Four Amino Acid Residues in Acyltransferase Affect the Production of Epothilone A to Mixed Epothilones A and B

Xin Sun1, Pengfei Li2, Yuezhong Li2 and Lin Zhao1,*

1State Key Laboratory of Biobased Material and Green Papermaking, school of bioengineering, Qilu University of Technology, Shandong Academy of Sciences, Jinan, 250353, China
2State Key Laboratory of Microbial Technology, College of Life Science, Shandong University, Jinan 250100, P. R. China

*Corresponding author

Abstract. Surveying more than 100 native epothilone producing strains were isolated from different locales revealed that the production ratio of epothilones A and B was sorted into two groups, almost single epothilone A or approximate 2:1 ratio of A and B. Sequencing the second acyltransferase in the epoC module (ATModC2) of different Sorangium producers indicated that four amino acid residues at sites 90, 91, 95 and 196 in the enzyme active centre altered in parallel to the production ratio shift. In those producers with single epothilone A, ATModC2 had almost the same sequence and computational structure as ATModC1, which selectively accepts malonyl-CoA as the extender units.

Keywords: epothilone; acyltransferase; aminoacid.

1. Introduction

Epothilones are taking many attentions because of their Taxol-like microtubule stabilization activity and potential large-scale fermentation production [1]. At present, there are at least six epothilones used in anti-cancer trials [2], and ixabepilones has been approved by the USFDA. The reported three strains, i.e. So ce90, SMP44, and So0157-2 [3, 4, 5] produce a mixture of epothilones A and B (approximately 2:1 in ratio) as the major product. There are differences of between epothilones A and B, and people studies to increase B ratio by either optimizing fermentation procedures, or engineering the producing strains and the biosynthetic genes[1].

Production of epothilones A and B is decided by the second acyltransferase domain of module C (ATModC2), via promiscuously accepting malonyl-CoA or methylmalonyl-CoA as the extender units [6, 7]. Although it is practical in other systems to swap AT domains to alter the substrate selection of acyltransferases, selective producing single epothilone B has never been realized in Sorangium strains [8, 9, 10].

In our screening work, we obtained many epothilone producing Sorangium strains from different locales. Although the yields of epothilones varied in different Sorangium producers, the ratios of epothilones A and B were sorted into two groups, either almost single epothilone A or approximate 2:1 ratio of A and B. Interestingly, the sequence and computational structure of the active centre of ATModC2 was significantly changed in parallel to the ratio shift. We amplified the ATModC2 of the strain So0157-2 which can produce both epothilone A and B. Then a four points mutation was introduced into this enzyme which made it transform into the ATModC2 of the strains with single epothilone A just like So0007-12.
2. Experimental Procedure

2.1. Strain and Condition
The Sorangium strains were in Table 1. Sorangium strains were cultured on CNST agar as previously described [11]. For DNA extraction, the Sorangium cells from 4-day CNST cultures were transferred into M26 liquid medium and incubated for 7 days [12]. For the production of epothilones, strains were inoculated on M26 for 4 or 5 days. Then XAD-16 resin were added, and incubated for additional 9 to 10 days as previously described [5]. E. coli cells were cultured in LB with Ampicillin. The Sorangium strains were incubated at 30°C, while E. coli was at 37°C.

Table 1. Production of epothilones in different Sorangium cellulosum strains.

| Isolate   | Accession number of ATModC2 | Epothilone A (mg/L) | Epothilone B (mg/L) | B/A  |
|-----------|----------------------------|---------------------|---------------------|------|
| So007-12-2A | EU554397                   | 1.52                | 0                   | −    |
| So007-12-2B | EU554398                   | 1.51                | 0                   | −    |
| So0087-2-1  | EU864255                   | 0.47                | 0                   | −    |
| So0087-2-2  | EU554399                   | 0.19                | 0                   | −    |
| So0087-2-3  | EU554400                   | 0.54                | 0                   | −    |
| So0087-7-1  | EU554401                   | 0.94                | ±                   | −    |
| So0087-14   | EU554402                   | 1.71                | 0                   | −    |
| So0157-2    | EU554406                   | 1.7                 | 0.8                 | 0.47 |
| So ce90\(^a\) |                       | 22                  | 11                  | 0.5  |
| SMP44\(^b\) |                       | 12.3                | 6.7                 | 0.5  |

a. cited from the reference of Gerth et al. 1996.
b. cited from the reference of Bollag et al. 1995.

2.2. Detection and Identification of Epothilones
Resins were extracted with methanol. After dried at 40°C, the extracts were stored at -20°C. For a typical HPLC-MS analysis, samples from 50 ml cultures were redissolved in 500 μL methanol. HPLC interfaced with a Finnigan MSQ classic quadruples mass spectrometer and MRC-ODS column with the mobile phase of 65% methanol and 35% buffer. Epothilone A and B were eluted at 12.5 and 14.2 min. To quantify the production titers, 10 μL aliquots of the samples were analyzed on Surveyor HPLC under the same chromatography condition, detected at 249 nm. They were quantified based on purified epothilone A and B.

2.3. Nucleic Acid Extraction and PCR Amplification
The genomic materials of Sorangium cells were extracted. The ATModC2 domain sequences in different Sorangium strains were amplified using the nest PCR protocol to avoid unspecific amplifications. The first route of PCR amplification was performed with the primer pair of 5' AGTGGACGCATGACGCTGAC-3' and 5'-ACGTCGATTTCCGTGGAATGC-3', targeting the KS and KR domains. The product is 2.7 kb, containing the ATModC2 domain. After electrophoresis, they were incubated for 15 min, rinsed with Milli-Q water for 10 min, and photographed using a BioRad GelDoc station (BioRad, California, USA). The target bands were excised and extracted using a Gel extract kit (Bioflux), and dissolved in 30 μL Milli-Q water. The second PCR route used a primer pair of 5'-CTG CGC GAG CAC CTG GAC ATG C-3' and 5'-GCT GCC GCT GCC ACG GAT AGG T-3'. The PCR fragments were extracted and ligated into T-vector and transformed into DH-5α.

2.4. Nucleic Bioinformatics Analysis
The epothilone biosynthetic gene sequences of So ce90 (AF210843) and SMP44 (AF217189) were obtained from the GenBank databases. The sequences were completely aligned using the ClastalW programs. Phylogenetic reconstruction was carried out in MEGA 3D structure of ATModC2 domain sequences was on-line predicted using swissmodel (http://swissmodel.expasy.org/), based on the acyltransferase of the DEBS biosynthetic gene cluster for erythromycin in the SWISS PDB databases.
The similarity of the sequences of the DEBS acyltransferase and AT\textsubscript{ModC2} is about 44%. The logP value and MR value were calculated by ChemDraw Pro6.0 (http://www.cambridgesoft.com/software/ChemDraw/).

3. Results

3.1. Epothilones A/B Ratios in Sorangium Producers

There have been reported three epothilone producers, Soce90, SMP44 and So0157-2 \cite{3, 4, 5}. These strains all had an approximately 2:1 ratio of mixed epo A and B. To obtain mutants that were able to selectively produce epothilone B, So0157-2 strain was mutated \cite{5}. Although mutants with increased proportion of epothilone B (but in low productivity of epothilones) were occasionally obtained, no strains were able to selectively produce epothilone B only, or had normal productivity with higher proportion of epothilone B than A. Our screening projects obtained more than 100 epothilone producing Sorangium strains. These native producers showed great differences in epothilone production, but the ratios of epothilones A and B in these strains were only sorted in two groups. Some strains produced a mixture of epothilones A and B with about 2:1 ratio (named epoB\textsuperscript{+} in this paper), while others had extremely low or even undetectable epothilone B (named epoB\textsuperscript{−}, which means they lost the ability of epothilone B production). However, none of these Sorangium strains were able to selectively yield epothilone B without A, or even higher proportion of B than A. In this study, seven natural epoB\textsuperscript{−} strains, together with three epoB\textsuperscript{+} strains were analyzed (Table 1 showed the productions of epothilones A and B in these strains).

3.2. AT\textsubscript{ModC2} Domain Sequences Analysis

Epothilones are biosynthesized in Sorangium cells by seven multifunctional proteins \cite{14, 15}. The acyltransferases (ATs) are responsible for the choice of extender units. Based on the incorporation of radioactively labeled precursors and the composition of biosynthetic enzymes \cite{6, 7}, the production of epothilone A or B is suggested to be attributed to the promiscuous substrate selection with malonyl-CoA or methylmalonyl-CoA. Besides AT\textsubscript{ModC2}, the other eight acyltransferases had a strict selection of malonyl-CoA or methylmalonyl-CoA. Fig. 1 exhibited the model of epothilones A and B biosynthesis. Because epothilone producing Sorangium strains are divided into epoB\textsuperscript{+} and epoB\textsuperscript{−} groups, it is then interesting to see changes of the AT\textsubscript{ModC2} gene in those epoB\textsuperscript{−} strains.

We amplified the AT\textsubscript{ModC2} of all strains. All of the ATs in nine modules of the three reported strains were phylogenetic analyzed and the results indicated that all sequences are clearly sorted into two groups: the malonyl-CoA accepting AT (MCAT) and the methylmalonyl-CoA accepting AT (MMCAT) (Fig. 2A). The AT\textsubscript{ModC2} domain belongs to the malonyl-CoA type, showing high identity to the first acyltransferase domain in epoC module (AT\textsubscript{ModC1}), which is only able to accept malonyl-CoA \cite{6, 7}. The AT\textsubscript{ModC2} sequences from different Sorangium strains were highly conserved with more than 98.6% similarity in nucleotide acid sequences and 97.7% similarity in amino acid sequences. AT\textsubscript{ModC2} were all located in the AT\textsubscript{ModC1}/AT\textsubscript{ModC2} subgroup (Fig.2A).
Figure 1. Substrate selection of AT\textsubscript{ModC2} domain (gray) in the biosynthesis of epothilones in \textit{Sorangium} strains.

Figure 2. (A) Phylogenetic analysis of the AT\textsubscript{ModC2} domains from different \textit{Sorangium} strains, with all of the acyltransferases in the biosynthetic gene clusters from So ce90 and SMP44 strains. (B) Alignment of AT\textsubscript{ModC2} sequences, with AT\textsubscript{ModC1} sequences from So ce90 and SMP44 strains. The active sites are marked with ‘*’; four sites that we analyzed are marked bold black).

Some particular amino acids (or motifs) of ATs are responsible for the substrate selection. We compared the 13 amino acids might be active sites of all the AT sequences in PKSs. (highlighted in Fig.2B) They were essential for the substrate selectivity and reaction\textsuperscript{[11]}. However, 11 of the 13 sites were the same in both epoB\textsuperscript{+} and epoB\textsuperscript{-} \textit{Sorangium} strains, except the 95\textsuperscript{th} and 196\textsuperscript{th} residues. In epothilone native
producers, this putative 4-aa decision motif in the epoB$^+$ strains is HASH, but HAFH in all epoB$^-$ strains (Fig. 2B). The results suggested that, besides the HASH/HAFH/YASH motif, there were some other potential residues cooperating with them for the substrate selection.

3.3. 3D Model of ATModC2 Domain and Substrate Selection Bioinformatics’ Analysis

In acyltransferases, the sites of Ser$_{94}$ and Arg$_{119}$ are responsible for the catalysis (the process was shown in Fig.3A). The computational 3D structure model of ATModC2 was constructed with the acyltransferase of DEBS in Streptomyces erythreus as template (pdb code: 2hg4a), which was highly consistent with the previous report of malonyl-CoA type acyltransferase [16]. Comparing the 3D structures of ATModC2 domains from different Sorangium producers revealed that four amino acid residues in the active centre (marked in the sequence alignment in Fig. 2B) were altered significantly. Two (residues 95 and 196) are among the 13 previously putative active sites in PKS acyltransferases [11], while the other two (residues 90 and 91) have not been mentioned. These four amino acids were close to the two catalyzing sites (Ser$_{94}$ and Arg$_{119}$) (Fig. 3B), suggesting their potential roles in the substrate selection. The four amino acids in epoB$^-$ (i.e. Leu$_{90}$, Val$_{91}$, Ile$_{95}$ and Phe$_{196}$) were much bigger than those in epoB$^+$ (Val$_{90}$, Ala$_{91}$, Ala$_{95}$, and Ser$_{196}$), increasing 29% (90th site), 146% (91th site), 218% (95th site), and 276% (196th site) in molecular volume. In addition, these four amino acids also showed more hydrophobic in epoB$^-$ strains than in epoB$^+$ strains. The substitution of the amino acid residues at these four sites formed a complicated stereo-hindrance for the entrance of methylmalonyl-CoA into the catalyzing centre of epoB$^-$ ATModC2. When we constructed the 3D-structure with acyltransferase of FADs in Saccharopolyspora erythraea as template (pdb code: 2qo3a), it showed that the substitution of 90, 91 site changed the position of neighboring Ser$_{94}$, which made the catalyzing sites of Ser$_{94}$ and Arg$_{119}$ to be closer in epoB$^-$ ATModC2 than in epoB$^+$ ATModC2 (Fig. 3B). The stereo-hindrance in the active centre of the epoB$^-$ ATModC2 had less effect for accepting Malonyl-CoA, which is smaller and less hydrophobic than methylmalonyl-CoA. Interestingly, the four amino acids of ATModC2 in those epoB$^-$ strains were identical to those of ATModC1 (Fig. 2B). Thus, it is suggested that these four amino acids are crucial and cooperative in the substrate selection of ATModC2 domain.

Figure 3. (A) Pathway of the transacylation. (B) The skin representation of epoB$^-$ ATModC2 and the catalyzing region (red is Ser$_{94}$; purple is Arg$_{119}$). (C) 3D structure of epoB$^+$ and epoB$^-$ constructed with template Pdb code: 2qo3a.
4. Conclusion
Myxobacteria are famous not only for their complicated multi-cellular social lifestyle [17], and their diverse secondary metabolites [18, 19]. The ratio of epo A and B in native Sorangium strains suggested that the biosynthetic genes are highly conserved. Acyltransferase is responsible for the choice of extender units for the PKS. In the enzymatic complex, there are nine acyltransferases either selectively accepting malonyl-CoA or methylmalonyl-CoA, except AT_{ModC2}. However, AT_{ModC2} had about 98% sequence identity to the malonyl-CoA type acyltransferase AT_{ModC1}. Paralleled to the production ratio shift, the amino acid residues at the sites 90, 91, 95 and 196 in the active region differed between epoB$^{-}$ and epoB$^{+}$ strains. The changes led to a high stereo-hindrance of AT_{ModC2} for the acceptance of methylmalonyl-CoA in epoB$^{-}$. In cells, malonyl-CoA is more easily accessible than methylmalonyl-CoA, because of their importance for cellular basic metabolisms. It is a disappointing suggestion that natural Sorangium producers probably cannot produce epothilone B without A. However, we assumed that it may be practical to swap the complete AT_{ModC2} domain with some other tolerable methylmalonyl-CoA type acyltransferases for the selective epothilone B without A. It will significantly improve the value of the strain for epothilone B.

Acknowledgments
The work was financially supported by Shandong Provincial Natural Science Foundation (NO. ZR2017BH021).

References
[1] Rowinsky E K, and Calvo E. Novel agents that target tubulin and related elements. Semin Oncol. 2006, 33: p421-435.
[2] Larkin J M, and Kaye S .B Epothilones in the treatment of cancer. Expert. Opin Investig Drugs. 2006, 15: p691–702.
[3] Gerth K, Bedorf N, Höfl e G, Irschik H, and Reichenbach H. Epothilones A and B: Antifungal and cytotoxic compounds from Sorangium cellulosum (Myxobacteria) -production, physico-chemical and biological properties. J Antibiot. 1996, 49: p560-564.
[4] Bollag D M, McQueney P A, Zhu J, Hensens O, Koupal L, Liesch J, Goetz M, Lazarides E, and Woods C M. Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. Cancer Res. 1995, 55: p2325-2333.
[5] Gong G L, Sun X, Liu X L, Hu W, Cao W R, Liu H. Mutation and high-throughput screening method for improving the production of Epothilones of Sorangium. J Ind Microbiol Biotechnol. 2007. 34: p615-623.
[6] Gerth K, Steinmetz H, Höfle G, and Reichenbach H. Studies on the biosynthesis of epothilones: the biosynthetic origin of the carbon skeleton. J Antibi ot. 2000, 53: p1373-1377.
[7] Molnár I, and Schupp T, Zirkle R, Milnamow M, Nowak-Thompson B, Engel N. The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from Sorangium cellulosum So ce90. Chem Biol. 2000, 7: p97-109.
[8] Petković H, Lill R E, Sheridan R M, Wilkinson B, McCormick E L, McArthur H A. A novel erythromycin, 6-desmethyl erythromycin D, made by substituting an acyltransferase domain of the erythromycin polyketide synthase. J Antibi ot. 2003, 56: p543-551.
[9] Regentin R, Cadapan L, Ou S, Zavala S, and Licari P. Production of a novel FK520 analog in Streptomyces hygroscopicus: improving titer while minimizing impurities. J Ind Microbiol Biotechnol.2002, 28: p12-16.
[10] Patel K, Piagentini M, Rascher A, Tian Z Q, Buchanan G O, Regentin R. Engineered biosynthesis of geldanamycin analogs for Hsp90 inhibition. Chem Biol. 2004, 11: p1625-1633.
[11] Yadav G, Gokhale R S, and Mohanty D. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. J Mol Biol. 2003, 8: p335-363.
[12] Nguimbi E, Li Y Z, Gao B L, Li Z F, Wang B, Wu Z H. 16S-23S ribosomal DNA intergenic spacer regions in cellulolytic myxobacteria and differentiation of closely related strains. Syst Appl Microbiol. 2003, 26: p262-268.
[13] Tang Y, Kim C Y, Mathews I I, Cane D E and Khosla C. The 2.7-Angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythonolide B synthase. Proc Natl Acad Sci. 2006, 03: p11124-11129.

[14] Julien B, Shah S, Ziermann R, Goldman R, Katz L, and Khosla C. Isolation and characterization of the epothilone biosynthetic gene cluster from Sorangium cellulosum. Gene. 2000, 249: p153-160.

[15] Molnár I, and Schupp T, Zirkle R, Milnamow M, Nowak-Thompson B, Engel N. The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from Sorangium cellulosum So ce90. Chem Biol. 2000, 7: p97-109.

[16] Oefner C, Schulz H, D'Arcy A, and Dale G E. Mapping the active site of Escherichia coli malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography. Acta Cryst. 2000, 62: p613-618.

[17] Whitworth D E. Myxobacteria: multicellularity and differentiation. ASM Press, 2007.

[18] Reichenbach H. Myxobacteria, producers of novel bioactive substances. J Ind Microbiol Biotechnol. 2001, 27: p149-156.

[19] Gerth K, Pradella S, Perlova O, Beyer S, and Müller R. Myxobacteria: proficient producers of novel natural products with various biological activities--past and future biotechnological aspects with the focus on the genus Sorangium. J Biotechnol. 2003, 106: p233-253.