Natural Products

Deutsche Ausgabe: DOI: 10.1002/ange.201507097
Internationale Ausgabe: DOI: 10.1002/anie.201507097

Development of Genetic Dereplication Strains in *Aspergillus nidulans* Results in the Discovery of Aspercryptin

*Yi-Ming Chiang, Manmeet Ahuja, C. Elizabeth Oakley, Ruth Entwistle, Anabanadam Asokan, Christoph Zutz, Clay C. C. Wang, and Berl R. Oakley**

Abstract: *To reduce the secondary metabolite background in Aspergillus nidulans and minimize the rediscovery of compounds and pathway intermediates, we created a "genetic dereplication" strain in which we deleted eight of the most highly expressed secondary metabolite gene clusters (more than 244,000 base pairs deleted in total). This strain allowed us to discover a novel compound that we designate aspercryptin and to propose a biosynthetic pathway for the compound. Interestingly, aspercryptin is formed from compounds produced by two separate gene clusters, one of which makes the well-known product cichorine. This raises the exciting possibility that fungi use differential regulation of expression of secondary metabolite gene clusters to increase the diversity of metabolites they produce.*

Genetic and molecular genetic approaches that up-regulate secondary metabolite (SM) production have dramatically facilitated the discovery of new fungal natural products.[1] However, these approaches often result in complex metabolite profiles owing to the production of a large number of compounds, including pathway intermediates. We thus devised a strategy we call "genetic dereplication" whereby the discovery of novel compounds is simplified by eliminating the major known SM biosynthetic pathways in *A. nidulans*, thereby reducing the complexity of SM profiles such that novel compounds are more easily detected. Elimination of highly expressed biosynthetic pathways might also reserve pools of SM precursors such as acetyl-CoA and malonyl-CoA for pathways expressed at low levels.

We have previously shown that we can delete entire SM biosynthetic gene clusters in *A. nidulans* while recycling a selectable marker.[2] We therefore set out to sequentially delete as many of the major SM clusters as possible. In this study, we engineered a strain (LO8030, genotype in Table S1) in which the clusters responsible for the biosynthesis of the following major SM producing clusters are deleted: sterigmatocystin,[3] the emericellamides,[4] asperfuranone,[5] monodictyphenone,[6] terrequinone,[7] F9775A and B,[8] asperthecin,[9] and both portions of the split cluster that makes austinol and dehydroaustinol.[10] Deletion of these clusters reduced the size of the *A. nidulans* genome by 244,061 base pairs (bp). LO8030 is surprisingly healthy (Figure 1). It forms aerial hyphae upon prolonged incubation, but it is not difficult to work with or to transform. As anticipated, production of the products of the deleted SM clusters was eliminated under all conditions, thus resulting in a low SM background.

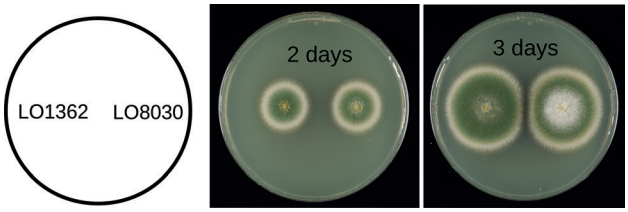


Figure 1. Growth of the parental strain LO1362 and a daughter strain, LO8030, in which eight secondary metabolism gene clusters have been deleted. Growth is on complete medium. After two days, the multicluster deletion strain grows as well as its parental strain. After three days, radial growth for the multicluster deletion strain is the same as for the parent. However, an increase in aerial hyphae makes the center of the colony white.

On glucose minimal medium (GMM) plates, sterigmatocystin (**1**) and the minor SMs terrequinone (**2**) and emericellamindes (**3**–**7**) were produced in a parental strain but were eliminated in LO8030 (Figure 2 A, traces i–iv). However, one MS-detectable peak (**8**, MW = 933) was identified specifically in LO8030, albeit at low yield (Figure 2 A, traces v and vi). The molecular formula of **8** was predicted to be $C_{47}H_{79}N_7O_{12}$ based on its HRESI-MS data, thus suggesting that **8** could be a metabolite biosynthesized by a non-

[*] Dr. Y. M. Chiang, Dr. C. C. C. Wang
Department of Pharmacology and Pharmaceutical Sciences
School of Pharmacy, University of Southern California
Los Angeles, CA 90089 (USA)
Dr. Y. M. Chiang
Department of Pharmacy
Chia Nan University of Pharmacy and Science
Tainan, 71710 (Taiwan)
Dr. M. Ahuja,[+] C. E. Oakley, Dr. R. Entwistle, Dr. B. R. Oakley
Department of Molecular Biosciences, University of Kansas
Lawrence, KS 66045 (USA)
E-mail: boakley@ku.edu
Dr. C. Zutz
Institute for Milk Hygiene, University of Veterinary Medicine Vienna
Vienna, 1210 (Austria)
Dr. C. C. C. Wang
Department of Chemistry, University of Southern California
Los Angeles, CA 90089 (USA)
Dr. A. Asokan
Del Shankel Structural Biology Center, University of Kansas, Lawrence, KS 66045, (USA)
[+] Current address: Industrial Biotechnology Division, Reliance Technology Group, Reliance Industries Limited, Reliance Corporate Park Thane Belapur Road, Ghansoli, Navi Mumbai 400701 (India)

Supporting information and ORCID(s) from the author(s) for this article are available on the WWW under http://dx.doi.org/10.1002/anie.201507097.

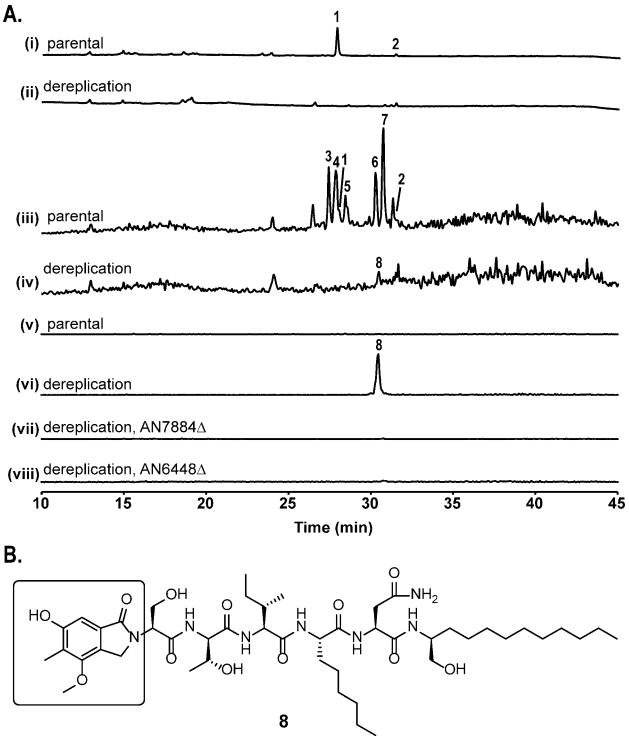
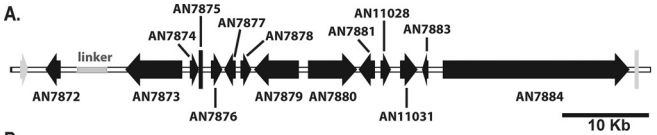


Figure 2. A) HPLC metabolic profiles of the parental and genetic dereplication strains. UV/Vis total scan from 200–600 nm (i, ii), total ion current (iii, iv), and extracted ion chromatogram at *m/z* 934 (v–viii) are shown. All strains were grown on GMM plates. **1** is sterigmatocystin, **2** is terrequinone, **3**–**7** are emericellamides C, D, A, E, and F, respectively, and **8** is aspercryptin. B) The chemical structure of aspercryptin (**8**). The box indicates the cichorine moiety. Strains: parental = LO1362; dereplication = LO8030; dereplication, AN7884Δ = LO8096; dereplication, AN6448Δ = LO9345. Genotypes are given in Table S1.

ribosomal peptide synthetase (NRPS) pathway. The MS/MS fragment data for **8** indicate that it contains several non-essential amino acids (see below). This is a common feature of metabolites produced from NRPS pathways since these mega-enzymes can incorporate unusual amino acids during the elongation steps. We were able to obtain 3.0 mg of **8** from a large-scale culture. Because of the large size and relatively low solubility of the compound, we took advantage of the high resolution and sensitivity of an 800 MHz spectrometer equipped with a cryoprobe to obtain NMR spectral data. From ^{13}C-^{1}H heteronuclear single quantum coherence total correlation spectroscopy (HSQC-TOCSY), six possible spin systems, namely threonine, isoleucine, aspartic acid/asparagine, serine, lysine-like, and an unknown spin system could be identified (Figure S1 in the Supporting Information). Analysis of the ^{1}H, ^{13}C, and 2D-NMR data (Figures S2, S7–S9, and Table S2) as well as MS/MS fragment data (Figure S3) revealed that compound **8** is a cichorine-derived hexapeptide with 2-aminododecanol and 2-aminocaprylic acid residues as shown in Figure 2B (for details of structure elucidation, see the Supporting Information). We named compound **8** aspercryptin.

The structure of aspercryptin indicates that it is an NRPS product. There is only one NRPS gene in the genome of *A. nidulans* that contains six adenylation (A) domains, AN7884.[11] Andersen et al.[12] found that the genes AN7872 to AN7884 are co-regulated, thus suggesting that they may form an SM gene cluster. Bioinformatic analyses of genes in this cluster indicate that they encode a fatty acid synthase (FAS), aminotransferases, a P450 hydroxylase, a short-chain dehydrogenase, and transporters, most of which can reasonably be predicted to be involved in the biosynthesis of aspercryptin (Figure 3).



B.

Gene	Predicted length gene/protein	Conserved domain/ Predicted function	Accession number of nearest BLAST hit (origin)	Protein coverage (%)	Similarity/ identity (%)
AN7872 *atnN*	1876/571	fungal specific transcription factor	KIA75995 (*A. ustus*)	98	63/52
AN7873 *atnM*	6500/2111	fatty acid synthase subunit beta	XP_002544941 (*U. reesii* 1704)	93	65/49
AN7874 *atnL*	1011/306	thioesterase-like superfamily	GAM42266 (*T. cellulolyticus*)	88	60/46
AN7875 *atnK*	524/125	hypothetical protein	none[a]		
AN7876 *atnJ*	1327/359	amino acid aminotransferase	XP_001273101 (*A. clavatus* NRRL 1)	99	76/61
AN7877 *atnI*	1305/383	RTA1 superfamily	KIA75996 (*A. ustus*)	88	83/70
AN7878 *atnH*	1237/396	amino acid aminotransferase	XP_006962560 (*T. reesei* QM6a)	93	74/58
AN7879 *atnG*	5099/1547	ABC transporter	EPS26282 (*P. oxalicum*114-2)	91	61/42
AN7880 *atnF*	5577/1663	fatty acid synthase subunit alpha	XP_001273102 (*A. clavatus* NRRL 1)	98	69/54
AN7881 *atnE*	1845/518	cytochrome P450	XP_001273104 (*A. clavatus* NRRL 1)	97	69/54
AN11028 *atnD*	1181/333	short chain dehydrogenase	KIN00738 (*O. maius* Zn)	95	66/47
AN11031 *atnC*	1985/508	MFS transporter	XP_008086094 (*G. lozoyensis* ATCC 20868)	92	57/38
AN7883 *atnB*	702/156	YCII-related domain	KFY79913 (*P. pannorum*VKM F-103)	81	50/65
AN7884 *atnA*	21390/7015	NRPS	CEJ80515 (*T. hemipterigena*)	94	34/52

Figure 3. A) Schematic representation of the AN7884 (*atn*) cluster. Each arrow indicates the direction of transcription and relative sizes of the open reading frames (ORFs). This cluster spans contig 134 and 135. The linker represents the connection between contig 134 and 135. B) The deduced function of each ORF and the amino acid sequence coverage and similarity/identity, as compared with the BLAST search of the NCBI nonredundent protein database. Genetic symbols we have assigned to members of the aspercryptin cluster are given below the AspGD gene numbers in column 1. [a]The nearest hit to AN7875 only has 40% of protein coverage and was considered to be irrelevant.

To determine whether this gene cluster encodes the aspercryptin biosynthetic pathway, we deleted the NRPS gene AN7884 and, indeed, deletion of AN7884 eliminated the production of aspercryptin (Figure 2A, trace vii), thus indicating that aspercryptin is synthesized by the AN7884 cluster. We designate AN7884 as *atnA* and give the other genes of the cluster the designations shown in Figure 3. Aspercryptin was found to be produced at a higher titer when LO8030 was grown on yeast extract, agar, and glucose (YAG) plates. Under these conditions, a new aspercryptin derivative **9** was identified, albeit at lower intensity than aspercryptin (Figure S4A). The MW of compound **9** is 14 Da less than that of aspercryptin and the MS/MS fragment data indicate that **9** is aspercryptin with Ile replaced by Val (Figure S5). To gain

insights into the functions of *atn* genes, we deleted 12 additional genes in the gene cluster (from *atnB* to *atnM*) and analyzed the metabolites produced on YAG plates (Figure S4). Deletion of the two fatty acid synthase genes *atnF* and *atnM* eliminated more than 99.5% of **8** production, thus indicating that the two FAS subunit genes are necessary for aspercryptin biosynthesis. The tiny amount of **8** detected in the FAS deletion strains suggests that fatty acids generated by other endogenous FASs can be incorporated, albeit at a significantly lower rate. Deletion of the dehydrogenase (*atnD*) and cytochrome P450 (*atnE*) genes abolished the production of **8**. AtnD and AtnE are likely involved in the oxidation of the α-carbon of the fatty acids. Deletion of the amino acid aminotransferase genes *atnJ* and *atnH* eliminated approximately 70% and 40% of **8** production, respectively (Figure S4B), thus suggesting that AtnJ and AtnH might be able to compensate for each other but with lower efficiency, or that other endogenous aminotransferases could partially compensate for their functions. Deletion of *atnI*, an RTA1 superfamily protein, eliminated more than 70% of **8** production. Deletion of three genes, an MFS transporter (*atnC*), an ABC transporter (*atnG*), and a thioesterase-like gene (*atnL*) did not alter the yield of **8** significantly. Interestingly, deletion of *atnB* and *atnK* decreased the titer of **8** by more than 80% but increased the titer of **9** more than six-fold, thus suggesting that AtnB and AtnK might participate in the selection of Ile versus Val residue in module 3 of AtnA. Without AtnB or AtnK, module 3 of AtnA prefers to incorporate Val and produces **9**. Analysis of the deletion mutants allowed us to propose a biosynthetic pathway for aspercryptin (Figure 4). The fatty acid synthase subunits (AtnF and AtnM), dehydrogenase (AtnD), cytochrome P450 (AtnE), and amino acid transferases (AtnJ and AtnH) could be involved in the biosynthesis of two unusual amino acids, 2-aminocaprylic and 2-aminododecanoic acids, which are then activated and incorporated into the growing peptide chain by AtnA. Inspection of the domain architecture of AtnA revealed that the second module has an epimerise (E) domain, thus suggesting that the second residue of **8** could be D-allo-Thr. Indeed, Marfey's analysis[13] showed that **8** does contain D-allo-Thr and the L form of the remaining amino acids (Figure S6). After condensation of the hexapeptide of **8**, the C-terminal reductase (R) domain might be involved in the reductive release and production of the aldehyde hexapeptide.[14] Further reduction would generate **8**.

Our proposed pathway postulates that aspercryptin synthesis uses chichorine-Ser as a precursor. Cichorine is a phytotoxin originally discovered as a natural product produced by members of the genus *Alternaria*. We have previously identified the cichorine biosynthetic cluster in *A. nidulans* and found that the non-reducing polyketide synthase AN6448 is essential for cichorine biosynthesis.[15] If cichorine-Ser produced by the cichorine pathway is the precursor of aspercryptin, deletion of AN6448 should eliminate aspercryptin production. We consequently deleted AN6448 and found that production of aspercryptin is indeed eliminated (Figure 2A, trace viii).

Our data demonstrate that genetic dereplication strains are valuable for discovering novel compounds from unknown

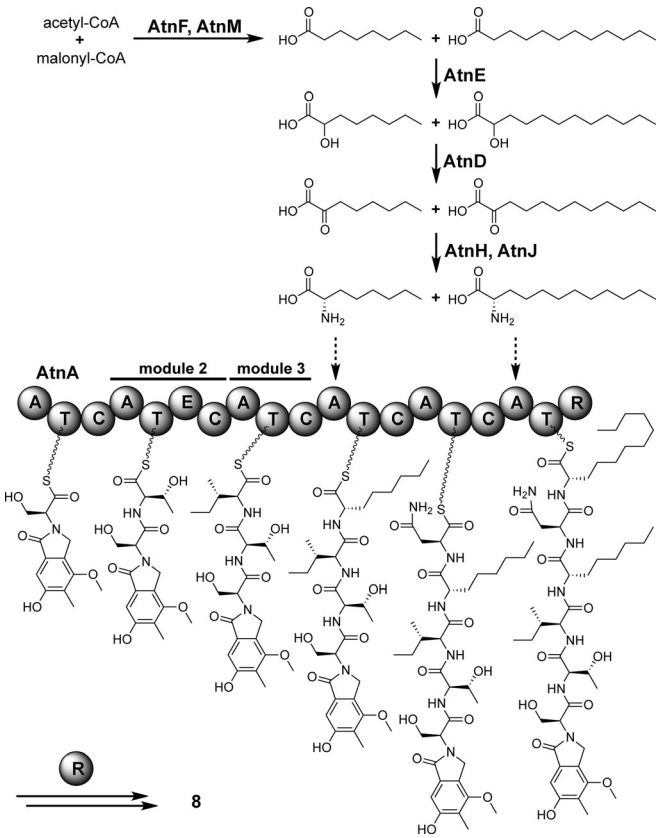


Figure 4. Proposed biosynthetic pathway for aspercryptin (**8**).

biosynthesis pathways. The fact that the SM background is lowered dramatically may be particularly valuable for detecting compounds produced through heterologous expression of SM genes from other fungi and in searches for genes that regulate cryptic SM clusters. They may have an additional advantage in the production of compounds or proteins for human and animal use because they are incapable of producing major toxic SMs such as sterigmatocystin (**1**). Our data also lead us to the fascinating conclusion that aspercryptin is made of building blocks from two distinct clusters that are physically separated in the genome, the AN6448 (cichorine) cluster and the *atn* cluster. This raises the interesting possibility that differential gene regulation could result in the production of cichorine or aspercryptin. This, in turn, raises the exciting possibility that *A. nidulans* (and by inference other fungi) may use differential regulation of SM gene cluster expression to expand their repertoire of natural products and tailor their SM arsenal to achieve maximum competitive advantage. Forseth et al. found that a somewhat similar situation occurs with the *lna* and *lnb* clusters in *Aspergillus flavus*.[16] These clusters are closely related to each other and contain homologous NRPS-like core biosynthetic genes. They are partially redundant in that when the *lna* NRPS-like gene (*lnaA*) is deleted, the *lnb* NRPS-like gene (*lnbA*) and perhaps other genes in the *lnb* cluster can modify intermediates produced by the *lna* cluster to produce two of the major products of the *lna* cluster, albeit in greatly reduced amounts. Forseth et al. also found evidence that intermediates

of the *lna* biosynthetic pathway are modified by the *lnb* biosynthetic pathway. The situation with the cichorine and aspercryptin clusters differs from the *lna* and *lnb* situation in significant ways, however. The cichorine and aspercryptin clusters are not homologous and the core biosynthetic enzyme for the cichorine pathway is a PKS whereas the core biosynthetic gene for the aspercryptin cluster is an NRPS. The cichorine and aspercryptin pathways are not redundant. Rather they function in a serial fashion, with the cichorine cluster able to produce an important metabolite on its own, and the *atn* cluster using that product to produce aspercryptin.

Experimental Section

Deletion of entire SM clusters was carried out as previously described.[2] Most clusters were deleted by using the loop-out recombination procedure based on the procedure of Takahashi et al.[17] Correct deletion of clusters was verified by diagnostic PCR amplifications with primers outside of the ends of the clusters. *Atn* cluster genes were deleted and deletions were verified by diagnostic PCR using the methods of Oakley et al.[18] Growth media and conditions, as well as extraction conditions, are detailed in the Supporting Information.

Acknowledgements

This work was supported by the H. L. Snyder Medical Foundation, the Irving S. Johnson Fund of the University of Kansas Endowment and by grant P01-GM084077 from the National Institute of General Medical Sciences. We thank the Bio-NMR Core Facility of the Del Shankel Structural Biology Center at the University of Kansas for 800 MHz NMR data.

Keywords: aspergillus · biosynthesis · genetic dereplication · natural products · NRPS

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 1662–1665
Angew. Chem. **2016**, *128*, 1694–1697

[1] N. P. Keller, G. Turner, J. W. Bennett, *Nat. Rev. Microbiol.* **2005**, *3*, 937; O. Bayram, G. H. Braus, *FEMS Microbiol. Rev.* **2012**, *36*, 1; A. A. Brakhage, *Nat. Rev. Microbiol.* **2013**, *11*, 21; J. H. Yu, N. Keller, *Annu. Rev. Phytopathol.* **2005**, *43*, 437; A. M. Calvo, *Fungal Genet. Biol.* **2008**, *45*, 1053; D. Hoffmeister, N. P. Keller, *Nat. Prod. Rep.* **2007**, *24*, 393; E. M. Fox, B. J. Howlett, *Curr. Opin. Microbiol.* **2008**, *11*, 481; C. Hertweck, *Nat. Chem. Biol.* **2009**, *5*, 450; D. R. Georgianna, N. D. Fedorova, J. L. Burroughs, A. L. Dolezal, J. W. Bok, S. Horowitz-Brown, C. P. Woloshuk, J. Yu, N. P. Keller, G. A. Payne, *Mol. Plant Pathol.* **2010**, *11*, 213; W. Yin, N. P. Keller, *J. Microbiol.* **2011**, *49*, 329; J. F. Sanchez, A. D. Somoza, N. P. Keller, C. C. Wang, *Nat. Prod. Rep.* **2012**, *29*, 351.

[2] Y. M. Chiang, C. E. Oakley, M. Ahuja, R. Entwistle, A. Schultz, S. L. Chang, C. T. Sung, C. C. C. Wang, B. R. Oakley, *J. Am. Chem. Soc.* **2013**, *135*, 7720.

[3] D. W. Brown, T. H. Adams, N. P. Keller, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 14873.

[4] Y. M. Chiang, et al., *Chem. Biol.* **2008**, *15*, 527.

[5] Y. M. Chiang, E. Szewczyk, A. D. Davidson, N. Keller, B. R. Oakley, C. C. C. Wang, *J. Am. Chem. Soc.* **2009**, *131*, 2965.

[6] Y. M. Chiang, E. Szewczyk, A. D. Davidson, R. Entwistle, N. P. Keller, C. C. C. Wang, B. R. Oakley, *Appl. Environ. Microbiol.* **2010**, *76*, 2067.

[7] J. W. Bok, D. Hoffmeister, L. A. Maggio-Hall, R. Murillo, J. D. Glasner, N. P. Keller, *Chem. Biol.* **2006**, *13*, 31–37.

[8] J. F. Sanchez, Y. M. Chiang, E. Szewczyk, A. D. Davidson, M. Ahuja, C. E. Oakley, J. W. Bok, N. Keller, B. R. Oakley, C. C. C. Wang, *Mol. Biosyst.* **2010**, *6*, 587.

[9] E. Szewczyk, Y. M. Chiang, C. E. Oakley, A. D. Davidson, C. C. C. Wang, B. R. Oakley, *Appl. Environ. Microbiol.* **2008**, *74*, 7607.

[10] H. C. Lo, R. Entwistle, C. J. Guo, M. Ahuja, E. Szewczyk, J. H. Hung, Y. M. Chiang, B. R. Oakley, C. C. Wang, *J. Am. Chem. Soc.* **2012**, *134*, 4709.

[11] H. von Döhren, *Fungal Genet. Biol.* **2009**, *46*, S45.

[12] M. R. Andersen, et al., *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E99.

[13] P. Marfey, *Carlberg Res. Commun.* **1984**, *49*, 591.

[14] N. Gaitatzis, B. Kunze, R. Müller, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11136; L. Du, L. Lou, *Nat. Prod. Rep.* **2010**, *27*, 255.

[15] J. F. Sanchez, R. Entwistle, D. Corcoran, B. R. Oakley, C. C. C. Wang, *MedChemComm* **2012**, *3*, 997.

[16] R. R. Forseth, S. Amaike, D. Schwenk, K. J. Affeldt, D. Hoffmeister, F. C. Schroeder, N. P. Keller, *Angew. Chem. Int. Ed.* **2013**, *52*, 1590; *Angew. Chem.* **2013**, *125*, 1632.

[17] T. Takahashi, F. J. Jin, M. Sunagawa, M. Machida, Y. Koyama, *Appl. Environ. Microbiol.* **2008**, *74*, 7684.

[18] C. E. Oakley, H. Edgerton-Morgan, B. R. Oakley, *Methods Mol. Biol.* **2012**, *944*, 143.

Received: July 31, 2015
Revised: October 4, 2015
Published online: November 13, 2015

