and **ON5/D1** triplexes at pH 6.0, insertion of the fluorophore positioned at 3'-carbon via the triazole linker tends to destabilize parallel triplexes. Introduction of a second 2'-5' thymidine with a pyrene residue led to further destabilization of the parallel triplexes (**ON3**, **ON6**, **ON7** toward **D1**). Destabilisation was found to be higher for triplexes formed with **ON6** and **ON7** having one or three nucleosides between two modified thymidine building blocks than observed for **ON3**, with two adjacent pyrene-containing thymidines. Moreover, it is interesting to mention that at pH 5.0 the T_m-value for the triplex with **ON4**, possessing three modified building blocks in the row, was higher than those for **ON6** and **ON7**. However, all these values were lower than the T_m-value for the unmodified triplex. Such behaviour can be explained by non-interrupted neighbouring 2'-5'-linkages in **ON4** compared to the separate 2'-5'-linkages in **ON6** and **ON7**. This can also be a reason for the surprising stabilization of parallel duplexes observed for **ON3** and **ON4** in comparison with the wild-type triplex **ON1**/**ON15** at pH 5.0 and 6.0 (Table 7.1). This is an interesting observation because both parallel triplexes and parallel duplexes are formed according to Hoogsteen base-pairing. However, a duplex can easier accommodate repetitive insertions of 2'-5' thymidine linkages with pyrene in the middle of one of the strands as it is a more flexible structure than a triplex. Unfortunately, this was not the case with Watson-Crick type duplexes, where a drop in thermal stability was detected for **ON2**-**ON7** toward **ON16** as compared to **ON1/ON16** at pH 5.0 and 6.0.

7.2.5.2 Fluorescence

A. Results

The presence of the pyrene residue in ONs resulted in the formation of an additional band in the UV-spectra with λ_{max} of 350 nm. The single stranded probes containing a single incorporation of the monomer **X** exhibited a fluorescence emission with $\lambda_{max} \approx 386$ (band I) and 402 nm (band III) upon excitation at 350 nm. This emission appears at a higher wavelength than the typical pyrene monomer emission with $\lambda_{max} \approx 378$ (band I) and 391 nm (band III), which can be explained by the conjugated triazole ring in the monomer **X**. To our surprise a very low intensity of fluorescence was observed for single-stranded ONs. Strikingly, the formation of parallel triplexes caused a huge increase of intensity. To study the fluorescence properties in more detail, we decided to determine fluorescence quantum yield (Φ_F) for ssONs and corresponding complexes with complementary strands. Φ_F is defined as the ratio of the number of photons emitted by fluorescence at a certain wavelength to the number of photons absorbed at this wavelength. In our study Φ_F was

determined in the thermal denaturation buffer at 10 °C using equimolar quantities of each strand (1.0 µM) upon excitation at 350 nm relative to anthracene (Φ_F = 0.36) and 9,10-diphenylanthracene (Φ_F =0.95) in cyclohexane[222] following standard procedures (Table 7.2).[223,224]

Table 7.2 Quantum Yields at λ = 350 nm for Single-Stranded Probes **ON2**, **ON5**-**ON7** and **ON10**, and the Corresponding Parallel Triplexes, Parallel and Antiparallel Duplexes in Aerated Thermal Denaturation Buffer at 10 °C.

		SS strand			Triplex[a] 5' GAC GGG GAA AGA AAA AA 3' CTG CCC CTT TCT TTT TT (**D1**)			Parallel duplex[a] 5' GAC GGG GAA AGA AAA AA (**ON 15**)			Antiparallel duplex[a] 3' GGG GAA AGA AAA AA (**ON 16**)		
pH		5.0	6.0	7.2	5.0	6.0	7.2	5.0	6.0	7.2	5.0	6.0	7.2
ON 2	3' TTT TTT CT**X** TCC CC	0.001	0.004	*n.d.*[b]	0.179	0.061	*n.d.*	0.001	*n.d.*	*n.d.*	0.032	0.036	*n.d.*
ON 5	3' TTT T**X**T CTT TCC CC	0.001	0.003	*n.d.*	0.138	0.023	*n.d.*	0.005	*n.d.*	*n.d.*	0.035	0.039	*n.d.*
ON 6	3' TTT TTT C**X**T **X**CC CC	0.001	0.003	*n.d.*	0.036	*n.d.*	*n.d.*	0.009	*n.d.*	*n.d.*	0.005	0.008	*n.d.*
ON 7	3' TTT T**X**T CT**X** TCC CC	0.002	0.003	*n.d.*	0.037	*n.d.*	*n.d.*	0.009	*n.d.*	*n.d.*	0.037	0.028	*n.d.*
ON 10	3' TT**ᴸ**T TT**ᴸ**T **C**ᴸT**X** TC**ᴸ**C **C**ᴸC	0.001	0.003	0.05	0.179	0.158	0.073[c]	0.008	0.032	0.081[c]	0.046	0.081	0.038[c]

[a] C = 1.0 µM of each strand in 20 mM sodium cacodylate, 100 mM NaCl, 10 mM $MgCl_2$, pH 5.0, 6.0 and 7.2; [b] *n.d.* not determined because of low thermal stability of triplexes; [c] exciplex band at 460 nm was observed.

B. Discussion

The fluorescence emitted by single stranded **ON2** and **ON5**-**ON7** was strongly quenched with Φ_F in a range from 0.001 to 0.004, which was most probably a result of communication with neighbouring bases including electron transfer.[225] Upon binding of these TFOs to the dsDNA (**D1**), an increase in fluorescence intensity was detected at pH 5.0. As parallel triplexes are pH sensitive and more thermally stable complexes are formed at acidic conditions due to required protonation of cytosine, fluorescence quantum yields were higher at pH 5.0 than at pH 6.0. For example, Φ_F for **ON2/D1** was almost three times higher at pH 5.0 (0.179) than at pH 6.0 (0.061). A correlation between thermal stability of parallel triplexes and fluorescence quantum yields can also be illustrated by comparison of **ON2** and **ON5**. Thus, T_m for **ON5/D1** at pH 6.0 was 9.0 °C lower than the T_m-value for **ON2/D1**, which

resulted in a lower Φ_F value (0.023 instead of 0.061). Surprisingly, a large difference was observed in fluorescence quantum yields between parallel triplexes, parallel and antiparallel duplexes formed by the same ON. Almost complete quenching of fluorescence was detected for all parallel duplexes, even at pH 5.0, when complexes were most stable. Despite the higher thermal stability, a lower Φ_F value for antiparallel duplex **ON2/ON16** was observed in comparison with triplex **ON2/D1** at pH 6.0. The contrary situation was observed for a different TFO with the single insertion of **X** (**ON5**) at the same pH, i.e. $\Phi_{F(ON5/ON16)} > \Phi_{F(ON5/D1)}$, which was an outcome of low triplex thermal stability ($T_{m(ON5/D1)}$ = 16.0 °C). With lowering the pH to 5.0, fluorescence intensities of triplexes were much higher than those for antiparallel duplexes (see **ON2**, **ON5** and **ON6** toward **D1,** Table 7.2). All measurements of Φ_F were performed at 10 °C. Because of partial melting of parallel triplexes, lower Φ_F values were detected at 20 °C (results are not shown). To ensure that these effects are not a consequence of interactions of the pyrene residue in the structure **X** with metal ions presented in cacodylate buffer solution, we performed measurements in medium salt sodium phosphate buffer at pH 5.0 and got the same tendency as described above in fluorescence spectra upon binding of **ON2** to **D1**, **ON15** and **ON16** (results are not shown).

a)

b)
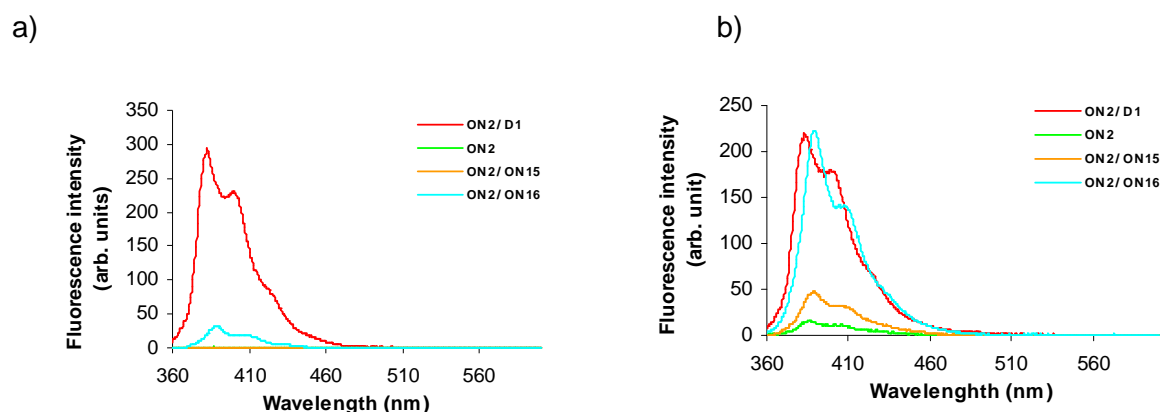
Figure 7.7. Steady-state fluorescence emission spectra of **ON2** and the corresponding complexes with D1, **ON15** and **ON16** upon excitation at 350 nm. Spectra of 1.0 μM solutions were recorded in thermal denaturation buffer at 10 °C at pH 5.0, emission slit 0.0 nm (Figure 7.7a) and pH 6.0, emission slit 2.5 nm (Figure 7.7b).

C. The applicability of the modified oligonucleotides as nucleic acid diagnostics

So far, the discrimination in fluorescence properties upon binding of ONs possessing pyrene residue to dsDNA and ssDNAs has never been reported. The strong increase of fluorescence intensity upon binding to dsDNA is a very important feature and an advantage of the described probe, as it opens novel perspectives for the detection of nucleic acids using triplex formation. Earlier described attempts were devoted to the development of

probes based on the principle of molecular beacons, i.e. a fluorophore and a quencher were placed separately at the 3' and 5'-ends of the TFO, which was extended with a number of complementary bases in order to make a hairpin, thus placing a fluorophore and a quencher in a close proximity.[226] No signal was observed for the probe alone, but the presence of dsDNA resulted in a probe hybridization followed by increase in fluorescence intensity. However, it is known that molecular beacons can sometimes open up in cell media after entering the cell in the absence of complementary targets, bringing false positives in the detection assays.[227] Existing nucleic acids staining dyes, based on small molecules, like TOTO (Figure 7.8) and others,[228] which are used nowadays for labelling of dsDNA, have two main disadvantages. First, such staining dyes bind to dsDNA unspecifically or have preferences only to certain regions, which usually are repetitive in the genome.[229] Second, these dyes have a very low affinity to triplexes as a consequence of not-fitting shape of molecules to the base triplet and due to electrostatic repulsion between C[+]•G-C triad and positively charged structure of dyes.[230] For this reason, a combination of TFO and a free staining dye cannot be used for a sequence selective labelling of dsDNA. Therefore, nucleic acids probes having one or several modified nucleotides with a fluorophore, which give a positive signal upon binding to complementary strands, are of high importance for any detection technique.
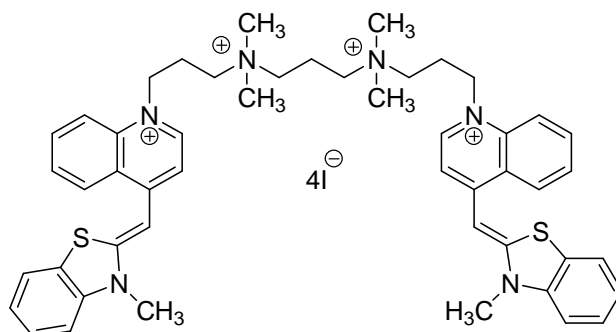
Figure 7.8 TOTO, a currently used quinolinium bichromophore nucleic acid stain

The main drawback of the monomer **X** is its destabilizing effect upon triplex formation. For this reason, attempts were made to develop a TFO, which can form stable triplexes at realistic cell pH's 6.0 and 7.2, and at the same time shows the same favourable fluorescence intensity increase observed for the above described ONs.

C.1 The incorporation of TINA

In **ON8** and **ON9**, we used bulged insertions of TINA monomer (Figure 7.6), known to stabilize parallel triplexes with ΔT_m up to 19.0 °C for a single insertion of **p**.[231] A disadvantage of TINA monomer in this particular case is the small difference between the excitation wavelengths for **X** and **p**, 350 and 373 nm respectively, meaning that irradiation of only one of these monomers is hardly achieved. We assumed that by placing **p** as left or right bulged neighbour to the modification **X** in **ON2**, strong monomer fluorescence coming from TINA in a single-stranded state could be transferred to an excimer band appearing at longer wavelengths by interaction of two pyrenes. In contrast to these considerations, further destabilization of parallel triplexes was observed (Table 1, **ON8** and **ON9** toward **D1** in comparison with **ON2/D1** at pH 6.0). Moreover, high monomer fluorescence was detected for ss**ON8** and ss**ON9** while only very small excimer band could be observed at 480 nm upon excitation at 350 nm. This hampered the detection of binding of these probes to dsDNA by fluorescence. However, these results do not exclude a use of TINA monomers in a combination with the monomer **X** in the future. For this purpose, fluorescently orthogonal intercalators to pyrene **X**, like naphthalene derivative,[231] should be used in TINA structure and could be placed three or four bases away from insertion of **X** in the TFO.

C.2 The incorporation of α-L-LNA

An alternative to intercalating nucleic acids is sugar modified nucleotides, which are fluorescently silent and give thermal stabilization upon insertion into TFOs. We decided to use α-L-LNA (Figure 7.6), which stabilizes triplexes up to 5 °C per insertion. As a general rule, every third or fourth nucleotide in the TFO should be substituted by α-L-LNA monomer to achieve high T_m values at neutral pH.[232] A use of 5-methyl-cytosine derivative of α-L-LNA is also an advantage for parallel triplex formation, because 5-methyl-cytosine is easier protonated than cytosine.[233] In the sequence **ON10**, five α-L-LNA nucleotides were incorporated leading to increased thermal stability in all cases (Table 7.1). Thus, at pH 6.0 a triplex stabilisation of 5 °C per insertion of α-L-LNA was observed and at pH 7.2, triplex **ON10/D1** melted at 35.5 °C. Parallel and antiparallel duplexes formed by **ON10** with **ON15** and **ON16,** respectively, were more thermally stable than the wild-type duplexes of **ON1**. Indeed at pH 7.2, the parallel duplex **ON10/ON15** showed higher T_m than the corresponding triplex **ON10/D1** with ΔT_m value of +3.0 ºC, which questions the formation of the triplex. In order to prove triplex formation of **ON10** with duplex **D1** we recorded CD spectra at different pH (Figure 7.9). The negative band at ~212 nm is considered to be a sign for a parallel homopyrimidine DNA triplex[234] and it has also been observed for triplexes formed by α-L-LNA modified TFOs.[235] As can be seen from Figure 7.9, a negative band at ~212 nm was

presented in CD spectra at all pH, which thereby confirmed the formation of a triple helix for
ON10. Melting of the triplexes at pH 7.2 and 6.0 was clearly seen on CD spectra recorded at
40 and 55 °C, respectively.
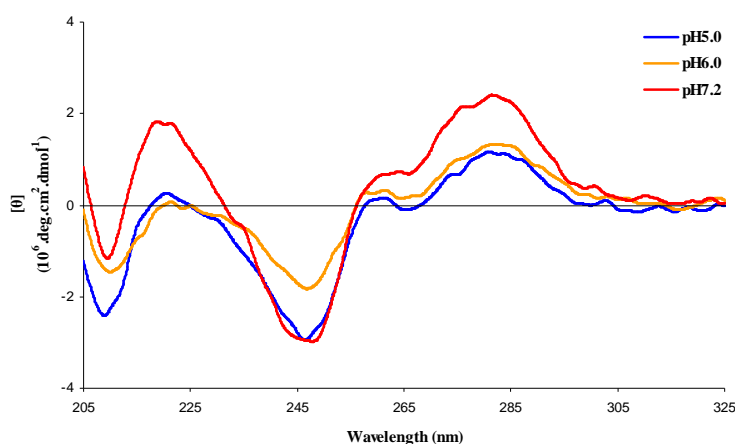
Figure 7.9 CD spectra of 1.0 µM solutions of **ON10/D1** were recorded in thermal denaturation buffer
at 10 ºC at pH 5.0, pH 6.0 and pH 7.2.

Concerning the fluorescence properties of **ON10**, a similar increase in fluorescence intensity
was observed upon triplex formation at pH 5.0 and 6.0 caused by monomer **X** (Figure 7.10).
Interestingly, triplexes formed by **ON2** and **ON10** at pH 5.0 showed the same fluorescence
quantum yield (0.179) (Table 7.2). A considerably lower decrease of Φ_F was detected for
ON10/D1 than for **ON2/D1** upon changing pH from 5.0 to 6.0, which is clearly a
consequence of higher triplex thermal stability for **ON10**. Importantly, the fluorescence
intensity for antiparallel duplex **ON10/ON16** was lower than that for the corresponding triplex
at pH 5.0 and 6.0, meaning that fluorescent discrimination caused by **X** in **ON2** between
triplex and antiparallel duplex remained after substitutions of some natural bases in TFO by
α-L-LNA.

a)
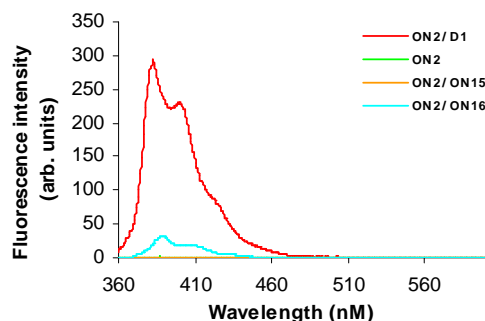
b)
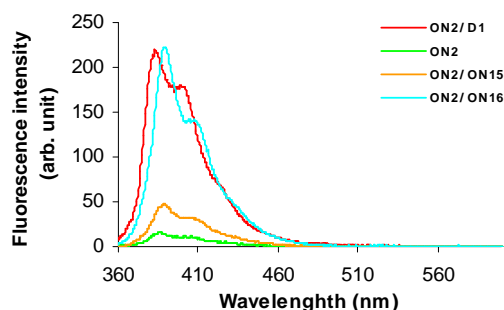
Figure 7.10 Steady-state fluorescence of ON10 and the corresponding complexes with **D1**, **ON15**
and **ON16** upon excitation at 350 nm. Spectra of 1.0 µM ON solutions were recorded in thermal
denaturation buffer at 10 °C at pH 5.0 (Figure 7.10a) and pH 6.0 (Figure 7.10b).

At pH 7.2, parallel duplex and triplex with **ON10** were less stable due to the deprotonation of cytosine (Table 7.1). Under these conditions **ON10** formed a triplex with T_m of 35.5 °C. However, deprotonation of cytosine had a large influence on the fluorescence properties of **ON10** and its complexes at pH 7.2. In the fluorescence spectra of ss**ON10** we observed a change in the ratio of band III/I (Figure 7.10), which was reversed in comparison with ss**ON2** (Figure 7.7). Moreover, fluorescence intensity of the probe alone was slightly increased upon changing the pH from 5.0 to 6.0 and then to 7.0 (ss**ON10**, Figure 7.10 and Figure 7.11). It is known that the ratio of band III/I is affected by local environmental polarity[236] and changes in fluorescence observed were associated with changing the pH and deprotonation of cytosine. Upon binding of **ON10** to complementary dsDNA and ssDNAs, a new broad emission appeared at longer wavelength (460 nm, Figure 7.10). However, it differed from an excimer emission band (480-500 nm), which usually is caused by interactions of at least two pyrene units positioned in a closed proximity. Similar results have been observed by Yamana et al.[237] upon duplex formation of pyrene labeled oligonucleotides. The increased intensity ratio of band III/I was associated with a decrease in local environmental polarity. In our case the pyrene residue was most probably transferred into the more hydrophobic base-pair pocket, which led to a new broad emission band around 460 nm. This emission can be ascribed to the exciplex of the pyrene and an adjacent base. This effect was observed upon triplex formation, but was even more obvious in the fluorescence spectrum for the parallel duplex at pH 7.2 (Figure 7.11). The latter, as we have already mentioned above, is less rigid than the triplex thus providing extra space for interaction of pyrene with neighboring bases.
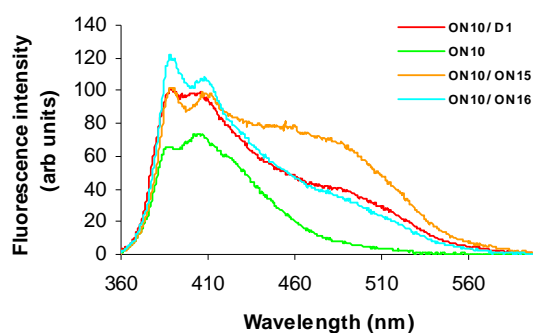
Figure 7.11 Steady-state fluorescence of **ssON10**, corresponding triplexes with **D1**, and parallel and antiparallel duplexes with **ON15** and **ON16**, respectively. Spectra were recorded in thermal denaturation buffer at pH 7.2 at 10 °C, using an excitation wavelength of 350 nm.

The melting temperatures of the single mismatched triplexes formed by **ON2** and **ON10** with duplexes **D2**-**D4** were determined at pH 5.0 and 6.0 and are shown in Table 7.3. As expected, the mismatched triplexes formed with **ON2** showed much lower thermal stability compared to matched triplexes (ΔT_m > - 25.0 °C) and no changes in fluorescent spectra of ss**ON2** were observed upon its hybridization to mismatched dsDNAs. The presence of α-L-

LNA in **ON10** resulted in formation of considerably more stable mismatched triplexes in comparison with **ON2**. Thus at pH 5.0 the difference in T_m values for matched and mismatched triplexes was not exceeding 7.0 °C showing higher drop in T_m toward duplex **D3**. Nevertheless, formation of mismatched triplexes with **ON10** was proved by the presences of a negative band at ~212 nm in theirs CD spectra at 10 °and T_m-values obtained by UV at 260 nm were verified by monitoring changes in θ at 210 nm in CD against increased temperature (Table 3 and Supporting Information). Interestingly that at pH 5.0 despite of the high thermal stability of triplexes **ON10/D2-D4**, where thymidine **X** was not matching with bases of duplexes, we observed a ~20 fold decrease in fluorescence intensity for all three mismatched triplexes in comparison with the perfectly matched **ON10/D1** (Figure 6). A high level of quenching in fluorescence and high thermal stability of mismatched triplexes with **ON10** can be a result of the interaction of pyrene moiety with adjacent nucleobases which can be realized if pyrene moved from outside the triplex into the triplex interior. The latter is possible because of imperfect base-pairing and consequent repulsion of thymine at mismatched positions making a room for pyrene intercalation. The excellent discrimination in fluorescence for **ON2** and **ON10** observed between matched and mismatched triplexes can be used for detection of dsDNA via formation of perfectly matched parallel triplexes.

Table 7.3 T_m [°C] Data for Mismatched Parallel Triplex of **ON2**, Taken from UV Melting Curves (λ= 260 nm)[a] and Fluorescence Quantum Yields at 350 nm (in brackets).

	3' TTT TTT CT**X** TCC CC (**ON2**)		3' TT**T**^{L}T TT**T**^{L}T **C**^{L}T**X** TC**C**^{L}C **C**^{L}C (**ON10**)	
	pH = 5.0	pH = 6.0	pH = 5.0	pH = 6.0
3' AAA AAA GA**C** AGG GGC AG 5' TTT TTT CT**G** TCC CCG TC	28.0 (0.001)	<5.0	62.0^b [62.0]^c	41.5^b [42.0]
3' AAA AAA GA**G** AGG GGC AG 5' TTT TTT CT**C** TCC CCG TC	21.0 (0.001)	<5.0	58.5^b [59.0]	34.5^b [35.0]
3' AAA AAA GA**T** AGG GGC AG 5' TTT TTT CT**A** TCC CCG TC	29.5 (0.001)	<5.0	61.0^b [*n.p*]^d	40.5^b [*n.p*]

For T_m measurement: C = 1.5 μM of **ON2** or **ON10** and 1.0 μM of each strand of mismatched dsDNA in 20 mM sodium cacodylate, 100 mM NaCl, 10 mM MgCl$_2$, pH 5.0 and 6.0; for quantum yield measurement: C = 1.0 μM of each strand in the same buffer at pH 5.0; [b] Third strand and duplex melting overlaid. T_m was determined at 350 nm; [c] Values in brackets are determined with CD T_m measurements; [d] *n.p.* not possible to determine.
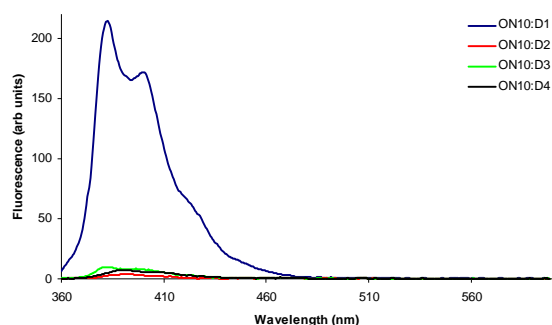
Figure 7.12. Steady-state fluorescence of **ON10** and the corresponding complexes with **D1**-**D4** upon excitation at 350 nm. Spectra of 1.0 μM dsDNA with 1.5 μM of **ON10** were recorded in thermal denaturation buffer at 10 °C at pH 5.0, emission slit 0.0 nm.

Although the formation of the exciplex band was unexpected at pH 7.2 for complexes of **ON10**, it opens new possibilities for the use of this probe for diagnostic purposes. Fluorescence intensity upon hybridization to dsDNA at pH 7.2 was much lower (ca. 100 arb.units, Figure 7.10) than at pH 6.0 (750 arb. units, Figure 7.7). Such pH-depending fluorescence signal can be useful for dsDNA detection in cells with lower pH values. Tumor cells have a natural tendency to overproduce acids, resulting in more acidic pH values in the tumor microenvironment, which can be down to 6.4 for human cancers, and down to 6.12 for mouse cancers, depending on the tumor type.[238] Therefore, the described probe is an ideal platform to develop sequence-specific detectors for tumor cells. Moreover, as COMBO-FISH experiments are so far performed under conditions with pH-values around 6.0,[209] probes based on the design of **ON10** are suitable for such experiments.

7.2.6   Mixed purine-pyrimidine oligonucleotides

7.2.6.1   Thermal stability

To study the influence of the monomer **X** on Watson–Crick base-pairing, a mixed sequence was synthesized with one, two or three neighbouring modified nucleotides incorporated. Table 7.4 shows the thermal stabilities of **ON11-14** upon hybridization to ssDNA/ssRNA and the fluorescence quantum yields for **ON12** and its complexes determined at 350 nm.

<div style="border:1px solid #000;">

Table 7.4 T_m [°C] Data for Duplex Melting of **ON11**-**14** and DNA/RNA Complements, taken from UV Melting Curves (λ= 260 nm)[a] and Fluorescence Quantum Yields for ON12 determined (in brackets).

</div>

		5' GTG AAA TGC DNA (**ON17**)	5' GUG AAA UGC RNA (**ON18**)
ON11	3' CAC TTT ACG	36.0	34.0
ON12	3' CAC T**X**T ACG (0.02)[b]	20.0 (0.04)	19.0 (0.07)
ON13	3' CAC T**XX** ACG	< 5.0	13.5
ON14	3' CAC **XXX** ACG	< 5.0	< 5.0

[a]Thermal denaturation temperatures measured as the maximum of the first derivative of the melting curve (A_{260} vs temperature) recorded in high salt buffer (1400 mM NaCl, 20 mM sodium phosphate, pH 7.0), using 1.0 μM concentrations of each complementary strands; [b] Fluorescence quantum yields were measured in the aerated thermal denaturation buffer at 10 °C, using 1.0 μM concentrations of each complementary strands at excitation wavelength of 350 nm.

Earlier reports for 2',5'-linked oligoribonucleotides showed a decrease of 1.3 °C for each 2',5'-linkage in ON in their thermal stability toward ssRNA and a much higher decrease toward ssDNA.[239] As no transition in melting curves was observed in a medium salt phosphate buffer for duplexes with **ON12**-**14,** we used a high salt concentration for this study. When adding a pyrene unit to the 3'-position of 2',5'-linkages (structure **X**), a considerable decrease in antiparallel duplex thermal stability, more than 15 °C per modification, was observed in comparison with the wild-type duplex (**ON11**/**ON17**). This indicates that pyrene could not stabilise duplexes by intercalation, but most probably was directed towards the outside of the duplex.

7.2.6.2 Fluorescence

The inability of pyrene to intercalate can also be seen in the steady-state fluorescence emission spectra of ON12 and the corresponding duplexes with ssDNA and ssRNA (Figure 7.13). Interesting to mention is that pyrene fluorescence was not quenched in ssON12 as much as we have already seen for single-stranded homopyrimidine sequences in Table 7.2. This can be a result of the presence of purines in the same sequence for ON12. Upon hybridisation of ON12 with ssDNA a two-fold increase in fluorescence intensity was observed. Binding to complementary ssRNA resulted in an approximately 3.5-fold increase in fluorescence intensity. Similar results have previously been described for 2'-pyrene-modified oligoribonucleotides. The 2'-N-position of 2'-amino-DNA[212] and –LNA[213] or the 2'-

O-position of RNA[214] have been substituted with pyrene via a short tether. In these cases the increase of fluorescence upon hybridisation to ssRNA was considerably stronger than in the case for binding to ssDNA.
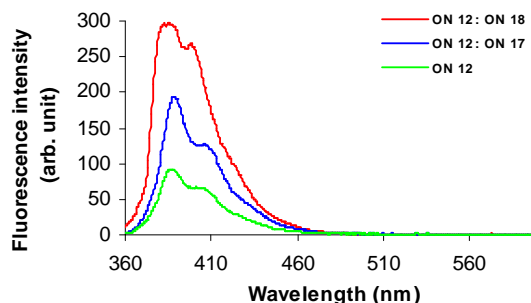
Figure 7.13 Steady-state fluorescence emission spectra of **ON12** and the corresponding duplexes with complementary ssDNA/RNA. Spectra were recorded in the thermal denaturation buffer at 10 °C using an excitation wavelength of 350 nm and an excitation slit of 4.0 nm and emission slit of 2.5 nm.

This short study about the influence of **X** on Watson-Crick base pairing shows that a much larger decrease in stability is obtained, while a smaller increase in fluorescence signal upon binding can be observed. Duplexes with RNA show a higher increase compared to duplexes with DNA, which is similar to previous reported analogue studies.

7.2.7 Modelling

Molecular modelling studies were performed on the complexation of a truncated dsDNA with **ON2** (3'-TTC T**X**T CCC-5') in order to visualize different conformations of the pyrene modified 2',5'-thymidine (Structure X). The parallel triplex was analysed with pyrene located either inside or outside the triplex. As can be seen from comparing Figure 7.14a and 7.14b a larger distortion of the triplex structure is observed when the pyrene was intercalating with bases of the triplex (Figure 7.14a). In agreement with our observation of very strong fluorescence signal for **ON2/D1**, Figure 7.14 b and 7.14c show the projected conformation of the pyrene modified 2',5'-thymidine. Despite of such a large distortion of the triplex structure upon intercalation of pyrene, this was most probably the case for triplex **ON10/D1** at pH 7.2 when exciplex band was observed in the fluorescence spectrum.

a)   b)

c)   d)

e)   f)
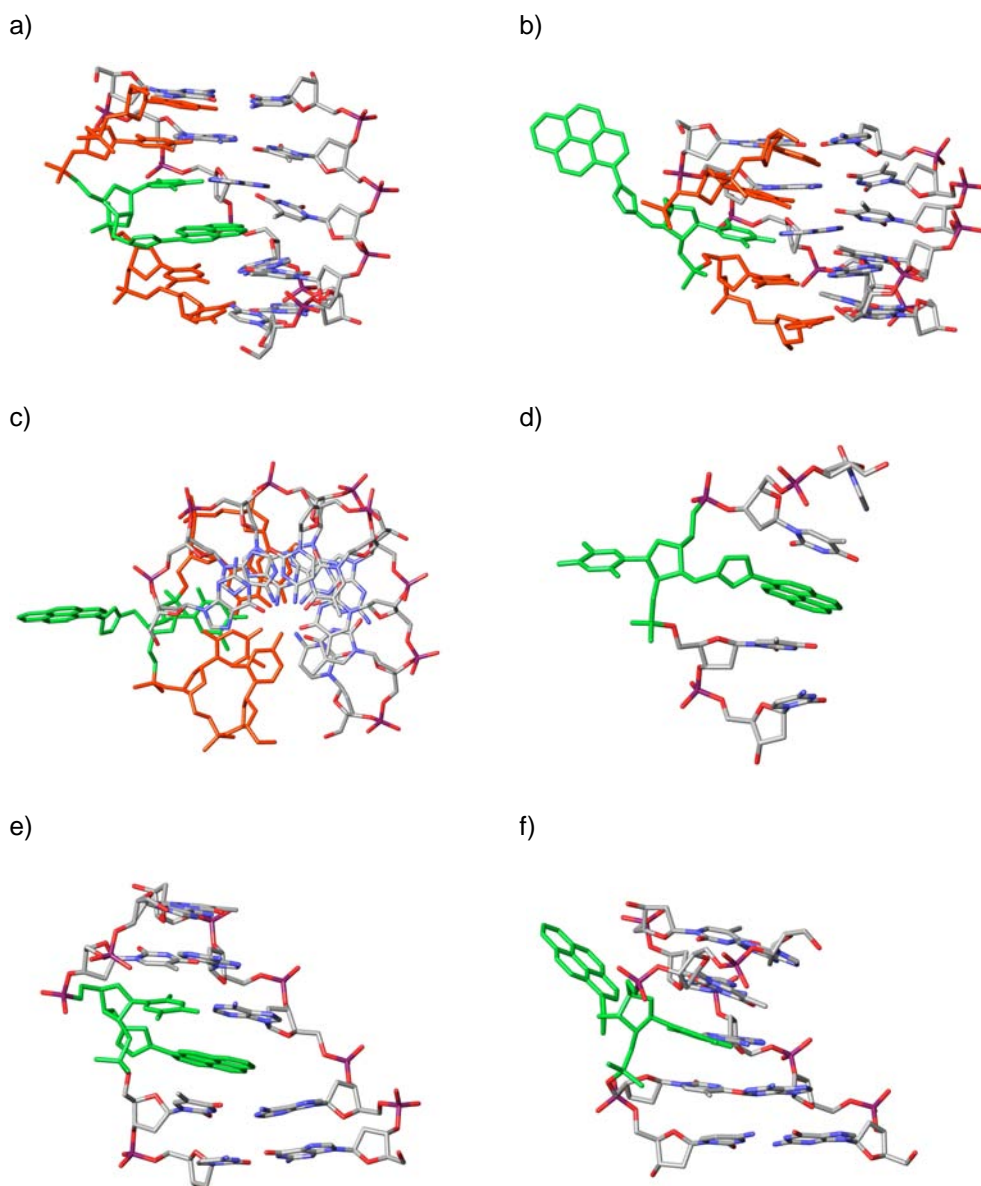
Figure 7.14 Representative truncated structures obtained by molecular modelling studies of the triplex, ssDNA and dsDNA. The modified nucleotide is shown in green and the TFO is shown in orange. a) Triplex formed by **ON2** with dsDNA with the pyrene moiety inside the triplex. b), c) The triplex formed by **ON2** with dsDNA with the pyrene moiety outside of the triplex, respectively sideview and topview. d) ssDNA, **ON2**, with the pyrene moiety placed between neighbouring nucleobases. e), f) The dsDNA formed by **ON12** with complementary DNA, with the pyrene moiety placed respectively inside and outside the duplex.

To explain the practically total quenching of fluorescence for ssONs possessing X, we performed molecular modelling analysis of the truncated single stranded **ON2** (3'-TTC T**X**T CCC-5'). Studies showed that rotation of the modified nucleotide, with the pyrene moiety placed between neighbouring nucleobases, resulted in stacking of pyrene with these nucleobases, giving a fluorescently silent oligonucleotide (Figure 7.14d). Upon hybridization

to complementary dsDNA, pyrene would be forced outside the triplex, thereby explaining the increase in the fluorescence signal observed for complexes of **ON2**, **ON5** and **ON10** with matched duplexes.

Modelling studies for the duplex formed between **ON12** and complementary ssDNA showed serious distortions of the phosphate backbone and the Watson–Crick base-pairs upon positioning of the pyrene moiety between bases of dsDNA (Figure 7.14e). The intercalation of pyrene led to a more stable duplex structure with an increase in stability of ~1966 kJ/mol compared to the duplex with the pyrene moiety placed outside the duplex (Figure 7.14f). However, the conformation with pyrene intercalated in the duplex can be rejected based on our experimental results. The formation of a less stable duplex, compared to an unmodified duplex and an increase of the fluorescence signal upon hybridization of **ON12**, indicates, in contradiction to what the modelling studies show, that the conformation of the duplex is one where the pyrene moiety is directed towards the outside of the duplex.

7.2.8 Conclusion

A novel 3'- substituted ribothymidine analogue possessing a pyrene moiety coupled to the sugar part via a triazolyl methyl linker, compound **X**, was synthesized and incorporated as a 2',5'-linker into different ONs. Thermal stability studies showed that single incorporation of **X** in the middle of ONs instead of thymidine resulted in lowering T_m values for either Watson–Crick or Hoogsteen type complexes. While the single stranded homopyrimidine probes were fluorescently silent with the quantum yield (Φ_F) of 0.001-0.004, formation of parallel triplexes with the complementary dsDNA led to the very strong increase of the fluorescence signal with the value of Φ_F in a range from 0.023 to 0.179 depending on the site of insertion and thermal stability. Lowering the pH resulted in higher thermal stability of Hoogsteen-type triplexes and higher fluorescence quantum yields. Antiparallel duplexes of the same probes had lower values of Φ_F than parallel triplexes under the similar conditions. Molecular modelling of triplexes showed that the pyrene-containing residue at the 3'-position of 2',5'-ribothymidine prefers to be positioned outside the triplex since it caused minimal distortion of the complex.

To compensate the low thermal stability induced by **X**, known triplex stabilizators were additionally incorporated into ONs. In a first attempt phenylethynylpyrene glycerol (TINA) was incorporated as a neighbouring bulge to the modification **X**. However, neither stabilisation nor the previously detected fluorescence enhancement was observed in triplexes formed by these ONs. For this reason α-L-LNA, a non-fluorescent nucleotide,

monomer was used to stabilize triplexes. As a result of five substitutions of native nucleic bases by α-L-LNA in the homopyrimidine sequence possessing **X**, enhancement in thermal stability and fluorescent quantum yield was observed at pH 6.0 and 5.0 upon parallel triplex formation. This oligonucleotide had very low fluorescence signal as single stranded probe and marginal changes in fluorescence were observed upon formation of mismatched triplexes, in which thymidine **X** was not matching with bases of dsDNAs. The latter means that fluorescence discrimination between homopurine dsDNA stretches and sequences with single base pair inversion is feasible. Moreover single mismatched triplexes showed an almost unchanged thermal stability, but with a high degree of quenching, which means that specific detection of matched dsDNA is possible. Upon binding to dsDNA at pH 6.0, a very specific 50-fold increase of the fluorescence signal was observed that was much stronger than that for antiparallel duplex formed by the same probe. However, at pH 7.2 a lower increase in fluorescence in combination with a shift to a longer wavelength, known as exciplex, was observed. This feature allows using that kind of probe to detect species of dsDNA in an environment of tumor cells, which are known to overproduce acids and have lower pH values than normal cells.

For the first time, a pyrene labelled oligonucleotide probe was developed that explores the increase of fluorescence signal to discriminate between dsDNA over ssDNAs. This is a promising platform for further development of fluorescent probes, which can bind a sequence selectively to dsDNA without prior denaturation of the target and give a positive signal upon hybridization. A perspective for such probes is their use in fluorescence *in vivo* hybridization under vital conditions and in living cells.

7.2.9 Experimental section

General Synthesis Procedures

NMR spectra were recorded on a Varian Gemini 2000 spectrometer at 300 MHz for ^{1}H using TMS (δ: 0.00) as an internal standard and at 75 MHz for ^{13}C using CDCl$_3$ (δ: 77.0) or DMSO (δ: 39.44) as an internal standard. Accurate ion mass determinations of the synthesized compounds were performed using the 4.7 T Ultima Fourier transform (FT) mass spectrometer (Ion Spec, Irvine, CA). The [M + Na]$^+$ ions were peak matched using ions derived from the 2,5-dihydroxybenzoic acid matrix. MALDI-TOF mass spectra of isolated oligodeoxynucleotides (ONs) were determined on a Voyager Elite biospectrometry research station (PerSeptive Biosystems). Thin-layer chromatography (TLC) analyses were carried out with use of TLC-plates 60 F$_{254}$ purchased from Merck and were visualized in UV light (254 nm). The silica gel (0.040-0.063 mm) used for column chromatography was purchased

from Merck. Solvents used for column chromatography were distilled prior to use, while reagents were used as purchased.

1-{3-[(4-(pyren-1-yl)-1*H*-1,2,3-triazol-1-yl)methylene]-3-deoxy-β-D-ribofuranosyl}thymine (7.1)

To a solution of compound **4.26** (420 mg, 1.41 mmol) in 8 mL of a DMF/H_2O 19:1 mixture in a microwave tube, freshly prepared 1M aq. $CuSO_4$ (283 μM, 0.28 mmol) was added. The solution was bubbled with Ar for 1 minute. Afterwards, a freshly prepared aqueous sodium ascorbate solution (424 μM, 0.42 mmol) and 1-ethynylpyrene (480 mg, 2.12 mmol, dissolved in 8 mL DMF/ H_2O 19:1) were added. The mixture was bubbled again with Ar for 3 minutes. The resulting mixture in a sealed microwave tube was placed in a microwave synthesiser Emrys Creator. The heating temperature was set to 125 °C, with a 30 s premixing time. The reaction mixture was irradiated for 15 min followed by N_2 cooling to 40 °C. Water was added and the mixture was kept at 6 °C during 5 h for complete precipitation of the triazol and unreacted alkyn. The precipitate was filtered off and washed with water to remove all catalysts. The solid on the filter was washed with methanol to dissolve the triazol. The collected organic phase was evaporated to dryness and the residue was purified by silica gel column chromatography (CH_2Cl_2/MeOH 95:5) to yield the pure triazol **7.1** (583 mg, 79%).
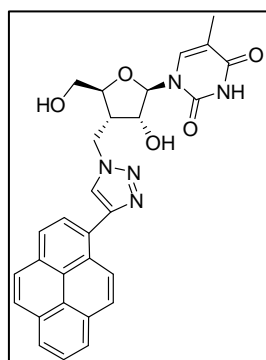
¹H NMR (300 MHz, DMSO): δ 1.79 (3H, s, 5-CH_3), 2.96 (1H, m, H-3'), 3.51 (1H, d, *J* = 12.3 Hz, H-5$_a$'), 3.81 (1H, d, *J* = 12.6 Hz, H-5$_b$'), 4.19 (1H, m, H-4'), 4.32 (1H, m, H-2'), 4.67 (1H, dd, *J* = 6.0 and 14.1 Hz, H-6$_a$'), 4.81 (1H, dd, *J* = 7.5 and 14.1 Hz, H-6$_b$'), 5.33 (1H, br s, 5'-OH), 5.78 (1H, br s, 2'-OH), 6.24 (1H, d, *J* = 5.1 Hz, H-1'), 8.04-8.41 (9H, m, pyrene-H and H-6), 8.77 (1H, s, triazol-H), 8.91 (1H, d, *J* = 6.3 Hz, pyrene-H), 11.31 (1H, br s, NH).

¹³C NMR (75 MHz, $CDCl_3$): δ 12.24 (5-CH_3), 41.17 (C-3'), 46.10 (C-6'), 60.12 (C-5'), 74.80 (C-2'), 82.77 (C-4'), 90.87 (C-1'), 108.30 (C-5), 123.86, 124.24, 124.89, 125.04, 125.10, 125.38, 125.46, 126.41, 127.01, 127.27, 127.43, 127.61, 127.92, 130.31, 130.49, 130.87 (pyrene and triazol C-5), 136.14 (C-6), 146.03 (triazol C-4), 150.39 (C-2), 163.80 (C-4).

Exact mass (ESI-MS) for $C_{29}H_{25}N_5O_5Na$ [M+Na]$^+$ found, 546.1769; calcd, 546.1753.

1-{3-deoxy-5-*O*-(4,4'-dimethoxytriphenylmethyl-3-[(4-(pyren-1-yl)-1*H*-1,2,3-triazol-1-yl)methylene]-β-D-ribofuranosyl}thymine (7.2)

Compound **7.1** (530 mg, 1.01 mmol) was coevaporated three times with pyridine and dissolved in anhydrous pyridine (4 mL). 4,4'-Dimethoxytrityl chloride (461 mg, 1.36 mmol) was added under N_2-atmosphere. After 16 h, MeOH (0.5 mL) and EtOAc (20 mL) were added and the mixture was quenched with aqueous $NaHCO_3$-solution (20 mL). The water layer was extracted with EtOAc (2 x 20 mL). The combined organic layers were dried over $MgSO_4$, evaporated to dryness and purified by column chromatography (CH_2Cl_2/MeOH/ pyridine 97/2.5/0.5). Compound **7.2** was obtained (760 mg, 91%) as a yellow foam.
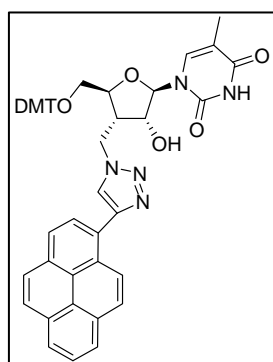
^{1}H NMR (300 MHz, $CDCl_3$): δ 1.42 (3H, s, 5-CH_3), 3.10 (1H, m, H-3'), 3.29 (2H, m, H-5$_a$' and H-5$_b$'), 3.67 (6H, s, 2 × OCH_3), 4.39 (3H, m, H-2', H-6$_a$' and H-6$_b$'), 4.71 (1H, m, H-4'), 5.71 (2H, m, H-1' and 2'-OH), 6.79 (4H, d, *J* = 8.7 Hz, DMT), 7.15-7.43 (10H, m, DMT and H-6), 7.72 (1H, s, triazol-H), 7.89-8.10 (8H, m, pyrene-H), 8.61 (1H, d, *J* = 9.3 Hz, pyrene-H), 10.91 (1H, br s, NH).

^{13}C NMR (75 MHz, $CDCl_3$): δ 12.06 (5-CH_3), 42.86 (C-3'), 49.69 (C-6'), 55.11 (OCH$_3$), 72.57 (C-2'), 75.95 (C-5'), 82.01 (C-4'), 87.53 and 87.08 (C-1' and *C*(Ar$_3$)), 110.64 (C-5), 113.35, 124.93-131.28, 135.40 (pyrene, DMT and triazol C-5), 136.11 (C-6), 144.18 (triazol C-4), 146.55 (DMT), 150.58 (C-2), 158.70 (DMT), 164.55 (C-4).

HRMS (ESI-MS) for $C_{50}H_{43}N_5O_7Na$ [M+Na]$^+$ found, 848.3084; calcd, 848.3060.

1-{2-*O*-[2-cyanoethoxy(diisopropylamino)phosphino]-3-deoxy-5-*O*-4,4'-dimethoxy-triphenylmethyl-3-[(4-(pyren-1-yl)-1*H*-1,2,3-triazol-1-yl)methylene]-β-D-ribofuranosyl}thymine (7.3)

Compound **7.2** (510 mg, 0.62 mmol) was dissolved under N_2-atmosphere in anhydrous CH_2Cl_2 (10 mL). *N*,*N*-Diisopropylammonium tetrazolide (159 mg, 0.93 mmol) was added followed by the dropwise addition of 2-cyanoethyl tetraisopropylphosphordiamidite (223 mg, 0.74 mmol) under external cooling with an ice-water bath. After 16h the reaction was quenched with H_2O (6 mL). The layers were separated and the organic phase was washed with H_2O (6 mL). The water layers were washed with CH_2Cl_2 and the resulting organic phases were combined, dried over $MgSO_4$ and filtered. The solvent was removed under reduced pressure, and the residue was purified using silica gel column chromatography ($CHCl_3$/MeOH/pyridine 99/0.5/0.5). The purified compound **7.3** (412 mg, 65%) was obtained as a foam and was further used in DNA synthesis.
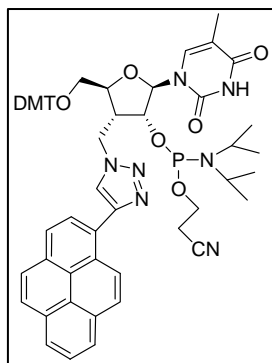
31P NMR (CDCl₃) δ 148.40, 150.34 in a ratio of 2:3.

Exact mass (ESI-MS) for $C_{59}H_{61}N_7O_8P$ [M+H]⁺ found, 1026.4338; calcd, 1026.4318.

Synthesis and Purification of Modified Oligonucleotides.

DMT-on oligodeoxynucleotides were synthesized in a 0.2 μmol scale on CPG supports using Expedite Nucleic Acid Synthesis System Model 8909 (Applied Biosystems). Standard procedures were used for phosphoramidite **7.3** with 4,5-dicyanoimidazole as an activator except for extended coupling time (10 min) and an increased deprotection time (100 s), resulting in step-wise coupling yields of 98% for monomer **7.3** and >99% for unmodified DNA phosphoramidites.

The obtained DMT-on oligonucleotides bound to CPG-supports were treated with 32% aqueous ammonia (1.3 mL) at room temperature for 2 hours and then at 55 °C overnight. Purification of the 5'-*O*-DMT-on ONs was carried out by using a reverse-phase semipreparative HPLC on a Waters Xterra MS C₁₈ column. DMT groups were cleaved by treatment with 80% aq. AcOH (100 μL) for 20 minutes, followed by addition of H₂O (100 μL) and 3 M aq NaOAc (50 μL). The ONs were precipitated from 99% EtOH (600 μL). The precipitate was washed with chilled 70% aqueous ethanol. The purity of the obtained ONs was checked by ion-exchange chromatography on a LaChrom system (Merck Hitachi) using a GenPak-Fax column (Waters) and it was found to be >82% for all ONs.

Table 7.5 MALDI-MS of synthesized ONs

Oligonucleotides		m/z [M+H]$^+$ Calcd (Da)	m/z [M+H]$^+$ Found (Da)	Purity
ON 2	3'-TTT TTT CT**X** TC CCC-5'	4403.2	4404.5	100%
ON 3	3'-TTT TTT C**XX** TC CCC-5'	4684.5	4684.8	100%
ON 4	3'-TTT TTT C**XX X**C CCC-5'	4965.8	4966.8	100%
ON 5	3'-TTT T**X**T CTT TC CCC-5'	4403.2	4404.3	100%
ON 6	3'-TTT TTT C**X**T **X**C CCC-5'	4684.5	4682.5	100%
ON 7	3'-TTT T**X**T C**X**T TC CCC-5'	4684.5	4683.0	100%
ON 8	3'-TTT TTT CT**pX** TC CCC-5'	4870.6	4871.1	100%
ON 9	3'-TTT TTT CT**Xp** TC CCC-5'	4870.6	4871.1	100%
ON 10	3' TT^{L}T TT^{L}T **C**^{L}T**X** TC^{L}C **C**^{L}C	4585.2	4584.9	100%
ON 12	3'-CAC T**X**T ACG-5'	2955.2	2953.6	100%
ON 13	3'-CAC T**XX** ACG-5'	3236.5	3236.9	100%
ON 14	3'-CAC **XXX** ACG-5'	3517.8	3518.0	82%

Fluorescence Measurements

Fluorescence measurements were performed on a Perkin-Elmer luminescence spectrometer LS-55 fitted with a Julabo F25 temperature controller using quartz optical cells with a pathlength of 1.0 cm. The spectra were recorded at 10 °C in the same buffer as for T_m studies using a 1.0 mM concentration of each ONs. In all cases, absorption in the range from 360 to 600 nm of the solutions did not exceed 0.10, in order to avoid inner-filter effects, and was never less than 0.01, to avoid uncertainties. Corrections were made for solvent background.

Steady-state fluorescence emission spectra (360-600 nm) were obtained as an average of five scans using an excitation wavelength of 350 nm, excitation slit of 4.0 nm, emission slit of 2.5 nm or 0.0 nm and scan speed of 120 nm/min. The fluorescence quantum yield (Φ_F) of anthracene (ANT) in cyclohexane in this experimental setting was measured to be 0.32 relative to 9,10-diphenylanthracene in cyclohexane (Φ_F =1.00), which is in excellent agreement with the reported value of 0.36.[222] Emission quantum yields of pyrene-labeled ONs Φ_F(ON) were determined according to following calculations: [223]

$$\Phi_F(ON) = \Phi_F(ANT) * [A\,(ON)/A_{350}(ON)] * [1/\alpha(ANT)] * [n(H_2O)^2/\,n(Cyclohexane)^2]$$

In this equation Φ_F(ANT) is the cross-calibrated value for the fluorescence emission quantum yield of anthracene in cyclohexane, A (ON) is the area of the fluorescence emission spectra

of the sample from 360 to 600 nm, A_{350} (ON) is the absorbance of the sample at the excitation wavelength (350 nm), α(ANT) is the slope of the fluorescence emission area vs. A350 calibration curve for ANT, and n(H_2O) and n(Cyclohexane) are the refractive indexes of water (1.3328) and cyclohexane (1,4266), respectively.
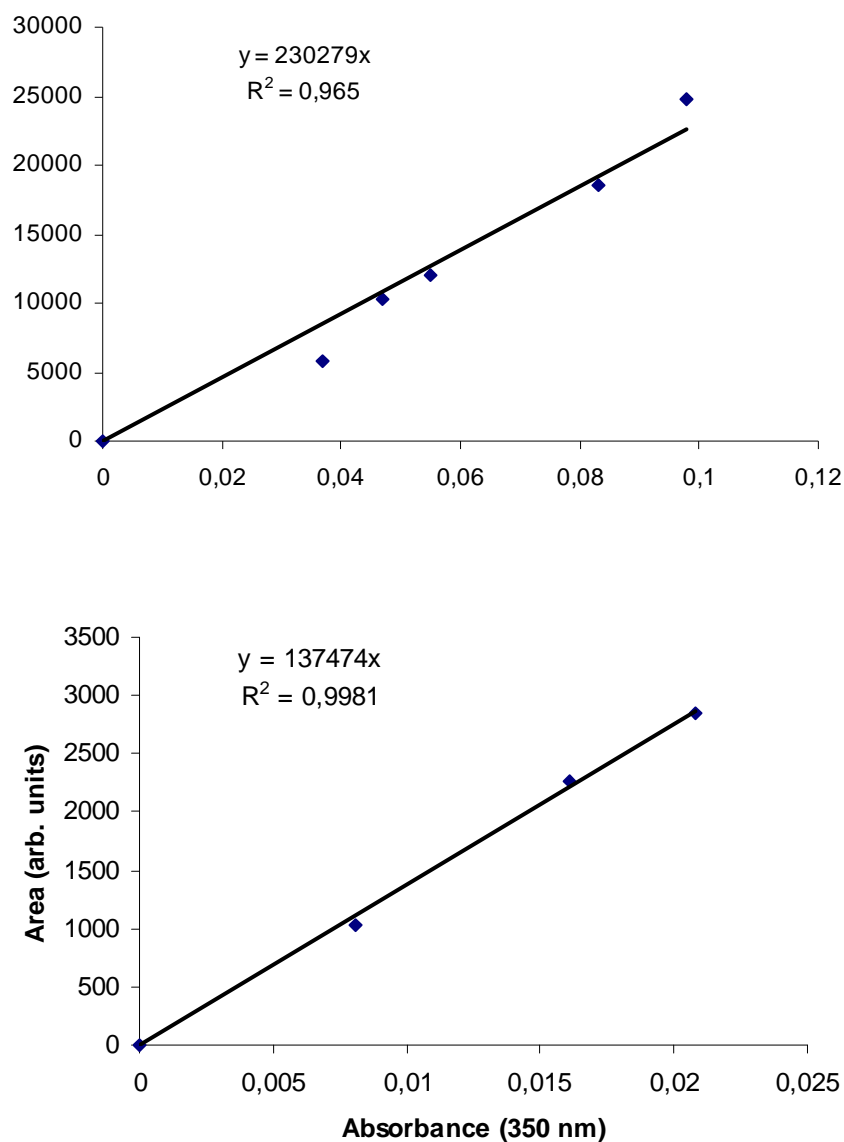

Figure 7.15 Area vs. A_{350} curves of anthracene in cyclohexane for calibration of emission quantum yield measurement (upper panel) and of triplex ON10/D1 in thermal denaturation buffer (lower panel) for determination of emission quantum yield. 'Area' refers to the area under the emission spectra (360-600 nm) using an excitation wavelength of 350 nm.

Melting Temperature Measurements

Melting profiles were measured on a Perkin-Elmer UV-vis spectrometer Lambda 35 fitted with a PTP-6 temperature programmer. The triplexes were formed by first mixing the two strands of the Watson–Crick duplex, each at a concentration of 1.0 μM, followed by the addition of the TFO at a concentration of 1.5 μM in the corresponding buffer solution. The solution was heated to 80 °C for 5 min and afterward cooled to 15 °C and kept at this temperature for 30 minutes. The duplexes were formed by mixing the two strands, each at a concentration of 1.0 mM in the corresponding buffer solution followed by heating to 80 °C for 5 minutes and then cooling to room temperature. The absorbance of both triplexes and duplexes was measured at 260 nm from 5 to 70 °C with a heating rate of 1.0 °C/min. Lower heating rate (0.5 °C/min) resulted in the same melting curves. Control experiments were also performed at 343 nm for **ON2**. The melting temperatures (T_m, °C) were determined as the maximum of the first derivative plots of the melting curves. All melting temperatures are within the uncertainty $\pm$ 0.5 °C as determined by repetitive experiments.

Molecular Modelling

Molecular modelling experiments were performed with Maestro v7.5 from Schrödinger. All calculations were conducted with AMBER* force field and the GB/SA water model. The dynamics simulations were performed with stochastic dynamics, a SHAKE algoritm to constrain bonds to hydrogen, time step 1.5 fs and simulation temperature of 300 K. Simulation for 0.5 ns with an equilibration time of 150 ps generated 250 structures, which all were minimized using the PRCG method with convergence threshold of 0.05 kJ/mol. The minimized structures were examined with Xcluster from Schrödinger, and representative low-energy structures were selected. The starting structures were generated with Insight II v97.2 from MSI, followed by incorporation of the modified nucleoside building block.

8 REFERENCES

1. World Health Organization (2007; http://www.who.int)
2. Global Alliance for TB Drug Development, *Developing a faster TB cure* (2007; http://www.tballiance.org)
3. Blumberg, H.M. et al. *Am. J. Respir. Crit. Care Med.* **2003**, *167*, 603.
4. Web of science, may 2007.
5. Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry, C.E.; Tekaia, F.; BAdcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R. et al. *Nature* **1998**, *393*, 537.
6. Minnikin, D.E. 1982: Complex lipids: their chemistry, biosynthesis and roles. In: The biology of the mycobacteria, Vol. 1, 95-184, Ratledge, C.; Stanford, J.L. (Eds), New York: Academic Press.
7. McNeil, M.R.; Brennan, P.J. *Research in Microbiology* **1991**, *142*, 451.
8. Basso, L.A.; Pereira da Silva, L.H.; Fett-Neto, A.G; Filgueira de Azevedo Jr, W.; Moreira, I.d.S.; Palma, M.S.; Calixto, J.B.; Filho, S.A.; Ribeiro dos Santos, R.; Soares, M.B.P.; Santos, D.S. *Mem. Inst. Oswaldo Cruz* **2005**, *100*, 575.
9. Trias, J.; Jarlier, V.; Benz, R. *Science* **1992**, *258*, 1479.
10. Trias, J.; Benz, R. *J. Biol. Chem.* **1993**, *268*, 6234.
11. Mukhopadhyay, S.; Basu, D.; Chakrabarti, P. *J. Bacteriol.* **1997**, *179*, 6205.
12. Engelhardt, H.; Heinz, C.; Niederweis, M. *J. Biol. Chem.* **2002**, *277*, 37567.
13. Kartmann, B.; Stengler, S.; Niederweis, M. *J. Bacteriol.* **1999**, *181*, 6543.
14. Bardou, F.; Raynaud, C.; Ramos, C.; Lanéele, M.A.; Lanéele, G. *Microbiology* **1998**, *144*, 2539.
15. Lipinski, C. A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3.
16. Ballel, L.; Field, R.A.; Duncan, K.; Young, R.J. *Antimicrob. Agents Chemother.* **2005**, *49*, 2153.
17. Ratledge, C. 1982: Nutrition, growth and metabolism. In: The biology of the mycobacteria, Vol. 1, 185, Ratledge, C.; Stanford, J.L. (Eds.), New York: Academic Press.
18. Ratledge, C. 1984: Metabolism of iron and other metals by mycobacteria. In Kubica, G.P. and Wayne, L.G. (Eds), The Mycobacteria: a sourcebook, Part A. New York: Marcel Dekker. 603.
19. Ferreras, J.A.; Ryu, J.-S.; Lello, F.D.; Tan, D.S.; Quadri, L.E.N. *Nat. Chem. Biol.* **2005**, *1*, 29.
20. Fjällbrant, H.; Ridell, M.; Larsson, L.O. *Respir. J.* **2001**, *18*, 376.
21. Fodor, T. *Tubercle* **1984**, *65*, 123.
22. Heifets, L. *Seminars in Respiratory Infections* **1986**, *4*, 242.
23. Anderson, P. *Trends in immunology* **2001**, *22*, 160.
24. Israel, H.L. *Respiratory Medicine* **1993**, *87*, 81.
25. Qunibi, W.Y.; Al-Sibai, M.B., Taher, S. et al. *Quarterly Journal of Medicine* **1990**, *77*, 1039.
26. Suzuki, Y.; Suzuki, A.; Tamaru, A.; Katsukawa, C.; Oda, H. *J. Clin. Microbiol.* **2002**, 501.
27. Healt, R.J.; White, S.W.; Rock, C.O. *Progress in lipid research* **2001**, *40*, 467.
28. Campbell, E.A.; Korzheva, N.; Mustaev, A.; Murakamin, K.; Nair, S.; Gordfarb, A.; Darst, S. *Cell* **2001**, *104*, 901.
29. Speirs, R.J.; Welch, J. T.; Cynnamon, M.H. *Antimicrob. Agents Chemother.* **1995**, *39*, 1269.
30. Zhang, Y.; Scorpio, A.; Nikaido, H.; Sun, Z. *J. Bacteriol.* **1999**, *181*, 2044.
31. Zimhony, O.; Cox, J.S.; Welch, J.; Vilchese, C.; Jacobs, W.R. *Nat. Med.* **2000**, *6*, 1043.
32. Boshoff, H.I.; Mizrahi, V.; Barry, C.E. 3rd *J. Bacteriol.* **2002**, *184*, 2167.
33. Bosne-David, S.; Barros, V.; Verde, S.C.; Portugal, C.; David, H.L. *J. Antimicrob. Chemother.* **2000**, *46*, 391.
34. Belanger, A.E.; Besra, G.S.; Ford, M.E.; Mikusova, K.; Belisle, J.T.; Brennan, P.J.; Inamine, J.M.; *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11919.
35. Takayama, K.; Kilburn, J.O. *Antimicrob. Agents Chemother.* **1989**, *33*, 1493.
36. Berg, S.; Starbuck, J.; Torrelles, J.B.; Vissa, V.D.; Crick, D.C.; Chatterjee, D.; Brennan, P.J. *J. Biol. Chem.* **2005**, *280*, 5651.
37. Liebermeister, K. *Deutsch Med. Wschr.* **1950**, *75*, 621.
38. Protivinsky, R. Antibiotics *Chemother.* **1971**, *17*, 101.
39. Mailaender, C.; Reiling, N.; Engelhardt, H.; Bossmann, S.; Ehlers, S.; Niederweis, M. *Microbiology* **2004**, *150*, 853.
40. Chacon, O.; Feng, Z.; Harris, N.B.; Careres, N.E.; Adams, L.G.; Barletta, R.G. *Antimicrob. Agents Chemother.* **2002**, *46*, 47.
41. De Logu, A.; Onnis, V.; Saddi, B.; Congiu, C.; Schivo, M.L.; Cocco, M.T. *J. Antimicrob. Chemotherap.* **2002**, *49*, 275.

[42] Kobarfard, F.; Kauffman, J.M. Univ.Sciences in Philadelphia, US6624153 B2 **2003**, 18 pp.

[43] Lee, R.E.; Protopopova, M.; Crooks, E.; Slayden, R.A.; Terrot, M.; Barry C.E.III *J. Comb. Chem.* **2003**, *5*, 172.

[44] Protopopova, M.N.; Lee, R.E.; Slayden, R.A.; Barry III, C.E.; Einck, L. National inst. Of Health, WO3096989 A2, 2003, 223 pp.

[45] Alland, D.; Steyn, A.J.; Weisbrod, T.; Aldrich, K.; Jacobs, W.R. *J. Bacteriol.* **2000**, *182*, 1802.

[46] Lounis, N.; Roscigno, G. *Curr. Pharm. Design* **2004**, *10*, 3229.

[47] Schuurmann, G.; Kühne, R.; Mekenyan, O.; Nedeyalkova, Z.A.; Andrew, P.W.; Shafi, J.; Scharpe, J.A.; Dimov, D.I.; Angelova, N.T.; Grozdanov, A.; Penev, P.N.; Haladgova, S.T.; Mincheva, D.T.; Paraskevova, S.T.; Dekin, S.I.; Atanasova, I.T. Umweltforschungs Zentrum Leipzig-Halle GMBH, WO03084965 A1, **2003**, 25 pp.

[48] Medeiros, M.; Costa, M.; Figueiredo, R.; Rosa, M.; Curto, M.; Santos, L.; Cruz, M.; Gaspar, M.; Feio, S. Instituto Nacional de Engenharia e Technologia Industrial, Portugal: WO04005298 A1, **2004**, 44 pp.

[49] Kato, H.; Yoshida, T.; Nishimoto, A.; Ohmoto, S. Hokuriku Seiaku Co. Ltd: WO0226753 A1, **2002**, 142 pp.

[50] Yoshida, T.; Takahashi, Y. Hokuriku Seiaku Co. Ltd: WO03022289 A1, **2003**, 38 pp.

[51] Broussy, S.; Coppel, Y.; Nguyen, M.; Bernadou, J.; Meunier, B. *Chem. Eur. J.* **2003**, *9*, 2034-2038.

[52] Yatvin, M.B.. Pederson, R.L. Patent US 2004224918 A1, **2004**, 24 pp.

[53] Van Gestel, J.F.E.; Venet, M.G.; Poignet, H.J.J.; Decrane, L.F.B.; Vernier, D.F.J. Janssen Pharma NV: WO04011436 A1, **2004**, 61 pp.

[54] Andries, K.; Verhasselt, P.; Guillemont, J.; Göhlmann, H. W. H.; Neefs, J.-M.; Winkler, H.; Van Gestel, J.; Timmerman, P.; Zhu, M.; Lee, E.; Williams, P.; de Chaffoy, D.; Huitric, E.; Hoffner, S.; Cambau, E.; Truffot-Pernot, C.; Lounis, N.; Jarlier, V. *Science* **2005**, *307*, 223.

[55] Clinical trials Ireland (2007; http://clinicaltrials.gov/show/NCT00449644)

[56] Hutchinson, D.K. *Expert Opin. Ther. Patents* **2004**, *14*, 1309.

[57] Mehta, A.; Rudra, S.; Raja Rao, A.V.S.; Yadav, A.S.; Rattan, A. Ranbaxy Labs Ltd: WO04014392 A1, **2004**, 62 pp.

[58] Arora, S.K.; Patil, V.J.; Nair, P.S.; Dixit, P.P.; Ajay, S.; Sinha, R.K. Lupin Ltd: WO04026848 A1, **2004**, 84 pp.

[59] Palittapongarnpim, P.; Kirdmanee, C.; Kittakoop, P. Rukseree, K. National Science and technology development, Thailand, US20020192262 A1, **2002**, 6 pp.

[60] Ascher, G.; Stauffer, F.; Berner, H.; Mang, R. SANDOZ GMBH: WO03082260 A2, **2003**, 74 pp.

[61] Deidda, D.; Lampis, G. Fioravanti, R.; Biava, M.; Porretta, G.C.; Zanetti, S.; Pompei, R. *Antimicrob. Agents Chemother.* **1998**, *42*, 3035.

[62] Arora, S.K.; Sinha, N.; Jain, S.; Upadhayaya, R.S.; Jana, G.; Ajay, S.; Sinha, R.K. LUPIN LTD.: WO04026828 A1, **2004**, 66 pp.

[63] Xu, Z.; Lin, Y.M.; Flavin, M.T. Sarawak Medichem Pharmaceuticals. US6268393, **2001**, 31 pp.

[64] Möllmann, U.; Makarov, V. Medac Gesellschaft für klinische spezialpräparate MBH. WO03042186 A1, **2003**, 39 pp.

[65] Isono, K. *J. Antibiot.* **1988**, *41*, 1711.

[66] Somu, R.V.; Boshoff, H.; Qiao, C.; Benett, E.M.; Barry III, C.E.; Aldrich, C.C. *J. Med. Chem.* **2006**, *49*, 31.

[67] Vannada, J.; Bennett, E.M.; Wilson, D.J.; Boshoff, H.I.; Barry III, C.E.; Aldrich, C.C. *Org. Lett.* **2006**, *8*, 4707.

[68] Rai, D.; Johar, M.; Manning, R.; Agrawal, B. Kunimoto, D.Y.; Kumar, R. *J. Med. Chem.* **2005**, *48*, 7012.

[69] Johar, M.; Manning, T.; Kunimoto, D.Y.; Kumar, R. *Bioorg. Med. Chem.* **2005**, *13*, 6663.

[70] Ballell, L.; Field, R.A.; Chung, G.A.C.; Young, R.J. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1736.

[71] Takeuchi, T.; Igarashi, M.; Naganawa, H.; Hamada, M. US6780616 B1 2004, 35 pp.

[72] Struve, W.G.; Sinha, R.K.; Neuhaus, F.C. *Biochemistry* **1966**, *5*, 82.

[73] Muramatsu, Y.; Muramatsu, A.; Ohnuki, T.; Ishii, M.M.; Kizuka, M.; Enokita, R.; Tsutsumi, S.; Arai, M.; Ogawa, Y.; Suzuki, T.; Takatsu, T.; Inukai, M. *J. Antibiot.* **2003**, *56*, 253.

[74] Hotoda, H.; Daigo, M.; Furukawa, M.; Murayama, K.; Hasegawa, C.A.; Kaneko, M.; Muramatsu, Y.; Ishii, M.M.; Miyakoshi, S.; Takatsu, T. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2833.

[75] Kimura, K.; Bugg, T.D.H. *Nat. Prod. Rep.* **2003**, *20*, 252.

[76] Coates, J.A.V.; Cammack, N.; Jenkinson, H.J.; Mutton, I.M.; Pearson, B.A.; Storer, R.; Cameron, J.M.; Penn, C.R. *Antimicrob. Agents Chemother.* **1992**, *36*, 202.

[77] Schinazi, R.F.; McMillan, A.; Cannon, D.; Mathis, R.; Lloyd, R.M.; Peck, A.; Sommadossi, J. P.; St. Clair, M.; Wilson, J.; Furman, P.A.; Painter, G.; Choi, W.B.; Liotta, D.C. *Antimicrob. Agents Chemother.* **1992**, *36*, 2423.

220

78 Lin, T.S.; Luo, M.-Z.; Liu, M.-C.; Zhu, Y.-L.; Gullen, E.; Dutschman, G.E.; Cheng, Y.-C. *J.Med. Chem.* **1996**, *39*, 1757.

79 Choo, H.; Chong, Y.; Choi, Y.; Mathew, J.; Schinazi, R.F.; Chu, C.K. *J. Med. Chem.* **2003**, *46*, 389.

80 Sabini, E.; Hazra, S.; Konrad, M.; Lavie, A. *J. Med. Chem.* **2007**, published online 27/05/2007

81 Girardet, J.L.; Drach, J.C.; Chamberlain, S.D.; Koslzalka, G.N.; Townsend, L.B. *Nucleosides Nucleotides* **1998**, *17*, 2389.

82 Migawa, M.T.; Girardet, J.L.; Walker II, J.W.; Koszalka, G.W.; Chamberlain, S.D.; Drach, J.C.; Townsend, L.B. *J. Med. Chem.* **1998**, *41*, 1242.

83 Schmidt, L.; Pedersen, E.B.; Nielsen, C. *Acta Chemica Scandinavica* **1994**, *48*, 215.

84 Morvan, F.; Rayner, B.; Imbach, J.-L.; Thenet, S.; Bertrand, J.-R.; Paoletti, J.; Malvy, C.; Paoletti, C. *Nucleic Acids Res.* **1987**, *15*, 3421.

85 Gagnor, C.; Rayner, B.; Leonetti, J.-P.; Imbach, J.-L.; Lebleu, B. *Nucleic Acids Res.* **1989**, *17*, 5107.

86 De Clercq, E. *J. Clin. Virol.* **2004**, *30*, 115.

87 Arner, E.S.; Eriksson, S. *Pharmacol. Ther.* **1995**, *67*, 155.

88 Sorbera, L.A.; Castaner, J.; Rabasseda, X. *Drugs Future* **2004**, *29*, 220.

89 Masters, P.A.; O'Bryan, T.A.; Zurlo, J.; Miller, D.Q.; Joshi, N. *Arch. Intern. Med.* **2003**, *163*, 402.

90 Andersen, E. 1973 Nucleoside and nucleotide kinases. In the enzymes (Boyer, P.D., ed.), 3rd edit, vol. 8, pp. 49-96, Academic Press, New York.

91 Jong, A.Y.S.; Kuo, C.L.; Campbell, J.L. *J. Biol. Chem.* **1984**, *259*, 11052.

92 Sclafani, R.A.; Fangman, W.L. *Pco. Natl. Acad. Sci.* **1984**, *81*, 5821.

93 Jong, A.Y.S.; Cambell, J.L. *J.Biol. Chem.* **1984**, *259*, 14394.

94 Munier-Lehmann, H.; Chafotte, A.; Pochet, S.; Labesse, G. *Protein Science* **2001**, *10*, 1195.

95 Li de la Sierra, I.; Munier-Lehmann, H.; Gilles, A.M.; Bârzu, O.; Delarue, M. *J. Mol. Biol.* **2001**, *311*, 87.

96 Saraste, M.; Sibbald, P.R.; Wittinghofer, A. *Trends Biochem. Sci.* **1990**, 430.

97 Ostermann, N.; Schlichting, I.; Brundiers, R.; Konrad, M.; Reinstein, J.; Veit, T.; Goody, R.S.; Lavie, A. *Structure* **2000**, *8*, 629.

98 Aparna, V.; Jeevan, J.; Ravi, M.; Desiraju, G.R.; Gopalakrishnann B. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1014.

99 Perryman, A.L.; Lin, J.-H.; McCammon, J.A. *Chem Biol Drug Des.* **2006**, *67*, 336.

100 Ghosh, A.K.; Pretzer E.; Cho, H.; Hussain, K.A.; Duzgunes, N. *Antiviral Res.* **2002,** *54,* 29.

101 Ghosh, A.K.; Sridhar, P.R.; Leshchenko, S.; Hussain, A.K.; Li, J.; Kovalevsky, A.Y.; Walters, D.E.; Wedekind, J.E.; Valerie Grum-Tokars, V.; Das, D.;Koh, Y.; Maeda, K.; Gatanaga, H.; Weber, I.T.; Mitsuya, H.; *J. Med. Chem.* **2006**, *49*, 5252.

102 Bennet, M.S.; Wien, F.; Champness, J.N.; Batuwangala, T.; Rutherford, T.; Summers, W.C.; Sun, H.; Wright, G.; Sanderson, M.R. *FEBS Lett.* **1999**, *443*, 121.

103 Haouz, A.; Vanheusden, V.; Munier-Lehmann, H.; Froeyen, M.; Herdewijn, P.; Van Calenbergh, S., Delarue, M. *J. Biol. Chem.* **2003**, *278*, 4963.

104 Fioravanti, E.; Adam, V.; Munier-Lehmann, H.; Bourgeois, D. *Biochemistry*, **2005**, *44*, 130.

105 Vanheusden, V.; Munier-Lehmann, H.; Pochet, S.; Herdewijn, P.; Van Calenbergh, S. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2695.

106 Vanheusden, V.; Munier-Lehmann, H.; Froeyen, M.; Dugué, L.; Heyerick, A.; De Keukeleire, D.; Pochet, S.; Herdewijn, P.; Van Calenbergh, S. *J. Med. Chem.* **2003**, *46*, 3811.

107 Vanheusden, V.; Munier-Lehmann, H.; Froeyen, M.; Busson, R.; Rozenski, J.; Herdewijn, P.; Van Calenbergh, S. *J. Med. Chem.* **2004**, *47*, 6187.

108 Pochet, S.; Dugué, L.; Douguet, D.; Labesse, G.; Munier-Lehmann, H. *Chembiochem.* **2002**, *3*, 108.

109 Van Rompaey, P.; Nauwelaerts, K.; Vanheusden, V.; Rozenski, J.; Munier-Lehmann, H.; Herdewijn, P.; Van Calenbergh, S. *Eur. J. Org. Chem.* **2003**, 2911.

110 Marquez, V.E. ; Ezzitouni, A. ; Russ, P. ; Siddiqui, M.A ; Ford, H.Jr.; Feldmann, R.J.; Mitsuya, H.; George, C.; Barchi, J.J.Jr. *J. Am. Chem. Soc.* **1998**, *120*, 2780.

111 Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1973**, *95*, 2333.

112 Saunders, M.; Houk, K.N.; Wu, Y.D.; Still, W.C.; Lipton, M.; Chang, G.; Guida, W.C. *J. Am. Chem. Soc.* **1990**, *112*, 1419.

113 Barbarich, T.J.; Rithner, C.D.; Miller, S.M.; Anderson, O.P.; Strauss, S.H. *J. Am. Chem. Soc.* **1999**, *121*, 4280.

114 Codington, J.F.; Doerr, I.L.; Fox, J.J. *J. Org. Chem.* **1964**, *29*, 558.

115 Damha, M.J.; Noronha, A.M.; Wilds, C.J.; Trempe, J.-F.; Denisov, A.; Pon, R.T.; Gehring, K. *Nucleosides, Nucleotides and Nucleic Acids* **2001**, *20*, 429.

116 Krol, M.C.; Huige, C.J.M; Altona, C. *J. Comput. Chem.* **1990**, *11*, 765.

117 Thibaudeau, C.; Plavec, J.; Chattopadhyaya, J. *J. Org. Chem.* **1996**, *61*, 266.

118 McGuigan, C.; Turner, S.; Mahmood, N.; Hay, A.J. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2445.

119 Blondin, C.; Serina, L.; Wiesmüller, L.; Gilles, A.M.; Bârzu, O. *Anal. Biochem.* **1994**, *220*, 219-222.

120 Donders, L.A.; De Leeuw, F.A.A.M.; Altona, C. *Magn. Res. Chem.* **1989**, *27*, 556.

121 Vorbrüggen, H.; Ruh-Pohlenz, C. Handbook of Nucleoside Synthesis, John Wiley & Sons, 2001.

122 Vorbrüggen, H.; Krolikiewicz, K.; Bennua, B. *Chem. Ber.* **1981**, *114*, 1234.

123 Parr, I.B.; Horenstein, B.A. *J. Org. Chem.* **1997**, *62*, 7489.

124 Lin, T.-S.; Zhu, J.-L.; Dutschman, G.E.; Cheng, Y.-C.; Prusoff, W.H. *J. Med. Chem.* **1993**, *36*, 353.

125 Filichev, V.V.; Brandt, M.; Pedersen, E.B. *Carbohydrate Res.* **2001**, *333*, 115.

126 Svansson, L.; Kvarnström, I.; Classon, B.; Samuelsson, B. *J. Org. Chem.* **1991**, *56*, 2993.

127 Chong, J.M.; Wong, S. *J. Org. Chem.* **1987**, *52*, 2596.

128 Faul, M.M.; Huff, B.E.; Dunlap, S.E.; Frank, S.A.; Fritz, J.E.; Kaldor, S.W.; LeTourneau, M.E.; Staszak, M.A.; Ward, J.A.; Werner, J.A.; Winneroski, L.L. *Tetrahedron* **1997**, *53*, 8085.

129 Lin, T.-S.; Mancini, W.R. *J. Med. Chem.* **1983**, *26*, 544.

130 Schreiber, S.L.; Ikemoto, N. *Tetrahadron Lett.* **1988**, *29*, 3211.

131 Greengrass, C.W.; Hoople, D.W.T.; Street, S.D.A.; Hamilton, F.; Mariott, M.S.; Bordner, J.; Dalgleish, A.G.; Mitsuya, H.; Broder, S. *J. Med. Chem.* **1989**, *32*, 618.

132 Saha, A.K.; Schairer, W.; Upson, D.A. *Tetrahedron Lett.* **1993**, *34*, 8411.

133 Ochem incorporation (2007; http://www.ocheminc.com)

134 Sanghvi, Y. S.; Bharadwaj, R.; Debart, F.; DeMesmaeker, A. *Synthesis* **1994**, 1163.

135 Parkes, K.E.B.; Taylor, K. *Tetrahedron Lett.* **1988**, *29*, 2995.

136 a) Bankston, D. D.; Almond, M. R. *J. Heterocyclic. Chem.* **1992**, *29*, 1405-1407; b) Prasad, C. V. C.; Caulfield, T. J.; Prouty, C. P.; Saha, A. K.; Schairer, W. C.; Yawman, A.; Upson, D. A.; Kruse, L. I. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 411-414.

137 Gröbel, B.-T.; Seebach, D. *Synthesis* **1977**, 357.

138 Yus, M.; Najera, C.; Foubelo, F. *Tetrahedron*, **2003**, *59*, 6147.

139 Munier-Lehmann, H. ; Pochet, S. ; Dugué, L. ; Dutruel, O. ; Labesse, G. ; Douguet, D. *Nucleosides Nucleotides* **2003**, *22*, 801-804.

140 Saenger, W. *Principles of Nucleic Acid Structure*, Springer, New York, **1984**

141 Kifli, N. ; Htar, T.T. ; De Clercq, E. ; Balzarini, J. ; Simons, C. *Bioorg. Med. Chem.* **2004**, *12*, 3247-3257.

142 Lin, T.-S.; Zhu, J.-L.; Dutschman, G. E.; Cheng, Y.-C.; Prusoff, W. H. *J. Med. Chem.* **1993**, *36*, 353-362.

143 Berndt, K.D.; Guntert, P.; Wuthrich, K. *J. Mol. Biol.* **1993**, *234*, 735-750.

144 Van Wijk, J.; Huckriede, B.D.; Ippel, J.H.; Altona, C. *Methods Enzymol.* **1992**, *211*, 286-306.

145 a) Farjon, J. ; Merlet, D. ; Lesot, P.; Courtieu, J. *J. Magn. Reson.* **2002**, *158*, 169-172; b) Fäcke, T.; Berger, S. *J. Magn. Reson.* Ser. A **1995**, *113*, 114-116.

146 Phetsuksiri, B.; Jackson, M.; Scherman, H.; McNeil, M.; Besra, G. S.; Baulard, A. R.; Slayden, R. A.; DeBarber, A. E.; Barry, C. E. III; Baird, M. S.; Crick, D. C.; Brennan, P. J. *J. Biol. Chem.* **2003**, *278*, 53123-53130.

147 Urbancik, B. *Tubercle* **1966**, *47*, 283-288.

148 Urbancik, B. *Antibiot. Chemother.* **1970**, *16*, 117-123.

149 Topliss, J.G. *J. Med. Chem.* **1972**, *15*, 1006.

150 Hansch, C.; Fujita, T. *J. Am.Chem. Soc.* **1964**, *86*, 1616.

151 Ward, D.I.; Jeffs, S.M.; Coe, P.L.; Walker, R.T. *et. Lett.* **1993**, *34*, 6779-6782.

152 Wang, J.; Choudhury, D.; Chattopadhyaya, J.; Eriksson, S. *Biochemistry* **1999**, *38*, 16993-16999.

153 Ferreras, J.; Ryu, J.-S.; Di Lello, F.; Tan, D.S.; Quadri, L.E.N. *Nat. Chem. Biol.* **2005**, *1*, 29-32.

154 Somu, R.V.; Boshoff, H.; Qiao, C.; Bennet, E.M.; Barry III, C.E.; Aldrich, C.C. *J. Med. Chem.* **2006**, *49*, 31-34.

155 Danel, K.; Larsen, E.; Pedersen, E.B. *Synthesis*. **1995**, *8*, 934-936.

156 Ciuffreda, P.; Loseto, A.; Santaniello, E. *Tetrahedron* **2002**, *58*, 5767-5771.

157 Wallace, A.C.; Laskowski, R.A.; Thornton, J.M. *Protein Eng. Des. Sel.* **1995**, *8*, 127-134.

158 McDonald, I.K.; Thornton, J.M. *J. Mol. Biol.* **1994**, *238*, 777-793.

159 Krause, S.; Willighagen, E.; Steinbeck, C. *Molecules* **2000**, *5*, 93-98.

160 Kraulis, P.J. *J. Appl. Cryst.* **1991**, *24*, 946-950.

161 Esnouf, R.M. *Acta Cryst D* **1999**, *55*, 938-940.

162 Merritt, E.A.; Bacon, D. J. *Methods Enzymol.* **1997**, *277*, 505-524.

163 Onderwater, R.C.A.; Commandeur, J.N.M.; Vermeulen, N.P.E. *Toxicology* **2004**, *197*, 80-90.

164 J. C. Palomino, A. Martin, M. Camacho, H. Guerrra, J. Swings, F. *Antimicrob. Agents Chemother.* **2002**, *46*, 2720-2722.

165 Collins, L.; Franzblau, S.G. *Antimicrob. Agents Chemother.* **1997**, *41*, 1004-1009.

166 Krause, S.; Willighagen, E.; Steinbeck, C. *Molecules* **2000**, *5*, 93-98.

167 Smith, D.H.; Gray, N.A.B.; Norse, J.G.; Crandell, C.W. *Anal. Chim. Acta* **1981**, *133*, 471-497.

168 Schmidt, M.; Baldridge, K.; Boatz, J.; Elbert, S.; Gordon, M.; Jensen, J.; Koseki, S.; Matsunaga, N.; Nguyen, K.; Su, S.; Windus, T.; Dupuis, M.; Montgomery, J. *J. Comput. Chem.* **1993**, *14*, 1347-1363.

169 Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. *Nucleic Acids Res.* **2000**, *28* 235-242.

170 Allen, F.H. *Acta Crystallogr.* **2002**, *B58*, 380-388.

171 Araki, K.; Wei Qin Yun; O'Toole, J.; Toscano, P.J.; Welch, J.T. *Carbohydr. Res.* **1993**, *249*, 139-161.

172 Rogers, J. P.; Beuscher, A.E.; Flajolet, M.; Mcavoy, T.; Nairn, A.C.; Olson, A.J.; Greengard, P. *J Med. Chem.* **2006**, *19*, 1658-1667.

173 Morris, G. M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E; Belew, R.K.; Olson, A.J. *J. Comput. Chem.* **1998**, *19*, 1639-1662.

174 Kosinska, U.; Carnrot, C.; Eriksson, S.; Wang, L.; Eklund, H. *FEBS J.* **2005**, *272*, 6365.

175 Ostermann, N.; Segura-Pena, D.; Meier, C.; Veit, T.; Monnerjahn, C.; Konrad, M.; Lavie, A. *Biochemistry* **2003**, *42*, 2568.

176 Kotaka, M.; Dhaliwal, B.; Ren, J.; Nichols, C.E.; Angell, R.; Lockyer, M.; Hawkins, A.R.; Stammers, D.K. *Protein Sci.* **2006**, *15*, 774.

177 Lavie, A.; Konrad, M. *Mini Rev. Med. Chem.* **2004**, *4*, 351.

178 Gustafson, E.A.; Chillemi, A.C.; Sage, D.R.; Fingeroth, J.D. *Antimicrob. Agents Chemother.* **1998**, *42*, 2923.

179 Suzutani, T.; Davies, L.C.; Honess, R.W. *Journal of General Virology* **1993**, *74*, 1011-1016.

180 Van Rompay, A.R.; Johansson, M.; Karlsson, A. *Pharmacol. Ther.* **2003**, *100*, 119-139.

181 Eriksson, S.; Munch-Petersen, B.; Johansson, K.; Eklund, H. *Cell Mol. Life Sci.* **2002**, *59*, 1327-1346.

182 Al-Madhoun, A.S.; Tjarks, W.; Eriksson, S. *Mini Rev. Med. Chem.* **2004**, *4*, 341-350.

183 Barroso, J.F.; Elholm, M.; Flatmark, T. *Biochemistry* **2003**, *42*, 15158-15169.

184 Johansson, K.; Ramaswamy, S.; Ljungcrantz, C.; Knecht, W.; Piskur, J.; Munch-Petersen, B.; Eklund, H. *Biochemistry* **2003**, *42*, 5706-5712.

185 Saada, A.; Shaag, A.; Mandel, H.; Nevo, Y.; Eriksson, S.; Elpeleg, O. *Nat. Genet.* **2001**, *29*, 342-344.

186 Elpeleg, O.; Mandel, H.; Saada, A. *J. Mol. Med.* **2002**, *80*, 389-396.

187 Lewis, W.; Dalakas, M.C. *Nat. Med.* **1995**, *1*, 417-422.

188 Lewis, W.; Day, B.J.; Copeland, W.C. *Nature Rev. Drug Discov.* **2003**, *1*, 812-822.

189 Balzarini, J.; Zhu, C.Y.; De Clercq, E.; Pérez-Pérez, M.J.; Chamorro, C.; Camarasa, M.J.; Karlsson, A. *Biochem. J.* **2000**, *351*, 167-171.

190 Kierdaszuk, B.; Krawiec, K.; Kazimierczuk, Z.; Jacobsson, U.; Johansson, N.G.; Munch-Petersen, B.; Eriksson, S.; Shugar, D. *Nucleos. Nucleot.* **1999**, *18*, 1883-1903.

191 Balzarini, J.; Degrève, B.; Zhu, C.Y.; Durini, E.; Porcu, L.; De Clercq, E.; Karlsson, A.; Manfredini, S. *Biochem. Pharmacol.* **2001**, *61*, 727-732.

192 Pérez-Pérez, M.J.; Hernandez, A.I.; Priego, E.M.; Rodriguez-Barrios, F.; Gago, F.;Camarasa, M.J.; Balzarini, J. *Curr. Top. Med. Chem.* **2005**, *5*, 1205-1219.

193 Hernandez, A.I.; Familiar, O.; Negri, A.; Rodriguez-Barrios, F.; Gago, F.; Karlsson, A.; Camarasa, M.J.; Balzarini, J.; Pérez-Pérez, M.J. *J. Med. Chem.* **2006**, *49*, 7766-7773.

194 Saenger, W. Principles of Nucleic Acid Structure; Springer-Verlag; Berlin, 1984.

195 Stryer, L. Biochemistry, 'th Ed.; W.H. Freeman and Company; New York, 1995.

196 http://www.answers.com/topic/dna

197 Faria, M.; Giovannangeli, C. *J. Gene. Med.* **2001**, *3*, 299-310.

198 Morrison, L. E. *J. Fluoresc.* **1999**, *9*, 187-196.

199 Kricka, L. *J. Ann. Clin. Biochem.* **2002**, *39*, 114-129.

200 Glazer, A.N.; Rye, H.S. *Nature* **1992**, *359*, 859-860.

201 Jenkins, Y.; Barton, J.K. *J. Am. Chem. Soc.* **1992**, *114*, 8736-8738.

202 Ishiguro, T.; Saitoh, J.; Yawata, H.; Otsuka, M.; Inoue, T.; Sugiura, Y. *Nucleic Acids Res.* **1996**, *24*, 4992-4997.

203 Svanvik, N.; Nygren, J.; Westman, G.; Kubista, M. *J. Am. Chem. Soc.* **2001**, *123*, 803-809.

204 French, D.J.; Archard, C.L., Brown, T.; McDowell, D.G. *Mol. Cell Probes* **2001**, *15*, 363-374.

205 Misra, A.; Mishra, S.; Misra, K. *Bioconjugate Chem.* **2004**, *15*, 638-646.

[206] Yamana, K.; Iwase, R.; Furutani, S.; Tsuchida, H.; Zako, H.; Yamaoka, T.; Murakami, A. *Nucleic Acids Res.* **1999**, *27*, 2387-2392.

[207] Yamana, K.; Zako, H.; Asazuma, K.; Iwase, R.; Nakano, H.; Murakami, A. *Angew. Chem. Int. Ed.* **2001**, *40*, 1104-1106.

[208] Manoharan, M.; Tivel, K.L.; Zhao, M.; Nafisi, K.; Netzel, T.L. *J. Phys. Chem.* **1995**, *99*, 17461-17472.

[209] Hausmann, M.; Winkler, R.; Hildenbrand, G.; Finsterle, J.; Weisel, A.; Rapp, A.; Schmitt, E.; Janz, S.; Cremer, C. *BioTechniques.* **2003**, *35*, 564-577.

[210] Radhakrishnan, I.; Patel, D. J. *Biochemistry*, **1994**, *33*, 11405-11416.

[211] Birks, J. B. Photophysics of Aromatic Molecules, Wiley-Interscience, London, **1970**.

[212] a) Kalra, N.; Parlato, M. C.; Parmar, V. S.; Wengel, J. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3166-3169; b) Kalra, N.; Babu, B. R.; Parmar, V. S.; Wengel, J. *Org. Biomol. Chem.* **2004**, *2*, 2885-2887.

[213] Hrdlicka, P. J.; Babu, B. R.; Sørensen, M. D.; Harrit, N.; Wengel, J. *J. Am. Chem. Soc.* **2005**, *127*, 13293-13299.

[214] Korshun, V. A.; Stetsenko, D. A.; Gait. M. J. *J. Chem. Soc., Perkin Trans I.* **2002**, 1092-1104.

[215] a) K. Yamana, R. L. Letsinger. *Nucl. Acids Symp. Ser.* **1985**, *16*, 169-172; b) J.S. Mann, Y. Shibata, T. Meehan. *Bioconjugate Chem.* **1992**, *3*, 554-558.

[216] Yamana, K.; Nunota, K.; Nakano, H.; Sangen, O. *Tetrahedron Lett.*, **1994**, *35*, 2555-2558.

[217] T. Bryld, T. Højland, J. Wengel. *Chem. Comm.* **2004**, 1064-1065.

[218] Filichev, V. V.; Pedersen, E. B. *J. Am. Chem. Soc.* **2005**, *127*, 14849-14858.

[219] Kashida, H.; Asanuma, H.; Komiyama, M. *Chem.Comm.* **2006**, 2768-2770

[220] Obika, S.; Hiroto, A.; Nakagawa, O.; Imanishi, T. *Chem. Comm.* **2005**, 2793-2795.

[221] Vanheusden, V.; Munier-Lehmann, H.; Froeyen, M.; Dugué, L.; Heyerick, A.; De Keukeleire, D.; Pochet, S.; Busson, R.; Herdewijn, P.; Van Calenbergh, S. *J. Med. Chem.* **2003**, *46*, 3811-3821.

[222] Berlman, I. B.Handbook of Fluorescence Spectra of Aromatic Molecules, Academic Press, New York 1965.

[223] Morris, J. V.; Mahaney, M. A.; Huber, J. R.; *J. Phys. Chem.* **1976**, *80*, 969-974.

[224] Du, H.; Fuh, R. A.; Li, J.; Corkan, A.; Lindsey, J. S. *Photochem. Photobiol.* **1998**, *68*, 141-142.

[225] Manoharan, M.; Tivel, K. L.; Zhao, M.; Nafisi, K.; Netzel, T. L. *J. Phys. Chem.* **1995**, *99*, 17461.

[226] Antony, T.; Thomas, T.; Sigal, L. H.; Shirahata, A.; Thomas, T. J. (2001) A Molecular Beacon Strategy for the Thermodynamic Characterization of Triplex DNA: Triplex Formation at the Promoter Region of Cyclin D1. *Biochemistry*, **40**, 9387-9395.

[227] Dirks, R. W.; Molenaar, C.; Tanke, H. J. *Histochem. Cell Biol.* **2001**, *115*, 3-11.

[228] Netzel, T. L.; Nafisi, K.; Zhao, M.; Lenhard, J.R.; Johnson, I. *J. Phys. Chem.* **1995**, *99*, 17936-17947.

[229] Jacobsen, J. P.; Pedersen, J. B.; Hansen, L. F.; Wemmer, D. E. *Nucleic Acids Res.* **1995**, *23*, 753.

[230] Keppler, M. D.; James, P. L.; Neidle, S.; Brown, T.; Fox, K.R. *Eur. J. Biochem.* **2003**, *270*, 4982.

[231] Filichev, V. V.; Pedersen, E. B. *J. Am. Chem. Soc.* **2005**, *127*, 14849-14858.

[232] Kumar, N.; Nielsen, K. E.; Maiti, S.; Petersen, M. *J. Am. Chem. Soc.* **2006**, *128*, 14-15.

[233] Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; Vandermarel, G. A.; Vanboom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625-5631.

[234] Manzini, G.; Xodo, L. E.; Gasparotto, D.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *J. Mol. Biol.* **1990**, *213*, 833-843.

[235] Kumar, N.; Nielsen, K. E.; Maiti, S.; Petersen, M. *J. Am. Chem. Soc.*, **2006**, *128*, 14-15.

[236] Kalyanasundaram, K.; Thomas, J. K. *J. Am. Chem. Soc.* **1977**, *99*, 2039-2044.

[237] Yamana, K.; Iwase, R.; Furutani, S.; Tsuchida, H.; Zako, H.; Yamaoka, T.;Murakami, A. *Nucleic Acid Res.* **1999**, *27*, 2387-2392.

[238] Gillies, R. J.; Raghunand, N.; Karczmar, G. S.; Bhujwalla, Z. M. *J. Magn. Reson. Imaging.* **2002**, *16*, 430-450.

[239] P.A. Giannaris, M. J. Damha. *Nucl. Acids Res*, **1993**, 4742-4749.

224