

RESEARCH

Open Access



# A bacterial negative transcription regulator binding on an inverted repeat in the promoter for epothilone biosynthesis

Xin-jing Yue, Xiao-wen Cui, Zheng Zhang, Ran Peng, Peng Zhang, Zhi-feng Li and Yue-zhong Li\* 

## Abstract

**Background:** Microbial secondary metabolism is regulated by a complex and mostly-unknown network of global and pathway-specific regulators. A dozen biosynthetic gene clusters for secondary metabolites have been reported in myxobacteria, but a few regulation factors have been identified.

**Results:** We identified a transcription regulator Esi for the biosynthesis of epothilones. Inactivation of *esi* promoted the epothilone production, while overexpression of the gene suppressed the production. The regulation was determined to be resulted from the transcriptional changes of epothilone genes. Esi was able to bind, probably via the N-terminus of the protein, to an inverted repeat sequence in the promoter of the epothilone biosynthetic gene cluster. The Esi-homologous sequences retrieved from the RefSeq database are all of the Proteobacteria. However, the Esi regulation is not universal in myxobacteria, because the *esi* gene exists only in a few myxobacterial genomes.

**Conclusions:** Esi binds to the epothilone promoter and down-regulates the transcriptional level of the whole gene cluster to affect the biosynthesis of epothilone. This is the first transcription regulator identified for epothilone biosynthesis.

**Keywords:** Negative transcription regulator, Transcription inhibition, Epothilone synthesis, Inverted repeat sequence in promoter, Myxobacteria, Proteobacteria

## Background

Many kinds of microorganisms, such as actinomycetes, bacilli and myxobacteria, are excellent producers of secondary metabolites with various biological activities, potentially intriguing in anti-infection, anti-cancer and other pharmaceutical applications. The production of secondary metabolites in microorganisms is normally limited, probably due to the complex and mostly-unknown network of global and pathway-specific regulators, as well as the large biosynthetic gene cluster containing many gene modules [1–4]. Myxobacteria are able to produce diverse secondary metabolites [5–8]. Although a dozen gene clusters for the biosynthesis of secondary metabolites have been reported in

myxobacteria, a few regulation factors have been identified. StiR is the first regulator identified for secondary mechanism in myxobacteria, which was found involving in the positive regulation of stigmatellin production in *Cystobacter fuscus* Cb f17.1 [9]. ChiR was found to serve as a pleiotropic regulator, which not only positively regulated the chivosazol biosynthesis, but also affected the fruiting body development in *Sorangium cellulosum* So ce56 [10]. NtcA, a famous global transcriptional regulator responsive to the concentration of nitrogen, was reported to be able to regulate the biosynthesis of chivosazol and etnangien negatively in *S. cellulosum* So ce56 [11].

Epothilones are originally produced by some strains of *S. cellulosum*, and are a group of antitumor compounds mimicking paclitaxel in the polymerization of tubulin [6, 12]. Epothilones are potential anticancer agents by their greater water solubility and efficacy against paclitaxel-resistant tumors [13, 14]. Researchers tried

\*Correspondence: lilab@sdu.edu.cn  
State Key Laboratory of Microbial Technology, School of Life Science,  
Shandong University, Jinan 250100, China

many methods to improve the epothilones yield, including co-cultivation of different strains [15], mutation and high-throughput screening of high-producing strains [16], immobilization of strains [17], heterologous expressions [18–24] and manipulation of related genes [25, 26]. However, the yield is still limited. The biosynthetic gene cluster of epothilone is a large operon of approximately 56 kb in size, containing seven homodromous open reading frames (ORFs) [27, 28]. Because of the lack of efficient protocols for genetic manipulation in *S. cellulosum*, the regulation mechanisms of this big operon have not yet been investigated and thus remains mostly unknown.

A key regulatory step that modulates bacterial gene expressions is the promoter recognition by RNA polymerase and transcription initiation [29]. A transcription factor can increase the rate of transcription initiation (activators) or prevent RNA polymerase from initiating transcription (repressors). Many repressors prevent transcription by binding DNA at positions directly interfering the binding of RNA polymerases. Thus, in the promoters subject to repression, operator sequences for a repressor are often found to overlap or be immediately adjacent to the transcription start site [30]. For example, HrcA, a classic heat-shock regulator, binds to the controlling inverted repeat of chaperone expression (CIRCE) to regulate expressions of the downstream *groEL-ES* operon and *dnaK* gene [31–33]. Transcription regulators achieve specific binding in normally a dimer or further multimer form; and most operators contain direct or inverted repeats of a 4–5 base pair sequence [34].

*Sorangium cellulosum* So0157-2 is an epothilone producing strain [35]. We previously sequenced the genome of So0157-2 [36], studied in details the characteristics of epothilone operon promoter in *Escherichia coli* (*E. coli*) [37], and randomly integrated the epothilone biosynthetic gene cluster into *M. xanthus* genome by transposition [22]. In this study, we reported the identification of a negative transcription regulator for the biosynthesis of epothilones.

## Results

### Inactivation of *esi* promotes the epothilone production in *S. cellulosum* So0157-2

Using the conjugation and subsequent homologous recombination protocols as previously described [38], we inactivated genes flanking the epothilone gene cluster in *S. cellulosum* So0157-2 to determine whether they were associated with the synthesis, regulation or modification of epothilones. The *SCE1572\_25030* gene locates approximately 500 bp upstream to the epothilone promoter (Fig. 1a). We amplified an internal fragment of the gene and cloned it into the pCC11 plasmid [38] to construct the inactivation vector, which was introduced into

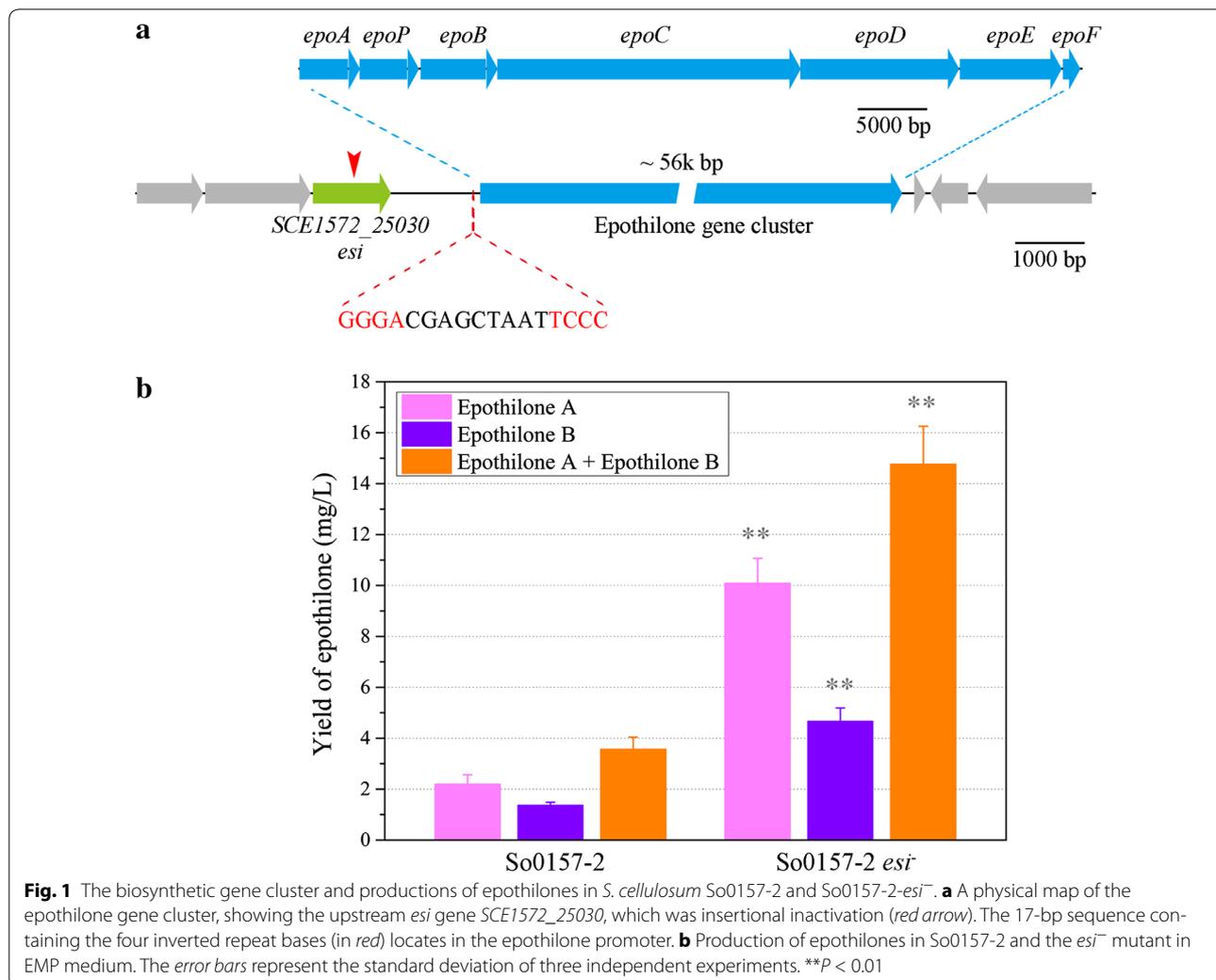
the genome of So0157-2. The gene was inactivated by insertional mutagenesis, which was confirmed by amplification using Tail-PCR and sequencing of the insertion region.

We fermented the mutant and the wild type strain in 50 mL of EMP medium supplemented with 2% of XAD-16 resin [16]. After 7 days of shaking cultivation, the resin was harvested, and the absorbed compounds were extracted for High Performance Liquid Chromatography (HPLC) analysis to determine the yields of epothilones. The results showed that the insertion mutation of *SCE1572\_25030* resulted in a remarkable increase in the yields of epothilones A and B, from 2.2 and 1.4 mg/L in the wild type strain to 4.7 and 10.1 mg/L in the mutant (Fig. 1b; *t* test, *p* value <0.01). The total yields of epothilones A and B increased approximately 4.1 times, from 3.6 to 14.8 mg/L. The results suggested that the *SCE1572\_25030* gene probably encoded an epothilone synthesis inhibitor (*Esi*). The mutant strain was termed as So0157-2-*esi*<sup>-</sup>.

### The presence of *esi* gene suppresses the epothilone production in *M. xanthus*

Due to the genetic manipulation difficulty in *S. cellulosum* cells, we investigated influences of the *esi* gene on the production of epothilones in *M. xanthus*. We previously integrated the So0157-2 epothilone biosynthetic gene cluster randomly into *M. xanthus* genome by transposition; and the *esi* gene upstream to the epothilone gene cluster was also included in the transferred fragment [22]. We chose the ZE9 transformant derived from *M. xanthus* DZ2 for the assay. Bioinformatics analysis indicated that there was no homologous sequence of *esi* in DZ2, which was further confirmed by PCR amplification of an *esi* fragment against the DZ2 genome. Firstly, we deleted the *esi* gene from ZE9, producing the knockout mutant ZE9Δ*esi*. The ZE9 and ZE9Δ*esi* strains had similar growth curves in CYE medium (Fig. 2a). Methyl oleate is able to greatly increase the epothilone production ability in *Myxococcus* cells [39]. When fermented in the CYE medium supplemented with methyl oleate (CMO medium), the two strains showed different production abilities of epothilones. After 6 days of incubation, the yields of epothilones A and B increased 72.4 and 24.2%, and the total yields of epothilones were 9.9 mg/L in ZE9 and 13.7 mg/L in ZE9Δ*esi*, respectively (Fig. 2b). The deletion of *esi* significantly increased the production of epothilones in *Myxococcus* cells (*t* test, *p* value <0.01).

Furthermore, we constructed an *esi*-overexpressing mutant strain ZE9 *att::esi* by using plasmid pSWU30-p630-*esi* to introduce an additional *esi* gene into the *attB* site of ZE9 genome. Using the same process, the empty plasmid pSWU30-p630 was also introduced into ZE9,



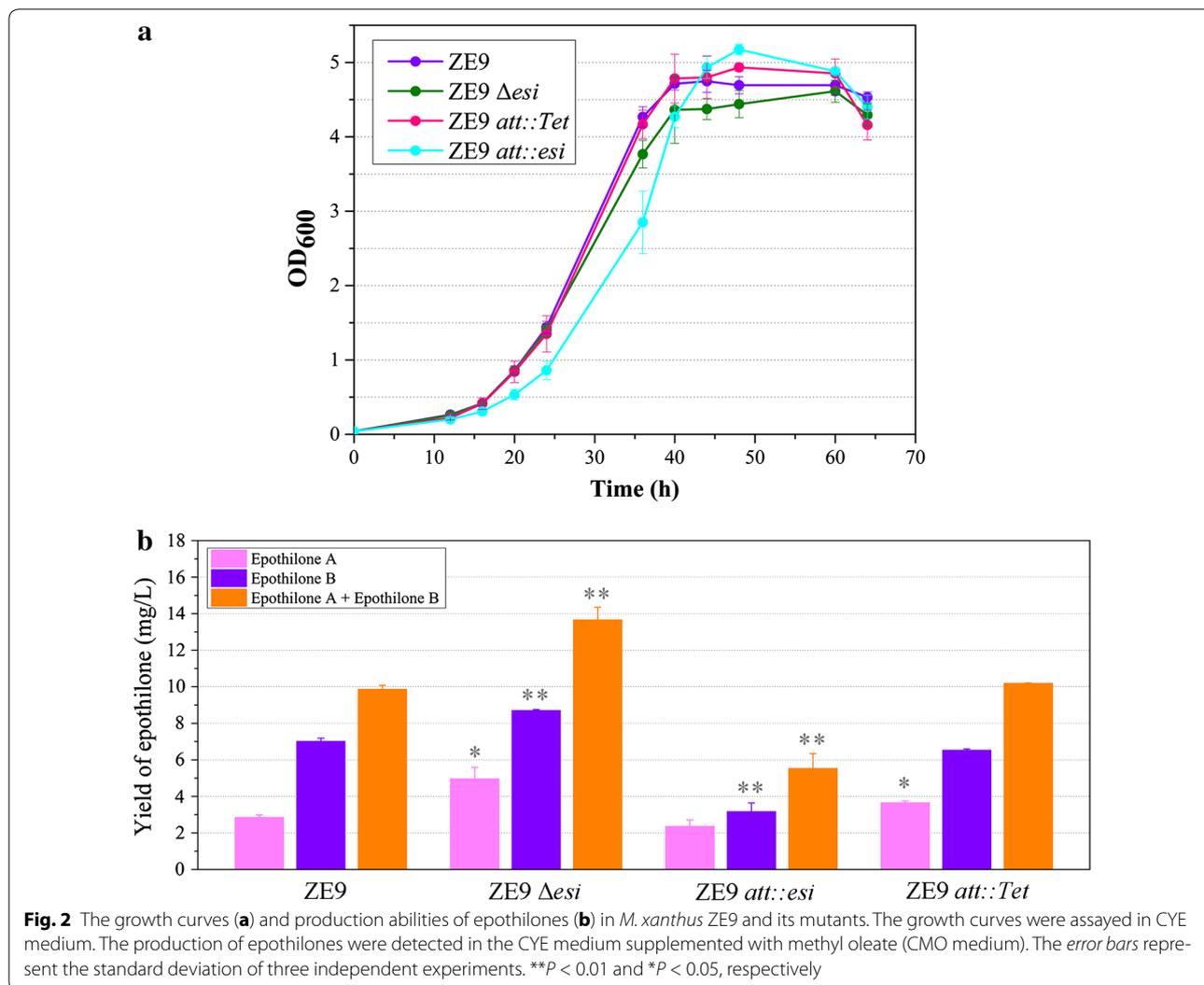
**Fig. 1** The biosynthetic gene cluster and productions of epothilones in *S. cellulosum* So0157-2 and So0157-2-*esi*<sup>-</sup>. **a** A physical map of the epothilone gene cluster, showing the upstream *esi* gene SCE1572\_25030, which was insertional inactivation (red arrow). The 17-bp sequence containing the four inverted repeat bases (in red) locates in the epothilone promoter. **b** Production of epothilones in So0157-2 and the *esi*<sup>-</sup> mutant in EMP medium. The error bars represent the standard deviation of three independent experiments. \*\**P* < 0.01

forming the ZE9 *att::Tet* strain as a control. These two strains were inoculated in CYE medium supplemented with 5 µg/mL tetracycline for 48 h, which were used as seeds to inoculate fresh CYE medium to assay their growth curves. The results showed that the ZE9 *att::esi* and ZE9 *att::Tet* strains also had a similar growth curve as the ZE9 strain (Fig. 2a). However, the introduction of an additional *esi* gene obviously suppressed the epothilone production (9.9 mg/L in ZE9 vs. 5.5 mg/L in ZE9 *att::esi*; *t* test, *p* value < 0.01), whereas the ZE9 *att::Tet* had a similar epothilone yield as ZE9 (*t* test, *p* value > 0.05) (Fig. 2b). Accordingly, Esi is a negative regulator for the production of epothilones, functioning not only in the producing *Sorangium* strains, but also in *Myxococcus*.

#### Esi negatively regulates the transcription of epothilone biosynthetic genes

To investigate the regulation mechanism of Esi on the production of epothilones, we comparatively analyzed

the transcriptional levels of epothilone biosynthetic genes in the *esi*-deletion and overexpression mutants with their parent strain *M. xanthus* ZE9 using quantitative real-time polymerase chain reaction (RT-qPCR). In ZE9 strain, compared with the expression of the *epoA* gene, the *esi* gene expressed at a low level (Fig. 3). In the early growth stage (24 h of incubation), expression of the *esi* gene was approximately 21% of the *epoA* gene; then weakly decreased with the increase of incubation. After the growth, the expression level of *esi* gene was greatly decreased, remaining less than 2% of the 24 h-expression at the 60 h of incubation (the stable stage of the growth curve, referred to Fig. 2b). Comparatively, in the ZE9 *att::esi* mutant, the additional *esi* gene was under the control of the 630-bp promoter of *pilA*, which held the total expression level of *esi* at high levels. The *esi* expressions at the 24, 48 and 60 h were 4.35-, 3.16- and 2.70-fold of that at the 24 h in the ZE9 strain (Fig. 3).



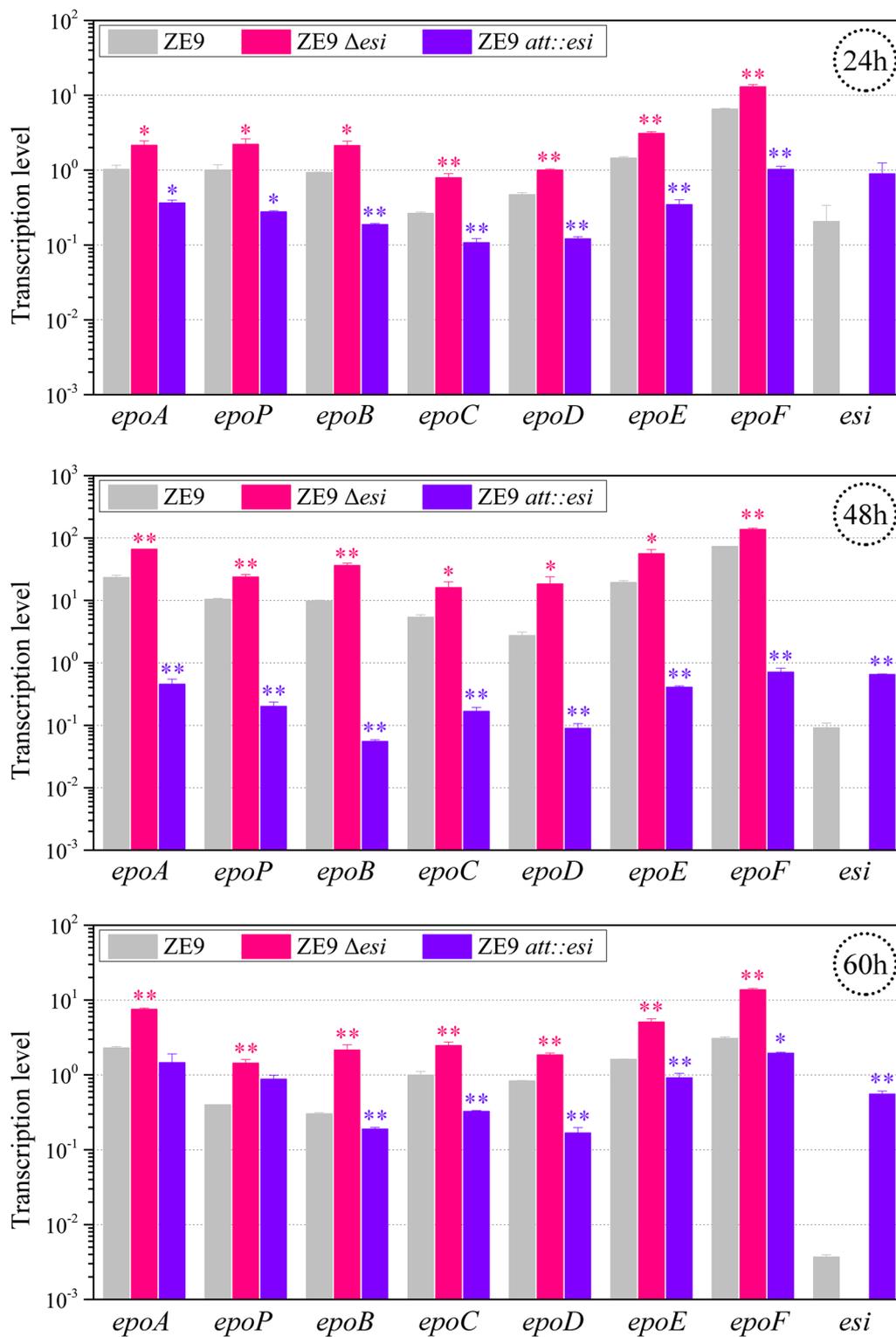
Similar to that in our previous report [22], transcriptional levels of the epothilone biosynthetic genes markedly varied in ZE9. With the increase of incubation time, transcriptions of the seven ORFs in the gene cluster each increased approximately ten times during the growth stage from 24 to 48 h, and then rapidly decreased in the following 12 h (Fig. 3). The transcriptional pattern was consistent with the production curves of epothilones in *M. xanthus* cells [22] or *S. cellulosum* cells [16], which was different from other microbial secondary metabolisms, usually expressed in stationary growth phase [2, 4].

After the deletion of the *esi* gene, the transcriptional level of each of the seven epothilone ORFs was markedly up-regulated at different incubation time points ( $t$  test,  $p$  value  $< 0.05$  or  $p$  value  $< 0.01$ ) (Fig. 3). The transcriptional increases ranged from 99% (*epoF*) to 200% (*epoC*) at 24 h, 88% (*epoF*) to 574% (*epoD*) at 48 h and 124% (*epoD*) to 611% (*epoB*) at 60 h of incubation. However, with the

overexpression of *esi*, the epothilone genes were correspondingly down-regulated in the ZE9 *att::esi* strain (Fig. 3). The above RT-qPCR analysis demonstrated that Esi negatively regulated the transcription of epothilone genes.

#### The structure model of Esi and its potential regulation mechanism

It is known that the repression of transcription initiation often occurs simply by steric hindrance [29]. From the above results, we suggested that the *esi* protein product probably functioned by directly binding on the epothilone promoter sequence to block the transcription. Esi (WP\_020736929.1) is a protein with 373 amino acids in size. Domain prediction by the SMART program showed that the protein had no signal peptide or transmembrane domain. We constructed the three-dimensional (3D) structure of the Esi protein using the



**Fig. 3** RT-qPCR analysis of expressions of the *esi* gene and the seven epothilone ORFs in the ZE9, ZE9  $\Delta esi$  and ZE9 *att::esi* strains after 24, 48 and 60 h of incubation in CMO medium. The expression level of *epoA* in the ZE9 strain at 24 h of incubation was set to be 1. The error bars represent the standard deviation of three independent experiments. \*\* $P < 0.01$  and \* $P < 0.05$ , respectively

I-TASSER and QUARK programs. All the structural models were refined at the atomic level by using the fragment-guided molecular dynamics (FG-MD) simulations. Based on the constructed model, the Esi protein contained three regions, i.e. an N-terminal domain, a linker region and a C-terminal domain, with the isoelectric points of 6.90, 5.87 and 5.95, respectively (Fig. 4a). The 106-aa N-terminal domain had three parallel  $\beta$ -pleated sheets and three alpha helices, the 248-aa C-terminal domain was a Cupin-like domain (pfam13621), and the linker region was an alpha helix (Fig. 4b). According to the quality assessment by Ramachandran plot (Additional file 1: Figure S1), the accuracy of protein structural models (80.6% of the residues in favored region) was acceptable [40].

The repressors, usually in dimer or further multimer form, bind to the operator sequence of promoter and thus block the transcription; and most operators contain direct or inverted repeats of a 4–5 base pair sequence for the binding [34]. We further constructed potential DNA–protein complex structure. We reconstructed a 3D structure of the epothilone promoter using w3DNA, and then calculated to determine the best complex structure of Esi-promoter by using nucleic acid-protein dock (NPDock). The results suggested that the N-terminal domain of Esi connected with the promoter DNA sequence (Fig. 4c). We checked the promoter sequence and found a special 17-bp DNA sequence (–137 to –153 bp relative to the transcriptional start site), which contains a 4-bp inverted repeat sequence (Fig. 1a). According to the DNA–protein complex structure, we suggested that the inverted repeat was the best Esi-binding region. The Esi protein might work as a dimer, and the two N-terminal domains bind specifically to the inverted GGGA of the double-strand promoter DNA sequence.

#### **Esi negatively regulates on epothilone promoter sequence**

In our previous report, we revealed that a 24-bp fragment in the epothilone promoter probably played a regulation role for the RNA polymerase functions [36]. Interestingly, this fragment includes the specific inverted repeat sequence. To determine the molecular mechanism of Esi regulation on the epothilone synthesis, we assayed the binding activity between Esi and the 17-bp operator DNA sequence of epothilone promoter. We constructed the *esi* gene into the expression vector pET-28a, forming pET-28a-*esi*, which was expressed in *E. coli* BL21 (DE3) cells. The Esi protein was tagged with a 6-His sequence. After induction with 0.1 mM IPTG, the Esi proteins were extracted from the cultivated cells, and purified based on the fused His-tag (Fig. 5a).

The double-strand 17-bp sequence was obtained by PCR annealing with the primers labeled with biotin.

Electrophoretic Mobility Shift Assay (EMSA) was performed to detect the specific binding activity of Esi and the 17-bp sequence. Each of the reaction system contained biotin-labeled DNA and LightShift™ Poly (dI–dC). The results clearly showed that the His-Esi protein was able to bind to the biotin-labeled DNA. If the molar ratio of protein and DNA was increased, the more DNA was blocked (Fig. 5b), which indicated that the His-Esi protein bound specifically with operator DNA, rather than the nonspecific competition DNA (Poly dI–dC).

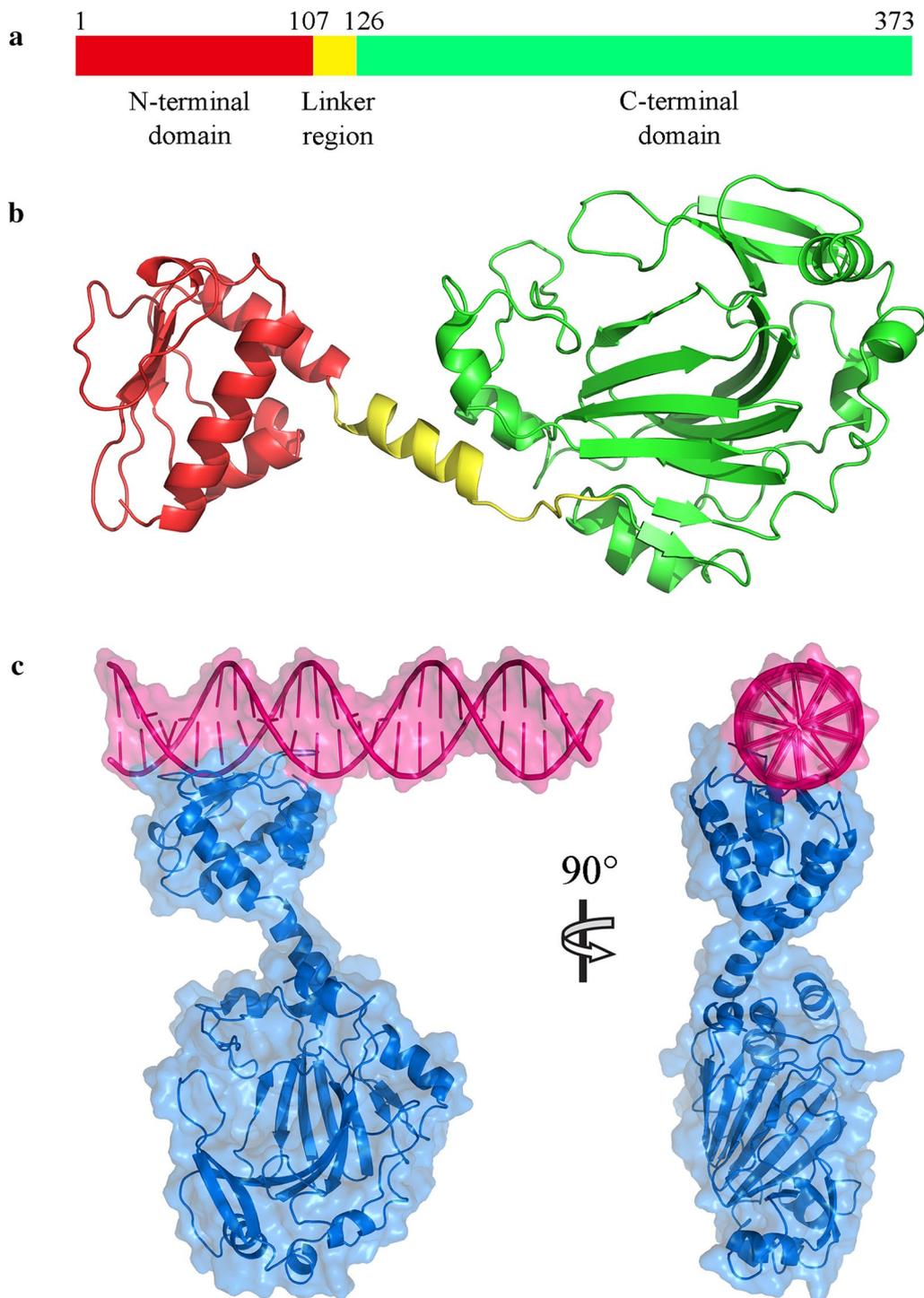
#### **Existence of *esi* gene in bacteria**

Using the Esi amino acid sequence as a seed, we iteratively searched the homologous sequences in RefSeq database using the psi-BLAST program, and found 156 Esi-homologues with more than 90% of alignment coverage. The genes encoding for the 156 Esi homologous proteins distributed in 203 different bacterial genomes (Additional file 2: Table S1), all belonging to the Proteobacteria. These genomes contained single *esi* gene copy, except *Pseudomonas fluorescens* FW300-N2C3, which contained two genes encoding for the Esi homologous proteins. The *esi* gene distributed mainly in *Pseudomonas* (143 of the 203 genomes), *Acinetobacter* (41 genomes) and *Burkholderia* (6 genomes). Surprisingly, among more than 20 sequenced myxobacterial genomes, the *esi* gene was only found in the *Stigmatella aurantiaca* strain DW4/3-1 and the *S. cellulosum* strains So0157-2 and So0163. The results indicated that the Esi regulation mechanism was not universal in myxobacteria.

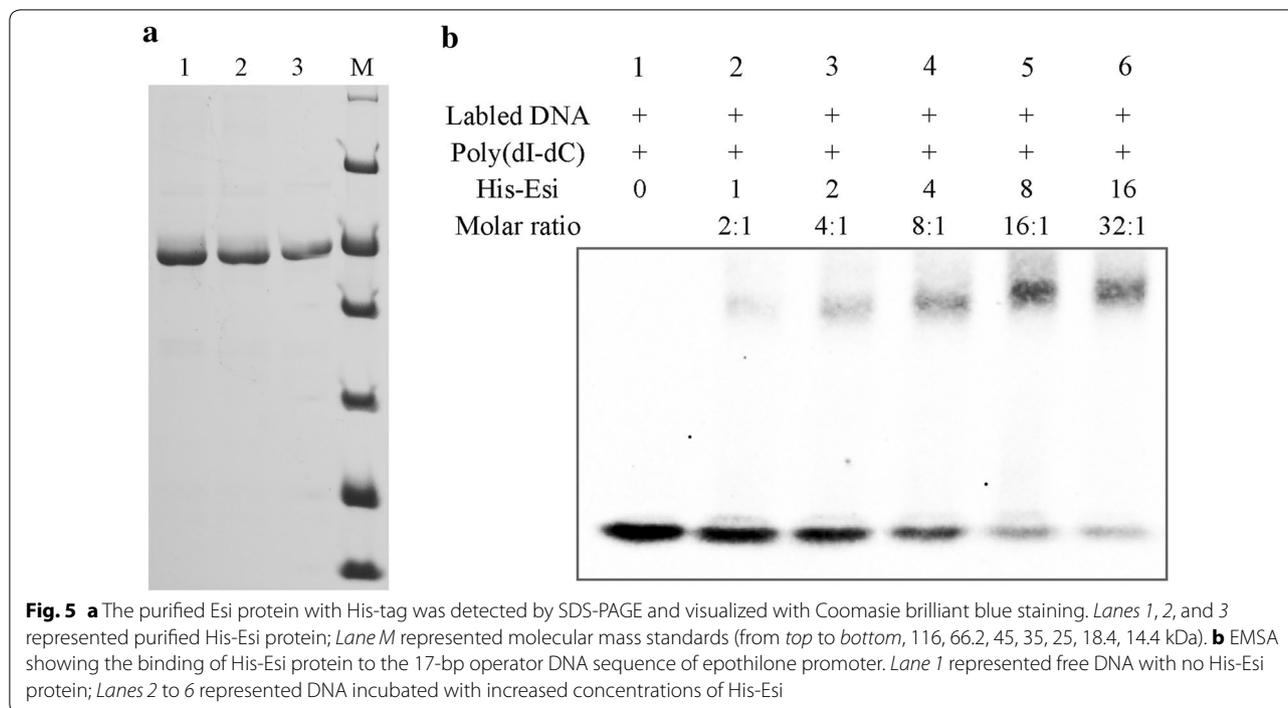
All of the Esi proteins had a length ranging from 356aa to 419aa, and their functions have not been reported, indicating that the *esi* gene codes a novel negative regulator. The evolutionary conservation of each amino acid residue in the Esi homologue proteins was calculated using ConSurf [41], which showed that the C-terminal domain had the highest conservation, followed by the N-terminal domain and the linker region (Additional file 3: Figure S2). Because of high similarities of their amino acid sequences, these homologous proteins were suggested to have similar 3D structures as the Esi protein described in this paper, probably playing functions as transcriptional regulation factors like the Esi protein.

#### **Conclusions**

In this paper, we determined that the Esi protein played as a transcription regulator for epothilone biosynthesis. Inactivation of *esi* gene markedly promotes epothilone production. The Esi binds, probably via the N-terminus of the protein, to the inverted repeat of epothilone promoter and down-regulates the transcription of each gene in the gene cluster to suppress epothilones synthesis. Esi regulation is not universal in myxobacteria. This is the



**Fig. 4** The structure modeling of Esi protein and the DNA-Esi binding complex. **a** Positions of N-terminal domain, linker region and C-terminal domain. **b** The structure model of Esi protein. The protein backbone is shown in Cartoon mode. There were three parallel  $\beta$ -pleated sheets and three alpha helices in the N-terminal domain and an alpha helix in the linker region. The C-terminal region is a cupin-like domain. **c** The DNA-protein complex structure of Esi protein (blue) and epothilone promoter (red). The top of the N-terminal domain of Esi binds with the promoter DNA sequence



first transcription regulator identified for the biosynthesis of epothilones.

## Discussion

The Esi protein contained three regions, i.e. an N-terminal domain, a linker region and a C-terminal domain. The Esi protein connected with an operator of the promoter for the biosynthesis of epothilones, and the operator sequence contained a 4-bp inverted repeat in a 17-bp DNA fragment. We speculated that Esi worked in a similar way as HrcA [32, 33]: two molecules of Esi protein interacted to each other through the C-terminal domain, forming a mirrored dimer and the two N-terminal domains bind to the 4-bp inverted repeat of each DNA strand to repress transcription initiation. Although bacterial Esi proteins have not yet been studied, many of the homologous proteins were annotated to be cupin, a diverse superfamily of proteins named after its conserved barrel domain [42, 43] or transcription factor jumonji, a protein having a DNA binding domain and two conserved *jmj* domains [44]. The JmjC domains have been identified to play transcription factors in numerous eukaryotic proteins, such as PHD, C2H2, ARID/BRIGHT and zinc fingers, functioning by a histone demethylation mechanism that is conserved from yeast to human [44, 45]. However, since there is no histone in prokaryotes, the functioning mechanism of Esi was not like the eukaryotic JmjC proteins, and detailed molecular functioning pattern requires further investigations.

## Methods

### Bacterial strains, plasmids and culture conditions

Bacterial strains and plasmids used in this study are listed in Additional file 4: Table S2. *E. coli* strains XL1-Blue, DH5 $\alpha$   $\lambda$  pir and DH5 $\alpha$  were used for cloning of pBJ113 plasmids, pSWU30 plasmids [46] and pET-28a plasmids, respectively. *E. coli* BL21 (DE3) was used to express the Esi protein. *E. coli* strains were grown at 37 °C in Luria-Broth (LB) medium (10 g/L peptone, 5 g/L yeast extract and 5 g/L NaCl, pH 7.2), and *M. xanthus* strains were grown at 30 °C in CYE medium [10 g/L casitone, 5 g/L yeast extract, 10 mM 3-(N-morpholino) propanesulfonic acid (MOPS) and 4 mM MgSO<sub>4</sub>, pH 7.6] or CMO medium [10 g/L casitone, 5 g/L yeast extract, 10 mM MOPS, 4 mM MgSO<sub>4</sub> and 7 mL/L methyl oleate, pH 7.6]. The medium were supplemented with the following different antibiotics if required: ampicillin [Amp], 100  $\mu$ g/mL; kanamycin [Km], 40  $\mu$ g/mL; Apramycin, [Apra], 30  $\mu$ g/mL, tetracycline [Tet], 5  $\mu$ g/mL, and Gentamycin [Gm]. The primers used in constructing vectors and mutant strains were list in Additional file 5: Table S3.

### Inactivation of *esi* gene in So0157-2

Inactivation of *esi* was performed by insertion mutation. Using genomic DNA of So0157-2 as template, the homologous arm was amplified by PCR with primers *esi* F1 and *esi* R1. The PCR fragment was phosphorylated with T4 Polynucleotide Kinase (Fermentas, Canada) and then cloned into the SmaI restriction site of pCC11 [38],

leading to the insertion plasmid pCC-*esi*. After verification by DNA sequence analysis, the pCC-*esi* was transformed into *E. coli* DH5 $\alpha$  ( $\lambda$  *pir*) harbouring the pRK2003(Km<sup>r</sup>) plasmid, the resulting strain containing double plasmids was designed as ET-*esi*. The transformation of the plasmid pCC-*esi* into the genome of the So0157-2 was performed by biparental mating and conjugation, as described previously [38]. The *E. coli* ET-*esi* and So0157-2 were mixed with a ratio of 1:50, inoculated on the nitrocellulose membrane on VY/2 plate (7  $\mu$ g/mL of Gm, 5  $\mu$ g/mL of Cm), incubated for 40 h and then transferred to the selection VY/2 plate (15  $\mu$ g/mL of Gm, 3  $\mu$ g/mL of Cm). The clones appeared after two weeks of incubation at 30 °C were verified by Tail-PCR (data not shown) as described previously [22] and purified on CNST plate (15  $\mu$ g/mL of Gm, 10  $\mu$ g/mL of Cm). A Genome Walking kit (TaKaRa, Japan) was used to amplified the flanking sequence of insertion sites with three specific primers, Cm1, Cm2 and Cm3, designed in this work and a random primer AP1 provided by the kit. The PCR products with appropriate size from the third PCR round were withdrawn by the DNA Extraction kit (Promega, USA) and sequenced after cloning to the pMD19-T vector (Takara, Japan).

#### Knockout and overexpression of *Esi* gene in ZE9

*Myxococcus xanthus* ZE9 was constructed with inserting the epothilone biosynthetic gene cluster and some flanking sequence, including the *esi* gene, into the genome of DZ2 by transposition in previous work [22]. The deletion of *esi* gene was performed by construction of knockout vector pBJ-*esi*. The arms for homologous recombination were amplified from the genome of So0157-2 by PCR using the primers *esi*-up F and *esi*-up R for upstream arm and *esi*-down F and *esi*-down R for downstream arm. The 1.6-Kb fragment of *esi*-ud was amplified by overlap PCR with *esi*-up mixed with *esi*-down as template and the primers *esi*-up F and *esi*-down R, digested with EcoRI and KpnI, and then cloned into pBJ113, constructing the knockout vectors pBJ-*esi*. The pBJ-*esi* was introduced into ZE9 by electroporation as described previously [22]. Resistant colonies that appeared after 6 days of incubation on CYE plates with 40  $\mu$ g/mL of Km were checked by colony PCR with primers KG-test F and KG-trst R. The Km-resistant colonies were resuspended in 2.5 mL-tube with CYE liquid medium and diluted in gradient for 10<sup>2</sup>–10<sup>5</sup> times, mixed with 2.5 mL of 0.5% soft agar and the mixture were spread on CYE selection agar plates containing 0.1% galactose to select the second recombination event. The galactose-resistant but Km-sensitive colonies were screened and checked by colony PCR with primers *Esi*-up F and *Esi*-down R. The knockout mutant, designated as ZE9 $\Delta$ *Esi*, was confirmed by sequence analysis of the PCR products.

To construct the overexpression mutant, the 630-bp promoter of the *pilA* gene (MXAN\_RS28035) was amplified from the genomic DNA of DK1622 by PCR using the primers p630 F and p630 R. The PCR fragment was digested with HindIII and XbaI and then ligated into the plasmid pSWU30 [46], generating the vector of pSWU30-p630. The 1.2-Kb *esi* gene was amplified from the genomic DNA of So0157-2 by PCR with the primers *esi* F2 and *esi* R2 and ligated into XbaI and KpnI sites of pSWU30-p630 to produce the overexpression vector pSWU30-p630-*esi*. The pSWU30-p630-*esi* was electroporated into ZE9 and Tet-resistant colonies were screened by colony PCR using the primer p630 F and *esi* R2. The mutant strain ZE9 *att::Esi* was purified by being gradient diluted, mixed with soft agar (0.5%), and spread on CYE plates with 5  $\mu$ g/mL of Tet. As a control, the empty plasmid pSWU30-p630 was also introduced into ZE9, producing ZE9 *att::Tet*.

#### Transcriptional analysis of *esi* gene and epothilone genes

The mutant strains and wild type strains were cultured in 50 mL of CYE liquid medium for 20 h (OD600 was about 1.5), then transplanted into fresh CMO medium with a start OD600 of 0.04, and harvested at the early stage, middle stage and stable stage of exponential growth successively. Total RNA were extracted according to the protocol provided by the BIOZOL Toal RNA Extraction Regent (BioFast, China), and transcribed reversely into cDNA with PrimeScript<sup>TM</sup> Regent Kit with DNAase (Takara, Japan). We analyzed the relative transcriptional level of the *esi* gene and epothilone gene cluster by RT-qPCR on LightCycler<sup>®</sup> 480 (Switzerland) with SYBR<sup>®</sup> Premix Ex Taq<sup>TM</sup> GC Dye (Takara, Japan). The *gapA* gene (MXAN\_RS13645), encoding glyceraldehyde-3-phosphate dehydrogenase, was used as the reference gene. The primers used in RT-qPCR were list in Additional file 6: Table S4.

#### Extraction and detection of epothilones

So0157-2 and So0157-2-Esi<sup>-</sup> were cultured in M26 medium (8 g/L potato starch, 2 g/L soy peptone, 2 g/L glucose, 2 g/L yeast powder, 1 g/L MgSO<sub>4</sub>, 1 g/L CaCl<sub>2</sub>, 1 mL/L EDTA-Fe<sup>3+</sup>, 1 mL/L microelement, pH 7.2) for 3 days and inoculated into 50 mL of EPM medium (4.7 g/L dextrin, 1.7 g/L soybean cake powder, 0.8 g/L glucose, 0.5 g/L saccharose, 2.2 g/L MgSO<sub>4</sub>·7H<sub>2</sub>O, 1 g/L CaCl<sub>2</sub>, 2 mL/L EDTA-Fe<sup>3+</sup>, 1 mL/L corn steep liquor, 1 mL/L microelement, pH 7.5) supplemented with 2% of resin XAD 16 for the absorbion of epothilone products. ZE9 and *esi* mutants were grown overnight in 50 mL of CYE medium supplemented with Apra (30  $\mu$ g/mL). The cultures were inoculated at a ratio of 2:100 into 50 mL of CMO medium containing 2% of the XAD-16 resin.

The resin were harvested with strainer after 9 days (for *S. cellulosum*) and 6 days (for *M. xanthus*) and extracted with 3 mL of methanol by shaking at room temperature overnight [16]. The supernatant was centrifuged for 10 min at 12,000 rpm and filtered with 0.22  $\mu$ m filter to remove the impurities. 20  $\mu$ L of the sample was injected into High Performance Liquid Chromatography (HPLC, SHIMADZU, Japan) and analyzed on a Shim-pack MRC-ODS RP C18 column (4.6 mm  $\times$  250 mm, 4.60  $\mu$ m; Shimadzu, Japan) and monitored at 250 nm, with a mobile phase of 60% of methanol (HPLC grade) and 40% of H<sub>2</sub>O at a flow rate of 1.0 mL/min. The yield of epothilone was quantified from the peak area in the UV chromatogram, by reference against a calibration standard.

### Bioinformatics analysis of Esi

The signal peptides and transmembrane regions were predicted with SignalP [47] and TMHMM [48], respectively. Esi protein domain was annotated by SMART [49] and PFAM [50]. We constructed the N-terminal domain of Esi protein by ab initio protein folding and protein structure prediction with QUARK program [51], and then modeled the three-dimensional structures model of the whole Esi protein using the I-TASSER program based on a threading approach [52]. All the structural models were refined in the atomic-level by the fragment-guided molecular dynamics (FG-MD) simulations [53]. According to quality assessment by Ramachandran plot [40], the accuracy of the protein structural models were acceptable. The three-dimensional nucleic-acid structures of the epothilone promoter sequence was reconstructed with w3DNA server [54]. The NPdock (Nucleic acid-Protein Dock) was used to model the complex structures of Esi protein and epothilone promoter [55].

The position-specific iterative basic local alignment search tool (PSI-BLAST) [56] was used to search for homologous proteins of Esi in the US National Center for Biotechnology Information (NCBI) reference sequence (RefSeq) database [57]. The amino acid sequence of Esi protein was used as query sequence. A multiple sequence alignment of the Esi protein and its homologous proteins was established using the MAFFT program [58]. The evolutionary conservation of amino acid positions in the Esi protein and its homologous sequences was estimated by using ConSurf algorithm [41]. The JTT substitution matrix was used and the computation was based on the empirical Bayesian paradigm. The conservation scale was defined from the most variable amino acid positions (grade 1), which were considered to be evolved rapidly, to the most conservative positions (grade 9), which were considered to be evolved slowly. The sequence and modeled structure of the Esi protein were shown in nine-color

conservation grades. The statistical analysis was conducted using IBM SPSS Statistics.

### Expression and purification of Esi protein in *E. coli*

Expression vector was constructed by amplifying the *esi* gene fragment, fused with a His-tag at N-terminus, with PCR using the primers of *esi* F3/*esi* R3 and ligating it into the XbaI and KpnI sites of the plasmid pET-28a, generating pET-28a-*esi*. *E. coli* BL21 (DE3) cells were transformed with plasmid pET-28a-*esi*. For the heterologous expression of Esi, 500 mL LB medium containing Km was inoculated with 5 mL of the overnight culture, prepared from a single colony, and grown at 37 °C for about 2 h. 0.1 mM IPTG was added at an optical density (OD<sub>600 nm</sub>) of 0.8. After 24 h of induced-expression at 16 °C, cells were harvested by centrifugation at 12,000 rpm for 5 min, resuspended with 1/10 volume of Lysis Buffer (25 mM Tris-HCl, 250 mM NaCl, pH 8.0) and disrupted by sonication. After centrifugation at 12,000 rpm for 30 min, the supernatant was incubated overnight with Ni-nitrilotriacetic acid HiszBind resin (Novagen), and then eluted with Lysis Buffer containing gradient imidazole ranging from 20 to 250 mM to collect soluble purified Esi protein.

### Electrophoretic mobility shift assay

The biotin-labeled DNA fragment of operator DNA sequence was generated by PCR annealing (95 °C, 3 min; 90 °C, 1 min; 83 °C, 1 min; 76 °C, 1 min; 70 °C, 1 min; 65 °C, 1 min; 60 °C, 30 min) using the following oligonucleotides: 5'-GGGACGAGCTAATTCCTCC-3' and 5'-GGGAATTAGCTCGTCCC-3'. The binding reactions were performed in 25  $\mu$ L final volume containing 2.5  $\mu$ L of reaction buffer (120 mM HEPES, 40 mM Tris, 600 mM KCl, 50 mM MgCl<sub>2</sub>, 1 mM EDTA, pH 8.0), 1  $\mu$ L of glycerol buffer (20% glycerol in 20 mM HEPES), 1  $\mu$ L of biotin-labeled DNA fragment (12 pmol), 1  $\mu$ L of Poly (dI-dC), N  $\mu$ L of protein and (17.5-N)  $\mu$ L of 50 mM NaH<sub>2</sub>PO<sub>4</sub>. The reaction samples were incubated at 30 °C for 30 min and then loaded into a 5% non-denaturing polyacrylamide gel to conduct electrophoresis with TGE buffer (14.3 g/L glycine, 3 g/L Tris, 1 mM EDTA, pH 8.0) for 2.5 h at 10 mA and 4 °C. The samples were electro-blotted onto an Amersham Hybond-N + membrane (GE, UK) by use of a mini transblot electrophoretic transfer cell (Bio-Rad). Hybridization and blotting was performed according to the protocol described by Chemiluminescent Nucleic Acid Detection Module (Thermo, USA). After completion of hybridization, the membrane was dried, autoradiographed, and the bands were detected by use of a Chemi Doc™ XRS + with Image Lab™ software (BIO-RAD).

## Additional files

**Additional file 1: Figure S1.** The ramachandran plot of the Esi structure model, which showed 80.6% residues in favored region, 13.5% residues in allowed region, 5.9% residues in outlier region.

**Additional file 2: Table S1.** Homologous amino acid sequences of Esi protein.

**Additional file 3: Figure S2.** Evolution conservative analysis of amino acid residues, calculated from the sequences of 156 Esi homologues. (A) Conservation scale is defined from the most variable residue sites (grade 1, color represented by turquoise; in rapid evolution) to conservative residue sites (grade 9, color represented by maroon; in slow evolution). (B) Conservation analysis of the Esi sequence sites in ConSurf grades.

**Additional file 4: Table S2.** Strains and plasmids used in this study.

**Additional file 5: Table S3.** Primers used in construction of vectors and mutant strains.

**Additional file 6: Table S4.** Primers used in RT-qPCR.

## Authors' contributions

XJY, ZFL and YZL designed the experiments. XJY, RP, PZ and XWC performed the experiments. XJY, ZZ, and YZL analyzed the data. XJY and YZL wrote the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

All data generated or analyzed during this study are included in this published article and in its supplementary information files.

## Funding

This work was financially supported by the National Natural Science Foundation of China (NSFC) (31471183 and 31670076) to YZL and (31370123) to ZFL.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 March 2017 Accepted: 15 May 2017

Published online: 23 May 2017

## References

- Macheleidt J, Mattern DJ, Fischer J, Netzker T, Weber J, Schroeckh V, Valiante V, Brakhage AA. Regulation and Role of fungal secondary metabolites. *Annu Rev Genet.* 2016;50:371–92.
- Bibb MJ. Regulation of secondary metabolism in streptomycetes. *Curr Opin Microbiol.* 2005;8:208–15.
- Bode HB, Muller R. Analysis of myxobacterial secondary metabolism goes molecular. *J Ind Microbiol Biotechnol.* 2006;33:577–88.
- Brakhage AA. Regulation of fungal secondary metabolism. *Nat Rev Microbiol.* 2013;11:21–32.
- Gerth K, Irschik H, Reichenbach H, Trowitzsch W. The myxovirescins, a family of antibiotics from *Myxococcus virescens* (Myxobacteriales). *J Antibiot (Tokyo).* 1982;35:1454–9.
- Gerth K, Bedorf N, Hofle G, Irschik H, Reichenbach H. Epothilons A and B: antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties. *J Antibiot (Tokyo).* 1996;49:560–3.
- Sasse F, Steinmetz H, Heil J, Hofle G, Reichenbach H. Tubulysins, new cytostatic peptides from myxobacteria acting on microtubuli. Production, isolation, physico-chemical and biological properties. *J Antibiot (Tokyo).* 2000;53:879–85.
- Gerth K, Pradella S, Perlova O, Beyer S, Muller 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;106:233–53.
- Rachid S, Sasse F, Beyer S, Muller R. Identification of StIR, the first regulator of secondary metabolite formation in the myxobacterium *Cystobacter fuscus* Cb f17.1. *J Biotechnol.* 2006;121:429–41.
- Rachid S, Gerth K, Kochems I, Muller R. Deciphering regulatory mechanisms for secondary metabolite production in the myxobacterium *Sorangium cellulosum* So ce56. *Mol Microbiol.* 2007;63:1783–96.
- Rachid S, Gerth K, Muller R. NtcA: a negative regulator of secondary metabolite biosynthesis in *Sorangium cellulosum*. *J Biotechnol.* 2009;140:135–42.
- Bollag DM, McQueney PA, Zhu J, Hensens O, Koupal L, Liesch J, Goetz M, Lazarides E, Woods CM. Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. *Cancer Res.* 1995;55:2325–33.
- Kowalski RJ, Giannakakou P, Hamel E. Activities of the microtubule-stabilizing agents epothilones A and B with purified tubulin and in cells resistant to paclitaxel (Taxol(R)). *J Biol Chem.* 1997;272:2534–41.
- Su DS, Balog A, Meng D, Bertinato P, Danishefsky SJ, Zheng YH, Chou TC, He L, Horwitz SB. Structure–activity relationship of the epothilones and the first in vivo comparison with paclitaxel. *Angewandte Chemie Int Ed Engl.* 1997;36:2093–6.
- Li PF, Li SG, Li ZF, Zhao L, Wang T, Pan HW, Liu H, Wu ZH, Li YZ. Co-cultivation of *Sorangium cellulosum* strains affects cellular growth and biosynthesis of secondary metabolite epothilones. *FEMS Microbiol Ecol.* 2013;85:358–68.
- Gong GL, Sun X, Liu XL, Hu W, Cao WR, Liu H, Liu WF, Li YZ. Mutation and a high-throughput screening method for improving the production of epothilones of *Sorangium*. *J Ind Microbiol Biotechnol.* 2007;34:615–23.
- Gong GL, Huang YY, Liu LL, Chen XF, Liu H. Enhanced production of epothilone B by immobilized *Sorangium cellulosum* in porous ceramics. *J Microbiol Biotechnol.* 2015;25:1653–9.
- Fu J, Wenzel SC, Perlova O, Wang J, Gross F, Tang Z, Yin Y, Stewart AF, Muller R, Zhang Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Res.* 2008;36:e113.
- Julien B, Shah S. Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*. *Antimicrobial Agents Chemother.* 2002;46:2772–8.
- Park SR, Park JW, Jung WS, Han AR, Ban YH, Kim EJ, Sohng JK, Sim SJ, Yoon YJ. Heterologous production of epothilones B and D in *Streptomyces venezuelae*. *Appl Microbiol Biotechnol.* 2008;81:109–17.
- Osswald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Muller R, Wenzel SC. Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth Biol.* 2014;3:759–72.
- Zhu LP, Yue XJ, Han K, Li ZF, Zheng LS, Yi XN, Wang HL, Zhang YM, Li YZ. Allopatric integrations selectively change host transcriptomes, leading to varied expression efficiencies of exotic genes in *Myxococcus xanthus*. *Microb Cell Fact.* 2015;14:105.
- Mutka SC, Carney JR, Liu Y, Kennedy J. Heterologous production of epothilone C and D in *Escherichia coli*. *Biochemistry.* 2006;45:1321–30.
- Tang L, Shah S, Chung L, Carney J, Katz L, Khosla C, Julien B. Cloning and heterologous expression of the epothilone gene cluster. *Science.* 2000;287:640–2.
- Ye W, Zhang W, Chen Y, Li H, Li S, Pan Q, Tan G, Liu T. A new approach for improving epothilone B yield in *Sorangium cellulosum* by the introduction of vgb epoF genes. *J Ind Microbiol Biotechnol.* 2016;43:641–50.
- Han SJ, Park SW, Park BW, Sim SJ. Selective production of epothilone B by heterologous expression of propionyl-CoA synthetase in *Sorangium cellulosum*. *J Microbiol Biotechnol.* 2008;18:135–7.
- Molnar I, Schupp T, Ono M, Zirkle R, Milnamow M, Nowak-Thompson B, Engel N, Toupet C, Stratmann A, Cyr DD, et al. The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. *Chem Biol.* 2000;7:97–109.

28. Julien B, Shah S, Ziermann R, Goldman R, Katz L, Khosla C. Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*. *Gene*. 2000;249:153–60.
29. Browning DF, Busby SJ. The regulation of bacterial transcription initiation. *Nat Rev Microbiol*. 2004;2:57–65.
30. Gralla JD. Activation and repression of *E. coli* promoters. *Curr Opin Genet Dev*. 1996;6:526–30.
31. Zuber U, Schumann W. CIRCE, a novel heat shock element involved in regulation of heat shock operon dnaK of *Bacillus subtilis*. *J Bacteriol*. 1994;176:1359–63.
32. Wiegert T, Schumann W. Analysis of a DNA-binding motif of the *Bacillus subtilis* HrcA repressor protein. *FEMS Microbiol Lett*. 2003;223:101–6.
33. Liu J, Huang C, Shin DH, Yokota H, Jancarik J, Kim JS, Adams PD, Kim R, Kim SH. Crystal structure of a heat-inducible transcriptional repressor HrcA from *Thermotoga maritima*: structural insight into DNA binding and dimerization. *J Mol Biol*. 2005;350:987–96.
34. Browning DF, Busby SJ. Local and global regulation of transcription initiation in bacteria. *Nat Rev Microbiol*. 2016;14:638–50.
35. Li ZF, Zhao JY, Xia ZJ, Shi J, Liu H, Wu ZH, Hu W, Liu WF, Li YZ. Evolutionary diversity of ketoacyl synthases in cellulolytic myxobacterium *Sorangium*. *Syst Appl Microbiol*. 2007;30:189–96.
36. Han K, Li ZF, Peng R, Zhu LP, Zhou T, Wang LG, Li SG, Zhang XB, Hu W, Wu ZH, et al. Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Sci Rep*. 2013;3:2101.
37. Zhu LP, Li ZF, Sun X, Li SG, Li YZ. Characteristics and activity analysis of epothilone operon promoters from *Sorangium cellulosum* strains in *Escherichia coli*. *Appl Microbiol Biotechnol*. 2013;97:6857–66.
38. Xia ZJ, Wang J, Hu W, Liu H, Gao XZ, Wu ZH, Zhang PY, Li YZ. Improving conjugation efficacy of *Sorangium cellulosum* by the addition of dual selection antibiotics. *J Ind Microbiol Biotechnol*. 2008;35:1157–63.
39. Lau J, Frykman S, Regentin R, Ou S, Tsuruta H, Licari P. Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*. *Biotechnol Bioeng*. 2002;78:280–8.
40. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins*. 2003;50:437–50.
41. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res*. 2010;38:W529–33.
42. Dunwell JM. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol Genet Eng Rev*. 1998;15:1–32.
43. Dunwell JM, Purvis A, Khuri S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry*. 2004;65:7–17.
44. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y. Histone demethylation by a family of JmjC domain-containing proteins. *Nature*. 2006;439:811–6.
45. Clissold PM, Ponting CP. JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2beta. *Trends Biochem Sci*. 2001;26:7–9.
46. Wu SS, Kaiser D. Genetic and functional evidence that type IV pili are required for social gliding motility in *Myxococcus xanthus*. *Mol Microbiol*. 1995;18:547–58.
47. Petersen TN, Brunak S, von Heijne G, von Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8:785–6.
48. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001;305:567–80.
49. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res*. 2015;43:D257–60.
50. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*. 2016;44:D279–85.
51. Xu D, Zhang Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins*. 2012;80:1715–35.
52. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 2015;12:7–8.
53. Zhang J, Liang Y, Zhang Y. Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. *Structure*. 2011;19:1784–95.
54. Zheng G, Lu XJ, Olson WK. Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res*. 2009;37:W240–6.
55. Tuszynska I, Magnus M, Jonak K, Dawson W, Bujnicki JM. NPdock: a web server for protein-nucleic acid docking. *Nucleic Acids Res*. 2015;43:W425–30.
56. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Res*. 2008;36:W5–9.
57. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. 2016;44:D733–45.
58. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–80.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

