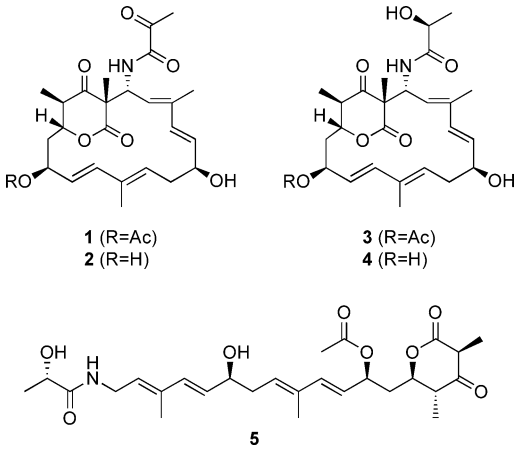
DOI: 10.1002/cbic.201100474

An Additional Dehydratase-Like Activity is Required for Lankacidin Antibiotic Biosynthesis

Jeroen S. Dickschat,[b, c] Olivia Vergnolle,[b, d] Hui Hong,[b] Stephen Garner,[b] Susanna R. Bidgood,[b] Hannah C. Dooley,[b] Zixin Deng,[a] Peter F. Leadlay,*[b] and Yuhui Sun*[a, b]

Lankacidins are 17-membered macrocyclic polyketides produced by several strains of *Streptomyces* bacteria, and show both antitumour[1] and antibacterial[2] activity. Lankacidin biosynthesis in *Streptomyces rochei* 7434AN4 is encoded by a gene cluster housed on a large linear plasmid, immediately adjacent to the gene cluster for the 14-membered macrolide antibiotic lankamycin.[3] Lankacidin and lankamycin inhibit protein synthesis synergistically by binding to adjacent binding sites on the ribosome.[4]

Lankacidins A (**1**) and C (**2**) contain an *N*-pyruvoyl side chain, whereas the lankacidinols A (**3**) and C (**4**) carry an *N*-lactoyl moiety (Scheme 1).[5] Their biosynthesis involves an unusual nonribosomal peptide synthetase (NRPS)/modular polyketide synthase (PKS) hybrid system, which shares many features with those non-canonical PKS multienzymes from diverse actinomycetes, myxobacteria and marine sources, in which chain-extension building blocks (almost always malonyl-CoA) are supplied to the growing polyketide chain via discrete acyltransferases (*trans*-AT systems).[6] In such systems, the co-linearity of canonical PKSs between the order (and enzyme domain content) of successive modules and the chemistry of the product, no longer holds: modules may either lack the expected domains, or contain tandem copies of domains; modules may be split between different polypeptides; and domain order within modules may be altered. Furthermore, lankacidin is one of a select group of PKSs where certain PKS modules catalyse a single cycle of chain extension (modular behaviour), while others act several times over (iterative behaviour, or "stuttering").[7]



Scheme 1. Structures of lankacidins (**1**–**4**) and of the biosynthetic intermediate LC-KA05 (**5**).

None of the lankacidin PKS modules (LkcA, C, F, G) are reported to exhibit dehydratase activity,[3,8] even though the antibiotic contains four carbon-carbon double bonds. A discrete dehydratase enzyme (LkcB) (*trans*-DH) is proposed instead to catalyse all the required dehydration steps. Based on an extensive genetic analysis,[3,8] Kinashi and colleagues have proposed the mechanism for lankacidin biosynthesis shown in Scheme 2, in which the PKS multienzyme subunit LkcA initiates synthesis and PKS subunit LkcC functions five times in succession, before the polyketide chain is passed on to subunits LkcF and finally LkcG, where chain termination and concomitant cyclization take place to generate intermediate **5**. In support of this mechanism, a mutated strain containing a version of the lankacidin PKS in which LkcF has been translationally fused to LkcG is reported to retain the ability to produce lankacidins.[8b]

A recent phylogenetic approach has proved very useful in predicting the structures of metabolites whose biosynthesis is controlled by *trans*-AT PKS systems.[9] Analysis of the ketosynthase (KS) domain sequences of 138 modules from *trans*-AT PKS systems (including the lankacidin PKS)[9c] has revealed that, unlike the KS domains of canonical PKSs, those of *trans*-AT PKS systems group strongly together according to their substrate specificity for the growing polyketide chain. The lankacidin PKS is a puzzling exception: the predicted specificities of the individual lankacidin KS domains match the mechanism shown in Scheme 2 in only three out of the eight extension cycles.[9a] The exceptional nature of the lankacidin PKS is underlined by a more recent estimate,[9d] for those KS domains that can be assigned to a specific clade, that the phylogenetic prediction is accurate in 96% of cases. Either KS2 of LkcC is unusually broad in its specificity, or the mechanism shown in Scheme 2 needs

[a] *Prof. Z. Deng, Prof. Y. Sun*
Key Laboratory of Combinatory Biosynthesis and Drug Discovery (Ministry of Education), and School of Pharmaceutical Sciences
Wuhan University
185 East Lake Road, Wuhan 430071 (China)
E-mail: yhsun@whu.edu.cn

[b] *Dr. J. S. Dickschat,[+] Dr. O. Vergnolle,[+] Dr. H. Hong, Dr. S. Garner, S. R. Bidgood, H. C. Dooley, Prof. P. F. Leadlay, Prof. Y. Sun*
Department of Biochemistry, University of Cambridge
80 Tennis Court Road, Cambridge CB2 1GA (UK)
E-mail: pfl10@cam.ac.uk

[c] *Dr. J. S. Dickschat[+]*
Institut für Organische Chemie, Technische Universität Braunschweig
Hagenring 30, 38106 Braunschweig (Germany)

[d] *Dr. O. Vergnolle[+]*
Biology Department, Brooklyn College
City University of New York
2900 Bedford Avenue, Brooklyn, NY 11210 (USA)

[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201100474.

Scheme 2. Proposed mechanism of polyketide chain assembly on the lankacidin NRPS-PKS in which only LkcC functions iteratively.[3,8] The individual domains are labelled as follows: C, condensation; A, adenylation; PCP, peptidylcarrier protein; KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; DH, dehydratase; MT, methyltransferase; TE, thioesterase/cyclase. The location of the putative DH domain is indicated by a question mark. The dotted circles denote where the discrete *trans*-AT LkcD is proposed to dock in each PKS module. Italicised labels denote modules where the respective enzymatic domains are presumed present, but do not appear to act. The MT of LkcC is hypothesised to act in the final extension cycle.

revision. An alternative arrangement for the lankacidin PKS, in which both LkcC and LkcF act iteratively in the sequence LkcA-LkcC-LkcF-LkcC-LkcF-LkcC-LkcG, would certainly match the phylogenetic predictions of KS specificity much more closely, but this mechanism (see the Supporting Information) is apparently ruled out by the fact that an LkcF-LkcG fusion retains activity.[8b]

We have re-examined the predicted domain structure of the lankacidin PKS, using the protein sequences derived from our independent sequencing of the biosynthetic genes on a giant linear plasmid (denoted here pSRV, Supporting Information) in the lankacidin-producing strain of *Streptomyces rochei* var. *volubilis* ATCC 21250. Since the sequence proved to be almost identical to that of the PKS from the better-studied *S. rochei* it is not analyzed here in detail. However, our analysis revealed that a significant (~400 amino acids) terminal portion of LkcC had not previously been ascribed any enzymatic function. Only the N-terminal 140 residues of this region have counterparts in LkcF, and this portion contains a fragment of an acyltransferase (AT) domain that presumably represents the docking site for the *trans*-acting AT LkcD in both multienzymes. The remaining ~260 residues are unique to LkcC.

We now report that the unique C-terminal portion of this unassigned region of LkcC houses a domain which has both the predicted protein fold and the conserved active-site features of an authentic dehydratase; and we have found that a

mutant strain of *S. rochei* var. *volubilis* in which the predicted active site histidine residue is specifically mutated to alanine no longer produces lankacidins. On this evidence, the dehydratase-like activity is therefore essential, and must be taken into account in any future revised mechanism for the lankacidin PKS. In parallel, we have carried out site-specific mutagenesis on the active site histidine residue of the previously-identified *trans*-DH (LkcB), and have shown that the analogous active site histidine is required for *trans*-DH activity, confirming that the activity of this enzyme is also required for lankacidin biosynthesis.

In the discrete homodimeric dehydratases of bacterial fatty acid biosynthesis, such as FabA from *Escherichia coli*[10] or FabZ from *Pseudomonas aeruginosa*,[11] each subunit (~170 amino acids) adopts a single "hot-dog" fold and contributes residues to two identical active sites at the intersubunit interface in the homodimer. The X-ray crystal structures reveal that the essential active site His residue is hydrogen-bonded to a backbone carbonyl, and a catalytic water molecule is positioned by hydrogen bonds to both a backbone amide and to an acidic residue (Asp84 in FabA) which in turn is oriented by hydrogen-bonding to a nearby polar residue (Gln88 in FabA). We first searched both the LkcB sequence (276 amino acids), and the ~260 amino acids at the C terminus of LkcC for secondary structural elements using the program PHYRE.[12] This revealed that both LkcB and the C-terminal region of LkcC comprise a double "hot-dog" fold, as seen in X-ray crystal structures of dehydratase domains from vertebrate fatty acid synthase[13] and from modular PKS systems,[14] and also found in enoyl-CoA hydratases[15] and thioesterases.[16] In these enzymes, a single active site is created at the interface of the two "hot-dog" folds. Mutagenesis of individual residues in the fatty acid synthase has confirmed that the N-terminal hot-dog contributes the essential His residue, while the positioning Asp and Gln residues are contributed by the C-terminal hot-dog.[17] Sequence alignment of the putative LkcC dehydratase domain with the lankacidin *trans*-DH (LkcB) and the erythromycin PKS DH domain[14a] showed (Figure 1) that the LkcC domain lacks the two β-strands from the N terminus of the first hot-dog. However, the key residues that are conserved in fatty acid synthase[13] and PKS[14,18] DH domain active sites are all present here: His1616 (predicted to be essential) in the N-terminal hot-dog, and Asp1754 and Gln1758 contributed by the C-terminal hot-dog. These residues, also conserved in LkcB, are indicated by triangles in the alignment of Figure 1.

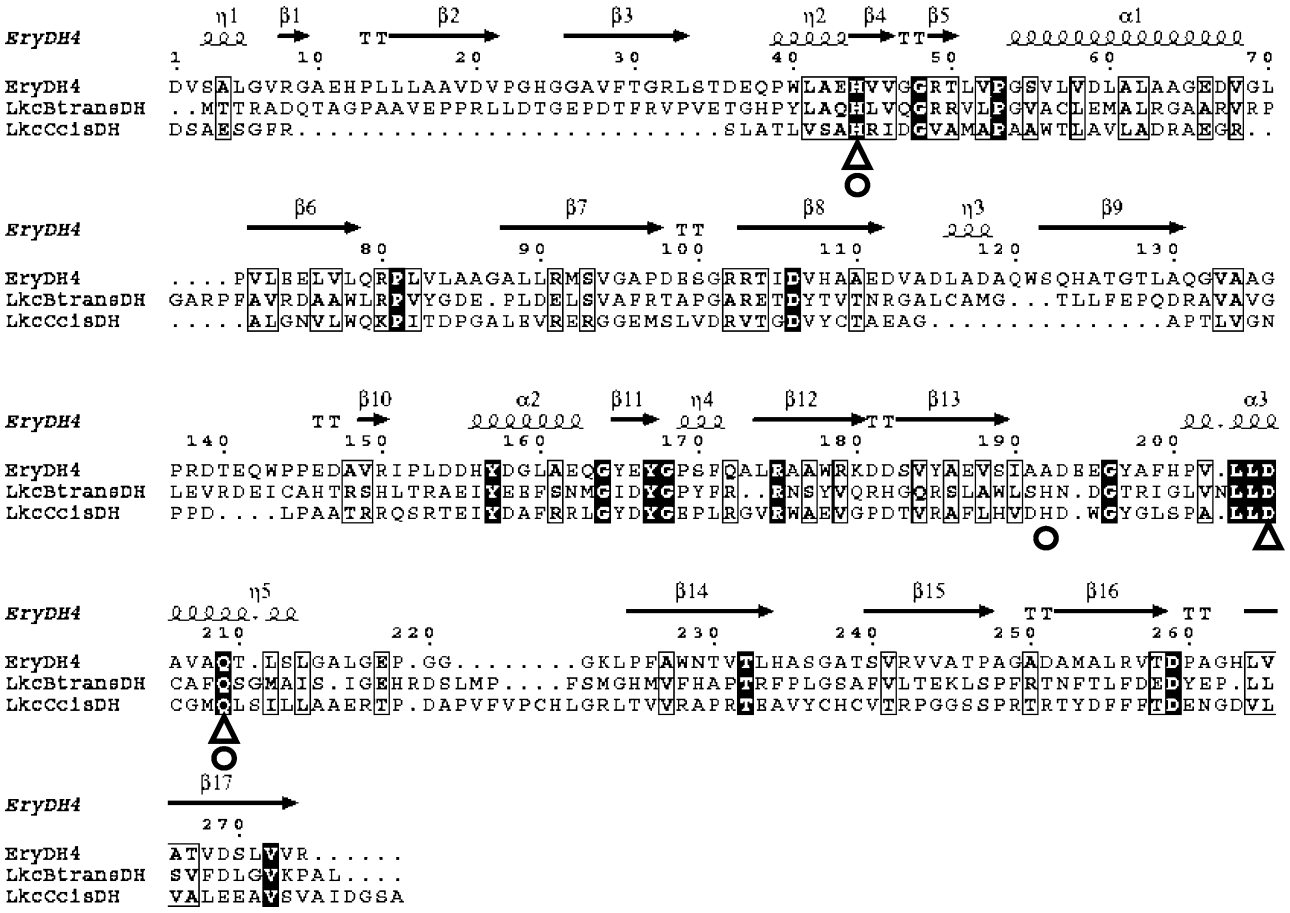


Figure 1. Sequence alignment of the putative dehydratase (DH) domain at the C terminus of lankacidin PKS multienzyme LkcC with the *trans*-DH LkcB and the DH domain from the erythromycin PKS. The secondary structure is also shown of the double hot-dog fold of the erythromycin DH domain, deduced from the X-ray crystal structure (3L6).[14a] Dehydratase active site residues are indicated by triangles. Mutated residues are indicated by circles.

Table 1. Lankacidin production by *S. rochei* var. *volubilis* strains bearing dehydratase mutations.

Strain	Dehydratase mutation	Relative lankacidin production[a]
ATCC 21250 wild-type	none	1.0 ± 0.7
ΔlkcB	In-frame deletion in *trans*-DH LkcB	no production
ΔlkcB complemented with plasmid pYH52 expressing wild-type *trans*-DH	none	1.1 ± 0.5
ΔlkcB complemented with plasmid pYH68	*trans*-DH (LkcB) His42Ala	no production
ΔlkcB complemented with plasmid pOV14	*trans*-DH (LkcB) Gln205Ala	0.4 ± 0.4
ATCC 21250 (H1616A LkcC)	DH domain (LkcC) His1616Ala	no production
ATCC21250 (H1742A LkcC)	DH domain (LkcC) His1742Ala	0.4 ± 0.2
ATCC21250 (Q1758A LkcC)	DH domain (LkcC) Gln1758Ala	0.8 ± 0.5

[a] The total lankacidin production was calculated by addition of the respective peak integrals for all four lankacidins **1**–**4**. Values represent means of at least three assays ± SD. The average of analyses of wild-type cultures was set to 1.0. The relative lankacidin production of all mutant strains was likewise calculated from the average value of the summed peak integrals of all four lankacidins. See the Supporting information for further details.

Active site mutants in each of the *trans*-DH enzyme and the putative DH domain of LkcC were created as follows: for LkcB, an in-frame deletion was engineered. This ΔlkcB strain was unable to synthesise lankacidins (Table 1). When an intact copy of the *lkcB* gene was introduced on an integrative plasmid under the control of a constitutive promoter, lankacidin production was restored. These results confirmed previously-published work in which a plasmid-borne copy of intact LkcB restored lankacidin production to a mutant disrupted in LkcB.[8b] We also introduced versions of LkcB altered at either His42 or Gln205 into the ΔlkcB strain. As shown in Table 1, lankacidin production was abolished in His42Ala LkcB consistent with His42 being an essential active site residue. The Gln205Ala LkcB mutant in contrast did show partially restored lankacidin production. To create mutations in the DH domain of LkcC, versions of this portion of the *lkcC* gene carrying the desired specific mutation were introduced into *S. rochei* var. *volubilis* by conjugation, and mutants were identified in which the required double crossover event had occurred. Candidate colonies were confirmed by PCR and sequencing. Lankacidin production was abolished in His1616Ala LkcC, but was substantially maintained in His1742Ala LkcC and Gln1758Ala versions of LkcC (Table 1). His1742 is remote from the active site and not expected to play a catalytic role in DH activity, while Gln1758, although conserved, is not an essential residue in the fatty acid synthase DH active site.[13] The results in Table S2 in the Supporting Information are therefore fully consistent with the idea that the dehydratase-like activity of the putative DH domain of LkcC is required for lankacidin biosynthesis.

It remains to be elucidated how the newly-identified domain would fit into the mechanism of lankacidin biosynthesis shown in Scheme 2. No evidence was obtained for the production of partially-assembled or variant polyketide chains, either in wild-type or mutant cultures, which might have shed additional light on the chain-building process. It is an appealing idea that the DH activities of LkcB and LkcC might help control the switch of the lankacidin PKS from an iterative to a modular mode of operation, as the growing polyketide chain reaches a critical chain length. Dehydratases have been previously implicated in switching the direction of biosynthesis: most clearly for FabA of *E. coli* fatty acid synthase, which operates to divert biosynthesis towards unsaturated fatty acid production only once the chain length reaches C10.[19] Clear chain-length selectivity has also been documented for the vertebrate fatty acid synthase DH[17] and for the (3*R*)-hydroxyacyl-ACP dehydratase of *Mycobacterium tuberculosis*.[17] Alternatively, the DH domain might catalyse a hydration reaction rather than a dehydration.[21] Further experiments are clearly needed to test the currently-proposed mechanism of chain extension on the lankacidin PKS. If indeed the KS domain of LkcC is uniquely broad in its specificity, as implied by Scheme 2, this domain would be potentially valuable in the future engineering of *trans*-AT PKS systems to produce novel bioactive metabolites.[6]

Experimental Section

Experimental details are given in the Supporting Information online.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council UK (BBSRC) through a project grant to P.F.L. and a PhD studentship to S.G., as well as a Royal Society China Fellowship and the National Basic Research Program of China to Y.S., by an EU Marie Curie Early Stage Researcher award to O.V., and by a fellowship from the Deutsche Akademie der Naturforscher Leopoldina to J.S.D.

Keywords: antibiotics · biosynthesis · dehydratase · polyketides · streptomyces

[1] K. Ootsu, T. Matsumoto, S. Harada, T. Kishi, *Cancer Chemother. Rep.* **1975**, *59*, 919 – 928.
[2] K. Tsuchiya, T. Yamazaki, Y. Takeuchi, T. Oishi, *J. Antibiot.* **1971**, *24*, 29 – 41.
[3] S. Mochizuki, K. Hiratsu, M. Suwa, T. Ishii, F. Sugino, K. Yamada, H. Kinashi, *Mol. Microbiol.* **2003**, *48*, 1501 – 1510.
[4] T. Auerbach, I. Mermershtain, C. Davidovich, A. Bashan, M. Belousoff, I. Wekselman, E. Zimmerman, L. Xiong, D. Klepacki, K. Arakawa, H. Kinashi, A. S. Mankin, A. Yonath, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1983 – 1988.
[5] S. Harada, T. Kishi, *Chem. Pharm. Bull.* **1974**, *22*, 99 – 108.
[6] J. Piel, *Nat. Prod. Rep.* **2010**, *27*, 996 – 1047.
[7] a) N. Gaitatzis, B. Silakowski, B. Kunze, G. Nordsiek, H. Blocker, G. Hofle, R. Muller, *J. Biol. Chem.* **2002**, *277*, 13082 – 13090; b) J. He, C. Hertweck,

ChemBioChem **2005**, *6*, 908 – 912; c) C. Olano, B. Wilkinson, S. J. Moss, A. F. Brana, C. Mendez, P. F. Leadlay, J. A. Salas, *Chem. Commun.* **2003**, 2780 – 2782; d) S. J. Moss, C. J. Martin, B. Wilkinson, *Nat. Prod. Rep.* **2004**, *21*, 575 – 593.

[8] a) K. Arakawa, F. Sugino, K. Kodama, T. Ishii, H. Kinashi, *Chem. Biol.* **2005**, *12*, 249 – 256; b) S. Tatsuno, K. Arakawa, H. Kinashi, *J. Antibiot.* **2007**, *60*, 700 – 708; c) S. Tatsuno, K. Arakawa, H. Kinashi, *Biosci. Biotechnol. Biochem.* **2009**, *73*, 2712 – 2719.

[9] a) J. Moldenhauer, X. Chen, R. Borriss, J. Piel, *Angew. Chem.* **2007**, *119*, 8343 – 8345; *Angew. Chem. Int. Ed.* **2007**, *46*, 8195 – 8197; b) T. Nguyen, K. Ishida, H. Jenke-Kodama, E. Dittmann, C. Gurgui, T. Hochmuth, S. Taudien, M. Platzer, C. Hertweck, J. Piel, *Nat. Biotechnol.* **2008**, *26*, 225 – 233; c) H. Irschik, M. Kopp, K. Weissman, K. Buntin, J. Piel, R. Müller, *ChemBioChem* **2010**, *11*, 1840 – 1849; d) R. Teta, M. Gurgui, E. J. N. Helfrich, S. Künne, A. Schneider, G. Van Echten-Deckert, A. Mangoni, J. Piel, *ChemBioChem* **2010**, *11*, 2506 – 2512.

[10] M. Leesong, B. S. Henderson, J. R. Gillig, J. M. Schwab, J. L. Smith, *Structure* **1996**, *4*, 253 – 264.

[11] M. S. Kimber, F. Martin, Y. Lu, S. Houston, M. Vedadi, A. Dharamsi, K. M. Fiebig, M. Schmid, C. O. Rock, *J. Biol. Chem.* **2004**, *279*, 52593 – 52602.

[12] L. A. Kelley, M. J. Sternberg, *Nat. Protoc.* **2009**, *4*, 363 – 371.

[13] T. Maier, S. Jenni, N. Ban, *Science* **2007**, *316*, 288 – 290.

[14] a) A. Keatinge-Clay, *J. Mol. Biol.* **2008**, *384*, 941 – 953; b) D. L. Akey, J. R. Razelun, J. Tehranisa, D. H. Sherman, W. H. Gerwick, J. L. Smith, *Structure* **2010**, *18*, 94 – 105.

[15] K. M. Koski, A. M. Haapalainen, J. K. Hiltonen, T. Glumoff, *J. Mol. Biol.* **2005**, *345*, 1157 – 1169.

[16] J. Li, U. Derewenda, Z. Dauter, S. Smith, Z. S. Derewenda, *Nat. Struct. Biol.* **2000**, *7*, 555 – 559.

[17] S. Pasta, A. Witkowski, A. K. Joshi, S. Smith, *Chem. Biol.* **2007**, *14*, 1377 – 1385.

[18] O. Vergnolle, F. Hahn, A. Baerga-Ortiz, P. F. Leadlay, J. N. Andexer, *ChemBioChem* **2011**, *12*, 1011 – 1014.

[19] R. J. Heath, C. O. Rock, *J. Biol. Chem.* **1996**, *271*, 27 795 – 27 801.

[20] E. L. Sacco, A. S. Covarrubias, H. M. O'Hare, P. Carroll, N. Eynard, T. A. Jones, T. Parish, M. Daffé, K. Bäckbro, A. Quémard, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14628 – 14633.

[21] A. J. Woo, W. R. Strohl, N. D. Priestley, *Antimicrob. Agents Chemother.* **1999**, *43*, 1662 – 1668.

Received: July 25, 2011
Published online on September 27, 2011

