Asymmetric Total Synthesis of Mycobacterial Diacyl Trehaloses Demonstrates a Role for Lipid Structure in Immunogenicity

Mira Holzheimer, Josephine F. Reijneveld, Alexandrea K. Ramnarine, Georgios Misiakos, David C. Young, Eri Ishikawa, Tan-Yun Cheng, Sho Yamasaki, D. Branch Moody, Ildiko Van Rhijn,* and Adriaan J. Minnaard*

Cite This: ACS Chem. Biol. 2020, 15, 1835−1841

ABSTRACT: The first asymmetric total synthesis of three structures proposed for mycobacterial diacyl trehaloses, DAT1, DAT2, and DAT3, is reported. The presence of two of these glycolipids, DAT1 and DAT3, within different strains of pathogenic M. tuberculosis was confirmed, and it was shown that their abundance varies significantly. In mass spectrometry, synthetic DAT2 possessed almost identical fragmentation patterns to presumptive DAT2 from Mycobacterium tuberculosis H37Rv, but did not coelute by HPLC, raising questions as the precise relationship of the synthetic and natural materials. The synthetic DATs were examined as agonists for signaling by the C-type lectin, Mincle. The small differences in the chemical structure of the lipidic parts of DAT1, DAT2, and DAT3 led to drastic differences of Mincle binding and activation, with DAT3 showing similar potency as the known Mincle agonist trehalose dimycolate (TDM). In the future, DAT3 could serve as basis for the design of vaccine adjuvants with simplified chemical structure.

INTRODUCTION

Mycobacterium tuberculosis (Mt), which is the causative agent of the disease tuberculosis (Tb), is responsible for the largest number of deaths worldwide by a single pathogen, killing an estimated 1.3 million people annually. The ability of Mt to survive and persist in the host is estimated to result in billions of latently infected individuals worldwide, with a high incidence of undiagnosed cases. After infection of macrophages, Mt is able to survive and replicate in host phagosomes, while withstanding the hostile acidic environment. The mycobacterial cell envelope is one factor that contributes to the resilience of Mt within host cells. It is a multilayered barrier, composed of many complex lipids, glycolipids, and glycoproteins, many of which are unique to Mt. In the last decades, it has been shown that many of these cell wall components have antigenic properties and/or possess immunomodulatory functions. One class of these mycobacterial cell wall components, which consists of diacylated and polyacylated trehaloses, is suggested to be located on the outer part of the mycobacterial cell wall. These trehalose-based glycolipids are esterified with palmitic or stearic acid at the 2- and 2' -position, as well as with the Mt-specific multimethyl-branched acyl residues phthioceranic acid, hydroxyphtioceranic acid, mycosanoic acid, mycolipanolic acid, and mycolipenic acid. Important examples are Ac2SGL,7−9 Sulfolipid-1,10−12 trehalose monomycolate and dimycolate,13,14 diacyl trehaloses (DAT),15−17 and pentaacyl trehaloses (PAT).15,18 Because of the chemical diversity of DAT and the potential for contamination of even small amounts of bioactive molecular variants, testing natural DAT compounds on cells for immune response is not reliable. To establish the molecular structure of these compounds and enable further biological studies, several of the compounds have been the target of total synthesis. In DAT and PAT, both of which have escaped total synthesis until now, the trehalose core is acylated with the methyl-branched fatty acids mycosanoic acid, mycolipanolic acid, and mycolipenic acid (see Figure 1).17,18

Received: January 14, 2020
Accepted: April 15, 2020
Published: April 15, 2020

https://dx.doi.org/10.1021/acschembio.0c00030
ACS Chem. Biol. 2020, 15, 1835−1841
DAT was first isolated in 1989 by Daffé et al. and was initially named SL-IV (Sulfolipid-IV), since the structure was first assigned as 2,3-diacyl-trehalose-2'-sulfate. The structure of this family of acyl trehaloses was eventually revisited and corrected to be 2,3-diacyl-trehalose, and depending on the nature of the 3-O-acyl group, were termed DAT1, DAT2, or DAT3 (see Figure 1A). In the following reports, these compounds were often referenced as just DAT, presenting a family of mycobacterial glycolipids rather than defined molecular structures. Many studies have asserted the antigenic properties of DAT glycolipids by ELISA, but these were tested mainly as mixtures rather than pure compounds. It was demonstrated multiple times that anti-DAT antibodies are present in blood sera of Tb patients but not of healthy controls. This raised great interest in using DAT for the detection and diagnosis of Tb in patients. The reports utilizing ELISA for the detection of anti-DAT antibodies, however, showed a huge variation in sensitivity and specificity, depending on assay design.

In recent years, research has focused on elucidating the biosynthesis of DAT and unravelling its effect on the immune system. It was shown that DAT partially inhibits the proliferation of murine T-cells, suggesting a role in immunosuppression and T-cell hyporesponsiveness associated with Tb. Mtb mutants incapable of synthesizing mycolipenic acid, and therefore deficient in DAT and PAT, show aggregation in liquid culture, because of defects in capsule attachment, indicating that one of the functions of DAT and PAT is anchoring the hydrophilic capsule to the hydrophobic mycolic acid layer of the mycobacterial cell envelope. However, in aerosol infection mouse models using DAT/PAT-deficient mutants, there were no observed differences in growth, compared to wild-type compounds, suggesting that DAT/PAT itself is not necessary for Mtb survival.

Recently, Mtb cell wall components—such as trehalose dimycolate (TDM, also known as cord factor)—have been identified as high-affinity ligands for macrophage-inducible C-type lectin (Mincle). The activation of Mincle results in downstream expression of cytokines, chemokines, and growth factors and leads to recruitment of inflammatory cells to the site of activation as a central part of the innate immune response to Mtb. Several other Mtb cell wall glycolipids have been identified as Mincle activators, and there is growing interest in using these Mincle ligands for the development of novel vaccine adjuvants.

In 2017, it was demonstrated that a DAT-containing extract from Mtb also activated Mincle. We realized that, apart from minute amounts of contaminants in natural isolates that can influence the results, the activation of Mincle could very well be dependent on the precise structure of the DAT. Therefore, we sought to synthesize three different DATs with precisely defined molecular structure and stereochemistry to study their Mincle activating properties and to assess the influence of the acyl substituents on Mincle binding. Furthermore, we aimed to confirm the presence of these three DATs in different strains of Mtb, including clinical isolates.

Scheme 1. Asymmetric Synthesis of Mycosanic Acid (2), Mycolipanolic Acid (3), and Mycolipenic Acid (4)
RESULTS AND DISCUSSION

Synthesis. DAT₁, DAT₂, and DAT₃ differ in their chiral acyl group esterified with the 3-OH of the trehalose core. Therefore, our synthesis plan involved the preparation of suitably protected 2-palmitoyl trehalose 1 and the three mycobacterial lipids 2, 3, and 4 as key intermediates necessary to construct the target diacyl trehaloses. Trehalose 1 could be obtained starting from α,α-trehalose by a desymmetrization approach previously applied in the synthesis of trehalose-based sulfoglycolipids. The mycobacterial lipids, on the other hand, can be traced back to the common precursor 5 (Figure 1B). The synthesis of mycolipanic and mycolipenic acid was previously reported by us and involves copper-catalyzed asymmetric conjugate addition (Cu-cat. ACA) and an Evans' aldol reaction to introduce the stereocenters. We sought to improve the current synthetic procedures to arrive at an efficient, high-yielding total synthesis.

The synthesis of the chiral enantiopure lipids 2, 3, and 4 (see Scheme 1) commenced with Cu-cat. ACA of methylmagnesium chloride in the presence copper(I) to install the linear alkyl tail of Grignard cross-coupling in excellent yield and diastereomeric ratio (dr). Removal of the silyl protecting group of the mycobacterial lipids, on the other hand, can be traced back to the common precursor 5 (Figure 1B). The synthesis of mycolipanic and mycolipenic acid was previously reported by us and involves copper-catalyzed asymmetric conjugate addition (Cu-cat. ACA) and an Evans’ aldol reaction to introduce the stereocenters. We sought to improve the current synthetic procedures to arrive at an efficient, high-yielding total synthesis.

The synthesis of the chiral enantiopure lipids 2, 3, and 4 (see Scheme 1) commenced with Cu-cat. ACA of methylmagnesium bromide to α,β-unsaturated thioester 5 giving 6 in 81% yield and 98% enantiomeric excess (ee). Reduction to the corresponding aldehyde, followed by Horner–Wadsworth–Emmons reaction, produced another α,β-unsaturated thioester 7. The second methyl stereocenter was again introduced by Cu-cat. ACA in excellent yield and diastereomeric ratio (dr). Double DIBAL-H reduction, followed by tosylation, gave 9 in 88% yield over three steps. Tosylate 9 was subjected to a Grignard cross-coupling in the presence copper(I) to install the linear alkyl tail of 10. Removal of the silyl protecting group, followed by Dess-Martin oxidation, gave aldehyde 11 in an excellent yield of 97% over two steps. From 11, all three mycobacterial lipids could be synthesized in a limited number of steps. Mycosanoic acid 2 was obtained in 92% yield after Pinnick oxidation of aldehyde 11. Mycolipenic acid 4 was prepared by first subjecting 11 to a Wittig reaction, followed by alkaline ester hydrolysis. To install the two remaining stereocenters present in 3, an Evans’ aldol reaction was performed, giving 13 in good yield and excellent dr. The aldol product 13 was then hydrolyzed to give mycolipanic acid 3. Compared to the previous syntheses of 3 and 4, the yields could be significantly improved by careful optimization of the reactions. For mycosanoic acid, mycolipanic acid, and mycolipenic acid, excellent overall yields were obtained with 53% over 10 steps, 47% (previously 2%) over 11 steps, and 46% (previously 5%) over 11 steps, respectively, making the synthesis of these chiral lipids highly efficient. In the synthesis of mycolipanic acid, oxidation, Wittig reaction, and ester hydrolysis were significantly improved, whereas in the case of mycolipanic acid, the Evans’ aldol reaction and the removal of the chiral auxiliary were optimized to give high yields.

With the enantiopure acids 2–4 in hand, the esterification of palmitoylated trehalose 1 was achieved by following the Yamaguchi procedure. In the case of mycosanoic acid and mycolipanic acid, the corresponding diacylated products were obtained in good yields; however, in order to reach that result for mycolipanic acid, the esterification procedure needed to be carefully optimized to avoid acyl migration and elimination of the β-hydroxyl of 3. By limiting the number of equivalents of base and keeping the time for acid activation at a minimum, synthesis of 14b could be achieved in good yield. Notably, in the case of 14a and 14c, no acyl migration was observed, indicating that the β-hydroxyl in 3 might play a role in acyl migration. Removal of the silyl protecting group of 14a–14c under buffered conditions gave the corresponding diols 15a–15c in good to excellent yields. The final deprotection—the removal of the benzylidene protecting group—was achieved by applying a procedure that was reported by Guiard et al., using aqueous sulfuric acid (DAT₁ and DAT₃) or by palladium hydrogenolysis (DAT₂) to prevent β-hydroxyl elimination and provided the three di-O-acyl trehaloses DAT₁, DAT₂, and DAT₃ in moderate to good yields. (See Scheme 2.) The spectral data of DAT₁ matched the reported NMR data of isolated DAT₁. Besra’s report describes the 1H NMR signals of the anomeric protons of DAT₁ at 5.25 and 5.05 ppm for the acylated and nonacylated glucose unit, respectively. The spectrum of synthetic DAT₁ shows these two anomeric signals at 5.24 and 5.06 ppm, which is in good agreement. Furthermore, H-2 and H-3 (at the positions bearing the acyl moieties) in natural DAT₁ appear at 4.83 and 5.40 ppm, respectively, and in synthetic DAT₁ at 4.82 and 5.39 ppm, respectively. The 13C signals of the anomeric carbons in natural DAT₁ are reported at 95.0 and 92.0 ppm. In the synthetic material, these signals can be found 94.6 and 91.7 ppm, again in good agreement. In addition, the carbonyl carbon signals in synthetic DAT₁ resonate at 173.5 and 177.6 ppm and the corresponding signals in natural DAT₁ can be found at 173.8 and 177.8 ppm. All in all, these data leave us confident that the structure of synthetic DAT₁ is identical to that of natural DAT₁, as described by Besra (for more detailed NMR signal comparison, see the Supporting Information). As for synthetic DAT₂ and DAT₃, the structural identity is beyond reasonable doubt, because the structures of the lipid components have been previously established and the nuclear magnetic resonance (1H NMR and 13C NMR) and mass spectra showed patterns very similar to those of synthetic and natural DAT₁.

Detection of DAT₁, DAT₂, and DAT₃ in Mtb Strains. Having completed the total synthesis of DAT₁, DAT₂, and DAT₃ with structures as described in the literature, we sought to determine if the synthesized glycolipids match the structures of natural products present in virulent strains of Mtb. Mycobacterial lipid extracts of the reference strain H37Rv and three clinical isolates j257, j011, and j117 were analyzed by...
The extracted-ion chromatograms (Figure 2A) suggest that all three DATs are produced by the laboratory strain H37Rv. Ions consistent with DAT1 and DAT2 were only reliably detected in the H37Rv strain, whereas DAT3 could be detected in all four strains. The corresponding mass spectra of the detected natural DATs matched the calculated m/z of each compound within the expected experimental error of 3−4 ppm. Collision-induced fragmentation (see the data given in the Supporting Information) of the natural and synthetic DATs yielded interpretable cleavages (Figure 2C, H-transfers not shown) that supported the general structures and connectivity. Co-injection (Figure 2B) of synthetic standards and natural lipid mixtures showed a chromatographic match for DAT1 and DAT3. However, synthetic DAT2 eluted more than a minute earlier than the natural compound. Thus, whereas the identity of DAT1 and DAT3 can be considered to have been established beyond a reasonable doubt, we conclude that material with the structure of synthetic DAT2 does not occur in the H37Rv strain. This may mean that an isomer of the proposed structure of DAT2 is present in this strain, or that the structure of natural DAT2 has been incorrectly assigned.15

Mincle Activation by DAT1, DAT2, and DAT3. We decided to assess the Mincle activating properties of the synthetic DATs as well (see Figure 3), keeping in mind that our synthetic DAT2 was not present in the studied Mtb strains. Mincle activation was compared to the known Mincle-agonist TDM, which is highly potent. Previous studies have identified various lipoidalike trehaloses that activate Mincle, so we expected that all three forms of DAT, which differ in small ways in their alkyl chains, were good candidate activators. Prior to the functional assays, TLC analysis of synthetic DAT1, DAT2, and DAT3 was performed to exclude the presence of glycolipid degradation products and quantification errors (Figure 3A). Mincle activation requires the adaptor protein FcRγ. Therefore, functional Mincle-activation assays were performed by treatment of NFAT-GFP reporter cells expressing murine Mincle and FcRγ (Figure 3B) or human Mincle and FcRγ (Figure 3C) with the synthetic DAT variants and TDM. In both assays, DAT3 was able to activate Mincle. In the case of human Mincle, DAT3 showed similar potency to the highly potent agonist TDM. DAT2 and, remarkably, DAT1 only weakly activated murine Mincle. When using human Mincle-expressing cells, DAT2 showed moderate activation, whereas DAT1 again barely induced Mincle stimulation. In an independent experiment, an ELISA-based technique was applied that was dependent not on cellular activation but only on the detection of physical interaction between DAT and soluble Mincle proteins (Figure 3D). Strong binding of murine Mincle to DAT3 was observed, but only minimal binding to DAT1 and DAT2, thereby confirming the results obtained in the cellular activation assay. These results provide evidence that the chemical structure of the 3-O-acyl substituent (either mycosanoic acid, mycolipanolic acid, or mycolipenic acid) strongly influences Mincle binding and activation.

CONCLUSIONS

In this study, we have accomplished the first total synthesis of three structurally related mycobacterial DATs. These synthetic DATs were used as a reference in the detection of natural DATs in Mtb by liquid chromatography−mass spectroscopy. This showed that the presence and abundance of DAT1 and DAT3 differs strongly, dependent on the Mtb strain. This has important consequences for the potential use of DATs as
markers for Tb infection. In addition, it might explain a posteriori the irreproducibility observed in the many attempts to reliably detect DAT by ELISA. It also showed that the proposed structure of DAT\(_2\) does not occur in the studied strains, including the H37Rv strain. An alternative explanation is that the structure of DAT\(_2\) has been assigned incorrectly, since, because of the lack of literature NMR data, a comparison with our synthetic material was not possible. This will be further investigated.

We found that small changes in the structure of the branched acyl chain in DAT\(_3\) lead to large differences in recognition by Mincle. It has been shown previously by Decout et al. that one of the molecular requirements for Mincle recognition, besides the trehalose or glucose scaffold, is the presence of two alkyl chains, either as two separate esters or as one fatty acid ester with an alkyl chain branched \(\alpha\) to the carbonyl. Moreover, it was previously demonstrated that the lipid chains can be significantly shorter than the C80 lipids present in TDM.\(^{13}\) In addition, in a previous report, Mincle activation by \(\beta\)-glucosylceramide, which also contains an unsaturation in the lipid chain, was demonstrated.\(^{40}\) Here, we show that the presence of the \(\alpha\beta\)-unsaturation in DAT\(_3\) enhances Mincle activation drastically, compared to the saturated counterpart DAT\(_1\). This leads us to speculate that the double bond either serves as a point of interaction (such as \(\pi\)-\(\pi\)-stacking) with parts of the Mincle binding pocket or induces a specific conformation beneficial for binding. All in all, one might conclude that DAT\(_3\) could be an alternative starting point for adjuvant design for TDM, given the higher complexity and lipophilicity of the latter. For future development of \(Mtb\) vaccine adjuvants, even simpler DAT analogues could be designed without chiral methyl branches or based on glucose rather than trehalose.

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.0c00030.

Synthetic procedures, LC-MS protocols, Mincle functional and binding assay protocols, compound data (\(^1\)H and \(^{13}\)C NMR, HRMS) (PDF)

**AUTHOR INFORMATION**

**Corresponding Authors**

Ildiko Van Rhijn — Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands; Email: ivanrhijn@uu.nl

Adriaan J. Minnaard — Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands; orcid.org/0000-0002-5966-1300; Email: a.j.minnaard@rug.nl

**Authors**

Mira Holzheimer — Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands; orcid.org/0000-0003-2157-0136

Josephine F. Reijnveld — Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands; Brigham and Women’s Hospital Division of Rheumatology, Immunology and Allergy and Harvard Medical School, Boston, Massachusetts 02115, United States; Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, 3584 CL Utrecht, The Netherlands

Alexandra K. Rammarine — Brigham and Women’s Hospital Division of Rheumatology, Immunology and Allergy and Harvard Medical School, Boston, Massachusetts 02115, United States

Georgios Misiakos — Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

David C. Young — Brigham and Women’s Hospital Division of Rheumatology, Immunology and Allergy and Harvard Medical School, Boston, Massachusetts 02115, United States

Eri Ishikawa — Department of Molecular Immunology, Research Institute for Microbial Diseases and Laboratory of Molecular Immunology, Immunology Frontier Research Center, Osaka University, Suita, Osaka 565-0871, Japan

Tan-Yun Cheng — Brigham and Women’s Hospital Division of Rheumatology, Immunology and Allergy and Harvard Medical School, Boston, Massachusetts 02115, United States

https://dx.doi.org/10.1021/acschembio.0c00030

Sho Yamasaki — Department of Molecular Immunology, Research Institute for Microbial Diseases and Laboratory of Molecular Immunology, Immunology Frontier Research Center, Osaka University, Suita, Osaka 565-0871, Japan
D. Branch Moody — Brigham and Women’s Hospital Division of Rheumatology, Immunology and Allergy and Harvard Medical School, Boston, Massachusetts 02115, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.0c00030

Author Contributions
These authors contributed equally.

Funding
This work was supported by the Dutch Science Foundation (NWO), the Bill and Melinda Gates Foundation, and AMED (Nos. JP19gm0991010 and JP19ak0101070).

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS
We thank T. Tiemersma, R. Sneep, and P. van der Meulen (University of Groningen) for assistance on analyses.

REFERENCES


