Biochemistry (Moscow), Vol. 66, No. 7, 2001, pp. 808-813. Translated from Biokhimiya, Vol. 66, No. 7, 2001, pp. 991-998.

Exploring the Properties of Thermostable *Clostridium thermocellum* Cellulase CelE for the Purpose of Its Expression in Plants

R. M. Abdeev*, I. V. Goldenkova, K. A. Musiychuk, and E. S. Piruzian

Vavilov Institute of General Genetics, Russian Academy of Sciences, ul. Gubkina 3, Moscow, 119991 Russia; fax: (095) 132-8962; E-mail: roustam@vigg.ru

Received January 25, 2001
Revision received April 4, 2001

Abstract—The main properties (pH and temperature range, stability, substrate specificity) of the modified cellulase CelE (endo-β-1,4-glucanase) from *Clostridium thermocellum* have been analyzed with the goal of its expression in plants. The modified enzyme is similar to plant cellulases. Deletions in the N-terminus of the enzyme do not affect its biochemical properties. Based on the present investigation, we conclude that the modified β-1,4-glucanase CelEM1, when expressed in plants, will be a good model to study the role of cellulases in plants.

Key words: thermostable cellulase (β-1,4-glucanase), *Clostridium thermocellum,* plant cellulases

Endo-β-1,4-glucanases (endo-β-1,4-glucan-D-glycosylhydrolase, EC 3.2.1.4) or cellulases are enzymes that hydrolyze polyglucans with β-1,4-glycosidic linkages. Various pro- and eukaryotic organisms produce these cellulases. Endo-β-1,4-glucanases of bacteria and fungi as well as other hydrolytic enzymes are involved in degradation of natural cellulose [1]. There is a particular class of organisms, bacterial and fungus phytopathogens, which hydrolyze plant cell wall during the first stages of invasion [2]. The functions of cellulases are not restricted only hydrolysis of cellulose. The activation of endo-β-1,4-glucanases occurs at the different stages of growth ant development [3-5]. However, the mechanisms of the participation of cellulases in plant vital functions are not still clear.

There are difficulties in studying plant endo-β-1,4-glucanase functioning. To begin with, the existence of numerous isoforms of cellulases makes difficult the isolation of a studied cellulase from the background of other plant endo-β-1,4-glucanases. The problem can be solved using some approaches. One is the expression of bacterial thermostable glucanase with analogous activity in plants. Such enzymes are stable and active at high temperature when the plant enzymes are inactivated. Data obtained earlier in our laboratory support the idea of creating transgenic plants expressing thermostable glucanases [6-8].

The thermophilic bacterium *Clostridium thermocellum*, whose cellulolitic enzymes are well studied [9, 10], was chosen as the source of the endo-β-1,4-glucanase gene. This anaerobic bacterium produces several intercellular cellulases including endo-β-1,4-glucanase E (CelE) [11], which was chosen for the experiments.

Endo-β-1,4-glucanase E of *C. thermocellum* is a modular protein having domain structure. The bacterial enzyme consists of a leader signal, catalytic domain, flexible linker (Pro-Thr-box), and C-end cellulosome-binding domain (CBD) [11]. The enzyme is secreted into intercellular fluid via the leader signal. The function of the CBD is simply to deliver its resident catalytic domain to the cellulosome. The cellulosome is multi-subunit complex including different types of glycosyl hydrolases demonstrating synergism in the hydrolysis of crystalline cellulose [12, 13]. The full-length polypeptide consists of 814 amino acid residues; the molecular weight is 90 kD [11] (Fig. 1a).

Elimination of specific bacterial sequences in the *cel*E gene is required for its effective expression in plants. However, for the bacterial endo-β-1,4-glucanase to imitate the action of numerous plant glucanases it must possess wide substrate specificity for plant cell wall polymers, broad pH optimum, high stability, and activity at 25°C (the optimal temperature for plant growth) corresponded to activity of plant enzymes. The purpose of this investigation was to modify the sequence of the *C. thermocellum* endo-β-1,4-glucanase gene, to study the biochemical

* To whom correspondence should be addressed.

0006-2979/01/6607-0808$25.00 ©2001 MAIK "Nauka/Interperiodica"

properties of modified enzyme, to compare their biochemical properties with plant cellulases, and to assess the possibility of using the modified enzyme for its expression in plants for studying the role of endo-β-1,4-glucanase.

MATERIALS AND METHODS

Strains and plasmids. The bacterial strain *Escherichia coli* XL1-Blue (F′ proAB lacI^{q}D(lacZ)M15 Tn10(Tetr)/ recA1 endA1 gyrA96 (Nalr) thi1 hsdR17(rk-mk+) supE44 relA1 lac) and vectors pQE-30 and pQE-32 [14] were used.

Molecular cloning was done according to the standard protocols [15]. Restriction endonucleases, T4 DNA-ligase, DNA-polymerase I (Klenow fragment), and alkaline phosphatase were used according to protocols of their manufacturers (Promega, USA; Fermentas, Latvia).

Deletion variants. Deletion variant *cel*EM1 was obtained using polymerase chain reaction (PCR) with primers: 5′-GCGTGGATCCATGGGAACAAAGCTT-TTGGA-3′ and 5′-GCGAGGTACCTCACGGCAT-TAATGTAGGTGT-3′ (Syntol, Russia) and plasmid pUC110 that bear a fragment of the *cel*E gene from *C. thermocellum* as the template [16]. The amplified fragment (1124 kb) was digested with *Bam*HI and *Kpn*I and cloned in the plasmid pQE-30 (previously digested by the same endonucleases) to create pQE-celEM1.

To obtain the second deletion variant *cel*EM2, the *Hind*III-fragment of pQE-celEM1 was cloned in pQE-30 (previously digested by the same endonuclease) to construct pQE-celEM2. The orientation of the glucanase gene to the LacZ-promoter was tested by selection of the cellulase active clones.

Deletion variant *cel*EM3 was obtained by cloning the *Sph*I-*Kpn*I fragment from the plasmid pQE-celEM1 in the plasmid pQE-32 (previously digested by the same endonucleases) to create plasmid pQE-celEM3.

Preparation and purification of protein extracts. The modified genes *cel*EM1, *cel*EM2, and *cel*EM3 were inducibly expressed according to the protocols of QIAexpress, USA. The bacterial extracts for testing enzyme activity of the cellulases were obtained by cell lysis in 8 M urea followed by dialysis against 50 mM citrate-phosphate buffer, pH 6.0. The protein extracts were heated at 65°C for 1 h followed by centrifugation for 30 min (4000*g*, Beckman J2-21 centrifuge, Ja20 rotor, Beckman, USA).

Cellulase activity was estimated using lichenan (Sigma, USA), carboxymethylcellulose (CMC) (medium viscosity, Sigma), xylan (Merck, Germany), and laminarin (Sigma) as the substrates. Reducing sugars released from the substrate were determined as described in [17]. One unit of activity is defined as the quantity of enzyme releasing 1 μmol of reducing sugars (as glucose equivalent) per minute. The enzyme activity versus pH was determined in 50 mM citrate-phosphate buffer. The dependence of activity on temperature and enzyme stability was estimated at pH 6.5.

Protein content was measured according to Bradford [18] using the dye reagent (BioRad, USA) and bovine serum albumin as standard (Sigma).

Electrophoresis of proteins was performed in denaturing polyacrylamide gel according to Laemmli [19]. Zymograms were obtained by staining of a lichenan (0.1%) containing gel after protein separation [20], with several modifications. Proteins were separated by electrophoresis in a 10% or gradient 8-16% polyacrylamide gel in the presence of sodium dodecyl sulfate (SDS). Lichenan (0.1%) was added to the separating gel before polymerization. Before application of the sample, loading buffer was added to bacterial extracts (1-10 μg) and the samples were incubated at 65°C for 40 min.

Cellulase activity on plates was tested done according to Beguin [21].

Thin layer chromatography of hydrolysis products was done according to Lake and Goodwin [22].

RESULTS

A clone carrying the pCU110 plasmid and possessing β-1,4-glucanase activity was chosen from the thermophilic bacterium *C. thermocellum* gene bank created in the laboratory of G. N. Velikodvorskaya. Based on knowledge of the *cel*E gene sequence [11] and using the method of physical mapping, it was determined that the plasmid pUC110 contains a part of the full-length *cel*E gene. The pCU110 plasmid carries the sequence of the leader signal, catalytic domain, and Pro-Thr-box (Fig. 1a). Because we planned to use only the catalytic domain for further investigation, it was first necessary to estimate the borders of the catalytic domain.

The three-dimensional structure of the cellulase was computer modeled using the interactive bioinformatic system Swiss-Model (http://www.expasy.ch/) and programs ProModII, Gromos96, and Swiss-PdbViewer [23]. The algorithm for determination of three-dimensional structure is based on the minimization of free energy and on existing data about the structure of proteins homologous to object of modeling. The amino acid sequence of CelE protein has identity 40.66% with CelCCA cellulase of *Clostridium cellulolyticum* and 26.35% identity with CelC cellulase of *Clostridium thermocellum*.

It should be emphasized that Swiss-Model can be used to predict the borders of catalytic domains. In our case, the borders of the CelE catalytic domain were determined as amino acid residues 41-388, and the three-dimensional structure of the catalytic domain was pre-

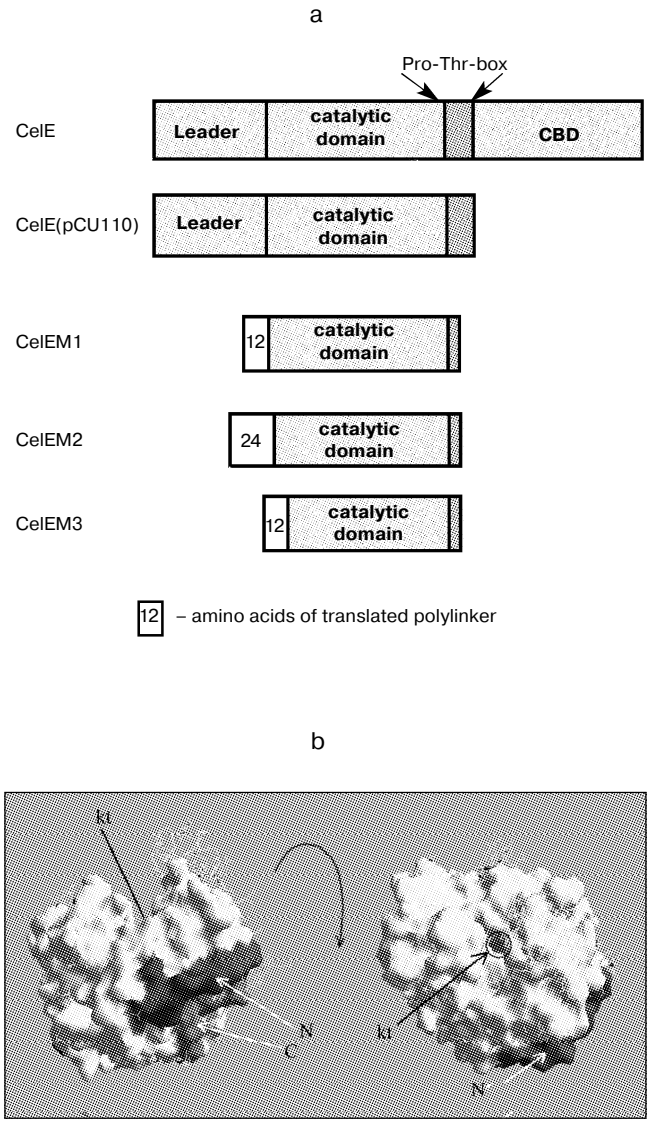

Fig. 1. a) Structure of full length CelE enzyme of *C. thermocellum* and its deletion variants (numbers are the position of amino acid as for the full length enzyme). b) Structural model of the catalytic domain CelE from *C. thermocellum* (kt, catalytic residues; N and C, NH_2- and COOH-termini of the catalytic domain).

dicted. The CelE catalytic domain has a cleft-like catalytic site [24] identical with Family 5 glycosylhydrolases. The catalytic amino acid residues are Glu193 and 316 that are present in the cleft (Fig. 1a).

To empirically check the theoretical prediction of the catalytic domain borders, the N-terminal region of the protein was modified. As a result, three deletion variants of the endo-β-1,4-glucanase, CelEM1, CelEM2, and CelEM3, were obtained (Fig. 1a). The sequence of CelEM1 consists of the catalytic domain from residue 35 with change of the first Ser to Met and part of the Pro-Thr-box (13 residues) (the numbering of amino acid residues is as for the full length protein). The native Pro-Thr-box consists of 22 residues. The second deletion variant of the cellulase, CelEM2, consists of the catalytic domain from residue 38. The third deletion variant, CelEM3, consists of the catalytic domain from residue 52.

The CelE deletion variants were expressed in *E. coli* using the pQE vector system. Together with the sequence of the catalytic domain, the translated genes contained additional amino acid residues from the translated polylinker. The sizes of the translated proteins CelEM1, CelEM2, and CelEM3 were 380, 388, and 363 residues, respectively.

It was first shown that the deletion variants of the cellulase demonstrate activity on lichenan, CMC, and xylan using the plate test method (Fig. 2). It should be emphasized that all the deletion variants have similar spot size and intensity of activity spot coloring.

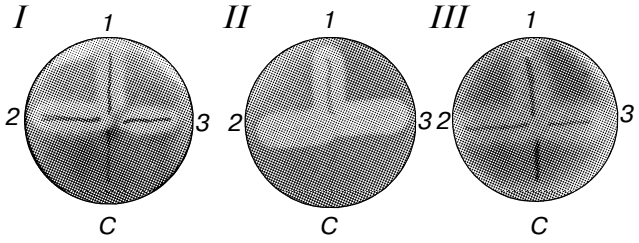

Fig. 2. Plate test of cellulase activity on carboxymethylcellulose (*I*), lichenan (*II*), and xylan (*III*). Strain of *E. coli* transformed by pQE-celEM1 (*1*), pQE-celEM2 (*2*), pQE-celEM3 (*3*), and pQE-30 (*C*) plasmids.

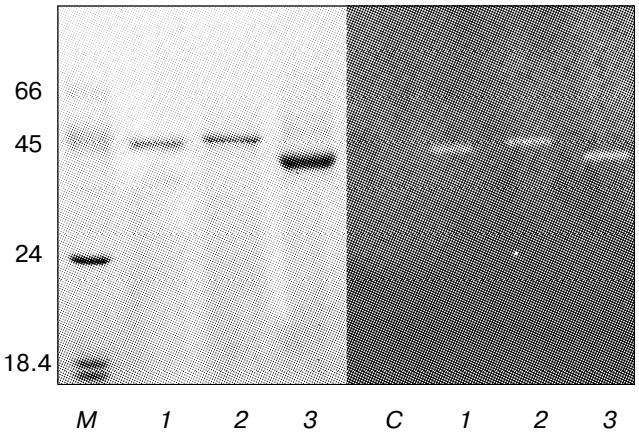

Fig. 3. Electrophoregram and zymogram of bacterial protein extracts over 0.1% lichenan as substrate. Lanes: *1*) CelEM1; *2*) CelEM2; *3*) CelEM3; *M*) molecular weight markers; *C*) control bacterial extract. Molecular weights of the markers are indicated on the left. The masses of the markers (in kD) are: bovine albumin, 66; egg albumin, 45; trypsinogen, 24; β-lactoglobulin, 18.4 (Sigma, USA).

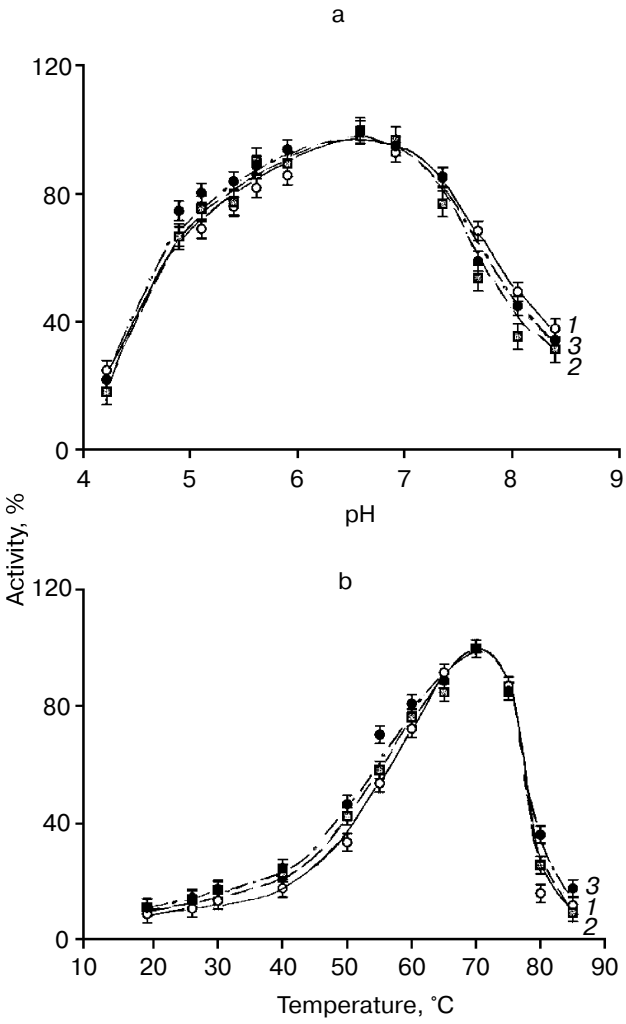

Fig. 4. Curves of dependence of specific activity of the deletion variants of cellulase on pH (a) and temperature (b): *1*) CelEM1; *2*) CelEM2; *3*) CelEM3.

The specific activity of CelEM1 estimated on different substrates was 124, 31, 6, and 0 U/mg on lichenan, CMC, xylan, and laminarin, respectively. The Michaelis constant on lichenan was 2.5 g/liter. Lichenan was used for further investigations. The specific activities of the deletion variants CelEM1, CelEM2, and CelEM3 were 124, 132, and 132 U/mg protein, respectively.

The deletion variants of the cellulase were analyzed by the zymogram method. The bands of activity of CelEM1, CelEM2, and CelEM3 are in agreement with theoretical predicted molecular masses of the enzymes, 42.8, 43.7, and 41.2 kD, respectively (Fig. 3).

Further, the pH optimum, temperature optimum, and thermostability of the deletion variants were investigated (Figs. 4 and 5). The curves of dependence of the relative activity versus pH have similar form for all variants of the protein. The enzyme retains 50% activity over a wide range of pH values (4.7-8.0), and the pH optimum is 6.5 (Fig. 4a). The dependences of the activities of the cellulase deletion variants on temperature are presented in Fig. 4b. The shape of curves and the temperature at which the enzymes show maximal activity under the chosen condition (70°C) are not significantly different. The stabilities of the deletion variants without substrate at different temperatures are presented in Figs. 5a and 5b. The curves of dependence of relative activity on incubation time at different temperature have similar form, and maximum stability was observed at 65°C.

Thin layer chromatography was used to analyze the products hydrolysis of lichenan, CMC, and xylan by cellulase (Fig. 6). The products of hydrolysis of the tested substrates have similar size. The smallest products of hydrolysis are disaccharides.

DISCUSSION

It is known from the literature that plant endo-β-1,4-glucanases do not have sequences participating in substrate binding as do bacterial glucanases [25]. This can be explained on the basis that plant cellulases act on non-crystalline cellulose and other non-cellulose substrates

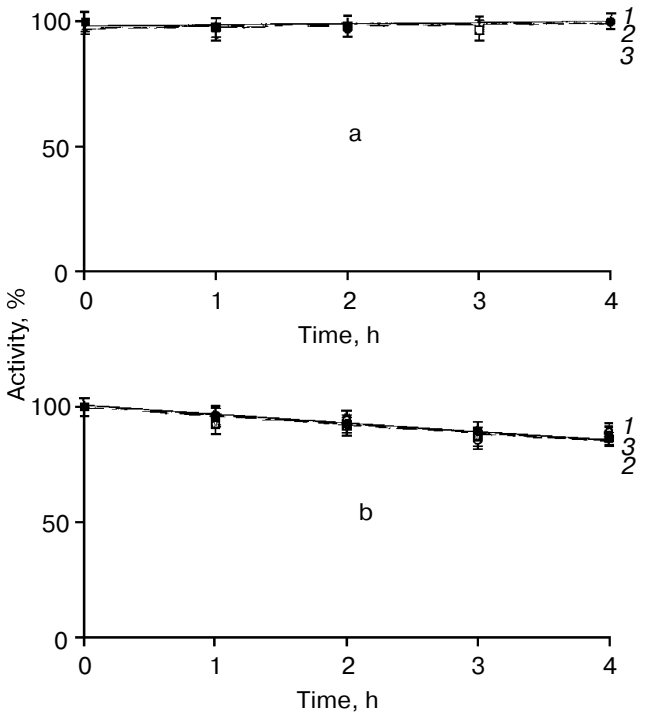

Fig. 5. Stability of cellulase deletion variants at 65 (a) and 70 and 80°C (b): *1*) CelEM1; *2*) CelEM2; *3*) CelEM3.

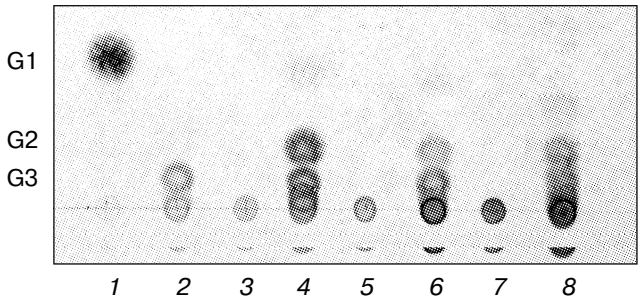

Fig. 6. Thin layer chromatography of oligosaccharide products released when CelEM1 cellulase was incubated with carboxymethylcellulose, lichenan, and xylan. G1, G2, G3) 1, 2, and 3 glucose residues, respectively. Lanes: *1*) glucose; *2*) three residues of glucose; *3*) unhydrolyzed lichenan; *4*) products of lichenan hydrolysis; *5*) unhydrolyzed CMC; *6*) products of CMC hydrolysis; *7*) unhydrolyzed xylan; *8*) products of xylan hydrolysis.

having β-1,4-glycosidic bonds and do not disrupt the structure of cellulose fibrils. As mentioned in the introduction, the purpose of this work was to obtain and investigate cellulase that can imitate the action of plant enzymes. Thus, only the catalytic domain of CelE was chosen for this study, and its biochemical properties have been investigated. We suggested that if the properties of the catalytic domain of CelE from *C. thermocellum* are similar to those of plant glucanases, it should be possible to use the expression of CelE glucanase for modeling the action of plant cellulase.

Hall and coworkers investigated the effect of C-terminus deletion on CelE activity and demonstrated the absence of significant effects on the cellulase catalytic activity [11]. These data are fully consistent with the results of computer modeling: the C-terminal border of the catalytic domain is a Pro-Thr-rich sequence. Hall with coworkers showed that the leader signal is subjected to proteinase hydrolysis at residue 34 in a heterological host. They assumed that catalytic domain begins with residue 35. This does not agree with the theoretically determined N-terminus border of the catalytic domain, which begins from residue 41. In this paper, we have demonstrated that the deletions (up to residue 52) and additional amino acid residues at the N-terminus of cellulase do not influence the biochemical properties of the enzyme. The data indicate that the N-terminal region of the enzyme is not sensitive to amino acid substitution. The N-terminal border of cellulase is conditional. The analysis of computer model of three-dimensional protein structure confirms this assumption (Fig. 1b). It is seen that the enzyme catalytic center is placed distally to the closely placed N- and C-terminuses. The enzyme retains activity with deletions in its N-terminus; anchoring of the protein in the membrane to expose it outside the cell does not change the cellulase activity if there is a flexible linker at the N- and C-terminuses of the protein. The catalytic center will be located outside, having maximal possibility for access to different substrates.

It should be emphasized that most plant cellulases are secreted to the intracellular space with the help of specific sequences usually placed in the N-terminus of the protein [26]. Also, it is known that some of plant cellulases act outside the plasma membrane, being bound to the membrane by a hydrophobic domain via a flexible linker [27].

The CelE cellulase demonstrates maximum activity on lichenan instead of carboxymethylcellulose. This can be explained by the fact that lichenan, in contrast to CMC, is a native substrate consisting of glucose monomers. It should be emphasized that true substrates of hydrolytic enzymes are poorly investigated. Potentially, cellulase may hydrolyze any substrates having β-1,4-glycoside bonds. We think the choice of lichenan as a substrate for all qualitative and quantitative measurements of cellulase activity is preferable.

The study of specific activity versus pH shows that the enzyme is active over a wide range of pH. It should be noted that plant cellulases include acidic and basic isoforms. The range of pH values at which β-1,4-glucanase retains more than 50% activity is optimal for acidic as well as basic isoforms of plant cellulase participating in plant vital functions [28].

It was important to determine the enzyme activity under the optimal living temperature of plants, 25°C. At this temperature the β-1,4-glucanase has ~10% of its activity (5-10 U/mg) (Fig. 4b). Such level of activity is close to the activity of plant cellulase induced by endogenous and exogenous factors [29]. The high thermostability of the enzyme allows testing of its activity over long incubation times. This can be important for testing the enzyme activity in plants when cellulase is expressed at a low level. We think the wide substrate specificity of cellulase should allow it to effectively participate in plant polyglucan metabolism.

The results of thin layer chromatography demonstrated that the modified enzyme hydrolyzes lichenan, carboxymethylcellulose, and xylan. The products of hydrolysis are disaccharides for the each investigated substrate. This means that hydrolysis of plant cell wall can produce oligosaccharides with different chain length, including those that may potentially carry out various regulatory functions in plants. It is well known that particular oligosaccharide fragments, named oligosaccharins, are regulator molecules and act at lower concentrations than phytohormones [30, 31].

There are conflicting data on the biochemical properties of CelE cellulase of *C. thermocellum* in the literature [11, 32]. Mosolova et al. demonstrated that deletion variants CelE without CBD demonstrate activity on CMC (98 U/mg), lichenan (162 U/mg), and xylan (14 U/mg) with temperature optimum 60°C and pH optimum 5.5-6.0 [32]. However, the data obtained of Hall et al. are different [11]. This can perhaps be explained by the use of different

methods of enzyme purification, conditions of enzyme reaction, and substrates obtained from various manufacturers. In the same way, we explain our data versus that of other authors [32] about the specific activity on different substrates and the optimum pH.

It should be emphasized that we used heated protein extracts and did not make additional steps of enzyme purification. This level of purification is enough to make correct estimation of specific activity. Our data are consistent with the data obtained by Hall et al. [11].

Based on the structure and biochemical properties of the modified enzyme, we think it is promising to use the CelEM1 cellulase deletion variant to study the role of plant endo-β-1,4-glucanases.

This work received financial support from the Russian Foundation for Basic Research, grant No. 01-04-48013.

REFERENCES

1. Bayer, E. A., Chanzy, H., Lamed, R., and Shoham, Y. (1998) *Curr. Opin. Struct. Biol.*, **8**, 548-557.
2. Ebel, J., and Mithofer, A. (1998) *Planta*, **20**, 335-348.
3. Wu, S., Blumer, J. M., Darvill, A. G., and Albersheim, P. (1996) *Plant Physiol.*, **110**, 163-170.
4. Matsumoto, T., Sakai, F., and Hayashi, T. (1997) *Plant Physiol.*, **114**, 661-667.
5. Yung, M., Schaffer, R., and Putterill, J. (1999) *Plant J.*, **17**, 203-208.
6. Piruzian, E. S., Goldenkova, I. V., Musiychuk, K. A., Abdeev, R. M., Volkova, L. V., and Kobets, N. S. (2000) *Fiziol. Rast.*, **37**, 382-389.
7. Piruzian, E. S., Monzavi-Karbassi, B., Darbinian, N. S., Goldenkova, I. V., Kobets, N. S., and Mochulsky, A. V. (1998) *Mol. Gen. Genet.*, **257**, 561-567.
8. Musiychuk, K. A., Goldenkova, I. V., Abdeev, R. M., Kobets, N. S., and Piruzian, E. S. (2000) *Biochemistry (Moscow)*, **65**, 1397-1402.
9. Gudliemli, G., and Beguin, P. (1998) *FEMS Microbiol. Lett.*, **161**, 209-215.
10. Lamed, R., and Bayer, E. A. (1988) *Adv. Appl. Microbiol.*, **33**, 1-46.
11. Hall, J., Hazlewood, G. P., Barker, P. J., and Gilbert, H. J. (1988) *Gene*, **69**, 29-38.
12. Boisset, C., Chanzy, H., Henrissat, B., Lamed, R., Shoham, Y., and Bayer, E. A. (1999) *Biochem. J.*, **340**, 829-835.
13. Shoham, Y., Lamed, R., and Bayer, E. A. (1999) *Trends Microbiol.*, **7**, 275-281.
14. *Protocol for Ni-NTA Spin Kit* (1994) *QIAGEN GmbH.*
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, N. Y.
16. Bumazkin, B. K., Velikodvorskaya, G. A., Tuka, K., Mogutov, M. A., and Strongin, A. Y. (1990) *Biochem. Biophys. Res. Commun.*, **167**, 1057-1063.
17. Wood, T. M., and Bhat, K. M. (1988) *Meth. Enzymol.*, **160**, 87-112.
18. Bradford, M. M. (1976) *Analyt. Biochem.*, **72**, 248-254.
19. Laemmli, U. K. (1970) *Nature*, **227**, 680-685.
20. Teather, R., and Wood, P. (1982) *Appl. Environ. Microbiol.*, **43**, 777-780.
21. Begiun, P. (1983) *Analyt. Biochem.*, **133**, 333-336.
22. Lake, B. D., and Goodwin, H. J. (1976) *Chromatogr. Electrophor. Techn.*, **1**, 345-366.
23. Peitsch, M. C. (1995) *Bio/Technology*, **13**, 658-660.
24. Henrissat, B., and Davies, G. (1997) *Curr. Opin. Struct. Biol.*, **7**, 637-644.
25. Brummel, D. A., Lashbrook, C. C., and Bennett, A. B. (1994) *Am. Chem. Soc. Symp. Ser.*, **566**, 100-129.
26. Llop-Tous, I., Domýnguez-Puigjaner, E., Palomer, X., and Vendrell, M. (1999) *Plant Physiol.*, **119**, 1415-1421.
27. Brummell, D. A., Catala, C., Lashbrook, C. C., and Bennett, A. B. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 4794-4799.
28. Nicol, F., and Hofte, H. (1998) *Curr. Opin. Plant Biol.*, **1**, 12-17.
29. Sela-Buurlage, M. B., Ponstein, A. S., Bres-Vloemans, S. A., Melchers, L. S., van den Elzen, P. J. M., and Cornelissen, B. J. C. (1993) *Plant Physiol.*, **101**, 857-863.
30. Farkas, V., and Maclachlan, G. (1988) *Carbohydrate Res.*, **184**, 213-219.
31. Elsbergeim, P., and Darvill, A. G. (1985) *V Mire Nauki*, **11**, 16-23.
32. Mosolova, T. P., Kalyuzhniy, S. V., Varfolomeev, S. D., and Velikodvorskaya, G. A. (1995) *Biochemistry (Moscow)*, **60**, 760-768 (Russ.).