Isolation and Characterization of the Versicolorin B Synthase Gene from Aspergillus parasiticus

EXPANSION OF THE AFLATOXIN B₁ BIOSYNTHETIC GENE CLUSTER*

Jeffrey C. Silva, Robert E. Minto, Clifford E. Barry III, Koren A. Holland, and Craig A. Townsend†

From the Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218

Versicolorin B synthase catalyzes the side chain cyclization of racemic versicalnoral hemiacetal (7) to the bisfuran ring system of (−)-versicolorin B (8), an essential transformation in the aflatoxin biosynthetic pathway of Aspergillus parasiticus. The dihydrobisfuran is key to the mutagenic nature of aflatoxin B₁ (1). The protein, which shows 58% similarity and 38% identity with glucose oxidase from Aspergillus niger, possesses an amino-terminal sequence homologous to the ADP-binding region of other flavoenzymes. However, this enzyme does not require flavin or nicotinamide cofactors for its cyclase activity. The 643-amino acid native enzyme contains three potential sites for N-linked glycosylation, Asn-Xaa-Thr or Asn-Xaa-Ser. The cDNA and genomic clones of versicolorin B synthase were isolated by screening the respective libraries with random-primed DNA probes generated from an exact copy of an internal vbs sequence. This probe was created through polymerase chain reaction by using nondegenerate polymerase chain reaction primers derived from the amino acid sequences of peptide fragments of the enzyme. The 1985-base genomic vbs DNA sequence is interrupted by one intron of 53 nucleotides. Southern blotting, nucleotide sequencing, and detailed restriction mapping of the vbs-containing genomic clones revealed the presence of vmsA, a methyltransferase active in the biosynthesis, 3.3 kilobases upstream of vbs and oriented in the opposite direction from vbs. The presence of vmsA in close proximity to vbs supports the theory that the genes encoding the aflatoxin biosynthetic enzymes in A. parasiticus are clustered.

Aflatoxin B₁ (see Scheme I, 1), the principal member of the aflatoxin family, is one of the most potent mycotoxins known to man. The imperfect fungi Aspergillus parasiticus, Aspergillus flavus, and Aspergillus nomius produce aflatoxins, and these fungi are known to infect corn, grains, and nuts during their growth and during storage leading to the introduction of aflatoxin into primary foodstuffs (2, 3). The natural product AFB₁,¹ itself does not pose a major health threat; however, renal and hepatic oxidative detoxification of AFB₁-contaminated foods by P450 enzymes yields aflatoxin-15,16-exo-epoxide (see Scheme I, 2), a highly toxic mutagen (4, 5). It has been shown that the epoxide targets guanine residues and selectively alkylates the N-7 position of this purine in double-stranded DNA (6, 7). Depurination of the alkylated base has been correlated to bladder cancer in laboratory mice (8–10), teratogenic effects in chicken embryos (11), and liver cancer in humans (12–14). A direct connection between DNA damage and the incidence of human cancer has been established to originate at mutational hot spots of the p53 gene, a tumor suppressor gene whose altered sequence has been associated with approximately 50% of all human cancers (15, 16). Aflatoxin B₁ has been found to be responsible in particular for G → T transversions at codon 249 of the p53 tumor suppressor gene in hepatocarcinogenesis (17, 18) (Scheme I).

The aflatoxin biosynthetic pathway is notably long and complex (Scheme II). Although the formation of polyketide natural products is initiated normally by acetate, a specialized fatty acid synthase apparently acts in the case of aflatoxin to generate a six-carbon hexanoyl starter unit. This primer is homologated by successive malonyl condensations to give, after intramolecular aldol condensation, cyclization, and oxidation, norsolorinic acid (3) (19–21). Simple redox changes in the hexanoyl side chain yield the internal ketal averufin (4) (22–24). Oxidation at C-2' of 4 induces migration of the anthraquinone nucleus from C-1' to C-2' to afford hydroxyversicolorolone (5) containing the first furan ring (25, 26). Preparatory to formation of the second furan ring, oxygen is inserted into the carbon chain of 5 by a proposed Baeyer Villiger-like reaction to give versicalnol acetate (6) (27, 28). Support for this mechanism has come from a fermentation conducted in an ¹⁸O₂-containing atmosphere in which the isotopic label (*) was specifically incorporated at the ester oxygen (Scheme II) as shown in 6 (28). A cell-free system of A. parasiticus has been described in which an esterase catalyzed the hydrolysis of this terminal acetate to give versicalnol (7), which was cyclized to (−)-versicolorin B (8) (29). Tracing the fate of ¹³C label (*) from 6, it was shown that heavy isotope was retained in 7 without loss in the critical cyclization to versicolorin B (8) (27, 28). In hemiacetals 5, 6, 7.

¹ This work was supported by National Institutes of Health Grant ES 01670. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U51327 and U51328 for the gDNA and cDNA sequences, respectively.

† To whom correspondence should be addressed. Tel.: 410-516-7444; Fax: 410-516-8420.

¹ The abbreviations used are: AFB₁, aflatoxin B₁; MeCN, acetonitrile; VBS, versicolorin B synthase; PCR, polymerase chain reaction; kb, kilobase(s); bp, base pair(s); HPLC, high pressure liquid chromatography; gDNA, genomic DNA.
and 7, the chiral C-2′ center is benzylic and adjacent to a masked aldehyde. This is an intrinsically labile stereocenter, and each of these three compounds is isolated as a racemate (25, 30, 31) (Scheme II).

The cyclization of versicolorin hemiacetal (7) can be carried out nonenzymatically in the presence of acid to yield versicolorin B (8) as its racemate, which is designated historically as versicolorin C (32, 33). At neutral pH this chemical process is slow and cyclization is catalyzed in vivo by versicolorin B synthase (VBS) to give optically active (+)-versicolorin B (8) (29–33).2 The absolute configuration installed in this cyclase-catalyzed step is preserved in the bisfuran throughout the remainder of the biosynthetic pathway to AFB$_1$. The stereochemical match of this structure when metabolically activated as the eno-epoxide 2 and intercalated into right-handed helical DNA is essential to successful covalent adduct formation (35). These are key events in the tumorgenesis of this natural product. Preliminary purifications of VBS have been reported (36–38), but an improved protocol yielding homogeneous protein has been achieved.2 Detailed kinetic analyses of the reaction catalyzed by this enzyme reveal that from the stereochemically equilibrating mixture of enantiomers of 7, the 2′S-configured hemiacetal is specifically cyclized by VBS to (+)-versicolorin B (8).2

Formation of the dihydrobisfuran is completed in the oxidative desaturation of versicolorin B (8) to versicolorin A (9) (29, 39). The subsequent steps of the biosynthetic pathway are significantly less well understood. Cleavage of the anthraquinone nucleus and cyclization, decarboxylation, and dehydration afford the xanthone 10 (40). Successive O-methylations are known to occur at C-5 and C-7 to give O-methylsterigmatocystin (11) (41–44). This intermediate is further cleared oxidatively, demethylated, cyclized, and decarboxylated to ultimately afford aflatoxin B$_1$ (1) (45–47).

Although the mechanisms of these deep-seated molecular rearrangements in the post-versicolorin A segment of the pathway are not known, important progress has been made recently to identify the first genes in A. parasiticus that encode proteins involved in the biosynthesis of aflatoxin (48–51). Preliminary evidence has been gathered to suggest that these genes are substantially clustered (48–51), contrary to earlier reports (52–55). A probable polyketide synthase (pksA) and two fatty acid synthase (fas-1A and fas-2A) genes have been identified by sequence homology and gene disruption experiments (51). The localization of these genes, a ketoreductase (nor-1) acting immediately after the formation of norsolorinic acid (3) and ver-1, whose gene product participates in the oxidative cleavage of versicolorin A (9), has been determined by gene disruption and complementation (48, 49, 56). Combined with the cloning of one of the purified O-methyltransferases (omtA), the direct linkage of these genes has been determined to be within 45 kb of one another (see Fig. 5) (51). In this paper we describe the isolation of the gene encoding versicolorin B synthase (vbs) from both cDNA and gDNA libraries derived from A. parasiticus. Comparison of the sequences reveals the presence of a single intron in the latter. Translation of the mature mRNA gives a protein of 70,226 Da, in modest agreement with the 78-kDa apparent molecular mass of wild-type VBS as judged by its relative electrophoretic mobility.2 Alignment of the translated amino acid sequence of VBS with protein sequences compiled in protein data bases revealed a marked homology to several flavin-dependent oxidases and dehydrogenases. This relationship was unexpected because VBS does not catalyze a redox reaction. Finally, mapping of vbs gDNA clones has allowed the locus of this gene to be established about 3.3 kb upstream of omtA and separated from it by an apparent cytochrome P450 monooxygenase approximately 1400 bp in length of unknown function. These findings expand the experimentally determined dimensions of the apparent aflatoxin gene cluster and unambiguously define the function and location of the gene encoding versicolorin B synthase.

**EXPERIMENTAL PROCEDURES**

Materials—Restriction endonucleases, calf alkaline phosphatase, T4 DNA ligase, and T4 polynucleotide kinase were purchased from New England Biolabs (Beverly, MA). Lys-C endoproteinase (sequencing grade) was purchased from Boehringer Mannheim. Modified T7 DNA polymerase (Sequenase-2.0®) was purchased from U.S. Biochemical Corp. [α-32P]dATP, [γ-32P]ATP, and [α-32P]dATP were obtained from Amer sham Corp.. The following were purchased from Life Technologies, Inc.: ultrapure urea, acrylamide, and N,N′-methylenebisacrylamide. Escherichia coli XL1 Blue cells, Lambda ZapII, Lambda FixII, helper phages VCSM13 and R408, Pfu DNA polymerase, and nitrocellulose membranes were obtained from Stratagene (La Jolla, CA). Cu2+ cheesecloth was obtained from the Kendall Co. (Wellesley Hills, MA). Maltose monohydrate, MgSO$_4$, and MgCl$_2$ were purchased from Aldrich. DNase I, RNase A, and hen egg lysozyme chloride were ob-

---

2 S. M. McGuire, J. C. Silva, E. G. Casillas, and C. A. Townsend, manuscript submitted.

3 R. E. Minto and C. A. Townsend, unpublished results.
Cloning of vbs from Aspergillus parasiticus

Peptide Fragment 1

- a) Actual VBS DNA seq.
- b) Degenerate VBS DNA seq.
- c) Primer 7NC
- d) Primer 10C
- e) Primer 13C

Fig. 1. Peptides isolated from Lys-C digestions of native vbs where R = A/C, Y = C/T, N = A/C/G/T, M = A/C, S = G/T, and H = A/C/T.

Peptide Fragment 2

- a) Actual VBS DNA seq.
- b) Degenerate VBS DNA seq.
- c) Primer 5NC
- d) Primer 5NC
- e) Primer 14C
- f) Primer 15C

and packaged using Gigapack® II Gold packaging extract (Stratagene). The packaged phage were then propagated in the restrictive F2 host E. coli (P2PLK-17) to an original titer of 5.4 × 10^8 pfu/ml containing 95% recombinant phage. The primary gDNA library was then amplified in E. coli cells (LE392) to 1 × 10^9 pfu/ml.

A. parasiticus SU-1 48h mRNA was used to prepare a cDNA library using the Uni-ZAP XR vector and packaged using the Gigapack® Gold II packaging extract (Stratagene). The packed phage were then propagated in E. coli cells (PLK-F) to a final titer of 4.4 × 10^9 pfu/ml containing 90% recombinant phage. The secondary cDNA library was subsequently amplified in E. coli cells (PLK-F') to 1 × 10^9 pfu/ml.

Hybridization Experiments with Degenerate Probes—The A. parasiticus cDNA and gDNA libraries were screened by plaque hybridization with seven radiolabeled degenerate probes (7NC, 8NC, 9NC, 10C, 13C, 14C, and 15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences. The radiolabeled probes were generated by end-labeling using T4 polynucleotide kinase (New England Biolabs) with ^14C, and ^15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.

The radiolabeled probes were generated by end-labeling using T4 polynucleotide kinase (New England Biolabs) with ^14C, and ^15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.

Hybridization Experiments with Degenerate Probes—The A. parasiticus cDNA and gDNA libraries were screened by plaque hybridization with seven radiolabeled degenerate probes (7NC, 8NC, 9NC, 10C, 13C, 14C, and 15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.

The radiolabeled probes were generated by end-labeling using T4 polynucleotide kinase (New England Biolabs) with ^14C, and ^15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.

The radiolabeled probes were generated by end-labeling using T4 polynucleotide kinase (New England Biolabs) with ^14C, and ^15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.

Hybridization Experiments with Degenerate Probes—The A. parasiticus cDNA and gDNA libraries were screened by plaque hybridization with seven radiolabeled degenerate probes (7NC, 8NC, 9NC, 10C, 13C, 14C, and 15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.

The radiolabeled probes were generated by end-labeling using T4 polynucleotide kinase (New England Biolabs) with ^14C, and ^15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.

The radiolabeled probes were generated by end-labeling using T4 polynucleotide kinase (New England Biolabs) with ^14C, and ^15C; Fig. 1). Degenerate probes for vbs were designed from sequenced fragments of Lys-C-endopeptidase-digested VBS, taking into account Aspergillus nidulans and Aspergillus niger codon preferences.
were carried out using purified recombinant λ DNA from the genomic library, temperature cycling program A (cycles 2–4), and normal PCR conditions as described by Ausubel et al. (65). The restriction product (615 bp) was examined on a 1.5% agarose gel. The same protocol was used to generate the cDNA vbs fragment; however, the cDNA library was used as the source of template DNA.

Screening of the cDNA Library Using PCR-generated vbs Probes—The A. parasiticus SU-1 cDNA library was screened by plaque hybridization with radiolabeled vbs. Single and double restriction digestions containing 0.5 μg of each enzyme and 500 ng of DNA were incubated at the lowest optimal temperature for 2 h. A second series of digestions using gel-purified restriction enzymes: SmaI, and NotI were generated using [α-32P]dCTP with the Random Primed DNA Labeling kit (Life Technologies, Inc.) using the 21C/22NC cDNA PCR fragment as template DNA. The library was transfected into E. coli (P2PLK-F) and transferred to Duralon membranes (Stratagene). The Duralon membranes were preincubated at 51°C in Quik Hyb solution (Stratagene) for 30 min before hybridizing with radiolabeled probe for 1 h at 51°C. The membranes were washed at 25°C for 15 min in 2× SSC followed by a 30-min stringent wash at 56°C in 0.1× SSC. The membranes were then placed on Kodak X-OMAT film for overnight exposure and identification of positive clones. Clones were further purified by conducting secondary and tertiary screens.

Plaques of positive clones were cored from the stock plates and placed in SM buffer (59). The positive clones were further verified by PCR using the phage stock solution as the source of template DNA and primers 21C and 22NC to amplify a 615-bp vbs fragment. A second gene fragment of approximately 750 bp was obtained from a PCR reaction using oligonucleotides BNC and 10C. Generation of the double-stranded cDNA plasmid clones by in vivo excision of plasmid particles with helper phage R408 was carried out as described by the manufacturer (Stratagene). Rescued cDNA plasmids were amplified in E. coli (XL1-Blue). Plasmid pCVBS241e (Fig. 2) was used as template DNA for versicolorin B synthase cDNA sequencing.

Screening of gDNA Library Using PCR-generated vbs Probes—The A. parasiticus SU-1 genomic library was screened by plaque hybridization with radiolabeled 21C/22NC vbs DNA probes by a method analogous to the one described for the cDNA library. Plaques of positive clones were cored from the stock plates and placed in SM buffer. The positive clones were further verified to contain vbs by PCR using the phage stock solution as the source of template DNA and primers 21C and 22NC by amplifying an appropriate length vbs fragment (615 bp). A second gene fragment of approximately 800 bp was obtained from a PCR reaction using oligonucleotides BNC and 10C. The gDNA plasmid clone, pGVBS45.5, was constructed by ligating the 4.5-kb NotI/KpnI fragment from a λ gDNA clone, λgel2b, into the NotI/KpnI site of pBluescriptII SK(−) following established procedures (Fig. 3) (59). pGVBS45.5 was amplified in E. coli (XL1-Blue) and served as template DNA for double-stranded DNA sequencing.

Restriction Mapping of gDNA Clones—Representative procedures for restriction mapping by single and multiple digests are described by Ausubel et al. (63). Lambda clones (λ52a, λ55c, λ56a, λ57a, λ62b, and λ63c) were individually digested with combinations of the following restriction enzymes: BamHI, EcoRI, HindIII, KpnI, NotI, NcoI, NotI, SacI, SalI, Smal, XbaI, and XhoI and separated on both 3.5% and 0.75% agarose gels. Single and double restriction digestions containing 0.5 μg of each enzyme and 500 ng of DNA were incubated at the lowest optimal temperature for 2 h. A second series of digestions using gel-purified XbaI fragments resolved further complexities.

The DNA from Salli and Xbali digests was resolved by agarose gel electrophoresis and transferred onto a nitrocellulose membrane by capillary (Southern) transfer and cross-linked to the surface using a UV cross-linker (Stratagene, 1200 μJ (62, 63). Multiple filters were probed with 5′-γ-32P-radiolabeled oligonucleotide probes (1.25 × 10−6 cpm/ml) for vbs (25C, Table II) and omtA (Omt1–2NC, Table II). Prehybridizations and hybridizations were conducted with Quik Hyb solution in sealed bags at 42°C for 1 and 7 h, respectively (Stratagene). The membranes were washed with 2 × SSC/0.1% SDS at 25°C for 15 min followed by a single wash in 0.1 × SSC/0.1% SDS at 42°C for 1–1.5 h. The membrane was autoradiographed at −80°C for 5–18 h using intensifying screens and Kodak X-OMAT film.

PCR Analysis of gDNA Clones—To test for the presence of vbs, ver-1, and omtA, polymerase chain reactions were assembled with appropriate primers using master mixes and a hot start protocol. The required oligonucleotide primers (2.5 μl, 50 pmol each), DNA template (1.0 μl, 100 ng), and mineral oil (60 μl) were loaded into microfuge tubes followed by the “lower” master mix (44 μl) and briefly centrifuged. The microfuge tubes were placed in the thermocycler, which had been previously heated to 95°C. After the initial denaturation step (5 min), the “upper” master mix (50 μl) was rapidly added, and the PCR cycling was initiated.

The lower master solution contained for each reaction deionized distilled water (32.6 μl), 10× cloned Pfu polymerase buffer (3.4 μl), and dNTPs (2.0 μl each, 10 mm). The upper master solution for each reaction contained cloned Pfu polymerase (1.0 μl, 2.5 units), 10× cloned Pfu polymerase buffer (6.6 μl), and deionized distilled water (42.4 μl). The PCR cycling parameters (program B) and oligonucleotide primers for Pfu polymerase reactions are shown in Tables I and II. Samples were resolved on agarose gels (0.4–1.5%) agarose.

Measurement of omtA-vbs Proximity—Measurement of the distance between omtA and vbs was accomplished using the Taq extender procedure reported by Nielsen et al. (64), modified to parallel the master solution/hot start protocol described above (65). Oligonucleotide primers (2.5 μl each, 20 μM, 50 pmol), λgel2b DNA template (1.0 μl, 100 ng, 0.6 μl final), and mineral oil (60 μl) were loaded into microfuge tubes followed by the lower master mix (42.5 μl) and briefly centrifuged. The above-described procedure (Cycling Method B, Table I) was followed by 20 cycles of amplification. The upper master solution that contained Taq Extender enzyme (1.0 μl, 5 units), AmpliTaq polymerase (1.0 μl, 5 units), 10× Taq Extender buffer (5 μl), dNTPs (2.5 μl each, 10 mm), and deionized distilled water (44.25 μl). Oligonucleotide primers employed are indicated in Tables I and II. The primer combinations were Omt1–1C + 56NC, Omt1–2NC + 56NC, Omt1–3C + 56NC, Omt1–4NC + 56NC, and Omt1–1C + Omt1–4NC. The PCR products were separated by 0.7% agarose gel electrophoresis. Migratory distances were correlated with digested λ DNA markers to determine fragment sizes.

Nucleotide Sequencing and Analysis—Specific restriction fragments were subcloned into the plasmid vector pBluescriptII SK(−), also containing the apparent cytochrome P-450 gene upstream of vbs.
DNA for double-stranded sequencing was purified by the Qiagen plasmid purification procedure. DNA sequencing was accomplished using Sequenase-2.0 DNA polymerase as described by U. S. Biochemical Corp. and either commercially available or custom-synthesized oligonucleotide primers. Sequence data were compiled manually and analyzed using the DNA Strider program.

RESULTS

VBS, a homodimeric protein of 78-kDa subunits, catalyzes the dehydrative cyclization of racemic versicolin hemiacetal (7) to optically active versicolorin B (8), the penultimate to desaturation of the tetrahydrobifuran to the dihydrobifuran present in (−)-versicolorin A (9) (29–33). This unique structural feature is conserved through the subsequent intermediates of aflatoxin B1 (1) biosynthesis (Scheme II) and is the seat of the progressively severe carcinogenic properties of these metabolites. VBS was purified to homogeneity from A. parasiticus SU-1 (ATCC 56775) by methods established in this laboratory, but failed to give amino-terminal sequence data by automated methods. Although homogenous enzyme was submitted for amino acid sequence analysis (250 pmol), amino acid intensities corresponding to ≤10 pmol of enzyme were observed, suggesting that the amino terminus of the native protein was post-translationally modified. To circumvent this problem, the protein was digested with Lys-C endopeptidase and two of approximately 20 VBS peptide fragments were isolated by reverse-phase HPLC. These two peptide fragments both gave reproducible amino acid sequence data and credible stoichiometry (Fig. 1).

The two peptide sequences were used to design seven oligonucleotide probes for hybridization and PCR experiments. The degeneracy of these probes was minimized by comparing codon usage in A. nidulans and A. niger structural genes to compile a table of codon biases. Plaque hybridization experiments did not provide reproducible results using the partially degenerate probes 7NC, 8NC, 9NC, 10NC, 13C, 14C, and 15NC synthesized by automated methods (Fig. 1). However, PCR-generated nondegenerate probes were later substituted for these degenerate probes in the hybridization experiments, as described under “Experimental Procedures,” to lead to the successful cloning of VBS.

A vbs cDNA gene fragment, obtained from the reverse transcriptase-mediated PCR, was amplified using PCR primers 8NC and 10NC and estimated to be approximately 750 bp in length. No PCR product was obtained using any other primer combination from the set of primers shown in Fig. 1. PCR amplification of the approximately 750-bp fragment was observed using both 48- and 60-h mRNA. This gene fragment was subcloned into pBluescript SK(−) generating clones with the insert oriented in both directions (pRPF1–17C and pRPF1–13NC). Single-stranded DNA was prepared by infecting the plasmid-borne XL1-Blue cells with R408 helper phage. Direct sequencing was performed using dideoxy sequencing methods in both directions using clones containing inserts in opposite orientations. Partial nucleotide sequences from the coding (pRPF1–17C) and the noncoding strands (pRPF1–13NC) can be seen in Fig. 4. Translation of the two nucleotide sequences of both pRPF1–17C and pRPF1–13NC concurred with the amino acid sequence data obtained from each of the two peptide fragments isolated from the Lys-C endopeptidase digestion of VBS and pRPF1–17C and pRPF1–13NC nucleotide sequences.

From these nucleotide sequences, nondegenerate primers 21C and 22NC were prepared to serve as oligonucleotide primers for PCR experiments with both the gDNA and cDNA libraries (Fig. 4). A 615-bp internal fragment was successfully amplified by PCR from both the cDNA and gDNA libraries. This PCR fragment was used to generate oligonucleotide probes for plaque hybridizations as described under “Experimental Procedures.” Further PCR analysis of the gDNA and cDNA clones with primers 8NC and 10NC afforded two discrete gene products from each set of clones approximately 800 and 750 bp in length, respectively. The approximately 50-bp difference between the PCR fragments derived from cDNA and gDNA templates was attributed to the presence of an intron within the gene fragment, which was later verified by DNA sequence comparison.

From the lambda Uni-ZAP XR cDNA library prepared from 48-h A. parasiticus mRNA, approximately 150,000 plaques were screened. Sixteen positive cDNA clones were isolated and verified to be identical through restriction mapping. Approximately 150,000 plaquees were screened from the lambda FixII gDNA library of A. parasiticus yielding six positive gDNA clones. Each was verified to contain vbs by PCR analysis, restriction mapping, Southern analysis, and/or nucleotide sequencing.

Further investigations were undertaken with the genomic lambda clones, which successfully assigned the orientation of vbs and a probable cytochrome P450 monooxygenase (cyp) with respect to earlier portions of the putative aflatoxin gene cluster. Southern and PCR analysis positively identified the presence of omtA in two clones (λ52a and λ62b) and the absence of ver1 in all of the isolated gDNA clones. Although the vbs gene was verified by PCR to be present in each clone following library screening, subsequent examinations attested to a significant truncation of the 5′ terminus of vbs in two clones (λ52a and λ55c). Together, the genomic clones λ52a and λ62b contained approximately 30 kb of overlapping genomic sequence, as measured by restriction mapping (Fig. 5). The distance between vbs and omtA was measured by PCR and verified by DNA sequence analysis. Primer combinations Om1–2NC + 56NC and Om1–4NC + 56NC (Table II) gave 2.78- and 4.35-kb PCR products, respectively, which are in agreement with the known 1.49-kb separation between the Om1 primers (66). Employing the four omtA primers (Tables I and II) of known orientations with primer 56NC (0.56 kb upstream of vbs), vbs and omtA were determined to be located within ap-
Cloning of vbs from Aspergillus parasiticus

**Fig. 5.** Further characterization of the A. parasiticus partial gene cluster for the aflatoxin B<sub>1</sub> biosynthetic pathway. a, recently published gene cluster of AFB<sub>1</sub> biosynthetic genes of approximately 60 kb, b, lambda clone λ62b, approximately 15 kb, extending the existing AFB<sub>1</sub> gene cluster to include versicolorin B synthase (vbs), and an apparent cytochrome P450 by amino acid homology of the translated mRNA sequence. c, lambda clone λ52a, approximately 18 kb, where vbs is truncated at the 5' end (;). The sizes of the EcoRI restriction fragments are indicated in bold above the mapped DNA.

**TABLE I**

| Program | Cycle 1 | Cycle 2 (30 repetitions) | Cycle 3 | Cycle 4 |
|---------|---------|--------------------------|---------|---------|
| A       | 95 °C/5 min<sup>a</sup> | 94 °C/1 min | 94 °C/1 min | maintain reaction at 10 °C |
| B       | 95 °C/5 min<sup>a</sup> | 93 °C/2 min | 93 °C/1 min | maintain reaction at 12 °C |
| C       | 95 °C/5 min<sup>a</sup> | 95 °C/3 min | 95 °C/1 min | maintain reaction at 12 °C |

<sup>a</sup> Polymerase added following the indicated step.

**TABLE II**

| Primers and oligonucleotide probes used for PCR and Southern analyses of genomic DNA clones |
|-----------------------------------------------|
| Omt1-1C | 5'-TACCGAGGACAGCCGGCC-3' |
| Omt1-2NC | 5'-GCTTTGTCGCTTGATGGCC-3' |
| Omt1-3C | 5'-GAGGAGATATGTGGCCGC-3' |
| Omt1-4NC | 5'-AACGGCCCGGTGAGACC-3' |
| Ver1-1C | 5'-GGGTTGGATGGTGCCGC-3' |
| Ver1-2NC | 5'-GGCCTGTGACCAAGGCCC-3' |
| Ver1-3C | 5'-GCATGTGGATATGACCG-3' |
| Ver1-4NC | 5'-GACGCCGCCGATTCCAGG-3' |
| 25C | 5'-CGGAGAATATTGGTCAC-3' |
| 30NC | 5'-GAAACGTTGCACTACGG-3' |
| 41C | 5'-GGGGTAGGATGGGACAGC-3' |
| 56NC | 5'-TACGGCACTAGCATTCTCC-3' |

pyrimidine-rich motif, commonly associated with fungal promoters, was located between -72 and -60 nucleotides upstream of the start codon (70, 71). A common trend found in this sequence and many other filamentous fungi genes was an adenine at the third nucleotide upstream of the start codon (70, 71). At the 3' terminus, a polyadenylation tail was appended at position +161 from the end of the stop codon. This site does not correspond to the canonical poly(A) site, although this is not unusual in fungal genes (73). The polyadenylation consensus sequence was represented by 5'-AAATTAATA-3', 126 nucleotides after the stop codon.

Translation of the coding sequence provides a protein of 643 amino acids with a molecular mass of 70,271 Da and a calculated isoelectric point of 5.06. These values differ from those observed for the native protein (36–38, 83).<sup>2</sup> The monomeric molecular mass of native VBS as estimated by SDS-polyacrylamide gel electrophoresis and size exclusion chromatography is approximately 78 kDa, with an experimentally determined isoelectric point of 4.7 ± 0.1. Recent work in our laboratory has demonstrated that the native protein is N-glycosylated (data not shown). From the translated amino acid sequence, there are three potential N-glycosylation sites with the motif Asn-Xaa-Thr or Asn-Xaa-Ser. The discrepancy in molecular mass and pI can be attributed to the post-translational modification of the native protein.

The amino acid sequence of the VBS protein was found to have significant homology to many flavin-dependent oxidases and dehydrogenases through BEAUTY (75) and BLAST (76) searches of the Brookhaven protein, SWISS-PROT, PIR, and GenBank<sup>®</sup> data bases (Table III). Choline dehydrogenase and glucose oxidase provided the highest correlations among the homologous proteins identified. Specifically, the BEAUTY search identified the greatest homology with proteins in the GMC oxidoreductase family (77) (cluster ID 3015). and other
flavin-dependent oxidases and dehydrogenases (Table III). Interestingly, glucose oxidase from *A. niger*, like VBS, is homodimeric and has a similar molecular mass. The former has eight potential *N*-glycosylation sites and is so modified in at least two of these (78). Glucose oxidase has a pI of 4.1 ± 0.1 (74).

The results from the GAP alignment of VBS (644 amino acids) and choline dehydrogenase (557 amino acids) showed 34% identity and 56% similarity over the entire VBS amino acid sequence. The GAP alignment to glucose oxidase (583 amino acids) showed 38% identity and 58% similarity over the entire VBS protein sequence (79). Strong regions of homology were observed in the nucleotide phosphate binding sites and the active sites of the GMC family of oxidoreductases (77). Alignments of each of these regions are shown in Fig. 7. An x-ray crystal structure of glucose oxidase from *A. niger* has been reported at 2.3 Å resolution (78). One FAD molecule is bound in each identical subunit, and these reside near the dimer interface in a βαβ-motif showing high structural conservation. Significant hydrogen bonding interactions are evident to the FAD, particularly to the ribose and phosphate groups. The principal interactions between the protein and the FAD are seen in Glu72, Gly49, and Gly123 in the amino-terminal region. The first two of these correspond to exact amino acid matches in the aligned VBS sequence, whereas the third does not. His102, thought to be hydrogen bonded to the ribose-2' oxygen in glucose oxidase, has been replaced by a tyrosine in VBS. The diphosphate group is involved in several hydrogen bonds, in part to water molecules and to Thr52, which has been replaced by alanine in the aligned VBS sequence, although the threonine can be found at the amino-terminal adjacent site. Although displaced by one residue, this threonine aligns with threonine or serine in all other members of the GMC oxidoreductase family (79). This sequence motif is associated with phosphate binding and is fully conserved in VBS.

So, although important interactions between glucose oxidase and FAD show strong correlations in the structure of VBS, a striking 23-amino acid gap exists between Gly138 and Phe139. This is a significant deletion in the middle of the potential FAD binding domain and a gap not present in any of the currently known GMC family of flavoproteins. Moreover, two amino acid contacts to FAD in glucose oxidase lying carboxyl-terminal to this gap do not map to identical residues in VBS. In this connection it is noteworthy that homogeneous VBS does not contain a bound flavin chromophore and preliminary kinetic evidence suggests that FAD, FMN, and glucose have little or no inhibitory effect on the cyclization of versiconal (7) to versicolorin B (8) (35, 38).2

**DISCUSSION**

VBS catalyzes the dehydrative cyclization of (±)-versiconal hemiacetal (7) to set the absolute configuration of (±)-versicolorin B (8) and, hence, aflatoxin B1 (4) (28, 29, 38). This key cyclization reaction in aflatoxin biosynthesis has been demonstrated by isolation and purification of the native protein (36–38) and by expression of vbs in *S. cerevisiae* to afford fully

**FIG. 6.** Nucleotide sequence for the gDNA and cDNA clones of vbs. The translated amino acid sequence is shown below the coding DNA sequence. Transcribed nucleotides are indicated by uppercase letters, whereas introns and nontranscribed nucleotides are in lowercase letters. Probable consensus sequences are as follows: underlining, Hogness box; , transcriptional start codon; double dashed underlining, polyadenylation signal sequence; and , polyadenylation site. The three possible sites for N-glycosylation are indicated by underlined and italicized amino acids.
active enzyme. Design of PCR probes from amino acid sequence data derived from the pure protein allowed the VBS gene to be isolated from both cDNA and gDNA libraries of A. parasiticus. The unexpectedly high homology of the translated sequence information (29) identified the organization of several demonstrated and presumed ketoreductase from the Streptomyces actII gene (48). The transformation of the anthraquinone 9 to the xanthone 10 involves oxidative ring cleavage, rearrangement, deoxygenation, and decarboxylation and may well require several enzymes (57). The degree to which these proteins act individually or in a tightly ordered or even physically associated manner remains to be established. No intermediates in this process have been isolated or, indeed, in the equally cryptic xanthone 11 transformation (45). Finally, the clear demonstration that 6-deoxyversicolorin A is not an intermediate in aflatoxin biosynthesis is at loggerheads with the function assumed for VER-1 based on protein sequence information (29).

Recent work by Linz, Bhatnagar, Payne and co-workers (51) has established partial organization of the aflatoxin B1 biosynthetic genes in A. parasiticus. In contrast to the earlier findings of Papa (53–55), whose work identified several linkage groups in A. parasiticus for AF B1-related genes, it appears that like a growing number of other secondary metabolites, aflatoxin is a further example of a natural product whose biosynthetic genes are clustered. These workers have identified the organization of several demonstrated and pre-

---

**Flavin Adenine Dinucleotide Binding Site:**

| NO. | High score | Sum probability (regions) p(N), N |
|-----|------------|----------------------------------|
| 1   | 121        | 5.3e-55 (8)                      |
| 2   | 112        | 3.2e-52 (7)                      |
| 3   | 127        | 1.2e-46 (7)                      |
| 4   | 136        | 3.6e-45 (7)                      |
| 5   | 126        | 2.0e-41 (7)                      |
| 6   | 96         | 1.4e-30 (7)                      |
| 7   | 120        | 2.0e-30 (6)                      |
| 8   | 108        | 1.9e-29 (6)                      |
| 9   | 97         | 2.2e-29 (6)                      |
| 10  | 97         | 2.2e-29 (6)                      |

**GM Oxidoreductase Active Site:**

| NO. | High score | Sum probability (regions) p(N), N |
|-----|------------|----------------------------------|
| 1   | 121        | 5.3e-55 (8)                      |
| 2   | 112        | 3.2e-52 (7)                      |
| 3   | 127        | 1.2e-46 (7)                      |
| 4   | 136        | 3.6e-45 (7)                      |
| 5   | 126        | 2.0e-41 (7)                      |
| 6   | 96         | 1.4e-30 (7)                      |
| 7   | 120        | 2.0e-30 (6)                      |
| 8   | 108        | 1.9e-29 (6)                      |
| 9   | 97         | 2.2e-29 (6)                      |
| 10  | 97         | 2.2e-29 (6)                      |

---

**Fig. 7. Sequence alignment of conserved flavin binding and GM oxidoreductase active sites as determined by BEAUTY.** The numbered lines correspond to the following proteins: 1, GM oxidoreductase conserved motif (cluster 3015); 2, glucose dehydrogenase (spP18173); 3, glucose dehydrogenase (spP18172); 4, choline dehydrogenase (spP17444); 5, alcohol dehydrogenase (spQ00593); 6, ORF2 (g576664); 7, mandelonitrile lyase (pirS32156); 8, glucose oxidase precursor (spP13006); and 9, versicolorin B synthase (the protein data banks and accession numbers are indicated in parentheses). An extended alphabet has been used to supplement the standard amino acid code in which combinations of amino acids observed at each aligned position are represented as defined by the Pattern-induced Multiple-sequence Alignment program multiple alignment: b, IL; c, FY; d, ST; f, LV; g, gap; h, AG; i, ILV; B, ND; j, IV; U, RK; X, wildcard; &; RS; $, IT; @, MV; \, EP (67).
sumed AF₂ genes: pksA, nor-1, fas-1A, fas-2A, afr, aad, ord1, ord2, and omtA, although only a few of these have been well characterized. The data presented in this paper further define the extent of clustering of the aflatoxin B₁ biosynthetic genes. Southern analysis and restriction mapping of the vbs gDNA clones resulted in the discovery that omtA, a later gene in the aflatoxin B₁ biosynthetic pathway involved in the S-adenosylmethionine-dependent formation of, was located within one of the vbs gDNA clones (Fig. 5) (J. Yu, 1993, GenBank accession number L25834). We have demonstrated that vbs and omtA are within 3.3 kb of each other by PCR and Southern analysis (Fig. 5). We have also identified a probable GenBank accession number L25834. We have demonstrated amino acid sequence analyses.

Acknowledgment—We thank J. Franklin of the Protein/Peptide/DNA Facility at The Johns Hopkins University School of Medicine, Department of Biological Chemistry, for custom oligonucleotide syntheses and Southern analysis (Rosenkranz, H. S., ed) pp. 141–154, Plenum Publishing Corp., New York.

REFERENCES

1. Lloyd, A. T., and Sharp, P. M. (1991) Mol. Gen. Genet. 230, 288–294
2. Lillegård, E. B., and Heseltine, C. W. (1977) in Mycotoxins in Human and Animal Health (Rodríguez, J. V., Heseltine, C. W., and Mehman, M. A., eds) Plenum Press, New York.
3. Dickens, J. W. (1977) in Mycotoxins in Human and Animal Health (Rodríguez, J. V., Heseltine, C. W., and Mehman, M. A., eds) pp. 99–105, Pathox Publishers, Park Forest, IL.
4. Baertschi, S. W., Raney, K. D., Stone, M. P., and Harris, T. M. (1988) J. Am. Chem. Soc. 110, 7923–7931
5. Martin, C. N., and Garner, R. C. (1977) Nature 267, 863–865
6. Goggin, J. M., and Holmes, H. A., Jr., Reinhold, V. N., Büchli, G., and Wogan, G. N. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 1970–1974
7. Locorriere, L. L., Teeter, M. M., and Whitlow, M. D. (1988) J. Bacteriol. Struct. Dyn. 5, 1237–1257
8. Neal, G. E. (1973) Nature 244, 432–435
9. Saunders, F. C., Barker, E. A., and Smucker, E. A. (1972) Cancer Res. 32, 2847–2849
10. Gelboin, H. V., Wortham, J. S., Wilson, R. G., Friedman, M., and Wogan, G. N. (1976) Science 194, 1205–1206
11. Lillehoj, E. B., and Hesseltine, C. W. (1977) in Mycotoxins in Human and Animal Health (Rodríguez, J. V., Heseltine, C. W., and Mehman, M. A., eds) Plenum Press, New York.
12. Ridley, B. S., and McLaughlin, B. M. (1961) Biochim. Biophys. Acta 27, 432–435
13. Manios, T. P., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
14. Chirico, R., and Saphire, P. S. (1980) Nucleic Acids Res. 8, 793–804
15. Horng, J. S., Linz, J. E., and Pesta, J. J. (1989) Appl. Environ. Microbiol. 55, 2561–2568
16. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Alibert, L. M., Coen, D. M., and Vakar, A. (1994) Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York.
17. Keller, N. P., Segner, S., Bhatnagar, D., and Adams, T. H. (1995) Appl. Environ. Microbiol. 61, 3628–3632
18. Ades, J., and Mateles, R. L. (1986) Biochim. Biophys. Acta 863, 418–420
19. Watanabe, C. M. H., and Townsend, C. A. (1996) J. Org. Chem. 61, 200–207
20. Mancini, P., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
21. Chirico, R., and Saphire, P. S. (1980) Nucleic Acids Res. 8, 793–804
22. Horng, J. S., Linz, J. E., and Pesta, J. J. (1989) Appl. Environ. Microbiol. 55, 2561–2568
23. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Alibert, L. M., Coen, D. M., and Vakar, A. (1994) Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York.
24. Keller, N. P., Segner, S., Bhatnagar, D., and Adams, T. H. (1995) Appl. Environ. Microbiol. 61, 3628–3632
25. Ades, J., and Mateles, R. L. (1986) Biochim. Biophys. Acta 863, 418–420
26. Watanabe, C. M. H., and Townsend, C. A. (1996) J. Org. Chem. 61, 200–207
27. Mancini, P., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
28. Chirico, R., and Saphire, P. S. (1980) Nucleic Acids Res. 8, 793–804
29. Horng, J. S., Linz, J. E., and Pesta, J. J. (1989) Appl. Environ. Microbiol. 55, 2561–2568
30. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Alibert, L. M., Coen, D. M., and Vakar, A. (1994) Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York.
31. Keller, N. P., Segner, S., Bhatnagar, D., and Adams, T. H. (1995) Appl. Environ. Microbiol. 61, 3628–3632
32. Ades, J., and Mateles, R. L. (1986) Biochim. Biophys. Acta 863, 418–420
33. Watanabe, C. M. H., and Townsend, C. A. (1996) J. Org. Chem. 61, 200–207
34. Mancini, P., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
35. Chirico, R., and Saphire, P. S. (1980) Nucleic Acids Res. 8, 793–804
36. Horng, J. S., Linz, J. E., and Pesta, J. J. (1989) Appl. Environ. Microbiol. 55, 2561–2568
37. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Alibert, L. M., Coen, D. M., and Vakar, A. (1994) Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York.
38. Keller, N. P., Segner, S., Bhatnagar, D., and Adams, T. H. (1995) Appl. Environ. Microbiol. 61, 3628–3632
39. Ades, J., and Mateles, R. L. (1986) Biochim. Biophys. Acta 863, 418–420
40. Watanabe, C. M. H., and Townsend, C. A. (1996) J. Org. Chem. 61, 200–207
41. Mancini, P., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
42. Chirico, R., and Saphire, P. S. (1980) Nucleic Acids Res. 8, 793–804
43. Horng, J. S., Linz, J. E., and Pesta, J. J. (1989) Appl. Environ. Microbiol. 55, 2561–2568
44. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Alibert, L. M., Coen, D. M., and Vakar, A. (1994) Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York.