Structure Elucidation

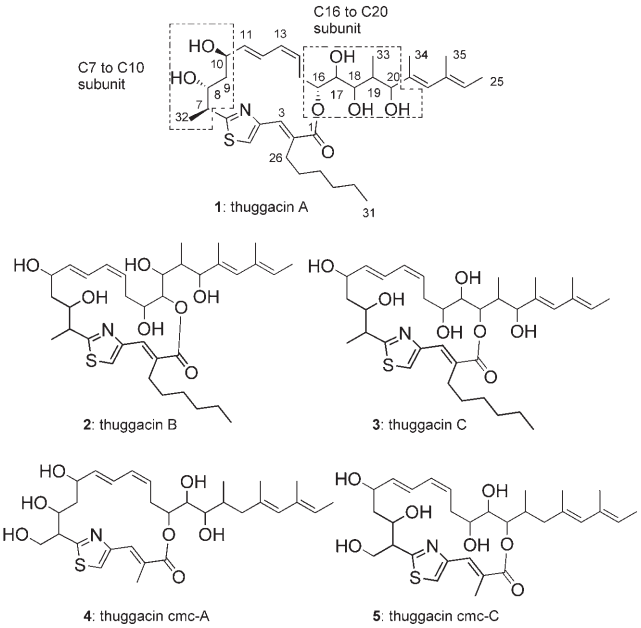
DOI: 10.1002/anie.200704897

Stereochemical Determination of Thuggacins A–C, Highly Active Antibiotics from the Myxobacterium *Sorangium cellulosum***

*Martin Bock, Kathrin Buntin, Rolf Müller, and Andreas Kirschning**

Since 1993, when the World Health Organization declared tuberculosis "a global health emergency", there has been new impetus to research on tuberculosis because it can no longer be sufficiently treated by currently available antibiotic therapy. The increasing multidrug resistance of *Mycobacterium tuberculosis* and its ability to persist as a latent infection makes the development of alternative antibiotics—preferably with novel modes of action—necessary.[1] The polyketide natural products thuggacin A (**1**), B (**2**), and C (**3**) recently isolated from the myxobacterium *Sorangium cellulosum* show strong antibiotic activity against *Mycobacterium tuberculosis* which targets the bacterial respiratory chain.[2] Recently, it was found that the myxobacterium *Chondromyces crocatus* Cm c5 also produces thuggacin derivatives, namely thuggacin cmc-A (**4**) and thuggacin cmc-C (**5**).

Thuggacin A (**1**; Scheme 1) features a 17-membered macrolactone with a thiazole ring, a diene (11*E*, 13*Z*), and an α,β-unsaturated lactone with a *n*-hexyl side chain at C2 plus a complex side chain at C16 bearing three hydroxy groups and a diene unit. Thuggacin B (**2**) shares the same structural features except for the ring size: the cyclic lactone is closed at O17 instead of O16 in thuggacin A. Finally, thuggacin C (**3**) is macrocyclized at O18. With two additional hydroxy groups and an aliphatic methyl group in the macrolactone ring, the thuggacins include eight stereogenic centers. The structure of the thuggacins, their constitution, and preliminary assignments of the relative stereochemistry were determined by Jansen et al. on the basis of NMR spectroscopic data.[2] However, as the analytical methods employed strongly relied on the rigid conformation of the ring system, they could not be extended to the stereocenters of the freely rotating side chain.



Scheme 1. Thuggacin A (**1**) (relative configuration at C7, C8, C10, and C16 as published by Jansen et al.[2]) and thuggacins B (**2**) C (**3**), cmc-A (**4**), and cmc-C (**5**).

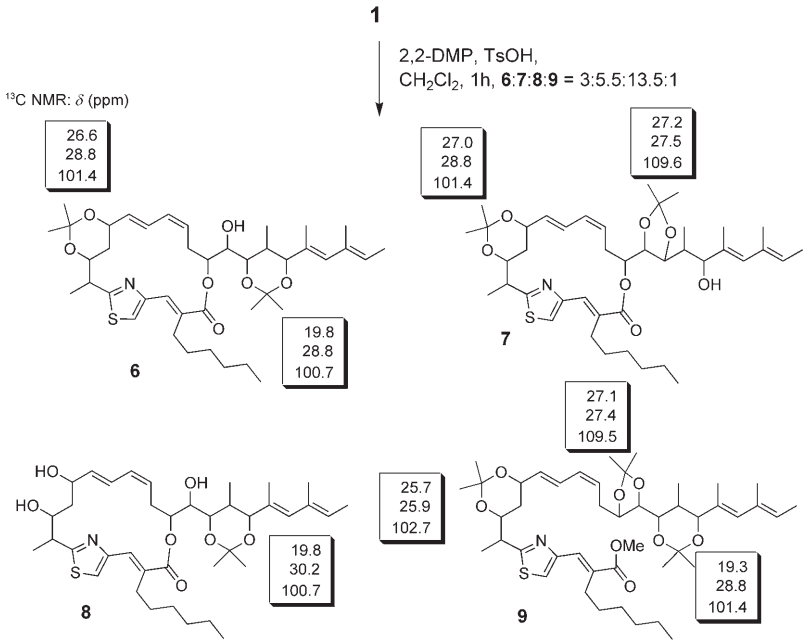
For this purpose, synthetic derivatization combined with NMR analysis and modeling was conducted. We also report that bioinformatic analysis of the thuggacin biosynthetic genes is a powerful tool for confirming the absolute configuration of the stereocenters first assigned by chemical methods. Thus, the relative configurations of both stereodomains (C7 to C10 and C16 to C20) were first determined separately, which was followed by establishing the correct sterochemical relationship between these stereoclusters.

Since six out of eight stereogenic centers bear a hydroxy group, a straightforward derivatization strategy could be developed based on rigidifying selected areas by protection of 1,3- and 1,2-diols as acetonides (Scheme 2). The constitutions of derivatives **6**–**9** were assigned by ^{1}H and ^{13}C NMR spectroscopy complemented with COSY, HMQC, and HMBC experiments and by comparison with the natural product **1**. Two chemical aspects are noteworthy, namely 1) the preferred formation of the 1,3-acetonide (C18/C20) over the 1,2-acetonide (C17/C18) and 2) formation of the acyclic methyl ester **9**, which was important to secure the NMR data of macrocyclic derivatives **6**–**8**. In fact, the δ (^{13}C) values found for the acetonide carbon atoms at C8/C10 and

[*] M. Bock, Prof. Dr. A. Kirschning
Institut für Organische Chemie
und Biomolekulares Wirkstoffzentrum (BMWZ)
Leibniz Universität Hannover
Schneiderberg 1B, 30167 Hannover (Germany)
Fax: (+49) 511-762-3011
E-mail: andreas.kirschning@oci.uni-hannover.de
K. Buntin, Prof. Dr. R. Müller
Institut für Pharmazeutische Biotechnologie,
Universität des Saarlandes
Postfach 151150, 66041 Saarbrücken (Germany)
Prof. Dr. R. Müller
Helmholtz-Zentrum für Infektionsforschung (HZI)
Inhoffenstrasse 7, 38124 Braunschweig (Germany)

[**] This work was supported by the Fonds der Chemischen Industrie. We are indebted to H. Irschik (HZI, Braunschweig) for the fermentation of *Sorangium cellulosium* and R. Jansen (HZI, Braunschweig) for helpful discussions.

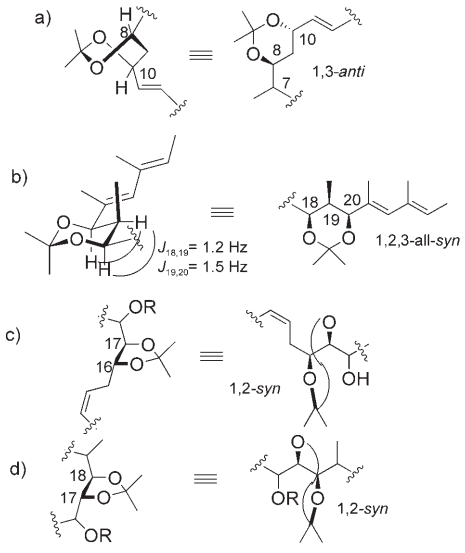
Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.



Scheme 2. Formation of acetonides **6**–**9** and relevant ^{13}C NMR signals of acetonide carbon atoms; 2,2-DMP = 2,2-dimethoxypropane, TsOH = *para*-toluenesulfonic acid.

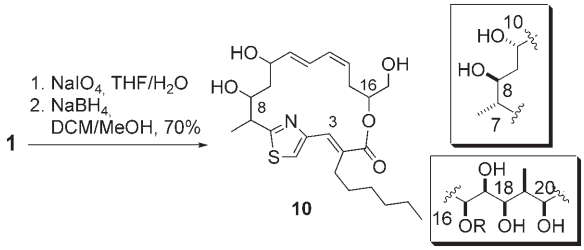
C18/C20 of methyl ester **9** are similar to those of the analogous macrocyclic derivatives **6**–**8**.

These data unambiguously indicate that the conformations in the 1,3-dioxane rings are not distorted by the macrocycle. Thus, Rychnovsky's method[3,4] could be applied reliably to assign the relative configuration of the 1,3-diols via their 1,3-dioxanes. Accordingly, the 1,3-diol at C8/C10 is configured *anti*. By including the information on the vicinal coupling constants ($^3J_{HH}$) between 18-H/19-H and 19-H/20-H, which both are in the range of 1.5 Hz for all derivatives, the relative stereochemistry of the stereotriade C18 to C20 was assigned to be all-*syn*. Moreover, the 1,2-diols protected in dioxolanes **7** and **9** provide information on the relative stereochemistry around C16/C17 and C17/C18 which were both determined to be 1,2-*syn*[5,6] (Scheme 3).



Scheme 3. Relative configuration of 1,3-dioxolanes and 1,3-dioxanes.

Final proof for the relative configurations of all stereogenic centers (C7, C8, C9, and C16) within the ring was possible after oxidative removal of the complex side chain. The intermediate aldehyde had to be reduced directly to the triol **10** in order to obtain sufficiently resolved NMR spectra (Scheme 4). Besides determination of vicinal coupling constants, molecular modeling studies (MMFFs) of the four possible diastereomers generated by variation of the C7 and C16 stereogenic centers in relation to the *anti* 1,8-diol unit of triol **10** were conducted.[7] To account for the observed NOE couplings, we constrained the distances between the atom pairs 7-H/10-H, 3-H/12-H, and 5-H/26-H to a range between 2 and 4.5 Å in the MonteCarlo searches. The calculated structures for all four diastereomers were analyzed for discrepancies between their structures and the observed



Scheme 4. Glycol cleavage of thuggacin A (**1**) and relative configuration of the two stereodomains.

NMR data (i.e. vicinal ^{1}H coupling constants and NOE couplings).

In both C7 *anti* structures, the 7-H would adopt an *exo* position However, the observed strong 7-H/10-H NOE contact is not consistent with this configuration (Table 1). This center is therefore assigned to be *syn* oriented with respect to the hydroxyl group at C8. This is further supported by a coupling constant of H-9 *exo* to H-10 of about 10 Hz. Similar results were obtained for the C16 center: 15-Ha has two vicinal coupling constants of about 10 Hz, matching dihedral angles of 160 to 170°. The modeled dihedral angle 15-Ha,C15,C16,16-H is about 165° for one configuration and 60°

Table 1: Selected structural features of the four modeled isomers of triol **10**.

Isomer	7-H position	Dihedral angle 15-Ha,C15,C16,16-H
10	*endo*	167°
7-*epi*-**10**	*exo*	162°
16-*epi*-**10**	*endo*	62°
7,16-*epi*-**10**	*exo*	64°

for the opposite one (Table 1).[8] This clearly points to the configuration shown in Figure 1. In fact, the 3D structure depicted in Figure 1 looks very similar to that of thuggacin A.[2] By assembling all of data collected, we could deduce the complete relative configuration of thuggacin A (Scheme 4 and Figure 1).

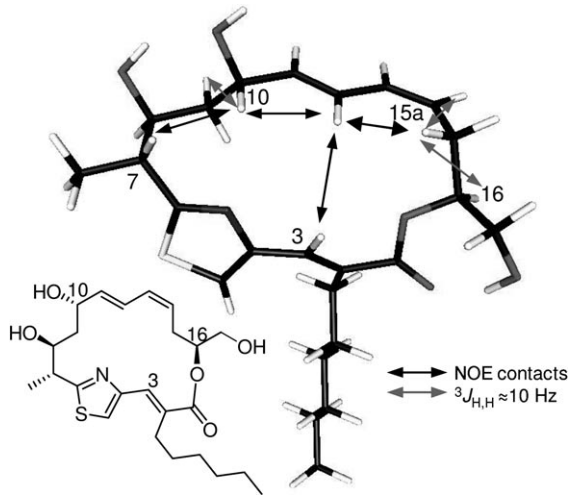


Figure 1. Modeling of triol **10** and selected NMR data; numbering refers to atom numbers given in Scheme 1.

Finally, the absolute configuration was determined by transformation of derivative **6** containing a free hydroxy group at C17 into esters **11a** and **11b** (Scheme 5) and applying the Mosher analysis rules.[9] From the calculated $\delta\Delta$ values the configuration at C17 was deduced to be *S*. Further support for the absolute configuration was obtained after removal of the isopropylidene protecting group yielding tetraols **12a** and **12b**.[10] Very similar $\delta\Delta$ values were determined, indicating that the adjacent 1,3-dioxane ring did not invalidate the Mosher analysis of **11a** and **11b**.

Recently, it has become possible to predict the stereochemistry resulting from keto reduction during polyketide biosynthesis based on a bioinformatic analysis of the respective domains in giant polyketide synthase modules.[11–13] This method is especially useful if the availability of the natural product is limited and it is not possible to generate degradation fragments in sufficient quantities for configurational analysis. However, the biosynthetic genes must be identified and analyzed. In the course of this study a method for gene inactivation in *Chondromyces crocatus* Cm c5 was applied, and the chondramide biosynthetic gene cluster was identified. As this myxobacterium also produces thuggacin derivatives thuggacin cmc-A (**4**) and thuggacin cmc-C (**5**) and since we assume that the biosynthesis for thugaccin A and thuggacin cmc-C is very similar, we set out to identify the corresponding gene cluster in *C. crocatus* Cm c5. Inactivation of a candidate polyketide synthase indeed resulted in the loss of thuggacin cmc-C production in *C. crocatus* Cm c5 and enabled the cloning and sequencing of the complete biosynthetic gene cluster which encodes a complex hybrid of polyketide synthases (PKSs) and a nonribosomal peptide synthetases. By following the analysis of the ketoreductase domains (KRs) within the PKS according to the core amino acid analysis described to predict the absolute configuration of secondary alcohols by Caffrey and Reid et al.[11] we could propose the configuration at positions C18, C16, C10, and C8 (Table 2).[14]

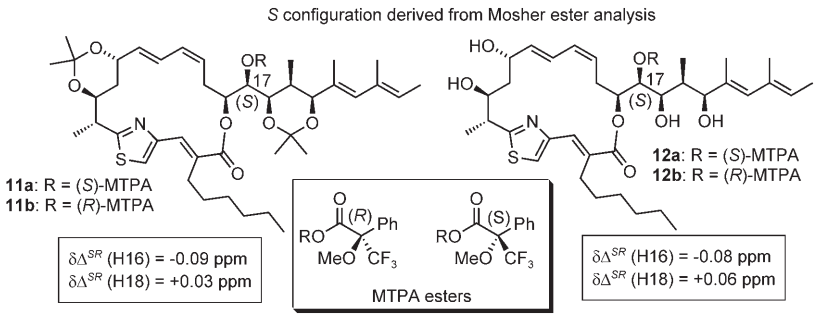
Table 2: Type of ketoreductases found in the polyketide synthase domains involved in thuggacin biosynthesis and the corresponding stereogenic centers with their proposed absolute configuration (C20, C18, C16, C10, C8).[a]

Ketoreductase	Type	Configuration
KR1 TugA	B-type	DB *trans*
KR2 TugA	B-type	DB *trans*
KR3 TugA	B-type	C20 (*S*)
KR1 TugB	B-type	C18 (*R*)
KR2 TugB	B-type	C16 (*S*)
KR1 TugC	A-type	DB *cis*
KR2 TugC	B-type	DB *trans*
KR1 TugD	A-type	C10 (*S*)
KR2 TugD	A-type	C8 (*S*)
KR3 TugD	B-type	DB *trans*

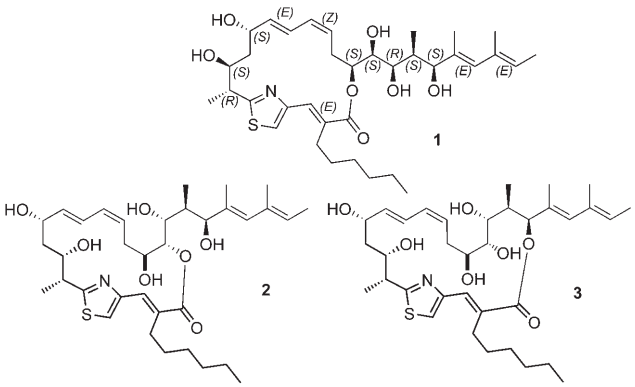
[a] Action of all other KR domains is followed by elimination of water and thus results in double bonds (DB) in the final molecule. Predicted double-bond configurations and absolute configurations of the alcohols are given (see the Supporting Information).

Only the assignment at C10 is somewhat ambiguous because the conserved amino acids are not in total agreement with the consensus. The application of these results to thuggacin A (**1**) confirmed the assigned configuration. Although the thuggacins produced by Cm c5 do not contain an OH group attached to C20, a corresponding intermediate must be produced during PKS assembly, and therefore the configuration at C20 can as well be predicted.

By considering all of the data presented, we could deduce the complete relative configuration of thuggacin A (Scheme 4 and Figure 1). As far as C7, C8, C10, and C16 are concerned, our relative assignments are in accordance with those reported by Jansen et al.[2] Scheme 6 summarizes the assignments of the complete relative and absolute config-



Scheme 5. Configurational assignment of thuggacin A by Mosher derivatization of **6** (the *R*-MPTA ester was prepared from *S*-MPTA chloride and vice versa; see the Supporting Information). MPTA = α-methoxy-α-(trifluoromethyl)phenylacetic acid.



Scheme 6. Absolute configuration of thuggacins A–C (**1**–**3**).

uration of thuggacin A. On the basis of a common biosynthesis, and the observed rearrangement of thuggacin A (**1**) into thuggacins B (**2**) and C (**3**), their configurations were assigned accordingly.[2] This assumption is in full agreement with the close similarities between the spectroscopic and physical data of thuggacins A–C.[2]

In conclusion, the complete stereochemical assignment for the antibiotic macrolides thuggacin A, B, and C has been determined as 2*E*,7*R*,8*S*,10*S*,11*E*,13*Z*,16*S*,17*S*,18*R*,19*S*,20*S*,-21*E*,23*E* on the basis of high-field NMR studies combined with extensive chemical derivatization and molecular modeling. Final proof of the absolute configuration of five stereogenic centers was collected from bioinformatic analysis of the biosynthetic proteins for which the present work provides one of the first examples.[12,13]

Received: October 22, 2007
Revised: November 26, 2007
Published online: February 14, 2008

Keywords: antibiotics · gene analysis · macrolactones · Mosher ester · structure elucidation

[1] J. D. McKinney, *Nat. Med.* **2000**, *6*, 1330 – 1333.
[2] H. Steinmetz, H. Irschik, B. Kunze, H. Reichenbach, G. Höfle, R. Jansen, *Chem. Eur. J.* **2007**, *13*, 5822 – 5832.
[3] S. Rychnovsky, D. J. Skalitzky, *Tetrahedron Lett.* **1990**, *31*, 945 – 948.
[4] Diagnostic δ (^{13}C) values for the 1,3-*syn* configuration: Me_{ax} $\approx$ 20 ppm, $Me_{eq} \approx$ 30 ppm; for 1,3-*anti*: both Me $\approx$ 25 ppm. Occasionally, also the δ (^{13}C) values for the quartenary C atoms are analyzed: $\delta < 100$ ppm (for 1,3-*syn*) and $\delta > 100$ ppm (for 1,3-*anti*).
[5] G. Dana, H. Denechpajouh, *Bull. Soc. Chim. Fr.* **1980**, 395 – 399.
[6] Diagnostic δ (^{13}C) values for the 1,2-*anti* configuration: Me_{β} $\approx$ 25 ppm, $Me_{\gamma} \approx$ 28 ppm; for 1,2-*syn* both Me $\approx$ 27 ppm.
[7] a) F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440 – 467; b) W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127 – 6129.
[8] Forcing a dihedral angle 15-Ha,C15,C16,16-H to be 150 – 180° in 16-*epi*-**10** is associated with a destabilization of about 10 kcal mol^{-1} and is not in agreement with the observed NMR data. (For vicinal ^{1}H coupling constants and NOE couplings, refer to the Supporting Information.)
[9] a) J. A. Dale, H. S. Mosher, *J. Am. Chem. Soc.* **1968**, *90*, 3732 – 3738; b) J. A. Dale, D. L. Dull, H. S. Mosher, *J. Org. Chem.* **1969**, *34*, 2543 – 2449; c) J. M. Seco, E. Quinoa, R. Riguera, *Chem. Rev.* **2004**, *104*, 17 – 117.
[10] One by-product from the formation of the *S* ester was assigned to be the thuggacin B derivative of diacetonide **6**. The NMR data of this acetonide were in agreement with the configuration deduced from the initial acetalization studies (refer to the Supporting Information).
[11] a) R. Reid, M. Piagentini, E. Rodriguez, G. Ashley, N. Viswanathan, J. Carney, D. V. Santi, C. R. Hutchinson, R. McDaniel, *Biochemistry* **2003**, *42*, 72 – 79; b) P. Caffrey, *ChemBioChem* **2003**, *4*, 654 – 657.
[12] a) D. Janssen, D. Albert, R. Jansen, R. Müller, M. Kalesse, *Angew. Chem.* **2007**, *119*, 4985 – 4988; *Angew. Chem. Int. Ed.* **2007**, *46*, 4898 – 4901.
[13] D. W. Udwary, L. Zeigler, R. N. Asolkar, V. Singan, A. Lapidus, W. Fenical, P. R. Jensen, B. S. Moore, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10376 – 10381.
[14] For analysis of the ketoreductase domains refer to the Supporting Information.

