Molecular BioSystems

Cite this: *Mol. BioSyst.*, 2011, **7**, 2459–2469

www.rsc.org/molecularbiosystems

PAPER

Cloning and functional analysis of the naphthomycin biosynthetic gene cluster in *Streptomyces* sp. CS†

Yingying Wu,‡[ab] Qianjin Kang,‡[a] Yuemao Shen,[b] Wenjin Su[b] and Linquan Bai*[a]

Received 28th January 2011, Accepted 28th April 2011
DOI: 10.1039/c1mb05036b

Naphthomycins (NATs) are 29-membered naphthalenic ansamacrolactam antibiotics with antimicrobial and antineoplastic activities. Their biosynthesis starts from 3-amino-5-hydroxy-benzoic acid (AHBA). By PCR amplification with primers for AHBA synthase and amino-dehydroquinate (aDHQ) synthase, a genomic region containing orthologs of these genes was identified in *Streptomyces* sp. CS. It was confirmed to be involved in naphthomycin biosynthesis by deletion of a large DNA fragment, resulting in abolishment of naphthomycin production. A 106 kb region was sequenced, and 32 complete ORFs were identified, including five polyketide synthase genes, eight genes for AHBA synthesis, and putative genes for modification, regulation, transport or resistance. Targeted inactivation and complementation experiments proved that the halogenase gene *nat1* is responsible for the chlorination of C-30 of NATs. The *nat1* mutant could also be complemented with *asm12*, the halogenase gene of ansamitocin biosynthesis. Likewise, an *asm12* mutant could be complemented with *nat1*, suggesting a similar catalytic mechanism for both halogenases. A putative hydroxylase gene, *nat2*, was also inactivated, whereupon the biosynthesis of NATs was completely abolished with a tetraketide desacetyl-SY4b accumulated, indicating the participation of *nat2* in the formation of the naphthalene ring. The information presented here expands our understanding of the biosynthesis of naphthalenic ansamycins, and may pave the way for engineering ansamacrolactams with improved pharmaceutical properties.

Introduction

The naphthomycins (NATs) (Fig. 1A) belong to the family of ansamycin antibiotics, whose characteristic structure comprises an aromatic moiety bridged at nonadjacent positions by an aliphatic chain.[1] The aromatic moiety can either be a naphthalene or naphthoquinone ring system, as in rifamycin[2] and naphthomycin, or a benzene or benzoquinone ring, as in geldanamycin[3] and ansamitocin.[4] The biosynthesis of ansamycin antibiotics involves the assembly of a polyketide from an unusual starter unit, 3-amino-5-hydroxy-benzoic acid (AHBA),[5] by chain extension with several acetate and propionate units catalyzed by a type I polyketide synthase (PKS). The starter unit AHBA is synthesized *via* a variant of the shikimate pathway for aromatic amino acid biosynthesis (Fig. 1B).[6,7]

NATs are 29-membered ansamacrolactams with the largest ring reported so far in the ansamycin family. Not only carrying the characteristic naphthalene ring and the shifted double bond at C12–13, NATs also have a featured triene and there is chlorination on the AHBA moiety of certain NATs. Moreover, a variety of nuclear substituents, including methyl, hydroxyl, methoxy, chloro and alkylthio groups, are present at the C-30 of NATs (Fig. 1A). They display significant antimicrobial and antineoplastic activities. Naphthomycin A and B are inhibitors of fatty acid synthesis with selective toxicity against prokaryotes,[8] whereas the cytotoxicity of naphthomycin A is exerted through the inhibition of various SH-enzymes, particularly those involved in nucleic acid biosynthesis.[9] NATs have been isolated from different *Streptomyces* strains, including NAT A from *S. diastatochromogenes* Tü 105,[10] NAT B from *S. galbus* subsp. *griseosporueus* Tü 353, NAT C from *S. collinus* Tü 1892,[11] NAT H from *Streptomyces* sp. Y83,[12] and NAT I and J from *Streptomyces* sp. E/784.[13] Recently, an endophytic *Streptomyces*, isolated from the medicinal plant *Maytenus hookeri*, namely *Streptomyces* sp. CS (hereafter called CS), was found to be the producer of naphthomycin A, C, E and K.[14]

Previous isotope-labeled feeding experiments clearly revealed that the backbone of NATs is assembled by a type I PKS from

[a] *State Key Laboratory of Microbial Metabolism, and School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai 200030, China. E-mail: bailq@sjtu.edu.cn; Fax: +86 21 62932418; Tel: +86 21 62932943*
[b] *State Key Laboratory of Stress Cell Biology and School of Life Sciences, Xiamen University, Xiamen 361005, China*
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c1mb05036b
‡ These two authors contributed equally to this work.

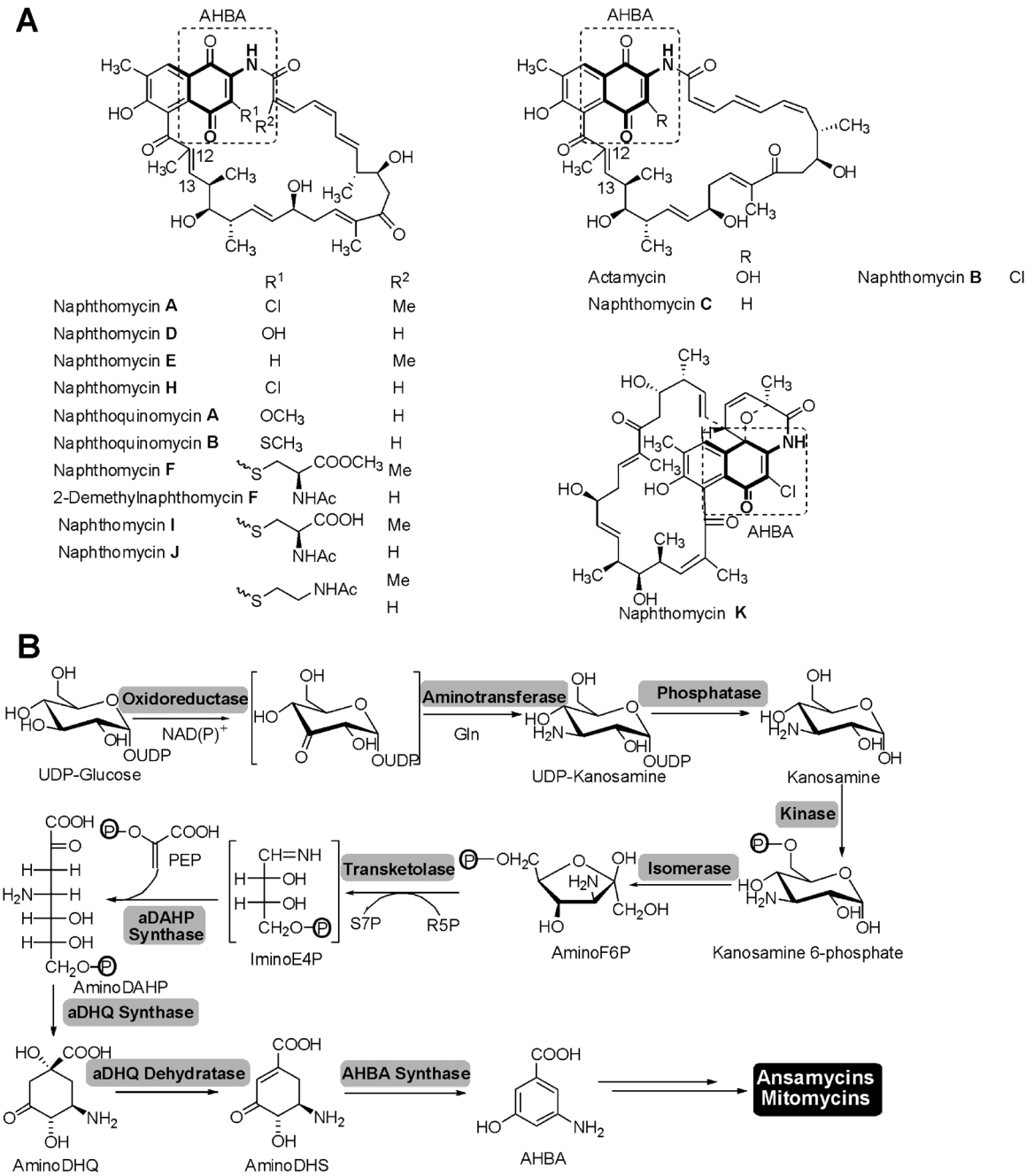

Fig. 1 Naphthomycins and the biosynthesis of AHBA. (A) Structures of naphthomycins. (B) Proposed AHBA biosynthetic pathway. aF6P, amino-fructose-6-phosphate; iminoE4P, imino-erythrose-4-phosphate; aDAHP, amino-3-deoxy-D-arabino-heptulosonate-7-phosphate; aDHQ, amino-dehydroquinate; aDHS, amino-dehydroshikimic acid; AHBA, 3-amino-5-hydroxy-benzoic acid.

seven propionate and six acetate chain extender units,[15] using AHBA as the starter unit. Through Southern hybridization with *rifK*, encoding the AHBA synthase of rifamycin biosynthesis[7] as a probe, a portion of PKS and two complete sets of AHBA biosynthetic genes were cloned from *S. collinus* Tü 1892.[16] However, the involvement of these genes in naphthomycin biosynthesis has not been confirmed genetically or biochemically. Herein, we report the cloning and sequencing of the naphthomycin biosynthetic gene cluster (*nat*) from CS, and the investigation of a halogenase gene and a hydroxylase gene required for naphthomycin biosynthesis.

Results

Cloning and identification of the naphthomycin biosynthetic gene cluster

The only reaction in the proposed biosynthetic pathway of AHBA, which has no analogy in the normal shikimate pathway, is the aromatization of aminoDHS by a postulated AHBA synthase. The AHBA synthase gene then served as a probe allowing the cloning of gene clusters for the biosynthesis of many AHBA-derived antibiotics, such as ansamitocin,[17] geldanamycin,[18] and mitomycin.[19] Since AHBA synthase is

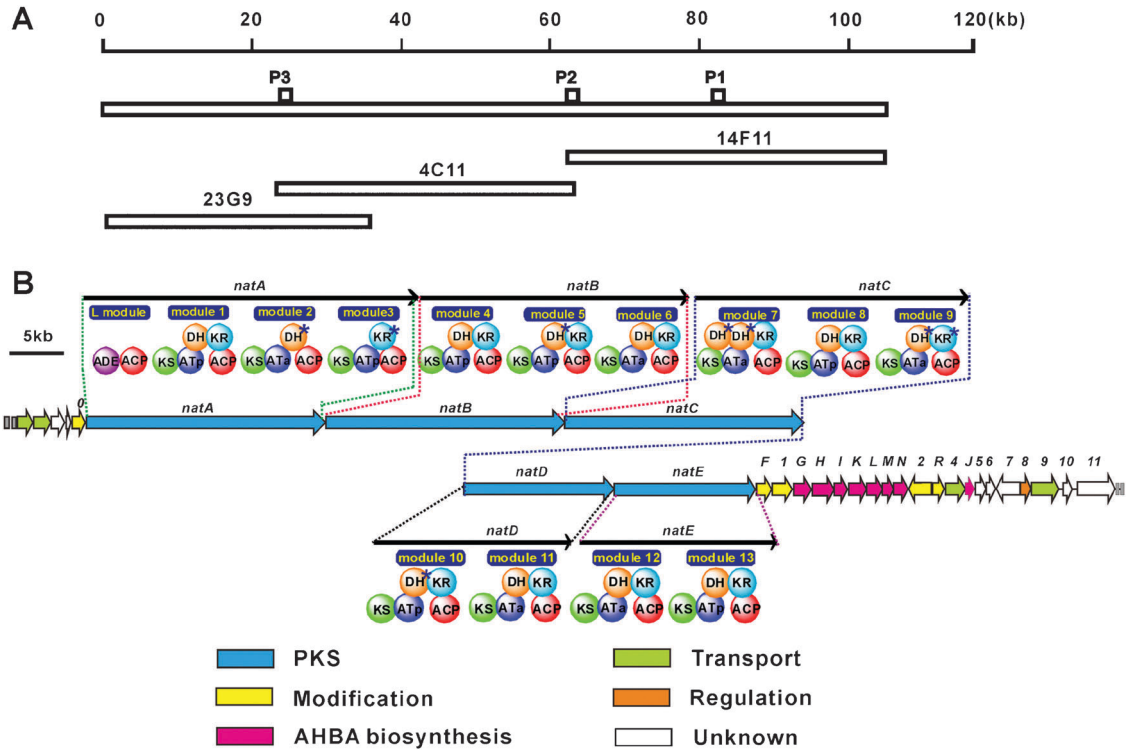

Fig. 2 Gene organization of the naphthomycin biosynthetic gene cluster. (A) overlapping inserts of 3 fosmid clones used for sequencing. P1, PCR probe used for screening of the genomic library; P2 and P3, PCR probes for sequential chromosome walking of naphthomycin biosynthetic gene cluster. (B) open reading frames deduced from the nucleotide sequence in naphthomycin biosynthetic gene cluster. KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein; DH, dehydratase; KR, ketoreductase. DH and KR domains labeled with * are assumed to be redundant.

highly conserved in the biosynthesis gene clusters of different ansamycins, degenerate primers for AHBA synthase were used to clone the naphthomycin gene cluster from *Streptomyces* sp. CS (Fig. S1, ESI†). Fifteen fosmids were identified from a genomic library of CS, which gave products of the expected size. These were further investigated by PCR using a pair of primers specific for *napC*, the amino-dehydroquinate (aDHQ) synthase gene expected to be involved in naphthomycin C biosynthesis.[16] Eventually, 11 overlapping fosmids including fosmid 14F11 were identified, carrying homologs of AHBA synthase and aDHQ synthase genes.

In order to confirm the involvement of the region covered by these fosmids in naphthomycin biosynthesis, a 7.2 kb region was replaced by the apramycin resistant gene *aac(3)IV* in CS. The mutants displayed a complete loss of naphthomycin production on ISP3 medium, confirming the involvement of this region in naphthomycin biosynthesis (Fig. S2, ESI†). Two more fosmids 4C11 and 23G9 were identified by sequential chromosome walking (Fig. 2A), and sequenced together with 14F11. The three fosmids covered a 106 kb region and contained 32 complete open reading frames (ORFs) (Fig. 2B). The predicted functions of these genes are shown in Table 1.

The *nat* PKS genes

Five modular type I PKS genes, *nat A–E*, are arranged in co-linearity with the predicted assembly of the polyketide backbone, totally encoding one loading module and thirteen extension modules as required for loading of the AHBA starter unit and incorporation of 7 propionate and 6 acetate units. However, two ketoreductase (KR) domains and six dehydratase (DH) domains are predicted to be redundant as deduced from the naphthomycin structure (Fig. 2B and 6).

NatA contains a loading domain, consisting of an acyl carrier protein (ACP) and an adenylation domain (ADE), presumably required for the activation of the AHBA starter unit, and the modules 1–3 for naphthomycin polyketide chain extension. The DH domains of module 1 and module 2 are unusually small with a size of 144 aa and 142 aa, respectively, suggesting both domains are inactive.[20] In the KR domain of module 3, the conserved catalytic Y residue[21] is mutated to S, which could abolish the ketoreductase activity.

NatB contains modules 4–6 for the incorporation of two propionate and one acetate units. However, even though the sequence of the DH domain of module 5 seems functional, no dehydration is required in this extension step, as judged from the structure of naphthomycin. NatC hosts modules 7–9 and is expected to catalyze the incorporation of an acetate, a propionate, and a further acetate unit, respectively. Similar to the DH domain of module 5, the DH domain of module 7 seems functional but is not required in naphthomycin biosynthesis. Moreover, even though the sequences of the KR and DH domains of module 9 seem functional, both of them are not required for this specific extension step.

NatD is a bi-modular protein, presumably for the incorporation of one propionate and one acetate unit.

Table 1 Deduced functions of ORFs in naphthomycin biosynthetic gene cluster

Polypeptide	Amino acids	Proposed function	Sequence similarity (protein, origin)	Identity/ similarity	Accession number
Orf1	567	ATP/GTP-binding protein	SAV_1815, *S. avermitilis* MA-4680	82/91	NP822991
Orf2	711	ATP/GTP-binding protein	SAV_1814, *S. avermitilis* MA-4680	68/79	NP822990
Orf3	47	Conserved hypothetical protein	SSEG_07125, *S. sviceus* ATCC 29083	62/70	YP002202776
Orf4	97	Protoporphyrinogen oxidase	Noca_3672, *Nocardioides* sp. JS614	37/1	YP924859
Orf5	442	Transcriptional regulator	ORF36, *A. mediterranei* S699	40/56	AAS00758
Orf6	211	Acyl carrier protein phosphodiesterase	acpD, *S. avermitilis* MA-4680	43/59	NP825521
Orf0	397	Cytochrome P450-like protein	ORF0, *A. mediterranei* S699	65/79	AAC01709
NatA	5698	PKS	RifA, *A. mediterranei* S699	66/75	ABB86419
Loading		ADE ACP			AAC01710
Module 1		KS AT DH* KR ACP			
Module 2		KS AT DH* ACP			
Module 3		KS AT KR* ACP			
NatB	5339	PKS	RifB, *A. mediterranei* S699	65/74	AAC01711
Module 4		KS AT DH KR ACP			
Module 5		KS AT DH* KR ACP			
Module 6		KS AT DH KR ACP			
NatC	5321	PKS	RifB, *A. mediterranei* S699	58/69	AAC01711
Module 7		KS AT DH* KR ACP			
Module 8		KS AT DH KR ACP			
Module 9		KS AT DH* KR* ACP			
NatD	3621	PKS	RifE, *A. mediterranei* S699	59/69	AAC01714
Module 10		KS AT DH* KR ACP			
Module 11		KS AT DH KR ACP			
NatE	3533	PKS	RifE, *A. mediterranei* S699	58/69	AAC01714
Module 12		KS AT DH KR ACP			
Module 13		KS AT DH KR ACP			
NatF	289	Amide synthase	RifF, *A. mediterranei* S699	51/64	AAC01715
Nat1	447	Halogenase	Asm12, *Ac. pretiosum* ATCC31565	61/73	AAM54090
NatG	353	3-Dehydroquinate synthase	RifG, *A. mediterranei* S699	74/80	AAC01717
NatH	478	DAHP synthase	RifH, *A. mediterranei* S699	58/70	AAF70331
NatI	281	Shikimate/quinate dehydrogenase	RifI, *A. mediterranei* S699	70/78	AAC01719
NatK	386	3-Amino-5-hydroxy benzoic acid synthase	RifK, *A. mediterranei* S699	82/88	AAC01720
NatL	364	Oxidoreductase	RifL, *A. mediterranei* S699	67/78	AAS07754
NatM	231	Phosphatase	RifM, *A. mediterranei* S699	71/83	AAC01721
NatN	306	Glucose kinase	RifN, *A. mediterranei* S699	60/73	AAC01722
Nat2	585	FAD-dependent oxidoreductase	ORF19, *A. mediterranei* S699	66/75	AAG52989
NatR	257	Thioesterase	RifR, *A. mediterranei* S699	6478	AAG52991
Nat4	461	Membrane ion antiporter	ORF7, *A. balhimycina*	58/74	CAC48373
NatJ	146	Aminodehydroquinate dehydratase	RifJ, *A. mediterranei* S699	78/88	AAS07762
Nat5	230	Alpha/beta fold family protein	UBAL2_80490237, *Leptospirillum* sp.	26/44	EAY56956
Nat6	187	Activator of Hsp90 ATPase	RHE_CH02426, *Rhizobium etli* CFN 42	37/54	YP469931
Nat7	476	Metallopeptidase	SAV_3952, *S. avermitilis* MA-4680	55/67	NP825129
Nat8	252	GntR family transcriptional regulator	SAV_3189, *S. avermitilis* MA-4680	72/81	NP824365
Nat9	529	Amino acid permease	YbeC, *Bacillus amyloliquefaciens* FZB42	56/76	YP001419888
Nat10	138	Hydrolase	SAV_3187, *S. avermitilis* MA-4680	71/82	NP824363
Nat11	870	Magnesium or manganese-dependent phosphatase	SAV_3185, *S. avermitilis* MA-4680	70/79	NP824361

* Redundant domains.

The DH domain of module 10 looks normal but apparently redundant. NatE comprises modules 12 and 13 and is expected to be responsible for the incorporation of one acetate and one propionate (as in NAT A) or acetate (as in NAT C) units to complete the polyketide assembly. While the typical propionate-specific motif is YASH for AT,[22] the motif of AT13 is YPSH, which may cause the leaky recognition of the acetate extender unit.

Located immediately downstream of *natE* is *natF*, which encodes a protein with a high homology to RifF, the enzyme responsible for the polyketide chain-release and cyclization of rifamycin biosynthesis.[23] The fully extended polyketide chain of naphthomycin is proposed to be transferred from NatE to NatF, followed by intra-molecular amide bond formation to release a 29-membered pronaphthomycin (Fig. 6).

The *nat* AHBA biosynthetic genes

Eight AHBA biosynthetic genes, *nat G-N*, were identified in the naphthomycin gene cluster from CS (Fig. 2B). Of these, the seven genes *natGHIKLMN* are apparently organized into a single operon, while the putative aDHQ dehydratase gene *natJ* is located 4.2 kb downstream of *natN* (Fig. 1B). These genes had been shown to be involved in naphthomycin biosynthesis through the aforementioned large fragment deletion. The organization and sequence of them show significant homology with other naphthalenic ansamycins, especially rifamycin and naphthomycin C (Fig. S3A, ESI†). According to the well-studied biosynthetic mechanism of AHBA in rifamycin (Fig. 1B), these homologous genes in naphthomycin may carry

out similar functions to catalyze the formation of AHBA from UDP-glucose.[24]

Genes for the post-PKS modifications

Compared to other ansamycins, the post-PKS modifications of naphthomycins are relatively simple. Nat1 shows 61% identity to the halogenase Asm12 for ansamitocin biosynthesis, and is considered to be responsible for the chlorination at C-30 in naphthomycin A formation. Nat2 shows 61% identity with Rif-orf19, a 3-(3-hydroxyphenyl) propionate hydroxylase in rifamycin, and is predicted to be involved in the formation of the naphthalene ring of naphthomycins during the polyketide chain elongation (Fig. 6). Orf0 shows sequence similarity to cytochrome P450 monooxygenase and may be involved in the oxidation steps required for the formation of some naphthomycins.

Inactivation of halogenase gene *nat1*

In order to confirm the possible role of *nat1* for halogenation at C-30, an 8.24 kb BamHI/EcoRI fragment from fosmid 14F11 carrying this gene was cloned. An internal 1344 bp region of *nat1* was replaced by an apramycin resistance gene, resulting in plasmid pJTU3245, with 3.24 kb left-flanking and 3.72 kb right-flanking sequences of the deletion (Fig. 3A). This plasmid was introduced into the wild-type CS strain by conjugation. Five double cross-over mutants (WYY2-1–WYY2-5) were obtained. PCR amplification using Nat1-Det-F and Nat1-Det-R as primers confirmed that they gave the expected 1.48 kb PCR product without NcoI restriction site, while the PCR fragment derived from the wild-type was 1.35 kb in size and could be digested into a 0.82 kb fragment and a 0.53 kb fragment by NcoI (Fig. 3B). LC-MS analysis of the fermentation broth showed that in mutant WYY2 the chlorinated component naphthomycin A

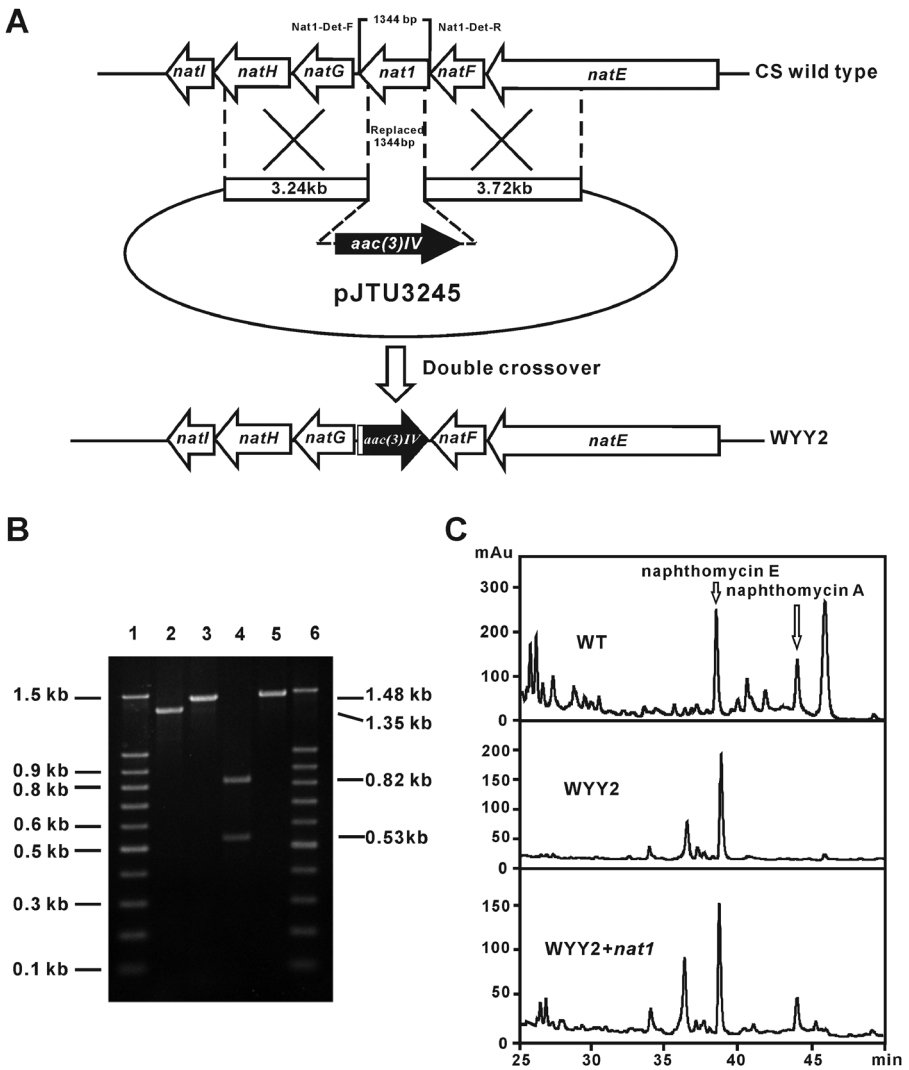

Fig. 3 Inactivation and complementation of halogenase gene *nat1*. (A) schematic representation of the *nat1* inactivation. (B) verification of the genotype through PCR amplification. Lane 1 and 6, DNA ladder; lane 2, PCR product of the wild-type; lane 3, PCR product of the mutant WYY2; lane 4, PCR product of the wild-type digested with NcoI; lane 5, PCR product of the mutant WYY2 digested with NcoI. (C), HPLC analysis of the fermentation broth of the wild-type, *nat1* mutant WYY2, and WYY2 complemented with cloned *nat1*.

was no longer produced, while naphthomycin E, the 30-deschloro-naphthomycin component, could clearly be detected (Fig. 3C).

Cross-complementation between the halogenase genes of naphthomycin and ansamitocin biosynthesis

The inactivation of the *nat1* gene had shown that *nat1* is responsible for the chlorination reaction at C-30 of naphthomycins (Fig. 4A). Nat1 shared high sequence identity (73%) with the previously identified halogenase Asm12 for the biosynthesis of ansamitocin.[17] Inactivation of *asm12* gave mutant HGF054, which no longer produced ansamitocin-P3, but accumulated a series of 19-deschloro-ansamitocin derivatives.[25] The reaction catalyzed by Asm12 was proposed to be modification of ansamitocins (Fig. 4A). Both *nat1* and *asm12* contain two conserved sequence motifs "GXGXXG" and "WXWXIP", typical of $FADH_2$-dependent halogenase (Fig. S4, ESI†).[26,27]

To further investigate whether Nat1 and Asm12 catalyze the similar reaction, a cross complementation experiment using

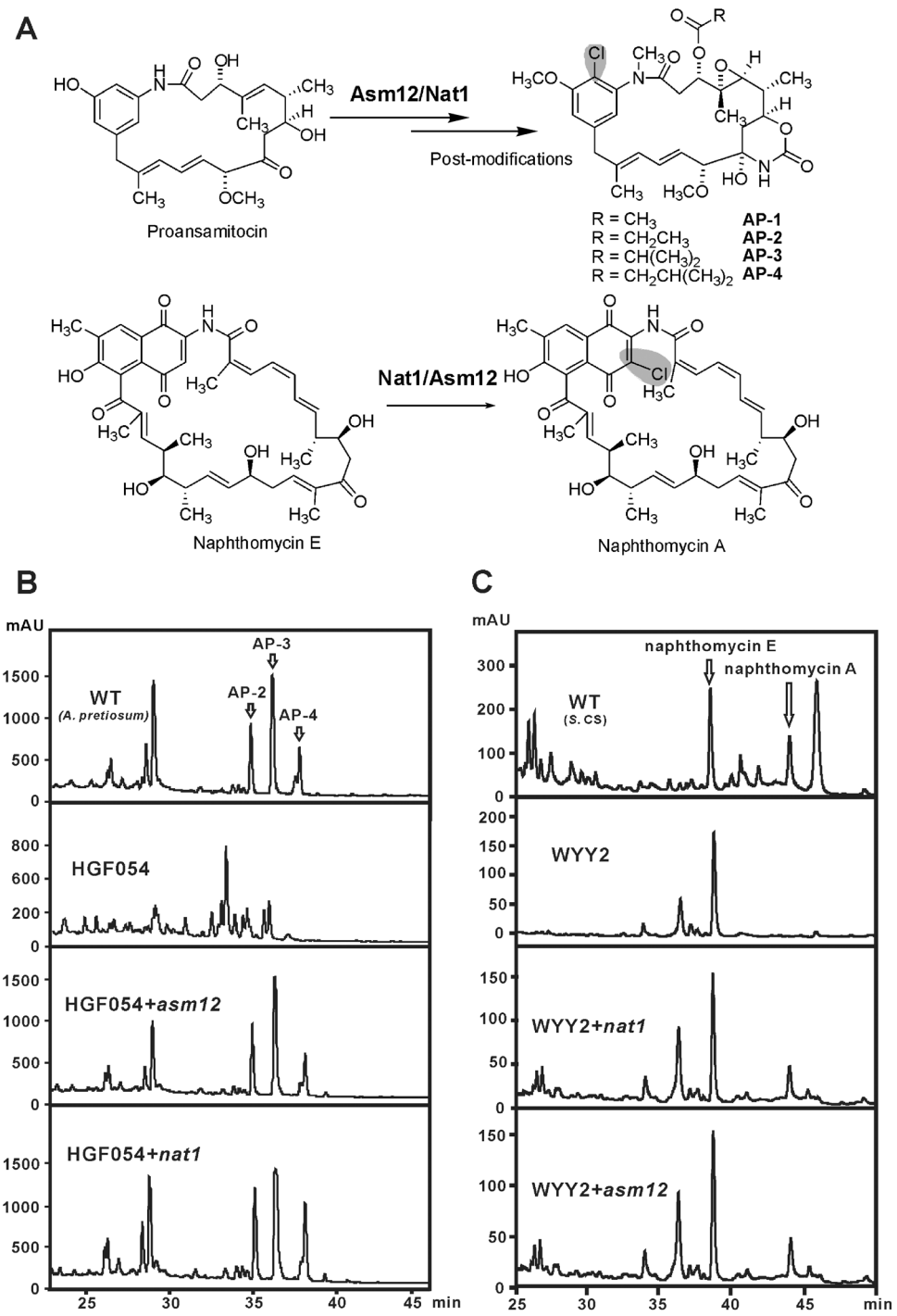

Fig. 4 Cross complementation between halogenase genes *nat1* and *asm12*. (A) proposed chlorination reactions in the biosynthesis of ansamitocins and naphthomycins. (B) complementation of the *asm12* mutant with cloned *asm12* or *nat1*. (C) complementation of the *nat1* mutant with cloned *nat1* or *asm12*.

nat1 and *asm12* was carried out. Plasmids pJTU3243 and pJTU3244, harboring *asm12* and *nat1*, respectively, under the control of the strong constitutive *PermE** promoter, were introduced into the *asm12* mutant HGF054 by conjugation. Both *nat1* and *asm12* complemented the production of ansamitocin AP-2, AP-3 and AP-4 to about 100% of the wild-type level (Fig. 4B). The plasmids pJTU3246 and pJTU3247, carrying *asm12* and *nat1*, respectively, were also introduced into the *nat1* mutant WYY2 similarly, and both complemented the naphthomycin A production to about 20% of the wild-type level (Fig. 4C).

Involvement of Nat2 in naphthalene ring formation

The possible involvement of *nat2* in the formation of a naphthalene ring was examined by gene replacement and complementation. In the plasmid pJTU1289, a 980 bp internal fragment of *nat2* was replaced by the apramycin-resistance gene *aac(3)IV*, inserted between a 3.94 kb left-flanking and a 3.98 kb right-flanking sequence, resulting in plasmid pJTU3250 (Fig. 5A). Four thiostrepton-sensitive and apramycin-resistant ($Thio^S$ Apr^R) recombinant mutants (WYY3-1–WYY3-4) were obtained. Total DNAs from these mutants and the wild-type were used as templates for PCR analysis with primers Nat2-Det-F and Nat2-Det-R. The wild-type gave the expected 1.0 kb PCR product without MluI site, whereas the mutants gave a 1.3 kb PCR product, which could be digested by MluI into a 1.0 kb fragment and a 0.3 kb fragment (Fig. 5B). LC-MS analysis of the fermentation extract of the mutant strains showed a complete abolishment of the production of all the naphthomycins (Fig. 5C). As expected, the mutant produced desacetyl-SY4b

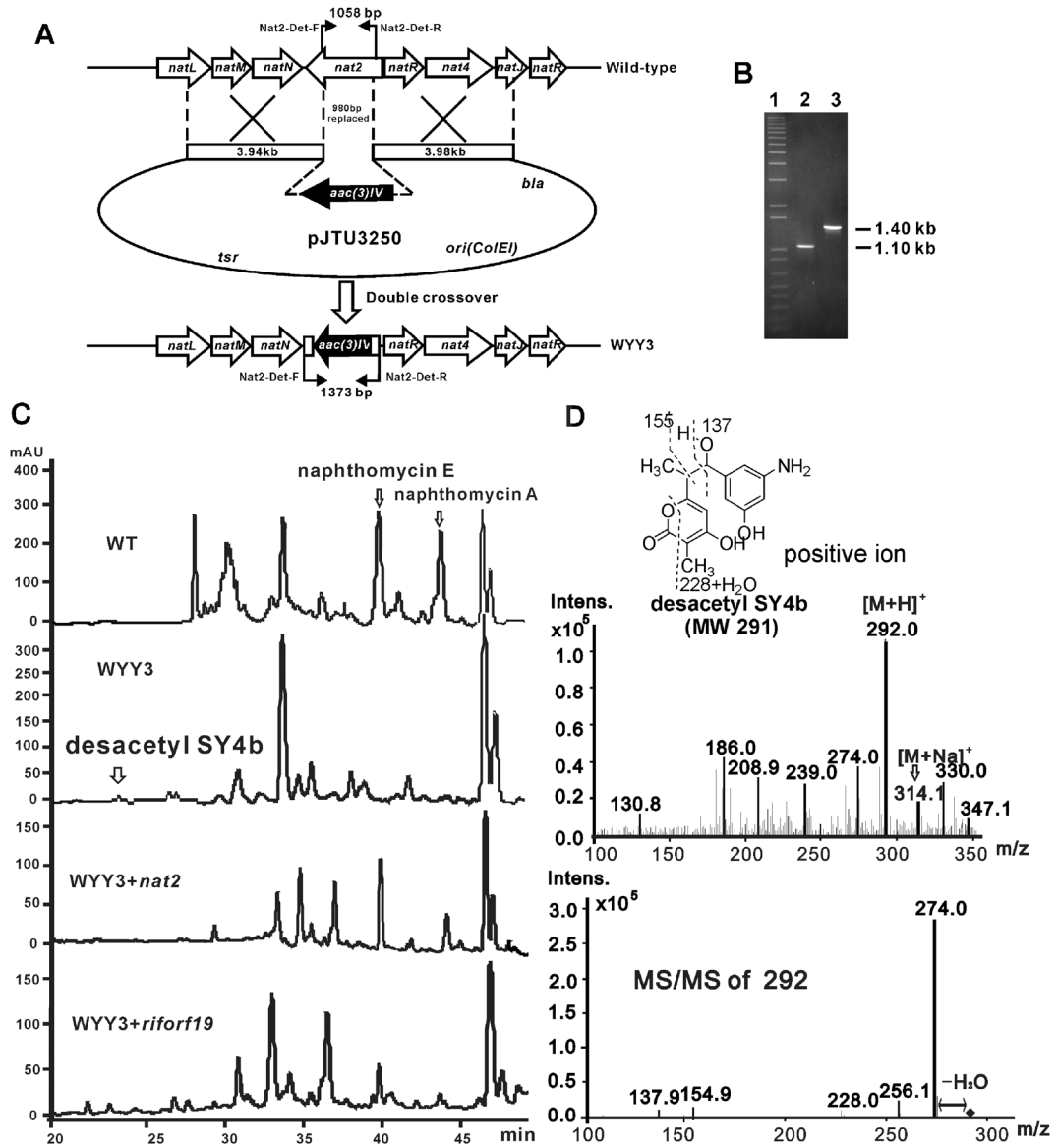

Fig. 5 Inactivation and complementation of the putative hydroxylase gene *nat2*. (A) schematic representation of the *nat2* inactivation. (B) verification of the genotype through PCR amplification. Lane 1, DNA ladder; lane 2, PCR product of the wild-type; lane 3, PCR product of the mutant WYY3. (C) HPLC analysis of the fermentation broth of the wild-type, *nat2* mutant WYY3, and WYY3 complemented with cloned *nat2* or *rif-orf19*. (D) ESI-MS/MS analysis of the mutant product desacetyl-SY4b.

as detected by LC-MS analysis (Fig. 5C & D), which is identical to the compound isolated from the *rif-orf19* mutant derived from rifamycin producer *Amycolatopsis mediterranei* S699.[28,29]

Further support for the involvement of Nat2 in naphthalene ring formation came from complementation of the mutant WYY3 with *nat2* and *rif-orf19*, respectively. LC-MS analysis of the extract of complemented mutants showed the production of naphthomycins was almost fully restored by *nat2*, and to about 30% of the wild-type level by *rif-orf19* (Fig. 5C). Additionally, *nat2* was introduced into HGF065, a proansamitocin producer with mutated *asm12* and *asm21*, to detect if there was any new derivative with a naphthanene ring or oxygenated derivative of proansamitocin. However, no detectable peak for expected compounds was observed in the complemented strains with *nat2* (data not shown).

Discussion

A few biosynthetic gene clusters of ansamycins had been identified and sequenced, including rifamycin and rubradirin in the naphthalenic family, and ansamitocin, geldanamycin, and herbimycin of the benzenic family. Since their AHBA synthase genes had been verified to be conserved in ansamycins (Fig. S1, ESI†), the naphthomycin biosynthetic gene cluster in CS was obtained by PCR screening using degenerate primers designed from *rifK* and other AHBA biosynthetic genes. Compared to other ansamycins, the arrangement of gene cluster and chemical structures of naphthomycins are very similar to rifamycin from *Amycolatopsis mediterranei* S699, suggesting a close genetic relationship between the biosynthesis of naphthomycin and rifamycin. As in rifamycin, eight genes are proposed to be involved in the biosynthesis of the AHBA unit for naphthomycin (Fig. 2). The gene *natI*, homologous with *rifI*, putatively encodes aminoquinate dehydrogenase to regulate the AHBA pool to prevent the uncontrolled accumulation of AHBA in the cell (Fig. S3A, ESI†).[24] Additionally, the organizations of AHBA biosynthetic genes and the AHBA synthases show prominent differences between naphthalenic and benzenic ansamycin biosynthetic genes (Fig. S3A and S3B, ESI†).

Surprisingly, extreme redundancy of reduction domains were identified in *nat* polyketide synthase as deduced from the structure of NATs. Among 13 elongation modules, 2 ketoreductase (KR) and 6 dehydratase (DH) domains were found to be redundant in 7 modules including module 1, 2, 3, 5, 7, 9, and 10 (Fig. 2B and 6). Interestingly, the redundancy of reductive domains is found to be prevalent in the PKS for naphthalenic ansamycin biosynthesis. In rifamycin PKS, five DH domains and 1 KR domain in 7 of the 10 elongation modules were found to be redundant for the biosynthesis of the 25-membered rifamycin B.[23] Moreover in rubradirin PKS, 3 DH domains and 2 KR domains in 4 of the 6 elongation modules were identified to be not required for the biosynthesis of the 15-membered rubradirin.[30] However, the redundant domain has so far not been found in the published PKSs for the biosynthesis of benzenic ansamycins including ansamitocin, geldanamycin, and herbimycin.[17,18,31] Although the redundant domains are not likely to be functional for the PKS structure, it is proposed that they could retain the structural and catalytic integrity as stand-alone PKS proteins.[32]

In nature, the presence of halogenation usually improves the biological activity of secondary metabolites.[33] In ansamycins, both Nat1 for naphthomycin and Asm12 for ansamitocin biosynthesis were categorized to be flavin-dependent

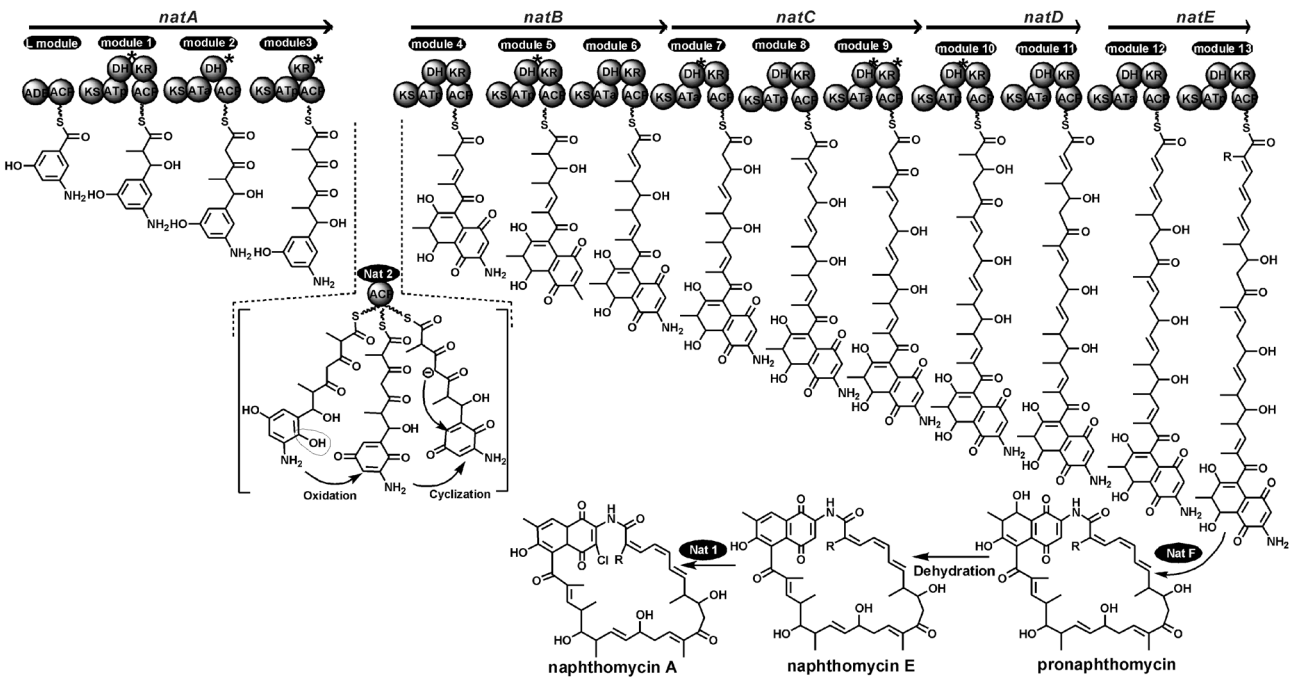

Fig. 6 Putative naphthomycin biosynthetic pathway. ADE, carboxylic acid: ACP ligase (loading domain); KS, β-ketoacyl-ACP synthase; DH, β-hydroxyacyl-thioester dehydratase; KR, β-ketoacyl-ACP reductase; ER, enoyl reductuase. The putative intermediates in chain-extension cycles and the *nat* genes involved in the various biosynthetic steps are indicated. The redundant domains are labeled with *. R, methyl or H.

halogenases. The flavin-dependent halogenases are usually two-component systems, in which a flavin reductase is required to regenerate the $FADH_2$ for catalytic turnover.[34] Since gene deletion study suggested Asm12 catalyzes the first step of ansamitocin post-polyketide modification, further research into the detailed function of Asm12 is needed. The successful cross complementation experiment points to a similar substrate for Nat1 and Asm12 in the biosynthesis of naphthomycins and ansamitocins, respectively. However, expected products of chloro-proansamitocin and naphthomycin A could not be detected *in vitro* using the predicted substrates of proansamitocin and naphthomycin E with purified Asm12 and Nat1, respectively (data not shown). One possibility is that the substrates for Nat1 and Asm12 could be ACP-bound intermediates. The other explanation is the rapid decomposition of free $FADH_2$ with molecular oxygen into FAD and H_2O_2, which leads to a significant decrease in halogenase activity.[34]

The *nat2* gene, putatively encoding a 3-(3-hydroxyphenyl)-propionate hydroxylase and homologous to Rif-Orf19 in rifamycin[23] and RubP1 in rubradirin,[30] was proposed to be involved in the naphthalene ring formation of naphthomycins. Since the disruption of *rif-orf19* ceased production of rifamycin B and caused the accumulation of the tetraketides SY4b and desacetyl-SY4b, Rif-Orf19 was assumed to introduce a hydroxyl group into the acyl carrier protein (ACP)-bound tetraketide. This sets the stage for further enzymatically catalyzed or spontaneous cyclization reaction to form the naphthalene ring. However, whether a single enzyme or a complex group of enzymes is responsible for the entire process remains to be investigated.[29] In *Streptomyces* CS, disruption of *nat2* abolished the production of all the naphthomycin components and accumulated tetraketide desacetyl-SY4b. Moreover, both *nat2* and *rif-orf19* could complement the production of naphthomycins in the *nat2* mutant, implicating the role of Nat2 in the biosynthesis of naphthalene ring formation. Besides naphthalenic ansamycins, *nat2* homologs had been reported to be involved in the oxidation of benzenic ansamycins, such as *gdmM* and *mcbM* from geldanamycin and macbecin biosynthetic gene cluster, respectively.[18,35] *In silico* analysis of the amino acid sequences revealed that all the proteins within this family contain typical nucleotide-binding motifs GXGXXG (motif I), DGXXSXXR (motif II) and GDXXH (motif III) for flavoprotein hydroxylases (Fig. S5A, ESI†).[36,37] Similar to the AHBA synthases, these hydrolases could be divided into two distinct clades either for naphthalenic or benzenic ansamycin biosynthesis by phylogenetic analysis, reflecting divergent biosynthetic routes for these two subgroups of ansamycins (Fig. S5B, ESI†).

Deduced from the abovementioned results, the biosynthetic pathway for naphthomycin A was proposed (Fig. 6). Initiated with the AHBA starter unit, the polyketide backbone of naphthomycin A is extended with the incorporation of 2 methyl-malonyl-CoA and 1 malonyl-CoA to generate a tetraketide tethered to the ACP of module 3. The discrete four-module NatA and three-module NatB may provide an inter-molecule space allowing the accessibility and the oxidation of the tetraketide intermediate by Nat2, followed with an spontaneous intra-molecular cyclization to form the naphthalene ring.[29] The tetraketide with naphthalene ring is then extended with 5 methyl-malonyl-CoA and 5 malonyl-CoA extender units, producing a linear polyketide chain tethered to the ACP of module 13. Subsequently, the nascent polyketide chain is transferred, hydrolyzed, and cyclized by the amide synthase NatF to form the 29-membered pronaphthomycin. Even though no obvious dehydratase gene is identified in the *nat* gene cluster, a dehydration step, rather than the proposed dehydrogenation in rifamycin biosynthesis, is required for the aromatization of the naphthalene ring to generate NAT E. Finally, NAT E is converted to NAT A through chlorination catalyzed by Nat1 (Fig. 6).

Experimental procedures

Materials and general methods

Streptomyces CS[14] was grown at 30 °C in TSBY medium (3% TSB, 1% yeast extract, 10.3% sucrose, pH 7.2, w/v) for liquid cultivation and on ISP3 agar plates (2% agar, 20% oatmeal, 1 ml trace element solution, pH 7.2, w/v) supplemented with 1% w/v leave extract of *Maytenus austroyunnanensis* and 0.6% (w/v) KNO_3. *E. coli* DH10B (Invitrogen) and ET12567/pUZ8002[38] were used throughout the study as a cloning host and transient host for conjugation, respectively. *E. coli* EPI300™ (EPICENTRE) was used for genomic library construction and routine fosmid preparations. *E. coli* strains were grown at 37 °C in LB broth and LB agar. Media were supplemented with appropriate antibiotic concentrations as follows: ampicillin (100 μg ml^{-1}), kanamycin (25 μg ml^{-1}), chloramphenicol (25 μg ml^{-1}), apramycin (30 μg ml^{-1}) and thiostrepton (25 μg ml^{-1}). The strains and plasmids used in this study are listed in Table S1, ESI.†

Synthesis of oligonucleotide primers and sequencing of PCR products were performed by Shanghai Sangon or Invitrogen Co., Ltd. Extraction of DNA fragments from agarose gel slices was performed with the Gel Recovery Kit (Tiangen). Sequence assembly into contigs was performed using Vector NTI 11.0 software (Invitrogen). ORF predictions were made using Framplot 3.0 Beta (http://watson.nih.go.jp/~jun/cgi-bin/frameplot-3.0b.pl) and the Genemark. hmm algorithm (http://exon.biology.gatech.edu/). Domain prediction of polyketide synthase was conducted with SBSPKS.[39] Conjugation between *E. coli* and *Streptomyces* CS were performed as described by Kieser *et al.*[40]

Fosmid library construction and screening

Total DNAs were isolated from *Streptomyces* strains as described by Kieser *et al.*[40] Total DNA was sheared and ligated into CopyControl pCC1FOS vector (EPICENTRE). The ligated mixture was Packed with MaxPlax Lambda Packaging Extracts and transduced into *E. coli* EPI300™ (EPICENTRE) to generate the genomic library. The library was first screened by PCR using degenerated primers for AHBA synthase gene, *i.e.*, AHBAF (5′-CCSGCCTTCACC-TTCATCTCCTC-3′) and AHBAR (5′-AYCCGGAACATS-GCCATGTAGTG-3′). In subsequent screening steps, specific primers based on the identified sequence of the AHBA synthase of naphthomycin C biosynthesis, napCF (5′-GTCA-GCGTGACGCTCAAGGC-3′) and napCR (5′-AGCCCG-TAGGAGGCGACCAC-3′) were used. The entire *nat* gene

cluster was obtained by sequential chromosome walking through PCR amplification.

Gene inactivation

Gene inactivation was carried out by REDIRECT® Technology. A 1.40 kb fragment of pIJ773 carrying the apramycin resistant gene *aac(3)IV* was routinely used as the template for the PCR amplification for gene disruption.[41] Detailed plasmid construction is described in the ESI.† The three constructs, pJTU3231, pJTU3245 and pJTU3250 were delivered into *Streptomyces* sp. CS by conjugation from *E. coli*. The apramycin-resistant and thiostrepton-sensitive recombinants were selected for further verification. Genomic DNA was isolated from the wild-type and the mutants and used as the templates for PCR amplification. The mutants of large fragment deletion, *nat1*, and *nat2* were named as WYY1, WYY2 and WYY3, respectively.

Cross complementation of halogenase genes *nat1* and *asm12*

The *nat1* gene was amplified with KOD-plus DNA polymerase using fosmid 14F11 as a template and primers Nat1-F (5′-AAAGGATCCTCATATGAGCGACAACTCGACCGA-3′, with engineered BamHI and NdeI underlined) and Nat1-R (5′-AAAGAATTCTACTGGCCTTGGGTCACCCGTC-3′, with engineered EcoRI underlined). The *asm12* gene was amplified from the genomic DNA of *A. pretiosum* using primers Asm12-F (5′-AAAGGATCCTCATATGCTCGACGCGATCGTGAT-3′, with engineered BamHI and NdeI underlined) and Asm12-R (5′-AAAGAATTCACCGGCCGACCGTGCCCG-3′, with engineered EcoRI underlined). The resultant PCR products were purified, ligated into pMD-18 (Takara), and sequenced to validate the inserted DNA sequence. For the complementation of the *nat1* mutant, *nat1* and *asm12* were transferred to pJTU824 digested with NdeI and EcoRI to generate pJTU3247 and pJTU3246, respectively. These two plasmids were individually introduced into WYY2 through conjugation, and thiostrepton-resistant exconjugants were selected. For the complementation of *asm12* mutant HGF054, *nat1* and *asm12* were transferred to pIJ139 digested with NdeI and EcoRI to generate pJTU3244 and pJTU3243, respectively. These two plasmids were individually introduced into HGF054 through conjugation, and thiostrepton-resistant exconjugants were selected.

Complementation of WYY3 or *asm12/21* mutant with cloned *nat2* or *rif-orf19*

The *nat2* gene was amplified using the following primers: Nat2-F (5′-AAAGGATCCTCATATGTCGCCGGGCCGTACGCTC-3′) with engineered BamHI and NdeI sites underlined and Nat2-R (5′-AAAGAATTCTCAGTCGGCGCGGCCCGCGGG-3′), with the engineered EcoRI site underlined. The *rif-orf19* gene was amplified using the following primers: Orf19-F (5′-AAAGGATCCTCATATGACGGACACGGACGTCGTC-3′), with engineered BamHI and NdeI sites underlined and Orf19-R (5′-AAAGAATTCCTAGCCGGTGGGCACCGCGGC-3′), with the engineered EcoRI site underlined.

For the complementation of the *nat2* mutant, *nat2* and *rif-orf19* were transferred to pJTU824 digested with NdeI and EcoRI to generate pJTU5156 and pJTU5157, respectively. These two plasmids were individually introduced into WYY3 through conjugation, and thiostrepton-resistant exconjugants were selected. For the complementation of the *asm12/21* mutant HGF065, *nat2* was cloned to pIJ139 digested with NdeI and EcoRI to generate pJTU5158. Plasmid pJTU5158 was introduced into HGF065 through conjugation, and apramycin-resistant exconjugants were selected.

Antibiotic fermentation, purification and analysis

Naphthomycin production by the wild-type and mutants was determined as follows. TSBY medium was inoculated with spore suspension as the seed culture. After incubation for 36 h at 30 °C and 220 rpm, the culture was used to inoculate ISP3 agar plates, which were incubated for 10 d at 30 °C. The culture was extracted twice with ethyl acetate–methanol–acetate (80 : 15 : 5, v/v/v) to afford the crude extract, which was centrifuged, dissolved in ethanol, and directly used for LC-MS analysis. High-performance liquid chromatography (HPLC) analysis was carried out using an Agilent 1100 HPLC system (Agilent Technologies). The samples were separated on a ZORBAX SB-C18 column (2.1 mm × 150 mm, particle size 3.5 μm). The column was equilibrated with 10% acetonitrile (ACN) and was developed with the following program: 0–8 min, 10–40% ACN; 8–42 min, 40–75% ACN; 42–43 min, 75–95% ACN; 43–52 min, 95% ACN; 52–54 min, 95–10% ACN; 54–58 min, 10% ACN at a flow rate of 0.1 ml min^{-1} and UV detection at 254 nm. Naphthomycin A is $C_{40}H_{46}ClNO_9$, with ESI-MS *m/z* of 720.3 $[M + H]^+$ or 742.3 $[M + Na]^+$, and an HPLC retention time of 43.9 min. Naphthomycin E is $C_{40}H_{47}NO_9$ with ESI-MS *m/z* of 686.2 $[M + H]^+$ or 708.3 $[M + Na]^+$, and an HPLC retention time of 38.7 min (Fig. S6). The tetraketide desacetyl SY4b, isolated from *nat2* mutant, is $C_{15}H_{17}NO_5$ with ESI-MS *m/z* of 292.0 $[M + H]^+$ or 314.1 $[M + Na]^+$, and an HPLC retention time of 23.8 min.

Ansamitocins were extracted from the solid cultures of *A. pretiosum*. After incubation in TSBY medium for 24 h, the seed culture was used to inoculate solid YMG medium (2% agar, 0.4% yeast extract, 1% malt extract, 0.4% glucose, pH 7.2, w/v), which was incubated for 7 d at 30 °C. HPLC analysis was carried out using an Agilent 1100 HPLC system (Agilent Technologies). The samples for HPLC analysis were separated by using gradient elution (0–10 min 10–55% B, 10–20 min 55–75% B, 20–35 min 5–95% B, 35–45 min 95% B, 45–50 min 95–10% B, buffer A: H_2O with 0.5% formic acid; buffer B: methanol). This was carried out at a flow rate of 0.1 ml min^{-1} and UV detection at 254 nm. AP-2 is $C_{31}H_{41}ClN_2O_9$, with ESI-MS *m/z* of 621.2 $[M + H]^+$, or 643.2 $[M + Na]^+$, or 659.1 $[M + K]^+$, and an HPLC retention time of 36.2 min. AP-3 is $C_{32}H_{43}ClN_2O_9$, with ESI-MS *m/z* of 635.3 $[M+H]^+$, or 657.2 $[M + Na]^+$, or 673.2 $[M + K]^+$, and an HPLC retention time of 37.6 min. AP-4 is $C_{31}H_{41}ClN_2O_9$, with ESI-MS *m/z* of 649.3 $[M + H]^+$, or 671.2 $[M + Na]^+$, or 687.2 $[M + K]^+$, and an HPLC retention time of 39.4 min (Fig. S6, ESI†).

Accession number

The nucleotide sequence of naphthomycin biosynthetic genes described in this paper was deposited in the GenBank with the accession number of GQ452266.

Acknowledgements

The authors would like to thank Lutz Heide for critical reading of the manuscript and helpful discussion. This work received financial support from the National Natural Science Foundation of China, the Ministry of Science and Technology (973 and 863 Programs), the Ministry of Education and the Shanghai Municipal Council of Science and Technology.

References

1 V. Prelog and W. Oppolzer, *Helv. Chim. Acta*, 1973, **56**, 2279.
2 P. Sensi, P. Margalith and M. T. Timbal, *Farmaco. Sci.*, 1959, **14**, 146–147.
3 C. DeBoer, P. A. Meulman, R. J. Wnuk and D. H. Peterson, *J. Antibiot. (Tokyo)*, 1970, **23**, 442–447.
4 E. Higashide, M. Asai, K. Ootsu, S. Tanida, Y. Kozai, T. Hasegawa, T. Kishi, Y. Sugino and M. Yoneda, *Nature*, 1977, **270**, 721–722.
5 J. J. Kibby, I. A. McDonald and R. W. Rickards, *J. Chem. Soc., Chem. Commun.*, 1980, 768–769.
6 C. G. Kim, A. Kirschning, P. Bergon, P. Zhou, E. Su, B. Sauerbrei, S. Ning, Y. Ahn, M. Breuer, E. Leistner and H. G. Floss, *J. Am. Chem. Soc.*, 1996, **118**, 7486–7491.
7 C. G. Kim, T. W. Yu, C. B. Fryhle, S. Handa and H. G. Floss, *J. Biol. Chem.*, 1998, **273**, 6030–6040.
8 J. Mochizuki, E. Kobayashi, K. Furihata, A. Kawaguchi, H. Seto and N. Otake, *J. Antibiot. (Tokyo)*, 1986, **39**, 157–161.
9 T. Okabe, B. D. Yuan, F. Isono, I. Sato, H. Fukazawa, T. Nishimura and N. Tanaka, *J. Antibiot. (Tokyo)*, 1985, **38**, 230–235.
10 M. Balerna, W. Keller-Schierlein, C. Martius, H. Wolf and H. Zahner, *Arch. Microbiol.*, 1969, **65**, 303–317.
11 W. Keller-Schierlein, M. Meyer, A. Zeeck, M. Damberg, R. Machinek, H. Zahner and G. Lazar, *J. Antibiot. (Tokyo)*, 1983, **36**, 484–492.
12 T. Mukhopadhyay, C. M. Franco, G. C. Reddy, B. N. Ganguli and H. W. Fehlhaber, *J. Antibiot. (Tokyo)*, 1985, **38**, 948–951.
13 R. R. Hooper AM, *J. Antibiot. (Tokyo)*, 1998, **51**, 845–851.
14 C. Lu and Y. Shen, *J. Antibiot. (Tokyo)*, 2007, **60**, 649–653.
15 J. P. Lee, S. W. Tsao, C. J. Chang, X. G. He and H. G. Floss, *Can. J. Chem.*, 1994, **72**, 182–187.
16 S. Chen, D. von Bamberg, V. Hale, M. Breuer, B. Hardt, R. Muller, H. G. Floss, K. A. Reynolds and E. Leistner, *Eur. J. Biochem.*, 1999, **261**, 98–107.
17 T. W. Yu, L. Bai, D. Clade, D. Hoffmann, S. Toelzer, K. Q. Trinh, J. Xu, S. J. Moss, E. Leistner and H. G. Floss, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 7968–7973.
18 A. Rascher, Z. Hu, N. Viswanathan, A. Schirmer, R. Reid, W. C. Nierman, M. Lewis and C. R. Hutchinson, *FEMS Microbiol. Lett.*, 2003, **218**, 223–230.
19 Y. Mao, M. Varoglu and D. H. Sherman, *J. Bacteriol.*, 1999, **181**, 2199–2208.
20 S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson and L. Katz, *Science*, 1991, **252**, 675–679.
21 R. Reid, M. Piagentini, E. Rodriguez, G. Ashley, N. Viswanathan, J. Carney, D. V. Santi, C. R. Hutchinson and R. McDaniel, *Biochemistry*, 2003, **42**, 72–79.
22 S. F. Haydock, J. F. Aparicio, I. Molnar, T. Schwecke, L. E. Khaw, A. Konig, A. F. Marsden, I. S. Galloway, J. Staunton and P. F. Leadlay, *FEBS Lett.*, 1995, **374**, 246–248.
23 P. R. August, L. Tang, Y. J. Yoon, S. Ning, R. Muller, T. W. Yu, M. Taylor, D. Hoffmann, C. G. Kim, X. Zhang, C. R. Hutchinson and H. G. Floss, *Chem. Biol.*, 1998, **5**, 69–79.
24 T. W. Yu, R. Muller, M. Muller, X. Zhang, G. Draeger, C. G. Kim, E. Leistner and H. G. Floss, *J. Biol. Chem.*, 2001, **276**, 12546–12555.
25 P. Spiteller, L. Bai, G. Shang, B. J. Carroll, T. W. Yu and H. G. Floss, *J. Am. Chem. Soc.*, 2003, **125**, 14236–14237.
26 N. S. Scrutton, A. Berry and R. N. Perham, *Nature*, 1990, **343**, 38–43.
27 K. H. van Pee and S. Unversucht, *Chemosphere*, 2003, **52**, 299–312.
28 T. W. Yu, Y. Shen, Y. Doi-Katayama, L. Tang, C. Park, B. S. Moore, C. Richard Hutchinson and H. G. Floss, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 9051–9056.
29 J. Xu, E. Wan, C. J. Kim, H. G. Floss and T. Mahmud, *Microbiology*, 2005, **151**, 2515–2528.
30 C. G. Kim, J. Lamichhane, K. I. Song, V. D. Nguyen, D. H. Kim, T. S. Jeong, S. H. Kang, K. W. Kim, J. Maharjan, Y. S. Hong, J. S. Kang, J. C. Yoo, J. J. Lee, T. J. Oh, K. Liou and J. K. Sohng, *Arch. Microbiol.*, 2008, **189**, 463–473.
31 A. Rascher, Z. Hu, G. O. Buchanan, R. Reid and C. R. Hutchinson, *Appl. Environ. Microbiol.*, 2005, **71**, 4862–4871.
32 C. Khosla, S. Kapur and D. E. Cane, *Curr. Opin. Chem. Biol.*, 2009, **13**, 135–143.
33 A. Butler and M. Sandy, *Nature*, 2009, **460**, 848–854.
34 C. Wagner, M. El Omari and G. M. Konig, *J. Nat. Prod.*, 2009, **72**, 540–553.
35 M. Q. Zhang, S. Gaisser, E. A. M. Nur, L. S. Sheehan, W. A. Vousden, N. Gaitatzis, G. Peck, N. J. Coates, S. J. Moss, M. Radzom, T. A. Foster, R. M. Sheridan, M. A. Gregory, S. M. Roe, C. Prodromou, L. Pearl, S. M. Boyd, B. Wilkinson and C. J. Martin, *J. Med. Chem.*, 2008, **51**, 5494–5497.
36 C. Enroth, H. Neujahr, G. Schneider and Y. Lindqvist, *Structure*, 1998, **6**, 605–617.
37 M. H. Eppink, H. A. Schreuder and W. J. Van Berkel, *Protein Sci.*, 1997, **6**, 2454–2458.
38 M. S. Paget, L. Chamberlin, A. Atrih, S. J. Foster and M. J. Buttner, *J. Bacteriol.*, 1999, **181**, 204–211.
39 S. Anand, M. V. Prasad, G. Yadav, N. Kumar, J. Shehara, M. Z. Ansari and D. Mohanty, *Nucleic Acids Res.*, 2010, **38**, W487–496.
40 *Practical Streptomyces Genetics*, ed. T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Charter and D. A. Hopwood, The John Innes Foundation, Norwich, United Kingdom, 2000.
41 B. Gust, G. L. Challis, K. Fowler, T. Kieser and K. F. Chater, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1541–1546.