

Characterization of the *Ustilago maydis* *sid2* Gene, Encoding a Multidomain Peptide Synthetase in the Ferrichrome Biosynthetic Gene Cluster

WALTER M. YUAN,^{1†} GUILLAUME D. GENTIL,^{1‡} ALLEN D. BUDDE,^{2§} AND SALLY A. LEONG^{1,2*}

Plant Disease Resistance Research Unit, Agricultural Research Service, U.S. Department of Agriculture,² and Department of Plant Pathology,¹ University of Wisconsin, Madison, Wisconsin 53706

Received 20 December 2000/Accepted 10 April 2001

Ustilago maydis, the causal agent of corn smut disease, acquires and transports ferric iron by producing the extracellular, cyclic peptide, hydroxamate siderophores ferrichrome and ferrichrome A. Ferrichrome biosynthesis likely proceeds by hydroxylation and acetylation of L-ornithine, and later steps likely involve covalently bound thioester intermediates on a multimodular, nonribosomal peptide synthetase. *sid1* encodes L-ornithine *N*⁵-oxygenase, which catalyzes hydroxylation of L-ornithine, the first committed step of ferrichrome and ferrichrome A biosynthesis in *U. maydis*. In this report we characterize *sid2*, another biosynthetic gene in the pathway, by gene complementation, gene replacement, DNA sequence, and Northern hybridization analysis. Nucleotide sequencing has revealed that *sid2* is located 3.7 kb upstream of *sid1* and encodes an intronless polypeptide of 3,947 amino acids with three iterated modules of an approximate length of 1,000 amino acids each. Multiple motifs characteristic of the nonribosomal peptide synthetase protein family were identified in each module. A corresponding iron-regulated *sid2* transcript of 11 kb was detected by Northern hybridization analysis. By contrast, constitutive accumulation of this large transcript was observed in a mutant carrying a disruption of *urbs1*, a zinc finger, GATA family transcription factor previously shown to regulate siderophore biosynthesis in *Ustilago*. Multiple GATA motifs are present in the intergenic region between *sid1* and *sid2*, suggesting bidirectional transcription regulation by *urbs1* of this pathway. Indeed, mutation of two of these motifs, known to be important to regulation of *sid1*, altered the differential regulation of *sid2* by iron.

Iron is an essential trace element that plays a key role in many cellular functions, such as respiration, DNA synthesis and damage repair, and detoxification of free radicals (13, 14). Despite iron being one of the most abundant elements on earth, its availability is largely limited to living organisms due to its extreme insolubility at physiological pH (19). Many specialized systems have evolved in microorganisms to efficiently assimilate iron from the environment. One such mechanism is the production and secretion of a family of low-molecular-weight, ferric ion-chelating compounds, termed siderophores, in response to iron starvation (27, 29, 30). Various studies have demonstrated that the production of siderophores is negatively regulated by iron (9).

Ustilago maydis, the causal agent of common smut of corn, is a basidiomycete belonging to the order *Ustilaginales*. In response to iron starvation, *U. maydis* produces two cyclic, hexapeptide, hydroxamate siderophores, ferrichrome and ferrichrome A (Fig. 1) (10). Two ferrichrome biosynthesis-related genes, *sid1* and *urbs1*, encoding L-ornithine *N*⁵-oxygenase and a GATA family transcription factor, respectively, have been

previously identified and characterized. *sid1* is responsible for the first committed step, the hydroxylation of L-ornithine, in ferrichrome and ferrichrome A biosynthesis (28, 42). Northern hybridization analysis has shown that the accumulation of *sid1* transcript is negatively affected by iron concentration in the growth medium (28). Furthermore, two lines of evidence suggest that this regulation requires the presence of both a functional copy of the *urbs1* gene and *cis* elements upstream of the *sid1* promoter. Disruption of *urbs1*, a zinc finger GATA family transcription factor, leads to constitutive expression of *sid1* and unregulated production of the ferrichromes (2). A similar deregulated phenotype can be generated by mutating the two GATA motifs 2.5 kb upstream of the *sid1* translation start site, suggesting that *urbs1* may directly interact with the *sid1* promoter to repress *sid1* expression when sufficient iron is present in the growth medium (2). In vitro, Urbs1 protein has been shown to specifically interact with one of the GATA motifs (2).

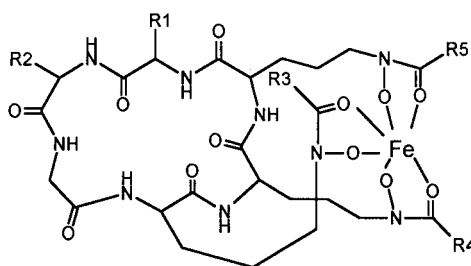
The hexapeptide ring structure of ferrichrome and ferrichrome A contains three δ -*N*-acyl-*N*-hydroxy-ornithine and three neutral amino acid residues consisting of glycines in ferrichrome and one glycine and two serines in ferrichrome A (10) (Fig. 1). Aside from the initial hydroxylation of L-ornithine, which is catalyzed by *sid1*, other steps of siderophore biosynthesis at the molecular level are poorly understood (31). Based on biochemical studies of several microorganisms, these steps have been assumed to proceed via thioester intermediates on a multimodular nonribosomal peptide synthetase complex. This assumption has been confirmed in *Aspergillus quadricintus* by purification of ferrichrome synthetase and its subsequent characterization as a peptide synthetase consisting

* Corresponding author. Mailing address: Department of Plant Pathology, University of Wisconsin, 1630 Linden Dr., Madison, WI 53706. Phone: (608) 262-5309. Fax: (608) 262-1541. E-mail: sal@plantpath.wisc.edu.

† Present address: Social Science Experimental Laboratory (SSEL), Division of Social Science, Pasadena, CA 91125.

‡ Present address: Department of Second Language Education, Faculty of Education, McGill University, Montreal, QC, Canada H3A1Y2.

§ Present address: USDA-ARS Malt and Barley Laboratory, Madison, WI 53706.



| | | |
|----------------------|--|--|
| Ferrichrome | R ₁ = R ₂ = H | R ₃ = R ₄ = R ₅ = CH ₃ |
| Ferrichrome A | R ₁ = R ₂ = CH ₂ OH | R ₃ = R ₄ = R ₅ = CH:C(CH ₃)CH ₂ CO ₂ H |

FIG. 1. Structure of ferrichrome and ferrichrome A. Both siderophores have a hexapeptide ring containing three residues of δ -*N*-acyl-*N*-hydroxy-ornithine and a tripeptide of neutral amino acids consisting of three glycine residues in ferrichrome or one glycine and two serine residues in ferrichrome A.

of repeated subunits (24). A phosphopantetheine cofactor was found to covalently link the substrate to each of the subunits, suggesting a thio-template mechanism for ferrichrome biosynthesis.

In this report, we describe the identification and characterization of the second biosynthetic gene, *sid2*, which encodes a nonribosomal peptide synthetase that is located in the ferrichrome biosynthetic gene cluster in *U. maydis*. The analysis of *sid2* regulation by iron indicates a common mechanism of regulation exists for the divergently transcribed *sid1* and *sid2* genes.

MATERIALS AND METHODS

Media and strains. Minimal, complete, and plate media used in these studies were adapted from Holliday (22) as previously described (41). Low-iron medium, E medium, and EM medium were previously described (42). *Escherichia coli* DH5alpha (Bethesda Research Laboratories, Bethesda, Md.) was used for all DNA manipulations. The *Salmonella enterica* serovar Typhimurium LT-2 mutant, *enb-7*, was a gift from J. B. Neilands (University of California, Berkeley). The *U. maydis* strains used in this study are listed in Table 1.

Genetic manipulations. Genetic crosses and construction of a diploid strain were done as described previously (26).

Assay for siderophores. Production of ferrichrome by *U. maydis* colonies was routinely assessed by the ability of colonies to cross feed the *Salmonella* serovar Typhimurium LT-2 mutant, as described by Wang (42). To confirm the production of ferrichrome, an XAD-4 resin-based extraction procedure was used. Briefly, a column was loaded with saturated XAD-4 resin (Sigma, St. Louis, Mo.), which was then washed with distilled H₂O, methanol, and distilled H₂O again. The iron complexes of siderophores in the culture supernatant were formed by the addition of 2% aqueous FeCl₃. The ferric complexes of siderophores were loaded onto the column and subsequently eluted with methanol (G. Winkelmann, personal communication). The volume of eluant was reduced by rotary evaporation and then analyzed by thin-layer chromatography with authentic standards (10).

DNA procedures. Plasmid DNA was purified from *E. coli* by the boiling miniprep protocol (18, 23) or by using a Qiagen midi-plasmid extraction kit (Qiagen, Valencia, Calif.). *U. maydis* genomic DNA was isolated by the glass bead method (26). Digestion of DNA and analysis of fragments were done by standard methods (6, 32). For subcloning, DNA fragments were isolated from SeaKem LE gels by running the DNA fragments through 3M filter paper, which was then centrifuged to extract DNA (6), or by using a QIAquick gel-extraction kit. Fragments to be labeled with ³²P were isolated with a Gene Clean kit (Bio 101, Inc., La Jolla, Calif.) or with a QIAquick gel-extraction kit. DNA fragments

TABLE 1. List of *U. maydis* strains

| Strain | Relevant characteristics | Source or reference |
|----------------------|--|---------------------|
| 518 | Wild type; <i>a2 b2</i> | R. Holliday |
| 521 | Wild type; <i>a1 b1</i> | R. Holliday |
| 288 | <i>a1 b1 pan1-1 inos-3 nar1-1 rec1-1</i> | R. Holliday |
| 227 | <i>a2 b2 ade1-1 met1-2 nar1-6 rec2-1</i> | R. Holliday |
| S1 | NTG mutant from 518; Fc ⁻ <i>a2 b2</i> | 42 |
| UMS049 | Derived from cross between S1 and 521; Fc ⁻ <i>a2 b2</i> | This work |
| UMS053 | Derived from cross between UMS049 and 288 | This work |
| UMS062 | Derived from cross between S1 and 521; Fc ⁻ <i>a2 b2</i> | This work |
| UMSD05 | Diploid UMS053 × 227 | This work |
| SH ⁺ 008 | Hyg ^r Fc ⁻ FcA ⁻ | 28 |
| UMS2 ⁺ H1 | Fc ⁻ <i>a2 b2 sid2</i> ::Hyg ^r (Hyg cassette in the 4.2-kb <i>KpnI</i> fragment) | This work |
| UMS2 ⁺ H2 | Fc ⁻ <i>a2 b2 sid2</i> ::Hyg ^r (Hyg cassette in the 4.2-kb <i>KpnI</i> fragment) | This work |
| UMS2 ⁺ H3 | Fc ⁻ <i>a2 b2 sid2</i> ::Hyg ^r (Hyg cassette in the 4.2-kb <i>KpnI</i> fragment) | This work |
| UMS2 ⁺ H5 | Fc ⁻ <i>a2 b2 sid2</i> ::Hyg ^r (Hyg cassette in the 4.8-kb <i>KpnI</i> fragment) | This work |
| UMS2 ⁺ H7 | Fc ⁻ <i>a2 b2 sid2</i> ::Hyg ^r (Hyg cassette in the 4.8-kb <i>KpnI</i> fragment) | This work |
| UMCO15 | <i>a2 b2 urbs1</i> ::phleo ^r | 2 |
| UMS204 | <i>a2 b2 sid1</i> ⁻ 204::GUS phleo ^r | 2 |

were labeled with a random oligonucleotide kit (Pharmacia, Piscataway, N.J.) or with a nick translation kit (Bethesda Research Laboratories). Transformation of *E. coli*, Southern hybridization, and colony blotting were done as described previously (28, 32). Vectors for subcloning were pHL1 (41), pJW42, pANUMV1 (1), pANUMV2 (1), pUC18, and pBluescript II KS(+) (Stratagene, La Jolla, Calif.). For complementation and gene replacement in *U. maydis*, self-replicative (pJW42, pANUMVII) and integrative (pHL1, pANUMVI) vectors were used. These vectors were also used to prepare genomic libraries of *U. maydis* wild-type strain 518. DNA transformation of *U. maydis* was carried out as described by Wang et al. (41) as modified by Voisard et al. (40).

DNA sequencing and sequence analysis. Nucleotide sequencing of double-stranded DNA fragments cloned in pANUMVI, pUC18, and pBluescript II KS(+) was done by the dideoxy method using the Sequenase kit (U.S. Biochemicals, Cleveland, Ohio) and by automated cycling sequencing using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit with AmpliTaq DNA polymerase FS (Perkin-Elmer, Foster City, Calif.). Both universal primers T7 and T3 and oligonucleotide primers designed from previously determined sequences and synthesized by the Applied Biosystems model 392 DNA/RNA synthesizer were used in sequencing. For cycle sequencing, we used the Perkin-Elmer model 480 DNA Thermal Cycler following the ABI PRISM Ready Reaction kit protocol, with the exception that the number of cycles was increased from 25 to 30. The initial sequence analysis was done using DNALysis (version 0.9) software. The sequence reads generated from automated sequencing reactions were also submitted for similarity searches to the BLAST (Best Local Alignment Search Tool) web server maintained at the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/BLAST>). A more detailed analysis was carried out using the GCG package (Genetics Computer Group, Madison, Wis.).

RNA isolation and Northern hybridization analysis. *U. maydis* total RNA was extracted by first grinding cell pellets in liquid nitrogen with a mortar and pestle. Total RNA was extracted from the cell powder by using the RNeasy Max kit (Qiagen). From 2 g (wet weight) of cells, 2 to 4 mg of total RNA can typically be obtained using this procedure. For Northern hybridization analysis, RNA samples were denatured with formaldehyde and fractionated by electrophoresis in a Reliant RNA gel system (FMC Bioproducts, Rockland, Maine). RNA was transferred to Hybond-XL nylon membrane (Amersham Pharmacia Biotech, Amersham, United Kingdom) and then hybridized and washed under stringent conditions (28, 32).

Primer extension mapping and cloning the 5' end of *sid2* mRNA. Primer extension with primers 1, 2, and 3 (see Fig. 4) was performed as described previously (28) on total RNA extracted from *U. maydis* cells grown in low (<5 μ M Fe detected) or high (~35 μ M Fe) iron media. The 5' end of the *sid2* gene was cloned using a 5' RACE (rapid amplification of cDNA ends) strategy (16, 17). The total RNA isolated from cells grown under the low iron condition was directly reverse transcribed using the *sid2* internal primer 4 (see Fig. 4). The 5' RACE system (GIBCO BRL/Life Technologies, Grand Island, Md.) was then used for both single-stranded cDNA isolation and tailing with poly(A) and a subsequent second-strand synthesis. The double-stranded cDNA was amplified by PCR using the adaptor from the deoxythymidine primer-adaptor and internal primer 5 (see Fig. 4). The PCR conditions were chosen as described in the GIBCO RACE system manual. The PCR products were purified using QIAquick PCR Purification kits (Qiagen) and subcloned in a pBluescript II KS(+) vector at the *SalI* and *SacI* sites.

RESULTS

Genetic analysis of the *sid2* mutant. Mutagenesis of *U. maydis* with nitrosoguanidine treatment previously led to the isolation of several classes of siderophore auxotrophic mutants (42). The ferrichrome⁻ (Fc⁻) phenotype was assigned to the mutant S1, a derivative of strain 518 (*a2b2*), because no ferrichrome was detected by high-pressure liquid chromatography in the culture supernatants when the mutant was grown in low iron medium for at least 4 days (42). To determine the genetic basis of the siderophore-nonproducing phenotype, S1 was independently crossed with the compatible wild-type strain 521 (*alb1*) and the auxotrophic mutant 288 (*alb1 pan1-1 inos1-3 nar1-1 rec 1-1*) and siderophore production of the basidiospore segregants was examined by assessing their ability to support the growth of the enterobactin-deficient *Salmonella* serovar

Typhimurium LT2 strain *enb-7* on E medium. The segregation ratio of 1:1 for the siderophore mutation versus the wild-type allele suggested that the Fc⁻ phenotype is controlled by a single gene lesion (data not shown).

Fc⁻ strain UMS049 was isolated after three backcrosses of strain S1 with strain 521. Fc⁻ strain UMS062 was isolated after two additional backcrosses with strain 521. Diploid cells (UMD05) were constructed by fusing an Fc⁻ progeny (UMS053) obtained from a cross between UMS049 and strain 288 with strain 227 (*a2b2 ade1-1 met1-2 nar1-6 rec2-1*). After purification on minimal medium, the resultant diploids were tested for siderophore production by using the *Salmonella* serovar Typhimurium LT2 bioassay and by high-pressure liquid chromatography. All diploids tested exhibited wild-type ferrichrome production, suggesting that the mutation in S1 and in its backcross derivatives is recessive and that the wild-type allele is dominant. The *sid2* genotype was assigned to this series of Fc⁻ mutants.

Genetic analysis of the linkage between *sid2* and *sid1* and location of *sid2*. We investigated possible linkage between *sid2* and *sid1* by crossing strain UMS062 with strain SH⁰⁰⁸, a mutant whose *sid1* gene has been disrupted by a gene cassette conferring resistance to hygromycin (28). Both strains are Fc⁻, and a ratio of 1:3 for segregation of the Fc⁻ phenotype versus the wild-type phenotype was expected if *sid1* and *sid2* were not genetically linked. Only 14 of 200 progenies tested were Fc⁺, and they were all Hyg^r, suggesting a tight linkage. We therefore investigated the possibility of complementing the *sid2* mutation with cosmid clones containing *sid1*. This strategy led to the successful cloning of the *a* mating-type locus of *U. maydis*, which is tightly linked to *pan-1* (5.3% recombination). These two loci were only a few kilobases apart on the same cosmid clone (15).

We first tested the cosmids from which we had originally isolated *sid1* (42) and found that these did not complement *sid2*. Overlapping cosmids were identified by using the ends of the pSid1 cosmid insert as probes to a genomic library generated in replicative vector pJW42 with DNA extracted from wild-type strain 518. Selected cosmids were used to transform *sid2*⁻ UMS062, and transformants were analyzed for siderophore production by using the *Salmonella* serovar Typhimurium LT2 bioassay and chemical analyses. Cosmid 8 was able to restore ferrichrome production in all transformants, suggesting that the mutated gene *sid2*⁻ had been complemented in *trans*. Cosmid 8 was partially mapped but the ends of the gene could not be defined. Another genomic DNA library of 518 was prepared in the improved replicative vector pANUMV2, which contains *NotI* sites flanking the cloning site (1). Additional cosmid clones were identified by colony hybridization and allowed us to place the gene within a 16-kb region (Fig. 2).

Subclones of cosmid 8 were generated in the integrative vector pANUMV1 (1) to define the minimal region required for complementation in *cis* of the mutation harbored in *sid2* by UMS062. DNA preparations of the subclones were digested with various restriction enzymes to release different-sized inserts from the cosmid vector. UMS062 was then cotransformed with the vector, which carries a hygromycin selectable marker and the released insert DNA. Transformants were screened for ferrichrome production.

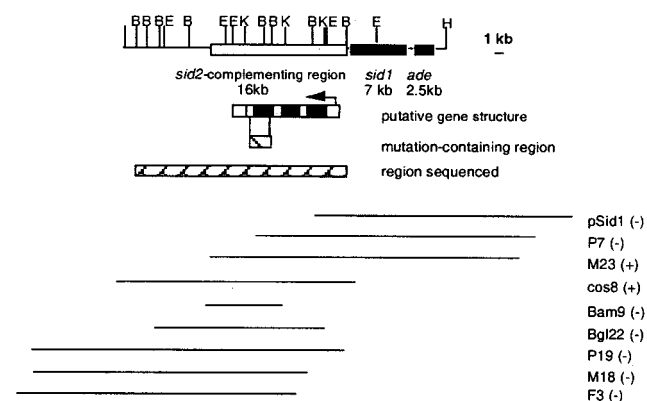


FIG. 2. The ferrichrome biosynthesis gene cluster of *U. maydis*. The cosmids used to localize the *sid2*-complementing region by transformation of the ferrichrome mutant are at the bottom. The hatched-box region contains the mutation. Sequence analysis demonstrates that *sid2* shares homology with members of the peptide synthetase protein family. The pSid1 DNA insert is in pJW42. All other inserts are in pANUMV2. *trans* complementation results are shown in parentheses: (+), complementation; (-), no complementation. Only *Bam*HI (B), *Eco*RV (E), and *Kpn*I (K) sites are indicated on the restriction map. One end of the map is defined by a *Hind*III site (H).

Molecular analyses of ferrichrome-positive transformants showed that a 23-kb *Bgl*II insert restored production of ferrichrome in mutant UMS062. Two deletion derivatives of the *Bgl*II subclone and an overlapping 9-kb *Bam*HI insert also restored the production of ferrichrome by homologous recombination. The minimum region capable of restoring ferrichrome production by homologous integration was localized to a 1.8-kb *Bam*HI-*Kpn*I fragment (Fig. 2).

Gene disruption. Two constructs were devised to demonstrate the isolation of *sid2* by gene disruption. The 4.5- and 4.8-kb *Kpn*I fragments of cosmid 8 were subcloned in pBS', a pBluescript II KS(+) derivative in which the *Bam*HI restriction site has been destroyed by Klenow fill-in. A *Bam*HI digest released the 0.7-kb *Bam*HI fragment contained in the 4.8-kb *Kpn*I fragment, and a *Xba*I-*Hind*III fragment of pHLI carrying the hygromycin B resistance gene was inserted in place of the missing *Bam*HI fragment, yielding a 6.4-kb *Kpn*I fragment

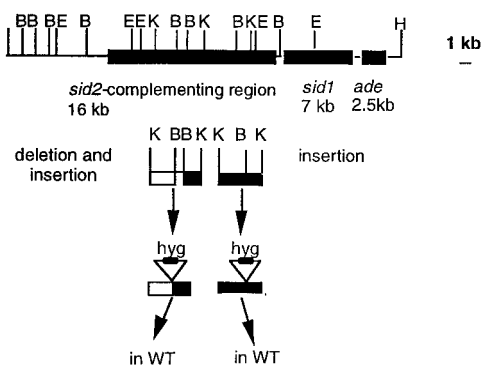


FIG. 3. Gene disruption of *sid2* in *U. maydis*. A hygromycin B resistance cassette was either inserted directly at the *Bam*HI site in the 4.2-kb *Kpn*I fragment (right-hand side) or was used to replace a 1-kb *Bam*HI in the 4.8-kb *Kpn*I fragment (left-hand side).

TABLE 2. Segregation of the hygromycin B resistance gene inserted into *sid2* and ferrichrome production

| Cross | Ratio of Fc ⁻ /Fc ⁺ segregants ^a | Ratio of Hyg ^r /Hyg ^s segregants |
|----------------------------|---|--|
| 518 × 521 | 0/100 | 0/100 |
| UMS2 ^r H2 × 521 | 46/54 ^b | 46/54 ^b |
| UMS2 ^r H5 × 521 | 41/59 ^b | 41/59 ^b |

^a Determined by bioassay with *Salmonella* serovar Typhimurium LT2 *enb-7*.

^b *P* > 0.05 for chi-square analysis of a 1:1 segregation model.

(Fig. 3) in the resulting subclone pKΔB::Hyg. Similarly, we inserted the hygromycin B resistance cassette into the unique *Bam*HI site of the 4.2-kb *Kpn*I fragment harbored in pBS', thus obtaining a 6.5-kb *Kpn*I fragment giving rise to the resulting subclone pK::Hyg. On the basis of the restriction map of the complementing DNA, this deletion and these insertions were expected to disrupt the putative *sid2* gene. Each of the two altered *Kpn*I fragments was transformed into wild-type strain 518.

Seven hygromycin B-resistant transformants were obtained with the 6.4-kb *Kpn*I fragment; three of them failed to feed the *Salmonella* serovar Typhimurium LT-2 mutant *enb-7*, whereas the other four displayed a wild-type phenotype for ferrichrome production. With the other construct, pK::Hyg, more than 200 hygromycin B-resistant transformants were obtained; 87 were tested for ferrichrome production using serovar Typhimurium LT-2 bioassay, and only 3 of 87 had a wild-type phenotype. Southern hybridization analysis of *Kpn*I-digested genomic DNA of a selected set of the potentially disrupted transformants with the 6.4-kb *Kpn*I fragment and the 6.5-kb *Kpn*I fragment as probes showed the correct exchange of the wild-type fragment with the mutated allele containing the hygromycin B resistance cassette (data not shown).

Genetic analysis of the disrupted mutants. Both *sid2* gene-disruption mutants were crossed with the compatible *U. maydis* prototrophic wild-type isolate 521 (UM002, *a1b1*). Basidiospore segregants were examined for ferrichrome production by assessing their ability to support the growth of serovar Typhimurium LT2 *enb-7* on E medium, and the segregation patterns (Table 2) showed a 1:1 ratio of Fc⁺ and Fc⁻ as would be expected for the segregation of alternative alleles of a single locus. Hygromycin B resistance cosegregated with the absence of ferrichrome biosynthesis. Southern hybridization analysis of genomic DNA from five segregants chosen from each cross confirmed that the mutant phenotype cosegregated with the expected restriction fragment polymorphisms (data not shown).

Mapping and cloning of the mRNA 5' ends. The transcription initiation sites of *sid2* were defined by primer extension with primers 1, 2, and 3 (Fig. 4). The *sid2* transcripts were shown to initiate at two different sites, -84 bp and -246 bp, upstream of the putative ATG start codon (Fig. 4). Interestingly, when total RNA extracted from iron-replete and starved cells was used, elevated levels of the extension product were observed from the iron-starved cells, suggesting that *sid2* transcription is iron regulated (data not shown). A 5' RACE strategy was used to confirm the results obtained from the primer extension experiments and to check for the presence of introns at the mRNA 5' ends. To clone the 5' end of *sid2* mRNA, the

-596 TCGATCGAGAAGAAAGAAAGAAAAAACAACAAAAAGGACGAGAAAGGGTCCGACAAACAGACTGCTGACCAACGTCACACGTCGCAA
 -504 CGTGTGTGTGCAACGTCGAAAGTCGAGACTACGAAATACGAAATCTGTCGAGGTGTGAGCGGCAACGAGAAAGATCTCAGTCCCTCCGAACTGAAATCTATCATCTCGTACCTCGACGG
 -384 GAAGAAACACGTTTGTGATCGTGTGAGACGAAAGGGAGCGCTGCTGCTCCGAGTCTGCTTGGTTCATAGGCAAGCGAACAGAAATCGTGAATACCTCCCTTTTGTGTGGGCAATCT
 -264 CTTTCCACTTTTCGCCCTCTACCTTGGCTTGGCTTTCCTCGACACCGTATCAACCGATCCCGTCTTGCAAGACATACCCATCTCGCTTTCGACTTCACAGCTCCATCTCTCGTGA
 -144 AAGTCCACCACAGTCCGAGGACTCACCTTGGCAACCTCGTACTCTGACCATCTCTAGCGCTGCTCGACCAACCCCGTCTCGTCAATTTAGACAAAGTACAGACTCTTTCTGCTTCATC
 ← primer 2
 1
 -24 TCGCAAGCCACGATCGTCCGCAACATGTCCTCCCAAGACGACATGCTCCCTTTCTGACGCTGACGGGTCTCCGAACTTCTGGACCCGACTTCTTCGCGTCCGCCGCAATGCAAG
 MetSerValSerLysAsnAlaHisAlaProPheProAspValThrGlyLeuArgThrSerGlyProAspPheSerArgValGlyProGlnCysLys
 ← primer 1
 +97 CTTCTGCGAGTTCAGCAGTCCCTCTTGTGCTCGTGCAGCGCTGAGCTTCTCCGCCAAGATGTCATCTCGCCCTACTCTCAGCATGGTGTGTCATCTCTTTCATTATGGACT
 LeuLeuArgValGlnGlnSerLeuHisLeuAspAlaArgAlaAlaSerSerPheSerGlyGlnAspLeuHisLeuAlaLeuLeuSerAlaTrpCysValIleLeuPheHisTyrGlySer
 ← primer 5
 +217 CACGAAGTACCAACATCGAGGTGCTTCCACATCACACACGCTCAGAGACCTTGTGAGCTCGACTTTGCGAAGACGCGGAGCTTGTCCATTACTCCCTCTCAACTCGCCGCA
 HisGluAlaThrAsnIleGluValLeuHisThrSerGlnHisGlyHisGluThrLeuValGluLeuAspPheAlaGluGlnArgGluLeuValProPheThrProSerGlnLeuAlaAla
 ← primer 4
 +337 TTTGTCCATCAAGCCAGGACGAAACAGCTTCTGCCAAGACCTCGATCATGGCTATGCTGCCTTTTATCAACCTAGCCACCAACGAGCCCTCACCTCACTTTCAACACTCCAACCTC
 PheValHisGlnAlaArgGlnGlnGlnLeuProAlaLysAspLeuAspHisGlyTyrAlaAlaPheTyrGlnProSerHisThrAsnGluProSerProHisPheGlnGlnLeuGlnLeu
 ← primer 3
 +457 TCGTCCAGGCCGACATCTCGCTCCGCTTCTCCTTACCTCCAAAGCTCGGCTCACTCGTTCCTGAAATGAAAGCAGCGCTTCTGTGACTCGCACAATCAGCTCAGCTGCAACTG
 ValLeuGluMetLysAlaAlaProSerValHisSerHisGlnSerAlaGlnLeuGlnLeuSerSerArgProAspIleSerLeuAlaLeuSerTyrHisLeuGlnGlySerAlaGlnLeu
 +577 CGTCAAGTGGCCGCTTCTCGAGTCTTACCACAGCGACCCCAACAGCAGCGCTTGTGCTGAGCGCTTTGACTGGAAGCTCAGAGCTCCGCAATCCCACTACCAGCACCTTCCC
 ValGluArgPheAspTrpLysLeuArgAlaSerAspAsnProAsnTyrGlnHisLeuProArgGlnValAlaAlaLeuLeuGluSerTyrHisSerAspThrGlnGlnHisAlaLeuSer
 +687 GACCCGAATCACATCCAAGTCTGATCCGATCGCTCGAGCCGAGTTCGAGTACTTTGCTGCTACTACTCCCGATCGCTGAGCCCTCGACTTTTCGCTTCGATTTCAAGACCTCAA
 AspProAsnHisIleGlnAspArgHisAlaAspArgLeuGluThrGluPheAlaAlaThrThrProAspAlaLeuAlaLeuAspPheAlaLeuAspLeuLysLeuLys
 +817 TCCACAAAATGGTCTATGTCGCAATGAATCAACGTCGCCAAAAGTCAAACACCTTCTTGGAGTCAAGGTCGCGCTCAGCTAGCTTACCACCAAGCCGACCACTTGTGGCTCTC
 SerThrLysTrpSerTyrAlaGluMetAsnGlnArgAlaGluLysValLysHisLeuLeuTrpSerHisGlyValGlySerAlaSerSerAspProGlnAlaAspHisIleValAlaLeu
 +937 TATCTCGAAAAAGTCCAGAGACTTACCTTCTTCCATTTGGTGTCTCAAAGCGGGCGCTGCATGGTCCCAATCGACACCGACTGGCTGCTTCCGCTAGACAGCGCTGCTCGCAAAG
 LeuHisLeuAlaIleGlnAlaMetArgAlaLeuArgIleTyrIleProTyrGlyLysAspLysIleGlyThrAspGlnIleArgTyrLeuAlaAspPheAlaLeuAspLeuLysLeuLys
 +1057 TCTAATGCAAAAATCGTATTGACCCAGCAGCACAAGATATCTGAGCAGCTTCGCTCAGGATCTCGAAAGCCAGCTGGTCAAGGCAAGGCGAAATCACAGCCATCCGCTTCGATCAGCTT
 SerAsnAlaLysIleValLeuThrHisAspAspLysIleSerGluGlnLeuArgHisAspLeuGluSerGlnLeuValLysAspLysGlyGluIleThrAlaIleArgLeuAspGlnLeu
 +1177 GAGCCGAGCTTTCACAAGTTCAGTCTGCTCCCGCAACGCCAATCGCTCGATACAACAGCTCGCTTACATGATCTGGACTTCAGGTACCACCGCTCCCAAAAGTGTGGCATC
 AspAlaGluLeuSerGlnValGlnValProProAlaAsnAlaAsnArgSerIleGlnGlnLeuAlaTyrMetIleTrpThrSerGlyThrThrGlyLeuProLysGlyValGlyIle
 +1297 CAGCATCTGCCATCTCAAAGCCATCGTGGCCCTCCGAATCTACATCCCGTACGGCAAGACAGATCGGCACCCAGCAGATTCTGTTATCTTCTAGTACTCGCCCTACAACCTTGGACCTC
 GlnHisLeuAlaIleIleGlnAlaMetArgAlaLeuArgIleTyrIleProTyrGlyLysAspLysIleGlyThrAspGlnIleArgTyrLeuAlaAspPheAlaLeuAspLeuLysLeuLys
 +1417 TCCATTATGGACTGCTTCTACACCTGGGGCTTGGCGCCACCATTTGCTGCTGCCCGCGCGCTTCTTCTCCAAGATCTCGTCAAGTGGGCAACAGCTTGCAGCCAACTCATACGCTC
 SerIleMetAspCysPheTyrThrTrpGlyLeuGlyGlyThrIleCysSerCysProArgGlyValLeuLeuGlnAspLeuValGluValGlyAsnSerLeuGlnProThrHisThrLeu
 +1537 CTGACACCAGCCGTCATGGCAATGACCAACGCCATCGTGTTCGAAGCTCAAGTGGTCTATCTCGGCGGTGAAAAGCTCAGTCAAGTCTGGCCGATGAATGGAGCAAGACTGTCTC
 LeuHisLeuAlaMetAlaMetThrGluArgHisValProSerLeuArgValProSerLeuHisValIleSerGlyGlyGluLysLeuSerGlnValIleAspGluLysLeuLysLeuLys
 +1657 TCTCTCAATCTCAGCCGCTGCCAGGCTACACTGATGCCATGAACCTGCTGTACCTTTGGCGATCGAGTCAAGGCAACCAACATTTGGTGTGCGCTCCCAACCGCTCTCTGTCTAC
 LeuLeuAsnLeuTyrGlyProAlaGluAlaThrLeuIleAlaMetAsnArgArgValProPheGlyAspArgValLysAlaProAsnIleGlyValAlaLeuProThrValSerCysHis
 +1777 GCCCTCGAAATACGATCAGTCTGATCAAGGTCGCTCCGAGCAACTGCTTGGAGTCTCAGCTTCTGCTGCTGCTTACGTCGCGGATCCGGTAAAGCGCCGCAAAATCTCTT
 AlaLeuAspLysTyrAspGlnIleValIleLysGlyAlaValGlyGluLeuValLeuGlyGlyProGlnLeuAlaArgGlyTyrValGlyAspProValLysThrAlaAspLysPhePhe
 +1897 CCTCATCCCGAGTTCGAGAGTCTACCCAGCCGCTGATCTCGTCTCACTGACAACCAAGAGTTGAATACCTTGGAGCATCGAGCATCAGGTCAAGATCAACGGCATCTCGTATC
 ProHisProGlnLeuGluTyrArgThrGlyAspLeuValArgGlnLeuAspAsnGlnGluPheGluTyrLeuGlyArgIleAspAspGlnValLysIleAsnGlyIleArgIle
 +2017 GAGTCTCGAGATCAATGCTGCCATCAAGACAGCCAGCAGAAGTCAAGGACTCCGAGACCATGGCTTCTCTAAGAAGGACAACGAAAGCGAACAGCAATATCAACTTCTCCGCT
 GluLeuLeuGluIleAsnAlaAlaIleLysAsnSerHisAspLysIleLysAspSerGluThrMetAlaPheSerLysLysAspAsnGluSerGluGlnGlnIleIleAsnPheSerAla
 +2137 CTCTCGCCGCGAGCCGGGACAGCTTCTCAGAACCATCAAGACCGGATGCTGTGGCTCGCGAGCTGACGGCCAATGCCAAGGACAGCTCCCTCCATGTTCCCAACCTCTTT
 LeuHisLeuProGlyGluProGlyGlnLeuLeuArgThrAspAlaIleAlaValAlaAlaArgGlnLeuGlnAlaAsnAlaLysAspSerLeuSerTyrMetLysLeuAlaAspPhe
 +2257 GTCACTTCTTCGCAATTCACAGCAGCATCCAGCGCTAAGATTGACGAGTCCGCGCTCAAATATGCTTGGCAGCTTCGATCAGCTCGACTGGGAAAACAACTTGCACACGAGGGCGAT
 ValIleLeuSerHisPheProArgThrSerSerAlaLysIleAspArgValAlaLeuLysAsnValLeuAlaSerPheAspGlnLeuAspTrpGluAsnLysLeuAlaAsnGluGlyAsp
 +2377 GATGACCAAGTGGACCGCCGCAACCGCACAAGCAGAAGCTTGCCTTCGAAAGTGGCTCGCCAGCTCTGCAACGTCGACGCCAGCAAGATCGGGCGCAAGACCCCTTTACAAGTGTGGC
 AspAspGlnValAspProAlaThrAlaGlnAlaGluAlaCysLeuArgLysTrpLeuAlaLysLeuCysAsnValAspAlaSerLysIleGlyArgLysThrProPheThrSerValGly
 +2497 CTCGACTCGATTCGTCATGATTTCTCGAAGCGCTCTCAGAAGAAGGATTCGCCGTGCTAGCTGAGCTGCGCGCTTCCCTACGCTCAAATCGCTCGGAGAACATCTCAAAGC
 LeuAspSerIleArgAlaMetMetPheSerLysArgValSerGluGluGlyPheAlaValIleSerGluLysAspValAlaArgPheProThrLeuLysSerAlaGlyIleHisLeuGlnSer
 +2617 TCTGGTCAAGCTCTGAGGAGCGTCCCAAGCGCCGCGAGCTTCTCGTGTACTTTGATGCCGCTTCCGCCAGTCTCTCATCTTGGTGAACAGCGCAAGGAGATGCTTCTGCG
 SerGlyAlaSerSerGluGluArgAlaLysArgAlaAlaSerPheLeuAlaAspPheAspAlaAlaPheArgProValValSerSerTrpValLysGlnArgLysAlaAspAlaSerAla
 +2737 ATCCAGTCCATCTTCCCTGATCCGCTTTCGAGGAAGGATATGCTGGCCGAGTCCGACGGGATAGCTCAAGCTACCGCATCCAGCGCAATATCGACTTGCCTGTGATTCGACCAAGCC
 IleGlnSerIleLeuProCysThrProLeuGlnGluGlyMetLeuAlaGluSerGlnGluArgAspSerSerTyrArgIleGlnArgLysTrpArgLeuAlaSerAspCysAlaAla
 +2857 CGCTTCTTAAAGCTTGTGACGAGACTTGTCTCACTTTGATAGTCTCCGACGCTCTTGGCGAGCTTGGCTCCCTGACGTTGGCTTCCACCAACGGAATGGCTTTCACGCTCAC
 ArgLeuSerLysAlaLeuIleGlnThrValAlaHisPheAspSerLeuArgThrSerPheAlaAspValGlySerLeuAspValGlyLeuHisGlnArgGluTrpProPheGlnProHis
 +2977 TTTTGGCAATCGTATGGAAGTCACTTCCCTGATCGAACAATGATGTCGACGACGCGATCGGATGCTGAAACAGCCATCTCTCTGACGCAAGACCAAGCTCGATTGGATCTC
 PheLeuGlnIleValTrpLysSerPheThrProSerIleGluGlnLeuAspValAspGlySerAspAlaGluGlnAlaIleLeuSerAlaAlaLysThrLysLeuAspLeuAspPro
 +3097 TTTGGAACAGTCCACCCGCTTCTTGTGTTGTCAGCAAGCACTTAGCAGGCTCTGGTGGTCTGTCATCATAGCAGTACGATGCCGATCGCTCGGCACTCTCGAGGATCAT
 PheGlyThrSerProProValAlaPheLeuPheValLysGlnAlaLeuSerGlnAlaLeuSerGluValValAlaHisHisSerThrTyrAspAlaArgSerLeuGlyIlePheGluAspHis
 +3217 GTCGAAGCTGTCTACAATCGCCAGAAACCCGACTTCACTCCAATTTTCGACCGCACTCGCCAGATTTCTCCATTTGACCAACCGGAGGCTCAGCGACATGCCGATGTTGGCAAAAG
 ValGluAlaValTyrAsnArgGlnLysProProThrSerLeuGlnPheSerThrAlaLeuAlaGlnIleLeuProIleAspGlnThrGluAlaGlnArgHisAlaAspValTrpGlnLys

FIG. 4. Nucleotide and deduced amino acid sequence of *sid2*, including 5' and 3' untranslated regions. The number, which refers to the nucleotide sequence, is based on the deduced ORF of *sid2* from +1 at the adenine of the translational initiation codon (ATG). The transcription start sites are marked in bold. Primers used for primer extension and 5' RACE are indicated as arrows with the respective primer number. Some putative transcription binding sites, including GATA boxes, are underlined.

+3337 GCGCTCTCAAACACCAGGAGCGGAATATGCCTCCCTCCGACTTTGTCTCTGACCAGGCTTCCGAGGCGACGACAGCCAGCCAGCCTGCACCAGTCGCGCTACCTGGAGCAAAAC
AlaLeuSerAsnTyrProArgGlyGluTyrAlaSerPheProThrLeuSerLeuThrArgProSerGluAlaAspAlaAspGlnAlaSerLeuHisGlnSerArgTyrLeuGluAlaAsn

+3457 ATCAACTGGGCGAAGTACGAGTCGGCTTGTGCGGAGCTCGGCGTCTGTCGACGCTCCCTTTGGTCCAGACGCTTGGGCTTAGTGCTCTCCGCTTCCAGAGTCGCAACACCTCATTCTG
IleAsnTrpAlaLysIleGluSerAlaCysArgGluLeuGlyValSerAlaArgLeuValGlnThrAlaTrpAlaLeuValLeuSerAlaPheThrGluSerGlnHisLeuIlePheLeu

+3577 GCGGACTCAGTCTCGGGTCTACGCTCTCAGCCGAACCTGACCTAGCATATGTCTCTGTCGACGCTGCTGACCTTTCATGCTCGCTCAGAGCAGAAAGCTCGCAATCTGATC
GlyAspSerValSerGlyArgThrLeuSerAlaGluLeuAspLeuAlaTyrGlyProValLeuSerThrValProValProPheMetLeuArgProGluGlnLysLeuGlyAsnLeuIle

+3697 AAACAGATGGATGATTTCCAAACCTCCATCATGGAAGCGCAGCATACAGACTTGGGTGCCATCCGTCGAATGCTCCAAGTCCACCACCGGGAATCACTTTTCCACTCCGCTTTGTCTC
LysGlnMetAspAspPheGlnThrSerIleMetGluAlaGlnHisThrAspLeuGlyAlaIleArgArgMetLeuGlnValProProArgGluSerLeuPheHisSerValPheValLeu

+3817 GAGCTTCCCTGAGCAGCCGAGGACATCGACTCGAGCCAGTTCGCTTGTCCAAGATGGCAGATCTCGGTGTGCCACCGAGCAGCTGCTCGGTGTCGAGGTGCTACCTGCCTCGGAT
GluProAlaProGluGlnProGlyIleAspSerSerGluProAlaPheArgLeuValSerLysMetAlaAspLeuGlyValAlaThrGluHisValLeuGlyValGluValLeuProLysSerAsp

+3937 GGATCTGTCAAGCTTGGCTTGAGCTGGCAGAAGAATCATCAGCGAAGGCTTCGGGGGCTCATACTCGAGCAATTTGATCGCAGCTCACCGCGCTATGCTTCTTGGATGCTGAC
GlySerValLysLeuGlyLeuSerTrpGlnLysAsnIleIleSerGluGlyPheGlyGlyLeuIleLeuGluGlnPheAspArgSerLeuThrAlaLeuCysSerSerLeuAspAlaAsp

+4057 GTTGGCTCTCTCTGACTGTACCTCCGATCCGACCAACAGGCTCATCTTTACTCGGTGACCAAGCGGCTTCCAAAGCAGAAGTGTTCATCGGCAACATTCAGGGTGTAGCCAGC
ThrGlySerLeuLysTyrCysHisProArgSerAspGlnGlnAlaSerSerPheThrValLysProValSerLysGlnLysCysSerSerAlaThrPheThrGlyValAlaSer

+4177 TCGCTGAACAAGCAAGCCATCTGTATAACAGCAACGCAAGCTGAGATCTATCAAGACATGGCAGACTCACCTCGTCCGAAAACCGCGGCAACAATGTCTCCAGAGCTCGAGCAG
SerLeuAsnLysGlnAlaIleSerAspAsnSerAsnAlaValGluIleTyrGlnAspMetAlaAspSerProSerSerArgLysProAlaAlaThrMetSerTyrProGluLeuGluGln

+4297 GCGTCTAACGGGCTGCCAACCTTTTTCGCCACCTCCACGCAACTCGGTCGTCGGTGTGTGCTCGAGCGTGGTCTAGAGAGCTACATTTGCCGCTGGTATCTTGAAGCTGGCCAT
AlaSerAsnGlyValAlaAsnLeuPheArgHisLeuProArgAsnSerValValGlyValCysLeuGluArgGlyLeuGluSerTyrIleValProLeuAlaIleLeuLysAlaGlyHis

+4417 GCTTACCTGCCCTCGATCGCAGCTTCCCTCGACCGAAGAAGAACTTGTCAAAGACAGCGGTGACGCTCTTATCGTCTGCTCAAGCAAGTTCACAGACTTTGACAGCTGACTGGA
AlaTyrLeuProLysGlnIleTrpGlySerGluLeuValLysLysLeuValLysLysAspSerGlyAlaAlaLeuIleValAlaSerSerLysPheThrAspPheAspSerLeuThrGly

+4537 GTGGAGTGTCCGACCGGATTCGAGGCAAGTCAAGGACCGCTTAAAGATGGAAGCGGACTGTTTCGGTAGAGTCTCGATCCGACGCTCGCCTTCATCATCTACACATCGGGCTCG
ValGluMetLeuGlyThrAspSerArgGlnPheLysAspAlaValLysAspGlyLysAlaThrValSerValGluSerArgSerAspAspValAlaPheIleIleTyrThrSerGlySer

+4657 ACGGAAAACCCAAAGGCTGCTGTGCTGACGCAAGCTTAAGTGGCTGCTGTGTCGAGGGCTTCTATTATAACTACGAAAAGAGCGCTGGCTCATTTGAGTCTCGCGCGCTTTCTG
ThrGlySerLeuLysLeuThrGlnAlaAsnLeuAlaAlaValGluGlyGlyTyrAsnTyrGlnLysGluAlaSerLysLeuThrValAlaSerLysLeuThrGly

+4777 GCTCGTCCGCGGAGGCTTCCAGCTTCACTGCTCGAACTTCCTCTCGTTCGGGGTGGGCGCCACCTCGTCACTGGTCCCGTCTGATTCATGACGATATTGCAAGACCATG
AlaArgSerAlaGluAlaPheAspValHisLeuLeuGluIlePheLeuSerLeuArgValGlyAlaThrIleValThrGlyProArgAlaLeuIleHisAspAspIleAlaLysThrMet

+4897 TCCACGCTGAGGTGACGCGCTTGCCTGTCCTCCCTTCTTTCAGCAAGGGCAAGCGGATTGACCTAGTGTGGTCCCTCGCTCGGAGTCTCATCATCGTGGCGAAGCTCTC
SerThrLeuGluValThrHisAlaCysValValProSerLeuPhePheSerLysGlyLysArgIleGluProSerValValProSerLeuArgValLeuIleIleGlyGlyGluAlaLeu

+5017 ACCAAGATCTTTGTGAGATTTGGGGGAGCGAGGCTCGGAACGACCCGCTGACTCAATGCCTATGGCCCAAGCGAGGCTACTATCGGTAACCTCGGTCGCTGTCAGCAAAAGTCT
ThrGlnAlaValLysGlnIleTrpGlySerGluLeuValLysGlySerGluArgProValLysGlySerGluLysGlnIleTyrGlyProSerGluAlaThrIleGlyAsnSerValAlaSerLysLysSer

+5137 CGACCAGCAACATTTGGTGCACCTTTCCCGGCACTCAATACCTCGTCTCAAAGACGCTCAATGGCAACTGGTCTTCTACGCTGCGAGGTGACCCAGGCGAGTGTACATCGGAGGAGAA
ArgProSerAsnIleGlyAlaProPheProGlyThrGlnTyrLeuValLeuLysAspValAsnGlyGlnLeuValProThrLeuArgGlyGluProGlyGluLeuTyrIleGlyGlyGlu

+5257 CAGTTCGCAAGGGCTACCTGCAAGCTTCCAGCTTCTGCTGCTCATACGTCACCGGTCAGCAAACTATCGCAGCTGGAGACATGGTTCGCTGCTCCAGCGCAGGCGCGAGTAT
GlnValAlaValLysLysLeuThrGlnAlaAsnArgProAspSerSerPheIleThrTyrGlnGlyGlnGlnIleTyrArgThrGlyAspMetIleAlaSerLysLeuThrValAlaSerAlaGlyTyr

+5377 CTTGGTGCATAGATGGCTCGAGGTCAAGGTGAGAGGCGCCGCTCGAGCTTCCGGAGGTGGATGCTGCCCTTCTCAGCTTCCCTCAACGAGAACCCTGGAAACCGTCCGCGCGGCTC
LeuGlyArgIleAspGlySerGlnValLysValArgGlyAlaArgLeuGluLeuAlaGluValAspAlaAlaLeuSerAlaSerLeuAsnGluAsnLeuGlyThrValGlyThrAlaVal

+5497 ACCATTCACCGCATCATCTTAAGTCAAGGAGCAGCGGCTTAGTGTCTTCTCGCTCAGGACTGTGTCCAGCAAGGCCAAGATTCGTGTGATCTGGCCCTCTGCTCTGACAA
ThrIleHisAlaAspHisProLysIleGluGlyAlaAlaArgLeuValSerPhePheAlaGlnAspCysValArgThrLysAlaGlnAspSerValAspProGlyAlaLeuLeuValGln

+5617 GCGCCGAAAGCGGTCAAGCAGTCTGCTGAGCTGAGGAGAAGCGTACGAGCTCGGCTGCCCTAGTACATGGTACCCTCTCTCGTCTCGCTCTCACCTATCTTCCATCTCACCGCTTTCT
AlaProGluAlaValLysLysAlaGluLeuArgArgSerValArgAlaAlaLeuArgLeuPheMetValProSerLeuValLeuAlaLeuAlaLeuThrLeuProLysLeuSer

+5737 GGCAAGCTGACCAAGACTGTGAAGGACTTTACCACCTCGATCGACCCAGTAAAGCTTCGACGCTTTCGGACAAGAAGCAGTCTCGTCAAAGAGAGCTCACTGACGAGCAGCAGACA
GlyLysAlaAspArgArgLeuLeuLysGluLeuTyrHisSerIleAspProSerLysLeuSerThrLeuSerAspLysAsnGluSerArgGlnArgGluLeuThrAspSerGluGlnThr

+5857 GTAGCCGAGCTGCTACGCTCCTCTGCTGCTTCCGATGTTCTATCATGATGATCTTGTATCTCATATGGCTGGCTTGGACAGCTTGCACCGTCTGCTCAGCTCGCAACAAGCTT
ValAlaGluLeuValLysSerValArgLeuSerSerAspValHisLeuSerLysLeuMetAlaGlyLeuAspLeuIleMetAlaGlyLeuAspLeuIleMetAlaSerLysLeuThrValAlaSerAlaGlyLeu

+5977 CGGAAGCAGGCTACGACGCAAGGCTCTGCTCATATGAAACCCACCTCGAAGCTATTGCGGCGCTCGCATCGCAAGCTTCTCCAAAGAACTAGCGATGTGGAATGGAAAG
ArgLysHisGlyTyrAspAlaThrValSerSerIleMetAsnGluProThrIleGluAlaIleAlaGlyArgArgIleAspLysLeuSerSerAsnGluThrSerAspValGluTrpLys

+6097 CAGACCGTCTCAACTCAGCACAAGTCCGATCTCTCCGCAAGTCCGCGGAGCTCAAATCGAGACTGTTGCCGTCGCTGCGATCCAAAGTTCGCTTCCAGCGGCTCTCC
GlnThrValSerGlnLeuThrAspLysValArgSerLeuProGlnTyrArgGlyThrGlnIleGluThrAlaLeuProCysValProIleGlnValAlaLeuValSerGlnAlaValSer

+6217 GAGCATCGCAGCAGCTCGTTACATTAATACCATCAGCATCGACTGTGAGCAACGAATCTCTGCCATCGGATTCGCAACGCTGGATGACTGCGCTGTCTAGGCACGAGATCTAC
AspAspArgSerThrProArgTyrIleThrThrIleSerIleAspLeuSerSerAsnGluPheSerAlaAspArgIleArgAsnAlaTrpMetThrAlaLeuSerArgHisGluIleTyr

+6337 CGCACCGTCTTTCGCGAGGTCGATCGCACGCTAGTCAGGTAGTGTCTCAGCGCGAAGCTTGAACCTCAAAGTGGTCCAGCCCTCTGAACCCATCTCCCTCGCCGATAGCTTGGCAGAC
ArgThrValPheAlaGluValAspArgThrLeuValGlnValValLeuSerAlaGluSerLeuThrSerAsnTrpSerAlaThrSerGluProIleProSerProAspSerLeuAlaAsp

+6457 TATCATGCAAGCAGCGAAGGACATCGTCGCAATATCAGCAGTCTCCCTGCTCTGCGCCTCAAGCTGTGGCAGGGCAGAAATGGCGCTCCGACGCTCACTTTGACATGCTCTCACGCG
TyrHisAlaSerThrAlaLysAspIleValAlaAsnIleSerSerValProAlaLeuArgLeuGlnIleTyrGlnGlyGluAsnGlyAlaProThrLeuThrLeuThrCysSerHisAla

+6577 ATCTACGCGGTGACTCGATCCGATGCTTCTCAAAGAGGCTCGGATTGCTTGTGACCAAGAGCAAGTCTGTCGCGGCTTGCCTTCCAAAGAAGCAGCAGTTCGATCTCGGTGAC
IleTyrAspGlyAspSerIleArgMetLeuLeuLysGluAlaSerAspCysLeuValThrLysSerLysValValProAlaLeuProPheGlnGluAlaAlaArgCysIleValGlyAsp

+6697 GCTGAAGCAGGAAAGCAAGCAATCTCGACCAAGCAGCTTCTGCTGCTTTTGCACAAGCAGCTGCGCAACCTAACTGGCGTTCGACCCGAGCACAAGCTCAGCCCGCGGAGGAGCTC
AlaGluAspGluAlaLysGlnPheTrpThrThrLeuAlaAspPheLeuProThrThrValProAsnLeuThrGlyValArgProGluHisAsnValSerArgGlyGluGluLeu

+6817 ACCATCGTTCACATCTTTTCACTCAGTGTGAGAAAGCGGCGGCTGCTTAAGTAAAGTAAAGTCCAGTCCAGTCCATCTTGGTAGCAGCTTCCGCAATCTCTGGTCTTTACGCTGGAGAG
ThrIleAlaSerHisLeuSerPheThrGlnLeuGluLysAlaAlaArgAlaAlaLysAlaValSerThrIleGlnSerIleLeuValAlaAlaPheAlaHisLeuAlaLeuThrValAlaGlyGlu

+6937 AGTGACGTCAGCTCGGCTTGTCTTCTGCTGCAAGCATTCGCTCGAGGTGTGAGTCTGATCCAGGACCTTGTGTCAACAAGTCCGCTGCGCTGACCGATGCTAGGTCAAAT
SerAspValThrLeuGlyLeuValLeuSerGlyArgSerIleProValAspGlyValGluSerIleHisGlyProCysValThrThrValProLeuArgLeuThrAspAlaArgSerAsn

+7057 GCTTCTCCGATTTTGAACAAGGCTCACAGGCGGTAATGCCATTTTGGCCCAACCAAGCTTCTCTGCGCAGTAAATGCGCTGGCTCGATTTGTCAAAGCGGCTTTTGAAGCC
AlaSerSerAspLeuLysLysArgAlaHisGlnAlaValAsnAlaIleLeuProHisGlnHisValSerLeuProGlnLeuMetArgTrpLeuAspLeuSerLysAlaProPheGluAla

+7177 CTCTTTCTGACCTCGCCAGCTGAGCGCTTCTGAGAAGCCCTACTTCTCAGAGAGGCTTCTCAGATGGAGAGATTATCTTTAGCGCTTGAAGTCAAGTCTGTTGGGGATGCC
LeuPheSerTyrLeuGlyGlnSerGluArgSerSerGluLysProTyrPheSerGluArgAlaSerGlnMetGluArgAspTyrSerLeuAlaLeuGluValSerAlaValGlyAspAla

FIG. 4—Continued.


```

+11257 GCGCGTGTCTCGCAGGTGGTGTCTCAAGGCAGACTCGGTCTCGGCCGGTTCGGGTTTTGCCGAGGTGCAGTGCAGGAGATGCAGAGCTGACCCCAAGGACGTCGTCGGTCTCGCGCA
GlyGlyAlaSerGlnValValLeuLysAlaAspSerValSerAlaGlySerGlyPheAlaGluValGlnValGluArgAspAlaGluLeuThrAlaLysAspValValGlyArgArgAla
+11377 GCCCAAGTTGGCCAGACCTCAAGCCGCCCTGTTGAGCTGACACTGGTGGGTGAGATCGCTGATTTCTCGCTGCACCCAGCTCCTATACGATGCCTTTTCGATCGAGTTTTGGCTCGC
AlaGlnGlyTrpProAspLeuThrSerProProValGluLeuThrLeuValGlyGluIleValValPheSerLeuHisHisValLeuTyrAspAlaPheSerIleGluPheLeuAlaArg
+11497 GATTTCCGGCTCGTGTACAACCGGGCAGTGTGGTCACTGCAATCAGTGGCCAGAGTGGTGCAGCACATTTGTCGAAGAGCAACAGCGAACGAGAGGGATGCGCAAGAGTATTGGTGC
AspPheGlySerLeuTyrAsnAlaGlyGluLeuValSerSerAsnGlnTrpProGluValValGluHisIleValGluGluGlnGlnArgThrArgGlyAspAlaGlnGluTyrTrpCys
+11617 AGAGCACTTGCACCCGGGAGCAGCGCTTGTTCGCCGATCGTCCAGGGGACACAGGAGAAGCGTGGCATGTGCAGCACAAACGCTATCTCCGCTCTCTTCGGCGGTGGACTGCGGATACG
ArgAlaLeuAlaProGlySerSerGlyLeuLeuAlaAspArgProGlyGluHisArgArgSerValAlaCysAlaAlaGlnArgTyrLeuAlaLeuPheGlyGlyArgAlaAspThr
+11737 CAAAGCAGGCTTGACGCTGGCTGGCCTCTGCTAGCGGCATGGTCGACTCTCTTTCAGAGCGAATGCACAGCGCTGCCCGGTGTTCCGGCTTGTACCAGCTCGGTAGATCGTCTCGT
GlnSerArgLeuAspAlaGlyTrpArgProAlaSerGlyMetValAspSerLeuValArgAlaAsnAlaArgArgValAlaArgValArgLeuValProAlaArgEnd
+11857 CGAGTCGATTGACAAGTACATGGACCGCTTACTCAACTGTCTACCCATACAGTTCGAGGTTGGTTCACGTGCTCGACAAGGCTCGAGCTGCAGTGCAGCGACTTGGGCTGCCCAAGTT
+11977 CGAACAGACCGACTTGCAAGATGCTCATCGATGGGCTGGGCTCAGCCAGCACCAAGCATGCTACAACACGTTTGTCAACATCTTGTTCGGTGTACAGTTGACCAACATCTCGAGATGCA
+12097 CCAGCTCGACCTCGTCAACCCGCTCGATTACTCGCACCACCTCGCAACTCCACTCATGATCGTACACCTCCAAGCACGCCACGTTGCTCTGCCCTGGCAACCAGACGTCACACTTGA
+12217 TGTCGTCCTCAAGAATCATGCTGTGATATC

```

FIG. 4—Continued.

cDNAs obtained from reverse transcription with primer 4 (Fig. 4) were amplified by PCR and subcloned into a pBluescript vector. Positive clones were identified by colony hybridization with a *sid2* DNA probe, and 12 of those clones were chosen for sequencing. Sequencing results showed that no introns were present at the 5' end of the *sid2* mRNA, as the DNA sequence of the cDNAs was identical to the corresponding regions of the genomic DNA fragment. Consistent with the primer extension data, there were two transcript sizes which differed by 162 bp in length at the 5' end in the cDNA population, suggesting that *sid2* transcription initiation sites were heterogeneous.

Nucleotide sequence and predicted protein structure of *sid2*. The DNA sequence of both strands of the 13.8-kb genomic DNA region corresponding to the *sid2* gene and the deduced amino acid sequence are shown in Fig. 4 (GenBank accession no. U062738). DNA sequence inspection revealed a single open reading frame (ORF) with a length of 11,841 bp which we have designated *sid2*. The putative translation initiation site (CAACATGTCCG) matches the consensus sequence [CCA(C/A)(C/A)ATGGC] for a fungal translation initiation site. The two most conserved nucleotides, an A at position -3 and a C at position +5 (underlined in the sequences above), relative to the ATG start codon (in bold) are present in the *sid2* sequence.

The CodonUsage program (CodonUse 3.5, Peter A. Stockwell, University of Otago, Box 56, Dunedin, New Zealand) was used to determine the coding preference of the deduced *sid2* ORF. Previously sequenced genes from *U. maydis* were used to compile a codon frequency table as a parameter to Codon Usage (Z. An and S. A. Leong, unpublished result). In this analysis, it was shown that the ORF exhibited *U. maydis* codon bias throughout the 11,841 bp (data not shown) and no evidence was found to support the presence of introns in the *sid2* gene. This coding potential was further confirmed with a second program, TestCode, which is part of the GCG (Genetics Computer Group, Madison, Wis.) sequence analysis package (data not shown). The putative *sid2* polypeptide contains 3,947 amino acid residues, with a predicted molecular mass of 432,650 Da and an isoelectric point of 6.42.

A database search using the deduced amino acid sequence of *sid2* led to the identification of multiple motifs which are characteristic of the nonribosomal peptide synthetase family. The number of appearances and the distribution of the motifs in the putative protein sequence led to the prediction that *sid2*

contains three similar modules of 1,000 amino acids each. To confirm this prediction, the first 1,500 amino acids of *sid2* were used to compare with the following residues by using the Compare program (GCG). The results showed that ~1,000 amino acids in the query sequence shared similarity with two distinct segments of approximately equivalent length in the remainder of the protein. The three segments, termed modules 1, 2, and 3, were graphically represented using the DotPlot program from GCG (data not shown). Pair-wise alignments between the three modules were carried out using the GCG programs Best-Fit and Gap. The three domains, known as adenylation, acyl carrier protein (ACP), and condensation domains, within each module were delineated (Fig. 5A). The core motifs in the modules were further demonstrated by their alignment using PileUp (GCG) and comparison to the compiled peptide synthetase consensus sequence (Fig. 5B). Interestingly, an additional ACP motif was identified at the end of the last module (Fig. 5A and B). This motif has a sequence (QLGLDS*ISAIR) that is almost identical to the default motif (RLGIDS*ISAIR). Blast searches of each ACP domain identified by pfam (Protein Families Database) (www.ebi.ac.uk) revealed similarity to both ACP domains found in polyketide synthetases and the peptidyl carrier protein domains found in nonribosomal peptide synthetases. In all cases, the majority of matches found among the top 10 were with the peptidyl carrier protein domain.

Transcriptional regulation of the *sid1* and *sid2* ferrichrome biosynthetic gene cluster in *U. maydis*. The two ferrichrome biosynthetic genes, *sid1* and *sid2*, are divergently transcribed from a 3.7-kb intergenic region (Fig. 6). Like *sid1*, the 1.6 kb of the 5' nontranscribed region of *sid2* does not contain an apparent TATA box, but an array of putative transcription factor binding sites is present (Find Pattern Analysis) (data not shown) (Fig. 4). Multiple GATA motifs are present in the 3.7-kb intergenic region, including a 12-bp palindromic motif (Fig. 6) which is present four times. The two palindromic GATA motifs in the middle have been previously shown to be essential for *urbs1*-mediated, iron-responsive regulation of the *sid1* gene, as either deletion or replacement of these motifs led to a constitutive expression of *sid1* (3). A purine track between -506 and -476 bp (Fig. 6) may contribute to the initiation of transcription, as is the case for the other genes in filamentous fungi (21).

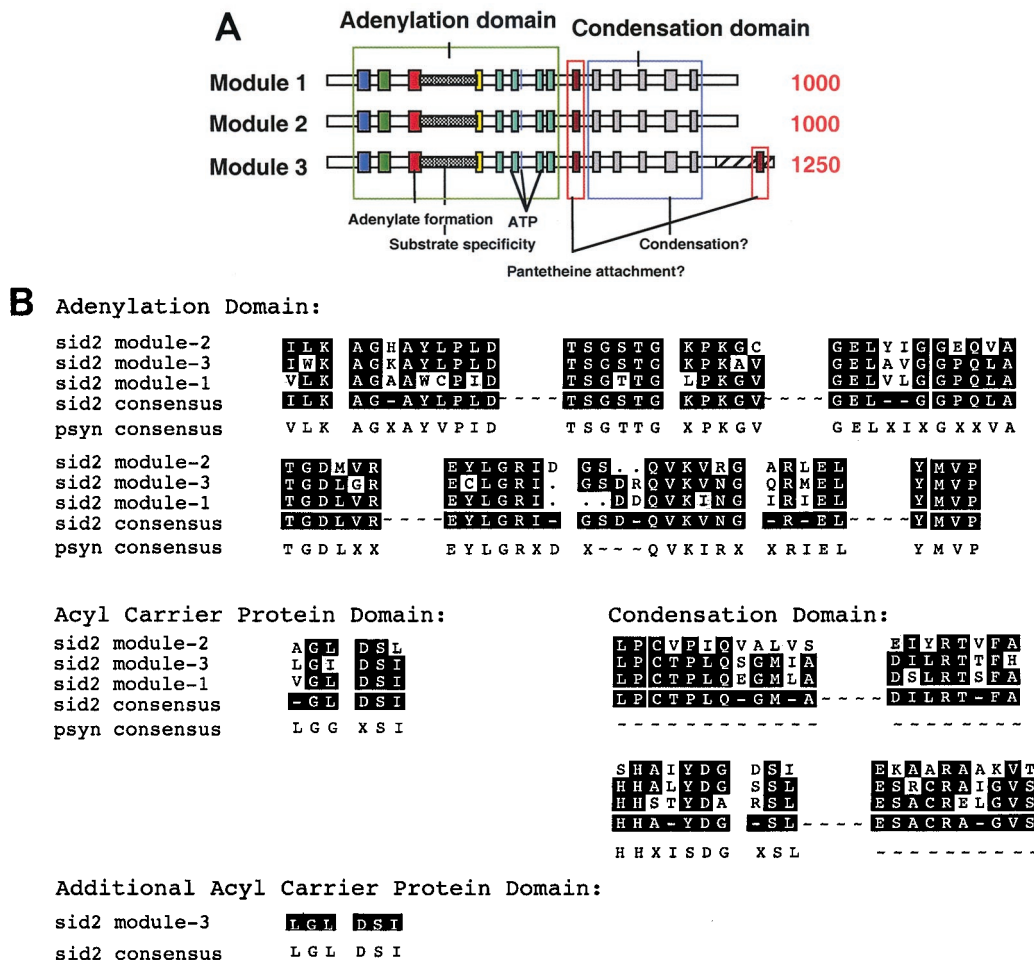


FIG. 5. Multimodular structure of *sid2* and multiple sequence alignment of key motifs in the modules. (A) A schematic representation of the highly conserved and ordered modular organization of the *sid2* gene. (B) Sequence alignment of the modules. Conserved motifs are shaded and grouped according to the domain they reside in. The sequences were aligned using the PileUp program in the GCG package, with a gap penalty of 8 and a gap extension penalty of 2. The consensus sequence of the alignment was further calculated using Pretty in GCG.

Regulation of *sid2* by iron. To determine whether iron affects the accumulation of *sid2* mRNA, Northern hybridization analysis of total RNA extracted from wild-type and three mutant strains grown in low-iron medium with or without 40 μ M FeSO₄ was performed (Fig. 7). A higher level (2.5-fold) of the 11-kb transcript was observed in wild-type cells grown in low-iron medium than in cells grown in the presence of iron. By contrast, approximately the same levels of the 2.3-kb transcript were observed in CO15 (2) (*urbs1*) and S204 (5' 3' distal GATA mutant AY4) (2) cells grown in either low- or high-iron conditions. A truncated transcript was detected in total RNA isolated from the *sid2* mutant, and the accumulation pattern of this truncated transcript was similar to that of the wild type (data not shown).

DISCUSSION

Following the initial hydroxylation of ornithine, ferrichrome biosynthesis has been postulated to proceed via covalently bound thioester intermediates on a multifunctional enzyme complex (33). This mechanism has been supported by biochemical data where hydroxyornithine-ATP-PP_i exchange activity was observed in extracts prepared from ornithine hydrox-

amate siderophore-producing fungi (4, 5). The genetic and molecular data presented herein indicate that *sid2* is essential for ferrichrome biosynthesis. *sid2* encodes a putative multifunctional nonribosomal peptide synthetase of a predicted molecular mass of 432,650 Da. As expected, the transcription accumulation of *sid2* in the cell is regulated by iron.

Based on several lines of evidence, *sid2* encodes a single genomic 11.8-kb ORF with no apparent introns. First, codon preference analysis showed a high codon potential throughout the entire 11.8-kb ORF. Second, the sequence of the cDNA which covers 800 bp of the 5' end of the *sid2* gene was identical to that of the corresponding genomic DNA. Third, no apparent intron consensus sequences were identified in the *sid2* sequence. Only one intron-containing nonribosomal peptide synthetase has been identified among the known peptide synthetases (7, 25). This finding is consistent with the hypothesis that many of the eukaryotic peptide synthetases are of prokaryotic origin.

Sequence analysis of *sid2* revealed the presence of three iterated modules of approximately 1,000 amino acid residues each. The three modules of deduced *sid2* protein sequence

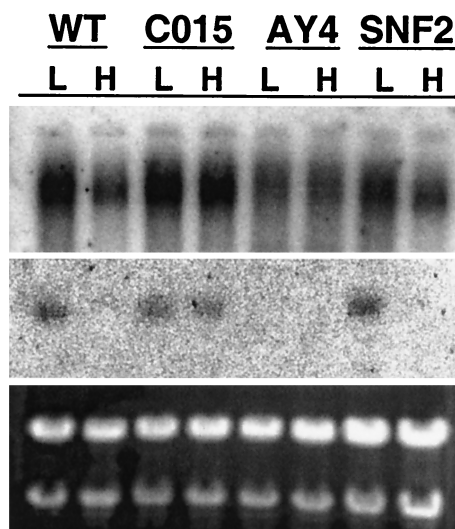


FIG. 7. Northern hybridization analysis of *sid2* total RNA. Total RNA (0.5 μ g/well) was denatured in formamide and electrophoresed on a 1.2% SeaKem gold agarose gel in MOPS (morpholinepropane-sulfonic acid) buffer with no denaturants. Ribosomal RNAs stained with ethidium bromide are shown in the lower panel to show equivalence of RNA loading. The RNA was transferred to nylon and hybridized sequentially with the *sid1* (middle panel) or *sid2* (upper panel) genomic DNA probes. Shown are RNA samples from L (low iron grown) and H (high iron grown): WT, wild type; C015, *urbs1* disruption mutant; AY4, *sid1* deletion mutant UMS204; SNF2, *umsnf2* disruption mutant of *U. maydis*.

communication). The repeated use of a single domain to charge different amino acids in order to complete the entire synthesis of the ferrichrome peptide remains an open possibility. It is noteworthy that a second potential peptidyl carrier protein domain is present in module 3 (Fig. 5). Future protein structure prediction, sequence alignment, and biochemical analysis in combination with site-directed mutagenesis will be required to provide a more definitive test of this repetitive module model.

Besides the additional putative acyl carrier protein domain motif, the third module of *sid2* is 250 amino acids larger than the other two (Fig. 5A). Since Orn is not only activated but also hydroxylated and acetylated, the additional sequence of the third domain could be involved in side chain modification such as the acetylation step. Interestingly, Hummel and Diekmann (24) first showed that a partially purified protein fraction also catalyzed transacetylation of hydroxyornithine from acetyl coenzyme A, hence implying that the peptide synthesizing activity may be part of a multienzyme complex. Later Siegmund et al. (33) managed to separate the transacetylase activity from the amino-acid-activating and the ferrichrome synthetase activities. Depending on the state of purification, the ferrichrome synthetase can thus be dissociated into functional subunits with partial activities. Efforts are under way to purify Sid2 and determine the biochemical activity of each module.

In many microbial systems genes responsible for secondary metabolites are often clustered (8, 25, 35, 36, 38). Here we have shown this also to be the case for ferrichrome biosynthesis in *U. maydis*. *sid1* and *sid2* are divergently transcribed from a 3.7-kb intergenic region which likely contains all the informa-

tion necessary for the *cis* regulation of both genes. Previous studies led to the identification and characterization of a zinc finger GATA family transcription factor, *urbs1*, which is responsible for iron-mediated regulation of *sid1* expression (2, 3, 40).

Northern hybridization analysis was performed to test whether *urbs1* also regulates *sid2* expression in a similar manner. High levels of a large transcript (>9.5 kb) were detected in wild-type cells grown in an iron-deficient medium, suggesting *sid2* is also transcriptionally regulated by iron (Fig. 7). To test the role of *urbs1* in this regulation, two mutant strains were included in this analysis, UMCO15 and UMS204 (Table 1). UMCO15 contains a nonfunctional copy of *urbs1*, whereas UMSO24 has two mutated *cis* elements involved in iron-mediated regulation of *sid1*. In contrast to the wild type, constitutive accumulation of high levels of the >9.5-kb transcript was detected, suggesting that *urbs1* may simultaneously regulate both *sid* genes in response to the iron concentration in the environment and the cell. A truncated transcript of about 9.0 kb was observed in the total RNA from the other mutant, UMS2H5, which has a hygromycin cassette inserted at the last domain of *sid2*. Future work involving site-directed mutagenesis will likely reveal other potential *cis* regulatory elements. In particular, the four highly conserved palindromic GATA motifs bear further examination. Preliminary efforts to determine the role of chromatin in the regulation of this gene cluster indicate that iron does modulate the pattern of nuclease hypersensitivity in the 3.7-kb intergenic region and that these changes also depend on Urbs1 (M. W. Yuan and S. A. Leong, unpublished results).

To delineate the borders of this ferrichrome biosynthetic gene cluster, the genomic DNA sequence immediately flanking the *sid1* and *sid2* genes has been determined. A yeast *ADE6* homolog was identified downstream of *sid1*, and disruption of this gene resulted in an adenine requirement on minimal medium (M. Warriner, A. D. Budde, and S. A. Leong, unpublished results). A yeast *SNF2* transcription factor homolog was identified immediately downstream of *sid2*, the function of which still needs to be determined. Disruption of this gene led to no obvious phenotypic effect on *sid* gene expression (M. W. Yuan, B. Bride, and S. A. Leong, unpublished results).

While our knowledge of the molecular mechanisms of iron uptake and storage in prokaryotes has greatly advanced in recent years, a similar understanding of these mechanisms in fungal systems is lacking. This is due to the greater difficulty of working with fungi and the lack of efficient cloning systems for many of these organisms. Using *U. maydis* as a model system, rapid progress has been made to genetically identify and characterize these mechanisms at the molecular level and hence iron-mediated transcriptional regulation is becoming clearer. Future studies will center on identifying other structural genes in the ferrichrome biosynthetic pathway, determining the substrate specificity of modules in *sid2*, and lastly, developing a better understanding of how the *sid* genes are regulated in response to iron conditions.

ACKNOWLEDGMENTS

We are very grateful to ZhiQiang An and Hans von Doehren for their helpful discussions and critical review of the manuscript and

Gunther Winkelmann for providing the figure of the structure of the ferrichromes.

This work was supported by the USDA-ARS and NIH grant GM33716 to S.A.L.

M. W. Yuan and Guillaume D. Gentil contributed equally to the work described in this paper.

REFERENCES

- An, Z., M. L. Farman, S. Taura, and S. A. Leong. 1996. New cosmid vectors for library construction, systematic chromosome walking and rapid restriction mapping in filamentous fungi. *Gene* **176**:93–96.
- An, Z., B. Mei, W. Yuan, and S. A. Leong. 1997. The distal GATA sequences of the *sid1* promoter of *Ustilago maydis* mediate repression of siderophore production and interact directly with *Urbs1*, a GATA family transcription factor. *EMBO J.* **16**:1742–1750.
- An, Z., Q. Zhao, J. McEvoy, W. Yuan, J. Markley, and S. A. Leong. 1997. The C-terminal finger domain of *Urbs1* is required for iron-mediated regulation of siderophore biosynthesis in *Ustilago maydis*. *Proc. Natl. Acad. Sci. USA* **94**:5882–5887.
- Anke, H., T. Anke, and H. Diekmann. 1973. Biosynthesis of sideramines in fungi. Fusigen synthetase from extracts of *Fusarium cubense*. *FEBS Lett.* **36**:323–325.
- Anke, T., and H. Diekmann. 1972. Biosynthesis of sideramines in fungi. Rhodotorulic acid synthetase from extracts of *Rhodotorula glutinis*. *FEBS Lett.* **27**:259–262.
- Ausubel, F. M., et al. (ed.). 1994. Current protocols in molecular biology. John Wiley, New York, N.Y.
- Bailey, A. M., M. J. Kershaw, B. A. Hunt, I. C. Patterson, A. K. Charnley, S. E. Reynolds, and J. M. Clarkson. 1996. Cloning and sequence analysis of an intron-containing domain from a peptide synthetase-encoding gene of the entomopathogenic fungus *Metarhizium anisopliae*. *Gene* **173**:195–197.
- Brakhage, A. A. 1997. Molecular regulation of penicillin biosynthesis in *Aspergillus (Emericella) nidulans*. *FEMS Microbiol. Lett.* **148**:1–10.
- Braun, V., and H. Killmann. 1999. Bacterial solutions to the iron-supply problem. *Trends Biochem. Sci.* **24**:104–109.
- Budde, A. D., and S. A. Leong. 1989. Characterization of siderophores from *Ustilago maydis*. *Mycopathologia* **108**:125–133.
- Challis, G. L., J. Ravel, and C. A. Townsend. 2000. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* **7**:211–224.
- Conti, E., T. Stachelhaus, M. A. Marahiel, and P. Brick. 1997. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **16**:4174–4183.
- Crichton, R. R., and R. J. Ward. 1992. Iron metabolism—new perspectives in view. *Biochemistry* **31**:11255–11264.
- Crichton, R. R., and R. J. Ward. 1998. Iron homeostasis. *Metal Ions Biol. Syst.* **35**:633–665.
- Froeliger, E. H., and S. A. Leong. 1991. The *a* mating-type alleles of *Ustilago maydis* are idiomorphs. *Gene* **100**:113–122.
- Frohman, M. A., T. R. Downs, P. Chomczynski, and L. A. Frohman. 1989. Cloning and characterization of mouse growth hormone-releasing hormone (GRH) complementary DNA: increased GRH messenger RNA levels in the growth hormone-deficient lit/lit mouse. *Mol. Endocrinol.* **3**:1529–1536.
- Frohman, M. A. 1994. On beyond classic RACE (rapid amplification of cDNA ends). *PCR Methods Appl.* **4**:S40–S58.
- Gomez-Marquez, J., M. Freire, and F. Segade. 1987. A simple procedure for large-scale purification of plasmid DNA. *Gene* **54**:255–259.
- Griffiths, E. 1987. Iron in biological systems, p. 1–25. *In* J. J. Bullen (ed.), *Iron and infection*. John Wiley & Sons, New York, N.Y.
- Guenzi, E., G. Galli, I. Grgunna, D. C. Gross, and G. Grandi. 1998. Characterization of the syringomycin synthetase gene cluster. A link between prokaryotic and eukaryotic peptide synthetases. *J. Biol. Chem.* **273**:32857–32863.
- Hamer, J. E., and W. E. Timberlake. 1987. Functional organization of the *Aspergillus nidulans* trpC promoter. *Mol. Cell. Biol.* **7**:2352–2359.
- Holliday, R. 1974. *Ustilago maydis*, p. 273–287. *In* R. C. King (ed.), *Handbook of genetics*. Plenum Press, New York, N.Y.
- Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**:193–197.
- Hummel, W., and H. Diekmann. 1981. Preliminary characterization of ferrichrome synthetase from *Aspergillus quadricinctus*. *Biochim. Biophys. Acta* **657**:313–320.
- Kleinkauf, H., and H. Von Doehren. 1996. A nonribosomal system of peptide biosynthesis. *Eur. J. Biochem.* **236**:335–351.
- Kronstad, J. W., and S. A. Leong. 1989. Isolation of two alleles of the *b* locus of *Ustilago maydis*. *Proc. Natl. Acad. Sci. USA* **86**:978–982.
- Leong, S. A., and J. B. Neilands. 1982. Siderophore production by phytopathogenic microbial species. *Arch. Biochem. Biophys.* **218**:351–359.
- Mei, B., A. D. Budde, and S. A. Leong. 1993. *sid1*, a gene initiating siderophore biosynthesis in *Ustilago maydis*: molecular characterization, regulation by iron, and role in phytopathogenicity. *Proc. Natl. Acad. Sci. USA* **90**:903–907.
- Neilands, J. B. 1993. Siderophores. *Arch. Biochem. Biophys.* **302**:1–3.
- Neilands, J. B. 1995. Siderophores: structure and function of microbial iron transport compounds. *J. Biol. Chem.* **270**:26723–26726.
- Ong, D. E., and T. F. Emery. 1972. Ferrichrome biosynthesis: enzyme catalyzed formation of the hydroxamic acid group. *Arch. Biochem. Biophys.* **148**:77–83.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Siegmund, K. D., H. J. Plattner, and H. Diekmann. 1991. Purification of ferrichrome synthetase from *Aspergillus quadricinctus* and characterization as a phosphopantetheine containing multienzyme complex. *Biochim. Biophys. Acta* **1076**:123–129.
- Smith, D. J., A. J. Earl, and G. Turner. 1990. The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *Penicillium chrysogenum* is a 421,073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases. *EMBO J.* **9**:2743–2750.
- Smith, D. J., M. K. Burnham, J. H. Bull, J. E. Hodgson, J. M. Ward, P. Browne, J. Brown, B. Barton, A. J. Earl, and G. Turner. 1990. Beta-lactam antibiotic biosynthetic genes have been conserved in clusters in prokaryotes and eukaryotes. *EMBO J.* **9**:741–747.
- Smith, D. J., M. K. Burnham, J. Edwards, A. J. Earl, and G. Turner. 1990. Cloning and heterologous expression of the penicillin biosynthetic gene cluster from *Penicillium chrysogenum*. *Biotechnology (New York)* **8**:39–41.
- Stachelhaus, T., H. D. Mootz, and M. A. Marahiel. 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**:493–505.
- Turner, G. 1992. Genes for the biosynthesis of beta-lactam compounds in microorganisms. *Ciba Found. Symp.* **171**:113–124.
- van Liempt, H., H. von Doehren, and H. Kleinkauf. 1989. delta-(L-alpha-aminoadipyl)-L-cysteinyl-D-valine synthetase from *Aspergillus nidulans*. The first enzyme in penicillin biosynthesis is a multifunctional peptide synthetase. *J. Biol. Chem.* **264**:3680–3684.
- Voisard, C., J. Wang, J. McEvoy, P. Xu, and S. A. Leong. 1993. *urbs1*, a gene regulating siderophore biosynthesis in *Ustilago maydis*, encodes a protein similar to the erythroid transcription factor GATA-1. *Mol. Cell. Biol.* **13**:7091–7100.
- Wang, J., D. W. Holden, and S. A. Leong. 1988. Gene transfer system for the phytopathogenic fungus *Ustilago maydis*. *Proc. Natl. Acad. Sci. USA* **85**:865–869.
- Wang, J., A. D. Budde, and S. A. Leong. 1989. Analysis of ferrichrome biosynthesis in the phytopathogenic fungus *Ustilago maydis*: cloning of an ornithine-N⁵-oxygenase gene. *J. Bacteriol.* **171**:2811–2818.