Chemical Science

EDGE ARTICLE

Oxidative dearomatisation: the key step of sorbicillinoid biosynthesis†

Ahmed al Fahad,[a] Amira Abood,[ae] Katja M. Fisch,[a] Anna Osipow,[a] Jack Davison,[a] Marija Avramović,[b] Craig P. Butts,[a] Jörn Piel,[bc] Thomas J. Simpson[a] and Russell J. Cox*[ad]

An FAD-dependent monooxygenase encoding gene (SorbC) was cloned from *Penicillium chrysogenum* E01-10/3 and expressed as a soluble protein in *Escherichia coli*. The enzyme efficiently performed the oxidative dearomatisation of sorbicillin and dihydrosorbicillin to give sorbicillinol and dihydrosorbicillinol respectively. Bioinformatic examination of the gene cluster surrounding SorbC indicated the presence of two polyketide synthase (PKS) encoding genes designated *sorbA* and *sorbB*. The gene *sorbA*-encodes a highly reducing iterative PKS while SorbB encodes a non-reducing iterative PKS which features a reductive release domain usually involved in the production of polyketide aldehydes. Using these observations and previously reported results from isotopic feeding experiments a new and simpler biosynthetic route to the sorbicillin class of secondary metabolites is proposed which is consistent with all reported experimental results.

Introduction

The sorbicillinoid natural products, such as sorbicillactones A **1a** and B **1b**, and dimeric and trimeric species **2–4**, comprise a large class of highly complex cyclic and polycyclic compounds, often with unusual biological activities (Fig. 1).[1] They occur in fungi, especially among species of *Penicillium*, and *Trichoderma* but also sporadically in other genera of ascomycetes from both terrestrial and marine environments. The biosynthesis of these compounds has been proven to be of polyketide origin through labelling experiments using stable isotopes. For example, bis-vertinolone **2** is a β-1,6-glucan biosynthesis inhibitor and radical scavenger which has been demonstrated to be derived from acetate.[2] Bisorbicillinol **3** is also a radical scavenger,[3] while trisorbicillinone C **4** is cytotoxic *vs.* P388 and HL60 cell lines.[4] It has been hypothesised that sorbicillinol **5a** forms a key pathway intermediate and is a precursor to almost all of the sorbicillinoids.[2,5] Labelling studies have been used to suggest the

biosynthetic route shown in Scheme 1, but until now little biochemical and no genetic information has been available.

The proposed biosynthetic route is rather unusual, featuring the formation of an unprecedented uncyclised tetracarbonyl **6** by a fungal polyketide synthase (PKS),[6] which is then hydroxylated to give **7** prior to cyclisation to form oxosorbicillinol **8**. This is then reduced to the alcohol **9** which has not been observed. Dehydration is then proposed to give sorbicillinol **5a**, which must be reduced to the observed sorbicillin **10a** (Scheme 1).

Support for this proposed route came from the incorporation of $[1,2^{-13}C_2]$-acetate into sorbicillinol **5a** which suggested that a symmetrical intermediate (*i.e.* **11**, **11′**) was not formed.[7] This conclusion was arrived at because free rotation of the symmetrical aromatic moiety of **11**, formed by cyclisation of **6**, would scramble the position of the *intact* acetate units (Scheme 1) and scrambling of label was not observed. The proposed pathway is rather unsatisfactory because of the hydroxylation prior to cyclisation; the fact that **9** has not been observed; and lengthy reductive chemistry required to synthesise sorbicillin **10a**.

The proposal that hydroxylation of the polyketide occurs *before* cyclisation rests upon the assumption that the sorbicillin PKS forms the putative thiolester precursor **6**. However, recent results from our laboratories have shown that some fungal PKS can release their products as aldehydes, instead of thiolesters.[8] Furthermore we have also recently shown that oxidative dearomatisation of PKS-products is a common early transformation during the biosynthesis of a number of fungal secondary metabolites[9] and this led us to reassess the biosynthetic

[a]*School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK. E-mail: r.j.cox@bris.ac.uk; Fax: +44 (0) 117 925 1295; Tel: +44 (0) 117 928 9184*

[b]*Institute of Microbiology, Eidgenössische Technische Hochschule (ETH) Zürich, Wolfgang-Pauli-Str. 10, 8093 Zürich, Switzerland*

[c]*Kekulé Institute of Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Str. 1, 53121 Bonn, Germany*

[d]*Institut für Organische Chemie, Leibniz Universität Hannover, Schneiderberg 1B, 30167 Hannover, Germany. E-mail: russell.cox@oci.uni-hannover.de*

[e]*Chemistry of Natural and Microbial Products, National Research Centre, Dokki, Egypt*

† Electronic supplementary information (ESI) available: Containing all experimental details. See DOI: 10.1039/c3sc52911h
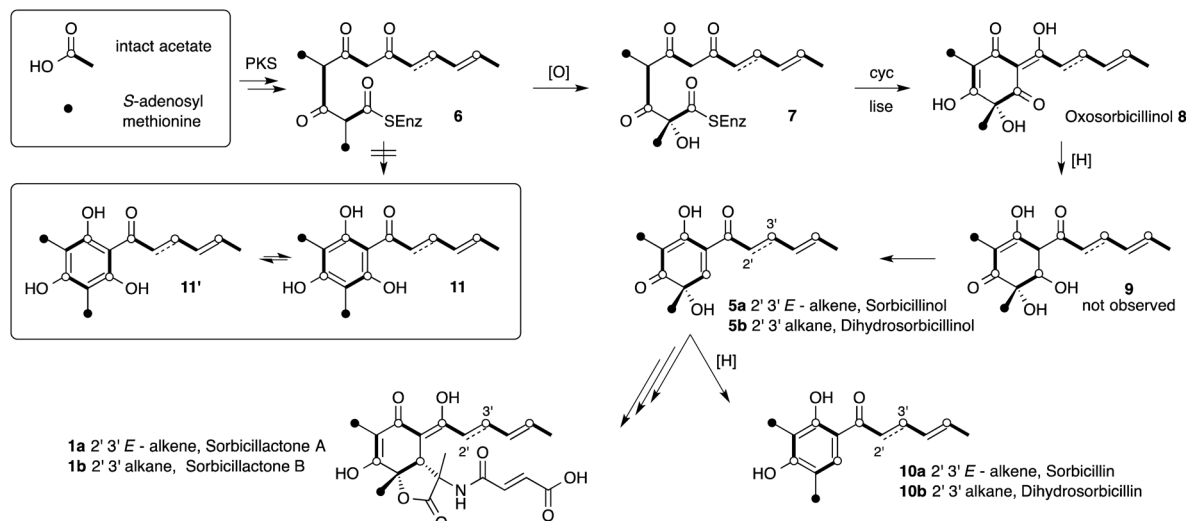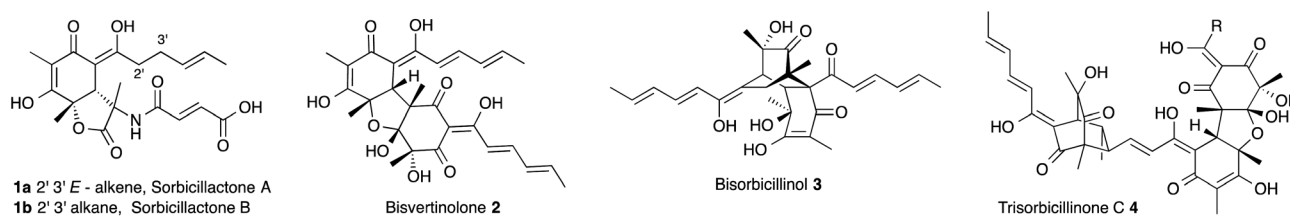
Scheme 1 Earlier proposed biosynthesis of sorbicillinol.[2]

Fig. 1 Structural diversity among the sorbicillinoid family. R = CH=CHCH=CHMe.

pathway to sorbicillinol **5a**. Here we describe molecular experiments which support a different route to sorbicillinol **5a** which is entirely consistent with the results from the previously described isotopic labelling experiments, and with recently reported results in other systems.

Results

Bringmann and coworkers reported the production of sorbicillactones A **1a** and B **1b** by the fungus *Penicillium chrysogenum* E01-10/3, as well as sorbicillin **10a** itself.[10] In our hands fermentation of this organism produced these compounds in good titres, but also produced 2′,3′-dihydrosorbicillin **10b**.[5,11] We hypothesised that **10a** and **10b** were likely to be precursors of sorbicillactones A and B respectively. **10a** and **10b** might be formed by synthesis of intermediate polyketide aldehydes which could undergo Knoevenagel cyclisation to form **5a** and **5b** directly. Aldehydes have recently been recognised as direct products of fungal non-reducing iterative polyketide synthases (NR-iPKS), for example during the biosynthesis of methyl-orcinaldehyde by the NR-iPKS methylorcinaldehyde synthase (MOS, encoded by *aspks*1), and also precursors of the azaphilones.[12] The formation of the aldehyde is achieved by a reductive release domain instead of the more usual hydrolytic release domain.[13] This class of NR-iPKS has been recognised in phylogenetic analyses and designated non-reducing clade III (NRcIII).[14] We have previously described the design of selective

degenerate PCR primers which can recognise and amplify fragments of NRcIII genes: for example KHKS2/KHKS3c primers were used in the successful cloning of *aspks*1.[8]

We thus applied PCR using the KHKS2/KHKS3c primers to gDNA isolated from *P. chrysogenum* E01-10/3. Three distinct 245 bp fragments of iPKS genes were amplified and sequenced, and one of these (KHKS32) was shown to be closely related to the iPKS citrinin synthase[15] which belongs to NRcIII. Next, a 4800-clone fosmid library was constructed from *P. chrysogenum* E01-10/3 gDNA. The fosmid library was screened by PCR using primers based on the sequence of KHKS32. This process yielded a number of potential fosmid clones, and the largest of these was subjected to random shotgun sequencing (GATC Biotech). This revealed the fosmid clone to be 48.8 kb in length. Analysis of the sequence (fgenesh[16]) revealed the presence of seven putative open reading frames possibly involved in secondary metabolism (*orfs* 1–7, Scheme 2) encoding: a transcriptional regulator; an FAD-dependent monooxygenase; an NR-iPKS; a highly-reducing iPKS (HR-iPKS); a transcription factor; an ABC transporter; and a second FAD-dependent oxidoreductase (Scheme 2). Interestingly the sequence is almost identical to a fragment of the sequenced genome of *P. chrysogenum* Wisconsin 54-1255,[17] but although this strain is a penicillin G producer it is not a known producer of sorbicillinoids.[18]

Domain analysis (see ESI†) of the NR-iPKS shows that it possesses typical starter-unit acyl transferase (SAT), β-ketoacylsynthase (KS), acyl transferase (AT), product template (PT),
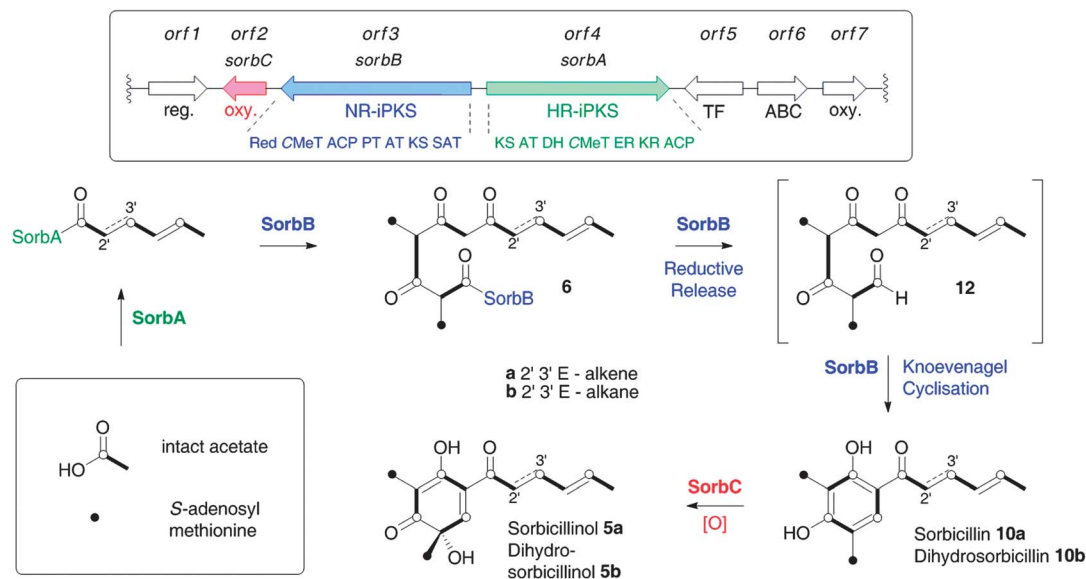
Scheme 2 New proposed biosynthetic route to sorbicillinols.

acyl carrier protein (ACP), *C*-methyl transferase (*C*MeT) and reductive release (Red) domains consistent with the chemistry required for the later steps of sorbicillin biosynthesis – and crucially for the release of an aldehyde intermediate. Similar analysis of the HR-iPKS indicated the presence of KS, AT, dehydratase (DH), *C*MeT, β-ketoacylreductase (KR), enoylreductase (ER) and ACP domains consistent with the early stages of sorbicillin biosynthesis. Attempts were made to knockout these PKS genes in *P. chrysogenum* E01-10/3, but without success.

In an alternative strategy to link these genes with the biosynthesis of the sorbicillins we considered the activity of the putative tailoring genes. *Orf*2 encodes an FAD dependent oxidase and is similar to *tropB* present in the tropolone biosynthetic cluster of *Talaromyces stipitatus*[9] and also similar to genes which encode proteins responsible for the oxidative dearomatisation reactions during azaphilone biosynthesis.[12] We hypothesised that this protein might hydroxylate sorbicillin **10a** and form the key intermediate sorbicillinol **5a**. In order to address this possibility we cloned *orf*2 from cDNA (to remove introns) and expressed it in *E. coli*. This produced the expected protein of 48.6 kDa (50.3 KDa including his₆ tag). Denaturation of the protein released the cofactor which was shown to be FAD by LCMS analysis. The catalytic activity of the protein was then examined by incubating it with purified sorbicillin **10a** (see ESI†) or dihydrosorbicillin **10b** and NADPH in 50 mM phosphate buffer pH 8.0. The reaction was followed by LCMS (Scheme 3) over a period of 30 min and this clearly showed the consumption of dihydrosorbicillin **10b** and the formation of a new peak which corresponded (m/z $[M]H^+ = 251$) with dihydrosorbicillinol **5b** which is chemically unstable (like its unsaturated analogue **5a**).[19] We converted **5b** to its *bis*-acetate **13** *in situ* by treating the reaction mixture with an excess of acetic anhydride and pyridine, followed by automated RP-HPLC purification of the product and full structural elucidation by 1D and 2D NMR (600 MHz) and HRMS.
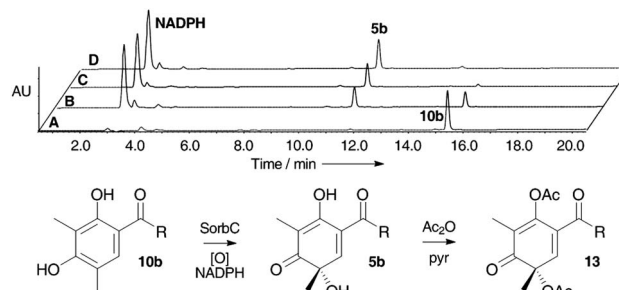
Scheme 3 Conversion of dihydrosorbicillin **10b** to dihydrosorbicillinol **5b** and its acetylation. DAD HPLC traces (200–400 nm): (A), 0 min (before addition of NADPH); (B), 10 min; (C), 20 min; (D), 30 min. R = $CH_2CH_2CH{=}CHCH_3$.

Discussion

Domain analysis of the two PKS genes suggest that they should work together to form the first intermediate in the pathway to the sorbicillinoids, sorbicillin **10a** itself. Similar sets of HR- and NR-iPKS are known to act together to form the resorcylic acid lactone[20] and azaphilone[21] classes of fungal polyketides both *in vivo* and *in vitro*.[22] We hypothesise that *orf*4 encodes a HR-iPKS which synthesises a triketide which is then passed to the *orf*3 encoded NR-iPKS which extends it three more times and methylates it twice. Alternatively, the *orf*3-encoded PKS may produce a tetraketide which is then chain extended and methylated twice by the *orf*4 PKS. We thus propose to name *orf*4 and *orf*3 *sorbA* and *sorbB* respectively. The SorbA iPKS is homologous with other HR-iPKS where the ER domain is functional, (the ER domain is 32% identical to the ER domain of the mammalian fatty acid synthase)[23] but here it appears that for the biosynthesis of sorbicillin **10a** the ER does not operate, although during the biosynthesis of dihydrosorbicillin **10b** it presumably operates correctly during the second cycle of chain

modification. Although SorbA contains a *C*MeT domain, this appears to be inactive similar to the compactin *C*MeT domain.[24] In contrast, SorbB must methylate the growing polyketide chain twice before release and cyclisation. This is the same as observed in MOS where methylation has been shown to be processive,[25] however MOS methylates once during 3 extension cycles, whereas SorbB appears to methylate during both extension cycles. This is similar to the austinol NR-iPKS which methylates twice, but which has a hydrolytic release domain resulting in the formation of a carboxylic acid product.[26]

After chain extension and methylation by SorbB, reductive release of **6** catalysed by the SorbB Red-domain would form an aldehyde **12** which would rapidly cyclise to form sorbicillin **10a**. The *orf*2 encoded FAD-dependent oxidase (hereafter referred to as SorbC) then hydroxylates sorbicillins **10a** and **10b** to form sorbicillinols **5a** and **5b** respectively, forming the precursors of **1a** and **1b**, and all other subsequent sorbicillins and sorbicillinoids.

Conclusions

The proposal that SorbA and SorbB produce sorbicillin **10a**, followed by oxidative dearomatisation performed by SorbC is fully consistent with the results of the previous isotopic labelling experiments since **10a** itself is not symmetrical. The proposed pathway is simpler, more direct and supported by biochemical evidence. The new pathway is also consistent with the biosynthesis of a number of other fungal secondary metabolites including the tropolones[9] and azaphilones[21] where non-reducing iPKS release simple aromatic aldehydes which are oxidatively dearomatised in early steps on pathways to more complex metabolites. Similar to the resorcyclic acid lactone[20] and azaphilone[21] classes of fungal polyketides, the sorbicillins also require the action of two PKS – a HR-iPKS to provide the starter unit for the NR-iPKS which catalyses the final chain extension steps. Further work to delineate the later steps of sorbicillactone biosynthesis are ongoing.

Acknowledgements

Notes and references

1 A. M. Harned and K. A. Volp, *Nat. Prod. Rep.*, 2011, **28**, 1790–1810.

2 N. Abe, T. Arakawa and A. Hirota, *Chem. Commun.*, 2002, 204–205.

3 N. Abe, T. Murata and A. Hirota, *Biosci., Biotechnol., Biochem.*, 1998, **62**, 2120–2126.

4 D. Li, F. Wang, X. Xiao, Y. Fang, T. Zhu, Q. Gu and W. Zhu, *Tetrahedron Lett.*, 2007, **48**, 5235–5238.

5 K. Sugaya, H. Koshino, Y. Hongo, K. Yasunaga, J.-I. Onose, K. Yoshikawa and N. Abe, *Tetrahedron Lett.*, 2008, **49**, 654–657; L. S. Trifonov, A. S. Dreiding, L. Hoesch and D. M. Rast, *Helv. Chim. Acta*, 1981, **64**, 1843–1846.

6 R. J. Cox, *Org. Biomol. Chem.*, 2007, **5**, 2010–2026.

7 N. Abe, T. Arakawa, K. Yamamoto and A. Hirota, *Biosci., Biotechnol., Biochem.*, 2002, **66**, 2090–2099.

8 A. M. Bailey, R. J. Cox, K. Harley, C. M. Lazarus, T. J. Simpson and E. Skellam, *Chem. Commun.*, 2007, 4053–4055.

9 J. Davison, A. Al Fahad, M. Cai, Z. Song, S. Y. Yehia, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7642–7647.

10 G. Bringmann, T. A. M. Gulder, G. Lang, S. Schmitt, R. Stöhr, J. Wiese, K. Nagel and J. F. Imhoff, *Mar. Drugs*, 2007, **5**, 23–30; G. Bringmann, G. Lang, T. A. M. Gulder, H. Tsuruta, J. Mühlbacher, K. Maksimenka, S. Steffens, K. Schaumann, R. Stöhr and J. Wiese, *Tetrahedron*, 2005, **61**, 7252–7265.

11 L. S. Trifonov, J. H. Bieri, R. Prewo, A. S. Dreiding, L. Hoesch and D. M. Rast, *Tetrahedron*, 1983, **39**, 4243–4256.

12 J. M. Winter, M. Sato, S. Sugimoto, G. Chiou, N. K. Garg, Y. Tang and K. Watanabe, *J. Am. Chem. Soc.*, 2012, **134**, 17900–17903.

13 L. Du and L. Lou, *Nat. Prod. Rep.*, 2010, **27**, 255–278.

14 S. Kroken, N. Glass, J. Taylor, O. C. Yoder and B. G. Turgeon, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 15670–15675.

15 K. Sakai, H. Kinoshita, T. Shimizu and T. Nihira, *J. Biosci. Bioeng.*, 2008, **106**, 466–472.

16 A. A. Salamov and V. V. Solovyev, *Genome Res.*, 2000, **10**, 516–522.

17 M. A. van den Berg, R. Albang, K. Albermann, J. H. Badger, J. M. Daran, A. J. Driessen, C. Garcia-Estrada, N. D. Fedorova, D. M. Harris, W. H. Heijne, V. Joardar, J. A. Kiel, A. Kovalchuk, J. F. Martín, W. C. Nierman, J. G. Nijland, J. T. Pronk, J. A. Roubos, I. J. van der Klei, N. N. van Peij, M. Veenhuis, H. von Döhren, C. Wagner, J. Wortman and R. A. Bovenberg, *Nat. Biotechnol.*, 2008, **26**, 1161–1168.

18 Other *P. chrysogenum* strains are known sorbicillinoid producers, see: R. F. Miller and S. Huang, *J. Antibiot.*, 1995, 520–521.

19 N. Abe, O. Sugimoto, K.-I. Tanji and A. Hirota, *J. Am. Chem. Soc.*, 2000, **122**, 12606–12607.

20 I. Gaffoor and F. Trail, *Appl. Environ. Microbiol.*, 2006, **72**, 1793–1799; H. Zhou, J. Zhan, K. Watanabe, X. Xie and Y. Tang, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 6249–6254.

21 Y.-M. Chiang, E. Szewczyk, A. D. Davidson, N. Keller, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2009, **131**, 2965–2970.

22 Y. Xu, P. Espinosa-Artiles, V. Schubert, Y.-M. Xu, W. Zhang, M. Lin, A. L. Gunatilaka, R. Süssmuth and I. Molnár, *Appl. Environ. Microbiol.*, 2013, **79**, 2038–2047; H. Zhou, K. Qiao, Z. Gao, J. C. Vederas and Y. Tang, *J. Biol. Chem.*, 2010, **285**, 41412–41421; Z. Gao, J. Wang, A. K. Norquay, K. Qiao,

Y. Tang and J. C. Vederas, *J. Am. Chem. Soc.*, 2013, **135**, 1735–1738.

23 M. Leibundgut, T. Maier, S. Jenni and N. Ban, *Curr. Opin. Struct. Biol.*, 2008, **18**, 714–725.

24 Y. Abe, T. Suzuki, C. Ono, K. Iwamoto, M. Hosobuchi and H. Yoshikawa, *Mol. Genet. Genomics*, 2002, **267**, 636–646.

25 K. M. Fisch, E. Skellam, D. Ivison, R. J. Cox, A. M. Bailey, C. M. Lazarus and T. J. Simpson, *Chem. Commun.*, 2010, **46**, 5331–5333.

26 H.-C. Lo, R. Entwistle, C.-J. Guo, M. Ahuja, E. Szewczyk, J.-H. Hung, Y.-M. Chiang, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2012, **134**, 4709–4720.