Pergamon

Tetrahedron Letters 42 (2001) 3653–3655

TETRAHEDRON LETTERS

Structure and absolute configuration of mycobactin J

Brett D. Schwartz and James J. De Voss*

Department of Chemistry, University of Queensland, Brisbane, Queensland 4072, Australia

Received 20 March 2001; revised 22 March 2001; accepted 29 March 2001

Abstract—Mycobactin J **1** is a commercially available siderophore isolated from *Mycobacterium avium* subsp. *paratuberculosis*. There are discrepancies between previous reports of its structure and none have addressed its absolute configuration. We report here the complete structure and stereochemistry of mycobactin J, along with methodology to enable the determination of the absolute configuration of other mycobactins on a small scale.

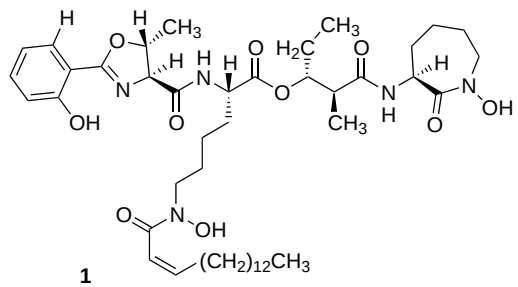
1. Introduction

Iron is the fourth most abundant element in the earth's crust and is an essential nutrient for most microorganisms. However, because the solubility of Fe(III) in aqueous media is so low (10^{-18} M at neutral pH) many bacteria have developed specialised mechanisms for its acquisition. In mycobacteria, iron uptake is generally mediated by siderophores (iron chelators) which fall broadly into two structural classes, the mycobactins and the exochelins.[1] The mycobactins contain a hydroxyphenyloxazoline moiety and were the first siderophores to be isolated from mycobacteria. Structurally, they are characterised by a core, which varies only slightly between different mycobactins, and by one to three very variable fatty acid derived alkyl chains. Mycobactins from any given species are usually isolated as a family of compounds differing in their alkyl substitution patterns. They were originally identified as a required growth factor for the in vitro culture of *Mycobacterium paratuberculosis* (now reclassified *M. avium* subsp. *paratuberculosis*[2]) the causative agents of Johne's disease in cattle.[3] Ironically, strains of *M. paratuberculosis* adapted to growth on laboratory media were subsequently shown to produce a mycobactin, mycobactin J,[4] which is now the only commercially available mycobactin. The structure of mycobactin J has been the subject of several reports which have arrived at slightly differing conclusions and none of which have addressed the stereochemistry of the system.[5–7] We report here the structure and absolute configuration of mycobactin J, along with methodology for rapidly determining the configuration of mycobactins on a small scale.

The major point of contention among the published mycobactin J structures is the identity of the β-hydroxy-acid that links the two lysine derived moieties of the siderophore. This was originally identified as being derived from the unusual 2,4-dimethyl-3-hydroxypentanoic acid.[5] Later, it was suggested that it was a 2-methyl-3-hydroxy carboxylic acid with a chain length of approximately 13 carbons.[6] Finally, on the basis of mass spectral evidence, it was recently assigned as the more common 2-methyl-3-hydroxypentanoic acid.[7] Additionally, the aromatic moiety has been suggested to be derived from either 6-methylsalicylate or salicylate and the oxazoline has been reported both with and without a methyl substituent.[5–7]

2. Results

Our analysis of the 400 MHz proton NMR spectrum of **1** clearly reveals the four protons expected for the unsubstituted hydroxyphenyl moiety shown in **1** (Fig. 1). The TOCSY spectrum reveals a spin system consisting of a methyl doublet at δ 1.6 (J 6 Hz), a multiplet at

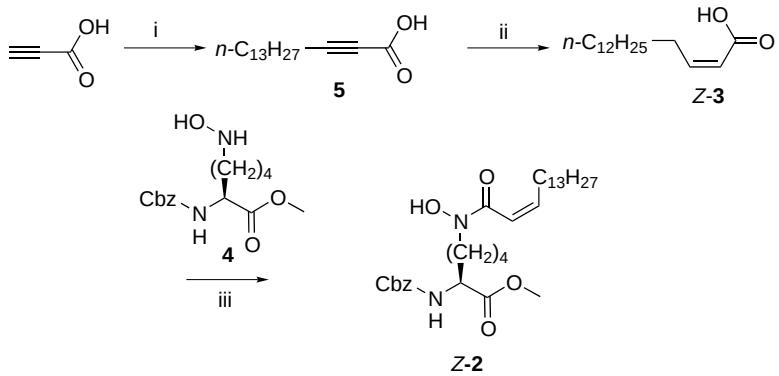


1

Figure 1. Structure and stereochemistry of the major component of mycobactin J.

* Corresponding author. E-mail: devoss@chemistry.uq.edu.au

0040-4039/01/$ - see front matter
PII: S0040-4039(01)00531-7



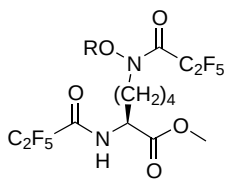
Scheme 1. Synthesis of model compound *Z*-**2**. (i) *n*-BuLi, HMPA, THF, *n*-$C_{13}H_{27}Br$, 65%; (ii) H_2 Lindlar catalyst, *n*-hexane, 72%; (iii) EDC, 18%.

δ 5.1 and a doublet at δ 4.5 (*J* 8 Hz) consistent with a methyl substituted oxazoline. Analysis of the TOCSY spectrum also suggested that the hydroxyacid was a 2-methyl-3-hydroxyacid as shown in **1**, although the number of carbons in the chain could not be determined. The geometry of the double bond of the long chain fatty acid moiety was difficult to determine from the NMR as the vicinal coupling constant cannot be measured because the vinyl protons appear as broad multiplets. We therefore synthesised and characterised *E*- and *Z*-**2** as model compounds. The *E* isomer was readily available from EDC mediated coupling of the known[8] acid *E*-**3** with the protected N^{ε}-hydroxylysine **4**.[9] The *cis* acid *Z*-**3** was available from the partial hydrogenation of the corresponding acetylene **5**, which in turn came from alkylation of the dianion of propiolic acid with 1-bromotridecane (Scheme 1). The chemical shifts of the vinylic and allylic protons seen in **3** and mycobactin J are given in Table 1, clearly indicating that the double bond geometry is *Z* as expected.

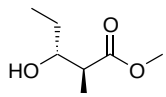
In order to establish the absolute configuration of mycobactin J we planned to employ the two step degradation procedure pioneered by Snow in the 1960s.[10] Base catalysed hydrolysis cleaves the only ester bond in the mycobactin and yields two fragments, a cobactin and a mycobactic acid that can be separated on the basis of their solubility.

Acid catalysed hydrolysis of the amide bonds then yields the constituent acids from each molecule. This route has the advantage of differentiating the two N^{ε}-hydroxylysines in the mycobactin, enabling their configuration to be determined separately. We wished to develop methodology for stereochemical determination that would be compatible with the small amount of other mycobactins usually available and thus explored GC analysis employing a chiral column. A Chirasil-Val® column is known to separate all of the naturally occurring amino acids when they are protected with pentafluoropropionyl groups,[11] but its resolving power with respect to N^{ε}-hydroxylysine was unknown.

N^{ε}-Hydroxylysine is somewhat unstable, but protected equivalents of it are readily available through Miller's dimethyldioxirane mediated oxidation of Cbz-lysine methyl ester.[9] Mild acid hydrolysis converted the nitrone to the free hydroxylamine which was immediately reacted with pentafluoropropionic anhydride (PFPA). The Cbz groups of the resulting amides were then removed by hydrogenolysis and the amines treated with PFPA to yield the triacyl derivative of N^{ε}-hydroxylysine **6** as a single product by GC–MS ($M^{+\bullet}$, 614). This compound was found to be somewhat unstable, slowly converting to the diacyl compound **7** (GC–MS $M^{+\bullet}$ 468) at room temperature. Treatment with PFPA regenerated the triacyl derivative **6**. Both the L- and DL-forms of the PFPA protected N^{ε}-hydroxylysine were synthesised from L- and DL-Cbz-lysine methyl ester, respectively. Analysis of the protected N^{ε}-hydroxylysines employing the Chirasil-Val® column showed that while the enantiomers of the diacyl derivative **7** did not separate well, those of the triacyl derivative **6** were clearly resolved. Thus, we felt that we could determine the stereochemistry of all the amino acid derived centres in mycobactin J by direct analysis of the hydrolysis products.



6 R = COC_2F_5
7 R = H



(2*S*,3*R*)-**8**

Hydrolysis of mycobactin J under basic conditions yielded the corresponding mycobactic acid and cobactin. Acid hydrolysis of the mycobactic acid was followed by esterification and PFPA derivatisation. GC–MS analysis utilising the Chirasil-Val® column

Table 1. Chemical shifts (400 MHz, $CDCl_3$) of characteristic protons in **1** and **2**

Proton	*E*-**2**	*Z*-**2**	Mycobactin J **1**
H2′	5.99	5.85	5.79
H3′	6.95	6.05	5.97
H4′	2.17	2.49	2.40

indicated that the only amino acids present in mycobactic acid were L-threonine and L-N^{ε}-hydroxylysine. In addition, methyl salicylate was the only aromatic acid identified in the diazomethane esterified crude acid hydrolysate.[†] The mycobactic acid fragment also contains the fatty acyl substituent which is the source of much of the structural variation observed in mycobactins. As expected therefore, GC–MS analysis revealed a range of fatty acids that included the methyl esters of 2-hexadecenoic, hexadecanoic and octadecanoic acids as the major components with traces of 2-octadecenoic and eicosanoic acids. Comparison with synthetic *E*- and *Z*-methyl-2-hexenoate revealed that the *E* isomer, rather than the expected *Z* form was produced in the degradation of mycobactin J. The ease of isomerisation of α,β-unsaturated fatty acids prompted us to investigate the periodate catalysed cleavage of the fatty acid moiety from the corresponding mycobactic acid.[10] As previously reported, no isomerisation occurs under these conditions and we identified only the *Z*-methyl-2-hexenoate by comparison with authentic material (GC coelution, MS fragmentation).

Similarly, analysis of the appropriately protected acid hydrolysate of cobactin yielded only L-N^{ε}-hydroxylysine and methyl 2-methyl-3-hydroxypentanoate **8**. All four stereoisomers of the latter were synthesised via double methylation of the dianion derived from methyl acetoacetate followed by selective reduction of the ketone of the ketoester with sodium borohydride. The known 2*S*,3*R* enantiomer of **8**[12] was available from propionic acid by utilisation of Evan's methodology. We found that all four stereoisomers could be separated by GC on a γ-cyclodextrin column and thus showed that the hydroxyacid derived from mycobactin J had the 2*S*,3*R* configuration. The *erythro* configuration of these two centres is consistent with that determined for all other mycobactins to date and the absolute configuration is the same as that seen in mycobactin R from *M. terrae*.[10]

3. Conclusion

We have clarified the structure of mycobactin J as that shown in **1**, in which the hydroxyacid is derived from (2*S*,3*R*)-3-hydroxy-2-methylpentanoate. In addition, we have also shown that the amino acid derived stereocentres all have the configuration found in the naturally occurring L-isomers. This has important implications for the mechanism of mycobactin biosynthesis. For example, retention of the 3*R* configuration of L-threonine in the oxazoline implies that cyclisation occurs via hydroxyl attack on the adjacent amide carbonyl rather than the reverse; the latter mechanism is well precedented in the synthesis of these heterocycles.[13] The gross structure of mycobactin J is the same as that recently proposed for mycobactin Av from *M. avium* as might be expected given the reclassification of *M. paratuberculosis* as a subspecies of *M. avium*.[2] No stereochemical information is yet available for mycobactin Av but it appears likely that mycobactin Av and J will prove identical. The development of GC protocols for determining the absolute stereochemistry of the mycobactins will be compatible with the small amounts of naturally occurring mycobactins available and be invaluable for comparing the recently identified water soluble mycobactins[7,14] with the cell associated ones.

[†] The standard methanol/HCl and PFPA derivatisation procedures resulted in the loss of the more volatile fatty esters. These were examined by GC–MS after esterification with ethereal diazomethane.

Acknowledgements

This work was supported by NH&MRC Grant 143031 to J.D.V.

References

1. De Voss, J. J.; Rutter, K.; Schroeder, B. G.; Barry, C. E. *J. Bacteriol.* **1999**, *181*, 4443–4451.
2. Thorel, M. F.; Krichevsky, M.; Levyfrebault, V. V. *Int. J. Syst. Bacteriol.* **1990**, *40*, 254–260.
3. Francis, J.; Macturk, H. M.; Madinaveitia, J.; Snow, G. A. *Biochem. J.* **1953**, *55*, 596–607.
4. Merkal, R. S.; McCullough, W. G. *Curr. Microbiol.* **1982**, *7*, 333–335.
5. McCullough, W. G.; Merkal, R. S. *Curr. Microbiol.* **1982**, *7*, 337–341.
6. Barclay, R.; Ewing, D. F.; Ratledge, C. *J. Bacteriol.* **1985**, *164*, 896–903.
7. Gobin, J.; Moore, C. H.; Reeve, Jr., J. R.; Wong, D. K.; Gibson, B. W.; Horwitz, M. A. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5189–5193.
8. Hino, T.; Nakakyama, K.; Taniguchi, M.; Nakagawa, M. *J. Chem. Soc., Perkin Trans. 1* **1986**, 1687–1690.
9. Hu, J. D.; Miller, M. J. *J. Org. Chem.* **1994**, *59*, 4858–4861.
10. Snow, G. A.; White, A. J. *Biochem. J.* **1969**, *115*, 1031–1045.
11. Frank, H.; Nicholson, G. J.; Bayer, E. *J. Chromatogr. Sci.* **1977**, *15*, 174–176.
12. Cane, D. E.; Yang, C. C. *J. Am. Chem. Soc.* **1987**, *109*, 1255–1257.
13. Wipf, P.; Miller, C. P. *Tetrahedron Lett.* **1992**, *33*, 907–910.
14. Lane, S. J.; Marshall, P. S.; Upton, R. J.; Ratledge, C.; Ewing, M. *Tetrahedron Lett.* **1995**, *36*, 4129–4132.