JOURNAL OF BACTERIOLOGY, Dec. 1991, p. 7956–7962
0021-9193/91/247956-07$02.00/0

Vol. 173, No. 24

Characterization of Endoglucanase A from *Clostridium cellulolyticum*

H.-P. FIEROBE,[1] C. GAUDIN,[1*] A. BELAICH,[1] M. LOUTFI,[1†] E. FAURE,[1‡] C. BAGNARA,[1]
D. BATY,[2] AND J.-P. BELAICH[1]

Laboratoire de Chimie Bactérienne,[1] and Centre de Biochimie et Biologie Moléculaire,[2] Centre National de la Recherche Scientifique, 31 Chemin Joseph Aiguier, BP71, 13277 Marseille Cedex 9, France

Received 18 June 1991/Accepted 2 October 1991

A construction was carried out to obtain a high level of expression in *Escherichia coli* of the gene *celCCA*, coding for the endoglucanase A from *Clostridium cellulolyticum* (EGCCA). The enzyme was purified in two forms with different molecular weights, 51,000 and 44,000. The smaller protein was probably the result of proteolysis, although great care was taken to prevent this process from occurring. Evidence was found for the loss of the conserved reiterated domains which are characteristic of *C. thermocellum* and *C. cellulolyticum* cellulases. The two forms were extensively studied, and it was demonstrated that although they had the same pH and temperature optima, they differed in their catalytic properties. The truncated protein gave the more efficient catalytic parameters on carboxymethyl cellulose and showed improved endoglucanase characteristics, whereas the intact enzyme showed truer cellulase characteristics. The possible role of clostridial reiterated domains in the hydrolytic activity toward crystalline cellulose is discussed.

Cellulolysis by mesophilic clostridia has been less extensively studied up to now than cellulolysis by thermophilic bacteria of the same genus (1, 4). These bacteria are of great interest, however (4). In particular, their mesophilic character can be an advantage for industrial processes such as the treatment of urban and industrial refuse or the production of chemicals. For several years, therefore, we have been investigating the cellulolytic system of *Clostridium cellulolyticum* (7, 8, 10, 11, 20). Once the characterization of the bacteria (20) and its physiological study were well under way (10, 11), we began to study its cellulolytic components by using a molecular genetics approach (7). Up to now, we have identified five genes coding for components of the cellulolytic system of *C. cellulolyticum*; one of these, *celCCA*, coding for endoglucanase A from *C. cellolyticum* (EGCCA), has been sequenced (8). This enzyme comprises a catalytic domain belonging to family A, according to the classification of cellulases proposed by Henrissat et al. (15), and a C-terminal domain of the type found so far only in clostridial cellulases, consisting of two reiterated sequences (1, 4).

The initial carboxymethyl cellulose (CMC)-positive clone obtained yielded only poor enzyme production, and in this report we describe the genetic construction carried out to obtain a higher level of EGCCA production. The subsequent purification yielded the separation of two active forms of the enzyme differing in size, the smaller (SS) form being the result of partial proteolysis which was responsible for the loss of reiterated domains. The two forms of the enzyme were characterized, and the most noteworthy differences in their catalytic properties are described. The significance of these differences from the point of view of the cellulolytic function and the possible role of reiterated domains will be discussed.

* Corresponding author.
† Present address: University Hassan II, Casablanca, Morocco.
‡ Present address: Institut de Chimie Biologique, Marseille, France.

MATERIALS AND METHODS

Bacteria, media, and reagents. *C. cellulolyticum* H10 was used (20). The main buffers used were TD buffer (50 mM Tris-HCl, pH 8.0), TDU buffer (TD plus 2 M urea and 1 mM phenylmethylsulfonyl fluoride), and TAK buffer (10 mM Tris-HCl, 1 M potassium acetate, pH 8.0).

Strains and plasmids. *Escherichia coli* TG1 (24) was used as a host for recombinant plasmids. BMM-71-18-*mut*-L (17) was the recipient strain for site-directed mutagenesis with M13mp18 bacteriophage. Plasmids pUC18 (26) and pJF 118EH (9) were used as the cloning vector and the expression vector, respectively. *E. coli* cells were grown on minimal medium or LB medium supplemented with ampicillin (100 μg/ml) when needed. Synthesis of endoglucanase was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) when the optical density of the cultures reached 1.0. Bacteria were harvested by centrifugation 3 h after induction.

celCCA, under the control of its own promoter, was poorly expressed in *E. coli*. To increase this expression, the gene was introduced into plasmid pJF118EH under the control of the P_{tac} promoter. The construction was carried out as follows (Fig. 1). The entire sequence of *celCCA* included in a 2-kb *Hpa*II-*Hpa*II fragment of plasmid pB52 (8) was cloned at the *Acc*I site of pUC18, generating plasmid pA24. Since no single restriction site was available in the beginning of the sequence, a *Pst*I site was introduced by site mutagenesis (Fig. 1, M13A24P). The construction was performed by cloning the *Pst*I-*Hin*dIII fragment of replicative-form M13A24P at the *Eco*RI and *Hin*dIII sites of pJF118EH, the junction being ensured by means of a synthetic linker restoring the beginning of the gene; an ATG codon was added to reconstitute the initiation region. The 400-bp *Hin*dIII fragment of pA24 including the stop codon and transcriptional termination sequences was replaced to complete the construction. The resulting plasmid was called pA2. Strain TG1 harboring this plasmid (strain TG1[pA2]) was used to obtain high quantities of EGCCA after induction (about 10 to 20% of the total proteins of *E. coli*).

Purification of EGCCA. Cells (180 g, wet weight) were

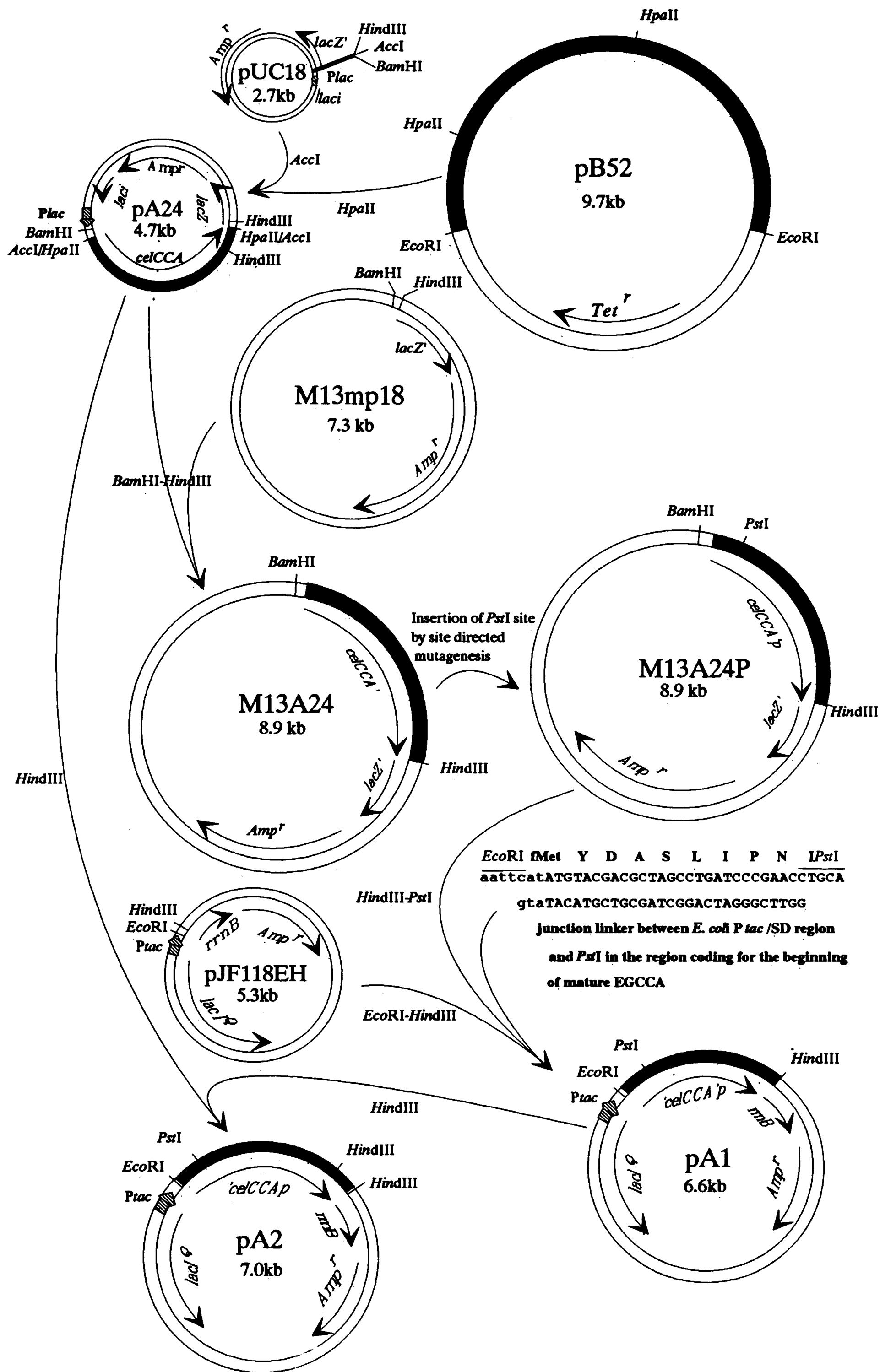

FIG. 1. Plasmid maps showing the generation of plasmid pA2, yielding high expression of the *celCCA* gene. Dark and open boxes indicate insert and vector, respectively. Hatched boxes indicate promoter *lac* or *tac* regions.

broken in a French press and centrifuged for 20 min at 10,400 × g; 15 g of streptomycin sulfate was then added to the supernatant, and the cells were centrifuged again to precipitate the nucleic acids. Two protein fractions were harvested by centrifugation after successive precipitations in the presence of ammonium sulfate at saturation rates of 30 and 60%, respectively. Each pellet (SA30 and SA60) was resuspended in 0.5 liter of TDU and loaded on a Q-Sepharose column (10 by 8 cm; Pharmacia) equilibrated with the same buffer. The resin was then washed with increasing concentrations of NaCl (0.05 to 0.5 M) in 1 liter of TDU. The most active fractions in each case were pooled, dialyzed against TD, and loaded on a DEAE-Trisacryl column (5 by 21 cm; IBF) equilibrated with the same buffer. Elution was performed with a linear gradient of NaCl (0 to 0.3 M in two 1-liter volumes of TD). This operation was repeated twice with each form. The samples showing the best specific activity were dialyzed against TAK and loaded on a phenyl-Sepharose column (Pharmacia) equilibrated with the same buffer. The elution was carried out with increasing dilutions of TAK. A final attempt at purification was made by performing gel filtration. A column (1.5 cm by 1 m) was filled with Ultrogel AcA44 (IBF) and equilibrated with 0.1 M NaCl–25 mM potassium phosphate, pH 7.0. The elution was carried out with the same buffer at a flow rate of 0.4 ml/min. At each step in the purification, the most interesting fractions were analyzed by sodium dodecyl sulfate–12.5% polyacrylamide gel electrophoresis (SDS-PAGE) (Pharmacia Phast system).

Enzyme assays. The protein concentration was determined by the method of Lowry et al. (18), with bovine serum albumin as the standard.

Carboxymethyl cellulase (CMCase) activity was assayed by mixing 1 ml of an appropriate enzyme dilution with 4 ml of CMC (medium viscosity; Sigma), 10 g/liter in 25 mM potassium phosphate (pH 7.0) (or 25 mM Tris [pH 7.0] when the activity was tested in the presence of Ca^{2+}). Aliquots of 1 ml were pipetted at specific intervals. The amounts of reducing sugars released were determined by the Park and Johnson ferricyanide method (19). One unit of activity (international unit) corresponds to 1 μmol of D-glucose equivalent liberated per min.

The increase in fluidity of the CMC solution was also used to characterize the CMCase activity. The procedure was the same as previously; 1-ml aliquots were boiled for 15 min, and the fluidity was measured at room temperature in a Prolabo viscosimeter adapted for dealing with small volumes. The specific fluidity was estimated by using the formula $F_{sp} = T_0/(T-T_0)$, where T_0 is the flow time of the water and T is the flow time of the CMC solution.

CMCase activity was determined in the presence of potential activator or inhibitor (EDTA, $MgCl_2$, $CaCl_2$, D-glucose, or cellobiose) by preincubating the enzyme with one of these components for 10 min at 37°C before addition of CMC.

The insoluble sugars barley glucan, lichenan, xylan (Sigma), and Avicel (Flucka) were washed twice with water to eliminate any soluble sugars and resuspended in 25 mM potassium phosphate (pH 7.0) at a concentration of 10 g/liter. The assays were performed under the same conditions as previously. The reducing sugars produced were measured at various incubation times, using the Park and Johnson method after removing the solids by centrifugation. The same procedure (except for the centrifugation) was used to determine the laminarinase activity, but instead of 1 ml, the volume of the aliquots pipetted was 10 μl (because laminarin itself reacts strongly with the ferricyanide method reactives). The kinetics were then recorded for several hours (instead of a few minutes). The activity units were defined as for CMCase.

p-Nitrophenol (pNP)-cellobiose, pNP-lactose, pNP-glucose, and pNP-xylose (Sigma) were tested. Nine hundred microliters of a 1-g/liter solution of chromophoric substrate (in 25 mM potassium phosphate, pH 7.0) was mixed with 100 μl of diluted enzyme at 37°C. Spectrophotometric measurements (400 nm) were performed to monitor the release of pNP. One unit of activity was defined as the liberation of 1 μmol of pNP (using pNP from Sigma as the standard) per min. Cellobiose and lactose (Sigma) were also tested; the catalytic activity was then estimated by performing glucose titrations at various times.

Binding experiments on Avicel. Eight hundred microliters of Avicel at various concentrations (ranging from 5 to 20 g/liter) in 25 mM phosphate or Tris (pH 7.0) was incubated with 200 μl of purified enzyme for 90 min at 37°C. Avicel was then pelleted and washed with the same buffer. The second pellet was resuspended in the same buffer, 8 M urea was added for 30 min, and the sample was centrifuged again. Each supernatant was tested to determine the protein titration and CMCase activity and then analyzed by SDS-PAGE (12.5% gel) after being concentrated on a Centricon (Amicon). Binding tests in the presence of Ca^{2+} were also carried out.

Amino acid composition and N-terminal sequencing. Amino acid compositions of the purified proteins were established by using a Beckman 6300 amino acid analyzer. N-terminal amino acid sequences were determined by using an Applied Biosystems 470 A sequence analyzer.

RESULTS

Preparation of purified endoglucanases. The purification pathway is summarized in Table 1. The main difficulty encountered was the cleavage of the protein in the early steps of separation. To prevent proteolysis, the first two purification steps were carried out in the presence of 1 mM phenylmethylsulfonyl fluoride.

After the fractioned ammonium sulfate precipitations, the precipitates were solubilized in 6 M urea and subjected to Q-Sepharose fractionation in the presence of 2 M urea. These initial precautions made it possible to limit the proteolysis. In this way, we obtained two forms of EGCCA which were largely separated during the ammonium sulfate precipitation steps. A larger (LS) form, apparently corresponding to the intact gene product, precipitated mainly in 30% ammonium sulfate. A SS form, resulting probably from a partial proteolysis, precipitated mainly in 60% ammonium sulfate. The subsequent steps were carried out in parallel on the two forms. After the Q-Sepharose steps, each form was found to be contaminated mainly by the other (Fig. 2). This made it even more difficult to achieve homogeneity. The large number of separation steps required as the result of this situation and the elimination of fractions which were not sufficiently homogeneous were mostly responsible for the poor yields obtained. Even when all separation steps had been carried out, the LS form remained partially contaminated by traces (<5%) of the SS form, and an ultimate gel permeation step did not improve the degree of homogeneity. It is worth noting the difference in specific activity: the SS form was 1.8 times more active than the LS form. A comparable but less marked difference was observed in the case of a *Cellulomonas* cellulase when it was deleted from its cellulose binding domain (12). In the case of EGCCA, the

TABLE 1. Purification of the two forms of EGCCA from *E. coli* extract

Fraction	CMCase (IU)	Protein (mg)	CMCase sp act (IU/mg)	Yield (%)	Purification (fold)
Crude extract	108,000	14,040	7.7	100.0	1.0
Streptomycin sulfate treatment	89,620	11,620	7.7	83.0	1.0
LS form					
30% ammonium sulfate precipitation	24,900	1,580	15.8	23.0	2.1
Q-Sepharose	14,080	396	35.5	13.0	4.6
DEAE-Trisacryl	6,600	145	45.6	6.1	5.9
Phenyl-Sepharose	4,330	67	64.3	4.0	8.4
SS form					
60% ammonium sulfate precipitation	40,700	3,720	11.1	37.7	1.4
Q-Sepharose	14,070	355	39.6	13.0	5.1
DEAE-Trisacryl	9,520	90	106.3	8.8	13.8
Phenyl-Sepharose	3,040	26	117.2	2.8	15.2

N-terminal sequences of the two purified forms were determined and found to be identical to (i.e., MYDASLIP) and in agreement with the sequence deduced from the nucleotide sequence (8). This result indicated that the proteolysis occurred in the C-terminal region. To locate the site of proteolysis, the amino acid compositions of both purified forms were determined. Analysis of the results (Table 2) indicated that 69 residues were probably deleted just before the first of the two reiterated domains observed in the C-terminal region. The molecular weights of the two forms were estimated from their relative mobility on SDS-PAGE, using the weight markers as standards. The LS form has an apparent mass of 51,000 Da, and the SS form has an apparent mass of 44,000 Da. These values were in good agreement with the molecular weights predicted from amino acid sequences for both forms (Table 2).

Enzyme kinetics. The enzyme kinetics of two forms of EGCCA were studied to determine their enzymatic proper-

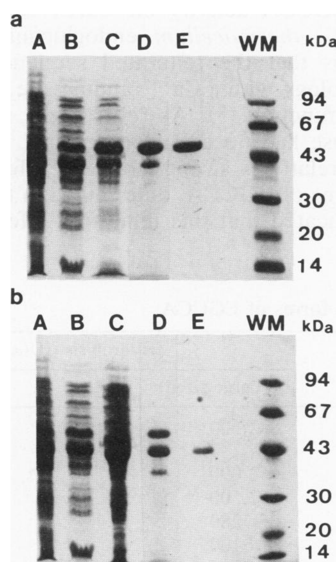

FIG. 2. SDS-PAGE analysis of EGCCA preparations at various purification steps. (a) LS protein; (b) SS protein. Lanes: A and B, the common steps (crude extract and streptomycin sulfate treatment, respectively); C, result of ammonium sulfate treatment (30% for panel a; 60% for panel b); D, result of DEAE-Trisacryl separation; E, result of final separation on phenyl-Sepharose; WM, migration of weight markers ranging from 14 to 94 kDa.

TABLE 2. Comparison of predicted and determined amino acid compositions of the two forms of EGCCA[a]

Amino acid	Amino acid composition			
	LS form		SS form	
	Predicted[b]	Determined[c]	Predicted	Determined
Ala	31	32	24	25
Cys	7	ND	7	ND
Asx	83	82	66	66
Glx	27	32	26	27
Phe	20	19	18	17
Gly	31	34	27	29
His	8	8	7	7
Ile	31	29	28	26
Lys	28	28	22	22
Leu	26	28	17	21
Met	15	16	13	12
Pro	17	18	14	15
Arg	13	14	13	13
Ser	26	23	24	22
Thr	19	19	19	19
Val	33	32	26	26
Trp	11	ND	11	ND
Tyr	24	22	19	17
Total	450	454[d]	381	382[d]
Mol wt	50,802	51,337	43,194	43,429
Mol wt by SDS-PAGE estimation	51,000		44,000	

[a] Amino acid sequence of EGCCA showing the probable proteolysis site and locations of reiterated domains. EGCCA sequence in clone pA2 (8):
MYDASLIPNLQIPQKNIPNNDGMNFVKGLRLGWNLGNTFDAFNGTNITNELDYE
TSWSGIKTTKQMIDAIKQKGFNTVRIPVSWHPHVSGSDYKISDVWMNRVQEVVN
YCIDNKMYVILNTHHDVDKVKGYFPSSQYMASSKKYITSVWAQIAARFANYDEH
LIFEGMNEPRLVGHANEWWPELTNSDVVDSINCINQLNQDFVNTVRATGGKNAS
RYLMCPGYVASPDGATNDYFRMPNDISGNNNNKIIVSVHAYCPWNFAGLAMADG
GTNAWNINDSKDQSEVTWFMDNIYNKYTSRGIPVIIGECGAVDKNNLKTRVEYM
SYYVAQAKARGILCILWDNNNFSGTGELFGFFDRRSCQFKFPEIIDGMVKYAFE
AKT(DP)DPVIVYGDYNNDGNVDALDFAGLKKYIMAADHAYVKNLDVNLDNEVNA
FDLAILKKYLLGMVSKLPSN

[b] Deduced from the published sequence (8). For the SS form, a deletion of 69 amino acids was assumed to occur at the C-terminal end, i.e., a cleavage between Asp-381 and Pro-382 (between parentheses in the EGCCA sequence) nine residues before the first reiterated sequence (underlined).

[c] Determined from analysis of the two purified forms. ND, not determined.

[d] To the total amino acids from the analysis were added predicted Cys and Trp residues which cannot be exactly estimated by analysis.

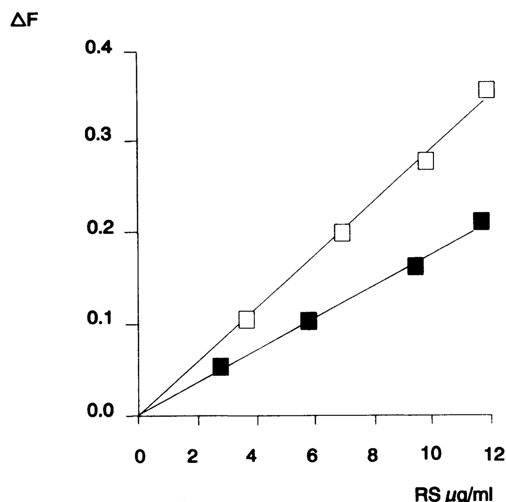

FIG. 3. Changes in relative CMC fluidity (ΔF) versus reducing sugars (RS) released from CMC by the LS (black squares) and the SS (white squares) forms of EGCCA.

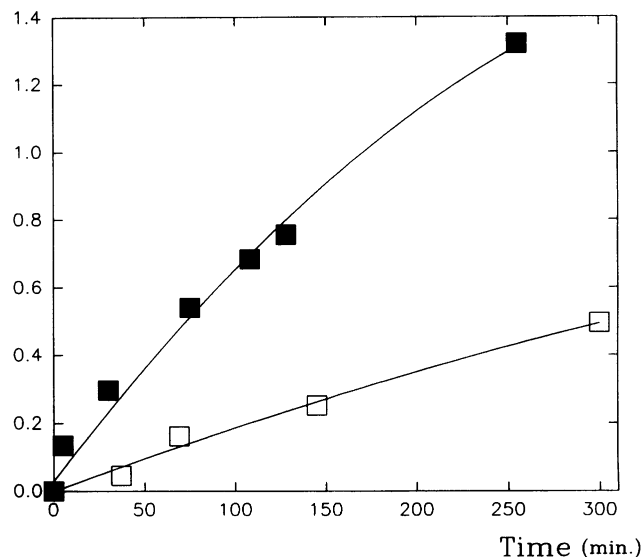

FIG. 4. Evolution of the specific release of reducing sugars from Avicel (8 g/liter) versus time by the two forms of EGCCA. Symbols are as in Fig. 3.

ties. The LS and SS forms differed in V_{max} (60.8 and 151.1 IU/mg, respectively) and K_m (1.97 and 3.96 g/liter, respectively). It thus emerged that the deletion of the reiterated domain of EGCCA resulted in an increase in the velocity accompanied by a relative loss of affinity for CMC.

As regards the hydrolytic mechanism of CMC, the changes in relative fluidity versus the production of reducing sugars were monitored in the case of both forms (Fig. 3). It is worth noting that the protein cleavage in the SS form resulted in an enhancement of the endo mode of action of EGCCA.

Table 3 shows the hydrolytic activities of the two forms of EGCCA toward various substrates. The hydrolysis was clearly of an endoglucanase character. The SS form was more efficient than the LS form on soluble celluloses and related compounds, but this situation was completely reversed with crystalline Avicel as the substrate (Fig. 4). The SS form appeared to have more pronounced endoglucanase characteristics than the LS form, but the latter seemed to be a more efficient cellulase than the former.

Effects of temperature and pH. Figures 5 and 6 show the effects of the temperature and pH on the CMCase activities of the two forms of the enzyme. No significant difference was observed between the behaviors of the two forms when these physicochemical conditions were varied. The temperature optimum was in the range of 37 to 51°C, and the pH optimum was from 5.5 to 7.2.

Sensitivity of EGCCA activity to inhibition or activation. The effects of various compounds on the activities of the two forms of EGCCA were tested. Neither EDTA nor Ca^{2+} affected the EGCCA activity on CMC, contrary to what occurs with *C. thermocellum* endoglucanase D (EGD), which confirms that the reiterated sequence, despite its sequence homology with a Ca^{2+} binding site, probably does not have this function (5). As regards the soluble sugars, although glucose had no effect on activity, cellobiose was found to have relatively strong inhibitory effects. The kinetics of both forms of EGCCA, determined in the presence of cellobiose, indicated that this inhibitory effect might be of

TABLE 3. Hydrolysis of various substrates by the two forms of EGCCA

Substrate	Bond type	Monomer	Hydrolysis (IU/µM)	
			LS (large-sized)	SS (small-sized)
CMC	β-1,4	Glucose	2,500	4,400
Laminarin	β-1,3	Glucose	0	0
Lichenan	β-1,3 and β-1,4	Glucose	2,090	3,440
Barley glucan	β-1,3 and β-1,4	Glucose	2,500	4,530
Xylan	β-1,4	Xylose	280	440
Swollen Avicel	β-1,4	Glucose	100	180
Avicel	β-1,4	Glucose	5.4	1.2
pNP-cellobiose	β-1,4		1.5	1.4
pNP-glucose	β-1,4		0.0	0.0
pNP-lactose	β-1,4		1.1	0.9
pNP-xylose	β-1,4		0.0	0.0
Cellobiose	β-1,4	Glucose	0.0	0.0
Lactose	β-1,4		0.0	0.0

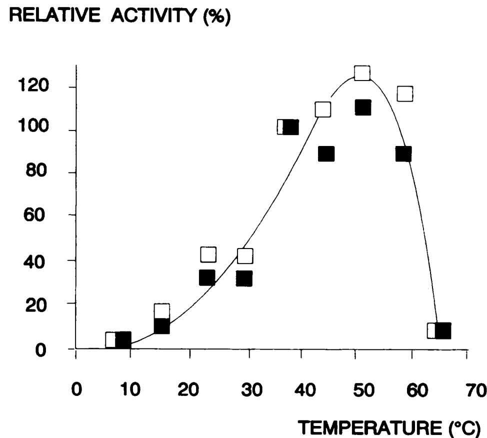

FIG. 5. Temperature dependence of the activities of both EGCCA forms on CMC. The relative activities are compared with those under standard conditions, i.e., pH 7.0 and 37°C. Symbols are as in Fig. 3.

the competitive type at an apparent K_i of around 11 mM (results not shown).

Binding to Avicel. Despite numerous attempts under various experimental conditions (relative proportions of Avicel and enzyme, presence or absence of Ca^{2+}), neither the LS nor the SS form of EGCCA was found to bind to Avicel.

DISCUSSION

Purification of EGCCA yielded two forms of enzyme differing in size as the result of a partial proteolysis occurring in the C-terminal region. The resulting forms differed in that the reiterated domains specific to the *C. thermocellum* and *C. cellulolyticum* cellulase sequences (1, 4) were present in the one case and absent in the other. The loss of these domains led to an enhancement of the EGCCA endo-type characteristics and a decrease in its catalytic properties on crystalline cellulose. In terms of catalysis, the substrates of EGCCA can therefore be divided into three categories: (i) the small substrates, the degradation of which does not depend on the size of the protein; (ii) the amorphous celluloses, on which the truncated protein is twice as efficient as the intact form; and (iii) crystalline cellulose, on which the integral protein is most active (fourfold).

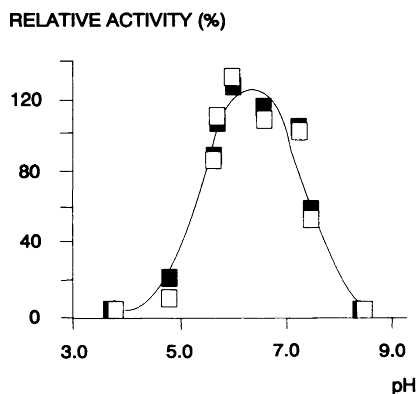

FIG. 6. pH dependence of the activities of both EGCCA forms on CMC. The relative activities are compared with those under standard conditions, i.e., pH 7.0 and 37°C. Symbols are as in Fig. 3.

The other parameters of the activities (temperature and pH) did not differ between the two forms. The differences between the catalytic properties of the two forms might reflect the function of the C-terminal reiterated domains, which have been found to exist only in *C. thermocellum* and *C. cellulolyticum* cellulases. Up to now, their role has been a matter of speculation (2), but it seems likely from the present study that they may have some influence on the mode of cellulose degradation by the enzyme. This may not be the general rule, however, since no such effects were observed with EGE (14), EGH (25), or XynZ (13) from *C. thermocellum*. It was suspected that these reiterated domains may constitute a cellulose binding domain such as those observed with *Cellulomonas* and *Trichoderma* cellulases (12, 21–23), but it was not possible to observe any binding of the intact form of EGCCA to Avicel. Another argument against this binding activity was recently put forward in the case of EGE from *C. thermocellum*, for which the existence of a cellulose binding domain which is separate from the reiterated domains was reported (6). Changes in the activity levels and/or specificities induced by protein truncation were observed, however, in the case of an endoglucanase from *Cellulomonas uda* (12) and in the case of β-glucosidase B from *C. thermocellum* (16). To conclude, it might be speculated that in the case of EGCCA, the presence of the reiterated domains may have facilitated the degradation of crystalline cellulose. Whether this effect was a property of the domains themselves or an indirect effect inducing a conformational change in the enzyme could not be determined from our experiments.

ACKNOWLEDGMENTS

We thank the Service de Séquençage LCB/CBM for performing amino acid composition analysis and the amino-terminal Edman sequencing of the two forms of EGCCA. We are grateful to Jessica Blanc for corrections to the manuscript.

This work was supported by grants from European Economic Community and from Gaz de France Co.

REFERENCES

1. **Aubert, J.-P., P. Béguin, and J. Millet.** 1991. Genes and proteins involved in cellulose and xylan degradation by *Clostridium thermocellum*. *In* M. Sebald (ed), Genetics and molecular biology of anaerobic bacteria, in press. Springer-Verlag, New York.
2. **Béguin, P.** 1990. Molecular biology of cellulose degradation. Annu. Rev. Microbiol. **44**:219–248.
3. **Béguin, P., P. Cornet, and J.-P. Aubert.** 1985. Sequence of a cellulase gene of the thermophilic bacterium *Clostridium thermocellum*. J. Bacteriol. **162**:102–105.
4. **Belaich, J.-P., A. Belaich, C. Gaudin, and C. Bagnara.** Genes and proteins involved in cellulose degradation by mesophilic *Clostridia*. *In* M. Sebald (ed.), Genetics and molecular biology of anaerobic bacteria, in press. Springer-Verlag, New York.
5. **Chauvaux, S., P. Beguin, J.-P. Aubert, K. M. Bhat, L. A. Gow, T. M. Wood, and A. Bairoch.** 1990. Calcium-binding affinity and calcium-enhanced activity of *Clostridium thermocellum* endoglucanase D. Biochem. J. **265**:261–265.
6. **Durrant, A. J., J. Hall, G. P. Hazlewood, and H. J. Gilbert.** 1991. The non-catalytic C-terminal region of endoglucanase E from *Clostridium thermocellum* contains a cellulose binding domain. Biochem. J. **273**:289–293.
7. **Faure, E., C. Bagnara, A. Belaich, and J.-P. Belaich.** 1988. Cloning and expression of two cellulase genes of *Clostridium cellulolyticum* in *Escherichia coli*. Gene **65**:51–58.
8. **Faure, E., A. Belaich, C. Bagnara, C. Gaudin, and J.-P. Belaich.** 1989. Sequence analysis of the *Clostridium cellulolyticum* en-

doglucanase-a-encoding gene, *celCCA*. Gene **84**:39–46.

9. **Fürste, J. P., W. Pansegrau, R. Frank, H. Blöker, P. Scholz, M. Bagdasarian, and E. Lanka.** 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. Gene **48**:119–131.
10. **Giallo, J., C. Gaudin, and J.-P. Belaich.** 1985. Metabolism and solubilization of cellulose by *Clostridium cellulolyticum* H10. Appl. Environ. Microbiol. **49**:1216–1221.
11. **Giallo, J., C. Gaudin, J.-P. Belaich, E. Petitdemange, and F. Caillet-Mangin.** 1983. Metabolism of glucose and cellobiose by cellulolytic mesophilic *Clostridium* sp. strain H10. Appl. Environ. Microbiol. **45**:843–849.
12. **Gilkes, N. R., R. A. J. Warren, R. C. Miller, Jr., and D. G. Kilburn.** 1988. Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. J. Biol. Chem. **263**:10401–10407.
13. **Grepinet, O., M.-C. Chebrou, and P. Beguin.** 1988. Nucleotide sequence and deletion analysis of the xylanase gene (*xynZ*) of *Clostridium thermocellum*. J. Bacteriol. **170**:4582–4588.
14. **Hall, J., G. P. Hazlewood, P. J. Barker, and H. J. Gilbert.** 1988. Conserved reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for catalytic activity. Gene **69**:29–38.
15. **Henrissat, B., M. Claeyssens, M. Tomme, P. Lemesle, and J.-P. Mornon.** 1989. Cellulase families revealed by hydrophobic cluster analysis. Gene **81**:83–95.
16. **Kadam, S., A. L. Demain, J. Millet, P. Beguin, and J.-P. Aubert.** 1988. Molecular cloning of a gene for a thermostable β-glucosidase from *Clostridium thermocellum* into *Escherichia coli*. Enzyme Microb. Technol. **10**:9–13.
17. **Kramer, B., W. Kramer, and H.-J. Fritz.** 1984. Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*. Cell **38**:879–887.
18. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurements with the Folin phenol reagent. J. Biol. Chem. **193**:265–275.
19. **Park, J. T., and M. S. Johnson.** 1949. A submicrodetermination of glucose. J. Biol. Chem. **181**:149–151.
20. **Petitdemange, E., F. Caillet, J. Giallo, and C. Gaudin.** 1984. *Clostridium cellulolyticum* sp. nov., a cellulolytic mesophilic species from decayed grass. Int. J. Syst. Bacteriol. **34**:155–159.
21. **Pilz, I., E. Schwarz, D. G. Kilburn, R. C. Miller, Jr., R. A. J. Warren, and N. R. Gilkes.** 1990. The tertiary structure of a bacterial cellulase determined by X-ray-scattering analysis. Biochem. J. **271**:277–280.
22. **Tomme, P., H. Van Tilbeurgh, G. Pettersson, J. Van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Clayssens.** 1988. Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Eur. J. Biochem. **170**:575–581.
23. **Van Tilbeurgh, H., P. Tomme, M. Clayssens, R. Bhikhabhai, and G. Pettersson.** 1986. Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*. FEBS Lett. **204**:223–227.
24. **Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon.** 1985. Nucleotide sequence of the AIDS virus, LAV. Cell **40**:9–17.
25. **Yagüe, E., P. Béguin, and J.-P. Aubert.** 1990. Nucleotide sequence and deletion analysis of the cellulase-encoding gene *celH* of *Clostridium thermocellum*. Gene **89**:61–67.
26. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.