Novel cellulose-binding domains, NodB homologues and conserved modular architecture in xylanases from the aerobic soil bacteria *Pseudomonas fluorescens* subsp. *cellulosa* and *Cellvibrio mixtus*

Sarah J. MILLWARD-SADLER,* Keith DAVIDSON,† Geoffrey P. HAZLEWOOD,†‡ Gary W. BLACK,* Harry J. GILBERT* and Jonathan H. CLARKE†

*Department of Biological and Nutritional Sciences, Faculty of Agriculture, The University, Newcastle upon Tyne NE1 7RU, U.K., and †Department of Cellular Physiology, The Babraham Institute, Babraham, Cambridge CB2 4AT, U.K.

To test the hypothesis that selective pressure has led to the retention of cellulose-binding domains (CBDs) by hemicellulase enzymes from aerobic bacteria, four new xylanase (*xyn*) genes from two cellulolytic soil bacteria, *Pseudomonas fluorescens* subsp. *cellulosa* and *Cellvibrio mixtus*, have been isolated and sequenced. *Pseudomonas* genes *xynE* and *xynF* encoded modular xylanases (XYLE and XYLF) with predicted M_r values of 68 600 and 65 000 respectively. XYLE contained a glycosyl hydrolase family 11 catalytic domain at its N-terminus, followed by three other domains; the second of these exhibited sequence identity with NodB from rhizobia. The C-terminal domain (40 residues) exhibited significant sequence identity with a non-catalytic domain of previously unknown function, conserved in all the

cellulases and one of the hemicellulases previously characterized from the pseudomonad, and was shown to function as a CBD when fused to the reporter protein glutathione-S-transferase. XYLF contained a C-terminal glycosyl hydrolase family 10 catalytic domain and a novel CBD at its N-terminus. *C. mixtus* genes *xynA* and *xynB* exhibited substantial sequence identity with *xynE* and *xynF* respectively, and encoded modular xylanases with the same molecular architecture and, by inference, the same functional properties. In the absence of extensive cross-hybridization between other multiple *cel* (cellulase) and *xyn* genes from *P. fluorescens* subsp. *cellulosa* and genomic DNA from *C. mixtus*, similarity between the two pairs of xylanases may indicate a recent transfer of genes between the two bacteria.

INTRODUCTION

Heteropolysaccharides termed β -1,4-D-xylans, consisting of a backbone structure of β -1,4-linked xylopyranose units substituted with acetyl, arabinosyl and glucuronosyl residues, are the most abundant polymeric components of the hemicellulose fraction of plant cell walls. Enzymic hydrolysis of the xylans is effected by a consortium of microbial enzymes, of which the endoxylanases (xylanases) are the most abundant and best characterized components. It is well established that xylanolytic micro-organisms produce isoenzymic xylanases that are encoded by multigene families, and furthermore that there is extensive sequence conservation within the catalytic domains of xylanases, indicative of evolution from a small number of ancestral genes [1,2]. Saprophytic soil bacteria like *Pseudomonas fluorescens* subsp. *cellulosa* and *Cellvibrio mixtus* play a key role in natural decay and composting by recycling carbon contained in the structural polysaccharides of plant cell walls, and are prolific producers of cellulases and xylan-degrading enzymes [3,4].

Relatively little is known about the xylan-degrading enzymes of *C. mixtus*, but previous work has shown that two xylanases produced by *P. fluorescens* subsp. *cellulosa* are modular in structure and contain highly conserved type I cellulose-binding domains (CBDs) that function independently of the catalytic domain by promoting strong and specific binding to cellulose, but not to other cell-wall polysaccharides [5,6]. That such domains are present in enzymes that are devoid of activity against cellulose suggests that they are not directly involved in the catalytic event, and we have argued in favour of a more

general role for the CBD in recognition and binding of native substrates [7]. In *P. fluorescens* subsp. *cellulosa* there is no evidence for the formation of multiprotein cellulase-hemicellulase complexes, such as occur in *Clostridium thermocellum* and some other anaerobic bacteria, and we can perceive the logic of evolving a mechanism whereby recognition and binding of substrates is vested in each enzyme of the cellulase-hemicellulase system, rather than being a property of the complex as a whole. It is, however, unclear whether selective pressure has led to the retention of CBDs in all of the xylanases produced by aerobic bacteria such as *P. fluorescens* subsp. *cellulosa*, or whether modular xylanases lacking the binding domain also have a role as part of the repertoire of xylan-degrading enzymes. It is also unclear at this stage whether the novel NodB homologous domain recently described in XYLD from *Cellulomonas fimi* [7] is unique to this enzyme and this organism, or whether it is widely distributed among xylanases.

To address these two issues we have undertaken a comprehensive investigation of the structures of xylanases from two aerobic soil bacteria, one of which is already known to produce modular xylanases containing CBDs. Our purpose was to determine whether these organisms produce modular xylanases that do not contain CBDs, and to establish whether domains homologous with NodB from nitrogen-fixing rhizobia are a common feature of xylanases.

Data presented in this paper describe the molecular architecture of four new xylanases, two from *P. fluorescens* subsp. *cellulosa* and two from *C. mixtus*. Each of the enzymes is modular and contains a novel CBD that is separate from the

Abbreviations used: CBD, cellulose-binding domain; GST, glutathione-S-transferase; LB, Luria broth; MUC, methylumbelliferyl cellobioside; ORF, open reading frame.

‡ To whom correspondence should be addressed.

catalytic domain. Two of the enzymes contain a domain that is homologous with NodB, and a fourth domain of unknown function. The paper therefore supports the notion that there has been strong selective pressure for the retention of CBDs in xylanases from saprophytic soil bacteria.

MATERIALS AND METHODS

Bacterial strains, vectors and culture conditions

Cellvibrio mixtus subsp. *mixtus* (N.C.I.M.B. 8633, A.T.C.C. 12120) was cultured aerobically at 20 °C in Dubos minimal salts medium or on Dubos agar plates [Dubos minimal salts broth containing agar (15 g/l) supplemented with glucose (0.25%)]. *Pseudomonas fluorescens* subsp. *cellulosa* (N.C.I.M.B. 10462) was cultured at 37 °C in Luria broth (LB) supplemented with glucose (0.25%) or on plates containing the above medium solidified by the addition of agar (15 g/l). *Escherichia coli* strains XL1-Blue and JM83 were cultured in LB or on LB plates; the medium was supplemented with ampicillin (100 µg/ml) or 5-bromo-4-chloro-3-indolyl β-D-galactoside (2 µg/ml) to select for *E. coli* transformants or recombinants, respectively. Soluble xylan (0.25%) was added to LB plates to select for xylanase⁺ clones. The vectors used in this study were λZAPII (Stratagene), pBluescript (Stratagene), pNM52 [8], pMTL22p and pMTL23p [9], and pGEX-2T (Pharmacia).

Construction and screening of gene libraries

Genomic DNA was extracted from *C. mixtus* essentially as described by Marmur [10], and from *P. fluorescens* subsp. *cellulosa* by using an adaptation of the Brij lysis method [8]. Gene libraries were constructed in λZAPII for *C. mixtus* by using the approach described by Clarke et al. [11] and for *P. fluorescens* subsp. *cellulosa* as documented by Braithwaite et al. [12]. Ligated phage DNA was packaged *in vitro* and infected into *E. coli* XL1-Blue to generate ~10⁶ recombinant phages. For each organism, a non-amplified library containing more than three genomes of DNA was screened for xylanase-producing clones by plating out recombinant phage, together with the indicator strain *E. coli* XL1-Blue, in soft agarose poured on to NZY plates. The plaques formed were then overlaid with agar containing soluble xylan (0.25%). After a period of incubation at 37 °C, xylanase-producing clones were identified by staining with Congo Red [13]. DNA inserts contained in these clones were excised and recircularized to form recombinants of pBluescript as described in the Stratagene protocol.

General recombinant DNA procedures

Agarose-gel electrophoresis, Southern hybridization and transformation of *E. coli* were carried out as described by Sambrook et al. [14]. Plasmid DNA was prepared by alkaline lysis either by the method of Birnboim and Doly [15], or by using Qiagen resin columns (Hybaid). DNA-modifying enzymes were used in accordance with the manufacturer's instructions.

Nucleotide sequencing

Double-stranded plasmid DNA was sequenced by the dideoxy-chain-termination method [16], with the protocol recommended for the Sequenase DNA Sequencing Kit (United States Biochemical). Nested deletions of the cloned DNA fragments were created by digesting with Exonuclease III (ExoIII). Where necessary, synthetic oligonucleotide primers (17-mers) were designed according to known sequence, and synthesized on a Biosearch 8750 DNA Synthesizer. DNA sequences were com-

pared and ordered with either the GCG (Wisconsin) sequence-handling package [17], or the computer programs described by Staden [18].

Assays

Production of xylanase activity by recombinants of *E. coli* XL1-Blue was detected by overlaying LB plates on which the bacteria had been cultured with agar [1.5% (w/v)] containing soluble xylan from oat spelts [0.25% (w/v)], incubating at 37 °C and staining with Congo Red [13]. Quantitative evaluation of enzyme activities was carried out at 37 °C in PC buffer (50 mM K₂HPO₄, 12 mM citric acid, pH 6.5) [19]. Reducing sugar was measured with the dinitrosalicylic acid reagent [20]. Protein was measured by dye binding [21] or by the Lowry method with BSA as standard. Xylo-oligosaccharides (333 µg/ml) were incubated in a total reaction volume of 4.5 ml with 1 unit of XYLE (60 ng/ml) or 0.9 unit of XYLF (87 ng/ml). The enzyme concentrations are based on the specific activities of XYLE (330 units per mg of protein) and XYLF (230 units per mg of protein) purified by cellulose affinity chromatography. Aliquots (450 µl) were removed at regular time intervals, boiled for 5 min to inactivate the enzyme, and applied to a Dionex PA1 column. Separation of xylo-oligosaccharides was achieved by isocratic elution with 100 mM NaOH from 0 to 5 min, a gradient of 0 to 75 mM sodium acetate in 100 mM NaOH from 5 to 16 min, and isocratic elution with 75 mM sodium acetate in 100 mM NaOH from 16 to 25 min. Sugars were detected with a pulsed amperometer. Products released during the hydrolysis of native xylans by XYLE and XYLF were analysed by Dionex HPLC, and were identified by reference to standard compounds.

Polyacrylamide gel electrophoresis

SDS/PAGE was carried out essentially as described by Laemmli [22]. Fractionated polypeptides were stained with Coomassie Blue or were analysed for activity *in situ* [23,24]. Glutathione-S-transferase (GST) activity was measured using 1-chloro-2,4-dinitrobenzene (Sigma Chemical Co.) as substrate [25].

Cellulose- and xylan-binding studies

The binding of xylanases and GST-xylanase derivatives to cellulose and xylan was assessed by incubating cell-free extract containing the appropriate enzyme with Avicel pH105, acid-swollen cellulose (Sigmacell type 50 treated according to Wood [26]) or insoluble xylan in PC buffer, pH 6.5, for 1 h at 0 °C [11,27]. The incubations, in a volume of 1 ml, contained either 0.5 ml of undiluted cell-free extract (for experiments where polypeptides bound to the polysaccharides were analysed by SDS/PAGE), or appropriately diluted cell-free extract (where enzyme assays were performed to evaluate the capacity of enzymes to bind to the polysaccharides) and 0.5 ml of a 2% (w/v) suspension of the relevant polysaccharide in PC buffer. Appropriate dilutions of cell-free extracts (0.5 ml) contained one of: 0.1 unit of full-length XYLE (encoded by pJSE11; Figure 1), 0.2 unit of truncated XYLE (encoded by pJSE44; Figure 1) lacking the 40-residue C-terminal domain, 0.2 unit of full-length XYLF (encoded by pJSF11; Figure 1), 0.1 unit of GST or 0.1 unit of GST fused to the C-terminal domain of XYLE (GST-XYLE'). After incubation, the polysaccharides were washed five times with 1 ml of ice-cold PC buffer, by centrifuging at 12000 g for 5 min, and were either assayed directly for the retention of xylanase or GST activity, or boiled for 5 min in 10% (w/v) SDS, after which the material released into the supernatant was analysed by SDS/PAGE.

Figure 1 Restriction maps of xylanase genes *xynE* (a) and *xynF* (c) from *P. fluorescens* subsp. *cellulosa*, and *xynA* (b) and *xynB* (d) from *C. mixtus*

Restriction enzyme sites are as follows: Ac, *AccI*; Be, *BstEII*; Bs, *BssHII*; Bx, *BstXI*; Cl, *Clal*; Dr, *DraII*; Er, *EcoRI*; Ev, *EcoRV*; Hc, *HincII*; Hf, *HinfI*; Kp, *KpnI*; Na, *NaeI*; Ps, *PstI*; Pv, *PvuII*; Sa, *SacI*; Sm, *SmaI*; Su, *SstI*; Sy, *SylI*. The locations of the signal peptides (vertical hatching), catalytic domains (diagonal hatching), NodB domains (stippling), cellulose-binding domains (horizontal hatching), domains of unknown function (white) and linker sequences (black), within the translated polypeptide sequences are shown. Truncated versions of each *xyn* gene, the extents of which are depicted beneath the restriction maps, were generated by making 3' and 5' deletions (exonuclease III/S1 nuclease) or by using known restriction sites. Where 5' deletions were made, the truncated gene was subcloned in-frame with the *lacZ* translational start in either pBluescript or pMTL22p. The capacity of each deleted gene to encode a functional xylanase is indicated with + or -.

RESULTS

Isolation of xylanase genes *xynA* and *xynB* from *C. mixtus* and *xynE* and *xynF* from *P. fluorescens* subsp. *cellulosa*

Xylanase-positive phages were isolated at a frequency of 1/120 and 1/200 from the *Cellvibrio* and *Pseudomonas* genomic libraries, respectively. The genomic inserts from these recombinant clones were excised into pBluescript SK⁻ and cross-hybridization experiments were performed to determine the number of distinct xylanase genes isolated from the two libraries. The results (not shown) revealed two different *C. mixtus* genes, designated *xynA* and *xynB*, while two new *Pseudomonas* xylanase genes, classified as *xynE* and *xynF*, were identified. Southern hybridization of *C. mixtus* and *P. fluorescens* subsp. *cellulosa* genomic DNA, with *xynA/xynB* and *xynE/xynF* respectively as probes, revealed that the four xylanase genes were present as single copies. Restriction maps of the four genes are shown in Figure 1.

Properties of xylanases A (XYLA) and B (XYLB) from *C. mixtus* and E (XYLE) and F (XYLF) from *P. fluorescens* subsp. *cellulosa*

The capacity of the four enzymes encoded by *xynA*, *xynB*, *xynE* and *xynF*, designated XYLA, XYLB, XYLE and XYLF, respectively, to hydrolyse a range of plant structural polysaccharides was evaluated. None of the enzymes attacked mannan, galactan, arabinan or any cellulosic substrates. In contrast, all four hydrolysed oat spelt xylan and the arabinoxylans from wheat and rye, respectively. The major products generated from both xylan and arabinoxylan by XYLE and XYLF were xylobiose and xylotriose, respectively. No arabinose was detected. All four enzymes hydrolysed xylan in a manner typical of an endoxylanase, with a rapid decrease in substrate viscosity accompanying the production of reducing end groups (results not shown). To evaluate whether the recombinant xylanases exhibited a preference for long-chain xylose polymers, xylo-oligosaccharides were incubated with XYLE and XYLF, and the rate of substrate cleavage and the nature of the products generated were determined. The results (Figure 2) showed that both enzymes exhibited a preference for xylopentaose, and displayed little activity against xylotriose. In general the products released from these xylo-oligosaccharides were consistent with the endo- mode of action of the enzymes. For example, xylopentaose was hydrolysed primarily to xylotriose and xylobiose, and xylotetraose was broken down to xylobiose, whereas xylose and xylobiose were very slowly released from xylotriose. Interestingly, during the early stages of xylotetraose and xylopentaose hydrolysis by XYLE, significant quantities of xylotriose and xylotetraose were released, respectively, without the associated production of xylose.

A possible explanation for the appearance of the triose and the tetraose is that XYLE mediates transglycosylation reactions generating long-chain oligosaccharides that are subsequently broken down to smaller oligosaccharides that accumulate transiently (xylotetraose), or for a prolonged period (xylotriose), owing to their relatively slow rate of hydrolysis. The transglycosylating activity of XYLE suggests that the enzyme hydrolyses glycosidic bonds by a double-displacement mechanism, and is therefore a retaining xylanase. XYLB and XYLF both hydrolysed methylumbelliferyl cellobioside (MUC), suggesting that their catalytic domains belong to glycosyl hydrolase family 10. In contrast, the lack of activity of XYLA and XYLF against MUC indicates that these enzymes probably belong to glycosyl hydrolase family 11. On the basis of the biochemical properties of

Figure 2 Hydrolysis of xylo-oligosaccharides by XYLE and XYL F

XYLE (a) and XYL F (b) were incubated with xylotriose (1), xylotetraose (2) or xylopentose (3) as described in the Materials and methods section. The productions of xylose (★), xylobiose (▲), xylotriose (▽), xylotetraose (◇) and xylopentose (■) was determined by HPLC analysis as described in the Materials and methods section. The variation of substrate hydrolysis with time is displayed as a percentage of the substrate present at the start of the reaction.

XYLA, XYLB, XYLE and XYL F, the four enzymes appear to be typical endo- β -1,4-xylanases.

Nucleotide sequences of *xynA*, *xynB*, *xynE* and *xynF*

The nucleotide sequences of the genomic inserts containing *xynA*, *xynB*, *xynE* and *xynF* were determined in both strands and revealed single open reading frames (ORFs) of 1968, 1863, 1983 and 1818 bp, respectively, encoding proteins with predicted M_r values of 68668, 64930, 68600 and 65000, respectively (Figure 3). Evidence supporting the notion that the four ORFs encode XYLA, XYLB, XYLE and XYL F is as follows. (i) Each putative ATG translational start codon was preceded by a

sequence similar to prokaryotic ribosome-binding sequences. (ii) Translational stop codons were observed in all three forward-reading frames immediately upstream and downstream of the four ORFs. (iii) The deduced N-terminal sequence of each enzyme displayed features typical of prokaryotic signal peptides; the positively charged N-terminus of each of the four xylanases was followed by regions containing mainly small hydrophobic residues capable of forming α -helical structures. The predicted cleavage sites for the putative signal peptides are between residues 23 and 24, 19 and 20, 25 and 26, and 19 and 20 for XYLA, XYLB, XYLE and XYL F, respectively. (iv) The M_r values of mature XYLA and XYLB, determined by zymogram analysis of the enzymes after SDS/PAGE, were 72000 and 65000, re-

The proposed ribosome-binding sites are shown by a dotted underline; linker sequences are boxed and the signal peptides are underlined. The 5' and 3' ends of the gene truncations depicted in Figure 1 are shown by black triangles.

Figure 4 Binding of XYLE, XYLF and their truncated derivatives to cellulose

(a) Cell-free extracts made from *E. coli* strains expressing XYLE (lane 1), XYLF (lane 2), truncated XYLF (XYLFtr) encoded by pJSF22 (lane 3), GST-XYLE, GST or from *E. coli* harbouring pBluescript alone (lane 4) were incubated with Avicel or acid-swollen cellulose and, after repeated washing of the polysaccharide, bound polypeptides were eluted in denaturing buffer and analysed by SDS/PAGE. (b) The polypeptides recovered from Avicel that had been incubated with XYLE (lane 1), XYLF (lane 2), XYLFtr (lane 3) or cell-free extract from *E. coli* alone (lane 4). (c) The polypeptides recovered from Avicel (i) or from acid-swollen cellulose (ii) that had been incubated with cell-free extract containing GST (lanes 1) or GST-XYLE (lanes 2). The relative molecular masses of cellulose-bound proteins are shown at the right.

spectively (results not shown), whereas XYLE and XYLF purified by cellulose affinity chromatography had M_r s of 72000 and 65000, respectively (Figure 4). Discrepancies between the calculated molecular masses and the M_r values observed by SDS/PAGE, such as those seen here for XYLA and XYLE, have been noted for other polysaccharide hydrolases, and are probably

attributable to the acidic nature of these enzymes [28,29].

Palindromic sequences of DNA capable of forming stem-loop structures with ΔG^0 values of -164 and -105 kJ, respectively, were noted downstream of *xynA* and *xynB*. Although these regions are not entirely characteristic of *rho*-independent transcription termination sequences, as they are not followed by

<i>Cellvibrio mixtus</i> XYLA	N N A T G T H N G F Y Y T F W K D S G D A - S M G L Q A G G R Y T S Q W S N G T N N W V - - - - - 74
<i>Pseudomonas fluorescens</i> subsp. <i>cellulosa</i> XYLE	S N S T G T N N G F Y Y T F W K D S G D A - S M T L L S G G R Y Q S S W G N S T N N W V - - - - - 74
<i>Cellulomonas fimi</i> XYLD	S N T T G T H D G Y F Y S F W T D S P G S V S M D L N S G G G Y T - - R W S N T G N F V - - - - - 87
<i>Clostridium acetobutylicum</i> XynB	S N E T G V N G G Y D Y E L W K D Y G N T S M T L K N - G G A F S C Q W S N I G N A L F - - - - - R107
<i>Bacillus circulans</i> XlnA	S A T A S A A S T D Y W Q N W T D G G I V N A V N G S G G N Y S V N W S N T G N - - F - - - - V 62
<i>Cellvibrio mixtus</i> XYLA	G G K G W N P G - G P K V V T - Y S G S Y N V D - - N S Q N S Y L A L Y G W T R S P L I E Y Y V I 119
<i>Pseudomonas fluorescens</i> subsp. <i>cellulosa</i> XYLE	G G K G W N P G N N S R V I S - Y S G S Y G V D - - S S Q N S Y L A L Y G W T R S P L I E Y Y V I 120
<i>Cellulomonas fimi</i> XYLD	A G K G W S T G - G R K T V S - Y S G Q F N - - - - P S R N A Y L T L Y G W T Q S P L V E Y Y I V 130
<i>Clostridium acetobutylicum</i> XynB	K G K K F N D T Q T Y K Q L G N T S V N Y D C N - Y Q P Y G N S Y L C V Y G W T S S P L V E Y Y I V 156
<i>Bacillus circulans</i> XlnA	V G K G W T T G S P F R - - - - T I N Y N A G V W A P N G N G Y L T L Y G W T R S P L I E Y Y V V 107
<i>Cellvibrio mixtus</i> XYLA	E S Y G S Y N P A S C S G G T D Y G S F Q S D G A T Y N V R R C Q R V Q Q P S I D G - T Q T F Y Q Y 168
<i>Pseudomonas fluorescens</i> subsp. <i>cellulosa</i> XYLE	E S Y G S Y N P A S C S G G T D Y G S F Q S D G A T Y N V R R C Q R V N Q P S I D G - T Q T F Y Q Y 169
<i>Cellulomonas fimi</i> XYLD	D S W G T Y R P - - - T G T F M G T V T S D G G T Y D I Y R T Q R V N K P S I E G D S S T F Y Q Y 176
<i>Clostridium acetobutylicum</i> XynB	D S W G S W R P - - - G G T S K G T I T V D G G T Y D I Y E T T R I N Q P S I Q G N - T T F K Q Y 202
<i>Bacillus circulans</i> XlnA	D S W G T Y R P T - - - G T Y - K G T V K S D G G T Y D I Y T T T R V N A P S I D G D R T T F T Q Y 153
<i>Cellvibrio mixtus</i> XYLA	F S V R S P K K G F G Q I S G T I T A N H F N F W A S K G L N L G N - H D Y M V L A T E G Y Q S R 217
<i>Pseudomonas fluorescens</i> subsp. <i>cellulosa</i> XYLE	F S V R N P K K G F G N I S G T I T F A N H V N F W A S K G L N L G N - H N Y Q V L A T E G Y Q S R 218
<i>Cellulomonas fimi</i> XYLD	W S V R Q Q K R T - - - G G T I T S G N H F D A W A S K G M N L G R - H N Y M I M A T E G Y Q S S 221
<i>Clostridium acetobutylicum</i> XynB	W S V R R T K R T S G - - - T I S V S K H F A W E S K G M P L G K M H E - T A F N T E G Y Q S S 247
<i>Bacillus circulans</i> XlnA	W S V R Q S K R P T G S - N A T I T F T N H V N A W K S H G M N L G S N W A Y Q V M A T E G Y Q S S 202
<i>Cellvibrio mixtus</i> XYLA	G S S D I T V S 225
<i>Pseudomonas fluorescens</i> subsp. <i>cellulosa</i> XYLE	G S S D I T V S 226
<i>Cellulomonas fimi</i> XYLD	G S S I T V S 229
<i>Clostridium acetobutylicum</i> XynB	G K A D V - N S 254
<i>Bacillus circulans</i> XlnA	G S S N V T V W 210

Figure 5 Comparison of the amino acid sequence (between residues 32 and 226) of XYLE from *P. fluorescens* subsp. *cellulosa* with the sequence of XYLA from *C. mixtus* and the sequences of the catalytic domains of *Cellulomonas fimi* XYLD [7], *Clostridium acetobutylicum* XynB [35] and *Bacillus circulans* XynA [36]

Regions containing identical residues or conservative substitutions are boxed. The numbers denote the positions of the residues within their respective full-length sequences.

Figure 6 Alignment of the amino acid sequences of the putative NodB domains of XYLE from *P. fluorescens* subsp. *cellulosa* and XYLA from *C. mixtus* with the sequence of the NodB domain of XYLD from *C. fimi* [7] and with the sequences of NodB proteins from *Rhizobium trifolii* [37], *Rhizobium meliloti* [38], *Rhizobium leguminosarum* [39] and *Bradyrhizobium* [40]

poly(T) regions, it is tempting to speculate that they play some role in terminating transcription. Inspection of the nucleotide sequences of the four genes revealed significant similarity between *xynA* and *xynE* (72%) and between *xynB* and *xynF* (64%) suggesting a close evolutionary relationship between the homologous DNA sequences. Codon utilization for the four genes was very similar, with a marked preference for G and C in the wobble position and little use made of the codons GGG, GGA, CGA, AGA, ATA and GTA. Overall, the codon bias was very similar to that seen previously in other *P. fluorescens* subsp. *cellulosa* genes encoding plant cell-wall hydrolases.

Similarity of XYLA, XYLB, XYLE and XYLF to other plant cell-wall hydrolases

Translation of the gene sequences revealed that XYLA and XYLE exhibited 75% identity (82% similarity) at the amino acid level, whereas XYLB and XYLF displayed 63% identity (77% similarity). Comparison of the sequences of these enzymes with sequences contained in the SWISSPROT database revealed evidence for a number of features in XYLA, XYLB, XYLE and XYLF that are also present in other plant cell-wall hydrolases. Residues 32–225 of *C. mixtus* XYLA and 32–226 *P. fluorescens* subsp. *cellulosa* XYLE displayed significant sequence identity to the catalytic domains of xylanases belonging to glycosyl hydrolase family 11 (Figure 5). Downstream of the catalytic domain of XYLA was a 120-residue sequence that exhibited 73% identity (80% similarity) with the corresponding region of XYLE but bore little similarity to any protein present in the SWISSPROT

database. Further downstream of this region in both XYLA and XYLE was a third domain, consisting of more than 140 residues, that was similar to an internal domain of *Cellulomonas fimi* XYLD and with the NodB proteins of *Rhizobium* spp. (Figure 6). The C-terminal regions of XYLA and XYLE, comprising about 40 residues, were closely similar and exhibited significant sequence identity with the middle domain described previously in two endoglucanases, a cellodextrinase and an endoxylanase from *P. fluorescens* subsp. *cellulosa* (Figure 7). Overall, these six sequences displayed 47% identity. Domains 1 and 2, and 3 and 4 of XYLA and XYLE, were joined by serine-rich linkers of the type seen previously in modular *Pseudomonas* cellulases and hemicellulases; a putative linker sequence rich in serine and glycine residues joined domains 2 and 3 (Figure 3). To summarize, comparison with known xylanase sequences indicates that both XYLA and XYLE are modular enzymes consisting of four discrete domains, of which the N-terminal region constitutes a glycosyl hydrolase family 11 catalytic domain.

A similar sequence analysis conducted with XYLB and XYLF indicated that both enzymes were composed of two domains. The putative signal peptide was separated by an extensive serine-rich linker sequence of more than 70 residues from a domain of more than 150 residues which, although similar in both XYLB and XYLF (59% identity; 75% similarity), exhibited no similarity to other proteins in the Protein Identification Resource (PIR), SWISSPROT, EMBL and Brookhaven databases. Towards the C-terminus from this domain in both XYLB and XYLF was a second domain of more than 350 residues, which displayed significant sequence identity with the catalytic domains

Figure 7 Comparison of the amino acid sequence (between residues 616 and 653) of XYLE from *P. fluorescens* subsp. *cellulosa* with the C-terminal region of XYLA from *C. mixtus* and with the conserved non-catalytic middle domain of endoglucanases A [41] and B [42], cellodextrinase C [43] and xylanase A [44] from *P. fluorescens* subsp. *cellulosa*

Figure 8 Comparison of the amino acid sequence (between residues 296 and 499) of XYL F from *P. fluorescens* subsp. *cellulosa* with the sequence of XYL B from *C. mixtus* (this work) and the sequences of the catalytic domains of *Thermoanaerobacterium saccharolyticum* XYLA [45], *Cellulomonas fimi* Cex [46] and *Clostridium thermocellum* XynZ [47]

of glycosyl hydrolase family 10 enzymes (Figure 8). Between the first domain and the putative catalytic domain of XYL B was a small serine-rich linker sequence (Figure 3).

Structure/function analysis of XYL A, XYL B, XYL E and XYL F

To verify the structures inferred by comparative sequence analysis, truncated derivatives of each of the four enzymes were generated by deleting the appropriate regions of *xynA*, *xynB*, *xynE* and *xynF*. The results, summarized in Figure 1, showed that removal of only 155 bp and 286 bp from the 5' ends of *xynA* and *xynE*, respectively, resulted in the generation of truncated genes (pERA218 for *xynA* and pJSE22 for *xynE*) which, when cloned in-frame with *lacZ'* of appropriate vectors, did not direct the synthesis of functional xylanases. In contrast, removal of up to 1242 bp (pEE91) and 1141 bp (pJSE33), respectively, from the 3' ends of *xynA* and *xynE* generated truncated genes that still encoded functional xylanases. These results confirm that the catalytic domains of XYLA and XYLE are located within the N-terminal regions that exhibit identity with glycosyl hydrolase family 11 enzymes.

To evaluate the role of the 40-residue C-terminal non-catalytic domains of XYLA and XYLE, the capacity of the latter enzyme to bind to cellulose was tested. The results, presented in Figure 4, showed that *E. coli* harbouring pJSE11, which contains a full-length copy of *xynE*, expressed a protein of M_r 72000 that was not a major component of the cell-free extract but could be concentrated significantly by binding to cellulose. This polypeptide was not expressed by *E. coli* containing pBluescript SK⁻ and was closely similar in size to mature XYLE (M_r 69000).

Furthermore, both Avicel and amorphous cellulose that had been incubated with cell-free extracts derived from *E. coli* containing pJSE11 retained 0.1 and 0.09 unit of xylanase activity, respectively, even after repeated vigorous washing with buffer. No xylanase was detected in the final washings of these polysaccharides. In contrast, *E. coli* harbouring a truncated form of XYLE, lacking the 40-residue C-terminal domain and part of the NodB sequence (pJSE44), did not produce an obvious polypeptide that bound to cellulose. Similarly, neither crystalline nor amorphous cellulose, when incubated with *E. coli* cell-free extract containing this XYLE derivative, retained significant xylanase activity. These data suggest that the small C-terminal domain of XYLE constitutes a CBD. To determine whether this C-terminal 40-residue domain can function as a CBD, independently of the rest of the enzyme, the 3' 140 bp of *xynE* were amplified by PCR and fused to *gex* (encodes GST). When transformed into *E. coli* the hybrid gene directed the synthesis of a polypeptide of M_r 32000 (GST-XYLE) that exhibited GST activity. When incubated with crystalline and amorphous cellulose, 0.05 unit and 0.04 unit of GST activity, respectively, were retained by the polysaccharide, as were polypeptides of M_r 32000. No GST activity was detected in the final washings of these polysaccharides. In contrast, when GST alone was incubated with the two forms of cellulose, no GST activity or specific polypeptides were retained by either of the polysaccharides (Figure 4). These data confirm that the small C-terminal domain of XYLE, and by inference that of XYL A also, constitutes an independent CBD.

To determine whether the C-terminal regions of XYLB and XYLF constitute catalytic domains, the capacity of derivatives of *xynB* and *xynF* to encode functional xylanases was evaluated.

The results, presented in Figure 1, showed that truncated forms of *xynB*, lacking 333 bp from the 3' end (p58EXA19), or 1311 bp from the 5' end (p58ERA71), did not encode a functional xylanase. However, deletion of 884 bp from the 5' end of *xynB* (p58ERA22) did not affect the capacity of the gene to direct the synthesis of a functional xylanase. A derivative (pJSF22) of *xynF* that lacked 360 bp of 3' sequence failed to express an active xylanase. These data support the view that the C-terminal domains of XYLB and XYLF, previously shown to exhibit sequence identity with glycosyl hydrolase family 10 enzymes, constitute functional catalytic domains. To determine whether the N-terminal domains of XYLB and XYLF are CBDs, the capacity of full-length XYLF to bind to cellulose was evaluated. The results, presented in Figure 4, show that a polypeptide of M_r 65000, which was contained in cell-free extract from *E. coli* expressing full-length XYLF (pJSF11), was concentrated significantly by binding to cellulose. Avicel, which had been incubated with the same *E. coli* cell-free extract, retained 0.1 unit of xylanase activity. *E. coli* harbouring a truncated form of *xynF* from which 360 bp of 3' sequence had been deleted (pJSF22, Figure 1), expressed a protein of M_r 55000 (XYLFtr) that bound to cellulose (Figure 4) but did not exhibit xylanase activity. *E. coli* cells lacking the recombinant plasmids failed to produce either of the cellulose-binding polypeptides with M_r values of 65000 and 55000, respectively. These data indicate that the N-terminal domain of XYLF, and by inference that of XYLB, constitutes a functional CBD.

It was recently shown that xylanase D from *Cellulomonas fimi* exhibits cellulose-binding capacity, and also contains a novel xylan-binding domain [27]. To determine whether XYLE and XYLF also bind to xylan, cell-free extracts from *E. coli* containing full-length *xynE* (pJSE11) and *xynF* (pJSF11) were incubated with insoluble xylan and, after washing, the polysaccharide was assessed for bound xylanase using a combination of SDS/PAGE and enzyme assay. The results (not shown) revealed that XYLE and XYLF were unable to bind to insoluble xylan and therefore did not contain xylan-binding domains.

DISCUSSION

Evidence presented here shows that two new xylanases from the aerobic soil bacterium *Pseudomonas fluorescens* subsp. *cellulosa* contain functional CBDs. In an earlier study [3], only 50 % of the extracellular xylanase activity produced by this organism was capable of associating with cellulose, suggesting either that the pseudomonad expresses some xylanases that do not contain CBDs, or that the *Pseudomonas* xylanases are subject to extensive proteolytic processing, resulting in the generation of enzymes that lack their endogenous CBDs. The former possibility now appears unlikely as the screening of an extensive gene library comprising 200000 recombinant clones did not reveal further novel xylanase genes. However, it remains a possibility that some *Pseudomonas* xylanase genes, encoding enzymes devoid of CBDs, were not expressed in *E. coli*, even though the gene library screened was an expressing library. Derivation of the structures of the two *Pseudomonas* xylanases XYLE and XYLF revealed that the two CBDs had not been previously identified in plant cell-wall hydrolases from this organism. In the case of XYLF the sequence of the CBD was not found in any of the protein databases searched, indicating that the enzyme XYLF contains a completely novel CBD. The CBD located in XYLE exhibited significant sequence identity with a domain also found in two endoglucanases (EGA and EGB), a cellodextrinase (CELC) and a xylanase (XYLA) from *P. fluorescens* subsp. *cellulosa*.

Interestingly, each of these four enzymes also contained an additional high-affinity type I CBD.

Assuming that the homologues of the XYLE CBD are functional in all of the plant cell-wall hydrolases, it suggests that *Pseudomonas* EGA, EGB, CELC and XYLA contain two distinct domains that bind to cellulose. Although endoglucanase B (CenB) from *Cellulomonas fimi* has been shown to contain multiple CBDs, this paper provides the first indication that single xylanases can also contain more than one type of CBD. In addition, this paper demonstrates for the first time that different xylanases from a single micro-organism can contain different CBDs. The basis for the evolution of hemicellulases containing diverse CBDs could be the different ligand specificities of these domains; for example, some domains may exhibit high affinity for crystalline cellulose, whereas others bind tightly only to amorphous cellulose. Alternatively, some CBDs may exhibit either high or low affinity for all forms of the polysaccharide. Support for this view is provided by other studies [30,31], which have shown that the two CBDs of CenB actually exhibit high and low affinity, respectively, for cellulose. In similar fashion, we have shown that the GST-CBD fusion protein containing the small C-terminal CBD from XYLE does not bind quantitatively to cellulose (results not shown), whereas the type I CBDs located in *Pseudomonas* cellulases and hemicellulases interact so tightly with the polysaccharide that they can be removed only by treatment with 5 % SDS at high temperature. In summary, data presented here support the notion that the 40-residue non-catalytic middle domain conserved in EGA, EGB, CELC and XYLA from *P. fluorescens* subsp. *cellulosa*, and now also shown to be present at the C-terminus of XYLE from the pseudomonad and a xylanase (XYLA) from *C. mixtus*, is a low-affinity CBD by virtue of its inability to bind quantitatively to cellulose. Thus EGA, EGB, CELC and XYLA from *Pseudomonas* contain both high-affinity (type I) and low-affinity CBDs.

In addition to the identification of novel CBDs, this paper has shown that NodB homologues are prevalent among xylanases from different soil bacteria. In *Rhizobium*, NodB deacetylates the *N*-acetylglucosamine residue at the non-reducing end of a range of chito-oligosaccharides, which are then acylated and function as signals initiating nodulation of the plant root [32]. The function of the NodB homologues in xylanases remains to be elucidated; we have found no evidence to suggest that the xylanases described in this report are capable of deacetylating chito-oligosaccharides or *p*-nitrophenyl acetate (results not shown). However, the high degree of conservation of the NodB domain in xylanases from three different bacteria suggests that it plays an important role in enzyme function.

Inspection of the molecular architecture of the four xylanases described here revealed that XYLA (*C. mixtus*) and XYLE (*P. fluorescens*) contain glycosyl hydrolase family 11 catalytic domains, whereas XYLB (*C. mixtus*), XYLF (*P. fluorescens*) and the two previously described *P. fluorescens* subsp. *cellulosa* xylanases (XYLA and XYLB) belong to family 10. It is therefore apparent that the multiple xylanases of *P. fluorescens* subsp. *cellulosa* and *C. mixtus* resulted from the transfer of at least two discrete xylanase genes into the two bacteria. In view of the fact that a third soil bacterium, *Cellulomonas fimi*, and an anaerobic fungus, *Neocallimastix patriciarum*, are also hosts to xylanases belonging to families 10 and 11, it seems that the evolution of multiple xylanases by the transfer of several discrete xylanase genes is a common phenomenon in both eukaryotic and prokaryotic micro-organisms.

As with most modular plant cell-wall hydrolases, XYLA, XYLB, XYLE and XYLF contained linker sequences between their different functional domains. Surprisingly, there were also

extensive linker sequences between the respective signal peptides and catalytic domains of XYLB and XYLF. There is no obvious rationale for the evolution of these N-terminal linker sequences, as the signal peptide should be removed during passage through the bacterial cell membrane. However, enzymes that are linked to the outer surface of cellular membranes via an uncleaved signal peptide often contain linker sequences between the signal peptide and the rest of the protein. For example, glucan-glucosylhydrolase D of *P. fluorescens*, which is membrane-associated in *E. coli* at least, contains a small linker sequence between its signal peptide and catalytic domain [33]. The location of linker sequences adjacent to the signal peptide in XYLB and XYLF may also indicate that these enzymes associate with bacterial membranes via an uncleaved signal peptide. However, this hypothesis must be viewed with some caution because the two enzymes do not seem to be associated with the *E. coli* cell envelope, and cellular location of these enzymes in their endogenous hosts remains to be elucidated.

Data presented in this report indicate extensive sequence identity between XYLA from *C. mixtus* and XYLE from *P. fluorescens* and between XYLB from *C. mixtus* and XYLF from *P. fluorescens*. The corresponding genes also exhibit extensive sequence identity and display the same bias in codon utilization. The presence of homologous genes in *C. mixtus* and *P. fluorescens* subsp. *cellulosa* suggests that these two soil bacteria are closely related. Support for this view is provided by morphological similarities; furthermore, both bacteria are Gram-negative rods and exhibit similar plant cell-wall degrading properties. The similarity between *Cellvibrio* and *Pseudomonas* was recognized in the 8th edition of Bergey's Manual of Determinative Bacteriology (1974) in which the genus *Cellvibrio* was relegated to a species of *Pseudomonas*, only to be revived subsequently as a separate genus [34]. It is apparent that there are major differences between *P. fluorescens* subsp. *cellulosa* and *C. mixtus*; the latter strain has a temperature optimum for growth of 25 °C whereas the pseudomonad exhibits optimal growth at 37 °C. In addition, probing of *C. mixtus* genomic DNA with the previously cloned pseudomonad cellulase and hemicellulase genes (*xynA*, *xynB*, *xynC*, *xynD*, *celA*, *celB*, *celC* and *celD*) did not reveal any homologous sequences, indicating that, in general, cellulase and hemicellulase genes found in *P. fluorescens* subsp. *cellulosa* do not occur in *C. mixtus* also. Thus it is possible that the similar xylanase genes reported here in *Cellvibrio* and *Pseudomonas* could be the result of recent lateral transfer of genes between the two bacteria. However, the possibility cannot be totally dismissed that the occurrence of homologous xylanase genes in the two prokaryotes is the result of a close evolutionary relationship between the two organisms.

To conclude, this paper confirms that modular xylanases are common in a range of different aerobic bacteria. The molecular architecture of these enzymes is very variable, with some enzymes containing more than one type of CBD, in addition to other non-catalytic domains that, although highly conserved, have no defined function as yet. Future work in our laboratories will focus on the role of these domains in the context of plant cell-wall hydrolysis.

We thank the Biotechnology and Biological Sciences Research Council for supporting this work (LE13/138) and Judith Laurie for skilled technical assistance.

REFERENCES

- Gilbert, H. J. and Hazlewood, G. P. (1993) *J. Gen. Microbiol.* **139**, 187–194
- Coughlan, M. P. and Hazlewood, G. P. (1993) *Biotechnol. Appl. Biochem.* **17**, 259–289
- Hazlewood, G. P., Laurie, J. I., Ferreira, L. M. A. and Gilbert, H. J. (1992) *J. Appl. Bact.* **72**, 244–251
- Wynne, E. C. and Pemberton, J. M. (1986) *Appl. Environ. Microbiol.* **52**, 1362–1367
- Ferreira, L. M. A., Durrant, A. J., Hall, J., Hazlewood, G. P. and Gilbert, H. J. (1990) *Biochem. J.* **269**, 261–264
- Kellett, L. E., Poole, D. M., Ferreira, L. M. A., Durrant, A. J., Hazlewood, G. P. and Gilbert, H. J. (1990) *Biochem. J.* **272**, 369–376
- Millward-Sadler, S. J., Poole, D. M., Henrissat, B., Hazlewood, G. P., Clarke, J. H. and Gilbert, H. J. (1994) *Mol. Microbiol.* **11**, 375–382
- Gilbert, H. J., Blazek, R., Bullman, H. M. S. and Minton, N. P. (1986) *J. Gen. Microbiol.* **132**, 151–160
- Chambers, S. P., Prior, S. E., Barstow, D. A. and Minton, N. P. (1988) *Gene* **68**, 139–149
- Marmur, J. (1961) *J. Mol. Biol.* **3**, 208–218
- Clarke, J. H., Laurie, J. I., Gilbert, H. J. and Hazlewood, G. P. (1991) *FEMS Microbiol. Lett.* **83**, 305–310
- Braithwaite, K. L., Black, G. W., Hazlewood, G. P., Ali, B. R. S. and Gilbert, H. J. (1995) *Biochem. J.* **305**, 1005–1010
- Teather, R. M. and Wood, P. J. (1982) *Appl. Environ. Microbiol.* **43**, 777–780
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (eds.) (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Birnboim, H. C. and Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Devereux, J., Haeblerli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
- Staden, R. (1980) *Nucleic Acids Res.* **16**, 3673–3694
- Hazlewood, G. P., Davidson, K., Laurie, J. I., Romaniec, M. P. M. and Gilbert, H. J. (1990) *J. Gen. Microbiol.* **136**, 2089–2097
- Miller, G. L. (1959) *Anal. Chem.* **31**, 426–428
- Sedmak, J. J. and Grossberg, J. R. (1977) *Anal. Biochem.* **79**, 544–552
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Schwarz, W. H., Bronnenmeier, K., Gräbner, F. and Staudenbauer, W. L. (1987) *Anal. Biochem.* **164**, 72–77
- Malburg, L. M. and Forsberg, C. W. (1993) *Can. J. Microbiol.* **39**, 882–891
- Habig, W. H., Pabst, M. J. and Jakoby, W. B. (1974) *J. Biol. Chem.* **249**, 7130–7139
- Wood, T. M. (1971) *Biochem. J.* **121**, 353–362
- Black, G. W., Hazlewood, G. P., Millward-Sadler, S. J., Laurie, J. I. and Gilbert, H. J. (1995) *Biochem. J.* **307**, 151–155
- Coutinho, J. B., Moser, B., Kilburn, D. G., Warren, R. A. J. and Miller, R. C., Jr. (1991) *Mol. Microbiol.* **5**, 1221–1233
- Coutinho, J. B., Gilkes, N. R., Warren, R. A. J., Kilburn, D. G. and Miller, R. C., Jr. (1992) *Mol. Microbiol.* **6**, 1243–1252
- Meinke, A., Braun, C., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr. and Warren, R. A. J. (1991) *J. Bact.* **173**, 308–314
- Meinke, A., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr. and Warren, R. A. J. (1991) *J. Bact.* **173**, 7126–7135
- John, M., Rohrig, H., Schmidt, J., Wieke, U. and Schell, J. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 625–629
- Rixon, J. E., Ferreira, L. M. A., Durrant, A. J., Laurie, J. I., Hazlewood, G. P. and Gilbert, H. J. (1992) *Biochem. J.* **285**, 947–955
- Blackall, L. L., Hayward, A. C. and Sly, L. I. (1985) *J. Appl. Bact.* **59**, 81–97
- Zappe, H., Jones, W. A. and Woods, D. R. (1990) *Nucleic Acids Res.* **18**, 2179
- Yang, R. C. A., MacKenzie, C. R., Bilous, D., Seligy, V. L. and Narang, S. A. (1988) *Appl. Environ. Microbiol.* **54**, 1023–1029
- Schofield, P. R. and Watson, J. M. (1986) *Nucleic Acids Res.* **14**, 2891–2903
- Torok, I., Kondorosi, E., Stepkowski, T., Posfai, J. and Kondorosi, A. (1984) *Nucleic Acids Res.* **12**, 9509–9524
- Rossen, L., Johnston, A. W. B. and Downie, J. A. (1984) *Nucleic Acids Res.* **12**, 9497–9508
- Scott, K. R. (1986) *Nucleic Acids Res.* **14**, 2905–2919
- Hall, J. and Gilbert, H. J. (1988) *Mol. Gen. Genet.* **213**, 112–117
- Gilbert, H. J., Hall, J., Hazlewood, G. P. and Ferreira, L. M. A. (1990) *Mol. Microbiol.* **4**, 759–767
- Ferreira, L. M. A., Hazlewood, G. P., Barker, P. J. and Gilbert, H. J. (1991) *Biochem. J.* **279**, 793–799
- Hall, J., Hazlewood, G. P., Huskisson, N. S., Durrant, A. J. and Gilbert, H. J. (1989) *Mol. Microbiol.* **3**, 1211–1219
- Lee, Y.-E., Lowe, S. E., Henrissat, B. and Zeikus, J. G. (1993) *J. Bacteriol.* **175**, 5890–5898
- O'Neill, G. P., Goh, S. H., Warren, R. A. J., Kilburn, D. G. and Miller, R. C., Jr. (1986) *Gene* **44**, 325–330
- Grépinet, O., Chebrou, M.-C. and Béguin, P. (1988) *J. Bacteriol.* **170**, 4582–4588