Appl Biochem Biotechnol (2008) 149:67–78
DOI 10.1007/s12010-007-8053-7

Isolation of New Polyketide Synthase Gene Fragments and a Partial Gene Cluster from East China Sea and Function Analysis of a New Acyltransfrase

Yu-Liang Jiao · Liang-Hua Wang · Xiao-Yi Dong · Yu-Feng Chen · Ying Zong · Yun Gao · Na Ren · Ai-Yun Guo · Xin-Qun Zhang · Bing-Hua Jiao

Received: 5 July 2007 / Accepted: 7 September 2007 /
Published online: 6 November 2007

Abstract Using the consensus-degenerate hybrid oligonucleotide primer polymerase chain reaction method, 26 new ketoacyl synthase (KS) fragments were isolated from a marine sediment sample in the East China Sea (ECS) and analyzed by construction of a phylogenetic tree. With a digoxigenin-labeled KS gene fragment used as a probe, a partial polyketide synthase (PKS) gene cluster was isolated and identified by hybridization screening of a marine sediment sample metagenome fosmid library constructed for this study. A new acyltransferase (AT) gene was cloned from the PKS gene cluster and heterogeneously expressed as a protein fused to maltose-binding protein (MBP). Ultraviolet spectrophotometry was used to study the binding of the MBP–AT fusion protein and single AT domain to substrates using MBP and bovine serum albumin as control proteins. Binding constants (*K*a, per micromolar) were calculated and used to analyze the substrate specificity of the acyltransferase. We concluded that there are many unrevealed new PKS gene clusters in marine sediments in the ECS. The acyltransferase is presumably an acetyltransferase from a new PKS gene cluster.

Y.-L. Jiao · L.-H. Wang · X.-Y. Dong · Y. Zong · Y. Gao · N. Ren · A.-Y. Guo · X.-Q. Zhang · B.-H. Jiao (✉)
Department of Biochemistry and Molecular Biology,
College of Basic Medical Science,
800 Xiangyin Road, Shanghai 200433,
People's Republic of China
e-mail: jiaobh@mail.uninet.com.cn

Y.-F. Chen
Department of Pharmacognosy, School of Pharmacy,
Second Military Medical University,
Shanghai 200433, People's Republic of China

Keywords Acetyltransferase · Acyltransferase · Ketoacyl synthase · Polyketide synthase · Substrate specificity

Abbreviations

CODEHOP PCR	consensus-degenerate hybrid oligonucleotide primers PCR
BSA	bovine serum albumin
PKS	polyketide synthase
ECS	the East China Sea
CoA	coenzyme A
DIG	digoxigenin
KS	ketoacyl synthase
PFGE	pulsed-field gel electrophoresis
AT	acyltransferase
ACP	acyl carrier protein
MBP	maltose-binding protein
PBS	phosphate-buffered saline
UV	ultraviolet
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
IPTG	isopropyl β-D-1-thiogalactopyranoside
NBT/BCIP	nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate

Introduction

Polyketide synthases (PKS) are multifunctional enzymes or multienzyme complexes that catalyze the formation of a polyketide assembled from sequential condensation of short-chain acyl coenzyme A (CoA) precursors or ACP-linked precursors [1–3]. Some of the polyketides have antitumor or antibiotic activity, and a large fraction of biologically active molecules and pharmaceuticals are synthesized by PKS pathways [4–6]. Isolation and characterization of new PKS gene clusters play an important role in studying and employing PKS for academic research and industrial development. Marine environment has the most enormous phylogenetic and metabolic diversity in the Earth but remain undersampled and essentially uncharted [7–9]. Given the great number of unknown marine microbes and their vast bioactive metabolites, the diversity of PKS gene clusters should be in a very high level. It is important to investigate the diversity of PKS in marine environment and also a shortcut to isolate new PKS gene clusters from it. Many new PKS gene clusters have been isolated from the metagenome of marine microbial symbionts, such as sponges (e.g., *Pseudoceratina clavata*) and bryozoans (e.g., *Bugula neritina*) [10–12], but related reports on marine sediments were few.

Combinatorial biosynthesis is the application of genetic engineering to modify biosynthetic pathways to natural products to produce new and altered structures using nature's biosynthetic machinery [13]. The combinatorial biosynthesis of polyketide has become a hot research spot in which the acyltransferase (AT) domains in charge of the specific loading of precursors are of great importance [14, 15]. However, the function identification of new AT domains or holo-PKS still limit the speed of research and development of PKS combinatorial biosynthesis [16, 17]. In this study, 26 new KS fragments were isolated from a marine sediment sample in the East China Sea (ECS), and a new acetyltransferase in a partial PKS gene cluster was identified.

Experimental Procedures

Extraction of Metagenome DNA from a Marine Sediment Sample

Ten milliliters of a marine sediment sample M1 was gathered from ECS and centrifuged at 10,000×*g* for 10 min then discarded the supernatant and rinsed the precipitation with sterile water. The procedure was repeated five times. The precipitation was freeze dried in vacuum and weighed 5.8 g approximately. Following the instruction of PowerSoil™ DNA isolation kit (Mo Bio Laboratories), 0.5 g of the precipitation was weighed for isolating the metagenome DNA. The step in the manufacturer's protocol of binding DNA to the silica spin filter membrane was not adopted in our procedure. The crude extract of DNA from the manufacturer's procedure was added 1/10 volume of 3 M sodium acetate (pH 7.0) and mixed gently, and then added 2.5 volumes of ethanol and centrifuged at 12,000×*g* for 10 min. The DNA pellet was washed two times by 70% ethanol and resuspended in STE buffer. The purity and quality of metagenome DNA was analyzed by pulsed-field gel electrophoresis (PFGE) with MidRange I PFG marker (New England Biolabs, range 15–240 kb). The DNA fragments of molecular size between 30–50 kb were excised and purified with GELase enzyme (Epicentre Technologies) for metagenome library construction. The PFGE condition consisted of 1% low-melting point (LMP) agarose gel, 6 V/cm, 15°C for 24 h, and switch time of 5 s.

Isolation of Ketoacyl Synthase Fragments by CODEHOP PCR

Ketoacyl synthase (KS) domain is conserved both in function and gene structure; many new PKS gene clusters have been found by amplifying the KS domain [18–20]. Based on the alignment of cd00833 in conserved domain database of GenBank, two most conserved motif, DPQQR and HGTGT, were selected as target site to design PCR primers. Primer design referred consensus-degenerate hybrid oligonucleotide primers (CODEHOP) method devised by Rose et al. [21]. Thirty-two primers were designed targeting motif DPQQR, and eight of them were synthesized and used as forward primer (KSCDFP). Twenty-four primers were designed targeting motif HGTGT, and 14 of them were synthesized and used as reverse primer (KSCDRP; as shown in Table 1). KSCDFP and KSCDRP were paired resulting totally to 112 pairs of primers. According to the distance between two motifs, the molecular sizes of PCR amplicons should range from 600 to 900 bp. Because of the high GC content of KS domain, GC buffer I (Takara) was used, and annealing temperature was around 65°C. Amplicons of correct molecular size were cloned into pMD18-T vector

Table 1 Primers and sequences.

Primer	Name	Sequence
CODEHOP for KS	KSCDFP	CGCTCCATGGAYCCSCARCA
	KSCDRP	GTCCCGGTSSCRTGSSHYTCSA
Primers for DQ924531	KSCDFP4	CGCTCCATGGACCCGCAGCA
	KSCDRP8	GTCCCGGTGCCATGCGCCTCCA
Fosmid sequencing primers	pCC2FSP	GTACAACGACACCTAGAC
	pCC2RSP	CAGGAAACAGCCTAGGAA
Primers for cloning and expression of AT-EF080951	ATFP	CTGAATTCATGGGCTTCGGCGAGGCGGGCTTC
	ATRP	AGGTCGACTCAGGGCTCCGCACGGAACCAG

(Takara) and sequenced by an ABI-3730 sequencer (Applied Biosystems) and BLASTed against GenBank both in nucleotide sequences and peptide sequences.

Metagenome Fosmid Library Construction of Marine Sediment Sample M1

CopyControl™ Fosmid Library Production Kit (Epicentre Technologies) was used to construct a library. The metagenome DNA of molecular size between 30–50 kb was end-repaired with end-repair enzyme and then cloned into pCC2FOS vector. The fosmid clones were packaged together by λ phage particles at 30°C by incubation for two consecutive 90 min. The packaged fosmid clones were serially diluted and incubated with host strain (EPI300™-T1^{R}) at 37°C for 20 min, and then the infected host strain cells were spread on an LB plate containing 12.5 mg/ml chloramphenicol and incubated at 37°C overnight. The colonies were counted, and the titer of packaged fosmid clones was calculated through the following formula.

$$\text{Titer} = \frac{(\text{Number of colonies})(\text{Dilution factor})}{(\text{Volume of phage plated})}$$

Colony Lift and Nonradioactive Screening of Fosmid Library

Positively charged nylon membranes (Osmonics Inc.) were used for colony lift. Approximately 500 clones transferred onto a 13-mm ∅ membrane were lysed by alkali and then baked at 80°C for 2 h for cross-linking of fosmid vectors to membranes. Using PCR DIG Probe synthesis kit (Roche), a KS fragment (GenBank accession no. DQ924531) was labeled with digoxigenin (DIG) for the screening of metagenome library. PCR primers for labeling are KSCDFP4 and KSCDRP8 (as shown in Table 1). The cycling parameter consisted of one cycle of 94°C for 3 min then 35 cycles of 94°C for 30 s, 65°C for 45 s, 72°C for 45 s, followed by a single 7-min cycle at 72°C for extension. The procedure of DIG-labeled probe hybridization and color detection of positive clones with chromogenic substrates (NBT/BCIP) was according to Roche's instruction.

Cloning, Expression, and Purification of Acyltransferase AT-EF080951

Cloning of Acyltransferase Gene AT-EF080951

PCR primers for cloning of AT-EF080951 were designed targeting at the nucleotide sequence of joining regions between KS domain's C-terminal region and ACP domain's N-terminal region. Forward primer ATFP was added an *Eco*RI site and a start codon at 5′-end. Reverse primer ATRP was added a *Sal*I site and a stop codon at 5′-end. AT-EF080951, a 1,048-bp fragment amplified from positive fosmid clone was cloned into pMD18-T and sequenced and verified. The joining region peptide that remained at the N-terminal and C-terminal regions of AT-EF080951 was GFGEAGFGSARRTA and GQVRWARWFRAEP respectively (as shown in Fig. 1).

Expression and Purification of Fusion Protein MBP-AT-EF080951

The AT-EF080951 gene fragment was double digested by *Eco*RI and *Sal*I (Takara), and then cloned into the pMAL-c2X vector of pMAL™ protein fusion and purification system

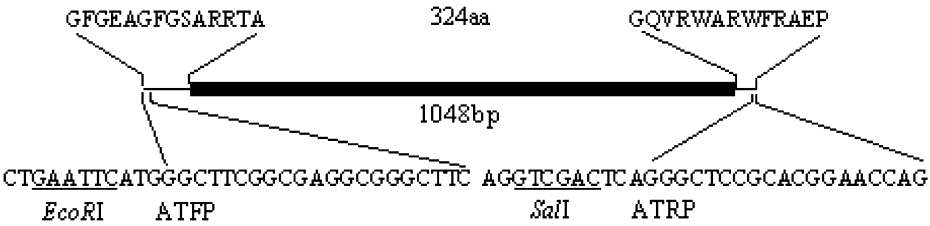


Fig. 1 Peptide sequence of linker and primer design of AT-EF080951 for pMAL expression system

(NEB) followed by transformation into *E. coli* BL21 (DE3; Novagen). Positive clones were selected by colony PCR followed by digestion with *Eco*RI and *Sal*I to verify the incorporation of inserts of the correct size. The correct clone was incubated with LB broth shaking at 37°C and added with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.3 mM when A_{600} was 0.5, then continued with vigorous shaking for 2 h. Cells were harvested by centrifugation at 4°C and rinsed two times by PBS (pH 7.4) buffer then resuspended and added with 5 ml column buffer per 0.1 g cells wet weight followed by sonication in short pulses of 15 s in ice-water bath, and then centrifuged at 9,000×*g* at 4°C for 10 min. The supernatant was loaded to amylose resin column at a flow rate of 1 ml/min. The fusion protein MBP-AT-EF080951 was eluted with column buffer containing 10 mM maltose and verified by running 1 μl of it on a 15% sodium dodecyl sulfate (SDS)-polyacrylamide gel. The rest were stored at −80°C. The protein concentrations were determined in triplicate with the Bradford assay (Bio-Rad) with bovine albumin as a standard.

Purification of Single AT-EF080951 Protein

MBP-AT-EF080951 was cleaved by Factor Xa at room temperature overnight. The cleavage reaction was confirmed by running 3 μl of the reaction on a 15% SDS-polyacrylamide gel. The single AT-EF080951 protein in the reaction mixture was purified by rebinding MBP to the amylose column. The trace Factor Xa that remained with AT-EF080951 was less than 0.1% (*w*/*w*).

Analysis of the Substrate Specificity of AT-EF080951

Immobilization of MBP-AT-EF080951, AT-EF080951, and Control Proteins

The coating buffer is just the elution buffer of MBP-AT-EF080951: 20 ml of 1.0 M Tris–HCl (pH 7.4), 11.7 g of NaCl, 2.0 ml of 0.5 M ethylenediaminetetraacetic acid (EDTA), 0.7 ml of β-mercaptoethanol, and the distilled water was added to a final volume of 1 l. A hundred microliters of MBP-AT-EF080951, MBP, and BSA all resolved in coating buffer were added into microwells of an enzyme-linked immunosorbent assay (ELISA) plate separately. The ELISA plate was incubated at 37°C for 2 h with stable moisture. The concentrations of the proteins were determined after 2 h of immobilization, and the amount of the proteins immobilized was calculated.

Determination of Absorption Curves of the Substrates at 260 nm Wavelength

Four substrates, acetyl coenzyme A, malonyl coenzyme A, coenzyme A, and methyl-malonyl coenzyme A (Sigma), were all resolved in PBS (7.4) buffer and used as substrates.

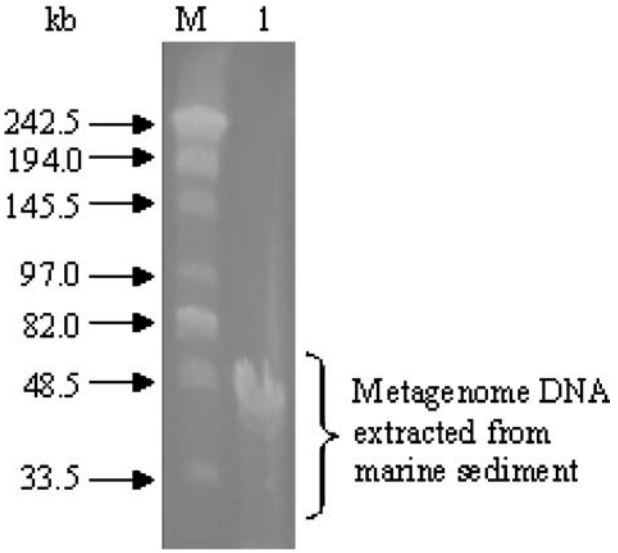


Fig. 2 PFGE analysis of metagenome DNA extracted from marine sediment sample M1. The molecular size of metagenome DNA ranges from 30 to 50 kb. *M* Marker and *1* metagenome DNA

Because the maximum absorbance wavelength of the four substrates is 260 nm and the limits of detection of Eppendorf biophotometer were from 5 to 120 μg/ml, concentrations of the substrates were selected at 10, 20, 40, 80, and 100 μg/ml in ultraviolet (UV) measurement at 260 nm, and each concentration was analyzed three times. The absorbances of substrates of the five concentrations at 260 nm wavelength were measured, and the linear regression equations of absorption curves were calculated.

Determination of Binding Constant of MBP-AT-EF080951 and AT-EF080951 for the Substrates

A hundred microliters of a substrate (80 μg/ml) was separately added into four microwells coated with quantified MBP-AT-EF080951, AT-EF080951, MBP, and BSA separately and a microwell coated with nothing as a control. The ELISA plate was incubated at 37°C with stable moisture for binding reaction, and then the absorbance at 260 nm wavelength of the substrate remaining in the solutions was measured by Eppendorf biophotometer. The amount of substrates bond to the proteins were calculated through the absorption curve equations of the substrates. The binding constant of a protein for a substrate can be calculated through the following formula:

$$Ka = \frac{[protein \cdot substrate]}{[protein][substrate]}$$

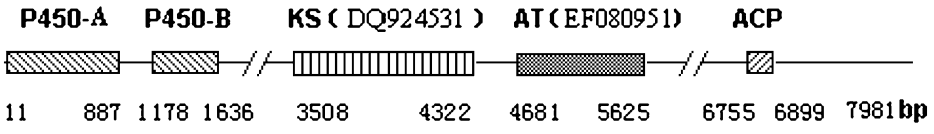


Fig. 3 Structure of new partial PKS gene cluster EF568935. The numbers under the figure show the location of the domains on the nucleotide sequence. *P450A and P450B* Two new cytochrome P450 oxidases

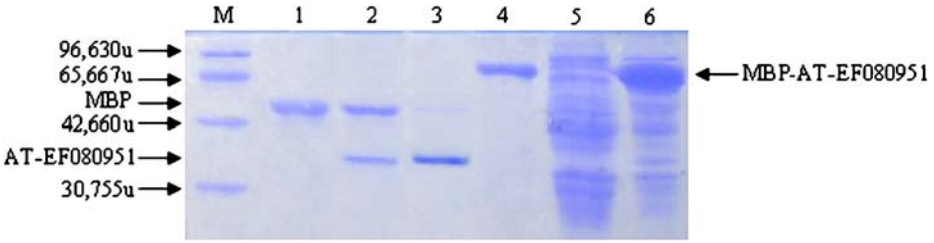


Fig. 4 SDS-PAGE analysis of expression and purity of fusion protein MBP-AT-EF080951 and single AT-EF080951. *M* Marker, *lane 1* purified MBP, *lane 2* MBP-AT-EF080951 cleaved by factor Xa, *lane 3* purified AT-EF080951 with trace factor Xa (MW similar to that of MBP), *lane 4* purified MBP-AT-EF080951, *lane 5* BL21-pMAL-c2X cell extracts, *lane 6* BL21-pMAL-c2X cell extracts 2 h after IPTG induction

Results

Extraction of Metagenome DNA from Marine Sediment Sample M1

Ten micrograms of metagenome DNA was extracted from 0.5 g marine sediment dry weight, and its molecular size ranged from 30 to 50 kb (as shown in Fig. 2).

Isolation of KS Fragments

Totally, 53 fragments of correct molecular size amplified using CODEHOP PCR were cloned into pMD18-T vector and sequenced, 26 sequences of which are confirmed as new KS domains and submitted to GenBank with the following accession numbers acquired: EF554859-EF554861, DQ837023, DQ837024, DQ640993, DQ640997, DQ641926, DQ641927, DQ673137-DQ673152, and DQ924531.

Construction of Metagenome Fosmid Library of the Marine Sediment Sample M1

The titer of packaged phage particles is 5×10^8 cfu/ml. The average molecular size of the inserts is 40 kb, and if the average size of a single clone in sediment sample M1 was 5 Mb, the library represents 10^6 clones in the sample calculated through the following formula:

$$\text{Clones represented} = \frac{\text{Titer}\bullet(\text{Vpackaged phage particles})}{\ln(1-P)/\ln(1-f)}$$

Table 2 Absorbance curve equations of substrates.

C-substrate	10	20	40	80	100	Curve equation
Substrate						
Malonyl coenzyme A	0.253	0.514	1.042	2.072	2.471	Y=0.025x+0.0225 R^2=0.9984
Acetyl coenzyme A	0.162	0.343	0.693	1.373	1.741	Y=0.0174x-0.0088 R^2=0.9999
Coenzyme A	0.129	0.26	0.522	1.048	1.158	Y=0.0119x+0.0304 R^2=0.9896
Methylmalonyl coenzyme A	0.125	0.252	0.503	1.010	1.230	Y=0.0124x+0.0056 R^2= 0.9996

Y A260 and *X and C-substrate* substrate concentrations (μg/ml)

Table 3 The binding constant (*K*a, per micromolar) of MBP-AT-EF080951, AT-EF080951, MBP, and BSA for malonyl coenzyme A, acetyl coenzyme A, coenzyme A, methylmalonyl coenzyme A.

Substrate	Malonyl coenzyme A	Acetyl coenzyme A	Coenzyme A	Methylmalonyl coenzyme A
Protein				
MBP-AT-EF080951	0.0049±0.003	0.107±0.002	0.005±0.003	0.0049±0.001
AT-EF080951	0.0041±0.002	0.120±0.001	0.0042±0.003	0.005±0.002
MBP	0.0049±0.004	0.0048±0.002	0.0051±0.002	0.0047±0.003
BSA	0.0055±0.004	0.005±0.003	0.004±0.004	0.005±0.002

Where P (=0.99) is the desired probability (expressed as a fraction), f (=40 kb/5Mb) is the proportion of the genome of average size contained in a single fosmid clone; V (volume of) packed phage particles is 1 ml.

Screening of the Fosmid Library

Four positive clones were detected by color reaction after the DIG-labeled probe hybridization. The fosmid vectors of the four clones were extracted and sequenced, but because of complicated structures existing in all of them, only a 7,981 bp partial PKS gene cluster was obtained and submitted to GenBank with accession number EF568935 acquired. The gene structure of EF568935 is shown in Fig. 3.

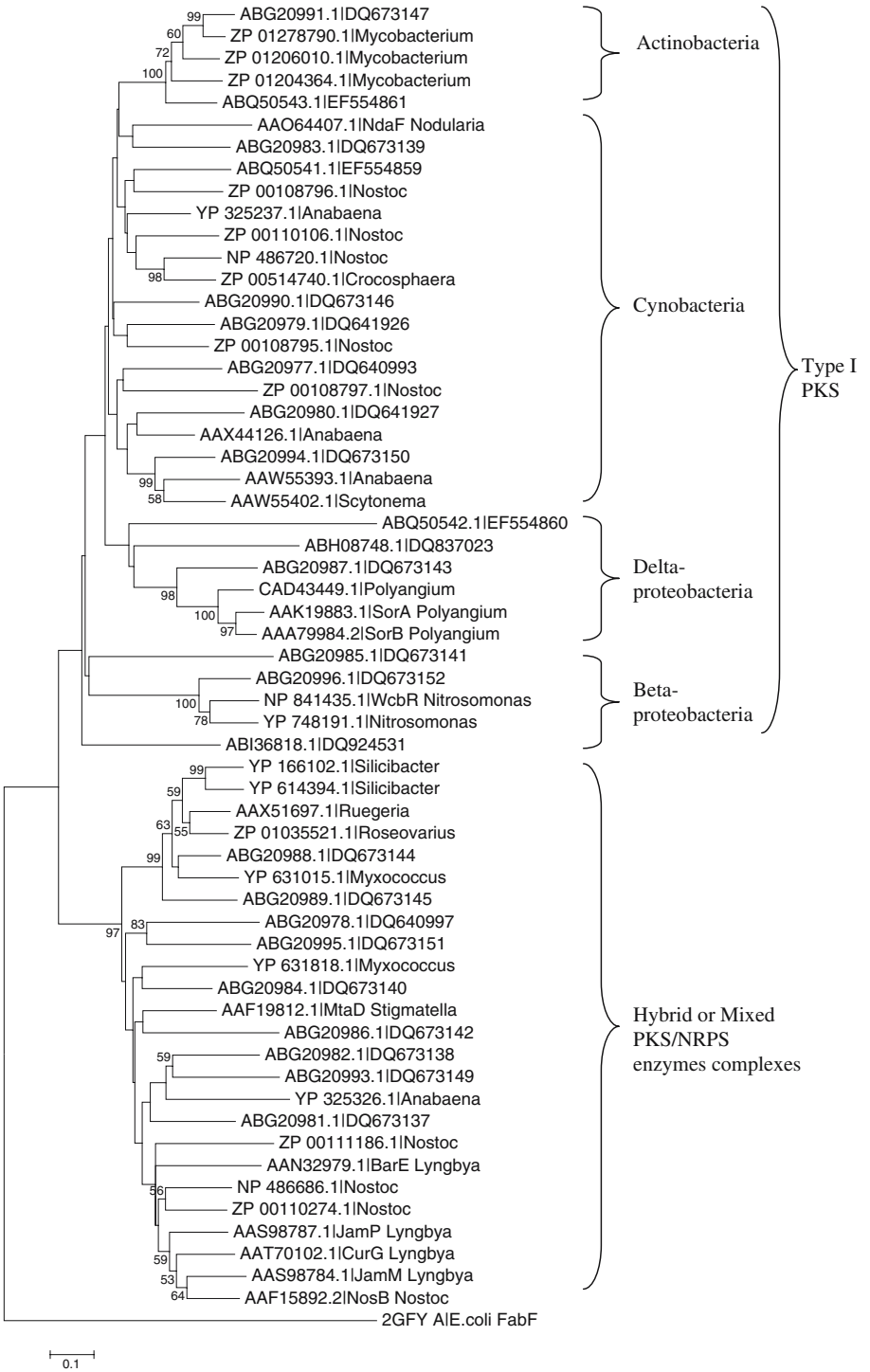
Expression and Purification of MBP-AT-EF080951 and AT-EF080951

The insertion of AT-EF080951 into pMAL-c2X vector was verified by sequencing and restriction enzymes digestion. The purity and the molecular size of the fusion protein were analyzed by SDS-PAGE. Concentration of the fusion protein and that of AT-EF080951 were 100 μg/ml and 80 μg/ml, respectively. There are 342 amino acids in protein AT-EF080951 and the molecular weight of it is 36,609 u. Then the molecular weight of the fusion protein MBP-AT-EF080951 should be 78,772 u. The result of the SDS-PAGE analysis is in good agreement with the prediction (as shown in Fig. 4).

Analysis on Substrate Specificity of AT-EF080951

The amounts of the proteins immobilized were all 100 μg per well. The concentrations of substrates in control wells did not change during the binding, so the substrates did not bind to the ELISA plate. UV spectrophotometry is highly specific and sensitive in the detection of the four substrates. The absorbance data and linear regression equations of absorption curves of substrates at 260 nm are shown in Table 2. The binding constants of

Fig. 5 Phylogenetic analysis of the protein sequences of 25 (of 26) new KS fragments from a marine sediment sample and the closest sequences derived from the GenBank (DQ673148 is not included because of its partial sequence). The construction is computed by the distance method (NJ, Poisson correction distance model). *E. coli* FabF is used as an outgroup. Only bootstrap values ≥50% are shown ▶


ABG20991.1|DQ673147
ZP 01278790.1|Mycobacterium
ZP 01206010.1|Mycobacterium
ZP 01204364.1|Mycobacterium
ABQ50543.1|EF554861
AAO64407.1|NdaF Nodularia
ABG20983.1|DQ673139
ABQ50541.1|EF554859
ZP 00108796.1|Nostoc
YP 325237.1|Anabaena
ZP 00110106.1|Nostoc
NP 486720.1|Nostoc
ZP 00514740.1|Crocosphaera
ABG20990.1|DQ673146
ABG20979.1|DQ641926
ZP 00108795.1|Nostoc
ABG20977.1|DQ640993
ZP 00108797.1|Nostoc
ABG20980.1|DQ641927
AAX44126.1|Anabaena
ABG20994.1|DQ673150
AAW55393.1|Anabaena
AAW55402.1|Scytonema
ABQ50542.1|EF554860
ABH08748.1|DQ837023
ABG20987.1|DQ673143
CAD43449.1|Polyangium
AAK19883.1|SorA Polyangium
AAA79984.2|SorB Polyangium
ABG20985.1|DQ673141
ABG20996.1|DQ673152
NP 841435.1|WcbR Nitrosomonas
YP 748191.1|Nitrosomonas
ABI36818.1|DQ924531
YP 166102.1|Silicibacter
YP 614394.1|Silicibacter
AAX51697.1|Ruegeria
ZP 01035521.1|Roseovarius
ABG20988.1|DQ673144
YP 631015.1|Myxococcus
ABG20989.1|DQ673145
ABG20978.1|DQ640997
ABG20995.1|DQ673151
YP 631818.1|Myxococcus
ABG20984.1|DQ673140
AAF19812.1|MtaD Stigmatella
ABG20986.1|DQ673142
ABG20982.1|DQ673138
ABG20993.1|DQ673149
YP 325326.1|Anabaena
ABG20981.1|DQ673137
ZP 00111186.1|Nostoc
AAN32979.1|BarE Lyngbya
NP 486686.1|Nostoc
ZP 00110274.1|Nostoc
AAS98787.1|JamP Lyngbya
AAT70102.1|CurG Lyngbya
AAS98784.1|JamM Lyngbya
AAF15892.2|NosB Nostoc
2GFY A|E.coli FabF
0.1
Actinobacteria
Cynobacteria
Delta-proteobacteria
Beta-proteobacteria
Type I PKS
Hybrid or Mixed PKS/NRPS enzymes complexes

MBP-AT-EF080951 and AT-EF080951 for methylmalonyl coenzyme A, malonyl coenzyme A, acetyl coenzyme A, and coenzyme A are shown in Table 3.

Discussion

The protocol for the extraction of the high-quality metagenome DNA is critical in the construction of a library and degenerate PCR amplification of KS fragments. The silica spin filter membrane in the kit for binding and eluting DNA was not used in our procedure. Compared with the kit, our protocol can increase the average molecular size by 10 kb longer. In constructing metagenome DNA fosmid library of the marine sediment sample, DNA of appropriate molecular size for constructing a library were separated by PFGE and then end-repaired, which, compared with the reverse procedure in the kit, could increase the titer of packaged phage particles by one magnitude.

Isolation and identification of new PKS gene clusters play an important role in studying and employing PKS. CODEHOP PCR amplification, screening of metagenome library, sequencing, and structure analysis of amplicons in this study are shown as effective ways to isolate new PKS gene clusters. Because of high GC content in PKS gene clusters and complicated structure existing in high frequency, the partial gene cluster EF568935 can not be sequenced through and the sequencing of holo-PKS gene clusters is, thus, a new problem urgent to be resolved.

Many new PKS gene clusters have been isolated from marine bacterial symbionts, but few reports have been seen on new PKS gene isolated from marine sediments. In this study, 26 new KS fragments isolated from a marine sediment sample in the ECS reveals that many PKS genes in marine environments that probably synthesize new bioactive polyketides were not recognized by people. It is important to isolate these new PKS gene clusters for combinatorial biosynthesis of new polyketide pharmaceuticals. The 26 new KS fragments belong to type I ketosyntase, as well as hybrid or mixed PKS/NRPS enzyme complexes primarily originating from marine actinobacteria, delta-proteobacteria, cyanobateria, and beta-proteobacteria by phylogenetic analysis using MEGA (4) and Clustal X (1.83) (as shown in Fig. 5), which also guides us to isolate and culture the original microbe of some certain KS fragment.

Soluble and specific substrate-binding active fusion protein MBP-AT-EF080951 and single AT-EF080951 acyltransferase are obtained through heterologous expression in pMAL system, which indicates the function of an acyltransferase selectively binding to a specific substrate is dependent, that is, the activity of an acyltransferase does not necessarily depend on the N-terminal KS domain and the C-terminal ACP domain. The joining regions at both

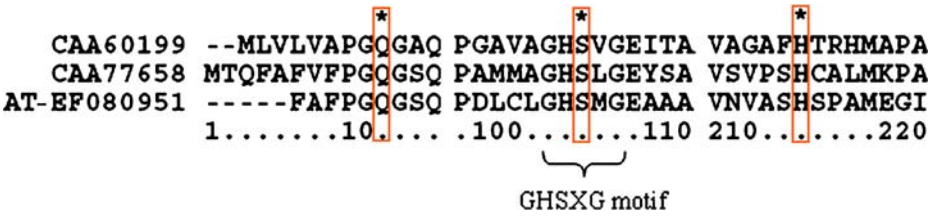


Fig. 6 Alignment of AT-EF080951 and tow malonyltransferases (GenBank accession no. CAA60199 and CAA77658). Active sites are indicated with *asterisks*

sides of AT-EF080951 remain no more than 14 amino acids, which also indicate that its function does not depend on the space structure of the joining regions. These conclusions can explain why in the PKS combinatorial biosynthesis engineering, when the sizes of foreign domains and joining regions are so different to that of the native domains swapped, a functional hybrid PKS can still be expressed. Therefore, it is relatively free to select the joining regions of a foreign domain as the homologous arms in constructing a targeting vector. This study contributes theoretical supports to PKS combinatorial biosynthesis engineering. The fusion protein MBP-AT-EF080951 and single AT-EF080951 can be used in studying the interaction with KS or ACP domain in vitro and researches in PKS combinatorial biosynthesis engineering.

In a study of Crawford et al. [22] on the substrate specificity of the starter unit acyltraferase of norsolorinic acid PKS, the binding constant of the specific substrate hexanoyl–CoA is 12-fold higher than that of octanoyl–CoA, which has the highest specificity constant in non-specific substrates. In this study, the binding constant of acetyl coenzyme A is 19-fold higher than that of the other two substrates compared with the control protein MBP and BSA. Acetyl coenzyme A is concluded as the specific substrate of acyltransferase AT-EF080951, which is then presumably an acetyltransferase. Comparison of the molecular structures of the four substrates suggests that AT-EF080951 specifically binds to acetyl group, which also explain why there are many high homologous AT domains sequentially repeatedly existing in a PKS gene cluster, but only one kind of polyketide product was synthesized by the PKS.

Crawford et al. have found that malonyltransferases had a GHSXG motif harboring the active site serine, which is covalently loaded by malonyl–CoA with participation of a conserved downstream histidine constituting a catalytic dyad. In this study, the protein sequence of AT-EF080951 is aligned with two typical malonyltransferases by Clustal X (1.83) and also found having the characters of malonyltransferases (as shown in Fig. 6). But AT-EF080951 does not bind to malonyl–CoA and methylmalonyl–CoA specifically, which suggests that the molecular catalytic mechanism of AT is more complicate than what we know.

The substrate specificity of an AT domain can also be studied by analyzing the structural alteration of the polyketide product of an AT-domain-swapped PKS [23, 24], but the method is complex and time-consuming. In this study, a functional single AT domain was obtained through cloning, and heterologous expression and its substrate specificity was studied directly in vitro. The method we described herein is simpler and more effective, which is also an important way to study and employ the function of PKS gene clusters in the future research.

Acknowledgment This work was sponsored by the Natural Science Funds of the People's Republic of China (grant number 30670043) and the National High Technology Research and Development Program of China (863 Program; grant number 2006AA09Z434).
Jiao, Wang and Dong contributed equally to this work.

References

1. Hopwood, D. A. (1997). *Chemical Reviews (Washington, DC)*, *97*, 2465–2497.
2. Staunton, J., & Weissman, K. J. (2001). *Natural Product Report*, *18*, 380–416.
3. Shen, B. (2000). *Topics in Current Chemistry*, *209*, 1–51.
4. Wawrik, B., Kerkhof, L., Zylstra, G. J., & Kukor, J. J. (2005). *Applied Environmental Microbiology*, *71* (5), 2232–2238.

5. Metsa-Ketela, M., Salo, V., Halo, L., Hautala, A., Hakala, J., & Mantsala, P., et al. (1999). *FEMS Microbiology Letters*, *180*, 1–6.
6. O'Hagan, D. (1991). Polyketides. Chichester, UK: Horwood.
7. Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R. et al. (2006). *Proceeding of the National Academy of Science of the United States of America*, *103*(32), 12115–12120.
8. Ward, A. C., & Bora, N. (2006). *Current Opinion of Microbiology*, *9*, 279–286.
9. Pathom-Aree, W., Stach, J. E., Ward, A. C., Horikoshi, K., Bull, A. T., & Goodfellow, M. (2006). *Extremophiles*, *10*(3), 181–189.
10. Kim, T. K., & Fuerst, J. A. (2006). *Environmental Microbiology*, *8*(8), 1460–1470.
11. Davidson, S. K., Allen, S. W., Lim, G. E., Anderson, C. M., & Haygood, M. G. (2001). *Applied Environmental Microbiology*, *67*, 4531–4537.
12. Lopanik, N. B., Targett, N. M., & Lindquist, N. (2006). *Applied Environmental Microbiology*, *72*(12), 7941–7944.
13. Floss, H. G. (2006). *Journal of Biotechnology*, *124*(1), 242–257.
14. Cheng, Y. Q., Tang, G. L., & Shen, B. (2003). *Proceeding of the National Academy of Science of the United States of America*, *100*(6), 3149–3154.
15. Zhang, W. J., Ames, B. D., Tsai, S. C., & Tang, Y. (2006). *Applied Environmental Microbiology*, *72*(4), 2573–2580.
16. Strohl, W. R. (2000). *Drug Discovery Today*, *5*, 39–41.
17. Strohl, W. R. (1997). Industrial antibiotics: today and the future. In W. R. Strohl (Ed.) Biotechnology of antibiotics (2nd ed.). Marcel DekkerNewYork, NY.
18. Ginolhac, A., Jarrin, C., Gillet, B., Robe, P., Pujic, P., & Tuphile, K., et al. (2004). *Applied Environmental Microbiology*, *70*(9), 5522–5527.
19. Kroken, S., Glass, N. L., Taylor, J. W., Yoder, O. C., & Turgeon, B. G. (2003). *Proceeding of the National Academy of Science of the United States of America*, *100*(26), 15670–15675.
20. Lopez, J. V. (2003). *Mol Genet Genomics*, *270*, 420–431.
21. Rose, T. M., & Henikoff, J. G. (2003). *Nucleic Acids Research*, *31*(13), 3763–3766.
22. Crawford, J. M., Dancy, B. C., Hill, E. A., Udwary, D. W., & Townsend, C. A. (2006). *Proceeding of the National Academy of Science of the United States of America*, *103*(45), 16728–16733.
23. Reeves, C. D., Chung, L. M., Liu, Y., Xue, Q., Carney, J. R., & Revill, W. P., et al. (2002). *Journal of Biological Chemistry*, *277*(11), 9155–9159.
24. Kim, B. S., Sherman, D. H., & Reynolds, K. A. (2004). *Protein Engineering Design and Selection*, *17* (3), 277–284.