0273-2289/91/3102—0119$03.20

Cloning and Expression of a *Clostridium thermocellum* DNA Fragment that Encodes a Protein Related to Cellulosome Component S_L

M. P. M. ROMANIEC,[1] T. KOBAYASHI,[2] U. FAUTH,[3] U. T. GERNGROSS, AND A. L. DEMAIN*

Fermentation Microbiology Laboratory, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; [1]Present address: Biological Laboratory, University of Kent, Canterbury, Kent CT2 7NJ, UK; [2]Kao Corporation, Tochigi Research Laboratories, Tochigi 321-34, Japan; and [3]Present address: Xenova Ltd, Slough, Berks SL1 4EQ, UK

Received March 29, 1991; Accepted April 30, 1991

ABSTRACT

Antibodies raised against the S_L subunit of the *Clostridium thermocellum* cellulosome were used to screen a library of *C. thermocellum* chromosomal DNA fragments constructed in the vector λgt11. A DNA fragment that encoded a polypeptide that crossreacted with the anti-S_L antibodies was isolated and its restriction map elucidated. No similarity with other previously cloned DNA fragments has been found. The anti-S_L crossreacting polypeptide was isolated from recombinant *Escherichia coli* and found to have a mol mass of 37,000 Da and to possess low levels of CMCase and Avicelase activity. Using CMC as the substrate, a temperature optimum of 55°C and a pH optimum of 6.6 were observed. These properties were compared to those of *C. thermocellum* S_L isolated by electroelution from an SDS gel, which was also found to possess low levels of CMCase and Avicelase activities. In addition, the S_L proteins produced in *C. thermocellum* and *E. coli* were able to interact positively against Avicel with an endoglucanase (S_S) purified from the *C. thermocellum* crude cellulase preparation, and with a recombinant protein that crossreacted with anti-S_S antibodies.

*Author to whom all correspondence and reprint requests should be addressed.

Index Entries: *Clostridium thermocellum;* gene cloning; protein isolation; cellulase; cellulosome.

INTRODUCTION

The anaerobic thermophile *Clostridium thermocellum* secretes a highly active cellulolytic complex that is capable of degrading crystalline cellulose, such as cotton and Avicel (*1,2*). A detailed study of its cellulase system is warranted because of the potential use of this bacterium in the direct conversion of cellulosic biomass to chemical feedstocks and ethanol (*3*). A cellulolytic structure, termed the cellulosome (*4*), exists on the extracellular surface of *C. thermocellum* vegetative cells (*5*) and is also found free in the extracellular medium. Cellulosomes have molecular masses ranging from 2 to 6.5×10^6 Da (*6–8*) and can aggregate into macromolecular complexes termed polycellulosomes, which range in size from 5 to 8×10^7 Da (*7,9*). These polycellulosomal complexes exist on the surface of *C. thermocellum* cells as discrete structures (protuberances) (*8*). Cellulosomes of different strains are reported to contain 14–50 proteins when analyzed by sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE). These proteins range in mol mass from 20 to 250 kDa (*7,8*). The major portion of the cellulase activity of *C. thermocellum* resides within the cellulosomal complexes (*6,8*).

Isolation of individual components from undenatured cellulosomes has proven to be extremely difficult. Thus far, these complexes have been dissociable only by SDS (*4,9,10*). The cloning of 28 or so individual genes encoding carboxymethylcellulases (CMCases), xylanases, and β-glucosidases from *C. thermocellum* into *Escherichia coli* has facilitated isolation of genes and production of individual cellulosome components (*11,12*). Our group has previously shown that the combination of two cellulosomal proteins, namely S_L and S_S, could hydrolyze crystalline cellulose (Avicel) to cellobiose, albeit at very low rates, since the proteins were isolated using denaturing conditions (*10*). S_L, the largest cellulosomal component, is glycosylated and contains 25–40% carbohydrate (*10,13*),which is predominantly galactose. S_L has a mol mass of 250–210 kDa (*8,10*), and is apparently the same as S_1 in the terminology of Lamed et al. (*4*). S_L was previously thought to have no enzymatic activity (*10*); however, the present paper reveals that it possesses some CMCase acatívity and is able to release small amounts of cellobiose from Avicel.

We now describe the isolation of a DNA fragment from *C. thermocellum*, encoding a protein that crossreacts with an antibody raised against its S_L component, and the preliminary characterization of the recombinant derived protein and its comparison with the *C. thermocellum* S_L protein.

MATERIALS AND METHODS

Bacterial Strains and Media

C. thermocellum ATCC 27405 was cultured anaerobically in CM3 medium (*14*), containing 0.5% (w/v) Avicel (PH 105 FMC, Marcus Hook, PA) as the carbon source, under strictly anaerobic conditions. *E. coli* Y1090 (*15*) and *E. coli* HB101 (*16*) were cultured in LB broth (*17*).

Isolation of DNA

Total DNA from *C. thermocellum* cultured in CM3 medium was isolated according to the method of Marmur (*18*), modified to include treatment with proteinase K (1 mg/mL; Sigma, St. Louis, MO) and RNAase (100 μg/mL; type II-A; Sigma). Bacteriophage λgt11 DNA was prepared from *E. coli* Y1090 using the large-scale method of purification (*17*). Large-scale preparations of plasmid DNA were made from *E. coli* HB101 by alkaline lysis (*17*). The method of Birnboim and Doly (*19*) was used for small-scale extraction and rapid screening of plasmid-containing clones.

Preparation of Anti-S_L Polyclonal Antibodies

The S_L protein of *C. thermocellum* was isolated by electrophoresis under denaturing conditions using SDS-PAGE and subsequent electroelution. The source of S_L was a *C. thermocellum* ATCC 27045 cellulase preparation resolved over a Sepharose 6B (Pharmacia, Piscataway, NJ) column; two peaks were observed. Peak I was used as the starting material. The peak I material was made up to 0.1% (w/v) with SDS and heated at 60°C for 30 min before loading onto an SDS-polyacrylamide gel. The gel was stained using ice-cold 0.25*M* KCl plus 1 m*M* dithiothreitol (DTT) for 30 min, and destained in ice-cold 1 m*M* DTT for 30 min (*20*). The S_L band was identified and cut out of the gel. The protein was electroeluted into 1*M* NaCl in a Pharmacia Extraphor unit. Before injection into a rabbit, the S_L protein was heated at 100°C for 15 min. Elution was done with a buffer containing 192 m*M* glycine, 0.1% SDS, and 25 m*M* Tris base. The buffer was removed by desalting over a 10DG (Bio-Rad Laboratories, Richmond, CA) desalting column. For the assay of cellulase, S_L was passed over an Extracti-Gel D column (Pierce, Rockford, IL) to remove SDS. The electroeluted S_L protein was injected into a white rabbit using the Ribi Adjuvant System (Ribi Immunochem Research, Hamilton, MT). A booster injection was done 12 d after the first injection and the rabbit was bled 7 d after the booster injection. The antiserum was centrifuged for 15 min at 11,000*g* at 4°C. The supernatant fluid was desalted over a 10DG desalting column and purified using a serum IgG purification column (Bio-Rad Laboratories).

Isolation of S_L DNA Fragment

For the construction of a genomic library, *C. thermocellum* DNA was partially digested with *Eco*RI and fragments in the 2- to 7-kb size range were recovered by electroelution after horizontal agarose gel electrophoresis. The DNA, partially digested with *C. thermocellum*, was ligated into the λgt11 vector using T_4 DNA ligase. The recombinant phage was packaged and transfected into the host strain *E. coli* Y1090, and plaques were generated using the Protoclone GT System (Promega Biotec, Madison, WI). The plaques were screened using anti-S_L antibodies as the primary antibody and alkaline phosphatase (AP) conjugated anti-rabbit IgG (Fc) antibodies as the secondary antibody. The blotting and immunodetection was based on the ProtoBlot Immunoblotting System (Promega). Presumptive "positive" plaques were isolated and purified using sequential plaque purification and immunodetection until all plaques on the plate gave a positive color reaction.

Digestion with Restriction Enzymes

Single- and double-enzyme digestions of recombinant bacteriophage and plasmids isolated from S_L-positive clones were carried out using the restriction endonucleases *Ava*I, *Bam*HI, *Eco*RI, *Hin*dIII, *Pst*I, *Pvu*II, *Sal*I, *Sma*I, and *Xho*I. Conditions used were those specified by the supplier of each enzyme. DNA fragments were fractionated by electrophoresis in 1% (w/v) agarose gels and their sizes estimated by reference to standards.

Southern Blot Hybridization

Chromosomal DNA from *C. thermocellum* was digested with restriction endonucleases as required and fractionated on 1% (w/v) agarose gels. DNA was then blotted on Hybond N (Amersham, Chicago, IL). ^{32}P-Labelled DNA probes (recombinant plasmids or restriction fragments therefrom) were prepared by nick translation (*21*) using α-^{32}P-dATP (Amersham). Hybridization was carried out as described by Southern (*22*).

Preparation of Cell-Free Extracts

E. coli clones containing recombinant plasmids were cultured on the shaker in baffled flasks for 16 h at 37°C in LB broth (100 mL) containing ampicillin (100 μg/mL). Cells were collected by centrifugation (4000*g*, 4°C, 10 min), washed once with ice-cold PC buffer (50 m*M* K_2HPO_4, 12 m*M* citric acid, pH 6.5) and resuspended in the same buffer to a final vol of 4 mL. The cells were disrupted by sonication at 4°C (Model 225 ultrasonic processor with microtip, setting output 7, 50% duty cycle, two applications of 5 min; Heat Systems-Ultrasonic Inc., Danbury, CT), and then centrifuged for 30 min at 10,000*g* and 4°C. The supernatant fluid was collected and warmed to 60°C and held at that temperature for 10 min.

The heat-denatured material was centrifuged for 30 min at 40,000*g* to remove debris. The clarified extract was stored at −20°C.

Enzyme Assays

All assays were performed in succinate buffer (60 m*M* sodium succinate, pH 5.8) at 60°C (unless otherwise stated). CMCase activity was measured according to Wu et al. (*10*). One unit of activity is defined as the amount of enzyme that releases 1 μmol of glucose equivalent per minute. Hydrolysis of Avicel (PH 105, FMC) was determined by the release of reducing sugar, measured using the residual-ferricyanide method (*23*). One unit of activity is the amount of enzyme which releases 1 μmol glucose equivalent per min. Hydrolysis of *p*-nitrophenyl-β-D-cellobioside was examined using 4 m*M* substrate in a total vol of 1.5 mL for 20 h. Color was developed by addition of 1.5 mL of 1*M* sodium carbonate, and absorbance was determined at 400 nm. Hydrolysis of cellodextrins was performed for up to 180 h in 1 mL of succinate buffer containing 1 mg/mL substrate. Samples were deionized by shaking with 100 mg of Bio-Rad RG501 x8 resin at 37°C for 30 min. Deionized material was resolved by thin-layer chromatography (TLC) on precoated silica-gel plates (F_{254}: Merck) using butanol/acetic acid/ether/water (5/6/3/3 by vol) as developing solvent. To locate products, plates were sprayed with 1% (w/v) 3:4-dinitrobenzoic acid in 1*M* Na_2CO_3 before heating at 160°C for 10 min.

Purification of Recombinant-Derived S_L

Unless otherwise stated, all manipulations were performed at 4°C under aerobic conditions. Column chromatography was performed on a Pharmacia FPLC system using Pharmacia chromatography media developed specifically for this purpose. An overnight culture of *E. coli* HB101 (pRV4.3) (5 L) grown in LB was centrifuged (10,000*g*, 15 min, 4°C). The pellet was resuspended in ice-cold PC buffer to a total vol of 65 mL, to which was added 1.5 mg RNAase I (Sigma) and 1.5 mg DNAase I (Sigma). The bacteria were disrupted by sonication for 10 min, and then centrifuged for 20 min at 14,000*g* and 4°C. The supernatant fluid is referred to as the crude extract. The crude extract (60 mL) was transferred to a 500-mL Erlenmeyer flask and heated with gentle swirling to 60°C and held at this temperature for 10 min. The flask was allowed to cool on ice, and the suspension was centrifuged for 20 min at 14,000*g* at 4°C. A 10% (w/v) streptomycin sulfate solution in PC buffer was added (10 mL/g of protein in the heat-treated protein). After stirring for 45 min at room temperature, the precipitate was removed by centrifugation (14,000*g*, 20 min, 4°C). Ammonium sulfate was added to the supernatant fluid to 70% of saturation. After stirring for 30 min at room temperature, the suspension was centrifuged (14,000*g*, 30 min, 16°C) and the precipitate was saved. Ammonium sulfate was added to the supernatant fluid to 100% of saturation. After

stirring for 1 h at room temperature, the suspension was centrifuged (14,000*g*, 30 min, 16°C). The two precipitates were pooled and dissolved in 10 mL of 20 m*M* Tris-HCl, pH 7.0. The solution was dialyzed against 2 L of 20 m*M* Tris-HCl, pH 7.0, for 24 h with one change of buffer. The precipitate was removed by centrifugation (20,000*g*, 30 min, 4°C). The supernatant fluid was adsorbed to a column (bed vol, 20 mL) of Q Sepharose FF previously equilibrated with 20 m*M* Tris-HCl, pH 7.0. The protein was eluted with a linear gradient of NaCl (0–1M) in the same buffer at a flow rate of 2mL/min. Positive fractions that eluted between 300 and 500 m*M* NaCl (assayed by crossreaction with anti-S_L antibody using a dot-blot format), were pooled and dialyzed overnight against 20 m*M* Tris-HCl, pH 7.0. The dialyzed sample was loaded onto a Mono Q HR5/5 column via a 50-mL Superloop. The protein was eluted with a linear $(NH_4)_2SO_4$ (0–200 m*M*) gradient in 20 m*M* Tris-HCl, pH 7.0, at a flow rate of 0.75 mL/min. Positive fractions eluted between 60 and 100 m*M* $(NH_4)_2SO_4$; these were pooled and $(NH_4)_2SO_4$ was added, to a concentration of 1.7*M*. The pooled material was adsorbed onto an Alkyl-Superose HR5/5 column. The protein was eluted utilizing a linear gradient (1.7–0*M*) of $(NH_4)_2SO_4$ in 20 m*M* Tris-HCl, pH 7.0, at a flow rate of 0.5 mL/min. The positive fractions, which eluted between 1.25*M* and 0.9*M* $(NH_4)_2SO_4$, were pooled anddialyzed against 20 m*M* Tris-HCl, pH 7.0.

The dialyzed material was adsorbed onto a Mono P HR5/5 column. The protein was eluted using a combined pH and salt gradient, i.e., 20 m*M* Tris-HCl, pH 7.0, and 50 m*M* triethylamine, pH 8.3, 1*M* NaCl. Positive fractions, which eluted at 400 m*M* NaCl, pH 7.35, and 500 m*M* NaCl, pH 7.5, were pooled and dialzyed overnight against 20 m*M* Tris-HCl, pH 7.0. The dialyzed sample was adsorbed back onto a Mono Q HR5/5 column. The column was eluted with a linear gradient of KSCN (0–1*M*) in 20 m*M* Tris-HCl, pH 7.0. Positive fractions, which eluted between 280 m*M* and 400 m*M* KSCN, were pooled and desalted over a Sephadex G-25 minicolumn (Pharmacia). The desalted material was pooled and lyophilized.

Protein S_S Preparations

C. thermocellum S_S protein was isolated from the culture supernatant fluid (*23a*). The recombinant-derived S_S protein was purified from *E. coli* extracts; the procedure will be published separately (Kobayashi et al., submitted).

Isoelectric Focusing

The recombinant-derived S_L protein was isoelectrically focused on Phast Gel IEF Medium (Pharmacia), using the broad-range pI calibration standards.

Protein Determinations

Protein was determined by the method of Bradford (*24*) using bovine serum albumin (fraction V) as standard.

Gel Electrophoresis of Proteins

Proteins were separated by SDS-PAGE using the method of Laemmli (*25*).

Western Blotting and Dot Blotting of Proteins

Western blotting from SDS-PAGE gels onto nitrocellulose membranes was done by the method of Towbin et al. (*26*). For dot blotting, samples from cell-free extracts or gel-chromatography fractions were dotted onto nitrocellulose membrane and allowed to dry in air. The membranes were blocked for 30 min at room temperature with TBST (10 m*M* Tris-HCl, pH 8.0; 0.15*M* NaCl; and 0.05% Tween 20) supplemented with 3% (w/v) nonfat dried milk (*27*). The antibody binding and the development of bands using donkey antirabbit serum conjugated to horseradish peroxidase (Amersham) were done according to the method of Rybicki and von Wechmar (*28*), except that TBST was used throughout.

Lectin Blotting

Blotting of proteins from SDS-PAGE gels onto nitrocellulose membranes was done as above (*26*). Glycoproteins were detected using the lectin Jacalin (Vector Labs, Burlingame, CA) and visualized using the Vectastain detection system (Vector Labs).

RESULTS

Cloning of the S_L DNA Fragment from *C. thermocellum* in *E. coli*

A total of 32 positive plaques were identified from a library of 11,200 plaques. Of the 32, eight were randomly chosen. After plaque purification, recombinant λDNA was isolated, and the inserts cut out with *Eco*RI. All eight clones were found to contain three identically sized fragments. Southern blotting showed that the three fragments cross-hybridized with the respective fragment in all eight clones. This strongly suggested that the eight clones carry a common *Eco*RI-ended fragment. The complete insert from one clone was subcloned into plasmid pBR325 for restriction mapping. The resultant clone was called plasmid pRV4.3, and pRV4.4 when the insert was carried in the opposite orientation.

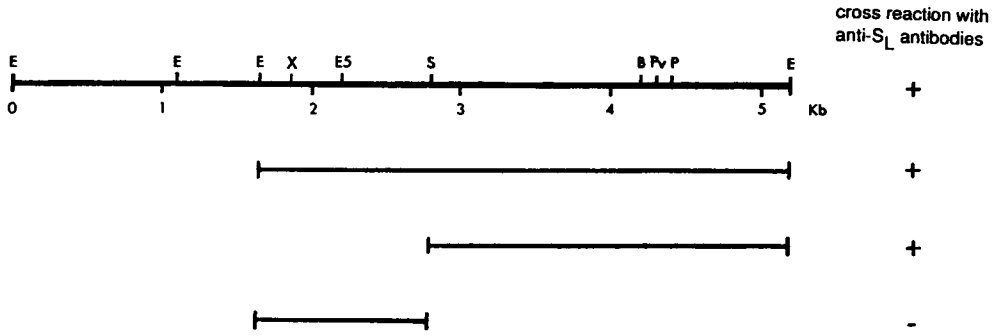

Fig. 1. Restriction map of pRV4.3 the 5.2-kb *Eco*RI–*Eco*RI DNA fragment from *C. thermocellum* cloned in pBR325 and the subcloned derivatives thereof. Shown in the last column are their ability, when introduced into *E. coli*, to express a protein that crossreacts with anti-S_L antibodies. B: *Bam*HI; E: *Eco*RI; P: *Pst*I; Pv: *Pvu*II; S: *Sma*I; X: *Xho*I.

Restriction Map of pRV4.3

The restriction map of plasmid pRV4.3 is shown in Fig. 1. No *Hin*dIII or *Sal*I sites were found. Southern blot hybridization was performed with total DNA from *C. thermocellum* digested with *Ava*I, *Bam*HI, *Eco*RI, *Pst*I, *Pvu*II, *Sma*I, and *Xho*I. The pattern of hybridization confirmed the physical map obtained for the insert in plasmid pRV4.3. This strongly suggested that the insert was a genuine *C. thermocellum* DNA fragment that had not been modified during the construction of the library. By subcloning (Fig. 1) we found the active portion of the fragment to lie between the *Eco*RI and the *Sma*I site. Plasmid pRV4.4 also was able to confer upon *E. coli* the ability to express a protein crossreacting with anti-S_L antibodies.

Purification of Recombinant S_L

The protein yields at various stages of the purification are summarized in Table 1, and analysis of various fractions by SDS-PAGE is shown in Fig. 2. Purification amounted to 137-fold. Recombinant S_L represented only a very small part of the total soluble proteins of *E. coli*. In addition, the enzyme was relatively difficult to purify, even though the heat-denaturation step eliminated 70% of the *E. coli* proteins, with a minimal loss of the recombinant protein.

Western Blotting of *C. thermocellum* S_L and Recombinant S_L

The anti-S_L antibodies had been raised against *C. thermocellum* glycosylated S_L which had been electroeluted from an SDS-polyacrylamide gel. To confirm that the antibodies did in fact crossreact with the recombinant S_L when it was resolved by SDS-PAGE (as opposed to a dot blot), a Western blot was performed against a crude cellulase sample from *C. thermo-*

Table 1
Purification of Recombinant S_L Produced by *E. coli* HB101 (pRV4.3)

Stage of purification	Total protein, mg	CMCase specific activity, U/mg	CMCase total units	Yield, (%)	Purification (fold)
Crude extract	1700	0.0003	0.51	100	1.0
Heat-treated extract	530	0.0009	0.48	94	3.0
Ammonium sulfate precipitate, dialyzed	325	0.0011	0.34	67	3.7
Q Sepharose	72	0.0024	0.17	33	8.0
Mono Q	18	0.0061	0.11	21	20
Alkyl Superose	4.3	0.010	0.043	8	33
Mono P	1.7	0.023	0.039	7	77
Mono Q, desalted	0.3	0.041	0.012	2	137

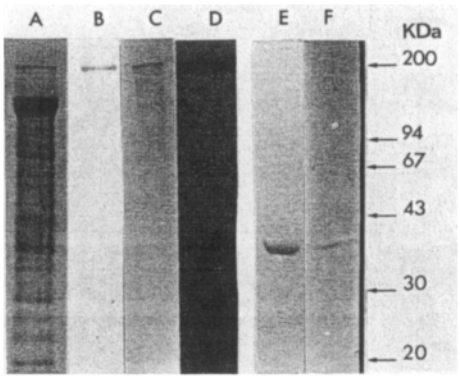

Fig. 2. SDS-PAGE gels: **A,** Coomassie blue stain of crude *C. thermocellum* cellulase (20 μg); **B,** Western blot of crude cellulase using anti-S_L antibodies; **C,** Lectin blot (Jacalin) of crude cellulase; **D,** Coomassie blue stain of pure electro-eluted S_L protein; **E,** Coomassie blue stain of recombinant S_L protein; **F,** Western blot of pure recombinant S_L using anti-S_L antibodies.

cellum, strain ATCC 27045, and the purified recombinant S_L fraction. As shown in Fig. 2, the antibodies are specific to the *C. thermocellum* S_L protein and crossreact with the homologous recombinant S_L protein. There was no crossreaction with *E. coli* protein. This confirms that the S_L gene, or at least part of it, had been cloned from *C. thermocellum.*

Properties of Recombinant S_L

By SDS-PAGE, the recombinant protein was calculated by have an M_r of 37,000 Da. With CMC as the substrate, the optimum pH is 6.6 (Fig. 3). By isoelectric focusing, recombinant S_L was found to have a pI of 5.1 (data

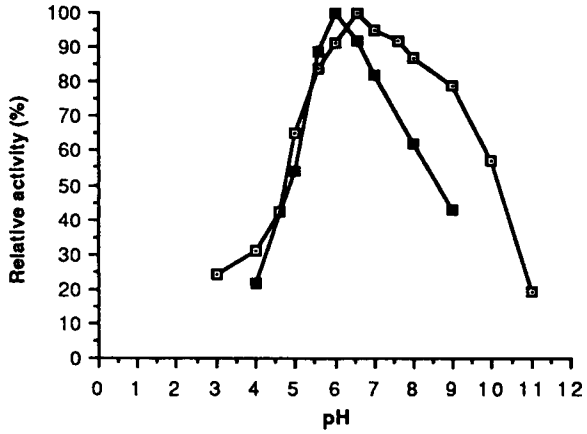

Fig. 3. The effect of pH on the CMCase activity of recombinant S_L (⊡) and *C. thermocellum* S_L (■). Assays were performed in 50 m*M* buffer; pH 3–7.6, phosphate-citrate; pH 8.0–9.0, Tris-HCl; pH 9.0–11.0, sodium carbonate. CMC was at 0.61% (w/v). Either 5 μg of recombinant S_L or 2 μg of *C. thermocellum* S_L was used. The incubation was carried out for 12 h at 60°C.

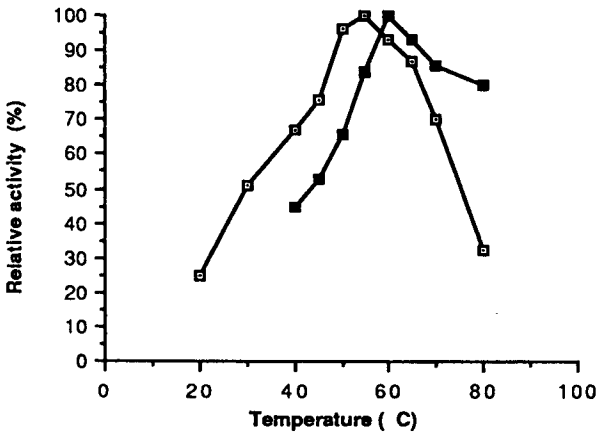

Fig. 4. The effect of temperature on CMCase activity of recombinant S_L (⊡) and *C. thermocellum* S_L (■). Assays were performed in 50 m*M* phosphate-citrate buffer, pH 6.0, containing 0.61% (w/v) CMC and 5 μg of recombinant S_L or 2 μg of *C. thermocellum* S_L. The incubation time was 12 h.

not shown). Using CMC as substrate, the optimum temperature was found to be 55°C (Fig. 4). The specific activity of recombinant S_L on CMC is 0.042 U/mg. In addition, recombinant S_L releases small amounts of cellobiose from Avicel, i.e., 20 μg of enzyme led to the release of 60 μg of cellobiose in 5 d. Recombinant S_L had no detectable activity on cellobiose or cellotriose, whereas both cellotetraose and cellopentaose were hydrolyzed, albeit at very low rates. Cellotetraose was hydrolyzed to cellobiose and cellopentaose to cellobiose and cellotriose. No hydrolysis was observed with *p*-nitrophenyl-β-D-cellobioside (data not shown).

Table 2
Action of S_L, Recombinant S_L (RS_L), S_S and Recombinant S_S (RS_S) on Avicel

Enzymes	Quantities, μg	Cellobiose released, μg
S_L	20	100
RS_L	20	60
S_S	20	0
RS_S	5	0
	15	0
	45	0
	100	0
$S_L + RS_S$	20+15	150
$S_L + S_S$	20+10	180
$RS_L + S_S$	20+10	65
$RS_L + RS_S$	5+15	35
	10+15	90
	20+15	170

S_L and S_S were prepared from *C. thermocellum*. S_L was eluted from an SDS gel. S_S was prepared without the use of denaturants (Fauth et al., 1991). RS_L and RS_S were isolated from *E. coli*. Assay mixtures contained 1 mg of Avicel in 60 m*M* succinate buffer, pH 5.8, in a total vol of 5 mL. Proteins were preincubated for 30 min at 37°C before addition into assay medium. Incubated for 5 d at 60°C, reducing sugar was measured using the ferricyanide assay with cellobiose as standard; no sugars other than cellobiose were observed by TLC.

Characterization of Electroeluted S_L from *C. thermocellum*

S_L was determined to have a mol mass of approx 210,000 Da (Fig. 2) and to be glycosylated. Using CMC as substrate, electroeluted purified S_L was found to have a pH optimum of 6.0 and a temperature optimum of 60°C (Figs. 3 and 4, respectively). The activity profiles for both temperature and pH are roughly similar to those found for recombinant S_L. Against CMC, S_L had a specific activity of 0.060 U/mg, and 20 μg was able to release 100 μg of cellobiose from Avicel over a 5-d period. Although hydrolysis of cellodextrins was not examined, there was no breakdown of *p*-nitrophenylcellobioside.

Properties of S_S Preparations

The recombinant-derived S_S protein was isolated from *E. coli* (Kobayashi et al., submitted). It was found to have only a small amount of activity against CMC, i.e., a specific activity of 0.02 U/mg. No activity could be found against Avicel. Similarly, *C. thermocellum* S_S had no activity on Avicel (Table 2).

Reconstitution Experiments

From Table 2 it can be seen that both *C. thermocellum* S_L and recombinant S_L are able to release cellobiose from Avicel, whereas *C. thermocellum* S_S and recombinant S_S were found to be completely inactive. When *C. thermocellum* S_L was combined with *C. thermocellum* S_S or recombinant S_S, there was an increase in the release of cellobiose. When recombinant S_L was combined with recombinant S_S, the increase in the amount of cellobiose was of the same order as that of *C. thermocellum* S_L plus *C. thermocellum* S_S. However, the addition of *C. thermocellum* S_S to recombinant S_L did not result in any increase in cellobiose release over that of recombinant S_L alone. The amount of cellobiose released declined with a decrease in the amounts of recombinant S_L added to the reaction.

DISCUSSION

Plasmids pRV4.3 and pRV4.4, which contain the same insert cloned in opposite orientation in pBR325, were each found to confer upon *E. coli* the same ability to crossreact with anti-S_L antibodies. Therefore, it appears that S_L can be expressed from a *C. thermocellum* DNA fragment that is recognized by *E. coli* RNA polymerase. Expression of *C. thermocellum* genes in *E. coli* is now widely recognized (*11,12,29*).

Southern blot analysis indicated that the DNA was cloned without rearrangement. In addition, the insert hybridized to a single DNA fragment of the *C. thermocellum* chromosome. The restriction map can be considered unique when compared to previously cloned fragments (*11,12,29*). However, some homology might be found after the nucleotide sequence of the S_L gene is determined. It has been reported that a number of *C. thermocellum* CMCase genes that do not show homology by cross-hybridization do in fact share common repeated regions (*30–33*). Nucleotide sequencing of the S_L gene is currently underway.

Western blot analysis demonstrated that the antibodies raised against the *C. thermocellum* glycosylated S_L protein (210,000 Da) crossreact with a recombinant-derived protein isolated from *E. coli*, which has a mol mass of 37,000 Da. Thus a considerable difference in molecular mass exists between the *C. thermocellum* S_L protein and the putative recombinant S_L protein. There are a number of possibilities that might explain this discrepancy. SDS-PAGE is a rapid technique for the determination of the molecular weights of proteins (*34*), and an accurate determination depends on two factors, as follows: (a) Proteins in general bind constant amounts of SDS per gram (*35*). When bound, the protein has an overall negative charge that masks its intrinsic charge, resulting in a constant charge-to-mass ratio for most proteins (*35*). (b) Proteins saturated with SDS take on a rodlike configuration, the length of the structure being pro-

portional to its polypeptide chain length, and thus its molecular weight (*35*). However, the SDS-PAGE procedure is not directly applicable to molecular-weight determinations of glycoproteins. Glycoproteins containing more than 10% carbohydrate behave anomalously during SDS-PAGE when compared to standard proteins (*36*). S_L is reported to contain between 25 and 40% carbohydrate (*13*). The cause of this anomalous behavior is a decreased binding of SDS per gram of glycoprotein (*37*). The lower binding of SDS results in a decreased charge-to-mass ratio for glycoproteins, a decreased mobility during SDS gel electrophoresis, and thus a higher apparent molecular weight (*37*). Therefore, the molecular mass estimated for S_L (210,000 Da) is an overestimate). However, whether this alone is able to explain the large discrepancy in molecular weight is unclear. A previous attempt to remove the carbohydrate from S_L chemically resulted in extensive hydrolysis of the protein core (*5*).

It is possible that the recombinant S_L protein is undergoing at least partial hydrolysis in *E. coli*. All recombinant-derived *C. thermocellum* enzymes isolated from *E. coli* so far have been smaller than their corresponding native proteins when compared by Western blot analysis (*38*).

A third possible explanation is that the protein core of S_L is polymeric, made up of a number of 37,000-Da subunits. However, S_L, isolated from *C. thermocellum* ATCC 27045, has never been dissociated into smaller subunits, even by boiling in SDS and mercaptoethanol. It is possible that the monomers are held together by a very stable linkage not broken by boiling in SDS and mercaptoethanol, a situation previously reported for porin proteins in Gram-negative bacteria (*39,40*). Morag et al. (*40a*) found that the S_1 (=S_L) subunit in a crude cellulase preparation from *C. thermocellum* YS can be made to dissociate, when dialyzed in water or against buffer of low pH (pH 4.5) or low ionic strength (0.3 m*M* NaCl), to a protein with a mol mass of approx 40,000 Da; however, we could not confirm this finding. It is also possible that S_L is composed of a number of tightly-linked subunits, each of which is encoded by a different gene. This is unlikely, however, as the eight randomly selected λ clones all contained an identical insert. Had the antibodies been monoclonal, this would have been understandable. However, the polyclonal antibodies were raised against an extensively denatured S_L protein. We were also able to demonstrate that the antibodies do not react solely with the highly antigenic carbohydrate moieity, but can react with a protein, isolated from *E. coli*, that is not glycosylated.

A fourth possibility is that the protein core of S_L is in fact encoded by one gene, but in this case the complete gene was not isolated in pRV4.3. The results from the subcloning experiment suggest that the gene encoding S_L is carried on the proximal end of the fragment between the *Eco*RI and the *Sma*I sites. Deletions from the *Eco*RI site resulted in subclones expressing a protein that crossreacted only weakly with the anti-S_L antibodies. Elucidation of this nucleotide sequence will reveal whether or not the entire gene has been isolated.

We found that our electroeluted S_L protein from *C. thermocellum* contains both CMCase and Avicelase activities, albeit very weak. This differs from the result published by Wu et al. (*10*). When we resolved the S_L protein by SDS-PAGE and used the silver stain, no other protein bands were detected, suggesting a high level of purity, and that the enzyme activity was not caused by a contaminating protein. The most likely explanation is that the conditions used to isolate the S_L protein were less harsh than those used by Wu et al. (*10*) and did not include an acetone precipitation step. Our results would suggest that S_L is not merely an anchor protein to which other proteins attach to gain access to the cellulose, but also an enzyme in its own right.

Recombinant S_L was able to release reducing sugar from CMC, with a very low specific activity, much lower than for any previously isolated recombinant cellulase protein (*41,42*). In addition, it was able to hydrolyze both cellotetraose and cellopentaoase, releasing cellobiose. However, it did not hydrolyze cellobiose, indicating that the enzyme does not possess β-glucosidase activity. A zone of clearing was not produced on CMC plates stained with Congo red, contrary to most endoglucanases. More important, since only cellobiose was released from Avicel, it appears that recombinant S_L possesses exoglucanase activity. On the other hand, contrary to most exoglucanases, there was no hydrolysis of *p*-nitrophenylcellobioside. The cellobiose-releasing activity on Avicel, and the lack of it on *p*-nitrophenylcellobioside, is consistent with the activity found in electroeluted purified S_L from *C. thermocellum*.

The combination of *C. thermocellum* S_L with *C. thermocellum* S_S or with recombinant S_S resulted in an increase in the release of cellobiose from Avicel. In addition, there was also an increase in the amount of cellobiose released when the recombinant S_S protein was added to recombinant S_L protein. However, *C. thermocellum* S_S was not able to enhance Avicel hydrolysis in the presence of recombinant S_L. This would suggest that (a) glycosylation is not important for at least part of S_L's enzymatic activity, but could play a role in the binding of the *C. thermocellum* S_S enzyme and (b) at least one active component of S_L had been cloned and purified.

ACKNOWLEDGMENTS

This work was funded by National Science Foundation Grant EET-8711725 and US Army Research Office Contract DAAL03-88-K-0064. Support for T. Kobayashi was supplied by the Kao Corporation. M. P. M. Romaniec was a Merck Postdoctoral Fellow and U. Fauth was a scholar of the Deutsche Forschungsgemeinschaft. Tillmann Gerngross is thankful to Joseph Krainer for his encouraging support. We thank Claus von der Osten, Anthony Sinskey, and members of the Biotechnology Process Engineering Center for use of their facilities. We acknowledge the gift of cello-

dextrins from Robert Longin of the Pasteur Institute, Paris. We thank Sara Tonini for technical assistance and Daniel Liberman for encouragement.

REFERENCES

1. Ait, N., Creuzet, N. and Forget, P. (1979), *J. Gen. Microbiol.* **113,** 399–402.
2. Johnson, E. A., Sakajoh, M., Halliwell, G., Madia, A., and Demain, A. L. (1982), *Appl. Environ. Microbiol.* **43,** 1125–1132.
3. Avgerinos, G. C. and Wang, D. I. C. (1980), *Annu. Rep. Ferment. Proc.* **4,** 165–191.
4. Lamed, R., Setter, E., Koenig, R., and Bayer, E. A. (1983), *Biotechnol. Bioeng. Symp.* **13,** 163–181.
5. Lamed, R. and Bayer, E. A. (1988), *Adv. Appl. Microbiol.* **33,** 1–46.
6. Lamed, R., Setter, E., and Bayer, E. A. (1983), *J. Bacteriol.* **156,** 828–836.
7. Mayer, F., Coughlan, M. P., Mori, Y., and Ljungdahl, L. G. (1987), *Appl. Environ. Microbiol.* **53,** 2785–2792.
8. Lamed, R., and Bayer, E. A. (1988), *Biochemistry and Genetics of Cellulose Degradation,* Aubert, J. P., Béguin, P., and Millet, J., eds., Academic, New York, pp. 101–116.
9. Coughlan, M. P., Hon-Nami, K., Hon-Nami, H., Ljungdahl, L. G., Paulin, J. J., and Rigsby, W. E. (1985), *Biochem. Biophys. Res. Commun.* **130,** 904–909.
10. Wu, J. H., Orme-Johnson, W. H., and Demain, A. L. (1988), *Biochemistry* **27,** 1703–1709.
11. Hazelwood, G. P., Romaniec, M. P. M., Davidson, K., Grépinet, O., Béguin, P., Millet, J., Raynaud, O., and Aubert, J. P. (1988), *FEMS Microbiol. Lett.* **51,** 231–236.
12. Schwarz, W. H., Bronnenmeier, K., and Staudenbauer, W. L. (1985), *Biotechnol. Lett.* **7,** 859–864.
13. Gerwing, G. J., de Waard, P., Kamerling, J. P., Vliegenthart, J. F. G., Morgenstern, E., Lamed, R. and Bayer, E. A. (1989), *J. Biol. Chem.* **264,** 1027–1035.
14. Weimer, P. J. and Zeikus, J. G. (1977), *Appl. Environ. Microbiol.* **33,** 289–297.
15. Young, R. A. and Davis, R. W. (1985), *Proc. Natl. Acad. Sci. USA* **80,** 1194–1198.
16. Boyer, H. W. and Roulland-Dussoix, D. (1969), *J. Mol. Biol.* **41,** 459–472.
17. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982), *Molecular Cloning (A Laboratory Manual),* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
18. Marmur, T. (1961), *J. Mol. Biol.* **3,** 208–218.
19. Birnboim, H. C. and Doly, J. (1979), *Nucleic Acids Res.* **7,** 1513–1523.
20. Nolte, A. and Mayer, F. (1989), *FEMS Microbiol. Lett.* **61,** 65–72.
21. Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977), *J. Mol. Biol.* **113,** 237–251.
22. Southern, E. (1975), *J. Mol. Biol.* **98,** 503–517.
23. Park, J. T. and Johnson, M. J. (1949), *J. Biol. Chem.* **181,** 149–151.
23a. Fauth, U., Romaniec, M. P. M., Kobayashi, T., and Demain, A. L. (1991), *Biochem. J.* **278,** in press.
24. Bradford, M. M. (1976), *Anal. Biochem.* **72,** 248–254.
25. Laemmli, U. K. (1970), *Nature* **227,** 680–685.
26. Towbin, H., Staehelin, T., and Gordon, J. (1979), *Proc. Natl. Acad. Sci. USA* **76,** 4350–4354.

27. Johnson, D. A., Gautsch, J. W., Sportsman, J. R., and Elder, J. H. (1984), *Gene Anal. Tech.* **13,** 3–8.
28. Rybicki, E. P. and von Wechmar, M. B. (1982), *J. Virol. Methods* **5,** 267–278.
29. Sakka, K., Furuse, S., and Shimada, K. (1989), *Agric. Biol. Chem.* **53,** 905–910.
30. Béguin, P., Cornet, P., and Aubert, J. P. (1985), *J. Bacteriol.* **162,** 102–105.
31. Grépinet, O. and Béguin, P. (1986), *Nucleic Acid. Res.* **14,** 1791–1799.
32. Joliff, G., Béguin, P., and Aubert, J. P. (1986), *Nucleic Acid Res.* **14,** 8605–8613.
33. Hall, J., Hazelwood, G. P., Barker, P. J., and Gilbert, H. J. (1988), *Gene* **69,** 29–38.
34. Shapiro, A. L., Viñuela, E., and Maizel, J. V. (1967), *Biochem. Biophys. Res. Commun.* **28,** 815–820.
35. Reynolds, J. A. and Tanford, C. (1970), *J. Biol. Chem.* **245,** 5161–5165.
36. Bretscher, M. S. (1971), *Nature New Biol.* **231,** 229–232.
37. Segrest, J. P., Jackson, R. L., Andrews, E. P., and Marchesi, V. T. (1971), *Biochem. Biophys. Res. Commun.* **44,** 390–395.
38. Nakae, T., Ishii, J., and Tokunagu, M. (1979), *J. Biol. Chem.* **254,** 1457–1461.
39. Nikaido, H. and Nakae, T. (1979), *Adv. Microb. Physiol.* **19,** 163–250.
40. Béguin, P., Cornet, P., and Millet, J. (1983), *Biochemie* **65,** 495–500.
40a. Morag, E., Bayer, E. A., and Lamed, R. (1991), *Biochem. Biotechnol.* in press.
41. Joliff, G., Béguin, P., Juy, M., Millet, J., Ryter, A., Poljak, R., and Aubert, J. P. (1986), *Bio/Technology* **4,** 896–900.
42. Pétré, D., Millet, J., Longin, R., Béguin, P., Girard, H., and Aubert, J. P. (1987), *Biochemie* **68,** 687–695.

Note added in proof: Further work has shown pRV4.3 to encode only part of the S_L protein. The entire gene encodes a protein of M_r 196,800 Da.