

Characterization of the Binding Protein-Dependent Cellobiose and Cellotriose Transport System of the Cellulose Degradar *Streptomyces reticuli*

ANDREAS SCHLÖSSER,* JENS JANTOS, KARL HACKMANN, AND HILDGUND SCHREMPF

FB Biologie/Chemie, Universität Osnabrück, D-49069 Osnabrück, Germany

Received 30 November 1998/Accepted 3 March 1999

***Streptomyces reticuli* has an inducible ATP-dependent uptake system specific for cellobiose and cellotriose. By reversed genetics a gene cluster encoding components of a binding protein-dependent cellobiose and cellotriose ABC transporter was cloned and sequenced. The deduced gene products comprise a regulatory protein (CebR), a cellobiose binding lipoprotein (CebE), two integral membrane proteins (CebF and CebG), and the NH₂-terminal part of an intracellular β -glucosidase (BglC). The gene for the ATP binding protein MsiK is not linked to the *ceb* operon. We have shown earlier that MsiK is part of two different ABC transport systems, one for maltose and one for cellobiose and cellotriose, in *S. reticuli* and *Streptomyces lividans*. Transcription of polycistronic *cebEFG* and *bglC* mRNAs is induced by cellobiose, whereas the *cebR* gene is transcribed independently. Immunological experiments showed that CebE is synthesized during growth with cellobiose and that MsiK is produced in the presence of several sugars at high or moderate levels. The described ABC transporter is the first one of its kind and is the only specific cellobiose/cellotriose uptake system of *S. reticuli*, since insertional inactivation of the *cebE* gene prevents high-affinity uptake of cellobiose.**

The ABC superfamily of transporters has been extensively studied, and members have been identified in most bacteria, archaea, and eukaryotes (9). Uptake of a large variety of nutrients seems to be the most obvious task of these systems. Moreover, ABC transporters are involved in the export of drugs or virulence factors, such as hemolysin, extracellular proteases, and toxins; signal transduction; plant host–bacterial-parasite interaction; antigen presentation in immune cells; transport of pheromones; and sporulation of gram-positive bacteria (3, 7).

Members of the family of binding protein-dependent systems have so far been identified only in prokaryotes (3, 7). These multicomponent systems consist of two membrane-inserted subunits, two components inside the cytoplasm that carry the ATP binding site, and the binding protein located outside the cytoplasm. The binding proteins are responsible for the substrate specificity of the ABC transporter. In gram-negative bacteria, the binding protein is a soluble periplasmic protein, whereas in gram-positive bacteria and archaea, it is a lipoprotein exposed to the cell surface (38).

Streptomyces reticuli is a soil bacterium which hydrolyzes crystalline cellulose (Avicel) due to the action of an exoglucanase (Avicelase, Cel1) (32, 33, 44). The generated cellobiose and cellotriose are taken up via an inducible, binding protein-dependent ABC transporter (cellobiose/cellotriose ABC transport system [34]). The corresponding cellobiose/cellotriose binding protein was shown to be a lipoprotein anchored to the cytoplasmic membrane. This protein was extracted from membranes of *S. reticuli* and purified to homogeneity; it binds to cellobiose and cellotriose with equally high affinities (34). The ATP-hydrolyzing subunit of the cellobiose/cellotriose ABC transporter is MsiK (35). The *msiK* gene is a homologue of a previously described *Streptomyces lividans msiK* (15). MsiK

from both *Streptomyces* species assists two different ABC transporters, one for maltose and one for cellobiose and cellotriose. Earlier studies have indicated that the chromosomally located *msiK* gene of *S. reticuli* is not situated in the vicinity of genes encoding the other components of the cellobiose/cellotriose ABC transport system (35).

In this report we characterize the additional clustered genes of the cellobiose/cellotriose transport system from *S. reticuli*. Further physiological, immunological, transcriptional, and mutational experiments elucidate details of the cellobiose/cellotriose ABC transport system.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. The wild-type strain *S. reticuli* Tü45 (DSM 40776; German Collection of Microorganism and Cell Cultures, Braunschweig, Germany) described by Wachinger et al. (44) was obtained from H. Zähler, Tübingen, Germany. The *Escherichia coli* plasmids pUC18 and pUC19 (47) were used as cloning vectors for DNA sequence analysis and for disruption of the *cebE* gene. The aminoglycoside 3'-phosphotransferase gene (25) from pUC4K, supplied by Pharmacia, was used for gene disruption experiments. The *E. coli* K-12 strain DH5 α containing the plasmids used in this study was grown in Luria-Bertani medium with 100 μ g of ampicillin ml⁻¹ (with pUC derivatives) or 50 μ g of kanamycin ml⁻¹ (with pUC4K derivatives) (29). *S. reticuli* was cultivated in pH-stable minimal medium (45) supplemented with the appropriate carbon source (1%, wt/vol) and 10 mM (NH₄)₂SO₄ as a nitrogen source. For protoplast preparation *S. reticuli* was grown in complete medium (Oxoid tryptone-soy broth [20 g liter⁻¹], yeast extract [5 g liter⁻¹], sucrose [100 g liter⁻¹], MgCl₂ [10 g liter⁻¹]).

DNA preparation and manipulations. Genomic DNA from *S. reticuli* was isolated as described previously (12). Plasmids were isolated from *E. coli* with the aid of a midi plasmid kit (Qiagen GmbH, Hilden, Germany). Restriction enzyme digestions, ligation reactions, and analyses of DNA with nucleases and polymerases were carried out by standard procedures (29). PCR amplification of *S. reticuli* chromosomal DNA was performed with the oligonucleotides CB1, 5'-G GACATCAACATCAAGGAGAA-3', and CB2, 5'-CTCCTTSCCSAGGTCSA CGAA-3', the former corresponding to the DINIKEN motif of amino acids located within the signature sequence of CebE (see Fig. 3). PCR was done under standard conditions (1) but in the presence of 5% dimethyl sulfoxide in a total volume of 50 μ l. The mixture was covered with 30 μ l of mineral oil and subjected to 30 cycles of 1 min at 96°C, 1 min at 52°C, and 1 min at 72°C. PCR products were purified with a QIAquick PCR purification kit (Qiagen GmbH), cloned into plasmid pUC18 with the aid of a Sure Clone ligation kit (Pharmacia, Freiburg, Germany), and subjected to nucleotide sequencing.

* Corresponding author. Mailing address: FB Biologie/Chemie, Universität Osnabrück, Barbarastraße 11, D-49069 Osnabrück, Germany. Phone: 49 541 969 2843. Fax: 49 541 969 2804. E-mail: schloesser@biologie.uni-osnabrueck.de.

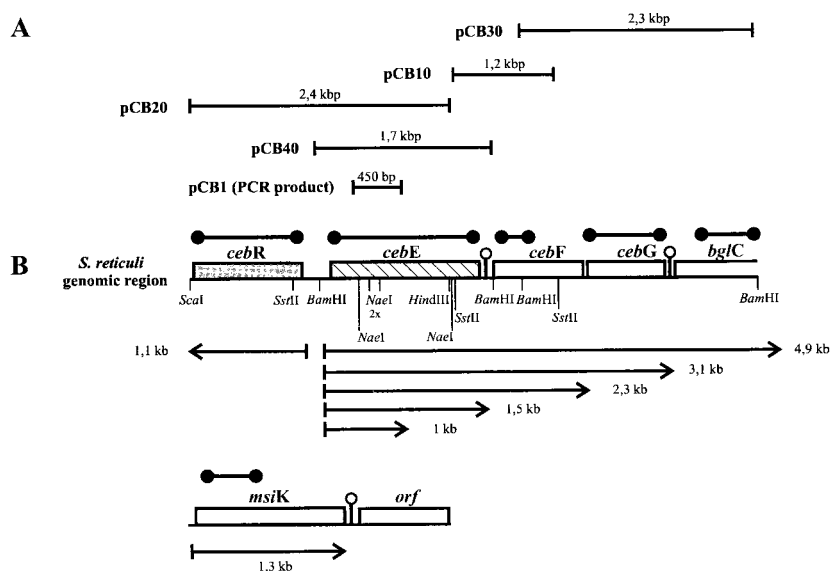


FIG. 1. Constructs and organization of genes. (A) Initial PCR product and maps of the cloned chromosomal fragments in pUC18. (B) Arrangement of genomic *cebREFG* and *bglC*. The restriction sites relevant for cloning are shown. The probes used for transcription analysis are shown above the genes (●—●). Arrows mark the positions of transcripts. Predicted transcription terminators are indicated by stem-loop structures. *orf*, open reading frame.

Preparation and screening of subgenomic DNA libraries. Total DNA (200 μ g) from *S. reticuli* was cleaved with *Bam*HI, and the resulting fragments were separated on an agarose gel. Fragments of about 1.5 to 2 kb were eluted with a QIAEX II gel extraction kit and ligated to *Bam*HI-digested and dephosphorylated pUC18. The ligation mixtures were transformed to *E. coli* DH5 α by electroporation with an Electroporator II from Invitrogen (NV Leek, The Netherlands). Ampicillin-resistant transformants were tested for the presence of the desired *cebE* gene by colony hybridization at 54°C overnight with the digoxigenin-labelled PCR fragment (DNA labelling and detection kit; Boehringer, Mannheim, Germany). Three additional subgenomic DNA libraries consisting of *Sst*II (1 to 1.5 kb), *ScaI*-*Hind*III (2 to 2.5 kb), and *Bam*HI (2 to 2.5 kb) were generated in pUC18 and screened as described above.

DNA sequence analysis. DNA sequencing of both strands of the *ceb* region was performed with double-stranded DNA based on the dideoxy chain termination method (30) with a Cy5 Autoread sequencing kit, Cy5-dATP labelling mix (Pharmacia), and universal or specific primers (MWG-Biotech, Ebersberg, Germany). Sequencing reactions were run on an ALFexpress sequencer from Pharmacia. The DNA and protein sequences were analyzed with the GENMON program (GBF, Braunschweig, Germany), as well as the Genetics Computer Group sequence analysis software package (version 8.0; Biotechnology Center, University of Wisconsin, Madison). Reading frames were determined with the GCWIND program (D. Shields, Dublin, Ireland) on the basis of the codon usage preference in *Streptomyces* DNA. The predicted proteins were scanned for similarities to sequences in the SWISS-PROT and EMBL databases with the FASTA and BLITZ algorithms (27). Membrane-spanning segments were predicted by the TMpred program according to the method of Hoffman and Stoffel (10).

Determination of amino acid sequences. Purified CebE protein (2 mg) was incubated overnight at 30°C with 20 mM CNBr in 80% (vol/vol) formic acid and subsequently evaporated to dryness. Resulting peptides were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (15%, wt/vol, acrylamide) and transferred onto a polyvinylidene difluoride membrane (Immobilon P; Millipore GmbH, Eschborn, Germany) as described previously (11). After Coomassie brilliant blue staining, different protein bands were excised and NH₂-terminal amino acids were determined by Edman degradation by R. Schmid, University of Osnabrück, Osnabrück, Germany.

Generation of MsiK antibodies and Western blot analysis. A DNA fragment (372 bp) encoding the C terminus of MsiK was amplified with the primers MsiK1 (CTGTCCAACCTGGACGCCAAG) and MsiK2 (GTGCTCGGGCGGACG CCGAC) with chromosomal DNA of *S. reticuli* and cloned into *Sma*I-restricted pUC18. From this plasmid the PCR fragment was isolated as a *Bam*HI-*Kpn*I fragment (383 bp) and recloned with pQE31 (Qiagen). The resulting construct, pMS131, contained the desired part of *msiK* with six codons encoding histidines at its 5' end. Strain SG13009 (Qiagen) transformed with pMS131 produced the His tag MsiK fusion protein in inclusion bodies. Once we obtained and solubilized the inclusion bodies, the fusion protein was purified by affinity chromatography with Ni²⁺-nitrilotriacetic acid according to the instructions of the manu-

facturer (Qiagen). Antiserum was obtained by immunization of a rabbit with the purified six-His-MsiK fusion protein (Eurogentec, Seraing, Belgium). Western blot analyses were conducted as described previously (40) with a 1:10,000 dilution of antibodies raised against His-tagged MsiK or a 1:100,000 dilution of antibodies raised against CebE (34).

Construction of a *cebE* disruption mutant. The *cebE* gene in pCB40 was cleaved by *Nae*I, resulting in an internal *cebE* deletion of 1,037 bp. Overhangs of the remaining plasmid were filled in with the Klenow enzyme and deoxynucleoside triphosphates and ligated to a *Pst*I fragment from pUC4K (Pharmacia) comprising the kanamycin resistance (*aph*I) gene (25). The resulting construct was named pCB41. Protoplasts of *S. reticuli* were generated (12) and transformed with pCB41, which was isolated from the *dam* and *dcm* methylation-deficient *E. coli* strain JM110. Transformants were selected by overlaying regenerating protoplasts with agar (0.75%) containing kanamycin (20 μ g/ml).

RNA isolation and Northern blot analysis. Total RNA was isolated from *S. reticuli* with acid guanidinium thiocyanate-phenol-chloroform (6). For Northern blot analysis, RNA (15 μ g per slot) was separated in formaldehyde gels (1) and transferred onto nylon membranes by vacuum blotting (LKB 2016 VacuGene; Pharmacia). RNA molecular weight marker I (0.3 to 6.9 kb) from Boehringer was used for size determination. The molecular weight markers were stained on the surface of the nylon filter with 0.2% methylene blue in 0.2 M sodium acetate (pH 4.7) (20). Hybridization was performed at 64°C for 20 h in a solution containing 5 \times SSC (1 \times SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 0.1% SDS, 100 μ g of salmon sperm DNA ml⁻¹, and 5 \times Denhardt's reagent (29) with randomly [α -³²P]dCTP-labelled DNA fragments (Rediprime DNA labelling system; Amersham, Freiburg, Germany). The membrane was washed twice in 2 \times SSC-0.1% SDS for 5 min and twice in 0.1 \times SSC-0.1% SDS for 30 min and subjected to autoradiography at -70°C.

S1 nuclease mapping of transcription start sites. The probes were generated by PCR amplification of a 396-bp fragment with pCB20 as the template and 10 pmol of the synthetic oligonucleotides E101 (for the *cebE* probe) and R103 (for the *cebR* probe) (see Fig. 2). Primers were labelled with [γ -³²P]ATP and T4 polynucleotide kinase (1) and used for PCRs with the corresponding unlabelled primer. Labelled PCR products were purified with a QIAquick PCR purification kit (Qiagen). For every S1 nuclease protection reaction, 50 μ g of RNA was hybridized to $\sim 10^5$ Cerenkov counts of the PCR fragment in Na-TCA buffer (23) min⁻¹ at 45°C for 6 h, after denaturation at 65°C for 15 min. Hybridization products were digested with S1 nuclease (Gibco, Life Technologies, Karlsruhe, Germany) as outlined by the manufacturer. Undigested radiolabelled DNA was precipitated with ethanol and run on a 6% polyacrylamide gel. The dried gel was subjected to autoradiography at -70°C. Sequence ladders were generated by the dideoxy chain termination method with the labelled primers R103 and E101 and with pCB20 as a template.

Nucleotide sequence accession number. The nucleotide sequence data reported in this article are stored in the EMBL database under the accession no. AJ009797 and AJ009798.

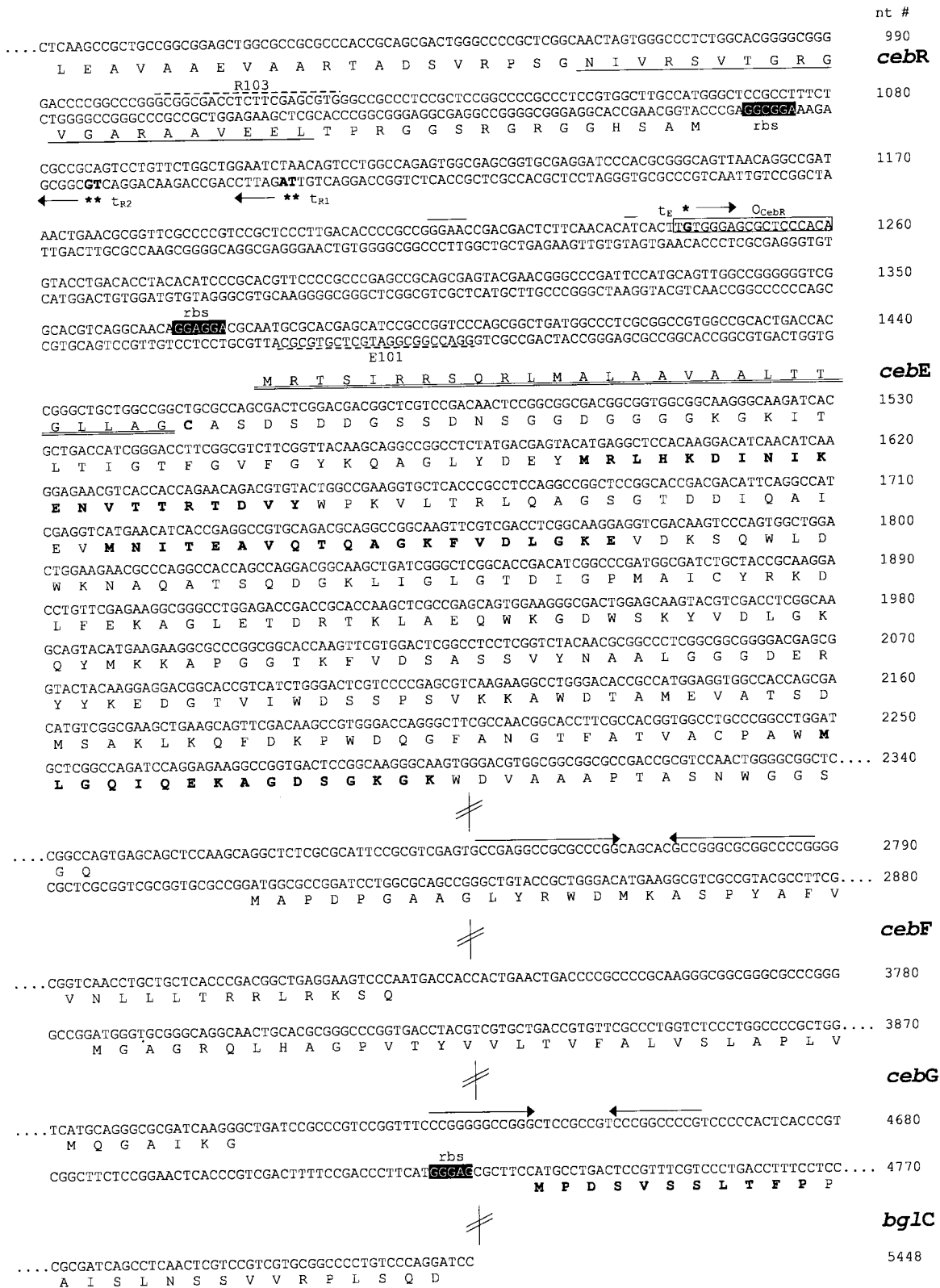


FIG. 2. Partial nucleotide sequences and deduced amino acid sequences of *cebREFG* and *bgIC*. The deduced amino acid sequences are given in the one-letter code below the nucleotide (nt) sequence, and nucleotide numbers are shown on the right. The putative ribosome binding sites (rbs) are in black letters on a black background, and predicted terminators are indicated by arrows above the sequence. The transcriptional start sites are in boldface type and are marked by asterisks followed by arrows indicating the direction of transcription (t_E , transcription start site for *cebE*; t_{R1} and t_{R2} , transcription start sites for *cebR*). The sequence similar to that of the *Streptomyces* class E promoter is marked by a line above the sequence. The oligonucleotides R103 and E101 used for S1 nuclease mapping are represented by broken lines. The predicted HTH motif in the deduced CebR protein is underscored, and the putative operator sequence for CebR binding is boxed (O_{CebR}). The signal peptide of CebE followed by the recognition sequence for the cleavage site of lipoprotein signal peptidase is double underscored. The NH_2 -terminal amino acids of peptides from the purified CebE and BglC proteins determined by Edman degradation are in boldface type. Most of the sequence within the structural genes has been omitted, as is indicated by dots and double-slashed bars.

Signature sequence	LXXL	GKXF	EXDX	GIK	VXV
	I	IAD	YT	E	NV
	V	VIQ	NA	DY	P
	A	WV		P	
CebE <i>Sr</i> (68)	<u>LY</u> DEYMR	LHKDI	NI <u>KEN</u> V	TTR	TDVY
MalE <i>Sc</i> (55)	LVKEFE	AANK	GV <u>KV</u> NF	VNV	FPDQ
MalE <i>Ec</i> (46)	LAEVG	KKFEKDT	GI <u>KV</u> TV	VEHP	DKLE
MalX <i>Sp</i> (57)	IEEVAK	AYEKEA	GV <u>KV</u> TL	TKTG	DALG
MsmE <i>Sm</i> (51)	LKKI	IKDFERE	NP <u>KI</u> HV	KMTS	VSPD
AmyE <i>Tt</i> (68)	LQNA	INTYMKEN	PV <u>KI</u> N	LETV	GGGQ
UgpB <i>Ec</i> (41)	VDSL	AQR	FNAEN	PDY	KI VPT
					YKGNVE

FIG. 3. Alignments of signature sequences from binding proteins. The numbers indicate the positions of the amino acids. The highly conserved lysine (R) residue (according to the work of Tam and Saier [39]) is given in boldface type, and residues conserved in CebE are underlined. *Sr*, *S. reticuli*; *Sc*, *S. coelicolor* A3(2); *Ec*, *E. coli*; *Sp*, *Streptococcus pneumoniae*; *Sm*, *Streptococcus mutans*; *Tt*, *Thermoanaerobacterium thermosulfurigenes*; MalE, MalX, and AmyE, maltose and maltodextrin binding proteins; UgpB, sn-glycerol-3-phosphate binding protein; MsmE, multiple-sugar binding protein.

RESULTS AND DISCUSSION

Cloning of the *cebREFG* cluster and the *bgIC* gene. As the NH₂-terminal amino acids of the purified CebE protein (34) could not be obtained by Edman degradation, several internal peptides were generated by CNBr treatment. After separation by SDS-PAGE, NH₂-terminal amino acids were determined from three peptides (see Fig. 2) and used to deduce corresponding oligonucleotides. These, in turn, were used to generate fragments from total *S. reticuli* DNA in PCRs. The obtained fragments were cloned into pUC18, and their nucleotide sequences were determined. From the construct which comprises a fragment encoding a portion of the CebE protein (pCB1) (Fig. 1A), the PCR fragment was reisolated, labelled, and used to screen a subgenomic *S. reticuli* DNA library in *E. coli* DH5 α (containing 1.5- to 2-kb *Bam*HI genomic *S. reticuli* fragments in pUC18). Several clones were identified by colony hybridization, and their plasmids were analyzed. Sequencing revealed that an inserted 1.7-kb *Bam*HI fragment was the desired one (Fig. 1). As genomic *S. reticuli* *Sst*II (1.2 kb) and *Hind*III-*Sca*I (2.4 kb) fragments hybridized with the 1.7-kb *Bam*HI fragment, corresponding subgenomic libraries were generated in pUC18 and the desired overlapping fragments were gained (Fig. 1). The genomic 2.3-kb *Bam*HI fragment was

obtained with the 1.2-kb *Sst*II probe in a manner similar to that described above.

Determination of the sequence and its analysis. The sequences of the cloned overlapping fragments were determined. FRAME analysis (2) of the assembled 5,448-bp DNA fragment revealed the presence of five reading frames, whose codon usage was found to be typical of GC-rich *Streptomyces* DNA. The sequence of the first reading frame comprises 1,056 bp with a G+C content of 74 mol%. A start codon (ATG), a putative ribosome binding site, and a stop codon (TGA) were identified (Fig. 2). The deduced amino acid sequence of this open reading frame encodes a 39-kDa protein of 351 amino acids (aa). It was named CebR, as within its NH₂ terminus a helix-turn-helix (HTH)-containing region characteristic of DNA binding proteins was identified (46). The deduced CebR is most closely related to GalR (38% of its amino acid residues are identical) and RbsR from *E. coli* (37% of its amino acid residues are identical), the latter representing a repressor of the ribose ABC transport operon (19). GalR and RbsR are both members of the LacI-GalR regulatory family. Interestingly, only 32% of CebR's amino acid residues are identical to those of the recently identified *Streptomyces coelicolor* A3(2) MalR regulator of the operon for a maltose ABC transporter (41).

The most conserved portion of HTH motifs from various members of the LacI-GalR family comprises the six residues ATVSRV, which make up the main portion of the recognition helix (GTVSRV in CebR) required for specific binding to the major groove within the operator sequences (36). By sequence alignments it was suggested that the C-terminal domain of GalR is homologous to the *E. coli* periplasmic D-galactose-D-glucose binding protein (46). Data from the crystal structures of several periplasmic binding proteins (22, 43) were used for molecular modelling of the binding domain for D-galactose-D-glucose in GalR (14). The amino acid side chains of Phe 73, Arg 194, Asn 245, and Asp 273 from GalR were predicted to be involved in creating the saccharide binding pocket. Three corresponding residues, Phe 73, Arg 194, and Asp 273, are found in the deduced CebR protein, and it appears likely that they interact with cellobiose.

The *cebR* gene is followed by the reading frame named *cebE* (1,334 bp), which has an opposite orientation. A start codon (ATG) with a putative Shine-Dalgarno sequence and a stop codon (TGA) were identified (Fig. 2). The deduced CebE protein (47.9 kDa) has 444 aa residues and contains the characterized internal peptides (determined by Edman degrada-

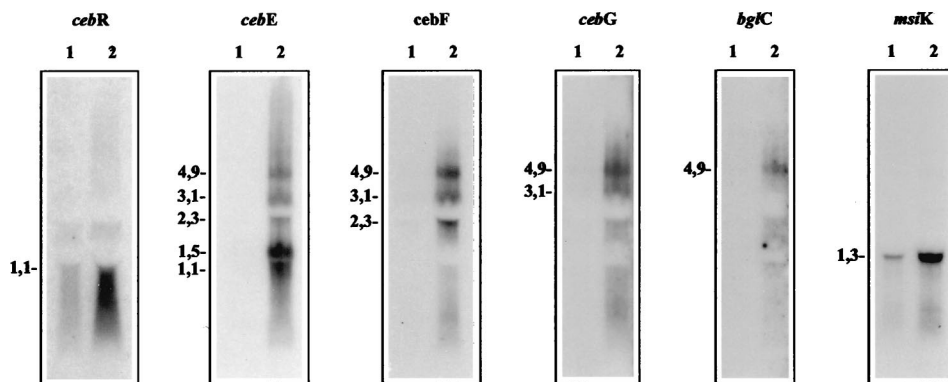


FIG. 4. Transcription studies of the *ceb* operon and *msiK* gene. Total RNA (15 μ g) from mycelia grown in the presence of glucose (lanes 1) and cellobiose (lanes 2) was separated electrophoretically. After transfer of the RNA to a nylon membrane, the strips were hybridized with the indicated ³²P-labelled probes and exposed to X-ray films. The X-ray film of the *cebR* Northern blot analysis was exposed eight times longer than the others. Transcript sizes (in kilobases) are given on the left.

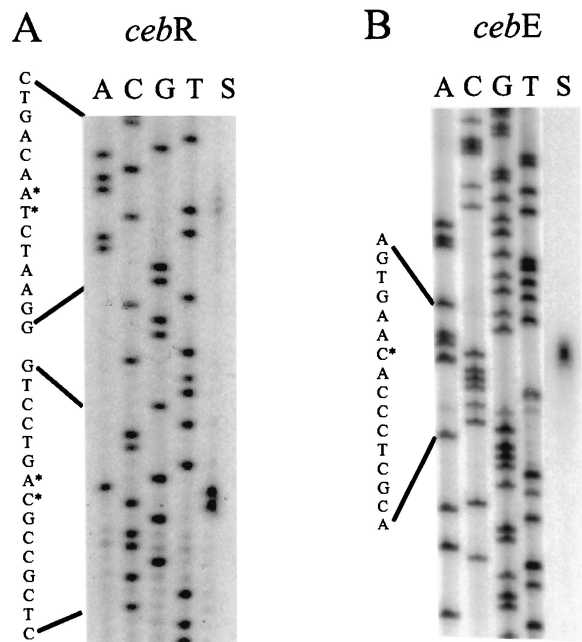


FIG. 5. Mapping of the transcription initiation sites of *cebE* and *cebR*. RNA (50 μ g) prepared from mycelia of *S. reticuli* grown in the presence of cellobiose was hybridized to 0.1 pmol of the 32 P-labelled *cebR* (A) or *cebE* (B) probes, and S1 nuclease treatment (lanes S) was done as described in Materials and Methods. ACGT indicates the *cebR* and *cebE* nucleotide sequence ladders. The asterisks mark the most probable transcription start sites.

tion) of the purified CebE protein. The first 26 aa residues of the deduced CebE include a positively charged NH₂-terminal hydrophobic core region and the sequence LLAGCA (the underscore indicates the cleavage site), which corresponds to the consensus (LLAGCS) of the lipoprotein signal peptidase cleavage site (38). This finding is in agreement with those of our previous biochemical experiments, which had revealed that *S. reticuli* produces CebE as a lipoprotein (34). Only 20% of the deduced amino acids of CebE are identical to those of binding proteins of cluster 1 as defined by Tam and Saier (39). The signature sequence (Fig. 3) comprising the highly conserved lysine residue is, however, conserved in CebE.

The *cebE* gene is followed by a putative transcription terminator and subsequently by the third (921-bp) and fourth (831-bp) reading frames. Neither reading frame is preceded by putative promoter and ribosome binding sites. The sizes of the deduced proteins for CebF and CebG are 276 and 306 aa, respectively, and both proteins contain six predicted membrane-spanning segments. Between the third and the fourth membrane-spanning segments, both proteins carry an EAA motif that matches the consensus EAAX₂DGAX₈IXLP sequence characteristic of membrane proteins from binding protein-dependent ABC transporters (31). In the *E. coli* MalF and MalG proteins the EAA motif has been identified as one site interacting with the predicted α -helical domain of the ATP binding protein MalK (21). Databank searches revealed that the deduced *S. reticuli* CebF and CebG proteins have highest identities with deduced lactose permeases from *Synechocystis* sp. (35% identity; EMBL accession no. P73352 and P73854) and deduced lactose permeases from *Thermus* sp. (36% identity; EMBL accession no. D1029300 and D1029301), all of which are presumed to be subunits of putative ABC transport systems.

The partially sequenced open reading frame following *cebG*

is preceded by a putative transcription terminator and a predicted ribosome binding site and encodes a portion (714 bp) of a protein (238 aa) named BglC. The sequence of the deduced NH₂ terminus is MPDSVSSLTFP (Fig. 2) and thus identical with the sequence of amino acids previously determined by Edman degradation for a purified *S. reticuli* intracellular β -glucosidase of 50 kDa (8). Alignments of the deduced BglC (calculated to correspond to about half of the purified β -glucosidase) have revealed that it has 72% (over 234 aa) and 61% (over 232 aa) identity with β -glucosidases deduced from *Streptomyces* sp. strain QM-B814 and *Microbispora bispora* nucleotide sequences, respectively.

The *ceb* operon lacks a gene encoding an ATP-hydrolyzing protein. Our previous experiments had shown that *S. reticuli* has a separately located *msiK* gene (35). In this context it is interesting that the *S. coelicolor* A3(2) *malEFG* operon also lacks a gene encoding an ATP binding protein (42). In the thermophilic archaea *Thermococcus litoralis* and *Thermococcus thermosulfurigenes*, the *malK* homologues are also not linked to ABC transport operons for maltose (28) and for maltose and trehalose (13), respectively. Inspection of complete genomic sequences revealed that the *msiK* homologues *msmX* from *Bacillus subtilis* (EMBL accession no. BG 11954) and *msiK* from *Synechocystis* sp. strain PCC 6803 (EMBL accession no. slr 0747) and an open reading frame from *Methanococcus jannaschii* encoding a homologue of the *E. coli* ATP binding protein UgpC of the *sn*-glycerol-3-phosphate ABC transporter (EMBL accession no. MJ0121) are located independently of the genes encoding other subunits of binding protein-dependent ABC transporters. Whether the above-cited MsiK homologues assist several ABC transporters as in *S. reticuli* remains to be elucidated.

Transcriptional experiments. Hybridization experiments with total RNA isolated from cellobiose-grown *S. reticuli* mycelia revealed the formation of a polycistronic transcript (4.9 kb) comprising *cebEFG* and *bglC*. Additionally, shorter transcripts corresponding to *cebEFG* (3.1 kb), *cebEF* (2.3 kb), and *cebE* (1.5 and 1 kb) were detected. The small transcript (1.5 kb) corresponding to the *cebE* gene attained up to 20-fold higher levels of transcription than the 4.9- and 3.1-kb transcripts. During cultivation with glucose, almost no *cebEFG* or *bglC* transcripts were found (Fig. 4). The large amount of the 1.5-kb *cebE* transcript corresponds to the high levels of CebE protein present within membranes of *S. reticuli* during cultivation on cellobiose (34). The size of the 1.5-kb *cebE* transcript

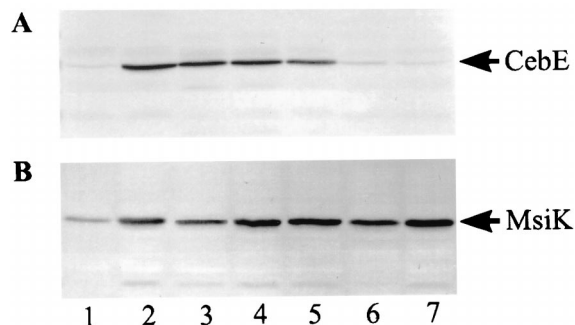


FIG. 6. Synthesis of CebE and MsiK proteins. Wild-type *S. reticuli* was grown in minimal medium with the following saccharide(s) at 0.5% (wt/vol) each: glucose (lane 1), cellobiose (lane 2), cellobiose plus glucose (lane 3), cellobiose plus maltose (lane 4), cellobiose plus sucrose (lane 5), maltose (lane 6), and sucrose (lane 7). Mycelia were disrupted by sonification, and 20 μ g of protein per lane was separated by SDS-PAGE (17). Immunodetection of CebE (A) and MsiK (B) proteins was performed as outlined previously (35).

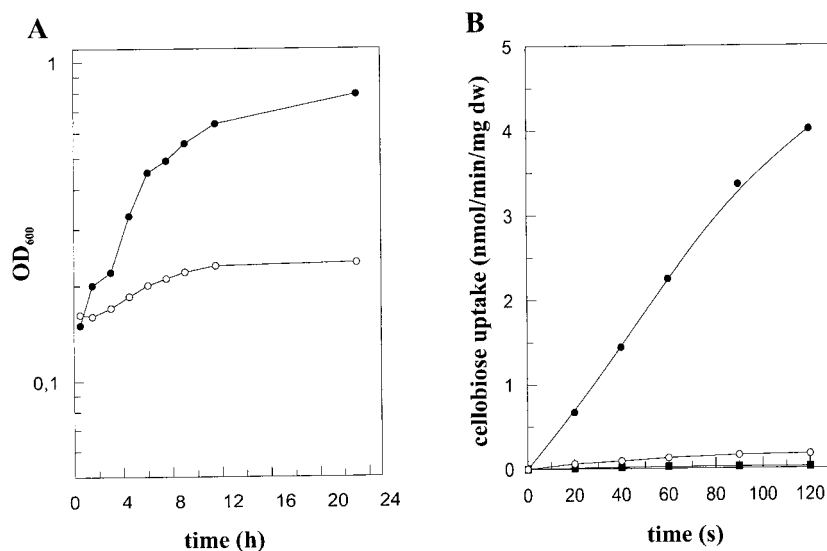


FIG. 7. Characterization of the *cebE* mutant. (A) Physiological experiments. Spores (10^6 /ml) of the wild type (●) and the *cebE* disruption mutant (○) were inoculated into minimal medium containing 0.5% (wt/vol) Casamino Acids and glucose or cellobiose. After 24 h, the mycelia were washed twice with minimal medium lacking Casamino Acids, and after the addition of cellobiose, the growth could be monitored photometrically (optical density at 600 nm [OD₆₀₀]), due to the finely dispersed hyphae. (B) Cellobiose uptake experiments. Mycelia from the *S. reticuli* wild type (●) or the *cebE* mutant (○) were grown in minimal medium containing Casamino Acids (0.5%) and cellobiose for 16 h, and then the mycelia were washed twice with 50 mM potassium phosphate buffer, pH 7.0. Uptake of [¹⁴C]cellobiose (5 μ M) was determined as described previously (34). As a control, the uptake of cellobiose was determined in wild-type (○) mycelia cultivated with glucose. dw, dry weight.

is in agreement with the distance from the *cebE* transcription start site (nucleotide 1244) (Fig. 2) to the predicted terminator following the stop codon of *cebE* (nucleotides 2749 to 2787). An additionally predicted terminator within the *cebE* gene corresponds to the 1.0-kb *cebE* transcript. However, a truncated CebE protein of the corresponding size (30 kDa) was not identified immunologically, suggesting that this small transcript is not translated under the conditions used or that it represents a degraded RNA form.

With the *cebR*-specific DNA probe, a monocistronic transcript of 1.1 kb was detected (Fig. 4). In glucose-grown *S. reticuli* mycelia, the level of *cebR* transcript was extremely low but it increased about 10 times during cultivation with cellobiose. A transcript of 1.3 kb corresponded to the monocistronic *msiK* gene. The quantity of the *msiK* transcript from cellobiose-grown cultures exceeded that from glucose-grown cultures by about 15-fold.

S1 nuclease mapping revealed that the *cebE* transcript starts 132 nucleotides 5' upstream of the translational start codon of *cebE*. The sequence GGAAC is located 24 nucleotides 5' upstream of the transcription start codon (Fig. 5). This motif matches the previously identified -35 region of the p2 promoter of the agarase gene *dagA* from *S. coelicolor* A3(2) (5, 37), a weak promoter of class E lacking a typical -10 region (4). The 5' upstream region of CebE contains the motif GGA GCGCTCC (Fig. 2), which has similarity to the optimal consensus sequence of the *E. coli* GalR operator [(G/T)AA(A/C)CGNTT(A/C)] (24, 46). Transcription of *cebR* starts at the T and G 21 nucleotides 5' upstream of the ATG start codon. Rarely used transcription start codons of *cebR* were found 22 nucleotides 5' upstream.

Production of CebE and MsiK. Antibodies raised against the C-terminal part of MsiK and the mature CebE (34) were used to monitor the levels of corresponding proteins during cultivation on minimal media containing different saccharides (Fig. 6). *S. reticuli* was found to produce CebE only if it was grown

in the presence of cellobiose (not with glucose, maltose, or sucrose). When, however, cellobiose and one of the above-mentioned saccharides were added together to the culture medium, the level of CebE attained was nearly that ascertained for mycelia grown only with cellobiose.

The quantity of MsiK was highest in mycelia grown with cellobiose, maltose, or sucrose. The addition of glucose to mycelia growing with cellobiose, maltose, or sucrose led to a decrease of MsiK synthesis. The relative amounts of *msiK* transcripts corresponded to the quantities of MsiK protein obtained from mycelia grown under comparable conditions (data not shown). These data show that MsiK synthesis is regulated differently from that of the CebE protein.

Construction of an *S. reticuli cebE* mutant. The *cebE* gene was inactivated by insertion mutagenesis. The *aphI* gene from Tn903 (25) was inserted into the pUC18-borne *cebE* gene with an internal deletion of an *NaeI* fragment (1,037 bp). After the *S. reticuli* wild type was transformed with the resulting plasmid, pCB41, several resistant colonies growing with 20 μ g of kanamycin per ml were found. Following a double crossover between the genomic *cebE* gene and the *cebE*-homologous flanking portions of the *aphI* gene on pCB41, the *aphI* gene was found to disrupt the reading frame of the *cebE* gene, which was confirmed by Southern hybridization. The mRNA of this kanamycin-resistant *S. reticuli* mutant lacks all transcripts corresponding to *cebEFG* and *bglC*, showing that the insertion of the *aphI* gene in *cebE* had an effect on the transcription of genes located downstream from it. As expected, the CebE protein was not detectable immunologically within cell extracts prepared from the *cebE* mutant. In minimal medium supplemented with glucose, the *cebE* mutant grew with the same doubling time as that of the wild-type strain (3 h). In cellobiose-containing minimal medium, the growth rate of the wild-type strain was not affected but the *cebE* mutant strain grew quite poorly (Fig. 7A). Unlike with the wild type, no or only very little cellobiose was taken up by mycelia of the *cebE*

mutant during short-term uptake experiments (Fig. 7B). However, after prolonged incubation (exceeding 10 min), small amounts of ^{14}C label were detected. Previously (8) we had shown that *S. reticuli* produces extracellular and intracellular β -glucosidase activities (detected by hydrolysis of *p*-nitrophenyl- β -D-glucopyranoside). Inspection of the *cebE* mutant revealed that the β -glucosidase activities were comparable to those of the progenitor strain (data not shown). Thus, we suspect that part of the ^{14}C -labelled cellobiose was cleaved to glucose.

Unlike with *S. reticuli*, within *E. coli* (26) and *Bacillus stearothermophilus* XL-65-6 (18) cellobiose is taken up via phosphoenolpyruvate-dependent phosphotransferase systems, and phosphorylated cellobiose, in turn, is cleaved by the action of intracellular phospho- β -glucosidases. In the cellulose degrader *Trichoderma reesei*, cellobiose uptake is mediated by a constitutively synthesized permease that is specific for different β -glucosides such as sophorose, gentiobiose, and cellobiose (16). The newly identified *S. reticuli* genes are so far the only known genes encoding a functional binding protein-dependent ABC transporter for cellobiose and cellotriose. Thus, this ABC transporter is an excellent model system for more-detailed studies to elucidate its high specificity.

ACKNOWLEDGMENTS

We are grateful to J. Ritz for training J. Jantos to isolate RNA, to T. Aldekamp for performing some of the Western blot analyses, and to R. Schmid, Department of Microbiology, University of Osnabrück, for determining NH_2 -terminal amino acids of CebE.

M. Lemme supported the writing of the manuscript. The work was initially financed by the SFB (grant 171/C14 to H. Schrempf) and then by the DFG (grant Schl 498/1-1 to A. Schlösser).

REFERENCES

- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.). 1987. Current protocols in molecular biology. Greene Publishing Associates, New York, N.Y.
- Bibb, M. J., P. R. Findlay, and M. W. Johnson. 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* 30:157-166.
- Boos, W., and J. M. Lucht. 1996. Periplasmic binding protein-dependent ABC transporters, p. 1175-1209. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
- Bourn, W. R., and B. Babb. 1995. Computer assisted identification and classification of streptomycete promoters. *Nucleic Acids Res.* 23:3696-3703.
- Buttner, M. J., I. M. Fearney, and M. J. Bibb. 1987. The agarase gene (*dagA*) of *Streptomyces coelicolor* A3(2): nucleotide sequence and transcriptional analysis. *Mol. Gen. Genet.* 209:101-109.
- Chromczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159.
- Dean, M., and R. Allikmets. 1995. Evolution of ATP-binding cassette transporter genes. *Curr. Opin. Genet. Dev.* 5:779-785.
- Heupel, C., A. Schlochtermeier, and H. Schrempf. 1993. Characterization of an intracellular β -glucosidase from *Streptomyces reticuli*. *Enzyme Microb. Technol.* 15:127-132.
- Higgins, C. F. 1992. ABC transporter: from microorganisms to man. *Annu. Rev. Cell Biol.* 8:67-113.
- Hoffman, K., and W. Stoffel. 1993. TMbase—a database of membrane spanning protein segments. *Biol. Chem.* 347:166.
- Höner zu Bentrup, K., R. Schmid, and E. Schneider. 1994. Maltose transport in *Aeromonas hydrophila*: purification, biochemical characterization and partial protein sequence analysis of a periplasmic maltose-binding protein. *Microbiol. J.* 140:945-951.
- Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf. 1985. Genetic manipulation of *Streptomyces*: a laboratory manual. John Innes Foundation, Norwich, United Kingdom.
- Horlacher, R., K. B. Xavier, H. Santos, J. DiRuggiero, M. Kossman, and W. Boos. 1998. Archaeal binding protein-dependent ABC transporter: molecular and biochemical analysis of the trehalose/maltose transport system of the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Bacteriol.* 180:680-689.
- Hsieh, M., P. Hensley, M. Brenowitz, and J. S. Fetrow. 1994. A molecular model of the inducer binding domain of the galactose repressor of *Escherichia coli*. *J. Biol. Chem.* 269:13825-13835.
- Hurtubise, Y., F. Shareck, D. Kluepfel, and R. Morosoli. 1995. A cellulase/xylanase-negative mutant of *Streptomyces lividans* 1326 defective in cellobiose and xylobiose uptake is mutated in a gene encoding a protein homologous to ATP-binding proteins. *Mol. Microbiol.* 17:367-377.
- Kubicek, C. P., R. Messner, F. Gruber, M. Mandels, and E. M. Kubicek-Pranz. 1993. Triggering of cellulase biosynthesis by cellulose in *Trichoderma reesei*. *J. Biol. Chem.* 268:19364-19368.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Lai, X., and L. O. Ingram. 1993. Cloning and sequencing of a cellobiose phosphotransferase system operon from *Bacillus stearothermophilus* XL-65-6 and functional expression in *Escherichia coli*. *J. Bacteriol.* 175:6441-6450.
- Lopilato, J. E., J. L. Garwin, S. D. Emr, T. J. Silhavy, and J. R. Beckwith. 1984. *D*-Ribose metabolism in *Escherichia coli* K-12: genetics, regulation, and transport. *J. Bacteriol.* 158:665-673.
- Miller, K. 1987. Gel electrophoresis of RNA. *Focus* 9:14-15.
- Mourez, M., M. Hofnung, and E. Dassa. 1997. Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits. *EMBO J.* 16:3066-3077.
- Mowbray, S. L., and L. B. Cole. 1992. 1.7 Å X-ray structure of the periplasmic ribose receptor from *Escherichia coli*. *J. Mol. Biol.* 225:155-175.
- Murray, M. 1986. Use of sodium trichloroacetate and mung bean nuclease to increase sensitivity and precision during transcript mapping. *Anal. Biochem.* 158:165-170.
- Nieto, C., M. Espinosa, and A. Puyet. 1997. The maltose/maltodextrin regulation of *Streptococcus pneumoniae*. *J. Biol. Chem.* 272:30860-30865.
- Oka, A., H. Sugisaki, and M. Takanami. 1981. Nucleotide sequence of the kanamycin resistance transposon *Tn903*. *J. Mol. Biol.* 147:217-226.
- Parker, L. L., and B. G. Hall. 1990. Characterization and nucleotide sequence of the cryptic *cel* operon of *Escherichia coli* K12. *Genetics* 124:455-471.
- Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biology sequence comparison. *Proc. Natl. Acad. Sci. USA* 85:2444-2448.
- Sahm, K., M. Matuschek, H. Müller, W. J. Mitchell, and H. Bahl. 1996. Molecular analysis of the *amy* gene locus of *Thermoanaerobacterium thermosulfurigenes* EM1 encoding starch-degrading enzymes and a binding protein-dependent maltose transport system. *J. Bacteriol.* 178:1039-1046.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
- Saurin, W., W. Köster, and E. Dassa. 1994. Bacterial binding protein-dependent permeases: characterization of distinctive signatures for functionally related integral cytoplasmic membrane proteins. *Mol. Microbiol.* 12:993-1004.
- Schlochtermeier, A., F. Niemeier, and H. Schrempf. 1992. Biochemical and electron microscopic studies of the *Streptomyces reticuli* cellulase (Avicelase) in its mycelium-associated and extracellular forms. *Appl. Environ. Microbiol.* 58:3240-3248.
- Schlochtermeier, A., S. Walter, J. Schröder, M. Moormann, and H. Schrempf. 1992. The gene encoding the cellulase (Avicelase) cell1 from *Streptomyces reticuli* and analysis of protein domains. *Mol. Microbiol.* 6:3611-3621.
- Schlösser, A., and H. Schrempf. 1996. A lipid-anchored binding protein is a component of an ATP-dependent cellobiose/triose transport system from the cellulose degrader *Streptomyces reticuli*. *Eur. J. Biochem.* 242:332-338.
- Schlösser, A., T. Kampers, and H. Schrempf. 1997. The *Streptomyces* ATP-binding component MsiK assists in cellobiose and maltose transport. *J. Bacteriol.* 179:2092-2095.
- Schumacher, M. A., K. Y. Choi, H. Zalkin, and R. G. Brennan. 1994. Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by α helices. *Nature* 266:763-770.
- Stroh, W. R. 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* 20:961-974.
- Sutcliffe, J. C., and R. R. B. Russell. 1995. Lipoproteins of gram-positive bacteria. *J. Bacteriol.* 177:1123-1128.
- Tam, R., and M. H. Saier. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* 57:320-346.
- Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* 76:4350-4354.
- van Wezel, G. P., J. White, P. Young, P. W. Postma, and M. J. Bibb. 1997. Substrate induction and glucose repression of maltose utilization by *Streptomyces coelicolor* A3(2) is controlled by *malR*, a member of the *lacI-galR*

- family of regulatory genes. *Mol. Microbiol.* **23**:537–549.
42. **van Wezel, G. P., J. White, M. J. Bibb, and P. W. Postma.** 1997. The *malEFG* gene cluster of *Streptomyces coelicolor* A3(2): characterization, disruption and transcriptional analysis. *Mol. Gen. Genet.* **254**:604–608.
 43. **Vyas, N. K., M. N. Vyas, and F. A. Quijcho.** 1991. Comparison of the periplasmic receptors for L-arabinose, D-glucose/D-galactose, and D-ribose. Structural and functional similarity. *J. Biol. Chem.* **266**:5226–5237.
 44. **Wachinger, G., K. Bronnenmeier, W. L. Staudenbauer, and H. Schrempf.** 1989. Identification of mycelium-associated cellulase from *Streptomyces reticuli*. *Appl. Environ. Microbiol.* **55**:2653–2657.
 45. **Walter, S., and H. Schrempf.** 1996. Physiological studies of cellulase (Avicelase) synthesis in *Streptomyces reticuli*. *Appl. Environ. Microbiol.* **62**:1065–1069.
 46. **Weickert, M. J., and S. Adhya.** 1992. A family of bacterial regulators homologous to Gal and Lac repressors. *J. Biol. Chem.* **267**:15869–15874.
 47. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103–119.