ELSEVIER

Comparative Biochemistry and Physiology Part B 137 (2004) 169–178

CBP

www.elsevier.com/locate/cbpb

Molecular cloning and characterization of an endo-1,3-β-D-glucanase from the mollusk *Spisula sachalinensis*

Valeri B. Kozhemyako[a,*], Denis V. Rebrikov[b], Sergey A. Lukyanov[b], Ekaterina A. Bogdanova[b], Antoine Marin[c], Alexey K. Mazur[c], Svetlana N. Kovalchuk[a], Elena V. Agafonova[a], Victoria V. Sova[a], Ludmila A. Elyakova[a], Valeri A. Rasskazov[a]

[a]*Pacific Institute of Bioorganic Chemistry, Far East Branch of the Russian Academy of Sciences, pr.100 let Vladivostoku, 159, 690022, Vladivostok, Russia*

[b]*Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia*

[c]*Institut de Biologie Physico-Chimique, Paris, France*

Received 18 July 2003; received in revised form 29 October 2003; accepted 30 October 2003

Abstract

cDNA encoding the endo-1,3-β-D-glucanase from *Spisula sachalinensis* (LIV) was amplified by PCR using oligonucleotides deduced from the N-terminal end peptide sequence. Predicted enzyme structure consists of 444 amino acids with a signal sequence. The mature enzyme has 316 amino acids and its deduced amino acid sequence coincides completely with the N-terminal end (38 amino acids) of the β-1,3-glucanase (LIV) isolated from the mollusk. The enzyme sequence from Val 121 to Met 441 reveals closest homology with *Pacifastacus leniusculus* lipopolysaccharide- and β-1,3-glucan-binding protein and with coelomic cytolytic factors from *Lumbricus terrestris*. The mollusk glucanase also shows 36% identity and 56% similarity with β-1,3-glucanase of the sea urchin *Strongylocentrotus purpuratus*. It is generally considered that invertebrate glucanase-like proteins containing the bacterial glucanase motif have evolved from an ancient β-1,3-glucanase gene, but most of them lost their glucanase activity in the course of evolution and retained only the glucan-binding activity. A more detailed evaluation of the protein folding elicited very interesting relationships between the active site of LIV and other enzymes, which hydrolyze native glucans.

Keywords: Active-site; Cloning; Endo-1,3-β-D-glucanase; Enzyme evolution; Mollusk; Sequence homology

1. Introduction

β-1,3-glucanases represent a well-known class of enzymes widespread in bacteria, fungi, plants and marine animals. These are hydrolases specific to *O*-glycoside bonds between 1,3-linked gluco-pyranose residues found in a variety of β-glucans. Functions of these enzymes are diverse and quite distinctive in kingdoms (Mackay et al., 1985). In pathogenic bacteria, they participate in digestion of cell walls. In higher plants they cleave glucans in seeds and also act as inducible defense enzymes. They are involved in autocatalysis of extracellular matrix glucans in fungi and also in the yeast cell development. In the animal kingdom such enzymes are commonly found in marine echinoderms where they take part in digestion of algal food and also play some important role in embryogenesis. Genes encoding β-1,3-glucanases have been cloned and sequenced from some bacteria, fungi and plants.

*Corresponding author. Tel.: +7-4232-310703; fax: +7-423-2-314050.

E-mail address: kozhemyako@piboc.dvo.ru (V.B. Kozhemyako).

1096-4959/04/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbpc.2003.10.018

To our knowledge, the only example of molecular cloning of an animal β-1,3-glucanase was reported for the β-1,3-glucanase from eggs of the sea urchin *Strongylocentrotus purpuratus* (Bachman and McClay, 1996).

The marine bivalve mollusk *Spisula sachalinensis* is a widespread species found in the northwest of the Sea of Japan, that utilizes β-1,3-glucans as the main source of glucose. The crystalline styles of this and similar bivalve mollusks were shown to be a rich source of β-1,3-glucanases (Sova et al., 1970a). The endo-1,3-β-D-glucanase (LIV) from *S. sachalinensis* was one of the first such enzymes isolated and its kinetics and catalytic mechanism were studied by various methods (Sova et al., 1970b; Isakov et al., 1971; Sova and Elyakova, 1972; Bezukladnikov and Elyakova, 1986, 1990). Herein we describe the gene cloning and report the putative amino acid sequence of LIV, which gives the second such example of a β-1,3-glucanase from animal tissue. The overall folding of the protein was recognized by standard software with a high credibility score. Sequence comparisons and analysis of known three-dimensional enzyme structures suggest that there is a significant spatial homology in the glucone part of the active-site between LIV and some beta-1,4 splitting enzymes.

2. Materials and methods

2.1. Animals

Mollusks (*S. sachalinensis*, Schrenck 1862) were collected in Trinity Bay (northwestern part of the Sea of Japan) near the Marine Experimental Station of the Pacific Institute of Bioorganic Chemistry.

2.2. Enzyme purification

Endo-1,3-β-D-glucanase was purified as described previously (Sova et al., 1970b).

Peptide sequencing of the N-terminal end (38 amino acids) was performed on an Edman Automated Sequencing Apparatus Beckman 890 C.

2.3. RNA isolation and cDNA synthesis

Total RNA was isolated from 0.3 g of hepatopancreas of live mollusk by the guanidine thiocyanate/phenol/chloroform method (Chomczynski and Sacchi 1987). A cDNA library was synthesized from total RNA and amplified by a SMART PCR cDNA Synthesis Kit (CLONTECH) using the provided protocol. cDNA amplification PCR was carried out for 22 cycles with 8 s at 94 °C, 10 s at 64 °C and 4 min at 68 °C.

2.4. Gene cloning and DNA sequencing

Degenerate primers were designed to the flanks of the N-terminal amino acid sequence GTVVF**RDDFNG**AFDPAGWNYEVSMYGGYN-**WEVQA**YVPD (bolded letters indicate the sites for primers V1: CGNGA(C/T)GA(C/T)TT(C/T)AA(C/T)GG and V2: GC(T/C)TGNAC(T/C)TCCCA(A/G)TT). The amplified cDNA was diluted 50-fold in water, and 1 μl of this dilution was used for PCR with V1 and V2 primers. After 31 cycles approximately 90-bp fragment was obtained. PCR product was cloned in pTAdv-cloning vector (CLONTECH) and sequenced using M13 direct and reverse universal primers by using a Beckman SEQ-2000 automated sequencer and the FS dye terminator chemistry. The nucleotide sequence obtained completely corresponded to endo-1,3-β-D-glucanase N-terminal amino acid sequence. Then primers for 5′- and 3′-RACE were designed (primer for 5′-RACE: 5′-TAACTC-TGCGTTGATACCAC-3′ and for 3′-RACE: 5′-CCTGCCGGCTGGAATTATGA-3′) and PCR was performed as described (Matz et al., 1999). Two obtained fragments (600 bp for 5′-RACE and 900 bp for 3′-RACE) were cloned and sequenced as described above.

2.5. Sequence analysis and comparisons

Sequences were analyzed using the BLAST website. Domain organization was verified using the web-based SMART tool (http://smart.embl-heidelberg.de). Multiple sequence alignment was performed with the Clustal W version 1.8 software (Thompson et al., 1994) with subsequent manual optimization. The phylogenetic tree was constructed by using a truncated sequence of the glucanase-like domain that appeared to be homologous to a number of invertebrate, fungal, plant and bacterial proteins. The quartet puzzling tree with maximum likelihood branch lengths was calculated using the TREE-PUZZLE program, version 5.0 (Strimmer and von Haeseler, 1997). Maximum likelihood trees were inferred by using quartet puzzling with

10 000 puzzling steps. We used the discrete G-distribution model (with eight categories) for site heterogeneity (Yang, 1996), and Whelan and Goldman model of substitution (Whelan and Goldman, 2001). The neighbor-joining tree was used to estimate the parameters. A possible direct or transitive homology between the LIV endo-1,3-β-D-glucanase and known three-dimensional protein structures was examined by using the BLAST (Altschul et al., 1997), the 3D-PSSM (Divne et al., 1994) and the FROST (Marin et al., 2002) software with the SwissProt databank (Bairoch and Apweiler, 2000), the protein databank (PBD) (Berman et al., 2000) and FSSP (Holm and Sander, 1996).

3. Results

The nucleotide and predicted amino acid sequences are shown in Fig. 1. The 5′ upstream leader contains 118 aa. The mature protein has 316 aa, with a predicted Mr of 38 kDa and IP of 7.0. Its amino acid composition agrees well with the earlier reported experimental results for the purified enzyme (Sova et al., 1993).

A database search performed with the BLAST software shows that the mollusk glucanase has a significant sequence homology with the lipopolysaccharide and β-1,3-glucan binding proteins and bacterial and fungal glucanases. Alignment of the part of the mollusk glucanase sequence, from Val121 to Met441, reveals closest homology with *Pacifastacus leniusculus* lipopolysaccharide and β-1,3-glucan binding protein (LGBP) (GenBank accession no. JC6141, Lee et al., 2000) (39% identity and 58% similarity), and with coelomic cytolytic factors from *Lumbricus terrestris* (GenBank accession no. AAL09587, Bilej et al., 2001) (38% identity and 55% similarity) and from *Eisenia fetida* (GenBank accession no. AAC35887, Beschin et al., 1998) (37% identity and 54% similarity). All these factors were reported as lipopolysaccharide and beta-glucan-binding proteins being able to cause activation of the proPO system. The mollusk β-1,3-glucanase also shows 36% identity and 56% similarity with β-1,3-glucanase of the sea urchin *S. purpuratus* (GenBank accession no. AAC47235, Bachman et al., 1996). A somewhat lower degree of similarity is observed with mixed-linked glucanase of the *Alternaria alternata* (GenBank accession no. AAK69516), with putative β-1,3-glucan binding protein of the *Neurospora crassa* (GenBank accession no. CAC28724) and with insect gram-negative bacteria-binding proteins from *Bombyx mori* (GenBank accession no. AB026441, Lee et al., 1996), *Anopheles gambiae* (GenBank accession no. CAA04496, Dimopoulos et al., 1997), *Hyphantria cunea* (GenBank accession no. AAD09290) (Shin et al., 1998) and *Drosophila melanogaster* (GenBank accession no. AAF33851, Kim et al., 2000).

When the predicted endo-1,3-β-D-glucanase sequence was analyzed with the SMART software (Schultz et al., 2000) it appeared that it contains a characteristic domain of family 16 of glycosyl hydrolases (glucanase-like domain) that starts at position 179 aa and ends at position 399 aa (Fig. 1). The corresponding *E*-value for the glucanase-like domain prediction was 9.2e-05. Comparison of this mollusk domain with the known glucanase-like domains shows that the mollusk endo-1,3-β-D-glucanase contains a putative polysaccharide binding motif and a putative glucanase active site highly homologous to that of bacterial β-1,3-glucanases and invertebrate glucanase-like proteins (Yahata et al., 1990; Yamamoto et al., 1993; Juncosa et al., 1994; Lee et al., 1996) (Figs. 1 and 2). The molluskan endo-1,3-β-D-glucanase also contains one short putative cell adhesion site and an integrin binding site, Arg-Gly-Asp (Ruoslahti, 1991) located in the mature protein sequence from Arg243 to Asp245.

Shown in Fig. 2 is the multiple alignment of the most conservative part of the mollusk β-1,3-glucanase amino acid sequence with those of other similar invertebrate, fungal, plant and bacterial proteins. Based upon this alignment, a maximum likelihood tree was constructed as shown in Fig. 3.

The mature protein sequence (residues 119–441) was used for a more detailed evaluation of the protein fold and a search of possible analogs in the PDB. The 3D-PSSM software (Kelley et al., 2000) reveals significant homology with the PDB entry 2AYH (*E*-value 2.2e-04). The next two entries in the ranking, 1CPM and 1DYP, are characterized by considerably higher *E*-values. All three are beta-sandwich proteins with closely similar folding topologies; moreover, they all are β-*O*-glucanases of family 16. The 2AYH and 1CPM are representatives of the same group (Hahn et al., 1994, 1995). These are hybrid chitinases constructed from native enzyme sequences of *Bacillus*

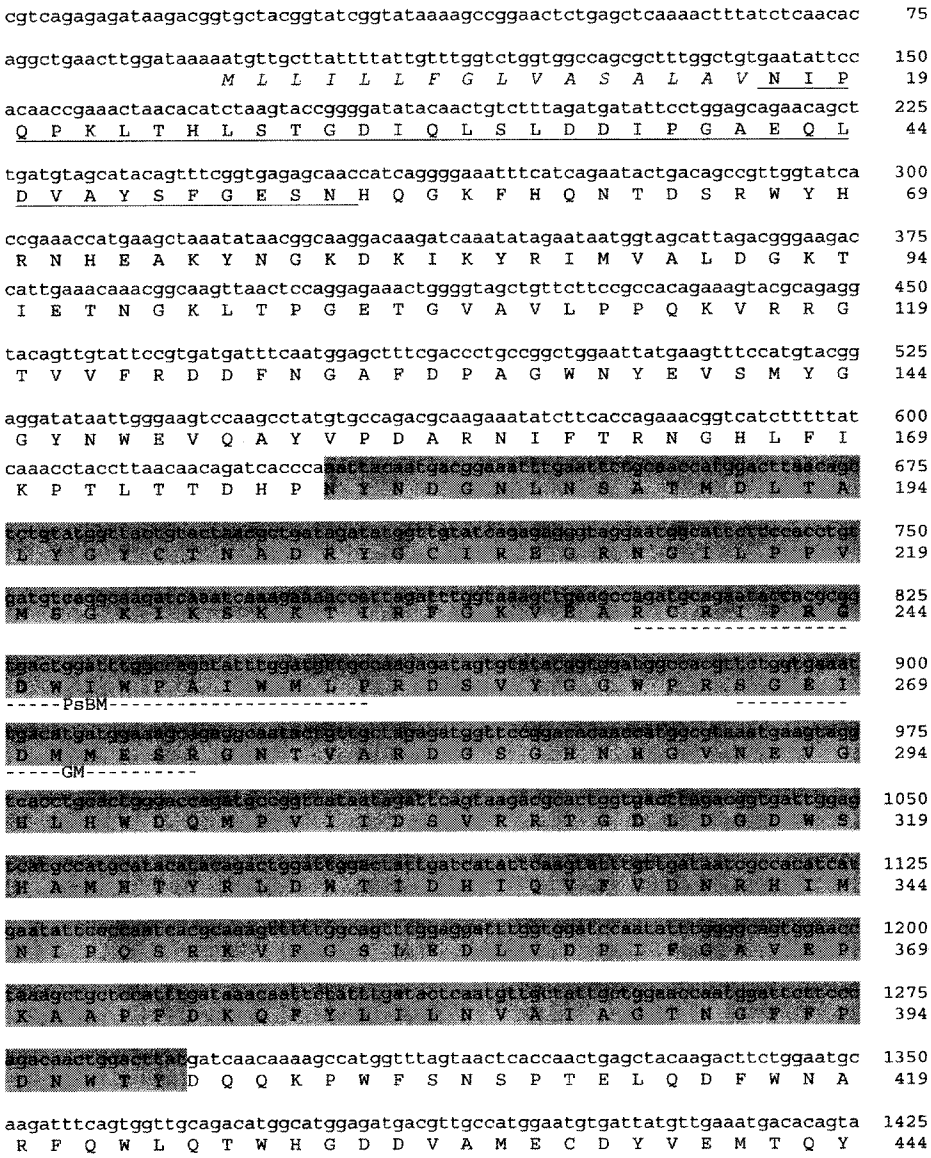

Fig. 1. Nucleotide sequence of the mollusk endo-1,3-β-D-glucanase cDNA and deduced amino acid sequence. Amino acids are numbered from the initial methionine. The underlined amino acid residues were confirmed by sequencing of the N-terminal end of the mature protein. Predicted signal peptide sequence is given in italics. glucanase-like domain is given in shaded box. Putative polysaccharide binding motif (PsBM) and glucanase active site (GM) are dot-lined. Putative cell adhesion site and integrin binding site is bold face.

macerans and *B. amyloliquefaciens*. In contrast, 1DYP is a native kappa-carrageenase from a marine bacterium *Pseudoalteromonas carrageenovora* (Michel et al., 2001). The last protein also exhibits relatively high sequence identity (20%) with the LIV glucanase. A broader search of three-dimensional structures similar to these three proteins performed with FSSP (Holm and Sander, 1996) showed that there is only one additional glucanase with very slight homology, namely, the family 7 cellobiohydrolase from *Trichoderma reesei* (PDB entry 1CEL) (Divne et al., 1994).

The four above mentioned enzymes differ from all other known glucanases by their reaction centers. In their three-dimensional structures, the three putative catalytic carboxyls exposed to the substrate binding cleft are not only close in space, but virtually consecutive in the sequence. In chitinases

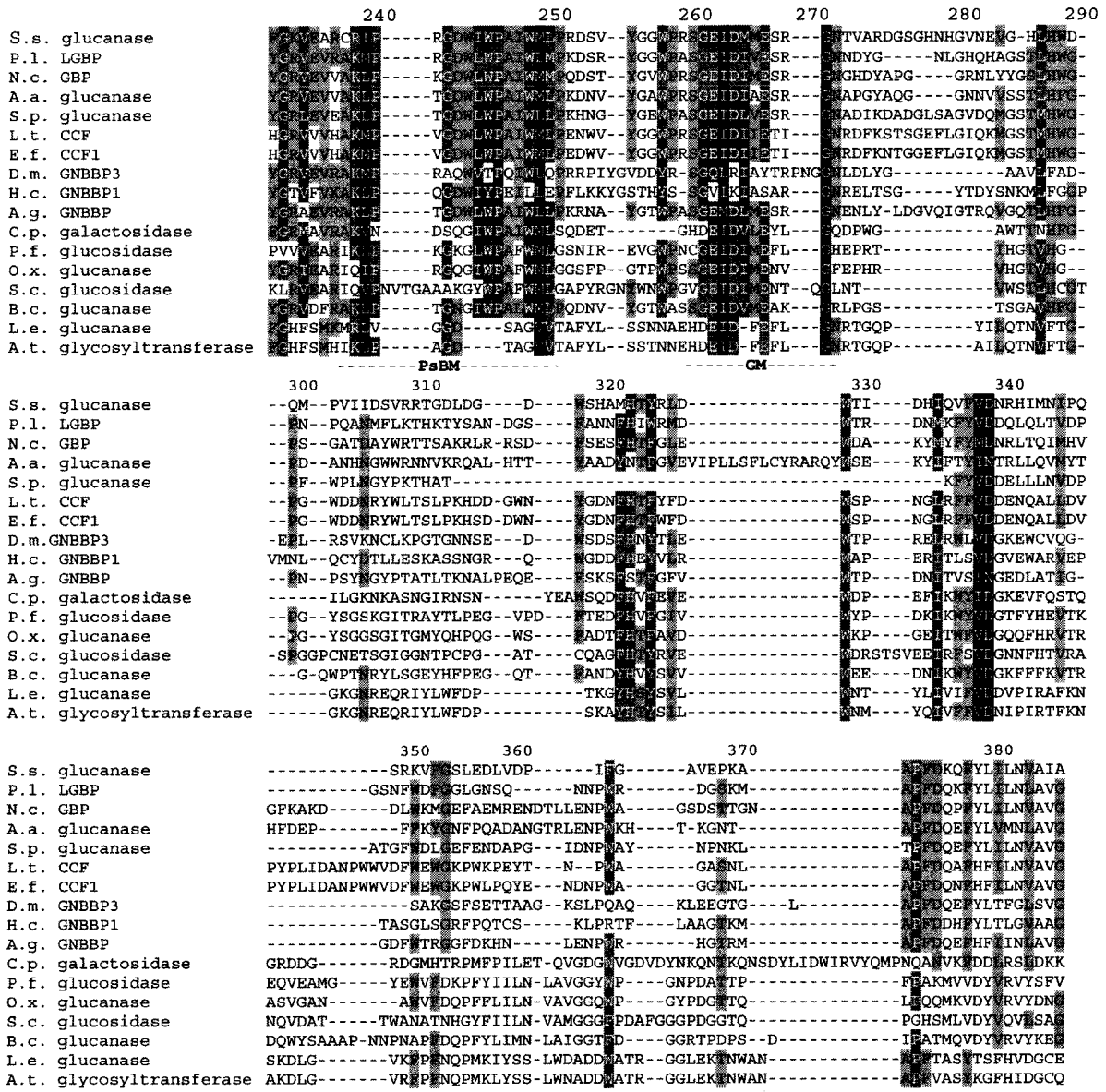

Fig. 2. Alignment of the part of the mollusk endo-1,3-β-D-glucanase amino acid sequence with those of other similar invertebrate, fungal, plant and bacterial proteins. The Clustal W version 1.8 software was used to align the sequence of mollusk β-1,3 glucanase with the similar proteins, which contain regions with similarity to bacterial glucanases. Conserved residues of the sequence are highlighted in black (strictly conserved) or in gray. Gaps were introduced to obtain maximal sequence alignment. Putative polysaccharide binding motif (PsBM) and glucanase active site (GM) are dot-lined. S.s. glucanase–mollusk *S. sachalinensis* β-1,3-glucanase (GenBank accession no. **AY308829**); P.l. LGBP–*P. leniusculus* LGBP (GenBank accession no. **JC6141**); N.c. GBP–*N. crassa* related to beta-1, 3-glucan binding protein (GenBank accession no. **CAC28724**); A.a. glucanase–*A. alternata* mixed-linked glucanase (GenBank accession no. **AAK69516**); S.p. glucanase–sea urchin *S. purpuratus* β-1,3-glucanase (GenBank accession no. **AAC47235**); L.t. CCF–*L. terrestris* coelomic cytolytic factor (GenBank accession no. **AAL09587**); E.f. CCF1–*E. fetida* coelomic cytolytic factor (GenBank accession no. **AAC35887**); D.m. GNBBP3–*D. melanogaster* gram-negative bacteria binding protein 3 (GenBank accession no. **AAF33851**); H.c. gnbbp1–*H. cunea* gram-negative binding protein (GenBank accession no. **AAD09290**); A.g. gnbbp–*A. gambiae* putative gram negative bacteria binding protein (GenBank accession no. **CAA04496**); C.p. galactosidase–*Clostridium perfringens* endo-beta-galactosidase C (GenBank accession no. **BAA97991**); P.f. glucosidase–*Pyrococcus furiosus* endo-beta-1,3-glucanase (GenBank accession no. **AAC25554**); O.x. glucanase–*Oerskovia xanthineolytica* beta-1,3-glucanase II (GenBank accession no. **AAC38290**); S.c. glucosidase–*Streptomyces coelicolor* secreted glucosidase (GenBank accession no. **T35164**); B.c. glucanase–*Bacillus circulans* beta-1,3-glucanase (GenBank accession no. **AAC60453**); L.e. glucanase—tomato *Lycopersicon esculentum* xyloglucan endo-1,4-β-D-glucanase (GenBank accession no. **D49539**); A.t. glycosyltransferase–*Arabidopsis thaliana* endo-xyloglucan transferase (GenBank accession no. **BAA03921**). Numbering corresponds to the mollusk glucanase sequence shown in Fig. 1.

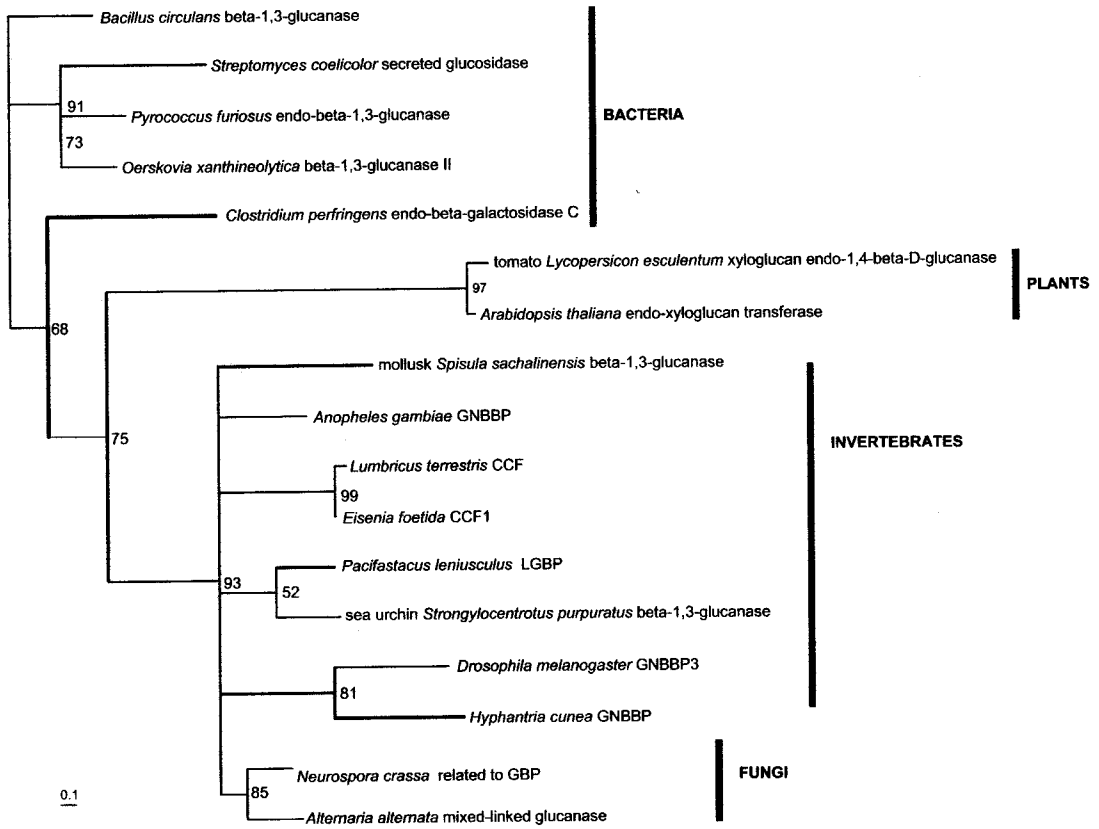

Fig. 3. Phylogenetic tree derived by comparison of the amino acid sequences of glucanase-like proteins. Alignment of glucanase-like amino acid sequences present at Fig. 2. The quartet puzzling tree with maximum likelihood branch lengths was calculated using the Tree-PUZZLE program. Numbers at nodes indicate support values for each branch. *B. circulans* β-1,3-glucanase sequence was selected as outgroup.

2AYH and 1CPM the catalytic motif is 'ExDxE'. It forms a fragment of a beta-strand, with the two hydrophobic side chains denoted as 'x' subtending towards the hydrophobic protein interior and the three carboxyl side chains facing the opposite direction. In 1DYP and 1CEL, the similar motif is 'ExDxxE', but the relative orientations of the three carboxyls remains the same owing to the so called beta-bulge which allows the two consecutive hydrophobic side chains in a beta-strand to subtend to interior without strongly disturbing the overall strand direction. A similar motif is also found in the mollusk β-1,3-glucanase (residues 268–273) as well as other laminarinases suggesting that the carrageenase 1DYP is possibly their closest relative in PDB. It appeared also that there is significant transitive homology relationship between our predicted sequence and 1DYP when a bacterial laminarinase sequence (SwissProt entry E13B_BACCI) was used as an intermediate. The intermediate has 30% identity and an *E*-value of 7e-22 with the mollusk sequence, and it also shows 26% identity and an *E*-value of 7e-6 with 1DYP. The transitive homology found covers residues from 121 to 387.

The supposed structural relationship between 1DYP and the LIV β-1,3-glucanase has been confirmed with the FROST fold recognition program (Marin et al., 2002). FROST uses a threading algorithm that estimates how difficult it would be to replace all amino acids in a given three-dimensional structure by those in the sequence in question. The normalized score of 5.8 obtained suggests a significant relationship between the LIV sequence and the 1DYP spatial structure, and promises more than 75% of correctly assigned residues (Marin et al., 2002). The FROST alignment together with the two transitive BLAST

```
k-carrageenase ----DEFN-VKDATKWNFQTENYGVWSWKN-------ENATVSKGKLKLTT---------
                   D FN   D   WN     YG   W          N     G L
S.s. glucanase VVFRDDFNGAFDPAGWNYEVSMYGGYNWEVQAYVPDARNIFTRNGHLFIKPTLTTDHPNY

k-carrageenase -------------KRESHQRTFWDGCNQQQVANYPLYYTSGVAKSRATGNYGYYEARIK
                                      GC              SG  KS  T   G  EAR
S.s. glucanase NDGNLNSATMDLTALYGYCTNADRYGCIREGRNGILPPVMSGKIKSKKTIRFGKVEARCR

k-carrageenase GASTFPGVSPAFWMYSTIDRSLTKEGDVQYSEIDVVELTQKSAVRESDHDLHNIVVKNGK
                        PA WM     R     G     EID  E       R      H      G
S.s. glucanase IP-RGDWIWPAIWML---PRDSVYGGWPRSGEIDMMESRGNTVARDGSGHNHGVNE-VGH

k-carrageenase PTWMRPGSFPQTNHNGYHLPFDPRNDFHTYGVNVTKDKITWYVDGEI---VGEKDNLYWH
                 W               L  D     HTY    T D I   VD
S.s. glucanase LHWDQMPVIIDSVRRTGDLDGDWSHAMHTYRLDWTIDHIQVFVDNRHIMNIPQSRKVFGS

k-carrageenase RQMNLTLSQGLRAPHTQWKCNQFYPSANKSAEG
                       G   P       QFY   N    G
S.s. glucanase LEDLVDPIFGAVEPKAAPFDKQFYLILNVAIAG
```

Fig. 4. Optimal alignment of the LIV and kappa-carrageenase homologous fragments identified by the FROST fold recognition program. Shown in boldface are the active site residues exhibited in Fig. 5. *k*-carrageenase–kappa-carrageenase of *P. carrageenovora*. S.s. glucanase-endo-1,3-β-D-glucanase of *S. sachalinensis*.

alignments was used to obtain the one shown in Fig. 4. We can see that the aligned fragment of the LIV sequence is longer by nearly 30 consecutive residues, which is close in length to one antiparallel beta-sheet in the same beta-sandwich. However, PSSM secondary structure prediction algorithm (Kelley et al., 2000) does not assign a definite structure here and it probably corresponds to an external loop.

Fig. 5 shows the three-dimensional fold of the kappa-carrageenase (PDB entry 1DYP; Michel et al., 2001) viewed from the opening of the putative substrate binding cleft. It is well known that, in the binding regions of glucanases, the spatial orientations of key chemical functions, like carboxyl groups or aromatic cycles, are more conservative than their sequences. Shown in Fig. 5 are the side chains protruding into the substrate binding cleft that are identical or similar in the two sequences according to the alignment in Fig. 4. Among them one finds the three catalytic carboxyls and the 'PAxW' motif in two neighboring beta-strands, which is a highly conservative feature in beta-glucanase sequences. The tryptophan residue

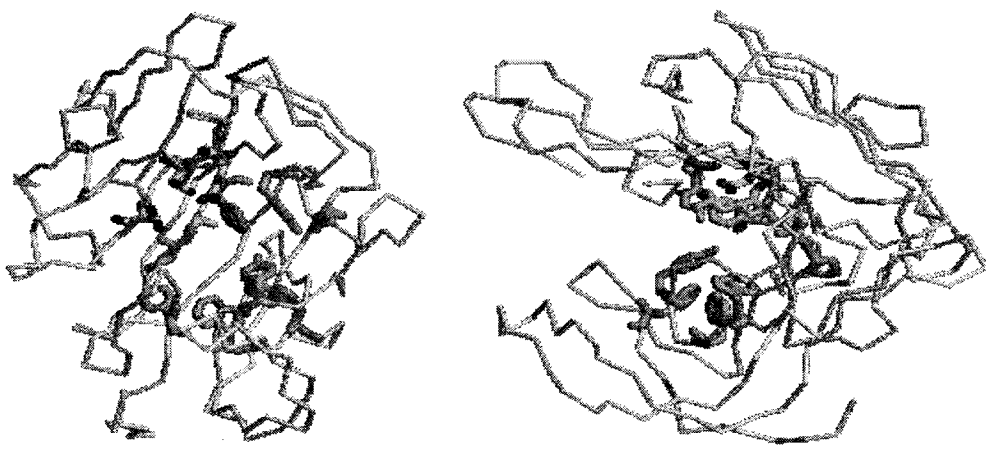

Fig. 5. Two orthogonal views of the kappa-carrageenase backbone structure (PDB entry 1DYP, Michel et al., 2001). The side chains shown are homologous with the LIV β-1,3-glucanase sequence according to the alignment in Fig. 4. The three catalytic residues are colored by atoms whereas all other side chains are shown in green Fig. 1.

supposedly binds the glycone sugar residue in the catalytic center (Michel et al., 2001), and it is clear from Fig. 5 that it is the glycone part of the substrate binding cleft where the homologous residues are localized. A similar conclusion is drawn when the substrate binding region of the 2AYH chitinase is considered, but in this case the number of residues homologous to LIV is smaller. The chitinase cleaves beta-1,4-*O*-glycosidic bonds between two glucoses while the carrageenase attacks 1,4-beta-D-linkages between D-galactose 4-sulfate and 3,6-anhydro-D-galactose. Despite these strong substrate differences, these enzymes as well as endo-beta-1,3-glucanases seem to have rather similar structures of the glycone parts of their substrate binding sites, apparently specific to the anomeric configuration of the glycone and the catalytic mechanism.

4. Discussion

Our earlier studies have shown that marine invertebrates, represented by a great number of types and species at various stages of evolution, are important sources of carbohydrases. The richest source of β-1,3-glucanases has been shown to be crystalline styles of bivalves (Sova et al., 1970a). Endo-1,3-β-D-glucanase from crystalline styles of *S. sachalinensis* have been studied in greater detail (Sova et al., 1970b; Isakov et al., 1971; Sova and Elyakova, 1972; Bezukladnikov and Elyakova, 1990). In this paper we report the cDNA cloning of the enzyme. The study of the protein fold reveals that the 'ExDxxE' catalytic motif found in LIV and other endo-laminarinases has been encountered in two families of three-dimensional enzyme structures, namely, the kappa-carrageenase (Michel et al., 2001) and the cellobiohydrolase 1CEL (Divne et al., 1994). Interestingly, these are the only examples of inverting glucanases with tunnel shaped active centers. In 1CEL, the tunnel is completely closed so that only endwise attacks of long cellulose chains are possible, with cellobiose released from the non-reducing ends. The tunnel shape of the active center suggests that this endwise cleavage is processive. The carrageenase is an endo-enzyme and its tunnel binding region is partially opened, but this enzyme is also known to degrade long substrate chains with a processive or multiple attack mechanism (Michel et al., 2001). As we have shown earlier, the LIV laminarinase is an endo-glucanase with a large degree of multiple attacks (Bezukladnikov and Elyakova, 1990). Owing to this feature, the enzyme releases a large amount of glucose and provides for a complete splitting of polymeric laminaran in mollusk digestive tracts where no other relevant activities could be detected. This observation agrees with the foregoing structural homology and suggests that the tunnel shaped active centers and the 'ExDxxE' catalytic motif are perhaps related, and possibly represent a general feature of strongly processive glucanases.

Analysis of a maximum likelihood tree (Fig. 3) constructed on the basis of the multiple alignment of the most conservative part of the mollusk β-1,3 glucanase amino acid sequence with those of other similar invertebrate, fungal, plant and bacterial proteins (Fig. 2) reveals that the glucanase-like proteins are divided into three major branches. The first two branches comprise bacterial and plant proteins, respectively, whereas the last branch unites invertebrate and fungal proteins. This result does not support the recent supposition by Bachman and McClay (1996) that sea urchin β-1,3-glucanase has arisen by horizontal transfer from a species similar to *B. circulans* and proves another suggestion of an early divergence of glucanase-like proteins in the prokaryotic/eukaryotic separation. Moreover, it agrees with the classical hypothesis of a sister relationship between fungi and metazoa (Baldauf et al., 2000). At the same time the third branch of our tree indicates that there are ancient divergences of the fungal and metazoan glucanase-like proteins. It is generally considered that invertebrate glucanase-like proteins containing the bacterial glucanase motif have evolved from an ancient beta-1,3-glucanase gene, but most of them lost their glucanase activity in the course of evolution and retained only the glucan-binding activity (Lee et al., 1996; Dimopoulos et al., 1997; Shin et al., 1998; Kim et al., 2000; Lee et al., 2000). However, the phylogenetic tree suggests that the divergence of these proteins could have occurred before the glucanase activity was lost and that further evolution including the loss of the glucanase activity proceeded independently.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Bachman, E.S., McClay, D.R., 1996. Molecular cloning of the first metazoan β-1,3-glucanase from eggs of the sea urchin *Strongylocentrotus purpuratus*. Proc. Natl. Acad. Sci. USA 93, 6808–6813.

Bairoch, A., Apweiler, R., 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res. 28, 45–48.

Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., Doolittle, W.F., 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290, 972–977.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., et al., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.

Beschin, A., Bilej, M., Hanssens, F., Raymakers, J., Van Dyck, E., Revets, H., et al., 1998. Identification and cloning of a glucan- and lipopolysaccharide-binding protein from *Eisenia foetida* earthworm involved in the activation of prophenoloxidase cascade. J. Biol. Chem. 273, 24 948–24 954.

Bezukladnikov, P.W., Elyakova, L.A., 1986. A study of the pattern of action of endo-(1→3)-β-D-glucanases from marine bivalves on a polymer substrate labelled at the reducing end. Carbohydr. Res. 152, 261–273.

Bezukladnikov, P.W., Elyakova, L.A., 1990. Transglycosylation and multiple attack of endo-(1→3)-β-D-glucanase L-IV from *Spisula sachalinensis*: a new approach to the evaluation of the degree of multiple attack on polysaccharides. Carbohydr. Res. 203, 119–127.

Bilej, M., Baetselier, P.D., Dijck, E., Stijlemans, B., Colige, A., Beschin, A., 2001. Distinct carbohydrate recognition domains of an invertebrate defense molecule recognize gram-negative and gram-positive bacteria. J. Biol. Chem. 276, 45 840–45 847.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chlorophorm extraction. Anal. Biochem. 162, 156–159.

Dimopoulos, G., Richman, A., Muller, H.M., Kafatos, F.C., 1997. Molecular immune responses of the mosquito *Anopheles gambiae* to bacteria and malaria parasites. Proc. Natl. Acad. Sci. USA 94, 11 508–11 513.

Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J., et al., 1994. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. Science 265, 524–528.

Hahn, M., Piotukh, K., Borriss, R., Heinemann, U., 1994. Native-like in vivo folding of a circularly permuted jellyroll protein shown by crystal structure analysis. Proc. Natl. Acad. Sci. USA 91, 10417–10421.

Hahn, M., Keitel, T., Heinemann, U., 1995. Crystal and molecular structure at 0.16-nm resolution of the hybrid *Bacillus endo*-1,3-1,4-beta-D-glucan 4-glucanohydrolase H(a16-m). Eur. J. Biochem. 232, 849–858.

Holm, L., Sander, C., 1996. Mapping the protein universe. Science 273, 595–602.

Isakov, V.V., Sova, V.V., Denisenko, V.A., Sakharovsky, V.G., Elyakova, L.A., Dzizenko, A.K., 1971. A study of the action of laminarinase from *Spisula sachalinensis* by the use of nuclear magnetic resonance. Biochim. Biophys. Acta. 268, 184–186.

Juncosa, M., Pons, J., Dot, T., Querol, E., Planas, A., 1994. Identification of active site carboxylic residues in *Bacillus licheniformis* 1,3-1,4-beta-D-glucan 4-glucanohydrolase by site-directed mutagenesis. J. Biol. Chem. 269, 14 530–14 535.

Kelley, L.A., MacCallum, R.M., Sternberg, M.J., 2000. Enhanced genome annotation using structural profiles in the program 3D-PSSM. J. Mol. Biol. 299, 499–520.

Kim, Y.S., Ryu, J.H., Han, S.J., Choi, K.H., Nam, K.B., Jang, I.H., et al., 2000. Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. J. Biol. Chem. 275, 32 721–32 727.

Lee, W.J., Lee, J.D., Kravchenko, V.V., Ulevitch, R.J., Brey, P.T., 1996. Purification and molecular cloning of an inducible gram-negative bacteria-binding protein from the silkworm, *Bombyx mori*. Proc. Natl. Acad. Sci. USA 93, 7888–7893.

Lee, S.Y., Wang, R., Soderhall, K., 2000. A lipopolysaccharide- and beta-1,3-glucan-binding protein from hemocytes of the freshwater crayfish *Pacifastacus leniusculus*. Purification, characterization and cDNA cloning. J. Biol. Chem. 275, 1337–1343.

Mackay, R.M., Baird, S., Dove, M.J., Erratt, J.A., Gines, M., Moranelli, F., et al., 1985. Glucanase gene diversity in prokaryotic and eukaryotic organisms. Biosystems. 18, 279–292.

Marin, A., Pothier, J., Zimmermann, K., Gibrat, J.-F., 2002. FROST: a filter-based fold recognition method. Proteins 49, 493–509.

Matz, M.V., Shagin, D.A., Bogdanova, O.V., Britanova, O., Lukyanov, S.A., Diatchenko, L., Chenchik, A., 1999. Amplification of cDNA ends based on template-switching effect and step-out PCR. Nucleic Acids Res. 27, 1558–1560.

Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., et al., 2001. The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. Structure (Camb) 9, 513–525.

Ruoslahti, E., 1991. Integrins. J. Clin. Invest. 87, 1–5.

Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., Bork, P., 2000. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. 28, 231–234.

Shin, S.W., Park, S.S., Park, D.S., Kim, M.G., Kim, S.C., Brey, P.T., et al., 1998. Isolation and characterization of immune-related genes from the fall webworm, *Hyphantria cunea*, using PCR-based differential display and subtractive cloning. Insect Biochem. Mol. Biol. 28, 827–837.

Sova, V.V., Elyakova, L.A., 1972. Some aspects of the specificity and action pattern of β-1,3-glucan glucanohydrolase from *Spisula sachalinensis*. Biochim. Biophys. Acta 258, 219–227.

Sova, V.V., Elyakova, L.A., Vaskovsky, V.E., 1970a. The distribution of laminarinases in marine invertebrates. Comp. Biochem. Physiol. 32, 459–464.

Sova, V.V., Elyakova, L.A., Vaskovsky, V.E., 1970b. Purification and some properties of β-(1→3)-glucan glucanohydrolase from the crystalline style of bivalvia, *Spisula sachalinensis*. Biochim. Biophys. Acta. 212, 111–115.

Sova, V.V., Svetasheva, T.G., Kostyuk, I.O., 1993. Molecular weight of endo-β-1,3-glucanase from the marine molusc *Spisula sachalinensis*. Biokhimya 57, 1344–1347.

Strimmer, K., von Haeseler, A., 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc. Natl. Acad. Sci. USA 94, 6815–6819.

Thompson, J., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18, 691–699.

Yahata, N., Watanabe, T., Nakamura, Y., Yamamoto, Y., Kamimiya, S., Tanaka, H., 1990. Structure of the gene encoding beta-1,3-glucanase A1 of *Bacillus circulans* WL-12. Gene 86, 113–117.

Yamamoto, M., Aono, R., Horikoshi, K., 1993. Structure of the 87-kDa beta-1,3-glucanase gene of *Bacillus circulans* IAM1165 and properties of the enzyme accumulated in the periplasm of *Escherichia coli* carrying the gene. Biosci. Biotechnol. Biochem. 57, 1518–1525.

Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. 11, 367–372.