Expression of a Butyrivibrio fibrisolvens E14 gene (cinB) encoding an enzyme with cinnamoyl ester hydrolase activity is negatively regulated by the product of an adjacent gene (cinR)

Brian P. Dalrymple and Yolande Swadling

A second cinnamoyl ester hydrolase (CEH) encoding gene (cinB) has been characterized from the ruminal bacterium Butyrivibrio fibrisolvens E14. CinB is more similar to CinA (previously named Cinl) (28% amino acid identity), the first CEH described from B. fibrisolvens E14, than either of the enzymes are to any other member of the family of hydrolases to which they belong. Upstream of cinB, and in the opposite orientation, is a gene (cinR) encoding a protein with substantial similarity to members of the MarR family of negative regulators of bacterial gene expression. By alignment of these sequences, a possible helix-turn-helix DNA-binding domain has been identified. CinR was expressed at a high level in Escherichia coli using the lac promoter. In E. coli CinR repressed the expression of CinB, but had no effect on the expression of CinA. In gel mobility-shift assays, CinR bound specifically to the cinR-cinB intergenic region. Two identical 16 nucleotide inverted repeats adjacent to the putative PcinR and PcinB promoters are likely binding sites for CinR. The addition of FAXX [O-(S-5-O-(trans-feruloyl)-a-L-arabinofuranosyl)-(1,3)-O-β-D-xylopyranosyl-(1,4)-β-D-xylopyranose] and Fara [S-5-O-(trans-feruloyl)-arabinofuranose], but not xylobiose, ferulic acid and a number of other soluble components of hemicellulose, inhibited the binding of CinR to DNA.

Keywords: Butyrivibrio fibrisolvens, cinnamoyl ester hydrolase, negative regulator, signal sequence, xylan

INTRODUCTION

The heterogeneous polysaccharide xylan is one of the major components of plant cell walls. The arabinose side chains of arabinoxylan are sometimes esterified to cinnamic acids (principally ferulic and p-coumaric acid). A number of cinnamoyl ester hydrolases (CEHs) able to release cinnamic acids from plant materials have been identified (see Christov & Prior, 1993). The action of the CEH is probably required to allow complete utilization of the esterified arabinose and adjacent xylose residues.

Abbreviations: CAT, chloramphenicol acetyltransferase; CEH, cinnamoyl ester hydrolase; Fara, 5-O-(trans-feruloyl)-arabinofuranose; FAXX, O-[5-O-(trans-feruloyl)-a-L-arabinofuranosyl-(1,3)-O-β-D-xylopyranosyl-(1,4)-β-D-xylopyranose]; MUTMAC, 4-methylumbelliferoyl(p-trimethylammonium cinnamate chloride).

The GenBank accession number for the nucleotide sequence reported in this paper is U64802.

The cloning of a gene encoding a CEH from the ruminal bacterium Butyrivibrio fibrisolvens E14, using the model substrate MUTMAC [4-methylumbelliferoyl(p-trimethylammonium cinnamate chloride)] has been described (Dalrymple et al., 1996). Although predicted to be an intracellular enzyme, Cinl (here renamed CinA) exhibited significant cleavage of Fara [S-5-O-(trans-feruloyl)-arabinofuranose] derived from wheat bran, demonstrating the utility of MUTMAC for screening for enzymes with activity against plant-derived cinnamoyl esters. Analysis of the amino acid sequence of CinA demonstrated that it was a member of a large and diverse family of serine/aspartic acid/histidine active site hydrolases. However, the nature of the enzyme encoded by the second group of genomic clones isolated using MUTMAC was not investigated.

The regulation of expression of CEHs has been studied in the non-ruminal bacterium Streptomyces olivochro-
mogenes. Oat spelt xylan, de-starched wheat bran and sugar cane bagasse were demonstrated to induce between 45 and 172-fold increases in the specific activity of a CEH (MacKenzie et al., 1987; Johnson et al., 1988). In a number of non-ruminal fungi, CEH activity was not induced by growth on oat spelt xylan (does not contain cinnamic acid esters), but was induced by growth on meadow fescue grass (Smith et al., 1991). In contrast, CEH production in Aspergillus oryzae was not related to the phenolic acid content of the substrates (Tenkanen et al., 1991). In a separate series of experiments, growth of the fungus Schizophyllum commune on both Avicel and wheat bran produced more CEH activity than growth on oat spelt xylan (MacKenzie & Bilous, 1988). Clearly the regulation of expression of CEHs is complex and may vary substantially from one organism to another. The precise nature of the compounds responsible for inducing the expression of CEHs is not known.

In this paper we describe the sequence of cinB, a second CEH-encoding gene isolated from B. fibrisolvens E14, and the role of a putative negative regulator of cinB expression that is encoded adjacent to cinB. We have also determined the nature of molecules that appear to be involved in inducing cinB expression.

**METHODS**

**Bacterial strains and vectors.** Host strains Escherichia coli XL1-Blue, E. coli JM109 and E. coli KW251 were used for the cloning and expression of genes. Subcloning of fragments and expression of the recombinant proteins were carried out in plasmids pGEM-T, pGEM3Zf(+) and pGEM7Zf(+). Promoter-probe vector pKK232-8 was obtained from Pharmacia.

**Media and chemicals.** Bacterial cells were cultivated in L broth (LB) or LB containing 1.5% agar for agar plates and 0.7% agar for top agar (Sambrook et al., 1989). Ampicillin was used at 50 μg ml⁻¹, IPTG at 1 mM and X-Gal at 40 μg ml⁻¹. Chloramphenicol was used at various concentrations between 25 μg ml⁻¹ and 1 mg ml⁻¹ as required. Fara and FAXX [O-[5-O-(trans-feruloyl)-α-1,arabinofuranosyl]-1-(3,5-O-β-D-xylopyranosyl)-β-D-xylopyranosyl] were prepared as described previously (Dalrymple et al., 1996). MUTMAC, lactose, cellobiose, methylated pectin, α-D-galacturonic acid, p-coumaric acid and L-arabinose were obtained from Sigma. Ferulic acid and chlorogenic acid were obtained from Fluka and alduronic acid and xylobiose were obtained from Megazyme.

**Preparation and manipulation of DNA.** Bacteriophage particles and DNA were prepared from plate lysates using lambdasorb (Promega) following the manufacturer's recommended procedure. For subcloning, bacteriophage and plasmid DNA was digested with restriction enzymes and fragments were separated by agarose gel electrophoresis following standard techniques (Sambrook et al., 1989). Prior to ligation, DNA samples were purified using the Magic PCR Preps DNA purification system (Promega). Computer analyses of DNA and protein sequences were carried out using MacVector (IBI) and the ANGIS (Australian National Genomic Information Service) facility. Database searches were carried out using the BLAST program (Altschul & Lipman, 1990) at The National Centre for Biotechnology Information and the BLOCKS (Henikoff & Henikoff, 1994) and PROSITE (Bairoch, 1993) databases at ANGIS. The cinR gene was amplified by PCR using the primers shown in Fig. 2 for 35 cycles at 94 °C (60 s)/50 °C (60 s)/72 °C (30 s). The 50 μl reaction mix contained approximately 1 ng plasmid DNA (containing the complete cloned B. fibrisolvens fragment), 50 mM KCl, 10 mM Tris/HCl, pH 9, 0.1% Triton X-100, 2 mM MgCl₂, 0.1 mM each dNTP and 1 unit Taq polymerase. Prior to ligation into pGEM-T the PCR product was purified using the Magic PCR Preps DNA purification system (Promega).

**Cell fractionation and analysis.** E. coli cultures were grown to an OD₆₀₀ of 0.5–0.8 in LB and, if required, protein expression was induced by the addition of IPTG to 1 mM. Uninduced and induced cultures were harvested by centrifugation after a further 16 h incubation at 37 °C. Cells were resuspended in PBS (Sambrook et al., 1989) and lysed by sonication, or fractionated as described by Hsiung et al. (1986). Culture supernatants, sucrose fractions, periplasmic fractions and sonicated cells were assayed for enzyme activity by spotting samples at the appropriate dilution and overlaying with MUTMAC-containing agarose as previously described (Dalrymple et al., 1996). Proteins were separated by SDS-PAGE by the method of Laemmli (1970). Prestained low-range molecular mass standards were obtained from Bio-Rad. The concentration of protein in samples was determined using the Bio-Rad protein assay kit with bovine serum albumin standards.

The first ten amino-terminal residues of the recombinant CinB protein were determined by automated Edman sequencing with a 471A gas-phase sequenator (Applied Biosystems) using periplasmic fractions separated by SDS-PAGE and transferred to a PVDF membrane by Western blotting.

**Assays of CinR-regulated expression of CinB and CinA in E. coli.** High-titre stocks of bacteriophage genomic clones were diluted to obtain either plate lysates or well-separated plaques on lawns of E. coli KW251 cells containing the appropriate plasmid by using standard methods described by Sambrook et al. (1989), except that the top agarose contained 10% clarified and autoclaved rumen fluid (see Dalrymple et al., 1996). To recover proteins from the plate lysates, 3 ml of 100 mM NaCl, 17 mM MgCl₂, 50 mM Tris/HCl, pH 7.5, was poured onto the surface and the plates rocked at room temperature for 2–3 h. CEH activity was measured as described previously (Dalrymple et al., 1996). On plates with separated plaques, CinB activity was visualized using MUTMAC as described previously (Dalrymple et al., 1996).

**Gel mobility-shift assays.** The method described by Tupper et al. (1994) was modified as follows. Approximately 20–50 ng DNA was incubated for 15–30 min at room temperature with approximately 1 μg protein from sonicated and centrifuged lysates of E. coli JM109. Reaction samples were made to a final volume of 20 μl with a final concentration of 100 mM Tris/HCl, pH 8, 1 mM EDTA and 15 mM KCl. Samples were incubated for 5 min at room temperature with all of the components except the DNA and for a further 30 min after the addition of the DNA prior to electrophoresis. The total sample volume was separated on either 1% DNA-grade agarose (Progen), or 2% Metaphor agarose (FMC) gels in

Oligo 1000 DNA Synthesizer (Beckman).
45 mM Tris/borate, 1 mM EDTA at 5 V cm⁻¹. Gels were stained with ethidium bromide and illuminated on a UV transilluminator to visualize the DNA.

RESULTS AND DISCUSSION

Subcloning and deletion analysis of cinB

Subclones of the insert in the genomic clone #11, a representative of the second of the two groups of genomic clones isolated from a library of B. fibrisolvens E14 using MUTMAC (Dalrymple et al., 1996), were obtained by partial digestion of the bacteriophage DNA with HindIII followed by ligation into the HindIII site of the plasmid vector pGEM7Zf(+) . All of the subclones that were positive for cleavage of MUTMAC contained a 7.5 kb HindIII fragment containing both λ vector and B. fibrisolvens E14 DNA (Fig. 1). Deletion analysis of the B. fibrisolvens DNA delineated a region between the third EcoRV site and the BalI site that contained sequences involved in the expression, or coding, of CinB (Fig. 1). Enzyme expression was increased by the addition of IPTG when transcription of the lac promoter was from the HindIII to the EcoRI site (see Fig. 1), but not in the reverse direction (data not shown).

The sequence of the subcloned region was determined between the HindIII site and the second DraI site (Fig. 1). One long open reading frame (ORF), cinB, encoding a potential protein of 285 amino acids (predicted molecular mass 31449 Da) and consistent with the results from the deletion analysis, was identified between nucleotides 1182 and 2836 (Fig. 2). The first methionine is preceded by a probable ribosome-binding site (RBS) and is followed by a probable signal peptide with a predicted site of cleavage after amino acid 17. However, an adjacent site also shows similarity to the consensus sequence for the cleavage of lipoprotein signal sequences, L(S,A)A(A,G)C(S,G), with the cysteine at residue 19 predicted to be the site of addition of the lipid moiety (Sutcliffe & Russell, 1995). An inverted repeat downstream of cinB, between residues 2097 and 2143, may form a transcription termination signal.

The amino acid sequence of CinB exhibits 28% amino acid identity with that of CinA (Fig. 3). CinB is more similar in sequence to CinA than either of the enzymes that are to any of the other members of the large and diverse family of serine/aspartic acid/histidine active site hydrolases to which they both belong (see Dalrymple et al., 1996). However, the probable substrate-binding domains of CinA and CinB, between amino acids 177 and 210 in CinB [by analogy with XylF, described by Diaz & Timmis (1995)], are not similar in sequence. It is unlikely that cinA and cinB have arisen through a recent gene duplication and it is possible that their natural substrates are quite different.

High-level expression of CinB in E. coli

Preliminary experiments with plasmid 13BS in E. coli strain JM109 identified a protein of approximately 33 kDa in the periplasm fraction that was not present in control cells. Enzyme activity was enriched in the periplasmic fraction relative to the cell fraction. To increase the level of expression of CinB, the DNA sequence between the Ecl136I site downstream of the plasmid lac promoter and the third EcoRV site in the inserted B. fibrisolvens E14 DNA was deleted to generate plasmid 13BEE (Fig. 1). A significant increase in expression of CinB was observed, with a significant proportion of the protein again located in the periplasmic fraction (Fig. 4). However, induction of the cultures with IPTG led to substantial cell lysis during growth, with little additional yield of protein (data not shown). By direct sequencing, the amino-terminus of the recombinant CinB in the periplasmic fraction was determined to be MKSDYTVNTE.

The predicted amino acid sequence preceding the experimentally determined amino-terminus was not the sequence predicted from the location of the putative signal sequence. The sequence immediately preceding the MKSDYTVNTE sequence does not fit with consensus sequences for bacterial signal peptidase cleavage sites, suggesting that CinB was not exported via a signal-sequence-dependent pathway involving cleavage to generate the observed amino-terminus. Inspection of the DNA sequence preceding the methionine codon identified a possible RBS (Fig. 2). In E. coli CinB appears to be translated from the second possible AUG in the open reading frame. In B. fibrisolvens it is possible that the first start codon is utilized to produce a protein that is secreted via a signal-sequence-dependent pathway, or even that both intra- and extracellular forms of the enzyme may be produced from the one gene.

A putative negative regulator of cinB expression is encoded by the cinR gene immediately upstream of cinB

Located 170 bp upstream of cinB is an ORF on the opposite strand (Fig. 1). This ORF is preceded by a probable RBS and it encodes a protein of 142 amino acids with a predicted size of 16 kDa (Fig. 2). The sequence of the encoded protein was used to search all of the protein sequence databases using the BLAST algorithm. A number of proteins with significant similarities were identified (Fig. 5). PecS (Reverchon et al., 1994), Hpr (Perego & Hoch, 1988), HpcR (Roper et al., 1993), EmR (del Castillo et al., 1991) and MarR (Cohen et al., 1993) have all been shown to be negative regulators of bacterial gene expression. The B. fibrisolvens sequence is most similar to HpcR, a negative regulator of the homoprotocatechuate degradative operon of E. coli C (Roper et al., 1993), and PecS, a negative regulator of pectinase, cellulase and blue pigment production in Eruvima chrysanthemi (Reverchon et al., 1994), with 31% and 26% amino acid identity respectively.

A section of the LexA family of regulatory proteins (see Garriga et al., 1992) was also identified as similar to a small region of CinR and other members of the family. Additional searches carried out using the BLOCKS database identified regions of the DeoR (BL00894A), GntR...
(BL00043) and LuxR (BL00622) families of regulatory proteins as similar to small regions of CinR and other members of the family. These regions of similarities covered the proposed helix–turn–helix DNA-binding domains of the LexA, DeoR, GntR and LuxR families. The similarities suggest that amino acids 48 to 69 of CinR, and the equivalent amino acids in the other members of the family, may constitute a helix–turn–helix DNA-binding domain (Fig. 5). The spacing of the hydrophobic residues in this region, which is identical to that conserved in such domains from many proteins (see Suzuki et al., 1995), also supports this proposal.

In E. coli, expression of CinB, but not CinA, is reduced by the over-expression of CinR

In order to test the hypothesis that CinR regulates the expression of CinB, over-expression of CinR was required. Using PCR primers flanking cinR (see Fig. 2), the gene was amplified and inserted into the vector pGEM-T. Both orientations of insertion of cinR were obtained. The constructs had translation stop codons in all three frames preceding the proposed start codon and included only half of a potential operator site. The cinR gene in pCINR#7 was sequenced and no differences from the previously determined DNA sequence of the original cloned B. fibrisolvens E14 DNA were detected. A protein of less than 22 kDa was observed in extracts of E. coli JM109 carrying pCINR#7, but not in cell extracts carrying a plasmid with the gene containing a deletion inserted in the opposite orientation (pCINR#4) (Fig. 4). A small induction in expression of CinR was observed from pCINR#7 in samples with IPTG added (Fig. 4).

Both plasmids were transformed into E. coli strain K2251 and each cell line infected with bacteriophage genomic clones containing either cinA (#14) or cinB (#11) to give well-separated plaques. Overlay with MUTMAC was used to obtain a rough estimate of the level of esterase expression. All samples were positive for esterase activity, but the plaques that contained genes expressing CinB and CinR exhibited substantially less fluorescence. To obtain a more accurate assessment, plate lysates of the combinations of the two plasmids and the two recombinant bacteriophages were made and assayed for cleavage of FAXX. The presence of CinR had no significant effect on the expression of CinA, but reduced the expression of CinB approximately sixfold.

CinR binds to sequences between cinR and cinB

Extracts of E. coli JM109 containing plasmids pCINR#4 (CinR⁻) and #7 (CinR⁺) were mixed with HinII-digested plasmid 13E1 DNA. Addition of extract #4 had no effect on the mobility of any of the fragments. Addition of extract #7 decreased the mobility of only the fragment that contained the cinR–cinB intragenic region (Fig. 6). Thus, the mobility-shift was specific for the presence of CinR and for the cinR–cinB intragenic region of the plasmid. No EcoRV fragments exhibited a gel-shift with either extract (data not shown), suggesting that the 16 nucleotide inverted repeats that lie upstream of cinR and cinB contain EcoRV sites (Fig. 2) are the binding sites for CinR.

DNA binding has also been demonstrated for Hpr (Kallio et al., 1991), MarR (Martin & Rosner, 1995; Seoane & Levy, 1995), SlyA (Oscarsson et al., 1996) and PecS (Praill et al., 1996). The MarR-binding sites are also inverted repeats with the two five-base half repeats separated by two bases (Martin & Rosner, 1995). A similar inverted repeat lies between the promoter and start codon of SlyA (see Ludwig et al., 1995). In contrast, the Hpr-binding sites appear to be a somewhat shorter and imperfect inverted repeat (Kallio et al., 1991), whilst the Pec5-binding sites appear to be much longer (Praill et al., 1996). Our identification of a possible DNA-binding helix–turn–helix motif in CinR and the related
Regulation of cinnamoyl ester hydrolase expression

proteins suggests that all members of this family may bind directly to DNA via this region.

**Binding of CinR abolished by the addition of Fara and FAXX, but not by any other compound tested**

The effect of addition of various sugars, cinnamic acids, cinnamoyl esters and products of arabinofuranosyl degradation to the gel mobility-shift assay was investigated. The following compounds had no effect on controls, nor on the binding of CinR to DNA: 1 and 10 mM IPTG, 0.01-10 mM cellobiose, 0.002-0.2% methylated pectin, 0.01-10 mM \( \alpha \)-D-galacturonic acid, 1 mM ferulic or \( p \)-coumaric acid, 0.01-100 mM \( \alpha \)-D-arabinose, 0.05-5 mg chlorogenic acid [1,3,4,5-tetrahydroxy-cyclohexanecarboxylic acid 3-(3,4-dihydroxycinnamate)]. In contrast, for FAXX

![Fig. 2. Sequence of cinB and flanking regions. ORFs reading left to right are shown above the DNA sequence and ORFs reading right to left are shown below the DNA sequence. Proposed ribosome-binding sites (RBS), potential –35 and –10 promoter sequences, transcription terminators and probable CinR-binding sites are underlined and annotated. The experimentally determined amino-terminal amino acid sequence of CinB is also underlined. The location and orientation of primers used for the synthesis of the cinR PCR product are indicated by arrows.](image-url)
and Fara, whilst no effect was seen on controls, a small effect was seen at 30 μM and with concentrations greater than 300 μM binding of CinR was almost completely abolished (Fig. 6). Addition of extracts of *E. coli* containing CinB to the assay system reversed the inhibition of binding of CinR by FAXX and Fara (data not shown).

The members of this family of negative regulators are quite diverse in sequence and are involved in regulating a wide range of different processes. However, 4-hydroxyphenylacetic acid (HpcR) (Roper et al., 1995), salicylate (MarR and ErmR) (Martin et al., 1993), & dinitrophenol (ErmR) (Lomovskaya et al., 1995) and may be a common feature of this family of negative regulators.

**Identification of promoters in the cinR–cinB intergenic region**

The promoters of RNA transcription of cinB and cinR are expected to lie between the two genes. In the absence of information concerning the nature of promoters in *B. fibrisolvens*, the intergenic region was searched for sequences homologous to the *E. coli* sigma-70 promoter consensus sequence TTGACA (−35) and TATAAT (−10). Two overlapping and divergent candidate promoters were identified (Fig. 2). To test the ability of these sequences to direct transcription in *E. coli*, the 76 bp EcoRV fragment containing the two potential promoters was cloned into the *E. coli*-specific promoter-probe plasmid pKK232-8. With CinR driving the CAT (chloramphenicol acetyl-transferase) gene, colonies were formed on plates containing 1 mg chloramphenicol ml⁻¹. With PcinB transcribing the CAT (chloramphenicol acetyl-transferase) gene, colonies were formed on plates containing 1 mg chloramphenicol ml⁻¹. With no promoter driving the CAT gene no colonies were visible on 25 μg chloramphenicol ml⁻¹. A single copy of the 16 nucleotide inverted repeat to which CinR probably binds lies between each of the two proposed divergent promoters, PcinR and PcinB (Fig. 2), and the corresponding ORF.

**A hypothesis for the regulation of CinB expression by CinR**

The location of the probable CinR-binding sites suggests that the expression of both CinB and CinR is regulated by CinR. Autoregulation has been observed for MarR (Martin & Rosner, 1995) and ErmR (del Castillo et al., 1991) and may be a common feature of this family of negative regulators.
The binding of CinR to DNA is specifically inhibited by cinnamic acid sugar esters and not by cinnamic acids, sugars, or chlorogenic acid (a non-sugar cinnamic acid ester). Thus, both the sugar and the cinnamic acid moieties of the inducer are required for induction. It is probable that the inducers bind to CinR, thereby altering its conformation and reducing its ability to bind to DNA. The role of Fara and FAXX in induction suggests that the inducer of CinR may act as a regulator of its own expression, thereby creating a feedback loop that can enhance or attenuate the expression of the cinnamoyl ester hydrolase gene.
they are indeed among the natural substrates for CinB, and that CinB is likely to be involved in the breakdown of plant materials by B. fibrisolvens E14.

ACKNOWLEDGEMENTS

This project was supported in part by grants from the Australian Meat Research Corporation. We would like to thank D. H. Cybinski for help with assays and PAGE, C. McSweeney for useful discussions and R. Pearson and G. Riding for amino acid sequence determination.

REFERENCES