University of Groningen

Asymmetric Total Synthesis and Structural Revision of DAT$_2$, an Antigenic Glycolipid from Mycobacterium tuberculosis

Lin, Zonghao; Kaniraj, Jeya Prathap; Holzheimer, Mira; Nigou, Jérôme; Gilleron, Martine; Hekelaar, Johan; Minnaard, Adriaan J

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:
2024

Link to publication in University of Groningen/UMCG research database

Accepted Article

Authors: Zonghao Lin, Jeya Prathap Kaniraj, Mira Holzheimer, Jérôme Nigou, Martine Gilleron, Johan Hekelaar, and Adriaan J. Minnaard

RESEARCH ARTICLE

Asymmetric Total Synthesis and Structural Revision of DAT₂, an Antigenic Glycolipid from *Mycobacterium tuberculosis*

Zonghao Lin[a], Jeya Prathap Kaniraj[a], Mira Holzheimer[a], Jérôme Nigou[b], Martine Gilleron[b], Johan Hekelaar[a], and Adriaan J. Minnaard*[a]

[a] M. Sc. Z. Lin, Dr. J. P. Kaniraj, Dr. M. Holzheimer, J. Hekelaar, Prof. Dr. A. J. Minnaard
     Stratingh Institute for Chemistry, University of Groningen
     Nijenborgh 7, 9747 AG, Groningen (The Netherlands)
     E-mail: a.j.minnaard@rug.nl

[b] Dr. M. Gilleron and Dr. J. Nigou
     Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse,
     CNRS, UPS, 205 route de Narbonne, F-31077 Toulouse (France)

     Supporting information for this article is given via a link at the end of the document.

Abstract: DAT₂ is a member of the diacyl trehalose family (DAT) of antigenic glycolipids located in the mycomembrane of *Mycobacterium tuberculosis* (*Mtb*). Recently it was shown that the molecular structure of DAT₂ had been incorrectly assigned, but the correct structure remained elusive. Herein, the correct molecular structure of DAT₂ and its methyl-branched acyl substituent mycolipanolic acid is determined. For this, four different stereoisomers of mycolipanolic acid were prepared in a stereoselective and unified manner, and incorporated into DAT₂. A rigorous comparison of the four isomers to the DAT isolated from *Mtb H37Rv* by NMR, HPLC, GC, and mass spectrometry allowed a structural revision of mycolipanolic acid and DAT₂. Activation of the macrophage inducible Ca²⁺-dependent lectin receptor (Mincle) with all four stereoisomers shows that the natural stereochemistry of mycolipanolic acid / DAT₂ provides the strongest activation, which indicates its high antigenicity and potential application in serodiagnostics and vaccine adjuvants.

Introduction

Mycobacterium tuberculosis (*Mtb*) is the causative agent of the disease tuberculosis (TB). In 2021, around 10.6 million people were diagnosed and 1.6 million people died from TB, the second leading infectious killer after COVID-19 (more than HIV/AIDS).[1] The mycobacterial cell envelope is a contributing factor to the resilience of *Mtb* in host cells.[2] About 40% of the dry weight of its cell wall is composed of lipids[3] and a large percentage of the coding capacity of the bacterial genome is used for lipid biosynthesis and degradation.[4] The lipid-rich cell wall of *Mtb* is involved in the regulation of the transport of nutrients, toxic host-cell effector molecules, and anti-tuberculosis drugs. In the last decades, it has been shown that many of these cell wall components have antigenic properties and/or possess immunomodulatory functions[5]. Long-chain multi-methyl-branched fatty esters containing trehalose make up a significant part of the membrane, are unique to *Mycobacteria,* and are involved in pathogenesis.[6] Di-*O*-acyl trehaloses include a family of glycolipids (Figure 1), DAT₁, DAT₂, DAT₃ and DAT₄.[7] DAT₃ and DAT₄ were isolated as minor components and DAT₄ remains unidentified.[8] "DAT" often refers to the family of the four

compounds because it is impossible to get the pure DAT compounds from the natural extract[9].
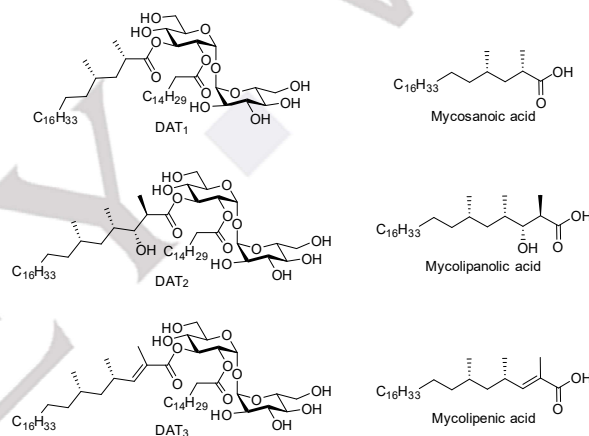
Figure 1. Chemical structures of the mycobacterial diacyl trehaloses DAT₁, DAT₂, DAT₃, and the corresponding fatty acids

The mycobacterial lipids component DAT, which is specifically found in *Mtb* and *M. fortuitum*, is crucial to the pathogenesis and structure of the cell envelope, and promotes bacterial intracellular survival.[10] DAT is capable of modulating host immune responses.[11],[12] Macrophage-inducible C-type lectin (Mincle) is an innate immune receptor, and its activation is a central part of the innate immune response to *Mtb*.[12b] DAT can act both as an antigen for T cells and an activator for the innate Mincle receptor, and small structural differences determine recognition by different parts of the immune system.[13] The use of DAT in serodiagnosis has been postulated.[14] In addition, there is a series of studies on the detection of antibodies against DAT and developing novel vaccine adjuvants from Mincle agonists.[15]

Whereas the structures of DAT₁ and DAT₃ are undisputed, the history of DAT₂ and its acyl substituent mycolipanolic acid is long and winding. Mycolipanolic acid was first described by Coles and Polgar in 1968 after its isolation from a massive culture of *Mtb*.[16] The compound, present as several homologues (e. g. length of

the chain), was coined mycolipanolic acid. The C4 and C6 methyl substituents of the deoxypropionate chain were established to have the *S* (at that time "L") configuration. Later it was shown that all methyl-branched fatty acids produced by the polyketide synthase PKS2, e.g. mycolipanolic acid, mycosanoic acid, mycolipenic acid, phthioceranic acid and hydroxy phthioceranic acid, indeed have an all-*S* configuration in the deoxypropionate chain.[8] Based on IR-spectroscopy, 100 MHz ^{1}H-NMR, and base-induced epimerization of the methyl group at C2 of the corresponding methyl ester, it was postulated that the C2 methyl substituent and the C3 hydroxy group in mycolipanolic acid had an *erythro* relationship. It was also explicitly mentioned that the absolute configuration at C2 and C3 could not be established.

Several years later, in an unrelated study the dimer of a compound closely related to mycolipanolic acid, albeit with a much shorter alkyl chain,[17] was found in a lichen. The compound was called bourgeanic acid (the monomer hemibourgeanic acid, Figure 2). The absolute stereochemistry of a derivative of hemibourgeanic was obtained by crystal structure determination.[18] Hemibourgeanic acid has the 2*S*, 3*S*, 4*R*, 6*R* stereochemistry and is therefore enantiomeric with mycolipanolic acid at the C4 and C6 methyl branches. Over the years, several syntheses of hemibourgeanic acid were disclosed[19] but no attempt was made to relate the spectral data to those of mycolipanolic acid.

In 1996, Wallace and Minnikin prepared mycolipanolic acid as a stereorandom mixture,[20] and wrote, referring to the papers of Coles and Polgar "*The absolute configuration of this mycolipanolic acid has been established as (2S, 3R, 4S, 6S)-3-hydroxy-2,4,6-trimethyltetracosanoic acid. A definite syn relationship has been proven between the C2 methyl and the C3 hydroxy group.*" This clearly was a too bold statement and in addition, the earlier proposed *erythro* relationship would translate into an *anti*-relationship. As the synthesis was stereorandom, the authors could not contribute further to a confirmation of the structure.
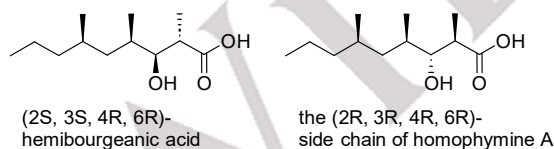

(2S, 3S, 4R, 6R)-
hemibourgeanic acid

the (2R, 3R, 4R, 6R)-
side chain of homophymine A

Figure 2. Hemibourgeanic acid and the side chain of homophymine A

In 2010, our group developed a stereoselective synthesis of mycolipanolic acid, with 2*S*, 3*R*, 4*S*, 6*S* configuration, and its related fatty acids.[21] We compared the NMR data with the numerical data available in literature and the optical rotation with the one reported by Coles and Polgar while the NMR spectra were not available. These seemed in agreement, although the magnitude of the optical rotation of mycolipanolic acid is small (7°) and allows no conclusion in terms of stereochemistry. At that time, we did not have access to natural mycolipanolic acid. In parallel, the group of Zampella disclosed a stereoselective synthesis of

(2*R*, 3*R*, 4*R*, 6*R*)-3-hydroxy-trimethyloctanoic acid (Figure 2) and showed it to be part of the depsipeptide homophymine A.[22] Hemibourgeanic acid and the side chain of homophymine A are therefore epimers at C2 and C3, but comparison of the NMR data showed that these are very similar. In retrospect, the stereochemistry of mycolipanolic acid had not been entirely solved, and comparison of numerical ^{1}H-NMR data present in literature was insufficient. This was underscored by the structure elucidation of (-)-SCH 64874 and hirsutellomycin which possess a 3-hydroxy-2,4-dimethyl hexanoate side chain (so without a methyl substituent at C6). In order to establish the stereochemistry, the authors had to prepare all four diastereomers which showed small but distinct differences in the ^{1}H- and ^{13}C-NMR spectra.[23]

In 1989, the mycolipanolic ester DAT$_2$ was isolated by Daffé *et al..* The compound was initially named sulfolipid-IV because it was described as a 2,3-diacyl-trehalose-2'-sulfate.[16a,24] The structure of this family of acyl trehaloses was corrected to be a group of 2,3-di-O-acyl-trehaloses.[7, 25] DAT$_2$ was shown to contain a mycolipanolic acid residue at the C3 position, but the stereochemistry of mycolipanolic acid was not further studied.[7]

Recently, we reported the first total synthesis of DAT$_1$, DAT$_2$, and DAT$_3$, in order to study the ability of the individual DAT family members to activate Mincle. Upon analysis by HPLC-MS, the retention time and mass spectra of DAT$_1$ and DAT$_3$ coincided with those of the natural material. The retention time of synthetic DAT$_2$, however, showed more than 1 min difference with that of the natural product (available as a mixture), although the mass spectra were virtually indistinguishable.[9] Thus, we suspected that the structure of DAT$_2$ in literature was incorrect.

The results forced us to re-evaluate the stereochemistry of mycolipanolic acid, because other errors in the assigned structure of DAT$_2$ would supposedly have led to differences in the mass spectra. In the biosynthesis of mycolipanolic acid, being a fatty acid-polyketide hybrid, the stereochemistry of methyl-branched C4 and C6 is determined by the enoyl reductase functional domain of PKS2, but the stereochemistry at the C3 and the C2 position is determined by the ketoreductase and the ketoacyl synthase domain, respectively.[26] So, there is no obvious connection between the stereochemistry of C6/C4 and of C2/C3, and there is insufficient information of the structure of PKS2 to predict the stereochemistry of its products.

To clarify this situation, we decided to embark on the synthesis of all four diastereomers of mycolipanolic acid, keeping the stereochemistry at C4 and C6 fixed in the *S* configuration. These stereoisomers would be used to prepare the corresponding stereoisomers of DAT$_2$ in order to explain the difference in HPLC retention time. Finally, we were interested to determine to what extent the stereochemistry of the mycolipanolic acid residue influences the potency of DAT$_2$ to activate Mincle. We have shown previously that the structure of the residue, mycosanoic acid, mycolipanolic acid or mycolipenic acid, significantly influences activation of Mincle.[9] Determination of the influence of stereochemistry would bring this fine-specificity one step further.

It is important to establish the correct structure of DAT$_2$, not just as an intellectual challenge. There is a need for better, in particular faster and low-cost, diagnostics for tuberculosis

RESEARCH ARTICLE

infection. The diacyl trehaloses could play a role in ELISA-type assays, and in order to achieve this, synthetic antigens of well-defined structure are essential. These lipids might also play a role in the development of novel adjuvants, given their activation of the Mincle receptor.

Results and Discussion

Synthesis of four stereoisomeric mycolipanolic acids
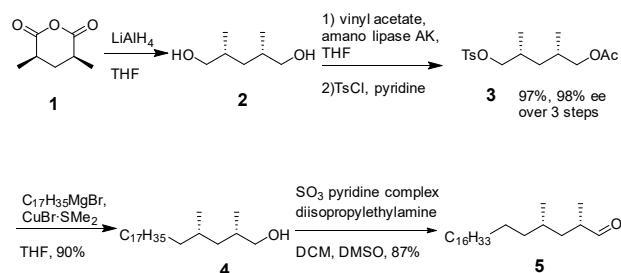
Scheme 1. Synthesis of the common intermediate **5** (THF = tetrahydrofuran, TsCl = p-tolylsulfonyl chloride, DCM = dichloromethane, DMSO = dimethylsulfoxide)

The synthesis of the literature-proposed structure of mycolipanolic acid involved copper-catalyzed asymmetric conjugate addition (Cu-cat. ACA) and an Evans' aldol reaction to introduce the stereocenters.[9,21] For the current study, we developed a considerably more efficient synthesis based on the work of Breit *et al.*,[27] (Scheme 1) to prepare aldehyde **5** by starting from *meso*-anhydride **1**. Reduction of **1** with LiAlH$_4$ provided the desired *meso*-diol **2**, and a subsequent enzymatic desymmetrisation with amano lipase AK provided the mono-acetate in excellent selectivity (98% *ee*) and near quantitative yield. The heptadecanyl chain was installed by a modified copper catalysed cross-coupling reaction[28] with tosylate **3**, with concomitant removal of the acetate group. This was followed by a Parikh–Doering oxidation, to give aldehyde **5** in high yield. In this way, the synthesis of **5** was shortened and provided a considerably higher yield.
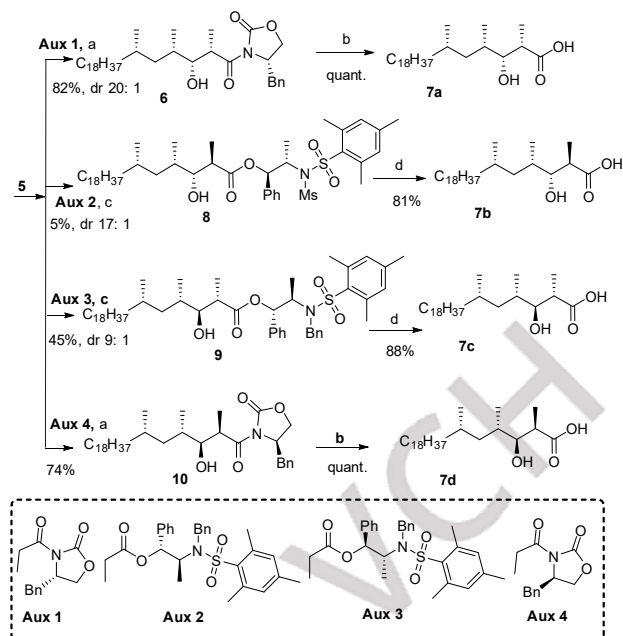
Scheme 2. Asymmetric synthesis of 4 stereoisomeric mycolipanolic acids. Stereocenters of the hydroxymethyl units were stereoselectively installed through Evans and Abiko-Masamune aldol reactions. Reagents and conditions: a) Bu$_2$BOTf, Et$_3$N, DCM, −78 °C to rt (room temperature); b) LiOH, H$_2$O$_2$, THF, H$_2$O; c) Cy$_2$BOTf, Et$_3$N, DCM, −78 °C to −10 °C; d) TBAOH, THF.

From **5**, all four stereoisomeric mycolipanolic acids could be synthesized in a limited number of steps (Scheme 2). To install the two remaining stereocenters, Evans aldol reactions[29] were performed for the preparation of the mycolipanolic acids with a *syn*-hydroxymethyl unit. Abiko-Masamune aldol reactions[30] were applied to prepare the *anti*-hydroxymethyl units. The different stereochemistry of the chiral auxiliaries led to the different relative configuration of the β-OH and α-methyl groups, and provided the aldol products **8-10** in an excellent diastereomeric ratio (dr). As the solubility of the substrates significantly decreased at lower temperature, the procedure of the aldol reactions leading to **9** and **10** had to be carefully optimized but finally were significantly improved to provide 45% and 74% yield, respectively. The 5% yield of **8** couldn't be improved, however. It was attempted to prepare **7b** through a magnesium chloride mediated Evans-type anti-aldol reaction[31] and alternatively a Mukaiyama aldol reaction,[32] but neither of these worked. The aldol products **8, 9** and **10** were finally hydrolyzed by lithium hydroxide (LiOH) or tetrabutylammonium hydroxide (TBAOH) to give diastereomers of mycolipanolic acid (**7b-d**).

Preparation of suitably protected mycolipanolic acids

In the initial total synthesis of DAT$_2$, now called DAT$_2$-a, we reported a direct esterification of mycolipanolic acid **7a** and suitably protected 2-palmitoyl trehalose **13**.[9] The hydroxy group in **7a** remained unprotected. However, this approach did not work for the esterification of the other stereo-isomers of mycolipanolic

3

RESEARCH ARTICLE

acid. Neither Yamaguchi esterification nor Shiina esterification of the other three acids **7b-d** with protected 2-palmitoyl trehalose **13** or **16** worked well. One of the problems turned out to be that the β-OH group in **7b-d** took part in intra- and intermolecular esterification reactions, aggravated by the slow esterification reaction of the 3-OH of the trehalose core. This hydroxy group is difficult to access, in particular for the voluminous fatty acids. So, we decided to prepare the mycolipanolic acids with the β-OH group protected as a tetrahydropyran (THP) acetal or a benzyl (Bn) ether, which are not too bulky (Scheme 3). The protection procedures were carefully optimized to avoid epimerization of the α methyl group or elimination of the β-OH group.
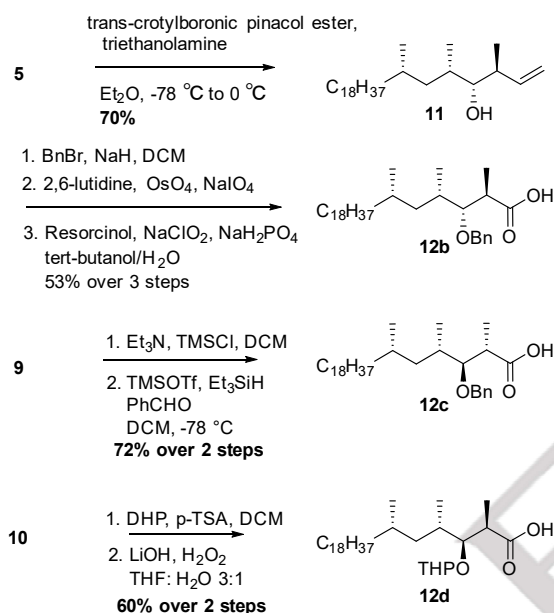

Scheme 3. Preparation of β-OH protected mycolipanolic acids **12b-12d** (Et₂O = diethylether, BnBr = benzyl bromide, TMSCl = trimethylsilyl chloride, TMSOTf = trimethylsilyl trifluoromethanesulfonate, DHP = dihydropyran, p-TSA = p-tolylsulfonic acid)

Abiko-Masamune aldol reaction gave a very low yield of **8** (5%), and THP protection[23] and benzylation[33] of **8** either gave very low yield or didn't proceed at all. Therefore we changed synthesis strategy for this particular isomer, and applied a Brown crotylation reaction with an E-crotylboronic ester to produce 2,3-anti-3,4-syn diastereomer **11**.[34] This provided a much higher yield and good dr. After benzylation, the double bond was cleaved successfully through Lemieux–Johnson oxidation with OsO₄/NaIO₄ and was further oxidized to the corresponding acid **12b** by Pinnick oxidation.

Silylation of **9** was followed by benzylation and removal of the auxiliary[35] to give the acid **12c.** The β-OH group in **10** was protected as a THP acetal and gave enantiopure acid **12d** after hydrolysis.

Esterification of mycolipanolic acid isomers with trehalose
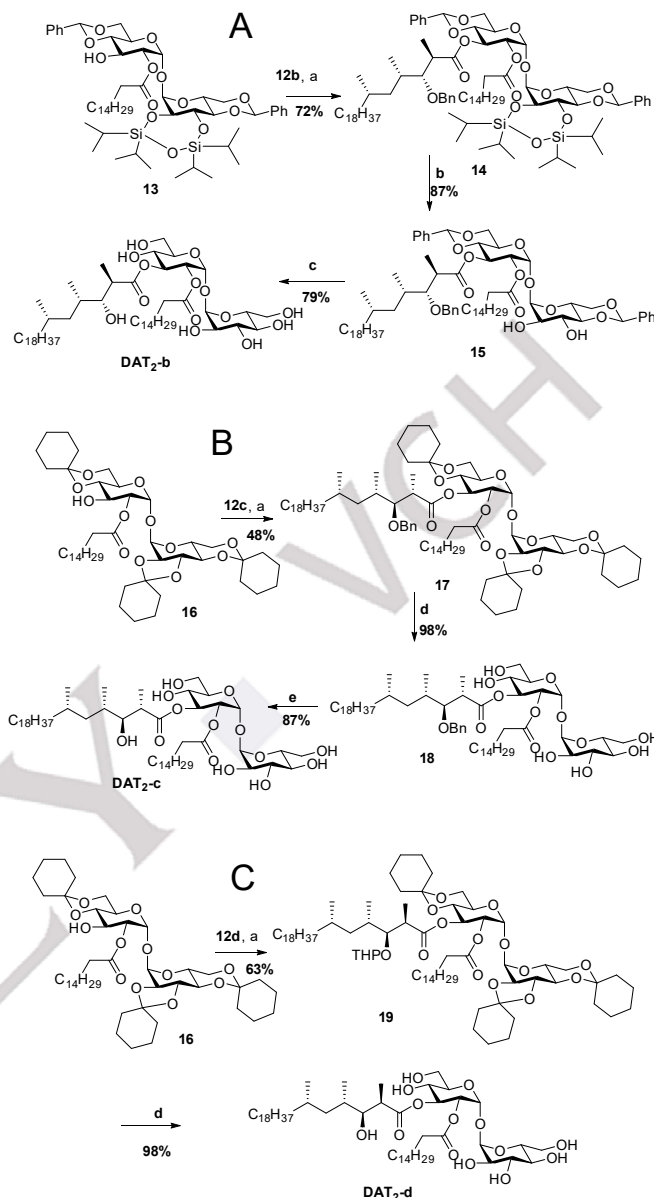
Scheme 4. Completion of the synthesis of the diastereomers of DAT₂. Suitably protected isomeric mycolipanolic acids were incorporated into protected palmitoyl trehaloses by Shiina esterification and gave the stereoisomers of DAT₂ after deprotection. Reagents and conditions: a) Et₃N, Dimethylamino pyridine (DMAP), MNBA, DCM; b) Tetrabutylammonium fluoride (TBAF), AcOH; c) Pd/C, Pd(OH)₂, H₂, THF; d) AcOH, DCM, 50 °C; e) H₂, Pd/C, EtOAc, EtOH.

Suitably protected trehalose **13** was obtained starting from α,α-trehalose through a desymmetrization strategy previously applied in the synthesis of trehalose-based sulfoglycolipids[36], whereas trehalose **16** has been used to prepare sulfolipid SL-1 and Ac₂SGL analogues.[37]

With the stereochemically pure acids **12b-d** in hand, the esterification of the palmitoylated trehaloses **13** and **16** was accomplished by following the Shiina procedure with the reagent 2-methyl-6-nitrobenzoic anhydride (MNBA). The reactions had to

RESEARCH ARTICLE

be carefully optimized by limiting the amount of base and the time for acid activation, and applying a concentration of the acid above 0.1 M. With the optimized conditions, esterification of **12b** with **13** produced **14** in 72% yield (Scheme 4A). Removal of the silyl protecting group followed by palladium-catalyzed hydrogenolysis[36a] provided DAT$_2$-b in good yield.

The esterification of **12c** and **12d** with **13** did not proceed, presumably because of the different conformations of these mycolipanolic acid stereoisomers. Better results, although for unknown reasons, were obtained with **16** to prepare **17** (Scheme 4B) and **19** (Scheme 4C) via Shiina esterification in 48% and 63% yield, respectively. The cyclohexylidene acetals of **17** were removed in aqueous AcOH (80%) at reflux and followed by hydrogenolysis to afford DAT$_2$-c. Removal of the cyclohexylidene and the THP acetals in **19** was carried out in one step to give DAT$_2$-d in 98% yield. This completed the synthesis of four stereoisomers of mycolipanolic acid and the corresponding DAT$_2$ isomers. The compounds were characterized by ^{1}H-NMR, ^{13}C-NMR and HRMS.

NMR analysis of synthetic and natural DAT$_2$

With four stereoisomers of DAT$_2$ in hand, we sought to determine whether one of the synthesized glycolipids would match the structure of natural DAT$_2$ (as a constituent of the diacyl trehalose family of compounds) present in pathogenic *Mtb*. To meticulously compare the ^{1}H NMR spectra of the four synthetic diastereomers with that of natural DAT (here called DAT$_n$), a purification of "DAT$_n$" was undertaken using the reference laboratory strain H37Rv. As expected, the ^{1}H NMR signals of the four stereoisomeric DAT$_2$ were mutually very similar and also similar to those of the authentic natural product. Small deviations, however, were observed in the chemical shifts of the hydrogens attached to C8–C15 and to C2, and C3 of the trehalose core. Hoye's differential chemical shift analysis[23, 38] was used to depict the differences between each of the synthetic isomers and natural DAT$_n$ and to determine which synthetic isomer would match closest to the natural sample (Figure 3, see also the ESI†). With respect to these signals, the ^{1}H NMR chemical shifts of DAT$_2$-b rather than the other 3 DAT$_2$ isomers were in excellent agreement with those present in the natural sample. Although the differences were small, these results indicated that the stereochemistry of the 3-hydroxy-2-methyl unit of DAT$_2$ corresponds to that of DAT$_2$-b. We were however not fully convinced, if only because it was surprising that such a small difference in stereochemistry between the originally prepared DAT$_2$-a and DAT$_2$-b/natural DAT$_2$ would lead to such a significant difference in retention time on HPLC, observed earlier.
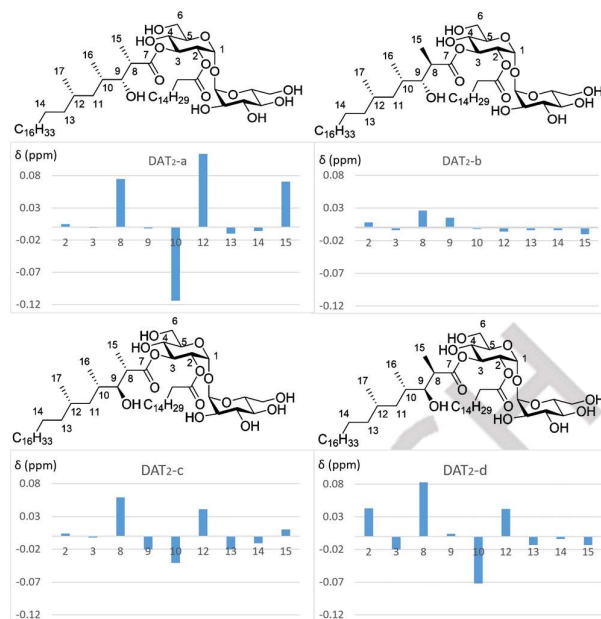
Figure 3. Comparison of the ^{1}H NMR (600 MHz, CDCl$_3$/CD$_3$OD 4/1) spectra according to Hoye. The columns represent differences between each of the synthetic isomers and the natural sample. DAT$_2$-b matches best with the signals from the natural sample (DAT$_n$).

Determination of natural DAT$_2$ by HPLC-MS/MS analysis

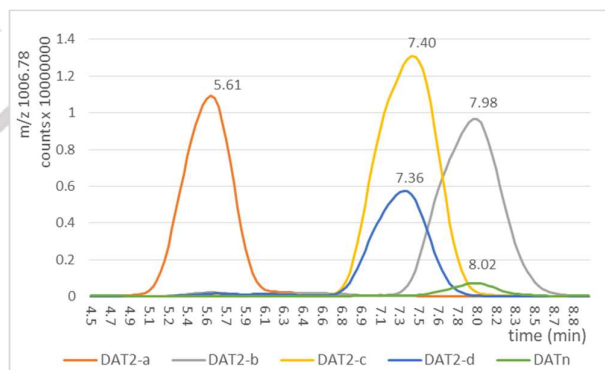**Figure 4.** Comparison of the four synthetic isomers of DAT$_2$ with the natural isolate DAT$_n$ by HPLC.

The next comparison of the synthetic DAT$_2$ isomers to natural DAT$_n$ was by means of HPLC-MS/MS. Careful optimization of the solvent system used for reverse-phase chromatography on a C-18 column revealed that large differences in retention time could be obtained with a gradient of a 28:12:51:9 methanol: water: 1-propanol: cyclohexane mixture as the polar component and a 85: 15 1-propanol: cyclohexane mixture as the nonpolar component. Both solvents contained 2.0 mM ammonium formate. The extracted-ion chromatograms (Figure 4) show a chromatographic match for DAT$_2$-b (the grey line) and the natural isomer (the green line) while the other three synthetic DAT$_2$ isomers show deviating

RESEARCH ARTICLE

retention times. In particular DAT2-a, possessing the structure proposed in literature, and initially prepared, showed a huge retention time difference with the natural isomer.

The mass of the natural isomer matched to those of the four DAT_2 isomers, as expected. Collision-Induced Dissociation fragmentation of the natural and synthetic DAT_2 isomers yielded interpretable fragmentations (see the ESI†) that supported the general structure and connectivity. Thus, we concluded that synthetic DAT_2-b is identical to the natural isomer produced by the laboratory H37Rv strain, which matches with the result of the NMR comparison.

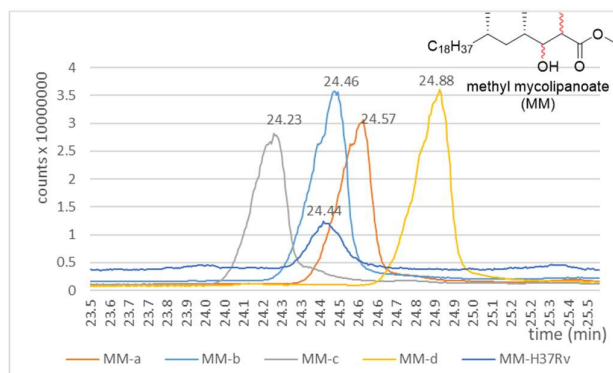## Comparison of the mycolipanolic acids by GC-MS

Figure 5. Comparison of retention times of the 4 isomeric methyl mycolipanoates with that of methyl mycolipanoate obtained from the natural isolate (*Mtb*-H37Rv) by GC chromatography on a polymethyl siloxane column.

To unequivocally determine the stereochemistry of natural mycolipanolic acid, a GC–MS analysis was carried out to confirm that acid **7b** from DAT2-b had the stereochemistry of natural mycolipanolic acid. After basic hydrolysis of a sample of the natural DAT_n extract, the four synthetic isomers of mycolipanolic acid, and the natural mycolipanolic acid were each converted into their corresponding methyl esters by treatment with trimethylsilyl diazomethane. GC–MS analysis (Figure 5) showed a chromatographic match of methyl mycolipanoate-b with the methyl mycolipanoate of H37Rv, which is fully consistent with the results of the NMR comparison and the HPLC-CID analysis. The retention time of the other isomers was sufficiently different. Peaks were identified by comparing the mass spectra ion fragments (see the ESI†).

Taken together, we established that the configuration of natural DAT_2 is represented by the structure **DAT2-b** and that of mycolipanolic acid by **7b**. The stereochemistry of the α-methyl group was incorrectly assigned in the literature.[20] This means that a number of studies provides an incorrect structure of mycolipanolic acid and DAT_2.[9, 20-21, 39]
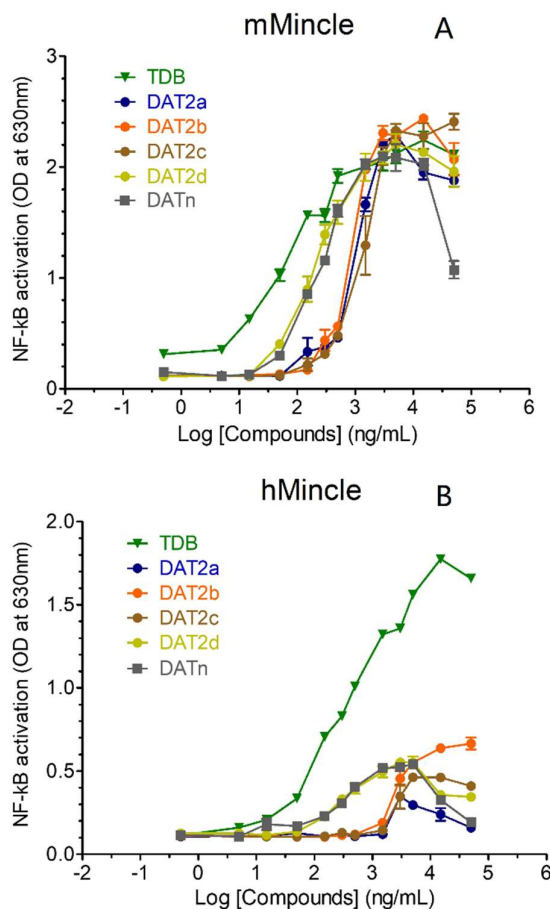
Mincle activation by synthetic stereoisomers of DAT2

Figure 6. Mincle activation by synthetic DAT_2 stereoisomers. HEK cells expressing murine (A) or human (B) Mincle and a NF-κB-inducible reporter system were stimulated with the indicated amount of DAT_2 diastereomers, Mtb H37Rv natural DAT (DATn), or TDB. After 24 h, NF-κB activation was determined by reading OD at 630 nm.

We next investigated the capacity of the different stereoisomers, in comparison to the natural DAT family (DATn) from *M. tuberculosis* H37Rv, to induce Mincle activation using HEK cells expressing human Mincle (HEK-hMincle) or murine Mincle (HEK-mMincle), and an NF-κB-inducible reporter system (secreted alkaline phosphatase) (Figure 6). DATn induced NF-κB activation in a dose-dependent manner in both reporter cell lines, although as previously reported it was less potent (i.e. had a higher EC50) than trehalose-6,6'-dibehenate (TDB), a synthetic analog of TDM in which simpler fatty acids replace the complex mycolic acids[9]. DATn showed an efficacy (i.e. a maximum response) similar to TDB in HEK-mMincle cells, but lower than TDB in HEK-hMincle cells. Of note, at high concentrations (>5 μg/ml), DATn showed toxicity.

Regarding the DAT_2 isomers, DAT_2-d showed the same potency as DATn in both reporter cell lines, whereas the other three DAT_2 isomers were less potent. All the DAT_2 isomers exhibited a similar efficacy, with the exception of DAT_2-b, which was systematically more effective in HEK-hMincle cells.

Conclusion

The asymmetric synthesis of four stereoisomers of mycolipanolic acid was achieved with full stereocontrol. It turned out that, with already two stereogenic methyl-substituents present in the substrate, different established methods for the preparation of aldol-type motives were required to prepare the various stereoisomers. A one-fits-all approach for the different syn- and anti-isomers could not be identified. For the syn-aldol products, the use of Evans chiral auxiliaries gave superior results. For the anti-aldol products, one was conveniently prepared using Abiko-Masamune's chiral auxiliary, the other one using Brown's crotylation reaction.

In this study, we accomplished the total synthesis of three diastereomers of mycobacterial DAT_2 and mycolipanolic acid. Together with the already prepared isomer, four stereoisomers of DAT_2 were compared with a natural sample from *Mtb* H37Rv, by NMR comparison and liquid chromatography-collision induced dissociation analysis. The corresponding synthetic and natural methyl mycolipanoates were analyzed through gas chromatography mass spectroscopy. Accordingly, the correct structure of natural DAT_2 and the absolute configuration of natural mycolipanolic acid were determined to be DAT_2-b and **7b** ($2R$, $3R$, $4S$, $6S$). This means that the structure of mycolipanolic acid and of DAT_2 has been re-assigned.

Immunological activity tests on mouse Mincle and human Mincle show that all four stereoisomers are potent activators, and natural DAT_2 / DAT_2-b has the highest efficacy in HEK-hMincle cells. The previously proposed diastereomer DAT_2-a is the least potent. DAT_2 is highly antigenic and can potentially be applied in the development of serodiagnostic tests and vaccine adjuvants.

We found that small changes in the structure of the branched acyl chain in DAT result in large differences in recognition by Mincle. Considering the varying ratios of DAT_1, DAT_2, and DAT_3 in natural "DAT", and the possibility of contamination with minute amounts of oligosaccharides or peptides, means that the use of natural extracts of DAT to study immune response is not reliable. The availability of the DAT constituents by chemical synthesis enables the development of a serodiagnostic test and the use of these compounds as vaccine adjuvants.

Supporting Information

The authors have cited additional references within the Supporting Information[9, 21, 27, 30e, 36a, 37, 40]

Acknowledgements

Keywords: Di-*O*-acyl trehalose •organic synthesis • stereochemistry • structural revision • antigenicity

[1]   a) F. Ntoumi, J. B. Nachega, E. Aklillu, J. Chakaya, I. Felker, F. Amanullah, D. Yeboah-Manu, K. G. Castro, A. Zumla, *Lancet Infect. Dis.* **2022**, *22*, 442-444; b, News Room of the World Health Organization, https://www.who.int/news-room/fact-sheets/detail/tuberculosis, **2022**.

[2]   S. Ehrt, D. Schnappinger, *Cell. Microbiol.* **2009**, *11*, 1170-1178.

[3]   M. Jackson, *Cold Spring Harb. Perspect. Med.* **2014**, *4*, 1-12.

[4]   S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, B. G. Barrell, *Nature* **1998**, *393*, 537-544.

[5]   O. Neyrolles, C. Guilhot, *Tuberculosis* **2011**, *91*, 187-195.

[6]   a) M. Jackson, *Cold Spring Harbor Perspectives in Medicine* **2014**, *4*, 25-30; b) D. E. Minnikin, P. J. Brennan, Lipids of Clinically Significant Mycobacteria. In: Goldfine, H. (ed.) Health Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids. Handbook of Hydrocarbon and Lipid Microbiology. Springer, **2020**.

[7]   G. S. Besra, R. C. Bolton, M. R. McNeil, M. Ridell, K. E. Simpson, J. Glushka, H. Van Halbeek, P. J. Brennan, D. E. Minnikin, *Biochem.* **1992**, *31*, 9832-9837.

[8]   D. Minnikin, G. Dobson, D. Sesardic, M. Ridell, *Microbiology* **1985**, *131*, 1369-1374.

[9]   M. Holzheimer, J. F. Reijneveld, A. K. Ramnarine, G. Misiakos, D. C. Young, E. Ishikawa, T. Y. Cheng, S. Yamasaki, D. B. Moody, I. Van Rhijn, *ACS Chem. Biol.* **2020**, *15*, 1835-1841.

[10]   R. Bailo, A. Bhatt, J. A. Aínsa, *Biochem. Pharmacol.* **2015**, *96*, 159-167.

[11]   K. S. Lee, V. S. Dubey, P. E. Kolattukudy, C. H. Song, A. R. Shin, S. B. Jung, C. S. Yang, S. Y. Kim, E. K. Jo, J. K. Park, *FEMS Microbiol. Lett.* **2007**, *267*, 121-128.

[12]   a) M. Correia-Neves, J. Nigou, Z. Mousavian, C. Sundling, G. Källenius, *Front. immunol.* **2022**, *13*, 1-20; b) S. J. Williams, *Front. immunol.* **2017**, *8*, 1-12; c) J. P. Palma-Nicolás, R. Hernández-Pando, E. Segura, M. J. Ibarra-Sánchez, I. Estrada-García, A. Zentella-Dehesa, L. M. López-Marín, *Immunobiology* **2010**, *215*, 143-152.

[13]   J. F. Reijneveld, M. Holzheimer, D. C. Young, K. Lopez, S. Suliman, J. Jimenez, R. Calderon, L. Lecca, M. B. Murray, E. Ishikawa, *Sci. Rep.* **2021**, *11*, 1-10.

[14]   M. T. Tórtola, M. A. Lanéelle, N. Martín-Casabona, *Clin. diagn. lab. immunol.* **1996**, *3*, 563-566.

[15]   H. Schoenen, B. Bodendorfer, K. Hitchens, S. Manzanero, K. Werninghaus, F. Nimmerjahn, E. M. Agger, S. Stenger, P. Andersen, J. Ruland, *J. Immunol.* **2010**, *184*, 2756-2760.

[16]   L. Coles, N. Polgar, *J. Chem. Soc. C* **1968**, 1412-1419.

[17]   B. Bodo, P. Hebrard, L. Molko, D. Molko, *Tetrahedron* Lett. **1973**, 14, 1631-1634.

[18]   B. Bodo, W. Trowitzsch-Kienast, D. Schomburg, *Tetrahedron Lett.* **1986**, *27*, 847-848.

[19]   a) J. D. White, A. T. Johnson, *J. Org. Chem.* **1990**, *55*, 5938-5940; b) J. D. White, A. T. Johnson, *J. Org. Chem.* **1994**, *59*, 3347-3358; c) J. S. Yadav, K. R. Rao, K. Ravindar, B. S. Reddy, *Eur. J. Org. Chem.* **2011**, 58-61.

[20]   P. A. Wallace, D. E. Minnikin, *Chem. Phys. Lipids* **1996**, *83*, 1-8.

RESEARCH ARTICLE

[21]   B. ter Horst, J. van Wermeskerken, B. L. Feringa, A. J. Minnaard, *Eur. J. Org. Chem.* **2010**, 38–41.

[22]   F. Bellotta, M. V. D'Auria, V. Sepe, A. Zampella, *Tetrahedron* **2009**, *65*, 3659-3663.

[23]   H. Tokuyama, K. Yamada, H. Fujiwara, J. Sakata, K. Okano, M. Sappan, M. Isaka, *J. Org. Chem.* **2017**, *82*, 353-371.

[24]   M. Daff, F. Papa, A. Laszlo, H. David, *Microbiology* **1989**, *135*, 2759-2766.

[25]   A. Lemassu, M. A. Lanéelle, M. Daffé, *FEMS Microbiol. Lett.* **1991**, *78*, 171-175.

[26]   V. T. Natarajan, D. Mohanty, R. S. Gokhale, *Biosynthesis of mycobacterial lipids by multifunctional polyketide synthases*, in: The mycobacterial Cell Envelope, M. Daffé, J.-M. Reyrat, Wiley, **2008**.

[27]   Y. Schmidt, K. Lehr, U. Breuninger, G. Brand, T. Reiss, B. Breit, *J. Org. Chem.* **2010**, *75*, 4424-4433.

[28]   D. H. Burns, J. D. Miller, H. K. Chan, M. O. Delaney, *J. Am. Chem. Soc.* **1997**, *119*, 2125-2133.

[29]   a) D. A. Evans, D. L. Rieger, M. T. Bilodeau, F. Urpi, *J. Am. Chem. Soc.* **1991**, *113*, 1047-1049; b) M. M. Heravi, V. Zadsirjan, *Tetrahedron: Asymmetry* **2013**, *24*, 1149-1188; c) D. A. Evans, J. Bartroli, T. Shih, *J. Am. Chem. Soc.* **1981**, *103*, 2127-2129.

[30]   a) A. Abiko, J. F. Liu, S. Masamune, *J. Am. Chem. Soc.* **1997**, *119*, 2586-2587; b) N. Tahiri, P. Fodran, D. Jayaraman, J. Buter, M. D. Witte, T. A. Ocampo, D. B. Moody, I. Van Rhijn, A. J. Minnaard, *Angew. Chem. Int. Ed.* **2020**, *59*, 7555-7560; c) C. F. Weise, M. C. Pischl, A. Pfaltz, C. Schneider, *J. Org. Chem.* **2012**, *77*, 1477-1488; d) F. Kleinbeck, E. M. Carreira, *Angew. Chem. Int. Ed.* **2009**, *48*, 578-581; e) T. Inoue, J. F. Liu, D. C. Buske, A. Abiko, *J. Org. Chem.* **2002**, *67*, 5250-5256.

[31]   D. A. Evans, J. S. Tetrow, J. T. Shaw, C. W. Downey, *J. Am. Chem. Soc.* **2002**, *124*, 392–393.

[32]   Y. Sun, Y. Ding, D. Li, R. Zhou, X. Su, J. Yang, X. Guo, C. Chong, J. Wang, W. Zhang, *Angew. Chem. Int. Ed.* **2017**, *129*, 14819-14823.

[33]   H. Liu, Y. Zhang, R. Wei, G. Andolina, X. Li, *J. Am. Chem. Soc.* **2017**, *139*, 13420-13428.

[34]   R. W. Hoffmann, H. J. Zeiss, *J. Org. Chem.* **1981**, *46*, 1309-1314.

[35]   M. Nishizawa, H. Yamamoto, H. Imagawa, V. Barbier-Chassefière, E. Petit, I. Azuma, D. Papy-Garcia, *J. Org. Chem.* **2007**, *72*, 1627-1633.

[36]   a) J. Guiard, A. Collmann, M. Gilleron, L. Mori, G. De Libero, J. Prandi, G. Puzo, *Angew. Chem.* **2008**, *120*, 9880-9884; b) C. H. Wu, C. C. Wang, *Org. Biomol. Chem.* **2014**, *12*, 5558-5562.

[37]   V. A. Sarpe, S. S. Kulkarni, *Org. Lett.* **2014**, *16*, 5732-5735.

[38]   a) E. Fleury, M. I. Lannou, O. Bistri, F. Sautel, G. Massiot, A. Pancrazi, J. Ardisson, *Eur. J. Org. Chem.* **2009**, 4992–5001; b) T. R. Hoye, J. C. Suhadolnik, *J. Am. Chem. Soc.* **1987**, *109*, 4402-4403.

[39]   a) M. Holzheimer, J. Buter, A. J. Minnaard, *Chem. Rev.* **2021**, *121*, 9554-9643; b) W. Wu, W. Cao, L. Hu, Z. Su, X. Liu, X. Feng, *Chem. Sci.* **2019**, *10*, 7003-7008.

[40]   a) L. C. Dias, E. C. Polo, *J. Org. Chem.* **2017**, *82*, 4072-4112; b) M. Das, Y. Du, J. S. Mortensen, M. Ramos, L. Ghani, H. J. Lee, H. E. Bae, B. Byrne, L. Guan, C. J. Loland, *Org. Biomol. Chem.* **2019**, *17*, 3249-3257; c) A. Decout, S. Silva-Gomes, D. Drocourt, S. Barbe, I. André, F. J. Cueto, T. Lioux, D. Sancho, E. Pérouzel, A. Vercellone, *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 2675-2680.

RESEARCH ARTICLE

Graphic abstract

The asymmetric synthesis of 4 stereoisomers of DAT$_2$, an important antigen of *M. tuberculosis*, showed that the commonly accepted structure of DAT$_2$ and its acyl substituent mycolipanolic acid had to be corrected. DAT$_2$ is an effective ligand activating Mincle, an important pattern recognition receptor of the immune system.

Proposed structure   Corrected structure