Extremophiles (1998) 2:53–60

REVIEW

Roy H. Doi · Jae-Seon Park · Chi-chi Liu
Laercio M. Malburg Jr. · Yutaka Tamaru
Akihiko Ichiishi · Atef Ibrahim

Cellulosome and noncellulosomal cellulases of *Clostridium cellulovorans*

Received: November 18, 1997 / Accepted: November 26, 1997

Abstract This paper reviews the properties of the cellulosome and noncellulosome cellulases produced by *Clostridium cellulovorans*, an anaerobic, mesophilic, spore-forming microorganism that produces copious amounts of cellulase. The three major subunits of the cellulosome, CbpA, exoglucanase S (ExgS), and P100, are described, as well as the properties of the functional domains of CbpA. The properties of two noncellulosomal endoglucanases, EngD and EngF, are compared. The functions of the cellulose-binding domain (CBD) of CbpA indicate its potential uses in biotechnology.

Key words Cellulase · Cellulosome · Scaffolding protein · Endoglucanase · Exoglucanase

Introduction

Much knowledge has been obtained about the cellulosome and noncellulosomal cellulases from the thermophilic *Clostridium thermocellum* (Beguin 1996), but less is known about the anaerobic, mesophilic *Clostridia* species and their cellulase activities (Doi et al. 1992, 1994; Gal et al. 1997). However, the general picture that is emerging about the *Clostridia* is that the cellulosome (Lamed and Bayer 1988) appears to be a common feature of all these organisms that are capable of degrading cellulose. The major feature of the cellulosome is the presence of a large, nonenzymatic scaffolding protein to which several enzyme subunits (perhaps as many as nine different subunits) are attached to form a stable, complex enzyme capable of degrading crystalline cellulose at a rapid rate. In this review we will emphasize the features of the cellulosome and noncellulosomal cellulases from the anaerobic mesophile, *Clostridium cellulovorans*.

Although the overall features of the cellulosome are reasonably clear, there are several properties of the cellulosome that require further investigation. These can be addressed by several questions concerning its properties and functions: How do the multimeric subunits assemble extracellularly to form the complex cellulosome structure? How is the expression of the multiple genes for the cellulosomal subunits regulated? How does the cellulosome degrade crystalline cellulose? What is the relationship between the cellulosome and the noncellulosomal cellulases? These questions pose problems that will have to be solved in a multidisciplinary fashion, since they concern protein–protein interactions, gene regulation, enzyme–solid substrate interactions, and synergistic actions of multiple enzymes.

We have approached this complex problem initially by studying the properties of the subunits of the *C. cellulovorans* cellulosome. This has included an analysis of the scaffolding protein, CbpA, the enzymes that bind to CbpA, and the properties of genetically engineered mini-cellulosomes. We are following these studies with an analysis of the regulation of expression of the genes for the cellulosomal subunits in order to understand the role of the cellulosome under different growth conditions and the synergistic interactions between the cellulosome and noncellulosomal cellulases.

The *C. cellulovorans* cellulosome

The *C. cellulovorans* cellulosome, with a mass of about 10^6 daltons, contains three major subunits with MWs of approximately 189000, 100000, and 70000 (Shoseyov and Doi 1990). The 189-kDa protein is the scaffolding protein CbpA (Shoseyov et al. 1992), the 100-kDa protein (P100) is an endoglucanase (Malburg, Liu, Matano, and Doi, unpub-

Communicated by K. Horikoshi

R.H. Doi (✉) · J-S Park · C-C Liu · L.M. Malburg Jr. · Y. Tamaru · A. Ichiishi · A. Ibrahim
Section of Molecular and Cellular Biology, University of California, Davis, California, USA
Tel. +1−530-752-3191; Fax +1−530-752-3085
e-mail: rhdoi@ucdavis.edu

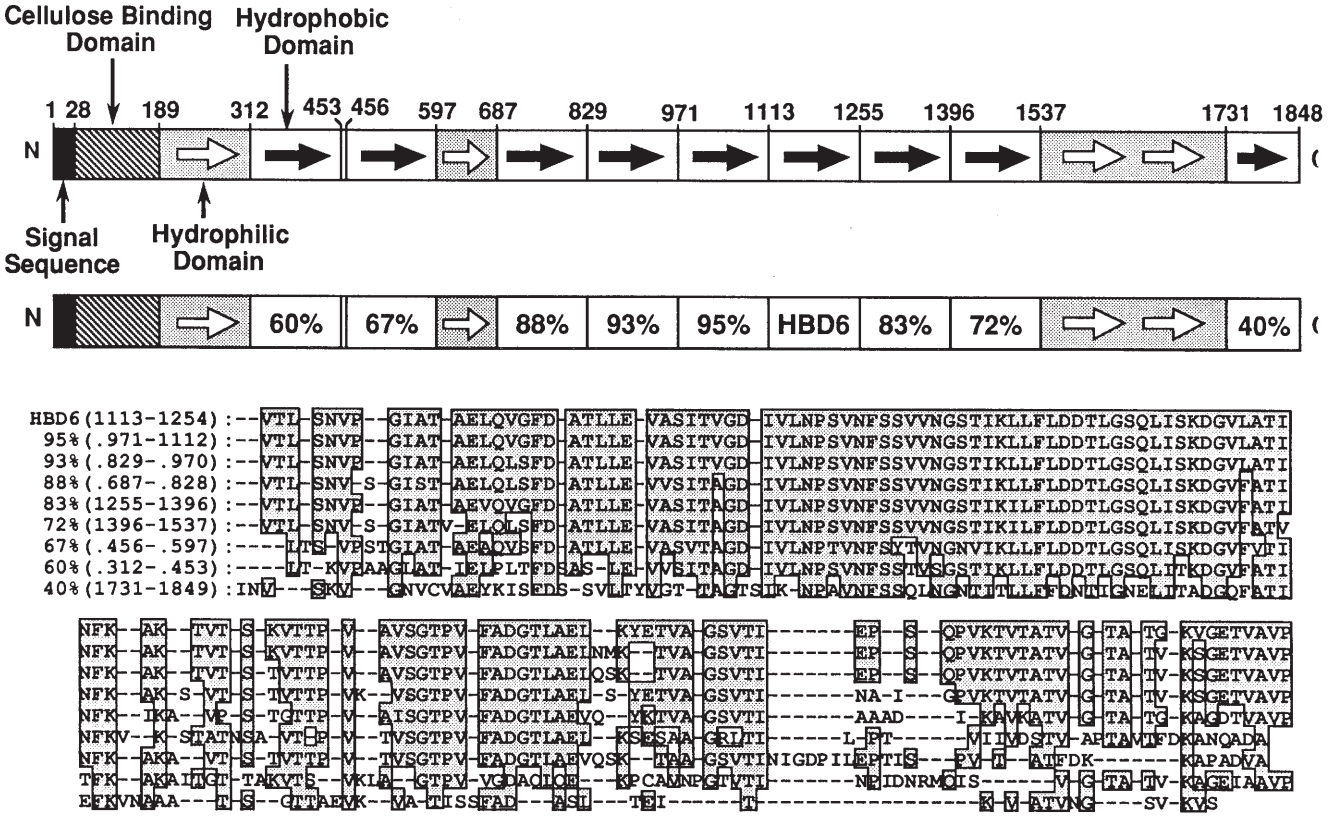

Fig. 1. The scaffolding protein, CbpA, of *Clostridium cellulovorans* and the homology between its hydrophobic domains (*HBDs*). The *top* of the figure shows the domains of CbpA including the nine hydrophobic domains (*black arrows*), the hydrophilic domains (*white arrows*), and the cellulose-binding domain (CBD) at the N-terminus. The *middle portion* of the figure indicates in per cent the homology between the hydrophobic domains. The *bottom part* of the figure shows the amino acid sequence and the high degree of identity (*shaded boxes*) between the hydrophobic domains

lished observations), and the 70-kDa protein (ExgS) is an exoglucanase (Liu and Doi, submitted for publication). The ratio of the subunits CbpA:P100:ExgS is approximately 1:2:3 and this composition is somewhat dependent on the particular preparation. Zymograms of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) preparations of the cellulosome subunits indicate the presence of up to 7–8 carboxymethylcellulase (CMCase) activity bands (Shoseyov and Doi 1990). Significantly, the CbpA band shows no enzyme activity. The cellulosome binds tenaciously to crystalline cellulose through its cellulose-binding domain (CBD) (Goldstein et al. 1993), and this feature can be used in the purification of cellulosomes.

The scaffolding protein CbpA

The interesting feature of the *C. cellulovorans* scaffolding protein, CbpA, is illustrated in Fig. 1 (Shoseyov et al. 1990). The sequence analysis of the gene for CbpA has revealed the presence of a signal peptide, a cellulose-binding domain (CBD), four hydrophilic domains (HLDs), and nine hydrophobic domains (HBDs or EBDs). The EBDs are the binding sites for the enzymatic subunits of the cellulosome (Takagi et al. 1993; Pages et al. 1996).

The analysis of the *C. cellulovorans* EBDs indicates the high identity that exists between the EBDs, ranging from 40% to 95% when compared with EBD6 (HBD6) (Fig. 1). The N-terminal halves of the EBDs (EBD1) are more conserved than the C-terminal halves (EBD2), and the N-terminal halves of EBD1 (EBD1A) are more conserved than the C-terminal halves of EBD1 (EBD1B) (Fig. 2). When CbpA is compared to the CipA from *C. thermocellum*, there is much less identity between the HBDs from these two scaffolding proteins, ranging from about 25% to 36% (Fig. 3). Thus, although the functions of CbpA and CipA are similar, they appear to accomplish their roles by using different EBD sequences.

When supernatants of acid-swollen cellulose (ASC) or cellobiose cultures are treated with 50 mg/ml of Avicel, and gently agitated at room temperature, nearly 95% binding of CbpA to the solid substrate is achieved within 10 min. Interestingly, CbpA undergoes fragmentation upon binding to Avicel, as revealed by a ladder of faint, immunologically cross-reacting bands which appeared in Western

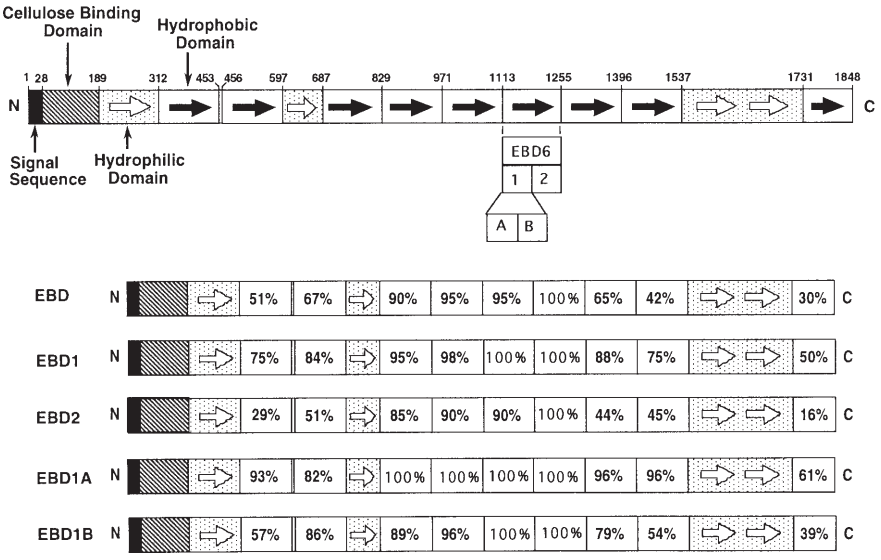

Fig. 2. Comparison of the hydrophobic domains (HBD or EBD) of the CbpA. *Top*, the nine hydrophobic domains (*dark arrows*); the hydrophobic domains are numbered in sequence from the N- to the C-terminus as EBD1 to EBD9. *Bottom*, the sequence of EBD6 is compared with the other eight EBDs (*line 1*); the sequence of the N-terminus (*EBD1*) and C-terminus (*EBD2*) of EBD6 were compared with the N- and C-termini of the other eight EBDs (*lines 2 and 3*). The N-terminal half of EBD1 and the C-terminal half of EBD1 were further compared with the same regions of the other eight EBDs (*lines 4 and 5*)

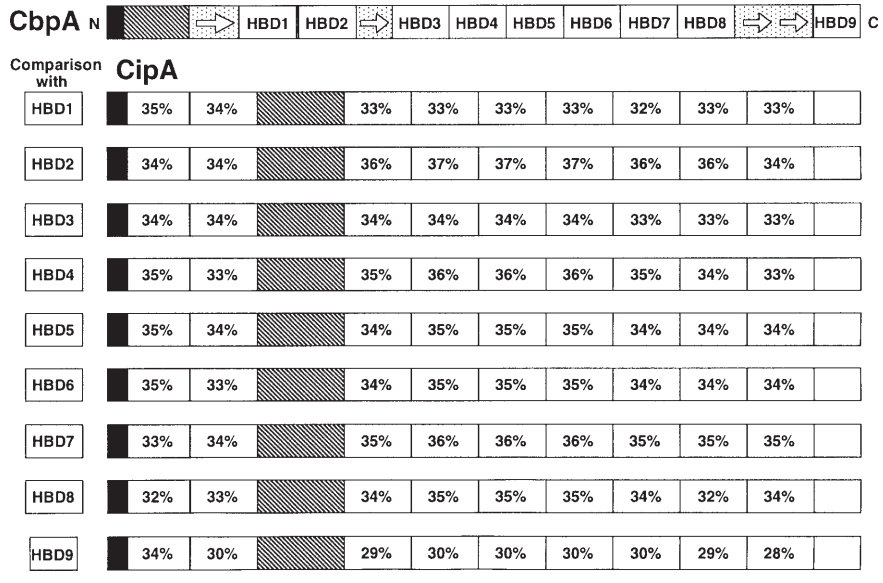

Fig. 3. Comparison of the homology of *C. cellulovorans* CbpA and *C. thermocellum* CipA HBD domains. A comparison of the nine hydrophobic domains of CbpA and CipA. The domain organization differs slightly between the two scaffolding proteins. The CipA HBDs were also numbered from the N-terminus as HBD1 through HBD9 for comparison purposes

immunoblots of Avicel-bound material from culture supernatants. The full-sized CbpA was obtained by elution from the corresponding protein band cut off from a SDS-PAGE gel. When the intact, gel-purified CbpA protein was rebound to Avicel and the bound fraction examined by Western immunoblots, it exhibited a similar ladder of fragments. Because the gel-purified CbpA was free of contaminating proteins, these results strongly suggest that CbpA fragmentation is not caused by protease (Malburg and Doi, unpublished observations).

When a similar Avicel-bound preparation of CbpA was examined by Western immuno-blots using an antiserum raised against the recombinant cellulose binding domain (CBD) of CbpA (Goldstein et al. 1993) as well as the anti-CbpA serum against the entire CbpA protein, the ladder of fragments was essentially identical. This indicates that the CbpA fragments observed in washed Avicel-bound preparations all contained CBD, and hence the N-terminus of CbpA (Malburg and Doi, unpublished observations).

Avicel samples containing bound cellulosome were washed with various detergents at 0.1% (w/v) concentration, with gentle agitation at 25°C for 30 min, and the proteins released were examined by SDS-PAGE and Western immunoblots using an anticellulosome antiserum. Only the anionic detergent SDS released the complete cellulosome with its major components CbpA (P170), P100, and ExgS (P70). Nonionic detergents such as Tween 20, Triton X-100, Nonidet P-40, and the zwitterionic 3[(3-

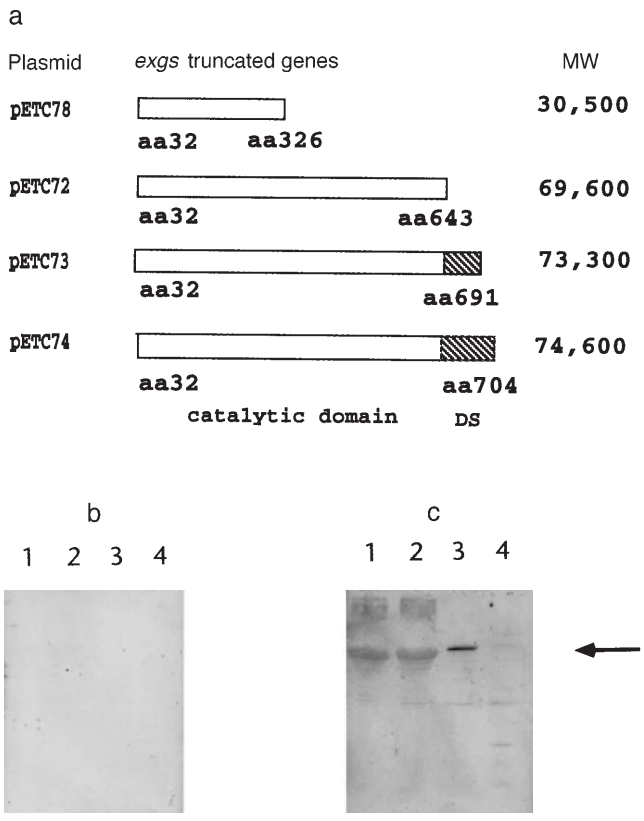

Fig. 4. Requirement for presence of duplicated sequence (*DS*) for binding of exoglucanase S (ExgS) to HBD. **a** Plasmids containing complete ExgS (pETC74) and truncated ExgSs (pETC78, pETC72, and pETC 73) as indicated. **b,c** Interaction between the CBD-HLD_1-HBD_1 of CbpA and ExgS derivatives. *Lanes 1*, *2*, *3*, and *4* contain the proteins encoded by pETC72, pETC73, pETC74, and pETC78, respectively. Interaction Western blotting was carried out as described previously (Takagi et al. 1993) . **b** Membrane was incubated with buffer only and without CBD-HLD_1-HBD_1 for 2 h as control; **c** Membrane was incubated with CBD-HLD_1-HBD_1 for 2 h. The *arrow* points to the strong interaction between the product of pETC74 and ExgS. Only pETC74 had a complete DS. CBD, cellulose-binding domain

cholamidopropyl)dimethylammonio]propanesulfonic acid (CHAPS) promoted elution of the P100 and P70 components only, and no CbpA could be detected. When elution was carried out with CHAPS at 37°C, an increase in the proportion of eluted proteins was observed, but CbpA was still absent from the eluate (Malburg and Doi, unpublished observations). These results indicated that the interaction between CbpA and cellulose was much stronger than the interaction between this anchor protein and the catalytic components of the cellulosome.

Exoglucanase S (ExgS)

The gene for ExgS, one of the major enzymatic subunits of the cellulosome, has been cloned and sequenced (Liu and Doi, submitted for publication). The gene *exgS* codes for a protein with a molecular weight of 77 700. The product has a signal peptide at its N-terminus and a short Pro-Thr-Pro linker that separates the N-terminal half from the C-terminal half. The C-terminal region also has a duplicated sequence (DS) that is involved in binding of ExgS to the EBD domain of CbpA. By comparing the binding of complete and truncated forms of ExgS to the hydrophobic domains (HBD) of CbpA, it was found that the presence of the DS was necessary for firm binding of ExgS to HBD (Fig. 4). Therefore, the function of the DS of *C. cellulovorans* appears to be similar to that found for *C. thermocellum* (Tokatlidis et al. 1991; Salamitou et al. 1992) and *C. cellulolyticum* (Pages et al. 1996; Gal et al. 1997) DS.

Recombinant ExgS (rExgS) containing no signal peptide actively digested several forms of cellulose, including Avicel, Sigmacell101, crystalline cellulose, and xylan, but not carboxymethyl cellulose (CMC), and cellotetraose was the smallest oligosaccharide substrate for rExgS (Liu and Doi, unpublished observations). The enzymatic studies indicated that ExgS was an exoglucanase. The sequence analysis revealed that ExgS had homology to CelS from *C. thermocellum* (Wang et al. 1993), CelF from *C. cellulolyticum* (Reverbel-Leroy et al. 1997), and enzymes from *C. josui* and *Caldocellum saccharolyticum* (Fig. 5).

P100

P100 is the other major protein of the *C. cellulovorans* cellulosome. Zymogram studies indicate that it has endoglucanase activity (Malburg, Liu, Matano, and Doi, unpublished observations). The *P100* gene has not been isolated as yet and thus the exact properties of P100 remain unknown. However, it is clear that the cellulosome contains a major exoglucanase, ExgS, a major endoglucanase, P100, and a scaffolding protein, CbpA, as the major core of the cellulosome. In addition, EngB, an endoglucanase/xylanase, has been identified and characterized (Foong et al. 1991; Foong and Doi 1992) as a cellulosomal enzyme.

Functions of CBD

The function of the *C. cellulovorans* CBD has been analyzed and it shows a high affinity for crystalline cellulose and for chitin (Goldstein et al. 1993). Mutations in the CBD indicate that most of the CBD sequence is necessary for maintaining its cellulose-binding ability (Goldstein and Doi 1994). A fusion protein was formed between the CBD domain and EngB, either at the N-terminus or the C-terminus of the EngB, with a factor Xa cutting site between the CBD and the EngB. These fusion proteins were capable of binding to cellulose; the EngB was active in these fusion proteins while the fusion proteins were bound to cellulose, and an active EngB could be released from the fusion proteins bound to cellulose by treatment with factor Xa (Park and Doi 1997). Potentially, CBD could be used to form many fusion proteins that could bind to cellulose. This is illustrated in Fig. 6.

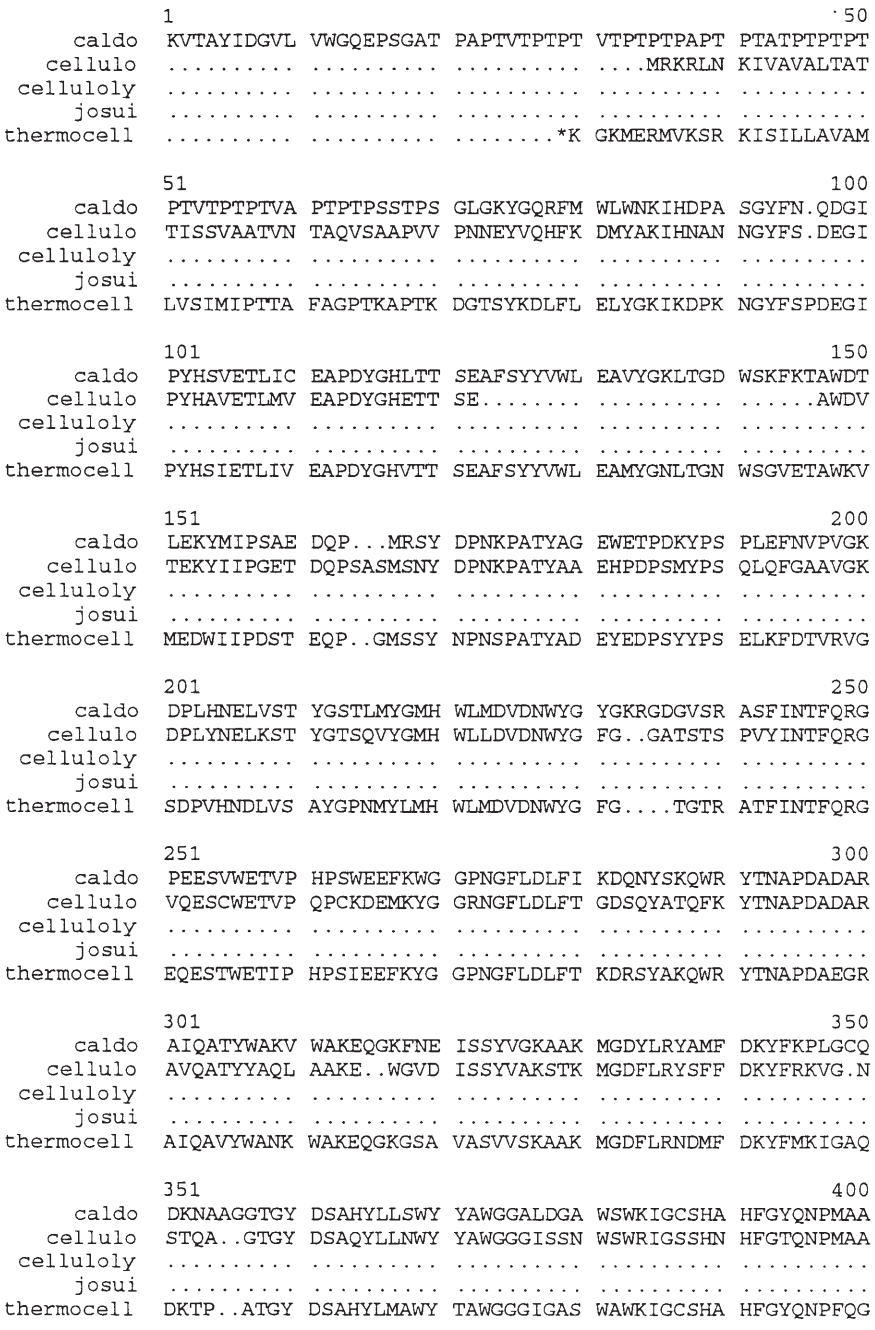

Fig. 5. Amino acid sequence comparison of *Clostridium cellulovorans* ExgS (*cellulo*) with CelS (Tokatlidis et al. 1993) from *C. thermocellum* (*thermocell*), the partial open reading frame (ORF) (*CCF*) (Bagnara-Tardif et al. 1992) of *C. cellulolyticum* (*celluloly*), the partial ORF (Fujino et al. 1993, 1996) of *C. josui* (*josui*), and CelA (Luthi et al. 1991) of *Caldocellum saccharolyticum* (*caldo*)

Noncellulosomal cellulases

Two noncellulosomal cellulases have been identified, EngD (Hamamoto et al. 1992) and EngF (Sheweita et al. 1996). They both contain a cellulose-binding domain that allows them to bind to cellulose. This differs from the cellulosomal enzymes which do not have a cellulose-binding domain, but depend on the CBD present on the CbpA for binding of the enzyme complex to the substrate.

EngD has a high endoglucanase/xylanase activity, in contrast to EngF, which has a low endoglucanase activity on carboxymethylcellulose (CMC). EngF, interestingly, is synthesized in relatively large amounts when *C. cellulovorans* is grown on glucose and cellobiose, when very little cellulosome is made, and is present at relatively smaller amounts when cells are grown on cellulose, when cellulosomes are very abundant (Malburg and Doi, unpublished observations). The types of cellulosomal and noncellulosomal *C. cellulovorans*

Fig. 5. *Continued*

```
                401                                                  450
     caldo  WALANDSDMK PKSPNGASDW AKSLKRQIEF YRWLQSAEGA IAGGATNSWN
   cellulo  WILSNTSDFK PKSPNAATDW NNSLKRQIEF YQWLQSAEGG IAGGASNSNG
 celluloly  .......... .......... ........EF YQWLQSAEGA IAGGATNSWN
     josui  .......... .......... ........EF YQWLQSSEGA IAGGATNSWN
thermocell  WVSATQSDFA PKSSNGKRDW TTSYKRQLEF YQWLQSAEGG IAGGATNSWN

                451                                                  500
     caldo  GRYEKYPAGT ATFYGMAYEP NPVYRDPGSN TWFGFQAWSM QRVAEYYYVT
   cellulo  GSYQAWPAGT RTFYGMGYTP HPVYEDPGSN EWFGMQAWSM QRVAEYYYSS
 celluloly  GRYEAVPSGT STFYGMGYVE NPVYADPGSN TWFGMQVWSM QRVAELYYKT
     josui  GRYESIPSGT STFYGMGYVE NPVYADPGSN TWFGMQVWSM QRVAELYYKT
thermocell  GRYEKYPAGT STFYGMAYVP HPVYADPGSN QWFGFQAWSM QRVMEYYLET

                501                                                  550
     caldo  GDKDAGTLLE KWVSWIKSVV KLNSDG.TFA IPSTLDWSGQ PDTWNGT..Y
   cellulo  KDPAAKSLLD KWAKWACANV QFDDAAKKFK IPAKLVWTGQ PDTW..TGSY
 celluloly  GDARAKKLLD KWAKWINGEI KFNADG.TFQ IPSTIDWEGQ PDTWNPTQGY
     josui  GDTRAKNLLD KWAKWVNSEI KFNADG.TFQ IPGTLDWEGQ PDTWDPTQGY
thermocell  GDSSVKNLIK KWVDWVMSEI KLYDDG.TFA IPSDLEWSGQ PDTW..TGTY

                551                                                  600
     caldo  TGNPNLHVKV VDYGTDLGIT ASLANALLYY SAGTKKYG..  ...VFDEEAK
   cellulo  TGNSNLHVKV EAYGEDLGVA GSLSNALSYY AKALESSTDA ADKVAYNTAK
 celluloly  TGNANLHVKV VNYGTDLGCA SSLANTLTYY AAKSG.....  .D....ETSR
     josui  TGNPNLHVKV VNYNTDLGCA SSLANTLTYY AAKSG.....  .D....TTSK
thermocell  TGNPNLHVRV TSYGTDLGVA GSLANALATY AAATERWEGK LD....TKAR

                601                                                  650
     caldo  NLAKELLDRM WKLY..RDEK GLSAPEKRAD YKRFFEQEVY IPAGWTGKMP
   cellulo  ETSRKILDYL WASY..QDDK GIAVTETRND RKRF.NQSVY IPSGWTGKMP
 celluloly  QNAQKLLDAM WNNY..SDSK GISTVEQRGD YHRFLDQEVF VPAGWTGKMP
     josui  ENAKKLLDAM WNNY..SDSK GISTIEQRGD YHRFLDQEVY VPAGWTGKMP
thermocell  DMAAELVNRA WYNFYCSEGK GVVTEEARAD YKRFFEQEVY VPAGWSGTMP

                651                                                  700
     caldo  NGDVIKSGVK FIDIRSKYKQ DPDWPKLEAA YKSGQVPEFR YHRFWAQCDI
   cellulo  NGDVIQSGAT FLSIRTKYKQ DPSWPKVEAA LANGTGVDMT YHRFWGQSDI
 celluloly  NGDVIKSGVK FIDIRSKYKQ DPEWQTMVAA LQAGQVPTQR LHRFWAQSEF
     josui  NGDVIKSGVK FIDIRSKYKQ DPEWQTMVAA LQAGQVPTQR LHRFWAQSEF
thermocell  NGDKIQPGIK FIDIRTKYRQ DPYYDIVYQA YLRGEAPVLN YHRFWHEVDL

                701                                                  750
     caldo  AIVNATYEIL FGNQ**
   cellulo  AIAFGTYGTL FTD.......  .PTPGLKGDV NSDAKVNAID LAILKKYILD
 celluloly  AVANGVYAIL FPD.......  QGPEKLLGDV NGDETVDAID LAILKKYLLN
     josui  AVANGVYAIL FPE.......  GSN...LGDV NGDETVDAID LAMLKKYLLN
thermocell  AVAMGVLATY FPDMTYKVPG TPSTKLYGDV NDDGKVNSTD AVALKRYVLR

                751                                        791
     caldo  ..........  ..........  ..........  ..........  .
   cellulo  STTKINTANS  DMNGDGKVNA  MDLALLKKAL  LA*.......  .
 celluloly  SSTTINTANA  DMNSDNAIDA  IDYALLKKAL  LSIQ*.....  .
     josui  SSTSIVAGNA  DMNGDGAIDA  IDYALLKKAL  LANQ*.....  .
thermocell  SGISINTDNA  DLNEDGRVNS  TDLGILKRYI  LKEIDTLPYK  N
```

enzymes that have been studied to date are shown in Table 1.

Minicellulosomes

Earlier studies had indicated that minicellulosomes containing mini-CbpAs and cellulosomal enzymatic subunits showed activity against pebble-milled cellulose (Doi et al. 1992). In these studies the mini-CbpA consisted of CBD-HLD1-HBD1, P100 (EngE), and ExgS. A mini-CbpA consisting of only CBD-HBD1 did not show activity with P100 and ExgS, suggesting that HLD1 may have some spacing role. However, the role of HLD still remains unknown and the CipA of *C. thermocellum* does not contain any HLDs (Gerngross et al. 1993).

In more recent studies, interaction was observed between CBD-HLD1-HBD1 and ExgS only when ExgS contained the duplicated sequence (DS) (Liu and Doi, unpublished observations). Truncated forms of ExgS missing the DS did not bind to the mini-CbpA. Furthermore,

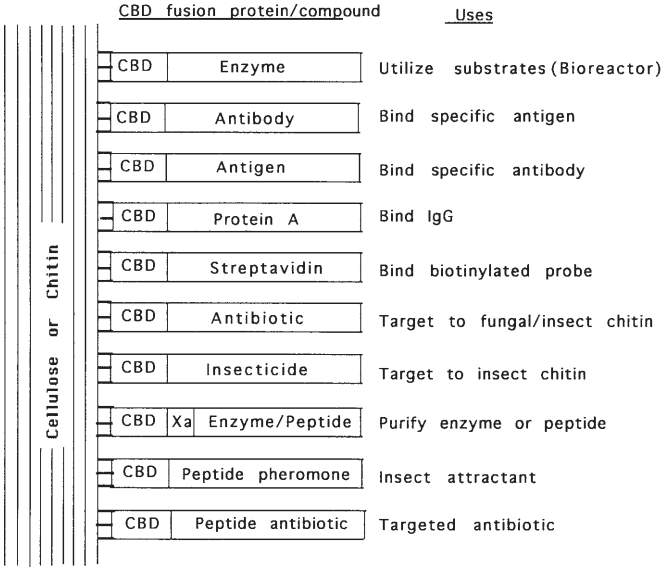

Fig. 6. Potential uses for the technology of binding putative CBD fusion proteins to cellulose. Xa, factor Xa

Table 1. Cellulase genes of *Clostridium cellulovorans*

Gene	Sequenced	Cellulosome[a]	Noncellulosome
engA	Partial	?	?
engB	Yes	Yes	No
engC[b]	Partial	?	?
engD	Yes	No	Yes
engE[b,c]	Partial	Yes	No
engF	Yes	No	Yes
engG[b]	Partial	?	?
engH[b]	Partial	?	?
engJ[b]	Partial	No	Yes
engY[b]	Partial	Yes	?
exgS[b,d]	Yes	Yes	No
cbpA	Yes	Yes	No

[a] Based on the presence of a duplicated sequence (DS) for enzymatic subunits.
[b] Unpublished.
[c] Also called *P100*.
[d] Also called *P70*.

interaction between ExgS and CBD-HLD1-HBD6 also occurred, indicating that both HBD1 and HBD6 could bind ExgS. Calcium ion and ethylenediaminetetraacetic acid (EDTA) had no effect on the binding ability.

Clustered cellulase genes

We have found one cluster of genes involved in cellulase degradation. The cluster includes the P-*regA*-P-*cbpA*-*exgS*-P-*engH* genes (Liu and Doi, unpublished observations). *cbpA* and *exgS* are in the same operon. There is a promoter between *regA* and *cbpA*. There is another promoter region between *exgS* and *engH*. The *P100* gene has not been linked with this gene cluster. Another interesting cluster has been found that includes a pectinase gene (*pec*), *engY*, and a DNA helicase gene (Tamaru and Doi, unpublished observations). All the other cellulase genes investigated to date in *C. cellulovorans* appear to exist as monocistronic units. Clustered genes for cellulase degradation have also been reported for *C. cellulolyticum* (Bagnara-Tardif et al. 1992).

Regulation of cellulosome synthesis

When *C. cellulovorans* is grown on glucose or cellobiose, a very different pattern of cellulase activity is observed. On glucose or cellobiose, very few or no cellulosomes are observed (Matano et al. 1994; Malburg and Doi, unpublished observations) and the cellulosomal subunits that are synthesized in very small quantities are usually present as individual subunits that can assemble into cellulosomes when the cell-free medium is incubated with cellulose (Matano et al. 1994). However, a relatively large amount of EngF is present in the medium under these conditions (Malburg and Doi, unpublished observations). On the other hand, when the cells are grown on cellulose, copious amounts of cellulosome are found (Shoseyov and Doi 1990) as well as EngD and EngF, the two noncellulosomal cellulases (Malburg and Doi, unpublished observations). Thus, the cell represses the synthesis of cellulosomes when glucose or cellobiose is present and derepresses the genes for the cellulosome when cells are grown on cellulose. EngF is made in the presence of either cellobiose or cellulose. When we compared the promoter regions of the *cbpA* operon, *engH,* and *engB*, we somewhat surprisingly did not find a common conserved promoter sequence which could be transcribed by a single type of RNA polymerase (Tamaru and Doi, unpublished observations).

Is there synergism between the cellulosome and noncellulosomal cellulases?

This is an important question, since cells growing on cellulose produce both cellulosome and non-cellulosomal cellulases such as EngD and EngF. The simultaneous presence of the two types of enzymes suggest that they could act in a synergistic fashion. The cellulosome itself is very active on crystalline cellulose. EngD is a very active endoglucanase/xylanase relative to EngF, which has about 100-fold less activity on various substrates (Foong and Doi 1992; Sheweita et al. 1996; Ichiishi, Sheweita, and Doi, submitted for publication). Preliminary results indicate that, when the activity of cellulosome alone, EngF alone, and mixtures of cellulosome and EngF at various ratios is determined, the activity is additive rather than synergistic (Ibrahim and Doi, unpublished observations). Further investigation with EngD may be more fruitful in this regard and is in progress.

Conclusions

The current status of our understanding of the *C. cellulovorans* cellulosome is that much useful information has been obtained about the cellulosome subunits and their interactions. However, there is much more to be learned about the assembly process, the regulation of expression of the genes for the cellulosome subunits, and any synergism that may exist between the cellulosome and noncellulosome cellulases. This work could lead to the genetic engineering of more effective cellulases, a more efficient cellulosome, and a cellulase-overproducing strain of *C. cellulovorans*.

Acknowledgments The research from our laboratory was supported in part by grant DE-DDF03-92ER20069 from the US Department of Energy and grant 94-37308-0399 (I) from the US Department of Agriculture.

References

Bagnara-Tardif C, Gaudin C, Belaich A, Hoest P, Citard T, Belaich JP (1992) Sequence analysis of a gene cluster encoding cellulases from *Clostridium cellulolyticum*. Gene 119:17–28

Beguin P (1996) The cellulosome – an exocellular, multiprotein complex specialized in cellulose degradation. Crit Rev Biochem Mol Biol 31:201–236

Doi RH, Goldstein MA, Hashida S, Park JS, Takagi M (1994) The *Clostridium cellulovorans* cellulosome. CRC Rev Microbiol 20:87–94

Doi RH, Goldstein MA, Park JS, Liu CC, Matano Y, Takagi M, Hashida S, Foong FCF, Hamamoto T, Segel I, Shoseyov O (1994) Structure and function of the subunits of the *Clostridium cellulovorans* cellulosome. In: Shimada K, Hoshino S, Ohmiya K, Sakka K, Kobayashi Y, Karita S (eds), Genetics, biochemistry and ecology of lignocellulose degradation. Uni Tokyo, Japan, pp 43–52

Foong F, Doi RH (1992) Characterization and comparison of *Clostridium cellulovorans* endoglucanases-xylanases EngB and EngD hyperexpressed in *Escherichia coli*. J Bacteriol 174:1403–1409

Foong F, Hamamoto T, Shoseyov O, Doi RH (1991) Nucleotide sequence and characteristics of endoglucanase gene *engB* from *Clostridium cellulovorans*. J Gen Microbiol 137:1729–1736

Fujino T, Beguin P, Aubert JP (1993) Organization of a *Clostridium thermocellum* gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in attachment of the cellulosome to the cell surface. J Bacteriol 175:1891–1899

Fujino T, Fujino E, Karita S, Ohmiya K (1996) GenBank accession number D85526

Gal L, Pages S, Gaudin C, Belaich A, Reverbel-Leroy C, Tardif C, Belaich JP (1997) Characterization of the cellulolytic complex (cellulosome) produced by *Clostridium cellulolyticum*. Appl Environ Microbiol 63:903–909

Gerngross U, Romaniec MPM, Kobayashi T, Huskisson NS, Demain AL (1993) Sequencing of a *Clostridium thermocellum* gene (*cipA*) encoding the cellulosomal SL-protein reveals an unusual degree of internal homology. Mol Microbiol 8:325–334

Goldstein MA, Doi RH (1994) Mutation analysis of the cellulose-binding domain of the *Clostridium cellulovorans* cellulose-binding protein A. J Bacteriol 176:7328–7334

Goldstein MA, Takagi M, Hashida S, Shoseyov O, Doi RH, Segel IH (1993) Characterization of the cellulose-binding domain of the *Clostridium cellulovorans* cellulose-binding protein A. J Bacteriol 175:5762–5768

Hamamoto T, Foong F, Shoseyov O, Doi RH (1992) Analysis of functional domains of endoglucanases from *Clostridium cellulovorans* by gene cloning, nucleotide sequencing and chimeric protein construction. Mol Gen Genet 231:472–479

Lamed R, Bayer EA (1988) The cellulosome of *Clostridium thermocellum*. Adv Appl Microbiol 33:1–46

Luthi E, Jasmat NB, Grayling RA, Love DR, Bergquist PL (1991) Cloning, sequence analysis, and expression in *Escherichia coli* of a gene coding for a β-mannanase from the extremely thermophilic bacterium *Caldocellum saccharolyticum*. Appl Environ Microbiol 57:694–700

Matano Y, Park JS, Goldstein MA, Doi RH (1994) Cellulose promotes extracellular assembly of *Clostridium cellulovorans* cellulosomes. J. Bacteriol 176:6952–6956

Pages S, Belaich A, Tardif C, Reverbel-Leroy C, Gaudin C, Belaich JP (1996) Interaction between the endoglucanase CelA and the scaffolding protein CipC of the *Clostridium cellulolyticum* cellulosome. J Bacteriol 178:2279–2286

Park JS, Doi RH (1997) *Clostridium cellulovorans* cellulase and its useful applications. Proceedings, international symposium on strain development for bioindustry, Korean Society of Applied Microbiology, pp 71–78

Reverbel-Leroy C, Pages S, Belaich A, Belaich JP, Tardif C (1997) The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: purification and characterization of the recombinant form. J Bacteriol 179:46–52

Salamitou S, Raynaud O, Lemaire M, Coughlan M, Beguin P, Aubert JP (1994) Recognition specificity of the duplicated segments present in *Clostridium thermocellum* endoglucanase CelD and in the cellulosome integrating protein CipA. J Bacteriol 176:2822–2827

Sheweita S, Ichi-ishi A, Park JS, Liu CC, Malburg LM, Doi RH (1996) Characterization of *engF*, a gene for a noncellulosomal *Clostridium cellulovorans* endoglucanase. Gene 182:163–167

Shoseyov O, Doi RH (1990) Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Proc Natl Acad Sci USA 87:2192–2195

Shoseyov O, Hamamoto T, Foong F, Doi RH (1990) Cloning of *Clostridium cellulovorans* endo-1,4-beta-glucanase genes. Biochem Biophys Res Commun 169:667–672

Takagi M, Hashida S, Goldstein MA, Doi RH (1993) The hydrophobic repeated domain of the *Clostridium cellulovorans* cellulose-binding protein (CbpA) has specific interactions with endoglucanases. J Bacteriol 175:7119–7122

Tokatlidis K, Salamitou S, Beguin P, Dhurjati P, Aubert JP (1991) Interaction of the duplicated segment carried by *Clostridium thermocellum* cellulases with cellulosome components. FEBS Lett 291:185–188

Wang WK, Kruus K, Wu JHD (1993) Cloning and DNA sequence of the gene coding for *Clostridium thermocellum* cellulase S_S (CelS), a major cellulosome component. J Bacteriol 175:1293–1302