On the total synthesis of terpenes containing quaternary stereocenters
Buter, Jeffrey

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 16-03-2019
*** CHAPTER 3 ***

‡ The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis* ‡

**ABSTRACT:** The discovery of new natural products from *Mycobacterium tuberculosis* is an important field of research as the natural isolates can shine light on the bacterium’s survival and virulence mechanism, and are potentially applicable as chemical markers for the tuberculosis disease. This chapter describes the isolation and structure elucidation of an abundant and *Mycobacterium tuberculosis* specific lipid. In addition, the elucidation of the biosynthetic pathway of the natural isolate and its development into a chemical marker are presented. The studies were supported by a racemic chemical synthesis of the natural product, allowing unambiguous assignment of the molecular structure and its biosynthesis.

This chapter has been published in part:


The content of this chapter resulted from a large collaboration involving several research groups, led by the Moody laboratory (Harvard medical school). All biochemical/biological work was performed by either Moody and co-workers or collaborating groups.
3.1 Introduction

Tuberculosis (TB) remains a leading cause of death worldwide, resulting in over 1.5 million deaths annually.[1] Surprisingly, yet no rapid, sensitive, and specific diagnostic test exists. Diagnosis based on detection of *Mycobacterium tuberculosis* (*Mtbc*) in patient samples mainly relies on sputum microscopy, which is insensitive and inaccurate, or on *in vitro* culture, which is slow, insensitive and infeasible in many clinics. T-cell antigen recall tests, such as intradermal injection of purified protein derivative (the PPD-skin test, also known as the Mantoux test) or interferon-γ release assays (IGRA) are in widespread use but give delayed results or are expensive and have suboptimal test characteristics related to sensitivity and specificity.[2] Vaccination with live Bacille Calmette–Guérin (BCG) in most parts of the world leads to antigen-specific T-cell responses, which create false-positive results, rendering the PPD test unusable in many populations. Accordingly, there is now strong consensus that developing better diagnostic tests for *M. tuberculosis* infection is the key issue for improved disease control through rapid initiation of antibiotics and categorization of patients for vaccine trials.[3]

Detection of pathogen-specific shed molecules or antigens provides rapid and specific diagnosis of many infectious diseases. Such antigen tests have long been a mainstay of diagnosis for infection by *Cryptococcus, Legionella*, and other pathogens.[4] The strengths of antigen test technology are high diagnostic specificity and rapid detection of molecules using a simple ELISA of urine or serum.[5] Therefore, the key criterion for discovery of chemical targets for antigen tests is specific expression of the target by the disease-causing pathogen, combined with lack of expression among other microbes, especially those that are abundant in the environment or cause diseases that mimic the disease of interest. Other desirable criteria related to test sensitivity involve identifying targets with broad expression among most infecting strains in clinical settings, expression of the antigen at high concentrations, expression *in vivo* under conditions of infection, and lack of host degradation or metabolism to unrecognizable chemical forms.

In addition to antigen capture, pathogen-specific molecules can be coated onto plastic and used to detect target-specific host antibodies, functioning as a serological test. Such tests have not yet moved into widespread clinical use for TB due in part to specificity concerns that may be related to immune responses to environmental mycobacteria, mildly pathogenic mycobacteria, or live vaccine strains[6] against TB.[7]

Despite the widespread use of antigen and serological tests in other infectious diseases, neither type of test is widely used for TB. Although the urine lipoarabinomannan (LAM) enzyme-linked immunosorbent assay (ELISA) has usefulness for TB-HIV co-infection,[6] no antigen or serological test has emerged as having widespread clinical usefulness for tuberculosis. The current chemical targets for testing were chosen based on their ready availability and represent only a small fraction of the candidate small molecules that could be developed. For example, among 169 subclasses of mycobacterial lipids in the MycoMass and Lipid DB databases, more than 90% are
The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*

expressed only by mycobacteria.\cite{8,9} Thus, the potential range of specific mycobacterial targets for diagnostics development is vast and largely unexplored.

To broadly compare the lipid profiles of virulent and avirulent mycobacteria, we took advantage of a recently validated metabolomics platform.\cite{8} This high performance liquid chromatography–mass spectrometry (HPLC-MS) system uses methods of extraction, chromatography, and databases that are specialized for mycobacteria. After extraction of total bacterial lipids into organic solvents, HPLC-MS enables massively parallel detection of thousands of ions corresponding to diverse lipids that range from apolar polyketides to polar phosphoglycolipids. Software-based (XCMS) ion finding algorithms report reproducibly detected ions as molecular features. Each feature is a 3D data point with linked mass, retention time, and intensity values from one detected molecule or isotope. All features with equivalent mass and retention time from two bacterial lipid extracts are aligned, allowing pairwise comparisons of MS signal intensity to enumerate molecules that are overproduced in one strain with a false-positive rate below 1%.\cite{8}

This comparative lipidomics system allowed an unbiased, organism-wide analysis of lipids from *M. tuberculosis* and the attenuated vaccine strain, *Mycobacterium bovis* Bacillus Calmette–Guerín (BCG). BCG was chosen because of its worldwide use as a vaccine and its genetic similarity to *M. tuberculosis*.\cite{10} These two species are evolutionarily related and share more than 99% sequence identity but only *M. tuberculosis* causes widespread disease. We reasoned that any features that are specifically detected in *M. tuberculosis* might be clinically useful as markers to distinguish tuberculosis-causing bacteria from vaccines. Furthermore, given the differing potential for productive infection by the two strains, any *M. tuberculosis*-specific compounds would be candidate virulence factors. Comparative genomics of *M. tuberculosis* and BCG successfully identified “regions of deletion” (RD) that encode genes that were subsequently proven to promote productive *M. tuberculosis* infection,\cite{11} including the 6-kDa early secreted antigenic target (ESAT-6) secretion system-1 (ESX-1).\cite{12,13} We reasoned that a metabolite-based screen might identify new virulence factors because not all functions of RD genes are known. Also, biologically important metabolites could emerge from complex biosynthetic pathways that cannot be predicted from single-gene analysis.

In this chapter we describe our collaborative efforts with the Moody laboratory on the isolation and structure elucidation of two novel terpene nucleosides from *M. tuberculosis*. The investigation led to a biosynthetic proposal which revised the function of the virulence associated enzyme Rv3378c. One of the natural isolates was found to be an abundant and specific tuberculosis produced compound which has considerable potential as a chemical marker in a diagnostic test for tuberculosis disease. The research was facilitated by chemical synthesis of the natural products. The synthetic material was used in cross-confirmation of the molecular structure of the terpene nucleosides and the biosynthesis thereof.
CHAPTER 3

3.2 Tuberculosinyl adenosines; novel terpene nucleosides from Mycobacterium tuberculosis

3.2.1 The isolation and structure elucidation of an unknown terpene nucleoside

Using HPLC-MS for comparative analysis of lipid extracts of *M. tuberculosis* H37Rv and BCG (Pasteur strain), 7852 molecular features were detected. By aligning datasets and seeking features that significantly differed in intensity, we identified 1845 features that were only detected in one bacterium or the other. Among these features, we focused on molecules selectively detected in *M. tuberculosis* that showed the highest fold-change ratios and intensity. We identified four molecular features corresponding to a singly charged molecular ion at \( m/z \) 540.357 (\( C_{30}H_{45}N_5O_4 \)) and its isotopes, but this chemical formula did not match entries in the MycoMass\(^{[8,9]} \) or other public databases. We named the unknown molecule substance A.

The molecular ion of substance A was one of the most intense ions in the *M. tuberculosis* lipidome, suggesting that it was produced in abundance. Identification of an apparently abundant molecule in a widely studied pathogen was unexpected, leading to questions about whether substance A was truly a natural product. However, this compound was absent in media, solvent blanks, and BCG lipid extracts but was reproducibly detected in three reference strains of *M. tuberculosis* (H37Ra, Erdman and H37Rv). As observed with cell associated compounds, culture filtrate yielded a conspicuous ion at \( m/z \) 540.357 whose intensity was higher than that of the abundantly secreted siderophore, carboxymycobactin. Its release into the extracellular space likely results from transmembrane transport, rather than budding of intact cell wall fragments, because cell wall-embedded lipids, trehalose monomycolate and mycobactin, were not detected in filtered supernatants. We detected substance A in *M. tuberculosis* during exponential or stationary phase and several types of media or when subject to acid stress. Thus, substance A is a natural product that is constitutively produced under many conditions and accumulates within and outside *M. tuberculosis*.

*M. tuberculosis* often compartmentalizes lipid biosynthesis so that lipids are assembled after transport across the plasma membrane. Sulfoglycolipids and phthiocerol dimycocerosates become undetectable when MmpL transporters are interrupted, even when biosynthetic genes are intact.\(^{[14-16]} \) Because ESX-1 is a transport system lacking in BCG, lack of export of an ESX-1–dependent lipid synthase might account for the loss of substance A. However, ESX-1–deficient *M. tuberculosis* lacking either the espA gene (Rv3616c) or the entire RD1 locus,\(^{[17]} \) which are both necessary for ESX-1 function, produces substance A at normal levels. After ruling out a major known species-specific difference in transport, we devised a screen to detect genes responsible for substance A biosynthesis.

Collision-induced mass spectrometry (CID-MS) identified the structural components of substance A as adenine ([M+H]\(^+ \), \( C_5H_6N_5 \), \( m/z \) 136.0618), adenosine ([M+H]\(^+ \), \( C_{10}H_{14}N_5O_4 \), \( m/z \) 268.1040), and a polyunsaturated C20 hydrocarbon ([M+H]\(^+ \), \( C_{20}H_{33} \), \( m/z \) 273.2576). A common C20 diterpene is geranylgeraniol, and *M. tuberculosis*
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis produces two C20 lipids containing bicyclic halimane skeletons: tuberculosinol and isotuberculosinol.\textsuperscript{[18,19]} Initially, CID-MS spectra could not distinguish among these three candidate diterpenes (Figure 1), and multistage CID-MS studies isolated the diterpene unit of substance A (m/z 273.3) and yielded collision patterns that matched both tuberculosinol and geranylgeraniol (Figure 1).

\textbf{Figure 1.} Mass fragmentations of the C\textsubscript{20}H\textsubscript{33} chain and MS \textsuperscript{+1}H-NMR assignments for 1-TbAd.

The lack of sufficient data impeded our structural assignment efforts, and NMR analysis of the unknown terpene nucleoside was preferential. After tedious purification of the natural product, we carried out NMR analyses using \textsuperscript{1}H-NMR, \textsuperscript{13}C-NMR, COSY, HMQC, and NOESY spectra, which unequivocally established the structure of substance A as 1-tuberculosinyl adenosine (1-TbAd, see Figure 1). The NMR signals of the diterpene moiety matched those of tuberculosinol\textsuperscript{[18b,c,19,20]} except for the expected difference in the side chain protons and carbons. The spectral data of the adenosine and adjacent atoms correspond closely to those of 1-prenyladenosine analogs\textsuperscript{[21]} The allylic methylene group absorbs downfield as a doublet at $\delta$ 4.92 ($J = 6.6$ Hz). A NOESY cross
peak between the adenine H-2 at \( \delta \) 8.53 and the alkene hydrogen and allylic methylene and methyl groups at \( \delta \) 5.46, 4.92, and 1.89, respectively, confirm that the tuberculosinyl group is attached to the adenine at the 1-position. Thus, \textit{M. tuberculosis} produces a previously unknown type of diterpene nucleoside.

3.2.2 \textbf{A total synthesis of racemic 1-TbAd}

Concurrent with the elucidation of the chemical structure of the isolated terpene nucleoside, we initialized a racemic synthesis to cross confirm the obtained data. As our synthesis was racemic in nature we decided to reproduce in a large part the recently reported racemic total synthesis of tuberculosinol by Sorensen and co-workers (Scheme 1).

\[\text{Scheme 1. The racemic total synthesis of 1-TbAd.}\]

The synthesis started with a Diels-Alder reaction between diene 1 and ethyl tiglate. The reaction provided a diastereomeric mixture of (2:1) of \textit{exo} and \textit{endo} Diels-Alder adducts 2 respectively. Reduction of the ester with \text{LiAlH}_4 then provided alcohol 3 in 70\% yield over the two steps. Oxidation was then followed by a Wittig olefination providing us with terminal alkene 5 in 95\% over the two steps. At this stage a Suzuki-Miyaura coupling was
The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*

performed,[24] however this reaction was somewhat troublesome in our hands as we were unable to perfectly reproduce the literature procedure. Whereas in the literature unprotected vinyl iodide \(6\) – TBS group was used to directly construct tuberculosinol, we only managed to achieve coupling with the protected vinyl iodide \(6\). After careful experimentation we managed to obtain TBS protected tuberculosinol \(7\), which was deprotected by the aid of TBAF, yielding tuberculosinol in \(67\%\) over the two steps. The spectral data of the obtained tuberculosinol were in agreement with the literature and the compound was subsequently used in the total synthesis of 1-TbAd. For that we first had to convert the tuberculosinol into its electrophilic counterpart for coupling with the nucleophilic adenosine. At first we performed a Corey-Kim chlorination as reported by the Poulter laboratory.[25] The chlorination could be performed rather smoothly, but we found the method showed some inconsistencies regarding the reproducibility. The chloride could be obtained but the yields were varying from \(65\%\) to \(96\%\) yield. As an alternative the chlorination was performed with mesyl chloride providing tuberculosinyl chloride \(8\), consistently, averaging to \(94\%\) yield. The chloride was then coupled with adenosine[26] under Finkelstein conditions, heated by microwave. Preparative HPLC-MS analysis (performed at Harvard medical school) of the reaction mixture afforded 1-TbAd in only \(17\%\) yield as a single diastereoisomer. To our surprise, a significant contaminant was found in the reaction mixture which corresponded to the tuberculosinyl triethylammonium ion. This side product was apparently formed in the chlorination reaction (method B) in which triethylamine substituted the chloride in \(8\). Because of inattentiveness during analysis of the \(^1\)H-NMR, the characteristic ethyl resonances were mistaken for residual diethyl ether. The low yield in the adenosine coupling and the aforementioned contamination problems were addressed and solved in later synthetic efforts (see chapter 4). The synthetic 1-TbAd cross-confirmed the chemical structure and provided the first quantities of 1-TbAd for further investigations.

### 3.2.3 The biosynthetic pathway of 1-TbAd

To identify the genes necessary for 1-TbAd production, an existing library of random transposon insertional mutants[27] was screened in high throughput (4196 mutants) for 1-TbAd production using a simplified 3-min HPLC-MS method. Thirty mutants showing low or absent signals were rescreened using the original, high-resolution lipidomic separation method. Reporting only mutants with complete signal loss of 1-TbAd signal in both assays, we identified two 1-TbAd–null mutants carrying transposons in Rv1796 (mutant 1) and Rv2867c (mutant 2). The concurrently performed structure elucidation described above identified the highly characteristic tuberculosinyl moiety as the terpene component of 1-TbAd, and the Rv3377c-Rv3378c locus was known to encode enzymes needed for tuberculosinol and isotuberculosinol production.[18,19,20,28] Sequencing identified spontaneous mutations in Rv3378c in both mutants.[18,19,20] Mutant 1 encoded a predicted Asp → Gly substitution at residue 34, and mutant 2 encoded a Pro → Ser substitution at residue 231. We generated complementation constructs to separately test whether the point mutations in Rv3378c or the transposon insertions were responsible
for 1-TbAd loss. Transfer of Rv1796 and Rv2867c failed to restore 1-TbAd production, but transfer of Rv3377c-Rv3378c reconstituted 1-TbAd production in both mutants. Thus, Rv3377c-Rv3378c genes are necessary for 1-TbAd biosynthesis in *M. tuberculosis*.

Furthermore, the known role of Rv3377c and Rv3378c enzymes in tuberculosinol production potentially provided a mechanism to connect Rv3377c and Rv3378c genes with the production of a nucleoside-modified tuberculosinol. Rv3377c is a terpene cyclase, which acts on geranylgeranyl pyrophosphate (GGPP) to generate tuberculosinyl pyrophosphate (TbPP).\(^{18}\) Rv3378c was thought to be a phosphatase, which converts TbPP to free tuberculosinol.\(^{19,20}\) Extending current models, 1-TbAd might result from downstream action of an unknown enzyme on free tuberculosinol to transfer it to adenosine. Polyprenol synthase genes and the Rv3377c-Rv3378c locus are coordinately regulated and encoded at adjacent sites on the chromosome.\(^{29}\)

Therefore, we searched *M. tuberculosis* databases for genes located near this locus that might plausibly function as adenosine transferases. We failed to find candidates and noted that no transposon insertion that blocked 1-TbAd production mapped to genes adjacent to this locus. Therefore, we considered a revised biosynthetic model in which Rv3378c protein is not a simple phosphatase, as currently believed, but instead acts with combined phosphatase and tuberculosinyl transferase functions, using adenosine as the nucleophilic substrate (Scheme 2). This model is mechanistically simple and might explain the lack of an apparent stand-alone transferase gene. Also, whereas current models predict that tuberculosinol is the end product of this pathway, we did not detect tuberculosinol in lipidomics experiments. The revised model posits that 1-TbAd is the end product of Rv3378c pathway, explaining why it accumulates to high levels as one of the brightest ions in the lipidome. After chemical synthesis of TbPP, by reaction of tuberculosinyl chloride with tris(tetra-n-butylammonium) hydrogen pyrophosphate, we tested TbPP and GGPP as substrates for the recombinant Rv3378c protein.\(^{18}\) Rv3378c catalyzed the condensation of adenosine and TbPP to generate 1-TbAd but produced no detectable product from GGPP. Free tuberculosinol was not detected in these assays. Thus, Rv3378c shows tuberculosinyl transferase activity, which rules in the revised biosynthetic pathway.

To test the sufficiency of this locus for 1-TbAd production in cells, we transferred the Rv3377c-Rv3378c locus to *Mycobacterium smegmatis*. In all three clones tested, expression of Rv3377c-Rv3378c transferred production of a molecule with the mass, retention time, and CID-MS spectrum of 1-TbAd. Thus, no other *M. tuberculosis*-specific co-factor or transporter is needed for 1-TbAd production. Rv3377c-Rv3378c is sufficient to synthesize 1-TbAd from ubiquitous cellular precursors present in most bacteria, likely GGPP and adenosine. This finding is particularly interesting and important as the gene cluster Rv3377c-Rv3378c have been shown to be crucial in the virulence and survival mechanism of *Mycobacterium tuberculosis*. Genetic deletion of the enzyme necessary for 1-TbAd biosynthesis (Rv3378c) rendered *M. tuberculosis* unable to grow in macrophages *in vivo*, suggesting a non-redundant role of this biosynthetic pathway in virulence.\(^{30}\)
The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*

Scheme 2. Extension of the biosynthetic function of the Rv3378c protein.

3.3 The development of 1-TbAd as a chemical marker for tuberculosis disease

3.3.1 1-TbAd is a major and specific lipid in *Mycobacterium tuberculosis*

Our described isolation efforts of 1-TbAd prompted us with the idea to develop 1-TbAd into a chemical marker for tuberculosis disease. In order to function as a successful marker a series of requirements have to be met, starting with abundancy and specificity. The molecule used as a marker has to be abundant and easily detectable to exclude false negative results. Additionally the molecule has been specific to the pathogen so that the chances of false positive results are minimized. To see whether 1-TbAd fulfills these initial requirements we set out to investigate these criteria.

As shown, 1-TbAd was identified in *M. tuberculosis* using an (ESI)-MS-based lipidomics platform. Among 7852 ions, 1-TbAd (C_{30}H_{46}N_{5}O_{4}^{+}; m/z 540.3545) was the second most intense ion in the lipidome. The high intensity might have resulted from 1-TbAd’s accumulation to high concentration. Yet it was difficult to imagine that any abundant class of molecule would have escaped detection over decades of study of a pathogen of worldwide importance. Alternatively, 1-TbAd might have been present only in trace amounts, yet its intrinsically charged nature and amphipathic character might have promoted particularly efficient ionization in ESI-MS. To measure the mass of 1-TbAd as a percentage of all lipids, we analyzed extracts from *M. tuberculosis* strain H37Rv using the method of standard additions. We compared the area under the curve of retention time versus intensity measured at the mass (m/z 540) of 1-TbAd (A_{540}) and...
phosphatidylethanolamine (A720), which controls for the efficiency of ESI-MS detection. As with other polar lipids, we observed a nearly linear relationship between mass input and 1-TbAd signal intensity. In three experiments, we determined that 1-TbAd comprises 1.1%, 1.4%, and 1.5% of the mycobacterial lipids. Thus, contrary to expectations that terpene nucleosides previously escaped detection due to scarcity, 1-TbAd is a highly abundant molecule that comprises a major class of lipid in *M. tuberculosis*. The basis for lack of prior detection remains unknown, but the near co-elution of 1-TbAd and abundant membrane phospholipids in normal-phase thin-layer chromatography and HPLC methods might have obscured 1-TbAd detection in the past.

To function as a sensitive marker of infection, molecules must be expressed in most strains of *M. tuberculosis* that infect patients and among the seven recognized *M. tuberculosis* lineages that exist worldwide. 1-TbAd derives from geranylgeranyl pyrophosphate (GGPP) and adenosine, which are present in nearly all organisms. Lipid cyclization and coupling to adenosine are performed by specialized enzymes, Rv3377c (prenyl cyclase) and Rv3378c (tuberculosinyl transferase), which together with polyprenyl synthases comprise a functional gene island that is necessary and sufficient for 1-TbAd biosynthesis. Knowledge of the essential TbAd biosynthetic genes allowed a survey of the genomes of 432 phylogeographically diverse clinical isolates of the *M. tuberculosis* complex (*Mtb* complex) for an intact biosynthetic locus. We found that TbAd genes were detected among all clinical strains examined and most clinical isolates have an intact Rv3377c-3378c locus and no clear evidence for selection driving the mutations identified.

To determine whether 1-TbAd biosynthesis exists in species other than *M. tuberculosis*, we sought genetic orthologues of Rv3377c and Rv3378c across the biological kingdoms. Considering all non-mycobacterial species, we did not identify candidate orthologues organized into a locus with both essential genes present. Among mycobacteria, orthologues of Rv3377c and Rv3378c could not be identified in most species, including disease-causing members of the *M. avium* complex, *M. kansasii*, and *M. marinum*. Instead, orthologues were present only within the *Mtb* complex: in *M. bovis*, *M. bovis* BCG (Pasteur strain), *M. cannetti*, and *M. africanum*. Despite our intensive investigation, gaps in the analysis remained. Some but not all of the Rv3378c mutations, including those present in *M. africanum*, have been proven to inactivate lipid biosynthesis. Second, BCG vaccine strains are of particular interest because they cause false-positive immunological tests. Although the locus in BCG strain Pasteur was inactivated through frameshift, other BCG or *M. bovis* strains might never have acquired these mutations, or they might have done so and then reverted to wild-type sequences. The 12 widely used BCG vaccine strains show variations in efficacy, so differential expression of TbAd might account for this. However, further analysis of all common vaccine strains used worldwide (Pasteur, Copenhagen, Japan, Mexican, Australian, Russia, Glaxo, Prague, Phipps, Connaught, Denmark, Tice) confirmed the presence of the frameshift mutation in all cases, so 1-TbAd is likely absent from all major vaccines used worldwide.
A general limitation of genetic analysis is that genes might exist that are functionally equivalent to Rv3377c and Rv3378c but lack the sequence identity needed for identification as orthologues. Therefore, we undertook biochemical analysis, using the HPLC lipidomics platform, for 1-TbAd production, focusing on non-tuberculous mycobacteria and microbes, whose infection can mimic tuberculosis disease.

In agreement with the genetic results, we did not detect 1-TbAd among disease-causing bacteria that are related to *M. tuberculosis* but lack identifiable orthologues of Rv3377c-3378c (*M. avium*, *M. marinum*) or those with orthologous loci containing a known frameshift mutation (*M. bovis*). Among all organisms tested to date, only *M. tuberculosis* produces 1-TbAd.

### 3.3.2 In vivo detection of 1-TbAd in infected mice

With the biosynthetic and genetic evidence gathered, confirming that 1-TbAd is *M. tuberculosis* specific and also abundant, an initial platform for 1-TbAd detection in context of a chemical marker could be developed. To determine if 1-TbAd is detectable in vivo, we infected BALB/c mice via inhalation for 21 days followed by collection of whole lung homogenates. Target validation in vivo does not require scanning of the complete lipidome, but instead focuses on sensitive and specific detection of predetermined targets, which are likely to be highly diluted within host tissues. We developed a reversed-phase method to increase chromatographic resolution, while taking advantage of the specificity and accuracy of triple quadrupole mass detection. This method detected a 1-TbAd standard at 11.8 min based on its known transition for 1-TbAd (*m/z* 540/408/136). Direct analysis of unfractionated lipids from whole-lung homogenates did not detect signal corresponding to 1-TbAd in uninfected lung but did detect a signal corresponding to 1-TbAd that was well above background signals in six of six infected mice tested, with representative data shown in scheme 3. Thus, 1-TbAd is produced in vivo and is readily detected ex vivo in a one-step HPLC-MS method.
3.3.3 Identification of another unknown terpene nucleoside

Ion chromatograms of the ion 540→408 also showed a second, later-eluting peak (~14.6 min) with transitions that were distinct from 1-TbAd in all six mice tested (Scheme 3). Unlike 1-TbAd, the unknown lipid contained a 540→148 ion and a relatively bright 540→408 transition that was equivalent in intensity to the 540→136 ion. Returning to normal-phase HPLC-quadrupole time-of-flight (Q-TOF)-MS experiments using *in vitro* grown *M. tuberculosis*, ion chromatograms measured at *m/z* 540.4 also showed two peaks at ~25 and ~7 min, which corresponded to late-eluting 1-TbAd and an early-eluting unknown lipid detected in mice. Collision-induced dissociation (CID)-MS analysis of the unknown showed fragment ions of *m/z* 148.062 and 408.313, matching the pattern of the late-eluting unknown from reversed-phase chromatography, suggesting that they were the same molecule. Thus, the unknown later-eluting compound detected in mice was likely derived from *M. tuberculosis* rather than the host. Also, separate tracking of the unknown and 1-TbAd in the bacterial pellet and conditioned supernatant demonstrated a much higher ratio of signal for the unknown in supernatants. Thus, the early-eluting unknown was likely generated during or after transit from the cytosol to the extracellular space.
The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*

![1-Tuberculosinyl adenosine (1-TbAd)](image)

**Figure 2. Differences in 1-TbAd and its unknown isomer.**

The unknown (m/z 540.3549, measured) and 1-TbAd (m/z 540.3545, calculated) had identical molecular ion masses (±2 ppm) and similar fragmentation patterns (Figure 2). The unknown had a much lower retention time in normal-phase chromatography. Therefore, the unknown was likely a much less polar isomer of 1-TbAd. CID-MS ions were assigned unequivocally as $C_{5}H_{6}N_{5}^{+}$ (m/z 136.0618) and $C_{10}H_{14}N_{5}O_{4}^{+}$ (m/z 268.1040), matching the mass of adenine and adenosine, respectively. The neutral loss leading to m/z 268.1040 suggested the loss of $C_{20}H_{32}$, likely a diterpene moiety. A fragment ion with the formula $C_{25}H_{38}N_{5}$ (calculated m/z 408.3122) matches the loss of a pentose-derived fragment ($C_{5}H_{8}O_{4}$, calculated 132.0423 Da). Therefore, the unknown was a diterpene-substituted adenosine ($C_{30}H_{45}N_{5}O_{4}$). Experiments were guided by two hypotheses. The unknown isomer and 1-TbAd might differ in the type of diterpene carried, or a tuberculosinyl group could have an alternate linkage to adenine. *M. tuberculosis* produces geranylgeranyl pyrophosphate and tuberculosinyl pyrophosphate is hydrolyzed, besides tuberculosinol, to isotuberculosinol in vitro\[^{19}\]. Therefore, 1-geranylgeranyl adenosine and 1-isoTbAd were candidate structures. However, they would be expected to have largely the same ionic properties as 1-TbAd. Alternatively, a tuberculosinyl linkage at the $N^{6}$-adenosine position would mimic common $N^{6}$-linked adenine compounds like zeatin\[^{36}\] and would be expected to alter the ionic properties observed in the unknown. Therefore, we considered $N^{6}$-tuberculosinyl adenosine as a candidate structure. Although during the structure elucidation of 1-TbAd we ruled out
the geranylgeranyl side chain as the $C_{20}H_{32}$ unit, primarily on the basis of NMR data, we now were in search for more definite evidence. As the unknown 1-TbAd isomer could not be isolated in sufficient quantities for NMR analysis we were dependent on MS-analysis alone to solve the structure. We therefore set out to synthesize $N^6$-geranylgeranyl adenosine (Scheme 4).

Commercially available farnesyl acetone 9 was reacted with acetonitrile under alkaline conditions to produce an isomeric mixture of $\alpha,\beta$-unsaturated nitrile 10 in 50% yield. The nitrile was reduced to geranylgeraniol 12 in two steps. A Mitsunobu reaction was performed to construct phthalimide 13 which in a Gabriel reaction was converted into geranylgeranyl amine 14. To complete the synthesis, geranylgeranyl amine 14 was coupled with inosine, with BOP as the coupling reagent, to provide $N^6$-geranylgeranyl adenosine in 72% yield. HPLC-MS analysis of $N^6$-geranylgeranyl adenosine showed divergent data compared to that of the natural isolate. The significant differences in the fragmentation pattern and signal intensities, together with an inconsonant retention time led us to conclude that the isomeric terpene nucleoside did not correspond to $N^6$-geranylgeranyl adenosine.

We also proposed $N^6$-tuberculosinyl adenosine to be the structure of the unknown $N^6$-terpene nucleoside. $N^6$-TbAd was assumed easily accessible via a Dimroth rearrangement of 1-TbAd. That process is facilitated under nucleophilic conditions with as a driving force the formation of neutral $N^6$-TbAd from positively charged 1-TbAd (Scheme 5). Subjecting 1-TbAd to a nucleophile like water, as is likely for in vivo
conversion of 1-TbAd to $N^6$-TbAd, induces ring-opening of the aminopyrimidine. Bond rotation places the amine functionality in proximity to the electrophilic carbon inducing a ring-closure. Water is then eliminated to give after tautomerization $N^6$-TbAd. It should be pointed out that, as 1-TbAd has a considerably higher pKa than $N^6$-TbAd, the Dimroth rearrangement is accompanied by proton loss and therefore formally not an isomerization.

Scheme 5. The Dimroth rearrangement mechanism in the synthesis of $N^6$-TbAd from 1-TbAd.

With the unknown structure now postulated to correspond to $N^6$-tuberculosinyl adenosine we decided to synthesize this compound to provide clarification (Scheme 6). In our first approach, $N^6$-TbAd was accessed via the inosine coupling reaction. Since at this stage of the research we only had tuberculosinyl chloride, the phthalimide was accessed by a substitution reaction with potassium phthalimide. Cleavage of the phthalimide with hydrazine provided tuberculosinyl amine $16$ in 65% yield over the two steps. Coupling of $16$ with BOP provided the product with good conversion however we failed to separate HMPA from the product. HMPA is formed from BOP as the coupling agent to activate the inosine. Repetitive flash chromatography produced a meager 7% yield of $N^6$-TbAd containing ~15% HMPA as an impurity.

With little 1-TbAd left from the racemic synthesis (Scheme 6) the Dimroth rearrangement was performed. Treatment of 1-TbAd with diethylamine in MeOH at room temperature smoothly provided $N^6$-TbAd in quantitative yield. Mass analysis and the retention time of the synthetic material matched perfectly with the natural isolate, thereby providing overwhelming evidence that the structure of the terpene nucleoside corresponds to $N^6$-TbAd.
A definite confirmation of the $N^6$-TbAd structure came about by isolation from the natural source. After growing *M. tuberculosis* cultures in roller bottles, ~50 mg of lipid was fractionated by normal-phase column chromatography followed by reversed-phase HPLC. Adding 0.2% trifluoroacetic acid (TFA) to the mobile phase increased the retention of 1-TbAd in reversed-phase chromatography but did not affect the unknown isomer, confirming that the two compounds differed in their ionic properties. This purification sequence produced ~300 μg of each isomer so that the combined yield (~1.2%) confirmed ESI-MS studies.

Analysis of the $^1$H-NMR, $^{13}$C-NMR, COSY, NOESY and HMQC spectra (800 MHz) established the structure of the unknown as $N^6$-TbAd. A geranylgeranyl side chain could not produce signals as low as the observed signals at 1 ppm. The single vinylic proton would be replaced by the two protons in the isotuberculosinyl group, also ruling out this possibility. The spectral data, especially the methyl resonances, for $N^6$-TbAd corresponded near perfectly to that of its isomer 1-TbAd. Additionally, the side chain protons correspond closely to those of $N^6$-(3-methyl-2-butenyl)adenosine.\(^{21}\) The adenine protons at $\delta$ 8.24 and 8.26 are characteristic of an $N^6$-substituted adenosine and are different from those of 1-TbAd at $\delta$ 8.53 and 8.66. The absorption of the allylic...
The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*

methylene group is a broad peak at 4.22–4.16 ppm due to slow rotation at room temperature around the C-N bond. All COSY and NOESY correlations are consistent with the assignment of the unknown as N^{6}-TbAd.

### 3.3.4 The biological origin of N^{6}-TbAd

A plausible mechanism to account for the biological origin of N^{6}-TbAd is that 1-TbAd converts to N^{6}-TbAd by a Dimroth rearrangement, as explained in scheme 5. Products from recombinant Rv3378c enzyme showed a strong signal for 1-TbAd and no signal for N^{6}-TbAd, demonstrating that the latter is not a primary product of this enzyme. Furthermore, we observed that treating 1-TbAd under conditions known to favor the Dimroth reaction (Me₂NH in water) led to clean conversion into N^{6}-TbAd (Scheme 6). This data suggest a two-step model of biosynthesis whereby Rv3378c mediates conjugation of tuberculosiny1 pyrophosphate and adenosine to produce 1-TbAd in the cytosol and 1-TbAd could later rearrange to N^{6}-TbAd in other compartments. However, *in vivo* transformation of 1-TbAd to N^{6}-TbAd in mice could not be unequivocally established because mouse lung analysis required harvest, homogenization, and HPLC-MS analysis at the bench. The *ex vivo* workup of lung might have inadvertently induced Dimroth rearrangement but this possibility was ruled out by workup of lung spiked with 1-TbAd, which detected only 1-TbAd. We therefore postulate, Rv3378c produces 1-TbAd, which is transformed to N^{6}-TbAd by the Dimroth rearrangement, and both compounds occur *in vivo* and represent specific markers of *M. tuberculosis* infection.

### 3.4 Conclusions

In summary we reported the identification of two novel terpene nucleosides from *Mycobacterium tuberculosis*, namely 1-TbAd and N^{6}-TbAd. The former compound was shown to be an abundant and *M. tuberculosis* specific lipid which rendered it suitable for the development of a diagnostic test. It was shown that 1-TbAd is produced *in vivo* in tuberculosis infected mice and can be readily detected *ex vivo* using a RP-HPLC-MS method. Concurrent with the detection of 1-TbAd, its pseudo-isomer N^{6}-TbAd was discovered which was postulated to arise from an *in vivo*, non-enzymatic, Dimroth rearrangement of 1-TbAd. This novel *in vivo* produced product was present in all infected mice and therefore can function, next to 1-TbAd, as a specific marker for the tuberculosis disease.

Another very interesting, and potentially groundbreaking, discovery was the fact that 1-TbAd, an abundant and specific *M. tuberculosis* natural product, is produced by the virulence associated enzyme Rv3378c. Although the function of 1-TbAd has not been shown, we do however suspect 1-TbAd to be pivotal in the virulence and survival mechanism of tuberculosis (*vide infra for a postulated mechanism of action*).

We also communicated the first chemical, racemic, total synthesis of 1-TbAd and N^{6}-TbAd. It was shown that synthetic chemistry and biology perfectly complemented each
CHAPTER 3

other, leading to a successful research project on tuberculosinyl adenosines, research which is still ongoing (see also chapter 4). The total synthesis of the TbAd compounds and tuberculosinyl pyrophosphate proved to be important as the chemical structures and biosynthesis could be unequivocally assigned.

3.5 Discussion: a hypothesis for the working mechanism of 1-TbAd

In this chapter we showed the identification of 1-TbAd and its potential application as an abundant and specific chemical marker for tuberculosis disease. Also its biosynthesis, which requires the virulent associated enzyme Rv3378c,[30] was elucidated. The abundance, specificity and production by Rv3378c makes 1-TbAd highly likely to be involved in virulence[40] and survival[41] of Mycobacterium tuberculosis. The key question however is, what role does 1-TbAd play and what is the mechanism of action?

As discussed in the first chapter of this dissertation, Mycobacterium tuberculosis is difficult to treat with antibiotics. One of the factors that contributes to this is the thick hydrophobic cell envelope (cell wall). Consisting of a complex array of (glyco)lipids, polysaccharides and peptidoglycans, the cell wall exhibits low permeability by drugs, forming an almost impenetrable fortress.[42] Another major contributor to virulence[40] and survival[41] of Mtb is its ability to successfully infect endosomal phagocytes. Residing in phagosomes, M. tuberculosis actively inhibits pH dependent killing mechanisms, protecting itself from both drugs and immune responses.[43] In order to understand the importance of this mechanism a simplistic and partial introduction into human immunology is given (Figure 3).

![Figure 3](image-url)  
Figure 3. A schematic representation of phagocytosis. © 2009 from of The Immune System, Third Edition by Parham. Reproduced and modified by permission of Garland Science/Taylor & Francis Group LLC.

When a bacterial pathogen enters the human host, an immune response is set in motion which recognizes the invader as foreign (step A in Figure 3). The recognized pathogen activates a macrophage (a type of white blood cell, step B) which starts to encapsulate the pathogen in a process called phagocytosis (step C). The compartmentalized pathogen now resides in a vesicle called an early phagosome (step D). In order to devour the pathogen, the phagosome has to mature from an early to a late phagosome which is governed by a complex cascade of events.[44] One crucial step however is
The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*

Acidification of the phagosomal space to set up the phagosome for fusion with the lysosome. The pH in an early stage phagosome is ~6.2 which gradually decreases to pH <5.0 in late stage phagosomes. The late stage phagosome then fuses with the lysosome, in which the lysosomal contents (acid lipases and esterases) are released (step E). The lipases and esterases are capable of breaking down the pathogen, thereby stopping infection of the human host.

Despite the intricacies of the human immune system, *Mycobacterium tuberculosis* evolved an ingenious mechanism to evade the immune response guaranteeing its survival in the human host. Upon encapsulation of the bacterium, the phagosome has to mature, and it is exactly this process which is blocked by *M. tuberculosis* leading to phagosome maturation arrest. In particular, the gradual acidification of the phagosomal space is inhibited, a process which involves the Rv3377c-Rv3378c locus.\(^{30}\) The early phagosome therefore does not mature to form an acidified late stage phagosome which can fuse with the lysosome. The fact that phagolysosome formation does not take place results in survival of the bacterium inside the human host. Although the bacteria are isolated, from the perspective of the pathogen it is more protected from drug penetration than only by its thick cell envelope. Additionally, the bacteria are shielded from further immune responses allowing it to await a weakening of the human immune system whereupon it will start multiplying. The involvement of the 1-TbAd producing Rv3377c-Rv3378c locus in phagosome maturation arrest raised our interest in the role 1-TbAd plays in this process. An obvious starting point to answer this question is the molecular structure of 1-TbAd, which was therefore examined.

Our immediate attention went out to the positively charged exocyclic NH\(_2\) and the implication it has on 1-TbAd’s acidic properties. The pKa of 1-TbAd’s exocyclic NH\(_2\) has not been established experimentally but from the literature it was known that the measured pKa of 1-dimethyl allyl adenosine in water is 8.47.\(^{45}\) Although we are well aware that 1-dimethyl allyl adenosine is structurally different than 1-TbAd and that the pKa was measured under non-physiological conditions, we assume the pKa of 1-TbAd is close to this value.

From this analysis we could conclude 1-TbAd is a weak acid and is therefore subjected to an acid-base equilibrium. Together with the knowledge that phagosome maturation requires phagosome acidification,\(^{44}\) and that *Mycobacterium tuberculosis* is effectively blocking this process,\(^{43}\) we could come up with testable hypothesis on the mode of action of 1-TbAd. The acidic properties of 1-TbAd led us to propose a mechanism in which the conjugate base of 1-TbAd blocks phagosome acidification (Figure 4).
Figure 4. A schematic representation of the 1-TbAd buffering hypothesis. a) Mtb in the phagosome. b) “Basic 1-TbAd” is passing the Mtb membrane, into the acidic phagosome. c) Buffering of the acidic phagosome by 1-TbAd’s free base. Maturation arrest follows, and no fusion with lysosomes (= pathogen survival).

Assuming that the pKa of 1-TbAd under physiological conditions is ~8.5 the Henderson-Hasselbach equation (\( pH = pK_a + \log_{10}[A]/[HA] \)) is used to calculate the ratio of 1-TbAd/1-TbAd free base. The pH of the \( M. tuberculosis \) cytosol, where 1-TbAd is produced, is 7.4 which gives an acid/base ratio of ~12:1 (step A in Figure 4). We hypothesize that 1-TbAd is too polar to cross the bacterial cell wall and that only the free base will do so (step B). Upon release in the phagosome 1-TbAd accepts a proton, raising the pH (step C) and thereby avoiding fusion between the phagosome and lysosome.

As discussed previously, phagolysosome formation is a crucial step in the immune response as it is responsible for breaking down the pathogen, stopping the infection of the host. The postulated hypothesis of 1-TbAd’s involvement in phagosome maturation arrest can therefore be an important step in unraveling this moderately understood mechanism.
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis

It goes without saying that the hypothesis itself raises some interesting questions. How is 1-TbAd free base passing the thick, hydrophobic, mycobacterial cell wall, is this by transport or passively? And what is the fate of the 1-TbAd formed in the phagosome? The second question is whether, and at which rate, 1-TbAd will undergo Dimroth rearrangement to produce N6-TbAd. We have shown the in vivo production of this compound in mice and as discussed in scheme 3 the Dimroth rearrangement produces a proton. Since the rearrangement is thermodynamically down-hill one can argue that over time 1-TbAd acts as an acid (the pKa of N6-TbAd is estimated to be 3.8). In other words, the buffering of the phagosome by 1-TbAd is probably only temporarily. Although this is potentially a weakness of the hypothesis we like to remind the reader that the virulence and survival mechanism of Mycobacterium tuberculosis is complex and does not only constitute one mechanism. It might therefore be that 1-TbAd buffers the phagosome, delaying phagolysosome formation, which is then by other factors prohibited indefinitely, making the Dimroth rearrangement not a concern.

3.6 Experimental section

General remarks:
All reactions were performed using oven-dried glassware under an atmosphere of nitrogen (unless otherwise specified) by standard Schlenk techniques, using dry solvents. Reaction temperature refers to the temperature of the oil bath/cooling bath. Solvents were taken from an MBraun solvent purification system (SPS-800). All other reagents were purchased from Sigma-Aldrich, Acros, TCI Europe or Fluorochem and used without further purification unless noted otherwise. TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using either Seebach’s stain (a mixture of phosphomolybdic acid (25 g), cerium (IV) sulfate (7.5 g), H2O (500 mL) and H2SO4 (25 mL)), a KMnO4 stain (K2CO3 (40 g), KMnO4 (6 g), H2O (600 mL) and 10% NaOH (5 mL)), or elemental iodine. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230 – 400 mesh) as obtained from Screening Devices. 1H- and 13C-NMR spectra were recorded on a Varian AMX400 or a Varian 400-MR (400 and 100.59 MHz, respectively) using CDCl3 or DMSO-d6 as solvent, unless stated otherwise. Chemical shift values are reported in ppm with the solvent resonance as the internal standard (CDCl3: δ 7.26 for 1H, δ 77.16 for 13C, MeOH-d4 δ 3.31 for 1H, δ 49.0 for 13C). Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, ddd = double double doublet, ddp = double double pentet, td = triple doublet, t = triplet, q = quartet, b = broad, m = multiplet), coupling constants J (Hz), and integration. GC-MS measurements were performed with an HP 6890 series gas chromatography system equipped with a HP 5973 mass sensitive detector. GC measurements were made...
CHAPTER 3

using a Shimadzu GC 2014 gas chromatograph system bearing an AT5 column (Grace Alltech) and FID detection. Enantiomeric excesses were determined by chiral HPLC analysis using a Shimadzu LC-10ADVP HPLC instrument equipped with a Shimadzu SPD-M10AVP diode-array detector. Integration at three different wavelengths (254, 225, 190 nm) was performed and the reported enantiomeric excess is an average of the three integrations. Retention times (tR) are given in min. High resolution mass spectra (HRMS) were recorded on a Thermo Scientific LTQ Orbitrap XL.
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis

NMR Analysis of 1-TbAd isolated from Mycobacterium tuberculosis

![Diagram of 1-Tuberculosinyl adenosine (1-TbAd)]

<table>
<thead>
<tr>
<th>Atom</th>
<th>Carbon</th>
<th>Hydrogen</th>
<th>COSY</th>
<th>NOESY</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>147.2</td>
<td>8.53</td>
<td>1&quot;, 15&quot;, 16&quot;</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>144.2</td>
<td>8.66</td>
<td>1&quot;, 2&quot;, 3&quot;</td>
<td></td>
</tr>
<tr>
<td>3'</td>
<td>90.2</td>
<td>6.10 (d, 5.2)</td>
<td>2&quot;, 8&quot;, 2&quot;</td>
<td></td>
</tr>
<tr>
<td>2'</td>
<td>76.4</td>
<td>4.61 (dd, 5, 2, 5.1)</td>
<td>1&quot;, 3&quot;</td>
<td></td>
</tr>
<tr>
<td>3'</td>
<td>71.8</td>
<td>4.34 (dd, 5.1, 4.1)</td>
<td>2&quot;, 4&quot;</td>
<td></td>
</tr>
<tr>
<td>4'</td>
<td>87.3</td>
<td>4.14 (dd, 4.1, 3.4, 3.1)</td>
<td>3&quot;, 5&quot;, 5&quot;</td>
<td></td>
</tr>
<tr>
<td>5'</td>
<td>62.4</td>
<td>3.87 (dd, 12.6, 3.1)</td>
<td>4&quot;, 5&quot;</td>
<td></td>
</tr>
<tr>
<td>1α&quot;</td>
<td>28.4</td>
<td>1.77</td>
<td>4&quot;, 5&quot;</td>
<td></td>
</tr>
<tr>
<td>1β&quot;</td>
<td>1.04</td>
<td>1&quot;, 2&quot;, 2&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2α&quot;</td>
<td>23.0</td>
<td>1.62</td>
<td>1&quot;, 2&quot;, 3&quot;, 3&quot;</td>
<td></td>
</tr>
<tr>
<td>2β&quot;</td>
<td>1.62</td>
<td>1&quot;, 1&quot;, 3&quot;, 3&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3α&quot;</td>
<td>41.9</td>
<td>1.41</td>
<td>3&quot;, 2&quot;, 2&quot;</td>
<td></td>
</tr>
<tr>
<td>3β&quot;</td>
<td>1.21 (dd, 13.1, 13.1, 4.8)</td>
<td>3&quot;, 2&quot;, 2&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6&quot;</td>
<td>117.2</td>
<td>5.48</td>
<td>7&quot;, 7&quot;, 10&quot;</td>
<td></td>
</tr>
<tr>
<td>7α&quot;</td>
<td>32.5</td>
<td>1.87</td>
<td>6&quot;, 7&quot;, 8&quot;, 6&quot;, 7&quot;, 8&quot;, 17&quot;</td>
<td></td>
</tr>
<tr>
<td>7β&quot;</td>
<td>1.77</td>
<td>7&quot;, 7&quot;, 8&quot;, 6&quot;, 7&quot;, 16&quot;, 17&quot;, 20&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8&quot;</td>
<td>34.5</td>
<td>1.54</td>
<td>7&quot;, 7&quot;, 17&quot;</td>
<td></td>
</tr>
<tr>
<td>9&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10&quot;</td>
<td>41.1</td>
<td>2.22 (br d, 12.9)</td>
<td>1&quot;, 1&quot;, 6&quot;</td>
<td></td>
</tr>
<tr>
<td>11&quot;</td>
<td>35.5</td>
<td>1.58</td>
<td>11&quot;, 12&quot;, 12&quot;, 8&quot;, 6&quot;, 12&quot;, 16&quot;, 17&quot;, 20&quot;</td>
<td></td>
</tr>
<tr>
<td>12&quot;</td>
<td>33.7</td>
<td>2.05</td>
<td>11&quot;, 11&quot;, 12&quot;, 10&quot;, 11&quot;, 11&quot;, 14&quot;, 16&quot;, 16&quot;</td>
<td></td>
</tr>
<tr>
<td>13&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14&quot;</td>
<td>115.2</td>
<td>5.46</td>
<td>12&quot;, 15&quot;, 16&quot;, 2&quot;, 12&quot;, 12&quot;</td>
<td></td>
</tr>
<tr>
<td>15&quot;</td>
<td>49.5</td>
<td>4.92&quot;</td>
<td>14&quot;, 2&quot;, 14&quot;, 16&quot;, 2&quot;, 14&quot;, 16&quot;</td>
<td></td>
</tr>
<tr>
<td>17&quot;</td>
<td>15.3</td>
<td>0.945 (d, 6.7)</td>
<td>8&quot;, 7&quot;, 8&quot;, 11&quot;, 14&quot;, 16&quot;, 20&quot;</td>
<td></td>
</tr>
<tr>
<td>18&quot;</td>
<td>30.1</td>
<td>1.06</td>
<td>18&quot;, 3&quot;, 3&quot;, 6&quot;, 19&quot;, 3&quot;, 10&quot;, 18&quot;</td>
<td></td>
</tr>
<tr>
<td>19&quot;</td>
<td>29.3</td>
<td>1.01</td>
<td>18&quot;, 3&quot;, 3&quot;, 18&quot;</td>
<td></td>
</tr>
<tr>
<td>20&quot;</td>
<td>16.5</td>
<td>0.66</td>
<td>1&quot;, 7&quot;, 11&quot;, 17&quot;</td>
<td></td>
</tr>
</tbody>
</table>

*The expected COSY and NOESY cross peaks are seen within the terpene, base and sugar fragments. NOESY cross peaks between the three fragments are seen between H₄ and the three sugar protons H₁₅, H₂₅ and H₃₂ and between H₃ and the terpene protons H₁₄, H₁₄, and H₁₆. The allylic methylene peak is the expected doublet.*
CHAPTER 3

NMR Analysis of N\(^6\)-TbAd isolated from *Mycobacterium tuberculosis*

\[ \text{N}^6\text{-Tuberculosisyl adenosine (N}^6\text{-TbAd)} \]

<table>
<thead>
<tr>
<th>Atom</th>
<th>Carbon</th>
<th>Hydrogen</th>
<th>COSY</th>
<th>NOESY</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>not observed</td>
<td>8.24</td>
<td>8.26</td>
<td>12</td>
</tr>
<tr>
<td>1'</td>
<td>90.8</td>
<td>5.95 (d, 6.6)</td>
<td>2'</td>
<td>8, 2'</td>
</tr>
<tr>
<td>2'</td>
<td>75.3</td>
<td>4.74 (dd, 6.6, 5.0)</td>
<td>1', 3'</td>
<td>1', 3'</td>
</tr>
<tr>
<td>3'</td>
<td>72.4</td>
<td>4.32 (dd, 5.0, 2.6)</td>
<td>2', 4'</td>
<td>2', 5'</td>
</tr>
<tr>
<td>4'</td>
<td>88.0</td>
<td>4.17 (ddd, 2.6, 2.4, 2.4)</td>
<td>3', 5', 5'</td>
<td>3', 5'</td>
</tr>
<tr>
<td>5'</td>
<td>63.3</td>
<td>5.69 (dd, 12.8, 2.4)</td>
<td>4', 5'</td>
<td>4'</td>
</tr>
<tr>
<td>1α&quot;</td>
<td>28.4</td>
<td>1.77</td>
<td>1&quot;, 2&quot;, 2&quot;</td>
<td>1&quot;, 2&quot;, 10&quot;</td>
</tr>
<tr>
<td>1β&quot;</td>
<td>1.04</td>
<td>1&quot;, 1&quot;, 2&quot;, 2&quot;</td>
<td>1&quot;, 2&quot;, 20&quot;</td>
<td></td>
</tr>
<tr>
<td>3α&quot;</td>
<td>23.0</td>
<td>1.62</td>
<td>1&quot;, 1&quot;, 3&quot;, 3&quot;</td>
<td>1&quot;, 3&quot;, 19&quot;</td>
</tr>
<tr>
<td>3β&quot;</td>
<td>1.62</td>
<td>1&quot;, 1&quot;, 3&quot;, 3&quot;</td>
<td>1&quot;, 3&quot;, 19&quot;</td>
<td></td>
</tr>
<tr>
<td>3α&quot;</td>
<td>41.8</td>
<td>1.41</td>
<td>3&quot;, 2&quot;, 2&quot;</td>
<td>2&quot;, 18&quot;, 19&quot;</td>
</tr>
<tr>
<td>3β&quot;</td>
<td>1.21 (dd, 13.1, 13.1, 4.8)</td>
<td>3&quot;, 2&quot;, 2&quot;</td>
<td>2&quot;, 18&quot;</td>
<td></td>
</tr>
<tr>
<td>4&quot;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>5&quot;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>6&quot;</td>
<td>117.2</td>
<td>5.48</td>
<td>7&quot;, 7&quot;, 16&quot;</td>
<td>7&quot;, 7&quot;, 18&quot;</td>
</tr>
<tr>
<td>7α&quot;</td>
<td>32.5</td>
<td>1.86 (br d, 17.3)</td>
<td>6&quot;, 7&quot;, 8&quot;</td>
<td>6&quot;, 7&quot;, 8&quot;, 17&quot;</td>
</tr>
<tr>
<td>7β&quot;</td>
<td>1.77</td>
<td>1&quot;, 1&quot;, 6&quot;</td>
<td>6&quot;, 7&quot;, 16&quot;, 17&quot;, 20&quot;</td>
<td></td>
</tr>
<tr>
<td>8&quot;</td>
<td>34.3</td>
<td>1.54</td>
<td>7&quot;, 7&quot;, 17&quot;</td>
<td>7&quot;, 7&quot;, 16&quot;, 17&quot;</td>
</tr>
<tr>
<td>9&quot;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>10&quot;</td>
<td>40.9</td>
<td>2.25 (br d, 12.9)</td>
<td>1&quot;, 6&quot;, 6&quot;</td>
<td>2&quot;, 7&quot;, 10&quot;, 12&quot;, 19&quot;</td>
</tr>
<tr>
<td>11&quot;</td>
<td>35.8</td>
<td>1.58</td>
<td>11&quot;, 11&quot;, 12&quot;</td>
<td>8&quot;, 10&quot;, 12&quot;, 16&quot;, 17&quot;, 20&quot;</td>
</tr>
<tr>
<td>12&quot;</td>
<td>33.7</td>
<td>1.42</td>
<td>11&quot;, 11&quot;, 12&quot;</td>
<td>8&quot;, 10&quot;, 12&quot;, 16&quot;, 17&quot;, 20&quot;</td>
</tr>
<tr>
<td>13&quot;</td>
<td>2.00</td>
<td>11&quot;, 11&quot;, 12&quot;</td>
<td>10&quot;, 11&quot;, 11&quot;, 14&quot;, 14&quot;, 16&quot;</td>
<td></td>
</tr>
<tr>
<td>17&quot;</td>
<td>15.2</td>
<td>0.85 (d, 6.7)</td>
<td>8&quot;</td>
<td>7&quot;, 7&quot;, 8&quot;, 11&quot;, 11&quot;, 20&quot;</td>
</tr>
<tr>
<td>18&quot;</td>
<td>30.0</td>
<td>1.06</td>
<td>3&quot;, 3&quot;, 6&quot;, 19&quot;</td>
<td></td>
</tr>
<tr>
<td>19&quot;</td>
<td>29.3</td>
<td>1.06</td>
<td>2&quot;, 3&quot;, 10&quot;, 18&quot;</td>
<td></td>
</tr>
<tr>
<td>20&quot;</td>
<td>16.4</td>
<td>0.65</td>
<td>1&quot;, 7&quot;, 11&quot;, 17&quot;</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) The expected COSY and NOESY cross peaks are seen within the terpene, base and sugar fragments. The only NOESY cross peak between the three fragments is between H\(_2\) and H\(_6\). \(^b\) The allylic methylene peak is broad rather than the expected doublet due to slow rotation about the C\(_4\)-N bond.
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis

Experimental procedures and data:

6,6-dimethyl-1-vinylcyclohex-1-ene (1):
For the synthesis of diene 1 see chapter 4.

(1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydropyran-1-yl)methanol (3):
6,6-dimethyl-1-vinylcyclohex-1-ene 1 (1.5 g, 11.0 mmol) was dissolved in (E)-ethyl 2-methylbut-2-enoate (4.55 ml, 33.0 mmol 3 eq) in a pressure vessel and heated at 160 °C for 3.5 d. After the reaction mixture was cooled to rt, the excess of ethyl tiglate was removed under reduced pressure. GC-MS indicated near pure ethyl 1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydropyran-1-carboxylate 2 as an exo:endo mixture of 2:1. The near pure mixture was used in the next step without further purification.

To a solution of crude Diels-Alder adduct 2 (2.91 g, 11.0 mmol) in dry THF (55 mL) was added portion wise lithium aluminum hydride (835 mg, 22.0 mmol). After addition the reaction was warmed up to 40 °C and stirred for 1 h. GC-MS indicated complete conversion of the starting material. The reaction mixture was cooled to -10 °C using an ice/salt bath, and quenched using water. After phase separation the etherial layer was washed with an aqueous 1 M HCl solution, water, and brine, where after it was dried over MgSO₄, filtered and concentrated under reduced pressure to afford 2.7 g of a yellow oil. Flash column chromatography using pentane : ether (4:1) gave (1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydropyran-1-yl)methanol 3 (1.7 g, 7.7 mmol, 70% yield over the two steps).

\[ ^1H-NMR (400 MHz, CDCl_3) \delta 5.43 (d, J = 5.2 Hz, 1H), 3.48 (d, J = 11.4 Hz, 1H), 3.39 (d, J = 11.3 Hz, 1H), 2.36 (d, J = 12.9 Hz, 1H), 1.91 – 1.81 (m, 1H), 1.81 – 1.71 (m, 2H), 1.70 – 1.63 (m, 2H), 1.61 – 1.51 (m, 2H), 1.38 (s, 1H), 1.28 – 1.14 (m, 2H), 1.05 (s, 3H), 1.00 (s, 3H), 0.86 (d, J = 6.5 Hz, 3H), 0.51 (s, 3H). \]

\[ ^13C-NMR (101 MHz, CDCl_3) \delta 146.12, 116.11, 65.78, 41.07, 39.33, 38.05, 36.27, 31.93, 31.33, 29.86, 28.86, 27.74, 22.26, 15.09, 11.63. \]

HRMS (ESI+): Calculated mass [M+H]^+ C_{15}H_{27}O = 223.2056; found: 223.2059

1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydronaphthalene-1-carbaldehyde (4):
Alcohol 3 (1.60 g, 7.2 mmol) was dissolved in CH$_2$Cl$_2$ (30 mL) whereafter TPAP (126 mg, 0.36 mmol, 5 mol%) and NMO (1.01 g, 8.64 mmol, 1.2 eq) were added, and the reaction mixture was stirred for 1 h. When TLC analysis confirmed completion of the reaction, the solvent was evaporated, and the crude was filtered through a short silica column, flushing with a 1:1 pentane : ether mixture, giving aldehyde 4 (1.58 g, 99% yield).

$^1$H-NMR (400 MHz, CDCl$_3$) $\delta$ 9.38 (s, 1H), 5.52 – 5.47 (m, 1H), 2.50 (d, $J = 13.1$ Hz, 1H), 2.00 – 1.91 (m, 1H), 1.90 – 1.80 (m, 1H), 1.80 – 1.69 (m, 1H), 1.57 – 1.49 (m, 2H), 1.45 – 1.37 (m, 1H), 1.36 – 1.28 (m, 1H), 1.27 – 1.14 (m, 2H), 1.08 – 1.04 (s, 3H), 0.80 (s, 3H), 0.78 (d, $J = 6.5$ Hz, 3H).

$^{13}$C-NMR (101 MHz, CDCl$_3$) $\delta$ 207.16, 143.92, 116.56, 52.26, 40.61, 38.15, 36.28, 32.65, 30.47, 29.66, 29.13, 28.68, 21.77, 16.14, 7.49.

Some of the characteristic signals of the endo product are:
$^1$H-NMR (400 MHz, CDCl$_3$) $\delta$ 9.65 (s, 1H), 1.06 (s, 3H), 1.04 (s, 3H), 1.02 (s, 3H), 0.99 (s, 3H), 0.92 (s, 3H), 0.82 (d, $J = 6.7$ Hz, 3H), 0.79 (s, 3H).


1,1,5,6-tetramethyl-5-vinyl-1,2,3,4,4a,5,6,7-octahydronaphthalene (5):
A stirred suspension of methyltriphenylphosphonium bromide (585 mg, 1.64 mmol, 2.2 eq) in dry THF (9 mL) was treated with a solution of LiHMDS (1.5 mL, 1.5 mmol, 1 M in THF, 2.0 eq) at 0 °C. The initial suspension was stirred for 30 min after which a clear bright yellow solution was obtained. To the solution, still at 0 °C, was added dropwise
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis

1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydonaphthalene-1-carbaldehyde 4 (164 mg, 0.744 mmol) in dry THF (9 mL). The resulting reaction mixture was allowed to warm up to rt and was stirred overnight. The reaction was quenched with a saturated solution of NaHCO$_3$ and diluted with ether. The phases were separated and the organic layer was washed with distilled water, brine and dried over Na$_2$SO$_4$. Flash chromatography employing pentane furnished 1,1,5,6-tetramethyl-5-vinyl-1,2,3,4,4a,5,6,7-octahydonaphthalene 5 (155 mg, 0.710 mmol, 95% yield) as a colorless oil.

$^1$H-NMR (400 MHz, CDCl$_3$) δ 5.55 (dd, $J = 17.5, 10.8$ Hz, 1H), 5.49 – 5.45 (m, 1H), 5.06 (dd, $J = 10.8, 1.6$ Hz, 1H), 4.92 (dd, $J = 17.5, 1.6$ Hz, 1H), 2.03 – 1.96 (m, 1H), 1.95 – 1.84 (m, 1H), 1.78 – 1.69 (m, 1H), 1.67 – 1.59 (m, 1H), 1.57 – 1.36 (m, 4H), 1.35 – 1.15 (m, 1H), 1.07 (s, 3H), 1.01 (s, 3H), 0.97 – 0.91 (m, 1H), 0.75 (d, $J = 6.8$ Hz, 3H), 0.71 (s, 3H).


(but-2-yn-1-yloxy)(tert-butyl)dimethylsilane (E):

To a solution of but-2-yn-1-ol D (1.0 ml, 13.4 mmol) and triethylamine (2.0 mL, 14.0 mmol, 1.05 eq) in CH$_2$Cl$_2$ (60 ml) at rt was added TBSCI (2.0 g, 13.4 mmol, 1.0 eq). The reaction mixture was refluxed overnight, cooled to rt, and quenched with a saturated NaHCO$_3$ solution (10 mL). The two phases were separated and the organic layer was washed with water (3x20 mL). The organic layer was dried over Na$_2$SO$_4$, filtered and concentrated under reduced pressure. Flash column chromatography employing pentane as an eluent furnished (but-2-yn-1-yloxy)(tert-butyl)dimethylsilane E (2.2 g, 11.9 mmol, 89% yield) as a colorless oil.

$^1$H-NMR (400 MHz, CDCl$_3$) δ 4.28 (s, 2H), 1.83 (t, $J = 2.4$ Hz, 3H), 0.91 (s, 9H), 0.86 (s, 6H).

The analytical data are in agreement with: D. F. Harvey, D. A. Neil, Tetrahedron 1993, 49, 2145.
Schwartz reagent:
To a suspension of ZrCp₂Cl₂ (2.0 gr, 6.84 mmol) in dry THF (15.6 mL) was added dropwise neat DIBAL-H (1.21 mL, 6.84 mmol, 1 eq). The reaction was stirred for 45 min where after the suspension was decanted. The white solid (the Schwartz reagent) was washed 3 times with THF (3x8 mL) and vacuum was applied to evaporate the residual solvent. The remaining white solid (0.8 g, 45% yield) was transferred into a fridge in a glovebox.

The low yield can be attributed to the solubility of Schwartz reagent in THF. Next experiments should be performed using cold THF and with a double Schlenk flask with a glass filter in between to make the washing and decantation more effective.

The analytical data are in agreement with: Z. Huang, E-I. Negishi, Org. Lett. 2006, 8, 3675.

(E)-tert-butyl((3-iodobut-2-en-1-yl)oxy)dimethylsilane (6):
To a suspension of Schwartz reagent (462 mg, 1.8 mmol, 1.1 eq) in dry THF (4 mL) was added (but-2-yn-1-yloxy)((tert-butyl)dimethylsilane E (300 mg, 1.62 mmol) in dry THF (2 mL). The reaction was stirred for 2.5 h. The resulting suspension was cooled to -78 °C where after I₂ (537 mg, 2.1 mmol, 1.3 eq) was added. The reaction was stirred for 90 min after which the reaction was quenched with an aqueous 1 M HCl solution (3 mL). The two phases were separated and the combined organic layers were washed thoroughly with an aqueous saturated Na₂S₂O₃ solution (3x5 mL) and brine (5 mL). The organic layer was dried using Na₂SO₄, filtered and concentrated under reduced pressure. Flash column chromatography employing pentane : ether (97:3) as the eluent afforded pure vinyl iodide 6 (405 mg, 1.2 mmol, 80% yield).

¹H-NMR (400 MHz, CDCl₃) δ 5.75 (t, J = 4.7 Hz, 1 H), 4.25 (d, J = 4.7 Hz, 2 H), 2.57 (s, 3 H), 0.98 (s, 9 H) 0.16 (s, 6 H).

¹³C-NMR (101 MHz, CDCl₃) δ 135.66, 99.30, 68.73, 33.84, 26.31 (3xC), 18.49, -4.96 (2xC).
The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*


(E)-3-methyl-5-(1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydropyridal-1-yl)pent-2-en-1-ol (tuberculosinol):

To alkene 5 (300 mg, 1.37 mmol) was added a solution of 9-BBN (4.2 mL, 0.5 M in THF, 1.5 eq). The resulting mixture was heated to 70 °C for 3 h, and cooled to rt, after TLC analysis indicated full consumption of the starting material. Iodo-alkene 6 (1.14 g, 2.74 mmol, 2 eq) was dissolved in degassed DMF (10 mL). Pd(dppf)Cl₂ (100 mg, 0.137 mmol, 0.1 eq), triphenyl arsine (42 mg, 0.137 mmol, 0.1 eq), cesium carbonate (110 mg, 0.34 mmol, 0.25 eq) and degassed distilled water (0.95 mL, 54 mmol, 40 eq) were added, yielding a dark solution with Cs₂CO₃ being suspended.

The hydroboration reaction was added to the DMF mixture mentioned above, using syringe. The combined reaction was immediately degassed after mixing, utilizing three freeze-pump-thaw cycles. After vigorously stirring the reaction overnight at rt, the reaction was checked by TLC and GC-MS. As the conversion was complete, the mixture was diluted with water, and extracted three times with diethyl ether. The combined organic layers were dried with magnesium sulfate, concentrated *in vacuo* and purified using column chromatography, using pentane : ether (95:5) as an eluent.

The obtained product was stirred with TBAF (3 mL, 1 M in THF, 3.0 mmol, 2.2 eq) in CH₂Cl₂ (10 mL). After 1 h, the reaction was complete, as indicated by TLC. The reaction mixture was concentrated *in vacuo*, and purified by flash chromatography (pentane : ether 7:3) yielding rac-tuberculosinol as a colorless oil (64% yield over two steps).

**¹H-NMR** (400 MHz, CDCl₃) δ 5.43 – 5.33 (m, 2H), 4.08 (d, *J* = 6.8 Hz, 2H), 2.57 (s, 1H), 2.13 (d, *J* = 12.6 Hz, 1H), 1.92 – 1.85 (m, 2H), 1.81 – 1.67 (m, 3H), 1.65 (s, 3H), 1.59 – 1.28 (m, 6H), 1.28 – 1.10 (m, 1H), 1.02 (s, 3H), 0.97 (s, 3H), 0.78 (d, *J* = 6.8 Hz, 3H), 0.59 (s, 3H).

**¹³C-NMR** (101 MHz, CDCl₃) δ 145.99, 140.10, 123.11, 116.13, 59.09, 40.94, 39.81, 36.92, 36.05, 34.97, 33.38, 32.76, 31.65, 29.78, 29.00, 27.44, 22.27, 16.46, 16.18, 15.10.
Some of the characteristic signals of the endo product are \(^1\)H-NMR (400 MHz, CDCl\(_3\)) \(\delta 5.29 - 5.15\) (m, 2H), \(4.02\) (d, \(J = 6.8\), 2H), \(2.4 - 0.7\) (m, 13H), \(1.56\) (s, 3H), \(0.93\) (s, 3H), \(0.87\) (s, 3H), \(0.68\) (d, \(J = 6.7\), 3H), \(0.49\) (s, 3H).


\[
\begin{align*}
\text{Tuberculosinyl chloride} \quad \text{Method A} \\
\text{NCS (20 mg, 0.15 mmol, 1.1 eq) was dissolved in CH}_2\text{Cl}_2 (5 \text{ mL}), and cooled to -30 \degree\text{C}. DMS (10 mg, 0.16 mmol, 1.16 eq) was added, and the mixture was allowed to warm to 0 \degree\text{C} for 5 min. Subsequently, the mixture was cooled to -40 \degree\text{C}, and tuberculosinol (40 mg, 0.14 mmol, 1 eq) in CH\(_2\)Cl\(_2\) (2 mL) was added over 3 min. The reaction mixture was removed from the cooling bath, and stirred at 0 \degree\text{C} for 15 min, before being stirred at rt for 1 h. TLC analysis showed complete conversion, and after evaporating the solvent, the crude mixture was purified by filtration on silica with pentane. Tuberculosinyl chloride 10 was obtained as a colorless oil (41 mg, 96% yield).*
\]

*Typical yields obtained employing method A varied from 65% to 96%.

\[
\begin{align*}
\text{Tuberculosinol} \quad \text{Method B} \\
\text{To a solution (0.2 M) of tuberculosinol (32 mg, 0.103 mmol) in dry CH}_2\text{Cl}_2 was added triethylamine (15.6 mg, 21.6 \mu\text{L, 0.154 mmol, 1.5 eq). The obtained solution was cooled with an ice/salt bath and mesyl chloride (13.0 mg, 8.8 \mu\text{L, 0.114 mmol, 1.1 eq) was added. The ice/salt bath was removed and the reaction was allowed to stir at rt for 45 min whereupon NMR analysis indicated that the reaction was complete. The reaction mixture was the diluted with Et}_2\text{O, transferred to a separatory funnel and washed with water (2 mL), HCl solution (2 mL, 2 M), saturated aqueous NaHCO}_3 \text{ solution (2 mL) and brine (2 mL). The organic layer was dried using Na}_2\text{SO}_4, filtered and concentrated under reduced pressure. Tuberculosinyl chloride 10 was isolated as a yellowish oil (32 mg, 94% yield).}
\]

\[
\begin{align*}
\text{1H-NMR (400 MHz, CDCl}_3\text{) } \delta 5.51 - 5.41\) (m, 2H), \(4.09\) (d, \(J = 7.9\) Hz, 2H), \(2.16\) (d, \(J = 12.8\) Hz, 1H), \(1.96\) (dt, \(J = 10.4,\) 4.6 Hz, 2H), \(1.83\) (ddd, \(J = 17.7, 14.5, 3.5\) Hz, 2H), \(1.75\) (s, 3H), \(1.72\) (s, 1H), \(1.63 - 1.25\) (m, 7H), \(1.20\) (td, \(J = 12.9, 4.8\) Hz, 1H), \(1.06\) (s, 3H), \(1.01\) (s, 3H), \(0.82\) (d, \(J = 6.8\) Hz, 3H), \(0.63\) (s, 3H).
\]
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis

$^{13}$C-NMR (101 MHz, CDCl$_3$) δ 146.11, 144.03, 119.92, 116.27, 41.34, 41.05, 39.93, 37.11, 36.20, 34.82, 33.51, 32.88, 31.75, 29.90, 29.13, 27.56, 22.37, 16.46, 16.28, 15.22.

HRMS (APCI): Calculated mass [M-Cl]$^+$ C$_{26}$H$_{36}$O$^+ = 273.2582; found: 273.2576.

**Important note:** The tuberculosinyl chloride 8 was purified using column chromatography. Later work showed purification by these means led to degradation of the product on the column. Therefore an alternative purification strategy was developed which is presented in the experimental section of chapter 4. This procedure is recommended as the stability of the tuberculosinyl chloride on silica is suspected to be batch dependent.

6-amino-9-((2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-1-1(E)-3-methyl-5-(1,2,5,5-tetramethyl-1,2,3,5,6,7,8a-octahydrobenzo[1,2]j)pent-2-en-1-yl)purin-1-ium (1-TbAd):

To a solution of 10 (11 mg, 35.6 μmol) in peptide grade DMF (0.7 mL) was added adenosine (11.4 mg, 42.7 μmol, 1.2 eq) and sodium iodide (8 mg, 53.4 μmol, 1.5 eq). The resulting reaction mixture was exposed to microwave heating (45 °C with an initial overshoot to 48 °C) for 65 h where after it was concentrated under reduced pressure. The resulting oil was purified employing preparative HPLC. 1-TbAd was obtained a colorless oil (3 mg, 17% yield).

$^1$H-NMR (400 MHz, CD$_3$OD) δ 8.62 (s, 1H), 8.49 (s, 1H), 6.08 (d, $J = 5.2$ Hz, 1H), 5.51 – 5.42 (m, 2H), 4.91 (d, $J = 6.6$ Hz, 2H), 4.62 (t, $J = 5.1$ Hz, 2H), 4.58 (s, 1H), 4.38 – 4.31 (m, 2H), 4.15 (q, $J = 3.3$ Hz, 1H), 3.87 (dd, $J = 12.3, 2.9$ Hz, 1H), 3.77 (dd, $J = 12.2, 3.4$ Hz, 1H), 2.23 (d, $J = 12.1$ Hz, 2H), 2.12 – 2.04 (m, 2H), 1.89 (s, 3H), 1.87 – 1.82 (m, 1H), 1.84 – 1.71 (m, 2H), 1.65 – 1.47 (m, 5H), 1.47 – 1.36 (m, 3H), 1.21 (ddd, $J = 12.7, 5.7$ Hz, 1H), 1.06 (s, 3H), 1.01 (s, 3H), 0.85 (d, $J = 6.7$ Hz, 3H), 0.66 (s, 3H).

$^{13}$C-NMR (101 MHz, CD$_3$OD) δ 152.52, 147.78, 147.58, 147.31, 147.08, 143.74, 121.49, 117.34, 115.89, 90.43, 87.41, 76.35, 71.79, 62.62, 49.39, 42.00, 41.06, 38.05, 36.94, 35.87, 34.52, 33.88, 32.62, 30.29, 29.49, 28.53, 23.18, 17.35, 16.58, 15.54.

HRMS (ESI+): Calculated mass [M+H]$^+$ C$_{30}$H$_{46}$N$_5$O$_4^+$ = 540.3544; found: 540.3542.
CHAPTER 3

Tuberculosinyl pyrophosphate:
To a solution of TTBAHPP (58 mg, 64.3 μmol, 2 eq.) in CH$_3$CN (1 mL), in an oven-dried Schlenk flask under nitrogen atmosphere, was added a solution of tuberculosinyl chloride 8 (10 mg, 32.2 μmol) in dry CH$_3$CN (0.5 mL). The solution was stirred for 3 h after which TLC analysis, using pentane as an eluent, indicated complete conversion of the starting material. The solvent was removed under reduced pressure where after the residue was dissolved in dry methanol and passed through a pre-washed column DOWEX$^®$ 50WX2 Na$^+$-form (50-100 mesh). This process was repeated twice after which the methanol was evaporated. The compound was used without further purification.

(2E,6E,10E)- and (2Z,6E,10E)-3,7,11,15-tetramethyl-2,6,10,14-hexadecatetraenitrile (10):
A solution of farnesylacetone 9 (5.0 g, 19.1 mmol) in MeCN (20 mL) was added to a stirred solution of powdered KOH (1.3 g, 23.2 mmol, 1.2 eq) in MeCN (30 mL) at reflux. The mixture was subsequently stirred overnight at reflux. The mixture was cooled down to rt, 30 mL of water was added, and the mixture was extracted with ether. The combined organic layers were washed with brine, dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (50:1) pentane : EtOAc to afford 10 (2.7 g, 50%, $E:Z = 2:1$) as a pale yellow oil.

$^1$H-NMR (400 MHz, CDCl$_3$) $\delta$ 5.12 – 5.03 (m, 4H), 2.46 – 2.41 (m, 1H), 2.24-2.15 (m, 3H), 2.11 – 1.96 (m, 10H), 1.91 – 1.90 (m, 1H), 1.68 (s, 6H), 1.60 (s, 6H).

(2E,6E,10E)- and (2Z,6E,10E)-3,7,11,15-tetramethyl-2,6,10,14-hexadecatetraenal (11):
To a solution of 10 (1.6 g, 5.6 mmol) in dry hexane (10 mL) at rt was slowly added DIBAL-H (10 mL, 1M in hexane, 10 mmol, 1.8 eq). The reaction mixture was stirred
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis

for 2 h at rt and carefully hydrolyzed by addition of 10% aq. HCl. The solid material was filtered off and washed with ether. The combined organic layers were washed with brine, dried over Na₂SO₄, filtered off and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (pentane : EtOAc, 20:1) to afford 11 (1.5 g, 90%, 2:1 E:Z = 2:1) as a pale yellow oil.

1H-NMR (400 MHz, CDCl₃) δ 10.00 – 9.89 (m, 1H), 5.88 (d, J = 8.0 Hz, 1H), 5.12 – 5.07 (m, 3H), 2.27 – 2.16 (m, 5H), 2.07 – 1.97 (m, 10H), 1.68 (s, 6H), 1.60 (s, 6H).

(2E,6E,10E)- and (2Z,6E,10E)-3,7,11,15-Tetramethyl-2,6,10,14-hexadecatetraenyl-1H-isoinodole-1,3(2H)-dione (13):
To a solution of 11 (129 mg, 0.45 mmol) in dry THF (2 mL) was added dropwise LiAlH₄ (0.23 mL, 1M in THF, 0.23 mmol, 0.5 eq) at 0 °C under nitrogen atmosphere. The mixture was stirred at this temperature for 1 h and ether (3 mL) was added, followed by 1 M NaOH (1 mL). The solid material was filtered off and washed with ether. The filtrate was concentrated to give crude geranylggeraniol (12) as a pale yellow oil, which was used for the next step with further purification.

1H-NMR (400 MHz, CDCl₃) δ 5.43 – 5.40 (m, 1H), 5.13 – 5.09 (m, 3H), 4.16 – 4.08 (m, 2H, E:Z = 4:1), 2.14 – 1.95 (m, 13H), 1.68 (s, 8H), 1.60 (2s, 6H).

To a solution of crude 12 in THF (3 mL) was added triphenylphosphine (154 mg, 0.59 mmol, 1.3 eq), diethyl azodicarboxylate (DEAD, 0.27 mL, 40% in toluene, 0.69 mmol, 1.5 eq) and phthalimide (87 mg, 0.59 mmol, 1.3 eq) at 0 °C. The mixture was stirred at this temperature for 1 h and subsequently concentrated under reduced pressure. The residue was dissolved in CH₂Cl₂ (10 mL) and the solution was washed with 1 M aq. NaOH and brine, dried over Na₂SO₄, filtered, and concentrated under reduced pressure. Crude 13 was purified by column chromatography on silica gel (40:1 pentane : EtOAc) to afford 13 (140 mg, 75% over two steps) as a pale yellow oil.

1H-NMR (400 MHz, CDCl₃) δ 7.83 – 7.81 (m, 2H), 7.71 – 7.68 (m, 2H), 5.29 – 5.25 (m, 1H), 5.10 – 5.04 (m, 3H), 4.27 (d, J = 7.2, 2H), 2.10 – 1.92 (m, 12H), 1.83 – 1.71 (m, 3H), 1.66 – 1.63 (m, 6H), 1.60 – 1.56 (m, 6H) The spectral data are identical to those previously reported.
(2E,6E,10E)- and (2Z,6E,10E)-3,7,11,15-tetramethyl-2,6,10,14-hexadecatetraenyl-1-amine (14): 
To a solution of 13 (140 mg, 0.33 mmol) in methanol (3 mL) was added hydrazine hydrate (24 μL, 0.50 mmol, 1.5 eq). The resulting mixture was heated at reflux for 1 h and then concentrated under reduced pressure. The residue was dissolved in 1 M aq. HCl and solid material was removed by filtration. The filtrate was made basic with 5 M aq. NaOH, filtered and concentrated under reduced pressure to afford 14 (68 mg, 70%), which was used for the next step without further purification.

N6-geranylgeranyladenosine: 
To a solution of 14 (30 mg, 0.104 mmol) in DMF (1.5 mL) was added inosine (28 mg, 0.104 mmol, 1 eq), BOP (benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate, 55 mg, 0.12 mmol, 1.2 eq) and DIPEA (27 μL, 0.16 mmol, 1.5 eq) at rt under nitrogen atmosphere. The reaction was stirred overnight at this temperature. The solvent was removed under reduced pressure and the crude product was purified by column chromatography on silica gel using CH2Cl2: CH3OH (20:1) to afford N6-geranylgeranyl adenosine (45 mg, 72%, mixture of E and Z isomers) as a dark yellow oil.

1H-NMR (400 MHz, CDCl3) δ 8.11 (br s, 1H), 7.75 (s, 1H), 5.92 (br s, 1H), 5.78 (d, J = 7.2 Hz, 1H), 5.3.6 – 5.32 (m, 1H), 5.09 (s, 3H), 5.00 – 4.97 (m, 1H), 4.42 (d, J = 4.8 Hz, 1H), 4.30 (s, 1H), 4.25 – 4.05 (2x br s, 2H, rotamers due to slow rotation about the C-N6 bond), 3.91 (d, J = 12.0 Hz, 1H), 3.71 (d, J = 12.0 Hz, 1H), 2.13-1.94 (m, 12H), 1.72 (m,3H), 1.66 (s, 6H), 1.59 (s, 6H), 1.48 – 1.41 (m, 3H, OH); (the peak at δ 2.55 (d, J = 9.6 Hz) is due to residual HMPA).

The discovery of novel terpene nucleosides from *Mycobacterium tuberculosis*

2-((E)-3-methyl-5-(1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydonaphthalen-1-yl)pent-2-en-1-yl)isoindoline-1,3-dione (15):
Chloride 8 (35 mg, 0.114 mmol) was dissolved in peptide grade DMF (10 mL) and potassium phthalimide (31.6 mg, 0.171 mmol, 1.5 eq) was added. The solution was stirred for 8 h at 100 °C. The resulting mixture was concentrated, and purified by flash chromatography with pentane : ether (4:1), yielding phthalimide 15 (34 mg, 72%).

\[ \text{H-NMR (400 MHz, CDCl}_3 \] \( \delta \) 7.83 (dd, \( J = 5.4, 3.0, 2H \)), 7.70 (dd, \( J = 5.3, 3.1, 2H \)) \( \delta \) 5.42 – 5.26 (m, 2H), 4.27 (d, \( J = 8.0, 2H \)), 2.19 – 1.86 (m, 5H), 1.85 (s, 3H) 1.84 – 1.10 (m, 9H), 1.03 (s, 3H), 0.97 (s, 3H), 0.78 (d, \( J = 6.7, 3H \)), 0.59 (s, 3H).

\[ \text{C-NMR (101 MHz, CDCl}_3 \] \( \delta \) 168.15, 146.07, 141.71, 133.77, 132.33, 123.13, 117.40, 116.10, 40.91, 39.74, 36.93, 36.06, 35.87, 34.66, 33.35, 32.69, 31.62, 29.75, 28.99, 27.38, 22.21, 16.63, 16.10, 15.07.

HRMS (ESI+): calculated for [M+H]^+ \( C_{28}H_{37}NO_2^+ \) = 420.2903 found 420.2891.

((E)-3-methyl-5-(1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydonaphthalen-1-yl)pent-2-en-1-yl)pent-2-en-1-amine (16):
Phthalimide 15 (13 mg, 0.032 mmol) was dissolved in methanol (2 mL), and treated with hydrazine hydrate (3.3 μL, 0.62 mmol, 2 eq). The reaction was heated to reflux for 20 min, after which TLC indicated the reaction was complete. The methanol was completely evaporated under reduced pressure, and the resulting oil was diluted with water, and extracted twice with pentane. The combined pentane layers were dried using magnesium sulfate, and concentrated under reduced pressure to yield amine 16 as slightly yellow oil (8 mg, 90%).

\[ \text{H-NMR (400 MHz, CDCl}_3 \] \( \delta \) 5.50 – 5.26 (m, 2H), 3.27 (d, \( J = 6.7, 2H \)), 2.20 – 1.68 (m, 5H), 1.65 (s, 3H), 1.60 – 1.15 (m, 11H), 1.06 (s, 3H), 1.00 (s, 3H), 0.81 (d, \( J = 6.7, 3H \)), 0.61 (s, 3H).

\[ \text{C-NMR (101 MHz, CDCl}_3 \] \( \delta \) 146.10, 137.48, 125.28, 116.11, 40.94, 39.79, 39.68, 36.93, 36.07, 35.08, 33.37, 32.69, 31.65, 29.77, 29.69, 29.01, 27.42, 22.26, 16.33, 16.17, 15.08.
HRMS (ESI+): calculated for [M+H]^+ C_{20}H_{35}N^+ = 290.2848 found 290.2843.

**CHAPTER 3**

(2R,3S,4R,5R)-2-(hydroxymethyl)-5-(6-(((E)-3-methyl-5-(1,2,5,5-tetramethyl-1,2,3,5,6,7,8,8a-octahydroronaphthalen-1-yl)pent-2-en-1-yl)amino)-9H-purin-9-yl)tetrahydrofuran-3,4-diol (N\textsuperscript{6}-TbAd):

Amine 16 (7 mg, 0.024 mmol), inosine (7 mg, 0.026 mmol, 1.1 eq), benzotriazol-1-yloxytris(dimethylamino)-phosphonium hexafluorophosphate (Castro’s reagent, BOP) (13.7 mg, 0.0312 mmol, 1.3 eq.) and DIPEA (7 μl, 0.038 mmol, 1.5 eq) were dissolved in peptide grade DMF (3 mL).* The reaction mixture was stirred overnight at rt, after which the solvent was removed under reduced pressure. The product was purified by repeated (3 times) flash chromatography using CH\textsubscript{2}Cl\textsubscript{2}: methanol (9:1).** N\textsuperscript{6}-TbAd was isolated as a white waxy solid (1 mg, 7%, ±15% HMPA impurity***). The NMR spectra of synthetic material matched those obtained from the natural source.

* It was observed that the use of peptide grade DMF significantly improved the reaction. Small quantities of dimethyl amine in standard DMF might interfere with the reaction.

** Due to the very similar R\textsubscript{f} of HMPA (0.22) and N\textsuperscript{6}-TbAd (0.21), repetitive column chromatography was required to separate the majority of the HMPA.

*** Appears as a doublet at 2.64 ppm in \textsuperscript{1}H-NMR.
The discovery of novel terpene nucleosides from Mycobacterium tuberculosis

(2R,3S,4R,5R)-2-(hydroxymethyl)-5-(6-((E)-3-methyl-5-(1,2,5,5-tetramethyl-1,2,3,5,6,7,8a-octahydropyridin-1-yl)pent-2-en-1-ylamino)-9H-purin-9-yl)tetrahydrofuran-3,4-diol (N6-TbAd):

A solution of 1-TbAd (5 mg, 0.01 mmol) in 60% Me2NH in water (5.5 mL) was stirred for 90 min. NMR analysis indicated complete conversion of the 1-TbAd. The reaction mixture was concentrated under reduced pressure and subsequently subjected to flash column chromatography, with 15% MeOH in CH2Cl2, to afford N6-TbAd as a white solid (5 mg, 0.01 mmol, quantitative yield) as yellowish oil.

Note: The rearrangement could be performed with similar results using Et3NH or iPr2NEt (2M) in MeOH.

1H-NMR (400 MHz, CD3OD) δ 8.25 (s, 1H), 8.23 (s, 1H), 5.95 (d, J = 6.4 Hz, 1H), 5.51 – 5.44 (m, 1H), 5.41 (t, J = 6.5 Hz, 1H), 4.74 (t, J = 5.6 Hz, 1H), 4.35 – 4.29 (m, 1H), 4.20 (s, 1H), 4.17 (s, 1H), 3.89 (dd, J = 12.6, 2.1 Hz, 1H), 3.74 (dd, J = 12.9, 2.3 Hz, 1H), 2.24 (d, J = 15.9 Hz, 1H), 2.07 – 1.93 (m, 3H), 1.91 – 1.82 (m, 2H), 1.80 (s, 3H), 1.65 – 1.49 (m, 4H), 1.49 – 1.36 (m, 3H), 1.21 (td, J = 11.9, 7.3 Hz, 2H), 1.06 (s, 3H), 1.02 (s, 3H), 0.85 (d, J = 6.6 Hz, 3H), 0.65 (s, 3H).

13C-NMR (101 MHz, CDCl3) δ 154.59, 152.44, 147.00, 146.09, 141.73, 140.05, 120.70, 118.96, 116.23, 91.08, 87.56, 73.89, 72.46, 63.15, 41.00, 39.89, 38.87, 37.04, 36.17, 35.00, 33.47, 32.80, 31.71, 29.87, 29.12, 27.51, 22.33, 16.83, 16.26, 15.22.


The spectral data are consistent with that of the natural isolate.
CHAPTER 3

3.7 References


The discovery of novel terpene nucleosides from Mycobacterium tuberculosis


CHAPTER 3


