Avermectin and its analogues, produced by the actinomycete *Streptomyces avermitilis* and are widely used in the field of animal health, agriculture, and human health. Here we have adopted a practical approach to successfully improve avermectin production in an industrial overproducer. Transcriptional levels of the wild-type strain and industrial overproducer in production cultures were monitored using microarray analysis. The avermectin biosynthetic genes, especially the pathway-specific regulatory gene, *aveR*, were up-regulated in the high-producing strain. The upstream promoter region of *aveR* was predicted and proved to directly be recognized by $\sigma^R$ in vitro. A mutant library of *hrdB* gene was constructed by error-prone PCR and selected by high-throughput screening. As a result of evolved *hrdB* expressed in the modified avermectin high-producing strain, 6.38 g/L of avermectin B1a was produced with over 50% yield improvement, in which the transcription level of *aveR* was significantly increased. The relevant residues were identified to center in the conserved regions. Engineering of the *hrdB* gene can not only elicit the overexpression of avermectin but also allows for simultaneous transcription of many other genes. The results indicate that manipulating the key genes revealed by reverse engineering can effectively improve the yield of the target metabolites, providing a route to optimize production in these complex regulatory systems.

Precision engineering | RNA polymerase | overproduction

Most drugs and drug precursors found in natural organisms need to be produced on a large scale by microbial fermentation in the pharmaceutical industry (1, 2). Because these compounds are produced in very small quantities under natural conditions and their total synthesis is difficult and impractical because of their complexity and diversity of chemical structure, strain improvement has become a critical part of the drug development process (3, 4). Random mutagenesis and screening has traditionally been used for strain improvement as it is technically simple to perform and little knowledge is needed about the strains. With these methods, the productivity of a penicillin strain used in industry was improved by 100,000-fold compared with the original Fleming strain (5). Even though a tremendous increase in production has been achieved, there are certain drawbacks. The traditional mutate-and-screen methods suffer from being time- and labor-intensive; iterative rounds of mutagenesis and selection are normally needed for yield improvement, which risk accumulation of unwanted mutations resulting in crippled strains; and the molecular genetic basis underlying such yield enhancement is largely unknown. Thus the overproduction mechanisms cannot be applied for the rational engineering of a better strain (6).

Since recombinant DNA technology has become increasingly sophisticated, the ability to modify specific genes and pathways for optimizing the production of commercial-significant microbial metabolites has advanced rapidly (7). The “omics”-guided technologies have also seen rapid development in recent years and are likely to become valuable tools for strain development (3, 6, 8–10). Reverse engineering an organism with a desirable property is carried out by identifying the genetic basis of the property and then directly engineering the key genes, the aim of which is to rapidly gain insight into the complex mechanisms controlling microbial metabolism. Reverse engineering has successfully improved the yield of lovastatin in *Aspergillus terreus* by association analysis of transcriptional and metabolite profiling data to elucidate the key genetic components and physiological traits that impact the production of lovastatin (11). This approach can be utilized to elucidate the interrelationships between physiological traits and more efficiently direct the engineering of target compound-producing strains to synthesize high yields of these natural products.

Avermectin and its analogues, produced by the actinomycete *Streptomyces avermitilis*, are a series of 16-membered pentacyclic macrolactone type I polypeptides that exhibit an excellent anthelmintic activity against a variety of nematode and arthropod parasites. With few side effects against the host organism, they are widely used in the field of animal health, agriculture, and human infections (12–14). With the published genomic information of *S. avermitilis*, the avermectin biosynthetic pathway was investigated, and it was deduced whether the strain had the ability to produce novel secondary metabolites (15, 16). The avermectin biosynthetic gene cluster (SAV935–SAV953) encodes four multifunctional modular polypeptide synthase components, AVES 1, 2, 3, and 4, which each catalyze a specific round of polypeptide chain elongation, and includes 15 additional proteins that are responsible for the biosynthesis of oleandrose, glycosylation, postpolyketide modification, and the regulation of avermectin production (17). Seven acetate units and five propionate units are condensed to the starting units derived from 2-methylbutyrate or isobutyrate, catalyzed by the polypeptide synthases, in a stepwise process. After cyclization, the aglycone undergoes further modification including dehydration at C-22 and C-23, O-methylation at C-5, and glycosylation at C-13 (18). The biosynthesis of avermectins is closely related to the production of metabolites in primary metabolic pathways, and the expression levels of genes involved in those precursor metabolic-related pathways will influence the metabolite flux for avermectin biosynthesis.

The goal of this study was to identify the genetic basis of the high yield of avermectin B1a in an industrial avermectin overproducer generated by a successful mutate-and-screen program for strain improvement. By analyzing thousands of genes in parallel, and are likely to become valuable tools for strain development (3, 6, 8–10). Reverse engineering an organism with a desirable property is carried out by identifying the genetic basis of the property and then directly engineering the key genes, the aim of which is to rapidly gain insight into the complex mechanisms controlling microbial metabolism. Reverse engineering has successfully improved the yield of lovastatin in *Aspergillus terreus* by association analysis of transcriptional and metabolite profiling data to elucidate the key genetic components and physiological traits that impact the production of lovastatin (11). This approach can be utilized to elucidate the interrelationships between physiological traits and more efficiently direct the engineering of target compound-producing strains to synthesize high yields of these natural products.
reverse engineering should greatly facilitate the characterization of existing overproducing strains and generate information for guiding of more efficient engineering to synthesize higher yields of avermectin B1a. In this paper, an industrial strain, which was obtained by traditional strain improvement with avermectin B1a production of 4.167 mg/L, was chosen as the parent strain for genetic manipulation. It is difficult to successfully or continuously improve its yield through the traditional mutate-and-select method. It was hoped that the genetic information obtained would lead to better understanding of the metabolic basis of the industrial avermectin overproducer and ultimately trigger further improvement.

**Results**

**Transcriptional Comparative Analysis Reveals Altered Gene Expression in the Overproducer.** The avermectin B1a overproducing strain of *S. avermitilis* 3-115 (19, 20) and the wild-type strain ATCC31267 were grown in a fermentation medium. Comparative analysis of the microarray data revealed that 125 genes were differentially expressed and 31 genes were up-regulated by at least twofold at the stationary phase of 3-115 growth (Table S1). The twofold differentially expressed genes were divided into seven groups according to their biological functions defined by the Kyoto Encyclopedia of Genes and Genomes Pathway Database. Most of them were found to be involved in sig-regulons (e.g., sig37, sig3, and rslW), gene transcription and regulation, signal transponder (e.g., pitH1), and avermectin biosynthesis (e.g., aveR, aveAI, and aveAIII) (see Table S2). Avermectin biosynthetic genes and their relative expression levels are listed in Table S2. It was noticed that the high expression of aveR, aveAI, and aveAIII was statistically significant in the overproducer, which was further confirmed by RT-PCR analysis (Fig. 1). Transcriptional analysis revealed that the pathway-specific regulatory gene, aveR, and its promoter region in the 3-115 strain was the same as that in the strain ATCC31267, which encouraged us to explore the upstream regulatory factors. Several studies revealed that the major sigma factor, σ^66^ (*E. coli*), in Streptomyces coelicolor (21–23) recognizes and binds to the consensus sequence: 5′-TGACA-N₇₋₂₀-AAN₃-3′. In *S. avermitilis* 3-115, five of these up-regulated genes were predicted to be transcribed by the major holoenzyme containing σ^66^ (Table S1), which was consistent with the pleiotropic regulating effect of σ^66^ (24). The sequence of the predicted aveRp promoter was quite similar to the −10 and −35 hexamer consensus sequences of mrdP2 and dagAp4 (Fig. 2B), which were proved to be transcribed by holoenzyme containing σ^66^ (σ^66^) in *S. coelicolor* A3(2). This result indicated that the transcription of aveR may be initiated by the principal and essential σ^66^.

**Transcription of aveR Was Activated by RNA Polymerase (RNAP) Containing σ^66^ in Vitro.** To provide a biochemical basis for the assertion (based on sequence alignments) that the transcription of aveR was specifically recognized and activated by σ^66^, it was necessary to purify σ^66^ RNA polymerase holoenzyme. Therefore, hrdB of *S. avermitilis* 3-115 was overexpressed in *Escherichia coli* using the pET28b vector. A prominent band of M, 57,000 corresponding closely to the M, predicted for σ^66^ was purified (Fig. S1). To determine whether the purified protein possessed σ^66^ factor activity, and to evaluate whether in vivo the aveR promoter is directly dependent on σ^66^, titrated amounts of σ^66^ were added to the purified core RNA polymerase from *E. coli* to reconstitute Eσ^66^ holoenzyme and were used for in vitro run-off transcription experiments with linear DNA templates harboring the predicted promoters. As shown in Fig. 2B, a template containing the *S. avermitilis* mrdP was used as a positive control. Transcription assays with a template containing the aveRp resulted in transcripts whose levels increased as the σ^66^ concentration increased. The identity of the transcript seen with core RNA polymerase alone is not clear, but it may indicate initiation at aveR promoters, perhaps by a contaminating holoenzyme form. The in vitro transcription assays demonstrated biochemically that aveRp is recognized by Eσ^66^ in the absence of other transcriptional activators.

**Fig. 1.** Expression levels of aveAI, aveAIII, and aveR in *S. avermitilis* 3-115 relative to those in ATCC 31267. Samples were collected from 3-115 and ATCC 31267 after 36 h (vegetative growth) and 96 h (stationary phase and avermectin producing). We used 16 s rDNA as an RNA integrity control. Standard deviations are marked with error bars (n = 3). All the genes were examined by relative quantification real-time RT-PCR with gene-specific primers. **P < 0.01 vs. control strain.

**Fig. 2.** Recognition of the aveRp promoter by σ^66^-containing RNA polymerase. (A) Comparison of the putative aveRp promoter of *S. avermitilis* with mrdP2-p2 and daga-p4 of *S. coelicolor*. The conserved nucleotides within the −10 and −35 regions as well as transcription initiation sites are marked in bold. (B) In vitro transcription assay of the aveRp promoter by RNA polymerase containing σ^66^ (Eσ^66^). The promoter of the mrd homologous gene in *S. avermitilis* was used as a positive control. In all lanes, the Eσ^66^ holoenzyme was reconstituted by 0.25 pmol *E. coli* RNA polymerase core enzyme (EPICENTRE) with the indicated amounts of σ^66^.
High-Throughput Screening of a hrdB Mutant Library. To identify the effect of \( \sigma^{\text{hrdB}} \) on \( \text{aveR} \) and avermectin biosynthesis, especially its function on the regulation of avermectin biosynthesis, a library was constructed by random mutagenesis of the hrdB gene at a moderate mutation frequency (4.5–9 mutations/kb), which contained approximately \( 10^6 \) white colonies. The first round high-throughput screening was performed in 96 deep-well microplates, and 403 viable single colonies on MS plates were screened, from which 89 mutants were selected for the second fermentation in Erlenmeyer flasks to confirm the productivity. Reference strains C1 and C2 were simultaneously cultivated to investigate the influence of the plasmid and the additional copy of hrdB. The selection results are shown in Fig. 3, in which the A56 strain (pSET152-hrdB\(^{\text{393}}\)) and the A393 strain (pSET152-hrdB\(^{\text{393}}\)) showed increased production of avermectin B1a compared with the control (100% line). Production of avermectin B1a in strain C1 was comparable to that in strain C2, with slight reduction when compared with the parental strain 3-115. Because the aim of this work was to obtain a yield-improved strain, we chose to use the avermectin B1a production of parent strain containing only the chromosomal hrdB gene as the control for further studies.

Characterization of the Mutant hrdB Genes. To identify whether the yield improvement of A56 or A393 was elicited by the introduction of hrdB, the mutant fragments were rescued and reconjugationally transferred into 3-115. Strains 3-115, A56, and A393 were used as reference strains (Table S3). Strains 3-115::hrdB\(^{\text{A56}}\) and 3-115::hrdB\(^{\text{A393}}\) have comparable levels of avermectin production to that of A56 and A393, respectively. The result revealed that the yield-improved character of A56, and A393, was directly related to the addition of the mutant hrdB fragment. The mutant HrdB\(^{\text{A56}}\) was overexpressed and purified from \( E. \text{coli} \). The transcriptional activity of holoenzyme reconstituted by the purified His6-HrdB\(^{\text{A56}}\) and core enzyme was analyzed. The reconstituted HrdB\(^{\text{A56}}\)-holoenzyme showed greater than threefold transcriptional activity compared with HrdB\(^{115}\)-holoenzyme at the same concentration (Fig. 4). These results revealed that the enhanced transcriptional activity was responsible for the yield improvement.

Sequence Analysis of the Mutant hrdB Genes. For the mutants A56 and A393 that showed the remarkable yield increase and stability, the fragments hrdB\(^{\text{A56}}\) and hrdB\(^{\text{A393}}\) were sequenced. As shown in Fig. 5, the mutant sites were concentrated in region 1.1 and region 2.4. In the 512 AA of hrdB\(^{\text{A56}}\), there occurred six site mutations of A137S, K139E, E163G, M356V, V357A, and M389I. For the hrdB\(^{\text{A393}}\), there were seven site mutations of M70I, V83L, D120V, K139E, E163G, D345H, and M356V. Among these mutation sites, three site mutations located in region 1.1 (K139E and E163G) and region 2.4 (M356V) were common in the two strains. The results suggested that the mutant hrdB genes from A56 and A393 might be responsible for the high avermectin production.

Increased Production of A56 in a Large-Scale Fermentor. To investigate the reproducibility of yield-improved mutants in large scale, A56 was cultivated in a 180-m\(^3\) fermentor, in which A56 exhibited an excellent productivity. As shown in Fig. 6, A56 had a higher avermectin-producing rate, and the yield of avermectin B1a in A56 has an increase of 53% more than that in the parent strain (3-115). This indicated that A56 maintained the industrial characteristic of the parent strain and could be well adapted to industrial-scale production.

Discussion
The complexity of cellular networks makes it difficult to build a complete understanding of multiple processes but also offers innumerable possibilities for modification to obtain yield-improved strains. Reverse engineering facilitates elucidation of overproduction mechanisms and the identification of genetic targets for modification. Because of the wide use of avermectins, many other efforts have been made to optimize production systems and maximize productivity, as it is not easy to improve the titer in an overproducing strain. In this study, we attempted to precisely engineer an industrial actinomycete strain and successfully obtained yield-improved mutants.

Fig. 3. Screening of avermectin B1a overproduction mutants transformed with optimal mutant hrdB gene. \( S. \text{avermitilis} \) 3-115 was used as the control (dashed line —100%), and the error bars represent standard deviations. All samples were cultured in Erlenmeyer flasks and measured in triplicate. C1: strain containing a blank pSET152; C2: strain containing an additional copy of wild-type hrdB gene.

Fig. 4. Comparative analysis of in vitro transcriptional assay of reconstituted holoenzyme with \( \sigma^{\text{hrdB}} \) protein or its mutant \( \sigma^{\text{hrdB}} \) protein on \( \text{aveR} \) promoter. In all lanes, the holoenzyme was reconstituted by 0.25 pmol \( E. \text{coli} \) RNA polymerase core enzyme (EPICENTRE) with the indicated amounts of \( \sigma^{\text{hrdB}} \). The ratio of sigma factor to the core RNA polymerase is indicated.

Fig. 5. Mutations for the best clone A56 and A393 isolated from the hrdB mutant library are shown mapped onto a schematic of critical functional components of \( \sigma^{\text{hrdB}} \). Conserved regions and their proposed functions are shown. The hrdB from parent strain 3-115 is also shown.
Our analysis of the differentially expressed level between the wild-type strain and the high producer identified 31 different genes that had diverse functions, including gene transcription and regulation, avermectin biosynthesis, sig-regulons, energy metabolism, and others. It is reasonable to assume that many of the observed alternations are somewhat related to avermectin high producing in strain 3-115. It has been reported that sigma factors play a central role in restructuring the transcription responses to environmental signals (25). Among these differentially expressed genes, although sig-regulons (such as sig37, sig3, and rsbW) have not been previously studied in detail, those genes may be involved in counteracting, e.g., the nutrient limitation, or may be potential mediators through which \( \sigma^{hrdB} \) exerts its control on downstream targets. Most of the up-regulated genes are involved in gene transcription and regulation, several of which are believed to be involved in carbohydrate metabolism (26), such as the LacI/GalR family transcriptional regulator (SAV1728), and in regulation of a wide range of cellular activities including antibiotic production and amino acid metabolism (27), such as TetR-family transcriptional regulator. Some other genes involved in energy metabolism, amino acid metabolism, phosphotransferase system, and ATP-binding cassette (ABC) transporter were also up-regulated in 3-115. The up-regulation of those genes is likely to be necessary to rapidly supply the cell with the essential elements (28, 29) (amino acids, carbohydrates, and nucleosides) for rapid growth and antibiotic biosynthesis. The high expression level of avermec- tin polyketide biosynthetic genes means more of the polyketide moiety produced in the overproducer. Enhanced expression of the pathway-specific regulator AveR in 3-115 indicated increased transcription of avermectin polyketide synthases genes, and therefore high avermectin production. This observation is consistent with its activity as a positive pathway-specific regulator (30).

The principal sigma factor, \( \sigma^{hrdB} \), is essential for morphogenesis and antibiotic production in \( S. \ coelicolor \) A3(2) (31). It is suspected that a cascade of other sigma factors, pleiotropic regulators and pathway-specific regulators, might be triggered to express by \( \sigma^{hrdB} \) (32). Engineering the binding factors of RNA polymerase has been successfully applied in the improvement of ethanol tolerance and productivity in \( S. \ cerevisiae \) (33), and in the yield improvement of products expressed in the heterogeneous host \( E. \ coli \) (34, 35). By binding to and affecting the promoter specificity of the RNA polymerase core enzyme, \( \sigma^{hrdB} \) directs the selective transcription of gene sets and acts cooperatively with one or more \( S. \ coelicolor \) extracytoplasmic function sigma factors to exert its control on downstream targets. So the expression level and the structural character of \( \sigma^{hrdB} \) can affect the expression of its functional protein. As none of the close homologues of \( \sigma^{hrdB} \) (\( \sigma^{hrdB} \), \( \sigma^{hrdC} \), and \( \sigma^{hrdD} \)) was required in vivo for differentiation and antibiotic production in \( S. \ coelicolor \), it is likely that the promoter of the pathway-specific regulator genes, actIÖRF4p and redDp, is recognized by RNP containing \( \sigma^{hrdB} \) (36). It has been reported that HrdB can recognize the proposed consensus sequences (TTGACN-16 to 18 bp-AN3T) of promoters (37). Based on the consensus promoter-binding sequence, aveRp was the candidate \( \sigma^{hrdB} \)-specific promoter. Since we detected a \( \sigma^{hrdB} \)-dose dependent level of in vitro transcription from aveRp with reconstituted Ez\( \sigma^{hrdB} \)-holoenzyme, we surmise that \( \sigma^{hrdB} \) perhaps recognizes aveRp in vivo; the potential 10 and 35 sequence of aveRp was similar to the proposed consensus sequence for the major holoenzyme. While the aveRp may be recognized in vivo by at least one other minor \( \sigma \) factor, it is possible that the ability of \( \sigma^{hrdB} \) to recognize aveRp in vitro is an artifact of the assay conditions used. Our data indicated that engineering of \( hrdB \) can affect the transcription activity of aveR, which induces positive regulation and high expression of the avermectin biosynthetic genes. Engineering of \( \sigma^{hrdB} \) may also make higher-level modifications, which tune the multiple, simultaneous alterations of diverse pathways and gene expression to obtain avermectin improvement.

While introduction of the wild-type \( hrdB \) ORF into the avermectin-high-producing strain 3-115 had no significant influence on avermectin production, the mutant \( hrdB \) ORF rescued from A56 and A393 introduced back into 3-115 endowed the host with a comparable level of production improvement as A56 and A393, suggesting that the mutant \( hrdB \) ORF was responsible for the yield improvement and played a dominant role when competing with the wild-type \( hrdB \) in vivo. It has been reported that the \( hrdB \) gene is essential for strain survival (31), so in our system, we enhanced the copy number of the \( hrdB \) gene in the parent strain 3-115 to examine its influence on avermectin production. The gene structure of \( hrdB \) was also optimized for avermectin production. The strain identified in this paper was cultured in a fermentor, and the yield indicated that the improved strain with mutant \( \sigma^{hrdB} \) could be successfully applied and well-adapted to industrial fermentation processes. This mutant A56 was genetically stable and has the potential to greatly improve the industrial productivity of avermectin B1a in industry.

\( \sigma^{hrdB} \) with mutations on the conserve region 1.2 was reported to result in antibiotic deficiency through reduced transcription of pathway-specific regulatory genes (38). In the high avermectin producing strain 3-115, we found three site mutations mainly located in the conserved region 1.1 and 1.2 in the predicted \( hrdB \) protein sequence. Region 1.1 can function as an autoinhibitory domain that masks DNA binding determinants in free \( \sigma^{hrdB} \). The high negative charge of region 1.1 allows it to act as a mimic of downstream DNA in its interactions with the active site channel, which might bind DNA recognition determinants in free \( \sigma \), thereby outcompeting promoter DNA (39). Depending on the promoter, the site mutation on the region 1.1 either promoted or inhibited transcription initiation by influencing open complex formation (38, 40, 41). Region 2.4 has been reported to be highly conserved and suggested to be involved in recognizing the promoters at 10 regions. We found two mutations occurred at the terminus of region 2.4 in \( hrdB^{A56} \) and \( hrdB^{A393} \). Amino acid substitution in this region might change the recognition of some promoters, therefore affecting the transcription of the functional proteins, which would be expected to cause avermectin overproduction in the mutant A56. As the components of global cellular transcription machinery, the site mutations in the conserved region of \( hrdB \) may directly affect the transcriptional level of aveR and thus the avermectin biosynthetic genes. The mechanism of the mutant \( hrdB \) effect on avermectin biosynthesis relative pathway genes needs further study by global transcriptional analysis.
In conclusion, this study represents preliminary attempts at reverse engineering to rapidly and effectively improve avermectin production in an industrial overproduc. The approach of reverse engineering facilitates whole-cell engineering through the selection and identification of key genes responsible for a variety of improved cellular phenotypes. This study has revealed that engineering on the target genes can effectively enhance the yield of desired natural products. The methods described in this paper have broad application to other actinomycetes in elucidating the correlation between the production and limitation steps in the biosynthetic pathway.

Materials and Methods

S. avermitilis ATCC31267 was used as the reference wild-type strain in microarray analysis. S. avermitilis 3-115 (CGMCC No. 3229), a high-level industrial producer of avermectin B1a (4,167 mg/L) obtained by traditional improvement programs with 35 × higher avermectin titers compared with ATCC31267, was collected in our laboratory. Strain 3-115 was used as the parent strains for genetic manipulation. Recombinant proteins for in vitro transcriptional assay were expressed and purified from E. coli. Detailed information on reagents, transcriptional analysis, expression and purification of recombinant proteins, data collection, high-throughput screening, and library construction is provided in SI Materials and Methods.

ACKNOWLEDGMENTS.

The authors are very grateful for the technical support provided by Huimin Yu from Tsinghua University. They also thank Elizabeth Ashforth, Huorang Tan, Zixin Deng, Linquan Bai, Zhiheng Liu, Jibin Sun, Keqian Yang, Xiaopeng Jia, and Yaqiao Li for critical reading and helpful discussions. This work was supported in part by grants from National Basic Project (2006AA02Z402, 2007AA02Z443), Key Project for International Cooperation (2007DFB31620), National Natural Science Foundation of China (30560001, 30600001, 30700015), National Key Technology R&D Program (2007BAI26B02), the CAS Pillar Program (KSCX2-YW-R-164), and Important National Science & Technology Specific Projects (2002ZX09401-05, 2009ZX09302-004). L.-X.Z. received funding from the Hundred Talents Program.